Volume 36 · Numbers 3–4 · April 2011

Social Choice and Welfare

355

Social Choice and Welfare · Volume 36 · Numbers 3–4 · April 2011 · pp. 355–650

Springer

Special Issue: Political Economy of Elections and Bargaining

Guest Editors: M. Gallego · N. Schofield · D.M. Kilgour

Further articles can be found at www.springerlink.com

Abstracted/Indexed in: *ABS Academic Journal Quality Guide, CSA Political Science & Government, CSA Sociological Abstracts, Current Contents/Social & Behavioral Sciences, EBSCO, ECONIS, EconLit, e-JEL, and JEL on CD, International Bibliography of the Social Sciences (IBSS), ISI Alerting Services, Mathematical Reviews Database, Research Papers in Economics (RePEc), SCOPUS, Social Science Research Network (SSRN), Social Sciences Citation Index, Social Scisearch, Zentralblatt für Mathematik*

Instructions for Authors for *Soc Choice Welf* are available at http://www.springer.com/00355

Soc Choice Welf
DOI 10.1007/s00355-010-0512-3

Guest editors' introduction to the special issue on the political economy of elections and bargaining

Maria Gallego · Norman Schofield · D. Marc Kilgour

1 Introduction

Modern political economy, a fusion of economics and political science, is a rapidly growing field. It naturally encompasses the bargaining that arises when self-interested actors with partially conflicting preferences interact so as to make joint decisions over public policies. In fact, bargaining occurs at many different stages of the policy-making process. The institutional context is crucial, as any polity, dictatorial, or democratic, needs a process by which political and economic decisions are made.

Gallego, Schofield, and De Donder ran a workshop on The Political Economy of Bargaining in April 2008 at Wilfrid Laurier University in Waterloo, Ontario, Canada (http://www.wlu.ca/sbe/gallego/PEB2008/). Some of the papers presented at that workshop are included in this volume.

The articles collected here examine policy decisions under various institutional arrangements at different levels of the policy-making process. In addition, some papers explicitly model the role of voters in altering the decision process. An extended summary of the following papers can be seen below:

1. "A Social Choice Theory of Legitimacy," by John Patty and Elizabeth Maggie Penn

M. Gallego (✉)
Department of Economics, Wilfrid Laurier University, Waterloo, Canada
e-mail: mgallego@wlu.ca

N. Schofield
Center in Political Economy, Washington University, St. Louis, MO, USA
e-mail: schofield.norman@gmail.com

D. M. Kilgour
Department of Mathematics, Wilfrid Laurier University, Waterloo, Canada
e-mail: mkilgour@wlu.ca

Published online: 21 December 2010

2. "A Theory of Income Taxation where Politicians Focus Upon Core and Swing Voters," by John Roemer.
3. "Bandwagon, Underdog, and Political Competition: The Uni-dimensional Case," by Woojin Lee.
4. "Competition for Popular Support: A Valence Model of Elections in Turkey" by Norman Schofield, Maria Gallego, Ugur Ozdemir, and Alexei Zakharov.
5. "Estimating the Effects of Activists in Two-Party and Multi-party Systems: Comparing the United States and Israel," by Norman Schofield, Chris Claassen, Ugur Ozdemir, and Alexei Zakharov.
6. "Secondary Issues and Party Politics An Application to Environmental Policy," by Vincent Anesi and Philippe De Donder.
7. "Omnibus or Not: Package Bills and Single-Issue Bills in a Legislative Bargaining Game" by Johanna Goertz.
8. "Bargaining over the Budget," by Daniel Diermeier and Pohan Fong.
9. "Intergovernmental Negotiations, Willingness to Compromise, and Voter Preference Reversals," by Maria Gallego and David Scoones.
10. "A Newton Collocation Method for Solving Dynamic Bargaining Games," by John Duggan and Tasos Kalandrakis.

1.1 Summary of "A Social Choice Theory of Legitimacy," by John Patty and Elizabeth Maggie Penn

Those dealing with democratic theory are also concerned with legitimacy: the belief that authorities, institutions, and social arrangements are appropriate, proper, and just. This article develops a theory of legitimacy in complete information environments with a finite set of alternatives, where the final choice is presented along with a principle justifying the choice and a sequence of intermediary steps. A decision is legitimate when presented along with reasons for the decision.

Legitimacy must satisfy "internal" and "external" consistency axioms. Minimal internal consistency requires that the decision sequence be acyclic with respect to the selected principle, i.e., the decision sequence and ultimate policy choice accord to the selected principle. External stability with respect to the selected principle requires that the exclusion of policies not considered be justified by the selected principle. So that, when an alternative not considered is proposed as a replacement for the final choice, a justification for the final choice can be constructed using the underlying decision sequence and the selected principle to obtain a counterargument. Legitimacy is always possible if there is agreement on the principle by which collective choices are governed. Democratic institutions that produce legitimate decisions (and the observed processes by which they are produced) necessarily obscures much, if not all, of the structure of the principles supporting their choices.

The theory shows how the process of deliberative decision making can render ill-behaved principles useful as sources of legitimation for policy choices. Even though it does not solve the problem of cyclic democratic choices through time, Patty and Penn's theory provides an example of how a given decision might be rationalized in the face of cyclic majority preferences.

Once principles are chosen, their theory offers a nonempty set of predicted policy choices as well as a prediction about the possible rationales justifying the chosen policy. When a decision sequence and the final policy choice are provided, it is descriptively analogous to many depictions of the operation of democratic institutions. Moreover, this sequence can be judged in relation to the final policy choices and the group's principles. Since achieving consonance between the decision sequence, the group's principles, and the final choice is a nontrivial task, the deliberative aspect of the theory not only requires an appropriate choice of structure (e.g., a proper sequencing of intermediary steps given the principle being utilized), but also refines the set of outcomes that might be chosen in a legitimate fashion.

Since principles need not be weak orders of the alternatives, this article presents a notion of legitimacy that can be satisfied even when the guiding principles are cyclic or incomplete. In addition to providing a means by which majoritarian decisions can be rendered legitimate, the theory highlights the power of majority preference in refining the set of legitimate outcomes.

1.2 Summary of "A Theory of Income Taxation where Politicians Focus Upon Core and Swing Voters," by John Roemer

This article develops a political economy model of voting over income taxation functions. A voter's "type" is defined by the pre-tax income of the agent or household. Two strategic parties compete for votes in a world where voters' utility is affected by a random component so that agents are probabilistically sincere when voting. This randomness implies that a party concerned about representing its constituents must attempt to represent every household type, since some fraction of every type will vote for each party. Each political party will propose a function which will define the post-tax and post-transfer income, for every possible realization of pre-tax income, and these functions will be chosen from a large space, constrained only by upper and lower bounds on what the marginal tax rates can be.

Each party has two factions: the guardians, concerned with satisfying the party's core constituency, its past voters; and the opportunists, interested in appealing to swing voters. While trying to maximize their vote share, parties propose tax policies that, if implemented, maximize the aggregate welfare of their core voters, subject to the constraint that the proposed tax policy makes swing voters no worse than the opponent's tax policy would. The party's constituency and the swing voters are endogenously determined.

Although equilibria are multiple, they all share the same properties, and especially the fact that both political parties propose the same tax schedule to an endogenously defined middle class, with the left party taxing more the rich and the right party taxing more the poor. There is an important sub-class of equilibria in which (i) each party proposes a piece-wise linear post-fisc policy with three pieces; (ii) the policy proposed by the left party entails an increasing average rate of taxation on the whole domain of incomes, and the policy proposed by the right party entails an average rate of taxation that increases up to a point, and then decreases; and (iii) these equilibria form a two-dimensional manifold, where each equilibrium is characterized by the relative

strength of "swing" versus "core" factions within the parties. In addition to the analytical solution of the model, the article also provides numerical examples and tests its predictions by comparing them to the evolution of US fiscal legislation since 1981, finding "mild support" for the main prediction of the model.

1.3 Summary of "Bandwagon, Underdog, and Political Competition: The Uni-dimensional Case," by Woojin Lee

Lee studies the interaction between voter conformism and equilibrium platforms when two parties compete in elections over a unidimensional policy space. Voter conformism is modeled as voters taking into account opinion polls when deciding which party to support. Voter conformism is either of the bandwagon or of the underdog type. The bandwagon effect refers to the tendency of opinion polls to prompt voters to support the winning party, which in turn increases the party's chances of winning the election. At the opposite extreme is the underdog effect, in which voters support instead the party faring worse in the polls. Voter preferences are modeled as a quasi-linear utility function representing the economic interests of voters which includes their after-tax income, a term that measures how much voters care about the public good, and a utility bonus/penalty for supporting the winning/losing party capturing the common bandwagon or underdog effect on each voter's utility.

Lee extends Roemer's factional bargaining model to allow for voter conformism. Parties are modeled as having two factions: the militants whose objective is to maximize the average well-being of the party members and the opportunists whose only goal is to win the election. Bargaining between factions in each party is modeled using Nash's (1950) framework where each faction may have different weights in the Nash product. The constituency of each party (the set of militants each party represents) is also endogenized as it represents those who vote for the party in equilibrium.

Lee then studies three special cases of his general model which correspond to three widely studied models of political competition. The first case is the one in which militants have no bargaining power within the party. There is then competition between two pure opportunistic parties, which corresponds to the Hotelling–Downs model adapted for voter conformism. In the second case, parties care only about their constituency, which produces an ideological-party equilibrium with voter conformism. Finally, there is an intermediate case where the two factions in both parties have the same bargaining weight. Because the two factions have equal bargaining power, this is the classical Wittman–Roemer model, adapted for endogenous party membership and voter conformism.

The influence of voter conformism is very different in the three cases. When parties care only about winning, voter conformism does not affect the equilibrium policies or vote shares. Voter conformism does impact equilibrium policies in the other two cases, but in opposite directions; equilibrium policies diverge starkly as the degree of conformism increases in the case of equal bargaining weight, while policies move in the same direction as conformism increases when parties care only about their constituency. One implication of Lee's theoretical exercise is that an empirical analyst, before testing for bandwagon and underdog effects, should be clear about the underlying model of party competition.

1.4 Summary of "Competition for Popular Support: A Valence Model of Elections in Turkey," by Norman Schofield, Maria Gallego, Ugur Ozdemir, and Alexei Zakharov and of "Estimating the Effects of Activists in Two-Party and Multi-party Systems: Comparing the United States and Israel," by Norman Schofield, Chris Claassen, Ugur Ozdemir, and Alexei Zakharov

The two articles, "Competition for Popular Support: A Valence Model of Elections in Turkey," and "Estimating the Effects of Activists in Two-Party and Multi-Party Systems: A Comparison of the United States in 2008 and Israel in 1996," use similar stochastic models to study elections in Turkey, the United States, and Israel.

Formal models of elections tend to predict that parties will maximize votes by converging to an electoral center. There is no empirical support for this prediction. In order to account for the phenomenon of political divergence, these two articles present an electoral model where party leaders or candidates are characterized by differing *valences*, the electoral perception of the quality of the party leader or candidate. If valence is simply exogenous, then it can be shown that there is a dimensionless "convergence coefficient," defined in terms of the empirical parameters of the model that determines whether convergent positions at the electoral origin constitute a vote maximizing equilibrium. A necessary condition for such an equilibrium is that this coefficient be bounded above by the dimension of the policy space. A sufficient condition for convergent equilibrium is that this coefficient be less that one.

In the first article, this model is applied to elections in Turkey in 1999 and 2002. It is shown that the convergence coefficient for the 2002 election was 5.94. Simulation of the model obtained two similar Nash equilibria, with all equilibrium positions located on a principal electoral axis. The valence model was then extended by considering sociodemographic valences, the differing perceptions of the party leaders by various sociodemographic groups in the society. Simulation of this sociodemographic model gave equilibrium predictions closer to the actual party positions. The idea of valence was then extended to include the possibility that activist groups contribute resources to their favored parties in response to policy concessions from the parties. For this model, it is shown that parties, in order to maximize vote share, must balance a *centripetal* electoral force against a *centrifugal* activist effect.The conclusion suggests that the differences between estimated party positions and equilibrium positions were due to the influence of activists associated with these sociodemographic groups.

The second article uses these valence models to model the US presidential election of 2008. The electoral data from the 2008 American National Election Survey is used to construct a pure spatial model with valence. The candidates' positions lead to an estimated value of 1.1 for the convergence coefficient, and simulation shows the joint electoral origin is a Nash equilibrium. The valence model is extended to include sociodemographic variables and voter perceptions of character traits. Simulation of the full traits model again showed that the unique vote maximizing equilibrium was one where the two candidates adopted convergent positions, close to the electoral center. This result conflicts with the estimated positions of the candidates in opposed quadrants of the policy. Again, the difference between estimated positions and equilibrium positions allows us to estimate the influence of activist groups on the candidates.

This convergence result for the United States is contrasted with a sociodemographic valence model for the Israel in the election of 1996. For Israel, the convergence coefficient was found to be 3.98, again implying divergence, just as in Turkey. The conclusion notes that related work shows that the models for the United States, Canada, and Britain (all polities with plurality electoral systems) have low convergence coefficients. In contrast, fragmented polities like Turkey, Israel, and Poland have high convergence coefficients, implying party divergence. It is suggested that under plurality rule, activist groups will tend to coalesce in order to increase the impact they have on the political process.

1.5 Summary of "Secondary Issues and Party Politics: An Application to Environmental Policy," by Vincent Anesi and Philippe De Donder

The objective of this article is to study the electoral incentives for political compromise in a two-dimensional, winner-takes-all model of electoral competition with endogenous party formation. The policy space is composed of a "front line issue," redistribution, and a "secondary issue," which is taken to be environmental policy. The environmental issue is considered to be secondary to the redistributive one in the sense that variations in the environmental policy affect less voters' utilities than variations in the redistributive policy. The article aims at answering questions such as: Under what conditions is the equilibrium environmental policy efficient? How is this policy affected by the proportion of voters who care for the environment? What are the conditions under which green parties form at equilibrium?

There are two goods in the economy, a numeraire and a polluting good. Voters differ both in their exogenous income (either low or high) and in whether they care about pollution (i.e., aggregate consumption of the polluting good). Thus, there are four groups of voters. Public policy consists of both a linear income tax and a linear environmental tax. Tax proceeds are redistributed in a uniform lump sum way to all voters, and with quasi-linear preferences only the linear income tax has redistributive consequences.

The modeling of political decision making relies on Levy's theory of endogenous party formation and it is viewed as a two-stage process. In the first stage, representatives of the four groups of voters form political parties. In the second stage, parties propose policies (pairs of tax rates) and compete in a winner-takes-all election with sincere voting. The crucial assumption is that parties are restricted, for credibility issues, to propose policies that belong to the Pareto set of their members. An equilibrium is a partition of representatives into parties and a vector of electoral platforms such that (i) no representative has an incentive to split up the party he or she belongs to, or to merge it with another party and (ii) no party can make its members better-off by choosing another electoral platform.

The results show that the equilibrium environmental policy is never socially optimal, and more surprisingly that it is larger when there is a minority of green voters. Stable green parties who offer the ideal environmental policy of green voters exist only if there is a minority of green voters and income polarization is large enough relative to the saliency of the environmental issue. Furthermore, as polarization increases, the

equilibrium environmental tax policy of a stable green party coexists with marked income redistribution.

1.6 Summary of "Omnibus or Not: Package Bills and Single-Issue Bills in a Legislative Bargaining Game," by Johanna Goertz

Legislators sometimes include unrelated issues in one bill (omnibus bill), and sometimes vote separately on different issues (single-issue bills). In Goetz's model, the main difference is that tradeoffs between issues are possible with omnibus bills but not with single-issue bills. In the model, legislators have to approve an ideological policy and a purely distributive policy. They can do so with an omnibus bill (vote on both policies at once) or by bargaining over the two issues separately.

According to most of the current literature, bargaining over single-issues is considered inefficient in the absence of bargaining frictions; Tradeoffs between bills cannot be made. In her article, Goertz argues that the choice of single-issue bills can be rational for the first proposal maker, if the underlying bargaining follows the rules of demand-bargaining (Morelli 1999). The previous literature have found omnibus bills to be the only possible rational choice for the first proposal maker because it uses a different bargaining game called: alternating-offers (Baron and Ferejohn 1989). The alternating-offers game yields a strong first-mover advantage. To benefit from this advantage in all dimensions, the first mover proposes an omnibus bill.

If the bargaining game does not give an advantage to the first mover—as, for example, in demand bargaining—the first mover can benefit by proposing single-issue bills. In a demand-bargaining game, it is costly for the first mover to make sure that a majority approves her proposal, costly in the sense that she has to propose policies that are not extremely favorable to her. And, as it turns out, omnibus bills are more costly than single-issue bills. So, first movers typically prefer to propose single-issue bills. Only if the asymmetry in ideological policy ideal points in the legislature is large and the first mover is an extreme legislator, she prefers to propose an omnibus bill. This is because an omnibus bill allows her to package a more extreme ideological policy with a distributive policy more favorable to the majority and get it passed.

1.7 Summary of "Bargaining over the Budget," by Daniel Diermeier and Pohan Fong

The article proposes an explanation for the persistence of inefficiently high levels of public spending in a context where the government is divided. The authors consider a model in which a legislature must choose both the amount and the distribution of public spending so that total spending is endogenously determined. Their model accommodates different institutions including corporatist politics, parliamentary systems, and presidential systems. By assuming party preferences have a linear-quadratic setup, they obtain a clean and intuitive closed-form solution.

The authors consider the agenda-setting model of Romer and Rosenthal (1978). The agenda setter first makes a policy proposal to the legislature. If at least a majority of parties supports the proposal, it is implemented; otherwise a status quo remains in place with each party always voting for the proposal if indifferent. In any subgame perfect

equilibrium, the agenda setter maximizes the utility by selecting a proposal from those policy alternatives that make its coalition partner indifferent to the status quo.

As expected, the authors' results show that the budget size in equilibrium depends on its initial size and on its allocation in a majoritarian institution. In particular, when the initial budget is sufficiently large and the initial allocation is sufficiently unequal, the equilibrium budget size can be greater than the coalitionally efficient amount, i.e., the amount that maximizes the joint utility of the coalition partners. On the other hand, regardless of the initial allocation, when the initial budget size is sufficiently small the equilibrium budget size always expands to the coalitionally efficient level.

The results show that government spending is easy to expand but difficult to cut. The legislative bargaining process can create a ratchet effect that keeps public spending at an inefficiently high level. This ratchet effect arises in corner solutions, as the agenda setter cannot propose policies in which the level of public spending for some group is negative.

Even though the ratchet effect holds across political systems, Diermeier and Fong do demonstrate that political institutions matter. Starting at the same status quo, it is more likely for the president in a presidential system to resist proposing a deep budget cut compared to a prime minister in a parliamentary system.

1.8 Summary of "Intergovernmental Negotiations, Willingness to Compromise, and Voter Preference Reversals," by Maria Gallego and David Scoones

Gallego and Scoones develop a model of multi-party elections in which policy formation requires intergovernmental negotiations and voters understand how willingness to compromise in those negotiations will affect policy outcomes. Voter choices depend both on parties' ideals, that is, what they would do were they to have free choice, and their willingness to compromise. The model examines voter choices among three parties where the winner of an election cannot set policy alone, but must negotiate with the party in control of another jurisdiction. The model takes as given a status quo policy. Parties in the election differ both in the location of their most preferred policy and in the utility they lose as policy deviates from that ideal, i.e., in the curvature of their utility functions. The less utility lost at the margin by a party, the more willing it is to compromise in negotiations. In standard bargaining models this implies, ceteris paribus, a final outcome further from the party's ideal.

Gallego and Scoones show that the resulting effect on voter choices can be significant. For example, when a right wing party that is extreme in its ideal is nonetheless more willing to compromise than the "moderate" center party, when negotiating with a left wing party the negotiated policy agreements may well be the reverse the parties' ideals. As a result, in the voting equilibrium, support for an extreme party will come from "moderate" voters and support for the center party will come from "extreme" voters.

Policy reversals of this kind do not depend on extreme assumptions about preferences or other elements of the model. Gallego and Scoones' model is very simple, with one election, complete information, a unidimensional policy space, and proportional

representation. Their results demonstrate that an assumption implicit in many political economy models, that for example the more right-leaning a party is, the more right-leaning will be the result of electing them, must be treated with care. It may be true, but if so it depends on special assumptions about the form of party preferences.

1.9 Summary of "A Newton Collocation Method for Solving Dynamic Bargaining Games," by John Duggan and Tasos Kalandrakis

This article studies dynamic bargaining games where the sequence of proposals and votes generates policy outcomes over time. In these dynamic games today's policy becomes tomorrow's status quo, so that the status quo evolves endogenously over time. Forward-looking players anticipate the future consequences of their decisions.

Players' preferences are subject to a transitory preference shock for each period capturing uncertainty about their future policy preferences. In addition, there is noise in the transition from today's outcome to tomorrow's status quo so that players have some small degree of uncertainty about the future state of the game, i.e., about how today's policy decision will be implemented in the future.

In the game, a randomly drawn player proposes a policy, and then all players vote simultaneously to either accept the proposal or maintain the status quo, with players voting to accept when indifferent. The proposal passes if a decisive coalition of players accepts it, and fails otherwise. The period $t+1$ status quo is drawn from a distribution that depends on the policy in period t. In the subgame perfect equilibrium, players use stationary Markov strategies.

This article develops and implements a collocation method to solve for an equilibrium in the dynamic legislative bargaining game of Duggan and Kalandrakis (2008). First the authors formulate stationary equilibria as solutions to a system of functional equations, the unknowns of which are essentially the future expected utilities (or "dynamic utilities") of the players. Then they solve these functional equations using a collocation method, a method for solving functional equations that belongs in the general family of projection methods.

The computational analysis then focuses on the "quasi-discrete" model, in which the uncertainty in the model is continuous (and so the state space of the bargaining game is infinite) but, given the status quo, only a finite number of alternatives may be proposed. This allows the exact solution of a proposer's optimization problem by an exhaustive grid search. The authors establish smoothness properties of the system of collocation equations for the quasi-discrete model, thereby supporting the use of Newton's method to solve for an equilibrium.

They illustrate these techniques for the dynamic core convergence theorem of Duggan and Kalandrakis (2008) in a nine-player, two-dimensional model with negative quadratic preferences. Despite the fact that proposals are endogenous in their model, they show that the equilibrium dynamic utilities can be approximated as the solution to a sufficiently smooth system of collocation equations.

References

Baron D, Ferejohn J (1989) Bargaining in legislatures. Am Political Sci Rev 83(4):1181–1206

Duggan J, Kalandrakis T (2008) Dynamic legislative policy making. Working paper: Rochester University.

Morelli M (1999) Demand competition and policy compromise in legislative bargaining. Am Political Sci Rev 93(4):809–820

Nash J (1950) The bargaining problem. Econometrica 18(2):155–162

Romer T, Rosenthal H (1978) Political resource allocation, controlled agendas, and the status quo. Public Choice 33(48):27–43

Soc Choice Welf
DOI 10.1007/s00355-010-0509-y

A social choice theory of legitimacy

John W. Patty · Elizabeth Maggie Penn

Abstract We develop a formal theory of legitimate collective choice. In our theory a policy choice is legitimate if the process through which the final choice was determined is consistent with some *principle* that can be used to (perhaps partially) rank the potential policy choices. The set of principles in any choice situation is taken to be exogenous, but a *decision-making process* is defined so as to deal with any nontrivial set of principles. Such a process is itself referred to as legitimate if it is guaranteed to select a legitimate outcome for each possible exogenous set of principles. We characterize the class of procedures that are legitimate, prove that legitimate policy decisions consistent with principles always exist and characterize the set of policy decisions that are legitimate for a any given set of principles. As we do not require the principles to be weak orders of the alternatives, our theory provides a notion of legitimacy that can be satisfied even when the guiding principles are potentially cyclic or incomplete. Accordingly, our theory illustrates one nontautological means by which majoritarian principles can be reconciled with legitimacy.

1 Introduction

Legitimacy, the belief that "authorities, institutions, and social arrangements are appropriate, proper, and just,"[1] is a central and longstanding concern of democratic theory. Recently, much attention has been focused on the notion that deliberation

[1] *C.f.* Tyler (2006).

J. W. Patty (✉) · E. M. Penn
Department of Political Science, Washington University in St. Louis, St. Louis, MO, USA
e-mail: jpatty@wustl.edu

E. M. Penn
e-mail: penn@wustl.edu

Published online: 15 January 2011

(or "reason-giving") provides a foundation for legitimacy.[2] While substantive differences exist between their arguments, most deliberative democratic theories agree that a reasonable requirement for a decision to be seen as legitimate is not only that reasons exist to justify the decision, but also that such reasons actually accompany the decision itself.[3] We follow this lead and consider the question of whether and how such justification can be achieved. While this "deliberative democratic" conception of collective decision making has attracted wide and deep interest,[4] an exact definition of what such decision making should produce—in terms of both the final choice and the deliberative process (i.e., the reason-giving) leading up to it—remains to be offered.[5]

In Sect. 2, we lay out the primitives for our theory. Essentially, the theory is a very abstract framework for thinking about the presentation of a final choice along with a principle justifying the choice, possibly combined with a sequence of intermediary steps. Within this framework, we then define two axioms that characterize legitimate decisions. These axioms link legitimacy with decision making that can be rationalized by one of the underlying principles. The first axiom requires that the decision sequence and ultimate policy choice accord with the selected principle and the second axiom requires that the exclusion of policies *not* considered be justified by the selected principle. Thus, legitimacy within our framework consists of the simultaneous two types of consistency between the decision, which we label "internal" and "external" consistency.

Our theory is concerned with legitimacy-provision in complete information environments. In other words, legitimacy in our theory is not based upon finding out which policy is *optimal* as much as it is about *why* the selected policy is legitimately claimed to be so. As our examples through the article are intended to clarify, this task is easy when a collectively accepted principle yields a unique prescription for what policy should be chosen. When such a principle exists, some conclusions are trivial (for example, the conclusion that a legitimate choice exists in that situation). One important conclusion of the article is that legitimacy within these environments does not require that the principles upon which legitimacy might be based are themselves

[2] In most work, the appeal of legitimacy is instrumental in the sense that legitimate policies are those that receive deference from the individuals subject to them. As Habermas (1998) and others have argued, though, deference alone is not sufficient for legitimacy, as the state might use its power to engender deference to illegitimate decisions. The difficulty with inferring legitimacy from deference alone implies a positive role for a normative linkage of legitimacy with reason-giving.

[3] For example, consider the arguments of Cohen (1986); Habermas (1987, 1998); Gutmann and Thompson (1996); Richardson (2003), and Farrar et al. (2010).

[4] For example, broad discussions of the topic are contained in two edited volumes: Bohman and Rehg (1997) and Elster (1998).

[5] This is not to imply that scholars have not attempted to provide formal definitions of deliberative decision-making processes. For example, Miller (1992); Knight and Johnson (1994, 1997); Johnson (1998); Dryzek and List (2003); Meirowitz (2006, 2007); Hafer and Landa (2007); Dickson et al. (2008), and Landa and Meirowitz (2008) all present important angles on the various aspects of a theory of deliberation. In a slightly different vein, the information aggregation work of Grofman and Owen (1986); Ladha (1992); Austen-Smith and Banks (1996); Ladha et al. (1996); Feddersen and Pesendorfer (1998), and others, has focused attention on the behavioral, informational, and institutional determinants of the epistemic performance of deliberation. Our theory focuses on the case of preference aggregation in a world of complete information, as opposed to the question of information aggregation. We return to one angle on this difference in Sect. 5.

"rational." Legitimacy is not trivial to obtain (for example, not all choices can be rendered legitimately, and legitimate decisions must be rendered in specific ways), but is always possible if there is an agreement on the principle by which the collective choice was governed.

As we discuss in Sect. 2.4, the existence and characterization results presented here offer a challenge to empirically minded democratic theorists: democratic institutions that produce legitimate decisions (and the observed processes by which they produce them) will necessarily obscure much, if not all, of the structure of the principles undergirding their choices. Very specifically, inference about the total structure of majority preferences from the observed processing, and ultimate resolution, of democratic choice is simply not possible when the decision is made in a way as to preserve legitimacy. Furthermore, the loss of information about the underlying preferences of individuals occurs for reasons different than those offered by, for example, Arrow (1951); Gibbard (1973), and Satterthwaite (1975), whose arguments are motivated by a desire to aggregate preferences in a manner that maintains consistence across different preference profiles. Here, we do not require that choice be consistent across different preference profiles. Rather, we impose requirements on the consistency of choice and principles (which can be thought of as representing preference profiles). While this change alleviates the impossibility conclusions found in Arrow, Gibbard, and Satterthwaite's seminal offerings, they do not controvert them. Put another way, our notion of legitimacy does not solve what one might call the "Arrovian problem." Democratic choice is fraught with the potential for dynamic inconsistency. Our theory does not solve the problem of cyclic democratic choices through time, but does provide an example of how a given decision might be rationalized in the face of cyclic majority preferences. In fact, we believe this is all that should be asked of a theory of democratic governance—there is simply no reason to expect that democracy is synonymous with the fullest definition of collective rationality, particularly once one grants that individual decision making might be expected to appear rational only to the extent that individuals themselves have well-ordered and stable preferences.

2 Theory

We consider choice in settings with a finite number of feasible alternatives. Let $\mathcal{X}$ denote the set of all finite nonempty sets and, for any set $X \in \mathcal{X}$, let $\mathcal{R}_X$ denote the set of all asymmetric binary relations on X and let $\Pi_X = 2^{\mathcal{R}_X} \backslash \emptyset$ denote the set of all nonempty finite sets of asymmetric binary relations on X.[6] A binary relation is a *tournament* if it is asymmetric and complete.[7] Let $\mathcal{T}_X$ denote the set of tournaments on X. For any binary relation $R \in \mathcal{R}_X$ and any $x \in X$, we use the following notation:

[6] A binary relation on a set X, R, is a subset of $X \times X$. If $(x, y) \in R$, we write xRy and say that "x is related to y through R." A binary relation is finite if it contains a finite number of ordered pairs.

[7] A binary relation $R \subseteq X^2$ is *asymmetric* if $(x, y) \in R$ and $(y, x) \in R$ implies that $x = y$ and it is *complete* if, for all $(x, y) \in X^2$ with $x \neq y$, $(x, y) \in R$ or $(y, x) \in R$.

Table 1 Example of competing principles

Three policies			
Policy	Wealth of i	Wealth of j	Wealth of k
x	10	10	10
y	8	15	11
z	9	9	13

Three possible principles		
Principle	Ranking of x, y, and z	Properties of principle
p_1: Wealth maximization	yp_1zp_1x, yp_1x	*Complete and transitive*
p_2: Majority rule	$zp_2yp_2xp_2z$	*Complete and cyclic*
p_3: "Makes both j and k better off"	yp_3x	*Incomplete*

$$
\begin{aligned}
R(x) &= \{y \in X : (x, y) \in R\} \\
R^{-1}(x) &= \{y \in X : (y, x) \in R\} \\
P(x) &= R(x) \backslash R^{-1}(x) \\
P^{-1}(x) &= R^{-1}(x) \backslash R(x)
\end{aligned}
$$

Finally, for any $Y \subseteq X$ and $P \in \mathcal{R}_X$, $P|_Y$ denotes the binary (sub)relation induced by P on the points in Y.

2.1 Principles

The binary relations contained in Π_X describe the set of *principles* that might be used to justify a collective decision. For the purposes of this article, the term "justify" can be replaced with "explain" or, equivalently, "provide a reason for." For example, consider the choice between three policies, x, y, and z, each representing a different distribution of wealth among the citizenry. For the sake of argument, suppose that there are three individuals, i, j, and k, the wealths of whom under each of the policies are displayed in Table 1. Table 1 also describes three potential principles, two of which are straightforward and yield complete binary relations (wealth maximization and majority preference) and a third (mutual gain for individuals j and k) that yields an incomplete relation: it does not give an unambiguous ranking for all pairs.[8] Supposing that one grants that these principles might be used as "reasons" to justify decisions within a deliberative democratic institution, one must allow for cyclic and/or incomplete principles within a theory of legitimate deliberative decision making.

[8] The third principle in this example yields an incomplete relation due to the fact that it relies essentially on a supermajoritarian rule (or, equivalently, gives "veto power" to some individuals). This type of reliance is frequently sufficient to generate incomplete orderings of the alternatives, but is of course not a necessary condition for such a principle. In particular, if one considers policies with multiple attributes (or characteristics) and principles that consider only strict differences between these characteristics (e.g., the "Elimination by Aspects" heuristic proposed by Tversky 1972), incompleteness results in many situations.

Understanding what principles "look like" in our theory, it is important to note that our definition of Π_X carries two important implications. First, being represented as binary relations is in some sense not without loss of generality: at the risk of belaboring the obvious, this choice implies that principles "rank" pairs of alternatives. Is this realistic? An argument that it is too restrictive is considered in Sect. 5. Nevertheless, we believe this restriction can be motivated by consideration of revealed collective preference when the principle in question is accepted as the proper determinant of choice. As an example, conduct the following thought experiment. Supposing that the collective agreed that some principle p was the correct way to make a choice between some pair of alternatives, x and y. Then if x is chosen, xpy. If y is chosen, ypx, and if neither is chosen, then p does not rank x and y (i.e., p is necessarily incomplete).[9] Using this revealed preference approach and considering all pairs under the operation of the principles p, one can in theory construct a binary relation representation of the principle.

The second important implication of our decision to not restrict the set of principles is normative in nature. For example, by not ruling out cyclic principles, we are allowing for principles that may not actually be well behaved as prescriptions for behavior. As mentioned earlier, majority rule represents a clear candidate for exactly such a principle. But more importantly, perhaps, is the fact that our theory shows how the process of deliberative decision making (at least as conceptualized here) can itself render ill-behaved principles useful as sources of legitimation for policy choices. From a technical standpoint, the inclusion of such principles "makes our job tougher" in the sense that it is obvious that a transitive principle will clearly be consistent with at least one "internally consistent" and "unobjectionable" choice (this is the central point of the classical theory of individual choice). A slightly different, but related objection to our decision not to restrict principles is that inclusion of incomplete principles allows for justifications that are themselves unsatisfying in the sense of not being very demanding. (This point will hopefully become crystal clear later in the article.) For the moment, we simply note that many well-understood principles (e.g., Pareto optimality) often fail to provide complete rankings.

2.2 Legitimate procedures

A pair (X, π) with $\pi \in \Pi_X$ is referred to as a *situation* but note that any binary relation $\pi \in \Pi_X$ for some $X \in \mathcal{X}$ essentially uniquely identifies X, so we will refer simply to $\pi \in \Pi_X$ in place of (X, π). A *decision sequence* is simply a finite sequence of elements within X. An arbitrary decision sequence is denoted by δ and the set of all decision sequences within a set X is denoted by Δ_X. For any $\delta \in \Delta_X$, we represent the ordering of δ by $x \succeq_\delta y$ if x comes after y in δ and $x \succ_\delta y$ if x comes after y in δ and y does not come after x in δ. The last element of δ, denoted by $C(\delta)$, is referred to as the *final decision*.[10] Then, given a finite nonempty set X, a *procedure* on X is

[9] For a theory in which such incompleteness is allowed for in individual preferences, a clear microfoundation for incompleteness of social preference (however aggregated), see Patty (2007).

[10] That is, the final decision is the unique policy $x \in \delta$ such that $x \succeq_\delta y$ for all $y \in \delta$.

a function $\varphi : \Pi_X \to \mathcal{R}_X \times \Delta_X$, with the ith component of $\varphi(\pi)$ being denoted by $\varphi_i(\pi)$,[11] satisfying $\varphi_1(\pi) \in \pi$. The set of all procedures on a set X is denoted by Φ_X. In words, a procedure picks out a principle, $\phi_1(\pi) \in \pi$ to justify the sequence of choices (including, obviously, the final choice), $\phi_2(\pi)$. The following axioms delimit the set of principles and choice sequences that are considered legitimate.

Axiom 1 (*Weak Internal Consistency*) For any $\pi \in \Pi_X$ and each pair $x, y \in \varphi_2(\pi)$ with $x \neq y$,

$$y \succ_{\varphi_2(\pi)} x \;\Rightarrow\; \neg\,[x\varphi_1(\pi)y].$$

Axiom 1 requires that the decision sequence $(\varphi_2(\pi))$ be acyclic with respect to the selected principle, $\varphi_1(\pi)$. It is well known that acyclicity is a minimal requirement for guaranteeing the existence of at least one maximal element (in this case, $C(\varphi_2(\pi))$ is maximal on $\varphi_2(\pi)$ with respect to the binary relation $\varphi_1(\pi)$). Acyclicity of the decision sequence with respect to the selected principle is accordingly a very minimal rationality requirement and, in this setting, can be viewed as a minimal internal consistency requirement (where consistency is with respect to the selected principle). We term this axiom "weak internal consistency" because the final decision is not inconsistent with the decision sequence according to the stated principle, but the final choice is not necessarily "superior" to every alternative in the decision sequence according to the principle, *per se*.

Axiom 2 (*Weak External Justification*) For any $\pi \in \Pi_X$ and any $y \notin \varphi(\pi)$,

$$y\varphi_1(\pi)C(\varphi_2(\pi)) \Rightarrow C(\varphi_2(\pi)) \succ_{\varphi_2(\pi)} z\varphi_1(\pi)y \quad \textit{for some } z \in \varphi_2(\pi).$$

Axiom 2 requires that the decision sequence be *externally stable* (von Neumann and Morgenstern 1947) with respect to the selected principle. As we return to later in the article, requiring external stability is attractive because it rules out attempts to satisfy Axiom 1 simply by selecting a single choice with no accompanying decision sequence. We term this "external justification" because it represents the requirement that, when an alternative not considered is proposed as a replacement for the final choice, a justification for the final choice can be constructed using the underlying decision sequence and the selected principle to obtain a special form of counterargument. We term this axiom "*weak* external justification" because it does not impose any length restriction on the justification. In particular, while the justification for not choosing an element not in the decision sequence will be finite, it might be quite long.

Finally, for any set $X \in \mathcal{X}$, let $\mathcal{L}(X) \subseteq \Phi_X$ denote the set of procedures for X that satisfy Axioms 1 and 2. We will refer to $\mathcal{L}(X)$ as the set of *legitimate procedures*. The next result establishes that the set of legitimate procedures is nonempty.

Theorem 3 *For any set $X \in \mathcal{X}$, $\mathcal{L}(X) \neq \emptyset$.*

[11] Thus, $\varphi(\pi) = (\varphi_1(\pi), \varphi_2(\pi))$, with $\varphi_1(\pi) \in \Pi_X$ and $\varphi_2(\pi) \in \Delta_X$.

Proof The proof is constructive. Fix any finite and nonempty set X and any nonempty set of binary relations $\pi \in \Pi_X$. Choose any binary relation $P \in \pi$ and any element $x_1 \in X$. If $P(x_1) = \emptyset$, let $\varphi(\pi) = (P, (x_1))$. It follows that $\varphi(\pi)$ satisfies Axioms 1 and 2: Axiom 1 is satisfied trivially and Axiom 2 is satisfied by the supposition that $P(x_1) = \emptyset$.

If $P(x_1) \neq \emptyset$, choose any element $x_2 \in P(x_1)$. If $P(x_2) \subset P^{-1}(x_1)$, let $\varphi(\pi) = (P, (x_1, x_2))$. Axiom 1 is satisfied since $x_2 P x_1$. To see that Axiom 2 is satisfied, notice that $z \in P(x_2) \Rightarrow x_1 P z$.

Otherwise, let x_3 be any element of $P(x_2) \backslash P^{-1}(x_1)$. If $P(x_3) \subset P^{-1}(x_1) \cup P^{-1}(x_2)$, let $\varphi(\pi) = (P, (x_1, x_2, x_3))$. Axiom 1 is satisfied since $x_3 P x_2$ and $x_3 \notin P^{-1}(x_1)$. To see that Axiom 2 is satisfied, notice that $z \in P(x_3) \Rightarrow [x_1 P z$ or $x_2 P z]$.

Finally, note that the iterative procedure can be continued at most a finite number of times since X is finite (and, furthermore, Π_X is finite, so that satisfaction of Axioms 1 and 2 at a single arbitrary π imply satisfaction for all π). Once this iterative construction is complete, we have constructed a φ satisfying Axioms 1 and 2. Accordingly, note that X and π are each arbitrary up to the requirement that $X \in \mathcal{X}$ (X is nonempty and finite) and $\pi \in \Pi_X$ is nonempty, so that the result follows: $\mathcal{L}(X) \neq \emptyset$. □

Before continuing to the characterizing legitimate procedures and what they produce, we note two simple facts about satisfying Axioms 1 and 2.[12]

Fact 4 *If a procedure φ returns exactly one policy in its decision sequence for all situations π, it necessarily satisfies Axiom 1.*

The notable feature of Fact 4 is what it doesn't say: internal justification is guaranteed to be satisfied by any one-element decision sequence *regardless of the principle chosen to justify it.* Logically speaking, this is a tautology. In terms of the substantive application of the theory, however, this extreme case indicates one type of "cost" of lengthy decision sequences. Allowing free-ranging discussion and deliberation within a collective can[13] create a situation in which no final choice can be reconciled with a principle. This is demonstrated by the following example.

Example 5 (Too Much Deliberation) Suppose that there are three policies, A, B, and C, and one principle, $\pi = \{p\}$, with $ApBpCpA$ (p is a Condorcet cycle among the three alternatives). Now consider a procedure φ that returns, say, $\varphi(\pi) = (p, (C, B, A))$. Clearly, such a procedure does not satisfy Axiom 1 at $\pi = \{p\}$, because the final choice, A, loses to an alternative considered earlier in the decision sequence (i.e., C). Furthermore, this is true for *any* procedure φ that returns a decision sequence containing all elements A, B, and C at $\pi = \{p\}$. As Theorem 3 implies, however, there are procedures that satisfy Axiom 1 at $\pi = \{p\}$: consider $\varphi' = (p, (B, A))$, which satisfies both Axioms 1 and 2 at $\pi = \{p\}$.

Example 5 encompasses a couple of important points. The first of these, of course, is that too much discussion can be "bad" for principle-based deliberative decision

[12] These two facts are in some sense analogous to the discussion by Buchanan and Tullock (1962) of the virtues of majoritarian versus those of more demanding supermajoritarian rules in collective choice.

[13] But need not, as illustrated in Example 8 below.

making because of the possibility that the discussion itself may extend so far as to make it impossible to rationalize *any* final choice. This implies that limits on the scope of discussion may be central to the provision of legitimate policy decisions. It also points out that achieving legitimate decision making may require abstraction from certain comparisons, similar to what Cass Sunstein terms an *incompletely theorized agreement* (Sunstein 2001).

Fact 6 *If a procedure φ returns every policy at least once in its decision sequence for all situations π, it necessarily satisfies Axiom 2.*

The import of Fact 6 is best seen by considering when a procedure φ can satisfy Axiom 2 through the offering of a unitary decision sequence (i.e., by simply pronouncing a decision, (x), for some $x \in X$). This is described in the following example.

Example 7 (Trivial Deliberation) Suppose that there are three policies, A, B, and C, and one principle, $\pi = \{p\}$, with $ApBpC$ and ApC (p is complete and transitive ordering of the three alternatives). Now consider a procedure φ that returns, say, $\varphi(\pi) = (p, (C, B, A))$. Clearly, such a procedure satisfies both Axioms 1 and 2 at $\pi = \{p\}$. Furthermore, note that the same is true for a procedure φ' for which $\varphi'(\pi) = (p, (A))$.

Example 7 demonstrates that when there is a complete and transitive ordering of the alternatives available as a principle, deliberation can yield a legitimate choice even when all of the alternatives are included but—precisely because of this fact—deliberation is also unnecessary for legitimation of the final choice. In other words, the principle itself yields a unique conclusion. Note, of course, that deliberative decision making is not necessarily trivial whenever there is some complete and transitive principle in the set of possible principles, π. Specifically, the choice of principle is still potentially nontrivial. We return to this point in Example 8.

2.3 Legitimate choices and rationales

A key element of our theory is its incorporation of the (potentially vacuous) decision sequence leading up to the final policy choice. Combined with the selected principle, the decision sequence and the final choice generate what we call a *rationale* for the final policy choice. In order to model rationale formally, define the following function for any procedure φ:

$$\rho(\varphi(\pi)) \equiv \varphi_1(\pi)|_{\varphi_2(\pi)}.$$

Thus, $\rho(\varphi)$ is a function that produces—for each nonempty $\pi \in \Pi_X$—the binary relation induced by the decision sequence, $\varphi_2(\pi)$, on the chosen principle $\varphi_1(\pi)$. Thus, $\rho(\varphi(\pi)) \in \Pi_X$ for all φ and all π. For any $X \in \mathcal{X}$, any $\pi \in \Pi_X$, and $\varphi \in \Phi_X$, we refer to $\rho(\varphi(\pi))$ as the rationale produced by φ at π. As this is a central component of our theory, it is useful to consider what a rationale looks like. We do this through three examples.

Example 8 (Three Complete & Transitive Principles) Suppose that there are three policies, A, B, and C, and three principles, $\pi = \{p_A, p_B, p_C\}$. Each of the three principles is complete and transitive, but they each have a different top-ranked alternative. In particular, principle p_A ranks alternative A the highest, principle p_B ranks B the highest, and principle p_C ranks C the highest. Suppose also that Bp_1C. Now consider a procedure φ^1 for which $\varphi^1(\pi) = (p_A, (B, A))$, in which the final choice in this situation is A, which can be phrased in words as "A is determined to be better than B by virtue of principle p_1." The corresponding rationale for the choice of A is simply $A\rho(\varphi^1(\pi))B$. A procedure that is equivalent in many ways to this φ in this situation is given by $\varphi^2(\pi) = (p_A, (C, B, A))$: the final choice is still A under this procedure, but the rationale for the decision is richer than the one provided by φ^1. Specifically, the rationale under φ^2 at π is $A\rho(\varphi^2(\pi))B\rho(\varphi^2(\pi))C \;\; A\rho(\varphi^2(\pi))C$, which can be described as "A is determined to be better than B and both A and B are determined to be better than C by virtue of principle p_1."

Both φ^1 and φ^2 satisfy Axioms 1 and 2 at π. Note, however, that satisfaction of these Axioms in this case can be done without provision of a rationale. In particular, note that $\varphi^3(\pi) = (p_A, (A))$, with $\rho(\varphi^3(\pi)) = \{\}$, satisfies both Axioms 1 and 2. This is because the chosen principle is a transitive ordering of the policies. Indeed, given that p_A is also a complete ordering of the alternatives, the choice of p_A as the principle is equivalent to selecting A as the final choice. Note that even with such well-behaved principles, the proper choice of principle is central to the satisfaction of Axioms 1 and 2: $\varphi^4(\pi) = (p_B, (A))$ does not satisfy Axiom 2 and $\varphi^5(\pi) = (p_B, (C, B, A))$ does not satisfy Axiom 1.

Of course, in many situations of interest, no complete ordering of the policies is available. In these situations, Axiom 2 is easier to satisfy, a point to which we return below (Fact 13). In fact, as Example 9 makes clear, reliance on a principle that is "far from complete" in the sense of ranking only a few of the pairs of policies will generally make it possible to rationalize a large set of final choices. The example also illustrates, however, the possibility of differentiating between legitimate policies on the extent of the rationale provided for it, where "extent" can be measured by the number of policies both included in the decision sequence and the number of policies not included in the decision sequence that are related to some element in the sequence through the chosen principle.

Example 9 (Rationales With An Incomplete Principle) Suppose that there are three policies, A, B, and C, and one principle, $\pi = \{p\}$, defined as follows: ApB. At π, any legitimate procedure φ must satisfy

$$\varphi(\pi) \in \{(p, (B, A)), (p, (A)), (p, (C)), (p, (C, B, A)), \\ (p, (B, C, A)), (p, (B, A, C))\}.$$

In other words, a legitimate procedure cannot return B as the final policy in this case, since A dominates B under principle p, but a legitimate procedure can return C as the final outcome.

Finally, when utilizing a cyclic principle, legitimate decision making involves the construction of a decision sequence that, put loosely, "makes sense" of the principle by choosing a sequence of policies that respect the principle's ordering while justifying the exclusion of all those policies not selected.

Example 10 (Rationales With A Cyclic Principle) Consider the situation described in Example 5. At π, any legitimate procedure φ must satisfy

$$\varphi(\pi) \in \{(p, (B, A)), (p, (C, B)), (p, (A, C))\}.$$

There are then three rationales that can be provided by a legitimate procedure in this canonical situation: one for each possible final choice, and each corresponding to one of the pairwise comparisons in the underlying principle, p. In this case, the final policy choice of a legitimate decision sequence is completely determined by the initial choice.

For any binary relations $\pi, \theta \in \mathcal{R}_X$, π is an *acyclic subgraph* of θ if π is acyclic and $\pi \subseteq \theta$. For any finite set X and binary relation $\pi \in \mathcal{R}_X$, $x \in X$ is a *maximal element* of π if there is no $y \in X$ such that $y\pi x$. For any $x \in X$ and $P \in \mathcal{R}_X$, let $M^x(P)$ denote the set of acyclic subgraphs of P in which x is maximal and define the following set:

$$\bar{M}^x(P) \equiv \{p \in M^x(P) : \forall y \in X, \forall q \in M^y(P), p \not\subset q\}.$$

The set $\bar{M}^x(P)$ (which may be empty, though not for every $x \in X$ unless $P = \emptyset$) contains each acyclic subgraph $p \in P$ in which x is the maximal element and such that any strict supergraph of p, $q \supset p$ is either not an acyclic subgraph of P or also has x as its maximal element (i.e., $q \in M^x(P)$). In words, these are the acyclic subgraphs of P such that x is top-ranked and the addition of any $y \in P(x)$ to the subgraph will create a new graph that is not acyclic.[14]

For any $X \in \mathcal{X}$ and any principle $p \in \mathcal{R}_X$, define the following covering notion, $\mathbf{C}_p$, induced by p on X: for all $(x, y) \in X^2$,

$$x\mathbf{C}_p y \Leftrightarrow x \in p(y) \text{ and } p(x) \subset p(y). \tag{1}$$

For any $p \in \mathcal{R}_X$, let $UC(p)$ denote the set of alternatives that are not covered by any other alternative under p. We refer to this as the *uncovered set* under p.[15] The next result states that legitimate procedures must return final choices within the uncovered set of their selected principle.

[14] For any binary relation p on a set X, a subset of $Y \subseteq X$ for which $p|_Y$ is *transitive* is referred to as a *chain*. Note that $M^x(P)$ contains all chains that are subgraphs of P with x top-ranked, and generally is a strict superset of that (possibly empty) set of subgraphs of P.

[15] As Bordes (1983); Banks and Bordes (1988); Dutta and Laslier (1999); Peris and Subiza (1999); Penn (2006), and others have noted, there are multiple closely related (but not equivalent) definitions of covering that one might use in this setting. The definition described by (1) corresponds to the C_u notion of covering as defined in Banks and Bordes (1988).

Theorem 11 *For any $X \in \mathcal{X}$ and any principle $p \in \mathcal{R}_X$,*

$$\bigcup_{\varphi \in \mathcal{L}(X)} C(\{p\}; \varphi) \subseteq UC(p).$$

Proof Fix $X \in \mathcal{X}$ and, for the purpose of obtaining a contradiction, suppose that there exists $\varphi \in \mathcal{L}(X)$ such that $x\mathbf{C}_{\varphi_1(\pi)}C(\{p\}; \varphi)$ for some $\pi \in \Pi_X$. Supposing that $x \in \varphi_2(\pi)$ contradicts satisfaction of Axiom 1 by φ. Supposing that $x \notin \varphi_2(\pi)$, note that $x\mathbf{C}_{\varphi_1(\pi)}C(\{p\}; \varphi)$ implies that there is no $y\varphi_1(\pi)x$, contradicting satisfaction of Axiom 2 by φ. Accordingly, supposing that there exists $x\mathbf{C}_{\varphi_1(\pi)}C(\{p\}; \varphi)$ for some $\pi \in \Pi_X$ implies that φ cannot satisfy both Axioms 1 and 2 and thereby contradicts the supposition that $\varphi \in \mathcal{L}(X)$, yielding the result. □

We can say more in the special case where π contains a single tournament. In words, when the unique principle that may be used to provide rationales is strict majority preference relation, the set of legitimate choices is identical to the Banks set (Banks 1985). For any $T \in \mathcal{T}$, let $B(T)$ denote the *Banks set* of π (Banks 1985).

Corollary 12 *For any $X \in \mathcal{X}$ and any tournament $T \in \mathcal{T}_X$,*

$$\bigcup_{\varphi \in \mathcal{L}(X)} C(\{T\}; \varphi) = B(T).$$

Corollary 12 implies two key facts about the outcomes that can be produced by a legitimate procedure. First, it is well-known that $B(T) \subsetneq UC(T)$ for some tournaments T.[16] In other words, Theorem 11 cannot be strengthened to provide an equivalence result. Furthermore, Theorem 11 is "tight" in a specific sense: for any principle p that is a "cyclone" on X, $UC(p) = B(p)$.[17] Thus, there are principles that are tournaments for which there is no "wiggle room" between Theorem 11 and Corollary 12.

In addition, Corollary 12 is illustrative in light of the motivations behind Banks's original work: a desire to characterize the set of outcomes that can be obtained through sophisticated voting over amendment agenda.[18] Applying the sophisticated voting algorithm to (ordered) amendment agenda essentially "creates" a "sophisticated equivalent agenda" that rational voters use to evaluate their vote choices at each stage.[19] For any amendment agenda that include all feasible alternatives, its sophisticated equivalent agenda satisfy both Axioms 1 and 2. Banks's seminal contribution was to show that these completely describe what might be achieved through sophisticated voting. In so doing, he also demonstrated one institutional approach to producing a legitimate majoritarian decision.

[16] For example, Banks (1985); Dutta (1988); Schwartz (1990), and Laslier (1997).

[17] A cyclone is a class of tournament that represents the generalization of the Condorcet cycle to sets of more than three alternatives. See Laslier (1997).

[18] For more on amendment agendas, see Ordeshook and Schwartz (1987) and Miller (1995).

[19] For a discussion of sophisticated voting, see Farquharson (1969). For the derivation and explanation of the sophisticated voting algorithm and "sophisticated equivalents," see Shepsle and Weingast (1984) and Banks (1985).

2.4 A brief defense of the legitimizing role of majority rule

While there have been many positive arguments in favor of majority rule institutions,[20] the possibility of majority rule cycles has nonetheless challenged those who promote the use of majoritarian principles in collective choice.[21] In a nutshell, the fact that majority preferences need not be transitive has led to extensive worries about "stability," "legitimacy," and/or "rationality" of democratic decision making.

What is somewhat surprising about Corollary 12 is that the case in which π consists of a tournament provides the *most* refined set of legitimate policy choices. In other words, in a specific sense, not only can potentially cyclic majority preference be rationalized—a priori reliance on majority preference represents the most demanding criterion that one might require for a collective choice to be considered legitimate.

Fact 13 *For any $p \in \Pi_X$ and $\varphi \in \mathcal{L}(X)$, there exists $t \in \mathcal{T}_X$ with $p \subseteq t$ and $\varphi' \in \mathcal{L}(X)$ such that $\varphi'(\pi) = (t, \varphi_2(\{p\}))$ for all π with $t \in \pi$.*

Proof Fix $p \in \Pi_X$ and $\varphi \in \mathcal{L}(X)$. Construct $t \in \mathcal{T}_X$ as follows:

1. For all $(y, z) \in p$, fix ytz if ypz and zpy otherwise.
2. For all (y, z) with $y \in \varphi_2(\{p\})$, $z \notin \varphi_2(\{p\})$, fix ytz.
3. For all (y, z) with $y, z \notin \varphi_2(\{p\})$, fix ytz or zty arbitrarily.

It can be verified that t so defined is a tournament and, more importantly, that $\varphi_2(\{p\})$ is an externally stable chain with respect to t. □

In addition to establishing that tournaments provide in a certain sense the most restriction on the set of outcomes that might be rendered legitimate, Fact 13 also implies something important about the observability of cycles when decisions are made by legitimate procedures. Namely, Fact 13 implies that legitimate decisions from acyclic principles are always consistent with some cyclic principles as well. Accordingly, the observation of seemingly transitive decisions is not sufficient evidence that the underlying principles guiding that choice are themselves transitive. This point represents a serious challenge to the claims of scholars that the well-ordered operation of democratic choice in the real world can be used to refute claims that majority preference cycles are present or, perhaps, relevant to the study of democracy or democratic institutions.[22]

[20] A few examples include May (1952); Downs (1957); Grofman and Feld (1988); Austen-Smith and Banks (1996); Feddersen and Pesendorfer (1998); List and Goodin (2001); Mackie (2003); Crombez et al. (2006); McGann (2006).

[21] For example, consider the different arguments in Buchanan and Tullock (1962); Riker (1980, 1986, 1996); Miller (1980); Rubinstein (1980); Shepsle and Weingast (1981, 1984); Tullock (1981); Niou and Ordeshook (1985), and McKelvey (1986).

[22] Foremost among such work are the arguments offered by Mackie (2003), who criticizes the focus on majority preference cycles by social choice and rational choice theorists. In many ways, we are sympathetic with Mackie's goal insofar as he is skeptical of claims by some that democratic choice fails to provide some form of legitimation for the selected policies. In spite of this, however, his arguments fail to provide proper accounting for the importance of the institutional details of democratic choice.

3 Principles as "structure"

Examples 7 and 8 both illustrate the fact that principle-driven deliberation per se (at least as portrayed here) is arguably less interesting when a "nice" principle—a principle that, once chosen, essentially completely structure and uniquely determine the outcome of deliberation—is available and used to guide the decision sequence. To the degree that Axioms 1 and 2 do not impose any consistency requirements across the subsets $\pi, \pi' \in \Pi_X$, legitimate procedures retain complete flexibility with respect to how principles are chosen in different situations. We think it is useful to consider for a moment the connections between this approach and the concept of *structure induced equilibrium* or, *SIE* (Shepsle 1979). The theory of SIE imposes institutional structure on "pure" preference-based collective choice and provides a transparent framework for reaching two central facts about collective choice. First, SIE demonstrates clearly that "stable" (or perhaps "predictable") collective choice is possible with the right choice of institutional structure. Second, and more importantly, the final policy choice will in general depend upon the details of the institutional structure imposed *ex ante*.

The choice of principle in any situation, $\pi, \varphi_1(\pi)$, is in a very real sense analogous to the choice of structure within the SIE framework. There are two differences between our theory and the SIE approach, one technical and one substantive. The technical difference is that an SIE prediction is a single policy choice, determined as the top of the collective preference relation (which will in general be incomplete) generated by filtering the individuals' preferences "through" the institutional structure. An implication of this is that the set of SIE outcomes may be empty, whereas Theorem 3 guarantees existence of legitimate outcomes. To be fair, the theory presented here offers little in terms of predicting which principle will be chosen by a legitimate procedure. Retaining the focus on the analogy with SIE, this point echoed the response of Riker (1980) to Shepsle (1979) regarding how much one can rely upon (presumably endogenously chosen) institutions to structure collective choice. Put slightly differently, just as Shepsle (1979) offers little guidance about how institutions *are* chosen, we offer little guidance about the choice of principles. One advantage of the approach here, however, is that once one determines how principles are (or perhaps ought to be) chosen in different situations, our theory offers a nonempty set of predicted policy choices as well as a prediction about the possible rationales justifying the chosen policy.

The predicted rationales represent the main substantive difference between the two concepts. SIE is silent about how the collective reaches its final decision. In the framework examined here, however, a decision sequence may be provided in conjunction with the final policy choice. This provision is motivated by two somewhat related purposes: first, the decision sequence is descriptively analogous to many depictions of the operation of a deliberative democratic institution. Second, for reasons not entirely dissonant with some of the arguments offered in favor of deliberative democracy, this accompanying decision sequence can be judged in relation to the final policy choices and the group's principles. Achieving consonance between the decision sequence, the group's principles, and the final choice is a nontrivial task (i.e., there are policy choices and decision sequences that do not agree with one another and, furthermore, there can exist a policy choice for which there is no principle and decision sequence that justify

its selection). Accordingly, the deliberative aspect of the theory not only requires an appropriate choice of structure (e.g., a proper sequencing of intermediary steps given the principle being utilized), but also refines the set of outcomes that might be chosen in a legitimate fashion.

4 Strengthened notions of legitimacy

As discussed earlier, Axioms 1 and 2 are not particularly strong. For example, Axiom 1 does not require that the rationale generated to justify the policy be transitive: this implies that the conclusion may not be related to the all of the steps of the rationale through the selected principle (though, of course, none of the steps of rationale may contradict the final policy choice). Somewhat similarly, Axiom 2 does not require that the justification offered for the failure to choose an alternative not considered be straight-forward in any sense. These points raise the question of whether Axioms 1 and 2 can be strengthened in pursuit of a more demanding (and, arguably, more appealing) notion of legitimacy. To answer this question (unfortunately in the negative), consider the following strengthened versions of Axioms 1 and 2.

Axiom 1* (*Strong Internal Consistency*) For any $\pi \in \Pi_X$ and each $x \in \varphi_2(\pi)$,

$$x \neq C(\varphi_2(\pi)) \;\Rightarrow\; C(\varphi_2(\pi))\varphi_1(\pi)x.$$

Axiom 2* (*Strong External Justification*) For any $\pi \in \Pi_X$ and any $y \notin \varphi(\pi)$,

$$y\varphi_1(\pi)C(\varphi_2(\pi)) \Rightarrow C(\varphi_2(\pi))\varphi_1(\pi)z\varphi_1(\pi)y \textit{ for some } z \in \varphi_2(\pi).$$

For any $X \in \mathcal{X}$, let $\mathcal{L}^*(X) \subset \mathcal{L}(X)$ denote the set of procedures satisfying Axioms 1* and 2*. The next result demonstrates that these axioms cannot be satisfied. In particular, some principles provide so little structure that *no* decision sequence will satisfy the conditions of both Axioms 1* and 2*.

Theorem 14 *For all $X \in \mathcal{X}$ such that $|X| > 3$, $\mathcal{L}^*(X) = \emptyset$.*

Proof Note that demonstrating the result for $X \in \mathcal{X}$ with $|X| = 4$ is sufficient, since $C(\varphi_2(\pi)) \in \mathcal{L}(X)$ must be in the minimal dominant set of $\varphi_1(\pi)$. Accordingly, for X with $|X| > 4$, simply set π to be a single principle with a four element minimal dominant set upon which the subgraph induced by the principle is equivalent to the principle examined below.

Let $X = \{a, b, c, d\}$ and consider the following principle P:

1. aPb, bPc, cPd, dPa, and
2. No other pairs are in P.

Let $\pi = \{P\}$. Any procedure φ for which $\varphi_2(\pi)$ is a singleton violates Axiom 2*. Similarly for any φ for which $\varphi_2(\pi)$ has two elements (in addition to possibly violating Axiom 1*). Since P is not itself acyclic on $\{a, b, c, d\}$, $\varphi(\pi)$ must contain exactly three

alternatives. However, none of the four candidates for φ_2 in this case are consistent with Axiom 1*. □

We now note that Axiom 1* alone is sufficient to obtain Theorem 14 in the sense that strengthening only Axiom 1 but leaving Axiom 2 unchanged results in an equivalent notion of legitimacy to imposing both Axioms 1* and 2*. For any $X \in \mathcal{X}$, let $\mathcal{L}_1^*(X) \subset \mathcal{L}(X)$ denote the set of procedures satisfying Axioms 1* and 2.

Proposition 15 *For any $X \in \mathcal{X}$,*

$$\mathcal{L}_1^*(X) = \mathcal{L}^*(X).$$

Proof Consider first a procedure $\varphi \in \mathcal{L}_1^*(X)$. We show that $\varphi \in \mathcal{L}^*(X)$ (the opposite inclusion follows immediately). Fix $X \in \mathcal{X}$, consider any $\pi \in \Pi_X$ and any $y \notin \varphi_2(\pi)$. Satisfaction of 1* implies that $C(\varphi_2(\pi))\varphi_1(\pi)z$ for all $z \in \varphi_2(\pi)$ such that $z \neq C(\varphi_2(\pi))$. Accordingly, by the presumed satisfaction of Axiom 2 by φ,

$$y\varphi_1(\pi)C(\varphi_2(\pi)) \Rightarrow C(\varphi_2(\pi))\varphi_1(\pi)z\varphi_1(\pi)y \text{ for some } z \in \varphi_2(\pi),$$

which is equivalent to Axiom 2*, so that $\varphi \in \mathcal{L}^*(X)$. □

Now let $\mathcal{L}_2^*(X) \subset \mathcal{L}(X)$ denote the set of procedures satisfying Axioms 1 and 2*. The following result states that imposing Axiom 2* along with Axiom 1 yields a notion of legitimacy that cannot be satisfied for all sets of principles. This is a slightly different result than Proposition 15, because $\mathcal{L}_2^*$ is actually a strict superset of $\mathcal{L}^*$, but the "enlargement" accomplished by relaxing Axiom 1* is not sufficient to alleviate the nonexistence conclusion of Theorem 14.

Proposition 16 *For all $X \in \mathcal{X}$ such that $|X| > 3$, $\mathcal{L}_2^*(X) = \emptyset$.*

Proof Consider the example provided in the proof of Theorem 14. Note that φ_2 must contain exactly three alternatives and then note that there is no alternative that is ranked above two other alternatives by the principle P. □

5 Conclusion

We have presented a formal theory of legitimacy in collective choice situations. We define a policy choice as legitimate if the process through which it was rendered is consistent with an exogenously given principle. Consistency of the process with the principle is based on the ability to partially rank the potential policy choices. We have shown that the set of legitimate choices is always nonempty and have characterized the class of procedures that are legitimate. Finally, we have argued that, because the principles need not be weak orders of the alternatives, the theory represents a notion of legitimacy that can be satisfied even when the guiding principles are potentially cyclic or incomplete. Indeed, in addition to providing a nontautological means by which majoritarian decisions can be rendered legitimate and in contrast with the tenor of the arguments of Riker (1988), the theory also highlights the power of majority preference in refining the set of legitimate outcomes. Of course, the theory offered here can, and we believe should, be extended. Accordingly, before concluding, we briefly discuss two potential directions for such work.

5.1 Incorporating uncertainty

As noted in the introduction, this theory considers the question of legitimate choice when the institution can in a sense announce a common recognized "principle" as the basis for justifying the choice. This is, in our minds at least, a reasonable place to start a search for a theory of legitimate policymaking. In particular, the supposition that a theory of legitimate decisions is possible would seem to require the acceptance that such principle can, at least in theory, be agreed upon by the group subjected to the decision. This presumption does not render the problem trivial by any means so long as one does not impose any a priori structure on the principles that might be available for use.

Nonetheless, we believe that situations in which the complete information assumption is implausible are not entirely outside of the theory presented here. In particular, one can conceive of a larger conception of deliberation as also including a process by which the set of principles themselves are discovered (or, perhaps, "constructed" in some way). Extensions of the theory in this direction raise some very interesting questions. Unfortunately, this interesting topic is beyond the scope of the current project. Accordingly, we simply note that these questions highlight an obvious but often overlooked aspect of theories like that provided here: not all real-world institutions are legitimate. Our theory is not "predictive" in the sense of positing anything to the contrary. Instead, the existence results here are properly thought of as possibility results: they indicate that deliberative choice can proceed in an orderly fashion without imposing the assumption that deliberation leads to orderly choice.[23]

5.2 Beyond binary relations

A possible complaint about our theory is that it conceptualizes principles as binary relations between alternatives. The main implication of this aspect of our theory is that the ranking of two alternatives by any given principle remains the same regardless of the set of feasible alternatives. This presumption, a traditional one in the theory of individual choice, has been repeatedly called into question on empirical grounds. However, our allowance for multiple incomplete and cyclic principles essentially eliminates the real restrictiveness of the assumption of a fixed binary relation in the classical setting of individual choice. As Examples 9 and 10 illustrate, the observation of phenomena such as context or path dependence is consistent with the theory here—indeed, our theory provides an argument as to why such phenomena might actually be necessary for legitimate governance. Shining a light on the central root of the dilemma facing those who pursue legitimate democratic governance, one has no guarantee that majority will is well-behaved *but it is clearly sensible to presume that it is well-defined on pairs of alternatives*. Finally, the theory does not even require the presumption that

[23] This point calls into question the need for what one might term "preference alignment arguments" forwarded by, among others, Grofman and Feld (1988); Miller (1992); Knight and Johnson (1994), and Dryzek and List (2003). In addition, Penn et al. (n.d.) provide several reasons to believe that such arguments, even if true, are not sufficient to guarantee sincere revelation of individual preferences in pursuit of the possibilities of well-ordered collective choice.

majority will is fixed for any given pair of policies: allowing for a vacillating majority will is within the framework defined here, because the set of admissible principles may contain more than one tournament.

Acknowledgments We thank Jake Bowers, Hans Noel, Norman Schofield, Ken Shotts, Cara Wong, and audience members at the University of California, Berkeley and University of Illinois for comments and suggestions. As usual, all remaining errors are our fault.

References

Arrow KJ (1951) Social choice and individual values. Wiley, New York
Austen-Smith D, Banks JS (1996) Information aggregation, rationality and the Condorcet Jury theorem. Am Polit Sci Rev 90:3445
Banks JS (1985) Sophisticated voting outcomes and agenda control. Soc Choice Welf 1:295–306
Banks JS, Bordes GA (1988) Voting games, indifference, and consistent sequential choice rules. Soc Choice Welf 5(1):31–44
Bohman J, Rehg W (1997) Deliberative democracy: essays on reason and politics. MIT Press, Cambridge
Bordes G (1983) On the possibility of reasonable consistent majoritarian choice: some positive results. J Econ Theory 31:122–132
Buchanan JM, Tullock G (1962) The calculus of consent: logical foundations of constitutional democracy. University of Michigan Press, Ann Arbor
Cohen J (1986) An epistemic conception of democracy. Ethics 97(1):6–25
Crombez C, Groseclose T, Krehbiel K (2006) Gatekeeping. J Polit 68:322–334
Dickson E, Hafer C, Landa D (2008) Cognition and strategy: a deliberation experiment. J Polit 70:974–989
Downs A (1957) An economic theory of democracy. Harper and Row, New York
Dryzek JS, List C (2003) Social choice theory and deliberative democracy: a reconciliation. Br J Polit Sci 33:1–28
Dutta B (1988) Covering sets and a new Condorcet choice correspondence. J Econ Theory 44:63–80
Dutta B, Laslier JF (1999) Comparison functions and choice correspondences. Soc Choice Welf 16:513–532
Elster J (1998) Deliberative democracy. Cambridge University Press, New York
Farquharson R (1969) The theory of voting. Yale University Press, New Haven
Farrar C, Fishkin J, Green D, List C, Luskin R, Paluck E (2010) Disaggregating deliberations effects: an experiment within a deliberative poll. Br J Polit Sci 40:333–347
Feddersen TJ, Pesendorfer W (1998) Convicting the innocent: the inferiority of unanimous jury verdicts. Am Polit Sci Rev 92:23–36
Gibbard A (1973) Manipulation of voting schemes: a general result. Econometrica 41(4):587–601
Grofman B, Feld SL (1988) Rousseau's general will: a Condorcetian perspective. Am Polit Sci Rev 82(2):567–576
Grofman BN, Owen G (eds) (1986) Information pooling and group decision making. JAI Press, Greenwich
Gutmann A, Thompson DF (1996) Democracy and disagreement. Harvard University Press, Cambridge
Habermas J (1987) The theory of communicative action. Beacon, Boston
Habermas J (1998) Between facts and norms: contributions to a discourse theory of law and democracy. MIT, Cambridge
Hafer C, Landa D (2007) Deliberation as self-discovery and the institutions for political speech. J Theor Polit 19:329–360
Johnson J (1998) Arguing for deliberation. In: Elster J (ed) Deliberative democracy. Cambridge University Press, New York
Knight J, Johnson J (1994) Aggregation and deliberation: on the possibility of democratic legitimacy. Polit Theory 22:277–296
Knight J, Johnson J (1997) What sort of equality does democratic deliberation require?. In: Bohman J, Rehg W (eds) Deliberative democracy. MIT, Cambridge
Ladha K (1992) The Condorcet Jury theorem, free speech and correlated votes. Am J Polit Sci 36:617–634
Ladha K, Miller G, Oppenheimer J (1996) Information aggregation by majority rule: theory and experiments. Typescript, University of Maryland

Landa D, Meirowitz A (2008) Game theory, information, and deliberative democracy. Am J Polit Sci 53:427–444
Laslier JF (1997) Tournament solutions and majority voting. Springer-Verlag, New York
List C, Goodin RE (2001) Epistemic democracy: generalizing the Condorcet Jury theorem. J Polit Philos 9:277–306
Mackie G (2003) Democracy defended. Cambridge University Press, New York, NY
May K (1952) A set of independent necessary and sufficient conditions for simple majority decision. Econometrica 20:680–684
McGann AJ (2006) The logic of democracy: reconciling equality, deliberation, and minority protection. University of Michigan Press, Ann Arbor
McKelvey RD (1986) Covering, dominance, and institution-free properties of social choice. Am J Polit Sci 30:283–314
Meirowitz A (2006) Designing institutions to aggregate preferences and information. Quart J Polit Sci 1:373–392
Meirowitz A (2007) In defense of exclusionary deliberation: communication and voting with private beliefs and values. J Theor Polit 19:301
Miller NR (1980) A new solution set for tournaments and majority voting: further graph-theoretical approaches to the theory of voting. Am J Polit Sci 24:68–96
Miller D (1992) Deliberative democracy and social choice. Polit Stud 40:54–67
Miller NR (1995) Committees, agendas, and voting. Harwood Academic, London
Niou EM, Ordeshook PC (1985) Universalism in congress. Am J Polit Sci 29:246–258
Ordeshook P, Schwartz T (1987) Agendas and the control of political outcomes. Am Polit Sci Rev 81(1):179–199
Patty JW (2007) Incommensurability and issue voting. J Theor Polit 19:115–131
Penn EM (2006) Alternate definitions of the uncovered set, and their implications. Soc Choice Welf 27: 83–87
Penn EM, Patty JW, Gailmard S (n.d.) Manipulation and Single-peakedness: a general result. Am J Polit Sci, Forthcoming
Peris JE, Subiza Bn (1999) Condorcet choice correspondences for weak tournaments. Soc Choice Welf 16:217–231
Richardson HS (2003) Democratic autonomy: public reasoning about the ends of policy. Oxford University Press, New York
Riker W (1980) Implications from the disequilibrium of majority rule for the study of institutions. Am Polit Sci Rev 74:432–446
Riker WH (1986) The art of political manipulation. Yale University Press, New Haven
Riker WH (1988) Liberalism against populism. Waveland Press, Prospect Heights
Riker WH (1996) The strategy of rhetoric. Yale University Press, New Haven
Rubinstein A (1980) Stability of decision systems under majority rule. J Econ Theory 23:150–159
Satterthwaite MA (1975) Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. J Econ Theory 10:187–217
Schwartz T (1990) Cyclic tournaments and cooperative majority voting: a solution. Soc Choice Welf 7(1):19–29
Shepsle KA (1979) Institutional arrangements and equilibrium in multidimensional voting models. Am J Polit Sci 23(1):27–59
Shepsle KA, Weingast BR (1981) Political preferences for the pork barrel: a generalization. Am J Polit Sci 25(1):96–111
Shepsle KA, Weingast BR (1984) Uncovered sets and sophisticated voting outcomes with implications for agenda institutions. Am J Polit Sci 28(1):49–74
Sunstein CR (2001) Designing democracy: what constitutions do. Oxford University Press, New York
Tullock G (1981) Why so much stability?. Public Choice 37(2):189–202
Tversky A (1972) Elimination by aspects: a theory of choice. Psychol Rev 79:281–299
Tyler TR (2006) Psychological perspectives on legitimacy and legitimation. Annu Rev Psychol 57:375–400
von Neumann J, Morgenstern O (1947) Theory of games and economic behavior, 2nd edn. Princeton University Press, Princeton

Soc Choice Welf
DOI 10.1007/s00355-010-0514-1

A theory of income taxation where politicians focus upon core and swing voters

John E. Roemer

Abstract We construct an equilibrium model of party competition, in which parties are especially concerned with their core and swing voters, concerns which political scientists have focused upon in their attempts to understand party behavior in general elections. Parties compete on an inifinite-dimensional space of possible income-tax policies. A policy is a function that maps pre-fisc income into post-fisc income. Only a fraction of each voter type will vote for each party, perhaps because of issues not modeled here or voter misperceptions of policies. Each party's policy makers comprise two factions, one concerned with maximizing the welfare of its constituency, or its *core*, and the other with winning over *swing* voters. An equilibrium is a pair of parties (endogenously determined), and a pair of policies, one for each party, in which no deviation to another policy will be assented to both its core and swing factions. We characterize the equilibria: they have the property that both parties propose *identical* treatment of a possibly large interval of middle-income voters, while the "left" party gives more to the poor and the "right" party more to the rich. An empirical section uses the data of Piketty and Saez on taxation in the US to assess the model's predictions. We argue that the model is roughly confirmed.

> The spirit of a people, its cultural level, its social structure, the deeds its policy may prepare—all this and more is written in its fiscal history, stripped of all phrases. He who knows how to listen to its message here discerns the thunder of world history more clearly than anywhere else.[1]

[1] Schumpeter (1954 [1918]), as quoted in Brownlee (2004).

J. E. Roemer (✉)
Yale University, New Haven, CT, USA
e-mail: john.roemer@yale.edu

Published online: 01 December 2010

1 Introduction

Formal political-economic analysis of taxation has been in the main of a schematic nature: that is, existing models of income taxation usually assume that taxation is an affine function of income.[2] In reality, income-tax policy is extremely complex, reflecting the fact that many competing interests must be satisfied or attended to. (For a useful "short history" of the income tax in the United States, see Brownlee (2004), which contains a full guide to the literature.) In this paper, I attempt to capture this complexity by modeling political competition over the income tax as taking place on an infinite-dimensional space of functions. Each political party will propose a function which will define the post-tax-and-transfer income, for every possible realization of pre-tax income, and these functions will be chosen from a large space, constrained only by upper and lower bounds on what the marginal tax rates can be.[3]

We will suppose that two parties are competing in a general election, and that the platform of each party consists in a proposal of such a "post-fisc" income function. The paper's positive aspect is to model the view that parties concentrate on *core* and *swing* voters, a view which is prominent in contemporary American political science.[4] A simple way of formalizing these aims of a party is to assume that there are intra-party factions concerned, respectively, with these two problems—of satisfying the core constituency, and of appealing to the swing voters. This model solves the problem of the existence of a (Nash-type) equilibrium in pure strategies in the game of party competition, even when the parties are choosing strategies from an infinite-dimensional space.

Besides modeling parties as complex organizations (in the sense that policy is set by intra-party bargaining, rather than by the maximization of a single payoff function), we depart from traditional formal approaches in the study of political competition in another way. In many—perhaps most—formal papers about political competition, parties represent *no* constituencies. This is the case with the Downs model, where each party is only the vehicle of a candidate who seeks election. (A classic paper by Lindbeck and Weibull (1987) studies competition over multi-dimensional distribution policies between two Downsian parties, which are interested only in maximizing vote share or the probability of victory.) It is as well essentially the case with the citizen–candidate models of Osborne and Slivinski (1996) and Besley and Coate (1997), where candidates run on their own ideal policies (each "party" represents a constituency of one type). In models where parties represent non-trivial constituencies—see,

[2] A number of papers, for example, Dixit and Londregan (1998), study "pork barrel politics," in which parties propose payments to each of a finite number of voter types. Here, the policy space is finite dimensional, but could be of high dimension. Carbonell-Nicolau and Ok (2007) study a political-economy model where tax schedules are drawn from a large space; their aim is to show that income tax policy is marginal-rate progressive. This is definitely *not* a property of the equilibrium policies in the present paper.

[3] I first studied taxation on this policy space in Roemer (2006). The present paper presents a new equilibrium concept. The optimization techniques used here are similar to the ones employed in that monograph. In particular, these techniques, although not simple, are *elementary,* in the sense that no knowledge of advanced optimization theory is required. A knowledge of calculus will suffice to understand the proofs.

[4] See Cox (2006) for a recent review of the formal literature which attempts to model parties' concerns with swing and core voters.

for instance, Dixit and Londregan (1998), Austen-Smith (2000), Levy (2004), Roemer (1999), Roemer (2001), and Roemer (2006)—it is supposed that each party represents an element of some partition of the polity. Here, we depart from this practice, by recognizing that, in reality, it is never the case that the sets of voters who support the various parties form an easily defined partition of the space of voter types. The polity, here, consists of a continuum of voter types, where "type" is defined by the pre-tax income of the agent or household. In American elections, a substantial fraction of voters *at every income level* supports each of the two parties: see Fig. 8 below for the details and McCarty et al. (2006, Chap. 3). Of course, one can say that if the *space of types* were modeled as having sufficiently large dimensionality, there would always be characteristics of voters that would enable us to define the set voting for a particular party as an element of a partition of that space. We prefer, however, to take a more statistical approach: to keep the space of types of small dimension and to say that, from the viewpoint of parties, there is a random element in voting, and therefore, if a party is concerned to represent its constituents, it has to attempt to represent every household type, at least to some extent, because *some* fraction of every type will vote for each party.

We will define an equilibrium concept reflecting these concerns and then characterize the equilibria in the income-tax competition game. The main characteristics of the equilibria are

1. In every equilibrium, there emerges a Left and a Right party. The two parties propose exactly the same tax treatment for what may be a substantial interval of middle-income voters. The greater the focus the parties place upon swing voters, the larger will be the size of this interval. The Left party proposes higher post-fisc incomes for poor voters than does the Right, and the Right party proposes higher post-fisc incomes for rich voters than does the Left.
   In addition, there is an important sub-class of equilibria in which:
2. Each party proposes a piece-wise linear post-fisc policy with three pieces;
3. The policy proposed by Left entails an increasing average rate of taxation on the whole domain of incomes; the policy proposed by the Right entails an average rate of taxation that increases up to a point, and then decreases;
4. There is a two-dimensional manifold of these equilibria, where a particular equilibrium can be viewed as being characterized by the relative strength of the "swing" versus "core" factions within each of the two parties;

In Sect. 2, we propose our equilibrium concepts, and justify them. In Sect. 3, we characterize "left–right" equilibria. In Sect. 4, we examine US income-tax data to see how well reality conforms to the model's predictions. Section 5 discusses and concludes. Lengthy proofs are gathered in Appendix 1.

2 A concept of political equilibrium in two-party politics

2.1 The policy space

A fiscal policy (or an income tax policy) is a mapping $X : \mathbb{R}_+ \to \mathbb{R}_+$ that associates to any pre-tax income $h \in \mathbb{R}_+$ a post-tax-and-transfer income. We assume that the

pre-tax income distribution is given by a distribution function F on $\mathbb{R}_+$; its mean is μ. The policy space $\Im$ consists of functions $X : \mathbb{R}_+ \to \mathbb{R}_+$ such that:

$(P0)$ X is continuous,

$(P1)$ $0 \leq \alpha \leq X' \leq 1$, some fixed α, where X is almost everywhere differentiable,

$(P2)$ $\int X(h)\mathrm{d}F(h) = \mu$

Condition (P1) states that the derivative of X, where it exists, lies between α and 1, and (P2) states that X redistributes pre-tax income fully. (P0) is best justified as a condition of *horizontal equity*: individuals of similar types are similarly treated.

If the policy is X, then the net taxes paid by an individual of type h are $t(h; X) = h - X(h)$. Hence the marginal tax rate for h at policy X is $1 - X'(h)$, which is bounded below and above by zero and $1 - \alpha$, respectively. Leisure is not an argument of the utility function for reasons of tractability: the equilibrium analysis would otherwise become unmanageable. I attempt, however, to partially recognize the issue of labor-supply elasticity by requiring that the marginal tax rate be at most $1-\alpha$: political parties agree not to consider policies that have very high marginal tax rates, because of the deleterious labor-supply effects (α is a parameter of the model). Alternatively put, we are assuming that when marginal tax rates lie in the interval $[0, 1 - \alpha]$, labor-supply elasticity is very small and can be ignored.[5]

Obviously, $\Im$ is a space of infinite dimension, so chosen to model the idea that political competition is ruthless, not being limited by arbitrary mathematical constraints (such as linearity).

2.2 Voter behavior

A voter's predicted utility at a policy X is her post-fisc income, $X(h)$.

However, voters behave stochastically. Facing two policies X^a and X^b from parties a and b, a voter of type (income) h votes for party a's policy when

$$X^a(h) > X^b(h) + \varepsilon,$$

where ε is the realization of a normally distributed random variable with mean zero. We fix this random variable and assume it is the same for all voter types (for purposes of simplicity) and denote its distribution function by $N(\cdot)$ and its density function by $n(\cdot)$. Assuming that the draws on the random variable are i.i.d. across all voters, the fraction of voters of type h who will vote for policy X^a is therefore $N(X^a(h) - X^b(h))$. We define the *share function*

$$S(x, y) \equiv N(x - y).$$

[5] Of course, it would be desirable to model labor-supply elasticity explicitly. In the models here that would necessitate solving a 'double Mirrlees' optimal tax problem; analytical tractability would be lost.

Notice that the function $S(x, y)$ has these properties:

(S1) $S : \mathbb{R}^2_+ \to [0, 1]$, S continuous
(S2) $S(x, y)$ is non-decreasing in x and non-increasing in y
(S3) $S(x, y) + S(y, x) = 1$
(S4) $S(\cdot, y)$ is a convex function of x when $x < y$ and a concave function of x when $x > y$.

Fact (S4) is true since $\frac{\partial S(x,y)}{\partial x} = n(x - y)$, which is an increasing function of x when $x < y$ and a decreasing function of x when $x > y$. In particular, the derivative $\frac{\partial S(x,y)}{\partial x}$ is maximized at $x = y$, a fact that will be important.

Thus, if two policies X^a, X^b are competing, the fraction of the vote received by X^a will be

$$\sigma(X^a, X^b) = \int N(X^a(h) - X^b(h))\mathrm{d}F(h). \tag{2.1}$$

Condition (S3) implies that $S(x, x) = 0.5$, so no income type is a priori biased towards one party. This assumption can be weakened, but we choose to keep the model as simple as possible.

The usual justifications may be given for why voting is stochastic at the level of the individual: voter perceptions of policies are noisy, there are issues not modeled here, and so on.

As I explain below, we may interpret some political entrepreneurs in parties as interested either in maximizing the party's vote share or the party's probability of victory. The first interpretation is more suitable to a system of proportional representation, the second to one of winner-take-all. In the voter behavior just described, individuals behave stochastically but independently, and with a continuum of voter types, there is no aggregate uncertainty. In reality, aggregate uncertainty exists in elections. My preferred way of modeling this is to say that, in the competition between policies X^a and X^b, the realized vote share will be $\sigma(X^a, X^b) + \mathbf{Z}$, where $\mathbf{Z}$ is a uniformly distributed random variable on a (fairly small) interval $(-\beta, \beta)$, for some $\beta > 0$. Thus the probability that policy X^a defeats X^b is

$$\pi(X^a, X^b) \equiv \Pr\left[\mathbf{Z} > \frac{1}{2} - \sigma(X^a, X^b)\right] = \frac{\beta + \sigma(X^a, X^b) - 0.5}{2\beta}, \tag{2.2}$$

where the last equality holds in the case that $-\beta \leq \sigma - 0.5 \leq \beta$. We can think of the macro-uncertainty represented by $\mathbf{Z}$ as occurring because of exogenous events that may shift large numbers of people in the same direction (e.g., the financial melt-down which occurred just before the 2008 US presidential elections, which shifted voter preferences towards the Democrats). Now a key point for the analysis in what follows is that with the "error-distribution model" (2.2), the probability of victory is strictly increasing in σ, when σ lies in the interval of interest (which is the interval where both parties have positive probability of winning). Therefore, *it does not matter* whether we model Opportunists as maximizing vote share or probability-of-victory. (Differences between these two objectives would only appear when the probability of victory is zero or one, in which case π is only *weakly* increasing in vote share.)

2.3 Political parties

We take as given that two political parties exist. The political entrepreneurs in these parties have one goal: to build their careers. However, in pursuing that end, they adopt different objectives. Some seek to maximize their party's vote share (or, in a variant, the probability of winning the election). Others attempt to create a reputation for championing the interests of the core constituency of the party, a constituency that develops over time. I model these politicians as attempting to propose policies which are as close as possible to the core members' ideal point. Others may attempt to maximize the *expected welfare* of their constituency (something which is not the same thing as "championing the interests," in my view). In previous work, I have called the first group the party's Opportunists, the second group the party's Militants or Guardians (they guard the interests of constituents), and the third group the party's Reformists. For discussion, including historical evidence of the existence of such groups or factions with the group of party entrepreneurs, see Roemer (2001, Chap. 8). Note that, in this formulation, *all* politicians are interested in holding office, but there are different paths to that end. I believe it is essential that parties contain either Guardians or Reformists—for without them, the party will come to seen by voters as entirely opportunist, having no attachment to any constituency, and voters will not trust them. I also believe it is essential that parties contain either Opportunists or Reformists: for without them, parties would not pursue strategies aimed at winning elections, and they would disappear.

In my previous work (e.g., Roemer (1999, 2001)), I assumed that each party contained all three of these factions, and an equilibrium is a pair of policies that is "Nash" in the sense that neither party can deviate to another policy while gaining the assent of *all three* of its factions of political entrepreneurs. This is akin to saying that each party's factions bargain which each other when facing the policy of the other party. So a "party-unanimity Nash equilibrium (PUNE)" is a pair of policies each of which is a solution of a bargaining problem in one party, when facing the opposing party's proposal. I showed, however, that the *set of such equilibria* is unchanged if we eliminate the Reformist faction. Reformists are, in a mathematical sense, redundant, once the party possesses both Guardians and Opportunists.

The model of parties that I propose in this paper is a version of the PUNE model. However, because I work on a very large space of policies, it turns out that attempting to maximize probability of victory—or vote share—is mathematically intractable (I will be more precise below). I treat this not simply as a problem for the modeler, but, more importantly, for the party's Opportunists. I therefore propose that the Opportunists do something much simpler than attempt to maximize vote share—they attempt to win over swing voters. I will argue below that doing so is a reasonable rule of thumb which substitutes for the intractable problem of finding policies that maximize vote share.

Thus, I will model parties as possessing Opportunists and Guardians. One may view this as modeling the fact that there are always two tendencies among political entrepreneurs: to propose policies that their constituents like and to propose policies that will win votes. I cannot argue that, with the formulation that is to follow, Reformists would be redundant in this model; I dispense with them. With this introduction, I proceed to a formal definition of equilibrium.

2.4 Political equilibrium

I propose a concept of political equilibrium in which party memberships are endogenous and each party bureaucracy contains political entrepreneurs who adopt different objectives. One such objective is to champion the interest of the *constituency* of the party; the other objective is to target *swing voters.* The constituency of the party and the swing voters are endogenous concepts.

We define the core or constituency of a party as an *historical* and *statistical* concept. Suppose in the last election, at date $t-1$, the set of voters who voted for the party is characterized by the function θ_{t-1} defined by

$$\theta_{t-1}(h) = S(X^a_{t-1}(h), X^b_{t-1}(h)).$$

The party at date t identifies its *core* as the voters so described: that is as a fraction $\theta_{t-1}(h)$ of voters of type h, for every h. From the party's viewpoint, its core at any date is not a well-defined *set*, but rather a *function* which describes the fraction of each voter type who supported the party in the last election.

We say the *swing voter types* comprise the set of income types $\{h|\theta_{t-1}(h) = \frac{1}{2}\}$.

We now discuss the behavior of political entrepreneurs, who set policy for the parties. Parties exist for a long time; they build reputations by representing certain constituencies. With stochastic voting, the constituency of a party is hard to define, because one can never be sure exactly who will vote for the party. Nevertheless, from a statistical viewpoint, the constituency of a party may be quite clear, as I have indicated.

We suppose at the present election (date t) those politicians (the Guardians) who attempt to champion the interests of the party's constituency want to choose the policy X to maximize the functional:

$$\int \theta_{t-1}(h)X(h)\mathrm{d}F(h).$$

That is, they attempt to maximize the average welfare (post-fisc income) of their statistical constituency, by weighting the welfare of every income type by the fraction of that type that comprise the constituency of the party. As stated above, we depart from the more familiar formulation that each party represents a distinct set of voter types.

We model the second faction of politicians, the Opportunists, or "swing voter faction," as insisting that the party promise at least as much to the swing voters as the other party is proposing to give them. As I wrote, their true interest is to maximize the probability of victory; the justification for their focusing upon swing voters will be provided in Sect. 2.5 below.

We first propose a concept of a sequence of political equilibria over time. Note, from the above, that the party's constituency is defined by the last election. We suppose that the distribution of types, F, is unchanging over time.

Definition 1 A *history of political equilibria given a function*[6] $\theta_0 : \mathbb{R}_+ \to [0, 1]$ is a sequence of policies $\{(X_t^{\mathrm{L}}, X_t^{\mathrm{R}}) \in \Im \times \Im | t = 1, 2, \ldots\}$, and a sequence of functions $\{\theta_t : \mathbb{R}_+ \to [0, 1] | t = 1, 2, \ldots\}$ such that:

(1a) for every $t = 1, 2, \ldots$, policy X_t^{L} solves the following program:

$$\max_{X \in \Im} \int \theta_{t-1}(h) X(h) \mathrm{d}F(h)$$
$$\text{subj. to } (\forall h)(\theta_{t-1}(h) = \frac{1}{2} \Rightarrow X(h) \geq X_t^{\mathrm{R}}(h)) \tag{L1}$$

(1b) for every $t = 1, 2, \ldots$, policy X_t^{R} solves the program:

$$\max_{X \in \Im} \int (1 - \theta_{t-1}(h)) X(h) \mathrm{d}F(h)$$
$$\text{subj. to } (\forall h)(\theta_{t-1}(h) = \frac{1}{2} \Rightarrow X(h) \geq X_t^{\mathrm{L}}(h)) \tag{R1}$$

(2) for every $t = 1, 2, \ldots$, and for all $h \in H$:

$$\theta_t(h) = S(X_t^{\mathrm{L}}(h), X_t^{\mathrm{R}}(h)).$$

The constraints (L1) and (R1) in the two programs are imposed by the factions concerned with swing voters: for instance, (L1) says that party L can propose no policy that provides lower utility to the swing voters than the R party proposes to provide them. In words, a history of political equilibria is a sequence of pairs of policies such that, given the party's conception of its statistical constituency from the election held at date $t - 1$, the policy of the party at date t cannot be dominated by any other policy with respect to weighted average post-fisc income (welfare) of the party's historical constituency, subject to providing at least as much as the opposition proposes to provide for swing voters.

The datum of the equilibrium concept is the pair of functions (F, θ_0). We may view θ_0 as the initial conjecture of the two parties concerning their statistical constituencies at the beginning of the process of political competition.

We can expect that there will be many histories of political equilibria. If one is interested in modeling general elections to understand the underlying long-range political conflicts in a society, then one should be interested in stationary points of these histories. An interest in stationary points must, of course, be justified by a view that the underlying distribution of preferences, represented by F, is changing slowly relative to the frequency of elections.

I propose a concept of stationarity which entails that the sequence of functions $\{\theta_t\}$ in a history of political equilibria converges to a function θ_*: thus, the constituency of each party becomes stable.

[6] The L party is the one whose vote share among type h voters at date t is given by $\theta_t(h)$; the analogous vote share of the R party is $1 - \theta_t(h)$.

Definition 2 A *stationary equilibrium* is a function $\theta_* : \mathrm{R}_+ \to [0, 1]$ and a pair of policies $(X_*^{\mathrm{L}}, X_*^{\mathrm{R}})$ such that:

(α1a) policy X_*^{L} solves the program:

$$\max_{X \in \Im} \int \theta_*(h) X(h) \mathrm{d}F(h)$$
$$\text{subj. to } (\forall h)(\theta_*(h) = \frac{1}{2} \Rightarrow X(h) \geq X_*^{\mathrm{R}}(h)) \tag{L2}$$

(α1b) policy X_*^{R} solves the program:

$$\max_{X \in \Im} \int (1 - \theta_*(h)) X(h) \mathrm{d}F(h)$$
$$\text{subj. to } (\forall h)(\theta_*(h) = \frac{1}{2} \Rightarrow X(h) \geq X_*^{\mathrm{L}}(h)) \tag{R2}$$

(α2) for all $h \in H, \theta_*(h) = S(X_*^{\mathrm{L}}(h), X_*^{\mathrm{R}}(h))$.

There is an equilibrium refinement of stationary equilibrium that plays an important role in the analysis. For the moment, view this as just a formal concept, which does not require a political justification.

Definition 3 A 1-*stationary equilibrium* is a function θ_*, a pair of policies $(X_*^{\mathrm{L}}, X_*^{\mathrm{R}})$, and an ordered pair $(h_*, y) \in \mathbb{R}_+^2$ such that:

(β1a) X_*^{L} solves the program:

$$\max_{X \in \Im} \int \theta_*(h) X(h) \mathrm{d}F(h)$$
$$\text{subj. to } X(h_*) \geq y \tag{L3}$$

(β2a) X_*^{R} solves the program

$$\max_{X \in \Im} \int (1 - \theta_*(h)) X(h) \mathrm{d}F(h)$$
$$\text{subj. to } X(h_*) \geq y \tag{R3}$$

(β3) for all $h \in H, \theta_*(h) = S(X_*^{\mathrm{L}}(h), X_*^{\mathrm{R}}(h))$
(β4) $X_*^{\mathrm{L}}(h_*) = y = X_*^{\mathrm{R}}(h_*)$.

In this concept, it is as if the vote-share-seeking faction is concentrating on not losing the loyalty of one swing voter type, namely $h*$. What is important is the relationship of 1-stationary equilibrium to stationary equilibrium.

Proposition 1 *Every 1- stationary equilibrium is a stationary equilibrium.*

Proof Let $(\theta_*, X_*^{\mathrm{L}}, X_*^{\mathrm{R}}, h_*, y)$ be a 1-stationary equilibrium. By (β4), we can write the constraint (L3) as $X(h_*) \geq X_*^{\mathrm{R}}(h_*)$. But by ($\beta$4), it also follows that $\theta_*(h_*) = \frac{1}{2}$. Therefore, constraint (L3) is weaker than constraint (L2). Hence the program in (β1a) has the same objective function but a larger opportunity set than the program in (α1a). However, X_*^{L} is a member of the opportunity set defined by (L2). It follows that X_*^{L} solves (α1a) by the Le Chatelier principle. In like manner, X_*^{R} solves (α1b), proving the claim. □

Now examine the program in condition (β1a). It is a concave programming problem. There is no interaction between the L and R policies—so it should be relatively easy to solve. The same goes for the program in (β1b). In other words, the refined equilibrium concept of 1-stationary equilibrium is quite tractable.

2.5 Further justification of the equilibrium concept

In this section, I wish to justify the constraint imposed by the swing-voter faction. I propose, as discussed in Sect. 2.3, that the swing-voter faction is really concerned with maximizing the party's vote share (or probability of victory). Thus, we might better model party L in the stationary-equilibrium concept as solving a program:

$$\begin{array}{ll} \max \int \theta(h) X(h) \mathrm{d}F(h) & \\ \text{s.t.} \quad X \in \Im & \cdot \text{ Program (P*)} \\ \quad \int S(X(h), X^{\mathrm{R}}(h)) \mathrm{d}F(h) \geq k^{\mathrm{L}} & \end{array}$$

This precisely models the conflict between the core-faction and the vote-share-maximizing faction (the "Opportunists") in party L, where the constant k^{L} reflects the bargaining power of the latter faction in intra-party bargaining. If the programs in the definition of stationary equilibrium were replaced with the programs (P*), then stationary equilibrium would just be a kind of party-unanimity Nash equilibrium (PUNE) that I have studied in the citations given earlier.

However, the function S is not a concave function of X (as I have pointed out: see property (S4) above), and so program (P*) is a non-concave program on an infinite-dimensional space. It is a very difficult problem to solve; I am unable to solve it. Thus, the Opportunists will not even be able to describe precisely the class of policies X that satisfy the constraint in (P*)! Hence, I propose, they will resort to using a rule of thumb. I argue that a reasonable rule of thumb is to focus upon giving income to the swing-voter types.

To see this, imagine that a policy X^a has been proposed by party a's core faction, and that the Opportunists had the option of taking some small per capita amount M of income and distributing it in a uniform manner to voters in some small interval of given length δ. They desire to do this in such a way as to maximize the increase in their party's vote share. (In this thought experiment, we do not worry about where M comes from or about the other constraints on the policy.) Thus, the Opportunists'

problem is to choose a voter type $h_1$and a number m to

$$\text{maximize} \int_{h_1}^{h_1+\delta} N(X^a(h)+m-X^b(h))\mathrm{d}F(h) - \int_{h_1}^{h_1+\delta} N(X^a(h)-X^b(h))\mathrm{d}F(h)$$

$$\text{subject to } \; m = \frac{M}{F(h_1+\delta)-F(h_1)}.$$

The ordered pair (δ, M) is given. We may rewrite the objective of this program as

$$m \int_{h_1}^{h_1+\delta} \frac{N(X^a(h)+m-X^b(h)) - N(X^a(h)-X^b(h))}{m} \mathrm{d}F(h).$$

Now let m approach zero. Then this objective approaches

$$m \int_{h_1}^{h_1+\delta} n(X^a(h)-X^b(h))\mathrm{d}F(h) = M \frac{\int_{h_1}^{h_1+\delta} n(X^a(h)-X^b(h))\mathrm{d}F(h)}{F(h_1+\delta)-F(h_1)}.$$

But this says that, for small M, the Opportunists' problem is to choose an interval $[h_1, h_1+\delta]$ to maximize the average value of the density function $n(X^a(h)-X^b(h))$ on the interval. That value is maximized when $X^a(h)-X^b(h)=0$: that is, the Opportunists should target an interval of swing voters (or an interval around the one swing voter type if it is unique).

To summarize, the mathematical fact that drives this argument is that the density function of a normal variate with mean zero is maximized at zero. The consequence of this fact, and of the stochastic voting rule in the model, is that a good rule of thumb for those concerned with maximizing vote share is to focus resources on the swing voters—at least if these resources are of an incremental nature. The swing-voter constraint in the equilibrium concepts presented here thus derives from the *bounded rationality* of the Opportunists, who (understandably) are unable to solve a non-concave programming problem on a very large space of possible tax policies.[7]

3 Analysis

The first theorem will characterize a two-dimensional family of 1-stationary equilibria. To do so, we define two families of piece-wise linear functions. Fix a number $h_* > 0$.

[7] In Roemer (2009), I fully characterize the equilibria of the model with a vote-share-maximizing faction $S(x, y)$ which *is* concave in x. I consider this, however, to be only a curiosum, because the reasonable stochastic voting model, used here, delivers a function $S(\cdot, y)$ which is neither concave nor convex.

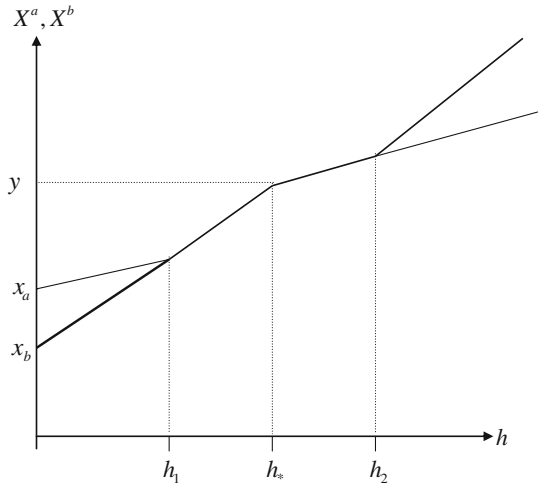

Fig. 1 Policies $X^a \in M_a(h_*)$ (*thin line*) and $X^b \in M_b(h_*)$ (*bold line*) which share a common value y at h_*

The first family is

$$M_a(h_*) = \left\{ \begin{array}{l} X \in \Im | \exists (x_a, h_1) \in \mathbb{R}^2_+ \text{ such that } h_1 \leq h_* \text{ and} \\ X(h) = \left\{ \begin{array}{l} x_a + \alpha h, \text{ if } h \leq h_1 \\ x_a + \alpha h_1 + (h - h_1), \text{ if } h_1 < h \leq h_* \\ x_a + \alpha h_1 + (h_* - h_1) + \alpha(h - h_*), \text{ if } h > h_* \end{array} \right. \end{array} \right\}.$$

A typical function in the family is graphed in Fig. 1. $M_a(h_*)$ is a uni-dimensional family of functions which we may view as being parameterized by the value $y \equiv X(h_*)$; that is, fixing the ordered pair (h_*, y) determines (at most) one policy in the family $M_a(h_*)$. By construction, the policies $X \in M_a(h_*)$ satisfy (P0) and (P1). The budget-balance condition (P2) gives one equation in the two unknowns (x_a, h_1): hence, the uni-dimensionality of this family.

The second family is

$$M_b(h_*) = \left\{ \begin{array}{l} X \in \Im | \exists (x_b, h_2) \in \mathbb{R}^2_+ \text{ such that } h_2 \geq h_* \text{ and} \\ X(h) = \left\{ \begin{array}{l} x_b + h, \text{ if } h \leq h_* \\ x_b + h_* + \alpha(h - h_*), \text{ if } h_* < h \leq h_2 \\ x_b + h_* + \alpha(h_2 - h_*) + (h - h_2), \text{ if } h > h_2 \end{array} \right. \end{array} \right\}.$$

Likewise, $M_b(h_*)$ is a uni-dimensional family of piece-wise linear policies, which is parameterized by $y \equiv X(h_*)$; a typical policy is also graphed in Fig. 1.

We will be interested in policy pairs $(X^a, X^b) \in M_a(h_*) \times M_b(h_*)$ which share a common value of $y = X^a(h_*) = X^b(h_*)$. The next proposition tells us exactly what the admissible range is for y.

Proposition 2 *Let $h_* > 0$, and let y lie in the interval*

$$\max[(1-\alpha)\mu + \alpha h_*, h_*] \leq y \leq h_* + (1-\alpha)\int_{h_*}^{\infty}(h-h_*)\mathrm{d}F(h). \tag{3.1}$$

Then

A. *There exist unique policies $X_a \in M_a(h_*)$, $X_b \in M_b(h_*)$ such that*

$$X_a(h_*) = y = X_b(h_*). \tag{3.2}$$

B. *Conversely, if y does not lie in the interval defined by (3.1), then there is no pair of policies in the two families for which (3.2) holds.*
C. *The number x_a is positive, and the number x_b is non-negative, and positive except in a singular case.*

All longer proofs, beginning with the proof of this proposition, appear in Appendix 1.

Define:

$$\Gamma = \{(h_*, y) \in \mathbb{R}^2_+ | y_{\min}(h_*) \leq y \leq y_{\max}(h_*)\}$$

where $y_{\min}(h_*) = \max[(1-\alpha)\mu + \alpha h_*, h_*]$, $y_{\max}(h_*) = h_* + (1-\alpha)\int_{h_*}^{\infty}(h-h_*)dF(h)$.

Proposition 2 tells us that for any $(h_*, y) \in \Gamma$, there exists a unique pair of policies $X^a \in M_a(h_*)$ and $X^b \in M_b(h_*)$ such that

$$X^a(h_*) = y = X^b(h_*).$$

To avoid notational complexity, let us fix $(h_*, y) \in \Gamma$ and denote these two functions simply by X^a and X^b. Figure 1 displays the graphs of a typical pair of such functions. Note, in particular, that these two policies coincide on the interval $[h_1, h_2]$. Suppose, now, that these two policies are being proposed by the parties, a and b, and define the vote-share function $\theta(h) = N(X^a(h) - X^b(h))$. We have the following proposition.

Proposition 3 *When $X^a \neq X^b$, the function $\theta(\cdot)$ is decreasing on the interval $[0, h_1]$, constant and equal to one-half on the interval $[h_1, h_2]$, and decreasing on the interval (h_2, ∞).*

Proof Easily verified from the definition of the functions X^a and X^b and the properties of N. □

We may now state our first main result.

Theorem 1 *Let $(h_*, y) \in \Gamma$ and let $X^a \in M_a(h_*)$, $X^b \in M_b(h_*)$ be such that $X^a(h_*) = y = X^b(h_*)$. Let $\theta(\cdot)$ be defined as above. Then (θ, X^a, X^b) is a 1- stationary equilibrium.*

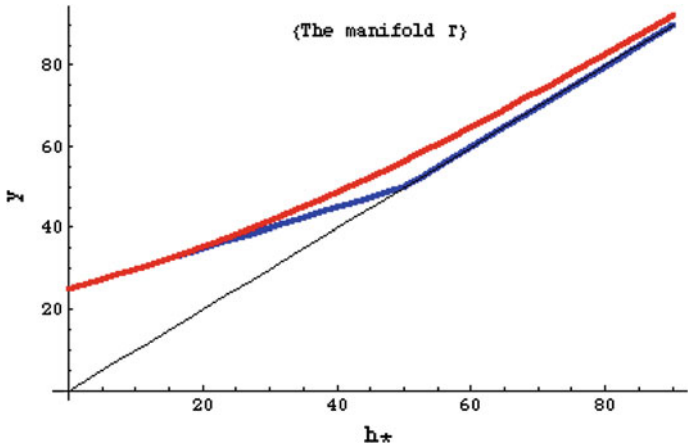

Fig. 2 Graphs of the functions $y^{\max}(\cdot)$, $y^{\min}(\cdot)$ and the ray $y = h_*$. The manifold Γ is the set bounded by the two *bold curves*

Theorem 1 gives us a stationary equilibrium for each $(h_*, y) \in \Gamma$: thus, it specifies a two-parameter family of equilibria. We propose an interpretation of the political nature of these various equilibria, which follows from the next result. In Fig. 2, we graph the manifold Γ for the case where F is the lognormal distribution of income with mean 50 and median 40, an approximation of the US household income distribution in 2000, in units of \$1000. Note that for $h_* \geq \mu$, the lower envelope of Γ coincides with the 45° ray.

Theorem 2 A. *Consider a point $(h_*, y_{\max}(h_*))$ on the upper envelope of the manifold Γ. Let the two policies of the 1-stationary equilibrium at this point be denoted X^{L} and X^{R}. Then $h_1 = 0, \quad h_2 = \infty$ and $X^{\mathrm{L}} = X^{\mathrm{R}} = X^*$, where X^* is the ideal policy in $\Im$ of voter h_*.*

B. *Consider a point $(h_*, y_{\min}(h_*))$ on the lower envelope of Γ, with its associated 1-stationary equilibrium $(X^{\mathrm{L}}, X^{\mathrm{R}})$. If $h_* \leq \mu$ then X^{L} is the ideal policy in $\Im$ of Left's constituency,*[8] *and if $h_* \geq \mu$ then X^{R} is the ideal policy in $\Im$ of Right's constituency.*

Fix the "pivot type" h_* and begin at the equilibrium on the upper envelope of Γ at h_*. In the stationary equilibrium at this point, both parties propose the ideal policy of the pivot type, h_*. Each party receives half the vote. Here we have politics where the concern for swing voters is very strong in both parties: the factions representing constituent interests have no pull. As we start to move vertically down the manifold Γ, decreasing y and holding h_* fixed, the two policies diverge. The factions concerned with core voters become more powerful in intra-party bargaining. When we reach the equilibrium on the lower envelope of Γ, if $h_* < \mu$, then this faction is entirely dominant in the Left party in the sense that the L party is playing as if it is only concerned with constituent interests; if $h_* > \mu$, then constituent interests are dictating policy in

[8] That is, X^{L} maximizes $\int \theta(h) X(h) dF(h)$, for $X \in \Im$, where $\theta(h) \equiv N(X^{\mathrm{L}}(h) - X^{\mathrm{R}}(h))$.

the Right party. In the singular case that $h_* = \mu$, both parties are maximizing over $\Im$ the average utility of their statistical constituencies.

For a policy X, define the average tax rate at h as:

$$a(h; X) = \frac{h - X(h)}{h}.$$

Define a policy as progressive if it unambiguously redistributes from the rich to the poor, in the following sense:

Definition 4 A policy X is *progressive* if there exists $\hat{h}$ such that:

$$h \leq \hat{h} \Rightarrow X(h) \geq h$$
$$h > \hat{h} \Rightarrow X(h) \leq h$$

and at least one (some) of these inequalities hold(s) strictly for some h.

It follows immediately from the definition of the set $\Im$ that all policies (except the laissez-faire policy) are progressive. We also have the following proposition.

Proposition 4 *Consider any 1-stationary equilibrium* $(X^{\mathrm{L}}, X^{\mathrm{R}})$ *of Theorem 1. Then* $a(\cdot; X^{\mathrm{L}})$ *is increasing on* $\mathbb{R}_+$*, and* $a(\cdot; X^{\mathrm{R}})$ *is increasing on* $[0, h_2)$ *and decreasing on* (h_2, ∞)*, except in the singular case that* X^{R} *is the laissez-faire policy.*

Proof The condition $\frac{\mathrm{d}}{\mathrm{d}h} a(h; X) > 0$ is equivalent to $hX'(h) < X(h)$. It is easy to check (for instance, examine Fig. 1) that this condition is true for X^{L}: formally, this follows from the fact that $x_a > 0$ and $x_b \geq 0$ (see Prop. 2(c)). For X^{R}, it is also easy to check that the condition holds if and only if $h < h_2$. Now note that if the segment of the graph of X^{R} on the domain $[h_2, \infty)$ is extended into a line, it passes below the origin. (For if it passed through or above the origin, the policy X^{R} would dominate the policy $X(h) \equiv h$, and would not be feasible, as it would integrate to more than μ.) This means that $hX' > X$ on $[h_2, \infty)$. □

We illustrate these stationary equilibria in Fig. 3. Here, F is the lognormal distribution with mean 50 and median 40. We graph the 1-stationary equilibrium of Theorem 1 for four values of y, holding h_* at the mode of the income distribution. When $y = y_{\max}(h_*)$ (Fig. 3a), the two policies coincide at the ideal policy of h_*; when $y = y_{\min}(h_*)$ (Fig. 3c), Left is playing the ideal policy of its average constituency, since $h_* < \mu$. Even on the lower envelope of the manifold, the policies agree for 16% of the polity.

In Fig. 4, we plot the average tax rate functions for the Left and Right policies, at a point in Γ. The Right imposes a higher average tax rate up to h_1; of course average tax rates of the two policies coincide on the interval $[h_1, h_2]$; the Left policy imposes a higher average tax rate on (h_2, ∞). Moreover, the Left average tax rate is monotone increasing on the whole domain, while the Right average tax rate is increasing until h_2, and then monotone decreasing asymptotically to zero thereafter.

We next ask what is the effect of a change in h_* on equilibrium policies. We graph some examples to show the contrast. In Fig. 5, we present the average tax rate functions

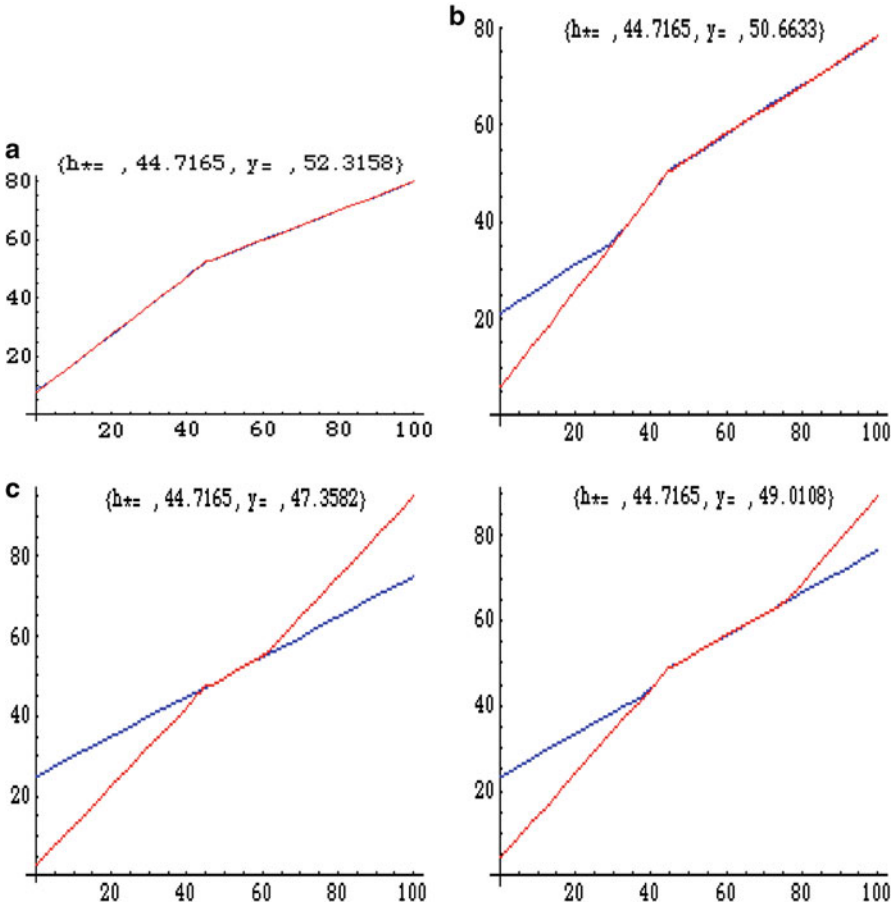

Fig. 3 **a** L and R policies on the upper envelope of the manifold Γ. **b** Graphs of stationary equilibria as y decreases in Γ at constant h_*. **c** Stationary equilibrium on the lower envelope of the Γ, where Left plays the ideal policy of its average constituency

associated with Left and Right policies for two values, $h_* \in \{20, 80\}$. In each case, we plot policies for $(h_*, \frac{y_{\min}(h_*)+y_{\max}(h_*)}{2}) \in \Gamma$. We see that the effect of increasing the "pivot" h_* is to flatten out the average tax rate functions. For large values of the pivot, both parties propose net tax rates of close to zero for middle-income voters. Further discussion of how to interpret the 2-manifold of equilibria appears in Sect. 5.

We next note the central role of 1-stationary equilibria in the theory.

Definition 5 A *left–right stationary equilibrium* is a stationary equilibrium where the function θ_* is weakly monotone decreasing, $\theta_*(0) \geq \frac{1}{2}$, and for sufficiently large h, $\theta_*(h) \leq \frac{1}{2}$.

In other words, such an equilibrium is one where one party (Left) gives more weight to voters the poorer they are, the other (Right) gives more weight to voters, the richer they are. (We include the case where $\theta_*(h) \equiv 1/2$ as an instance of "Left–Right"

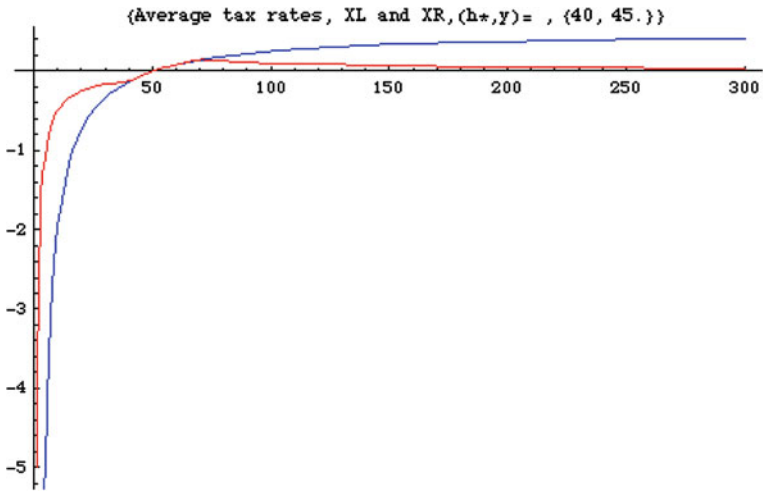

Fig. 4 The average tax rate functions for a pair of equilibrium policies

equilibrium only for semantic convenience, to simplify the statement of theorems.) The next theorem tells us how 1-stationary equilibria come about historically.

Theorem 3 *Suppose that $\theta_0(\cdot)$ is a weakly monotone decreasing function, and there is a unique income h_* such that $\theta_0(h_*) = \frac{1}{2}$. Let $y \in [y_{\min}(h_*), y_{\max}(h_*)]$ and let $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}}) \in M_a(h_*) \times M_b(h_*)$ be the unique policies associated with the ordered pair $(h_*, y) \in \Gamma$. Then $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$ is a stationary equilibrium reached at date 1 beginning from θ_0 in a sequence of historical equilibria. Conversely, let $(X^{\mathrm{L}}, X^{\mathrm{R}})$ be any equilibrium reached at date 1 beginning from θ_0, in a sequence of historical equilibria. Then $(X^{\mathrm{L}}, X^{\mathrm{R}})$ are precisely the policies in $M_a(h_*) \times M_b(h_*)$ associated with the ordered pair $(h_*, X^{\mathrm{L}}(h_*)) \in \Gamma$.*

If we assume, in a history of political equilibria, that once a stationary equilibrium is reached, it continues to be played at all future dates, then Theorem 3 says that all sequences of historical equilibria which begin with a vote share function θ_0 as specified in the premise end in one period, with a left–right 1-stationary equilibrium, as depicted in Fig. 1. Theorem 3 tells us that 1-equilibria occupy a special significance: they are the stationary equilibria that are attained from an initial historical vote-share function θ_0 which possesses a unique swing-voter type.

We next will describe the stationary equilibria reached by sequences of historical equilibria associated with historical share functions θ_0 which are weakly monotone decreasing and for which there is a non-degenerate interval $[h_*, h_{**}]$ upon which $\theta_0 = \frac{1}{2}$. Let $Z : [h_*, h_{**}] \to \mathbb{R}_+$ be an arbitrary, continuous function such that $\alpha \leq Z'(h) \leq 1$. If the integral $(\mathrm{d}F)$ of Z on $[h_*, h_{**}]$ is neither too small nor too large, then there exists a left–right stationary equilibrium, in which both L and R policies coincide with Z on the interval$[h_*, h_{**}]$, and are as depicted in Fig. 6a. On the intervals $[0, h_*]$ and $[h_{**}, \infty)$, the two policies behave just as the policies in a 1-equilibrium (see Fig. 1). Note that if $h_{**} = h_*$, then Fig. 6a becomes exactly Fig. 1.

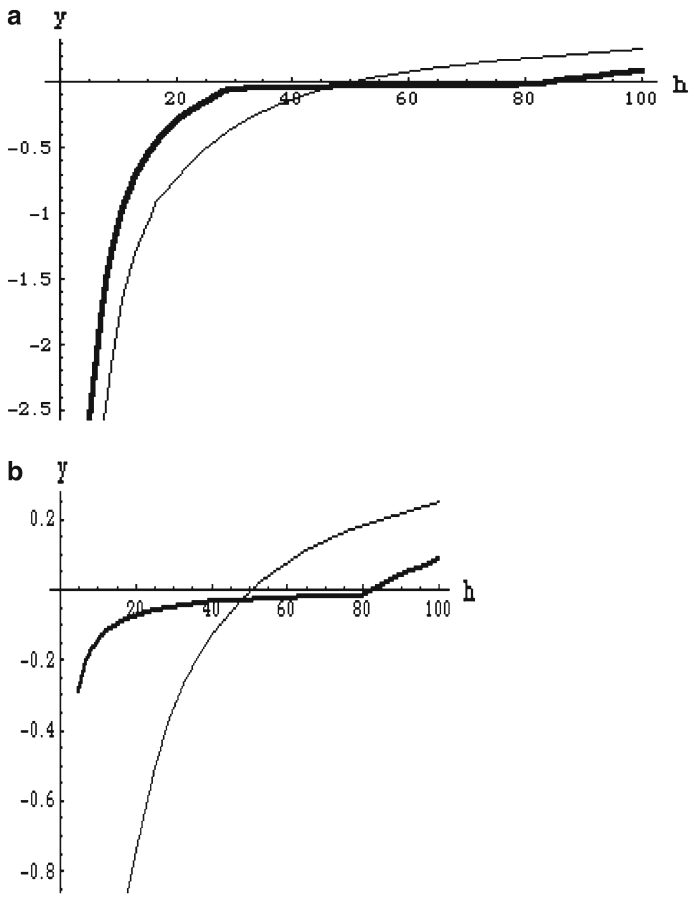

Fig. 5 Average tax rates. At $h_* = 20$ (*plain*) and $h_* = 80$ (*bold*). **a** Left policy, **b** right policy

To be precise:

Theorem 4 *Let θ_0 be weakly monotone decreasing such that $\theta_0(h) = \frac{1}{2}$ on $[h_*, h_{**}]$. Let Z be a continuous function defined on $[h_*, h_{**}]$ such that*

$$(\forall h \in [h_*, h_{**}])(\alpha \leq Z'(h) \leq 1).$$

Suppose that there exist numbers $h_1 \in [0, h_*]$ and $h_2 \in [h_{**}, \infty)$ and $x_a \geq 0,\ x_b \geq 0$ such that the functions $\hat{X}^{\mathrm{L}}$ and $\hat{X}^{\mathrm{R}}$, defined below, are continuous and integrate ($\mathrm{d}F$) to μ:

$$\hat{X}^{\mathrm{L}}(h) = \left\{ \begin{array}{l} x_a + \alpha h,\ 0 \leq h \leq h_1 \\ x_a + \alpha h_1 + (h - h_1),\ h_1 < h \leq h_* \\ Z(h),\ h_* < h \leq h_{**} \\ Z(h_{**}) + \alpha(h - h_{**}),\ h > h_{**} \end{array} \right\}$$

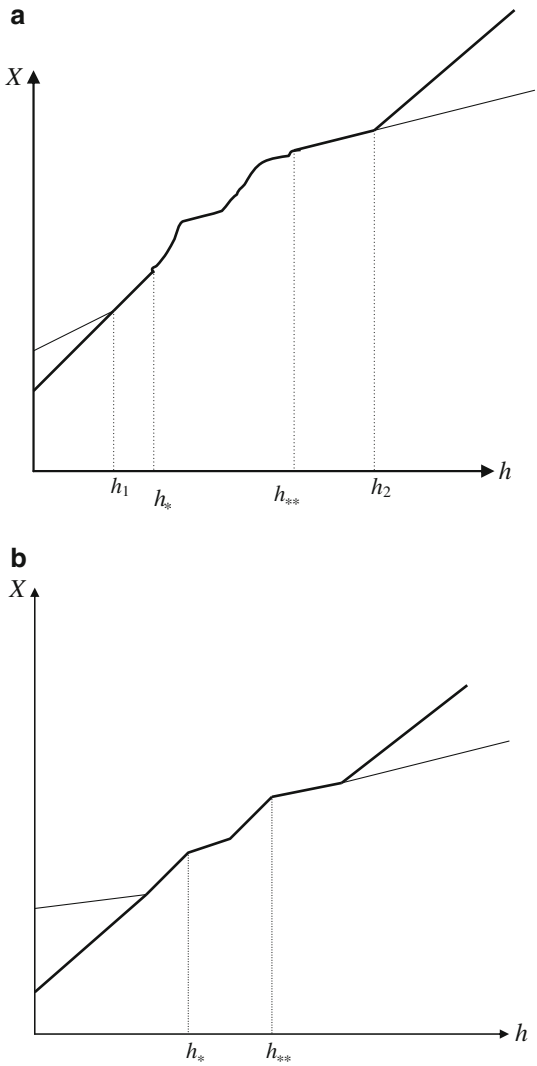

Fig. 6 **a** Left–right stationary equilibrium where X^{L} and X^{R} coincide with arbitrary Z on $[h_*, h_{**}]$ and coincide on $[h_1, h_2]$. **b** A generic 2-stationary equilibrium. Policies coincide on an interval including $[h_*, h_{**}]$. *R* policy in *bold*, *L* policy in *plain face*

$$\hat{X}^{\mathrm{R}}(h) = \begin{cases} x_b + h, \ 0 \le h \le h_* \\ Z(h), \ h_* \le h \le h_{**} \\ Z(h_{**}) + \alpha(h - h_{**}), \ h_{**} \le h \le h_2 \\ Z(h_{**}) + \alpha(h_2 - h_{**}) + h - h_2, \ h > h_2 \end{cases}$$

Then $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$ *is a left–right stationary equilibrium, reached in one date from the historical vote share function* θ_0. *Conversely, let* $(X^{\mathrm{L}}, X^{\mathrm{R}})$ *be any equilibrium beginning at* θ_0 *which is reached at the first date, and let* $Z(h) \equiv X^{\mathrm{L}}(h)$ *on* $[h_*, h_{**}]$. *Then the functions* $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$ *can be defined as in the statement, they integrate* $(\mathrm{d}F)$ *to* μ, *and*

$$(X^{\mathrm{L}}, X^{\mathrm{R}}) = (\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}}).$$

Theorems 1 and 4 characterize left–right stationary equilibria. Obviously, the 1-stationary equilibria are the simplest; there is a 2-manifold of them. The manifold of equilibria of the form described in Theorem 4 is infinite dimensional, for the function Z can be specified in an essentially arbitrary way.

We can locate a set of equilibria which lie "between" the 1-equilibria and the equilibria of Theorem 4. Let us define the concept of a 2-stationary equilibrium:

Definition 6 A 2-*stationary equilibrium* is a tuple $(h_*, y_*, h_{**}, y_{**})$ and a triple of functions $(\theta_*, X^{\mathrm{L}}, X^{\mathrm{R}})$ such that:

(γ1a) X^{L} solves

$$\begin{aligned}
&\max_{X \in \Im} \int \theta_*(h) X(h) \mathrm{d}F(h) \\
&\text{s.t.} \\
&X(h_*) \geq y_* \\
&X(h_{**}) \geq y_{**}
\end{aligned}$$

(γ2a) X^{R} solves

$$\begin{aligned}
&\max_{X \in \Im} \int (1 - \theta_*(h)) X(h) \mathrm{d}F(h) \\
&\text{s.t.} \\
&X(h_*) \geq y_* \\
&X(h_{**}) \geq y_{**}
\end{aligned}$$

(γ3) $\theta_*(h) = S(X^{\mathrm{L}}(h), X^{\mathrm{R}}(h))$
(γ4) $X^{\mathrm{L}}(h_*) = y_* = X^{\mathrm{R}}(h_*)$ and $X^{\mathrm{L}}(h_{**}) = y_{**} = X^{\mathrm{R}}(h_{**})$.

It will not surprise the reader that there is a 4-manifold of 2-stationary equilibria, parameterized by the choice of the vector $(h_*, y_*, h_{**}, y_{**})$; they comprise piece-wise linear policies, where each policy has five pieces. A typical one is depicted in Fig. 6b. The slopes of the line segments of each policy alternate between α and 1, beginning with slope α for the L policy and slope 1 for the R policy. Each of these is, of course, an equilibrium of the type described in Theorem 4, where the function Z is a piece-wise linear function with two pieces. The theorem characterizing 2-stationary equilibria is again proved by the same method as Theorem 1.

It is hard to imagine how the general left–right stationary equilibria of Fig. 6a might come about. In contrast, 1-stationary equilibria are easy to imagine, and we have given an historical explanation of them in Theorem 3. Even 2-stationary equilibria are imaginable, where the swing-voter factions are concentrating not on all the swing voters, but on two prominent income types. We can easily generalize this concept to n-stationary equilibrium, where the swing factions concentrate on not losing the loyalty of n

voter types: this generates stationary equilibria where the policies are each piece-wise linear with $2n+1$ pieces. Thus, in the Eisenhower administration, when the piece-wise linear income-tax schedule in the United States had 17 pieces, we can imagine that eight income types had sufficient clout to convince both parties that their votes were up for grabs.

A final remark may be in order, comparing the *stationary equilibria* of the present paper with the *quasi-PUNEs* of Roemer (2006, p. 46). In fact, the 1-stationary equilibria of this paper comprise post-fisc income policies that are identical to the post-fisc income policies of quasi-PUNEs, taking $\alpha = 0$. The definitions of the two concepts, however, are quite different. For instance, in quasi-PUNEs, each party's constituency is a discrete interval of types, whereas in stationary equilibria, the constituency is a statistical concept. The fact that the equilibrium policies look the same might be viewed as a kind of robustness. However, the general form of policies in stationary equilibria is quite different from the quasi-PUNE policies.

4 Income tax rates in the United States

> On the average, people with income below $100,000 would get more from Obama than from McCain. From $100,000 to $250,000 they'd be fairly even under Obama and McCain. For those over $250,000, Obama increases taxes—Tax Policy Center.[9]

In this section, we ask how well the model performs in light of the recent historical record in the United States. We use the data on income taxation assembled by Piketty and Saez (2006). In their research, Piketty and Saez have used the public micro-file tax data of the IRS, and have computed the sum of four federal taxes for US taxpayers: the income tax, the social security and medicare payroll taxes, the estate tax, and the corporate tax. The corporate tax is allocated to households in proportion to their holdings of corporate equity.[10] The authors then compile the distribution of taxes paid, annually for the years 1960–2004, and consequently the distribution of post-tax income.

Post-tax income, so computed, is not the theoretical concept that we used in Sects. 2 and 3 of *post-fisc* income. Thus, I have amended the Piketty–Saez data by including transfer payments, taken from the PSID, for the years 1974–2000. Thus, Piketty–Saez's tax rates are converted to average post-fisc tax rates of US taxpayers, by pre-fisc income quantile. This is defined as $\frac{T-t}{y}$, where T is the sum of the four taxes of Piketty–Saez, t is the value of transfer payments, and y is pre-fisc income. See Appendix 2 for the details of how the Piketty–Saez data were amended.[11]

[9] Cited in "For incomes below $100,000, a better tax break in Obama's plan," *New York Times*, October 31, 2008, p. A14.

[10] This is a questionable decision. General-equilibrium computations show that the incidence of corporate taxation falls in significant measure on workers, especially because of increased capital mobility in recent years.

[11] I am grateful to Kenneth Couch and his research assistants, who extracted the transfer-payment data from the PSID for this paper.

Finally, the model assumes a balanced budget for all policies; but in reality, the budget is rarely balanced. Therefore, I have amended the post-fisc tax rates computed above by allocating the budget deficit (surplus) to all voters via a proportional tax rate. I call the final statistic the "average adjusted post-fisc tax rate."

In recent US fiscal history, the main tax reforms were the following[12]:

- In 1981, the Economic Reform Tax Act was passed under R. Reagan, which reduced the top marginal income tax rate from 70 to 50%, and continued to cut rates over three years;
- In 1986, the Tax Reform Act was signed by Reagan;
- In 1993, the top personal income tax rate was raised under B. Clinton to 39.6%;
- In 1997, the Tax Payer Relief Act cut the top rate on capital gains from 28 to 20%;
- In 2001, under George W. Bush, the Economic Growth and Tax Relief Reconciliation Act reduced the tax rate in the lowest bracket to 10%, reduced the highest marginal rate to 35%, and reduced the marriage penalty. In addition, the estate tax was to be reduced over a ten year period to the vanishing point.

Our model supposes that each of the two political parties proposes a tax policy as part of its electoral strategy. This is, of course, a stylization of reality. In order to confront the data, we will assume that when a major tax reform occurs, the policy that is enacted is the equilibrium policy of the president's party. That policy continues to hold until the next major tax reform. Thus, for example, we will assume that the policy in force prior to 1981 was the Democratic Party's equilibrium policy; that the policy after 1986 was the Republican Party's equilibrium policy; that the policy after 1993 was the Democratic Party's equilibrium policy; and that the policy after 2001 was the Republican Party's equilibrium policy. Our method will be to examine the *de facto* changes in the distribution of average post-fisc tax rates, before and after these major tax reforms, identifying the results that we observe with equilibrium party policies as described.[13]

Figure 7 (various panels) presents the average post-fisc tax rates before and after major tax reforms. In Fig. 7a, we present the average tax rates for the various quantile groups reported by Piketty and Saez, before (1981) and after (1988) the Reagan tax reforms. Piketty and Saez are particularly interested in the tax treatment of the very rich; we see that they disaggregate the top decile of the income distribution into six quantile groups, where the top group refers to the top 0.01% of the income distribution (currently, about 13,000 households).

The main observations from Fig. 7a are that the Reagan tax reforms substantially reduced the tax rates on the top 0.5% of the income distribution, reduced tax rates on the top decile, left tax rates on the 60–90$^0/_{00}$ about the same, increased tax rates on the two quantiles occupying the 20–60%, and reduced net taxation (increased the transfer rate) on the bottom quintile. If we interpret the pre- and post-Reagan reform tax rates as

[12] See Brownlee (2004).

[13] It must be emphasized that this move is a very rough approximation to reality. Tax policy in the US is hammered out by Congressional compromise. In periods of divided government, it is difficult to maintain that the legislated policy is the president's policy. Taking the legislated policy to be the government leader's policy would be more satisfactory in a parliamentary democracy.

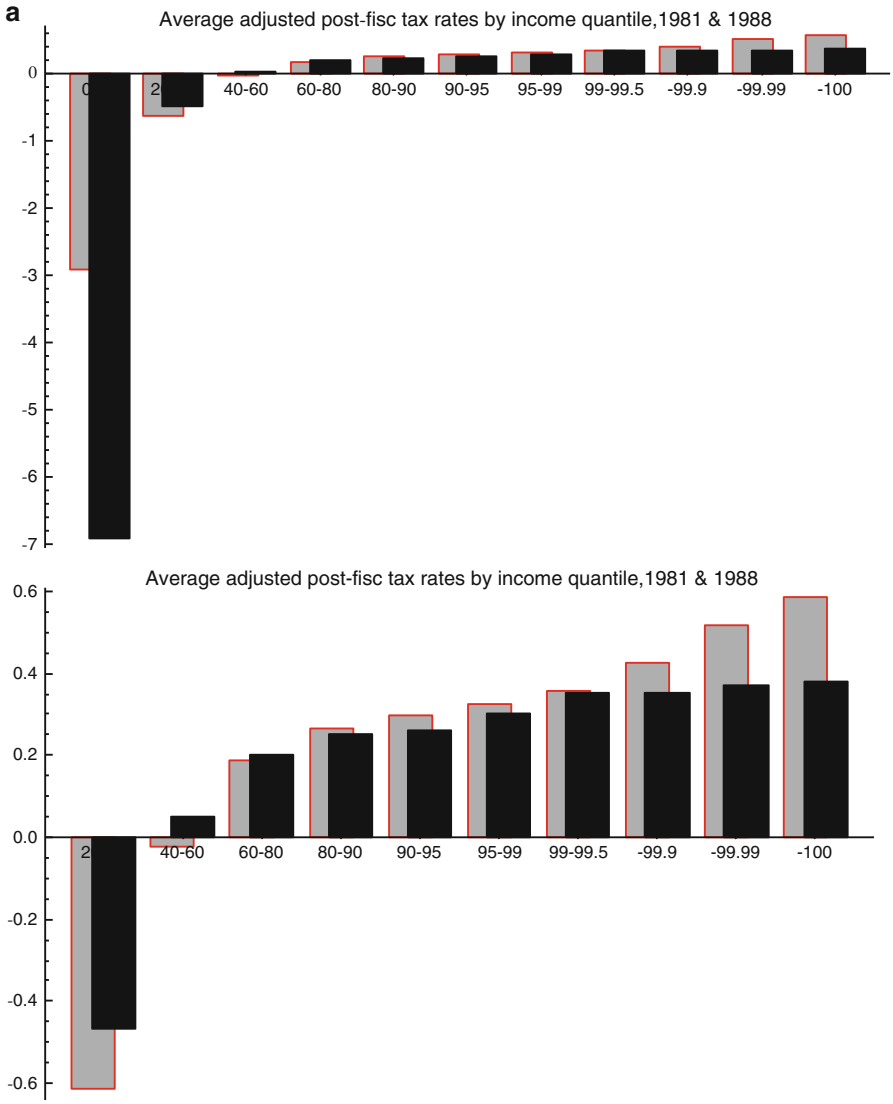

Fig. 7 **a** *Light bars* are 1981 (Carter), *dark bars* are 1988 (Reagan). **b** *Dark bars* are 1988 (Reagan), *light bars* are 1996 (Clinton). **c** *Dark bars* are 2004 (G.W. Bush), *light bars* are 1996 (Clinton). Lower panel excises the bottom quintile and re-scales. **d** *Light bars* are 1974, *dark bars* are 2004

associated with Democratic and Republican equilibrium policies, respectively, these characteristics conform to the model's predictions, except for the treatment of the bottom quintile (due to the earned income tax credit which was expanded significantly in 1986): in particular, there is a sizeable group of middle income voters who receive essentially the same tax treatment by both parties.

One characteristic of our equilibrium policies that does not conform to the data is the predicted decrease in the average tax rate proposed by the Right party for the

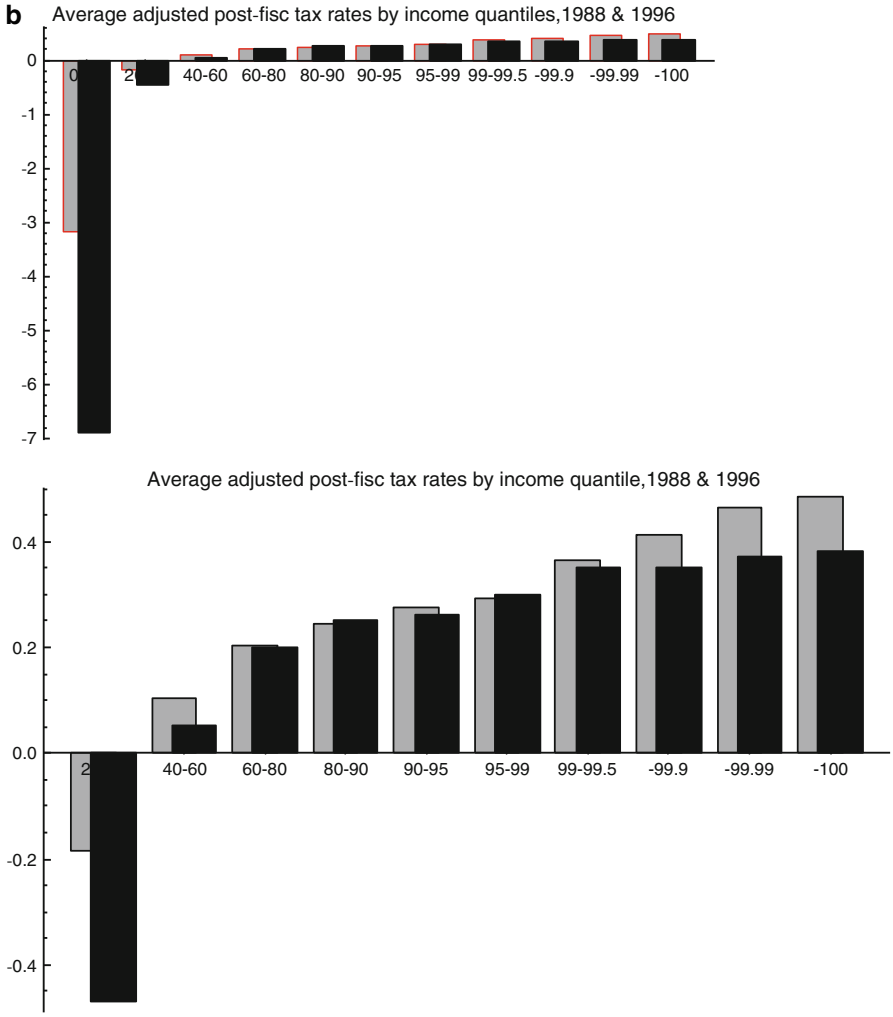

Fig. 7 continued

upper end of the distribution (incomes greater than h_2; recall Fig. 4). In addition, the equilibrium policies in our model either tax at the minimal marginal tax rate (zero) or the maximal marginal tax rate $(1 - \alpha)$; this is not a feature of observed tax rates.

Figure 7b presents the average post-fisc tax rates by quantile groups in 1988 and 1996, to attempt to capture the effect of the Clinton tax reform of 1993.[14] These reforms substantially increased tax rates on the top 0.5%; left tax rates approximately

[14] Of the tax reforms considered in this section, the Clinton reform is perhaps the one for which the heroic assumption—that the reform is the policy advocated by the president's party—is the most accurate. This tax reform passed both houses by the narrowest of partisan margins.

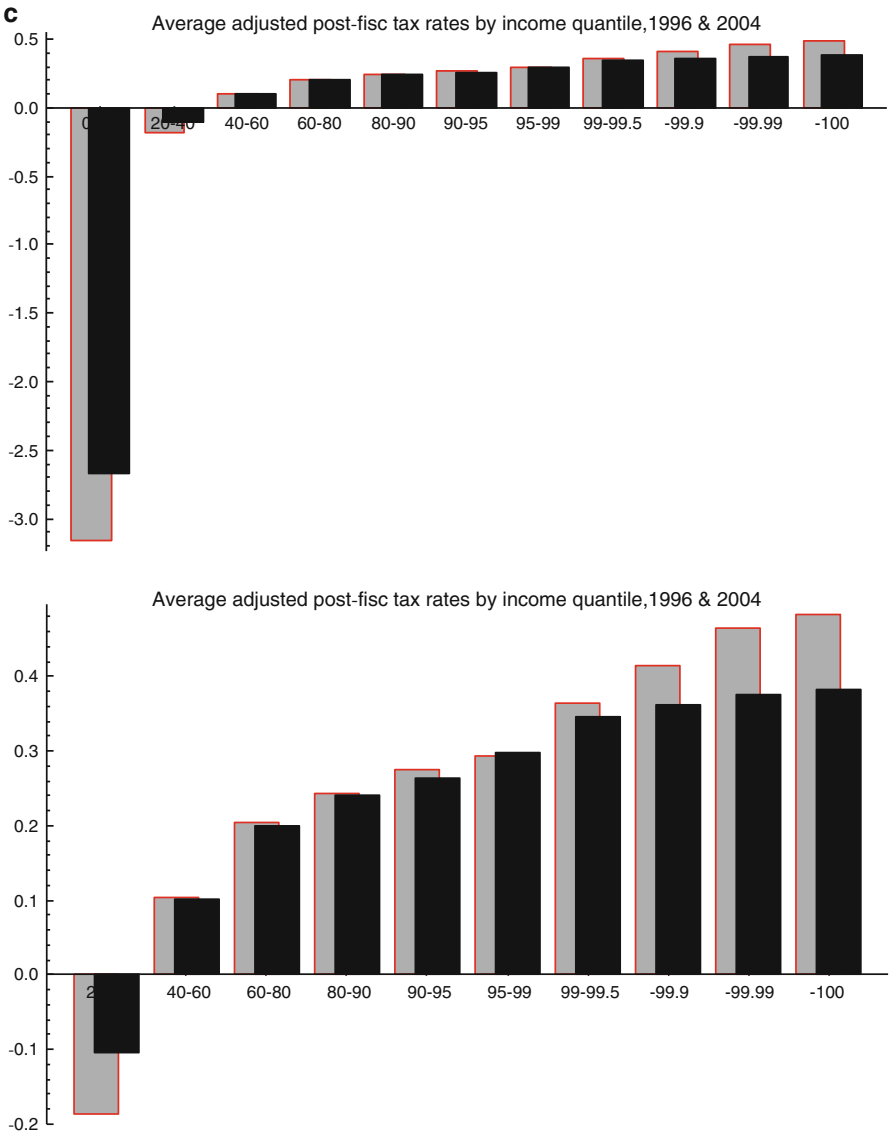

Fig. 7 continued

unchanged on the 60–99.5‰, and increased tax rates on the bottom three quintiles. Part of the effect on the poorest voters is probably due to the welfare reform.

Figure 7c presents the tax rates before and after the G.W. Bush tax bill of 2001. The Bush tax bill reduced post-fisc tax rates on the richest centile. There appears to be no significant change on treatment of the 40–99‰. Transfer payments to the bottom two quintiles were reduced. (Without the deficit financing built into to our adjusted tax rates, it turns out that G.W. Bush decreased the tax rates on *all* quantiles.)

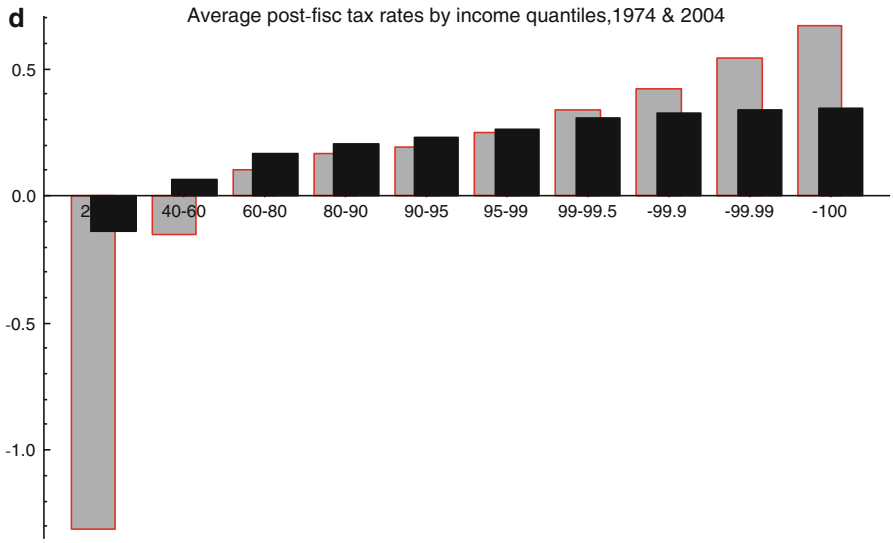

Fig. 7 continued

As a final contrast, we present in Fig. 7d tax rates in 1974 and 2004.[15] These years are too far apart for us to interpret the tax treatments as a pair of equilibrium policies. The figure shows that there has been a large shift of the tax burden from the top 1% of the income distribution to the bottom 99%. However, pre-tax income has also shifted from the bottom 99% to the top 1% over this period, and so the shift in net taxes paid by the very rich and the rest will not have shifted so dramatically as the figure might suggest.

In Fig. 8, we present data on the percentage of voters who voted for the Democratic presidential candidate, for various election years, by income quantile.[16] In other words, the graphs in Fig. 8 give us a discrete approximation of the function $\theta(\cdot)$ for various years. In 2000, 1996, and 1992, we can say that the swing voter types occupied a region between 34 and $95^0/_{00}$ of the income distribution; in 1980, swing voter types were between 17 and $33^0/_{00}$ of the income distribution. It certainly appears from these data that the US is characterized by what we have called a left–right equilibrium: the function θ is monotone decreasing. The general observation seems to be that over this period, the swing voter types have become more numerous, and have moved up in the income distribution. At least for the recent elections (since 1992), the regions of the income distribution where the two parties coincide in their tax treatment (from Fig. 7) seem to correspond roughly to where the swing voter types lie.

[15] These rates in Fig. 7d are not adjusted for the deficit.

[16] I thank Joseph Bafumi for providing me with these data, which he compiled from the American National Election Studies (ANES).

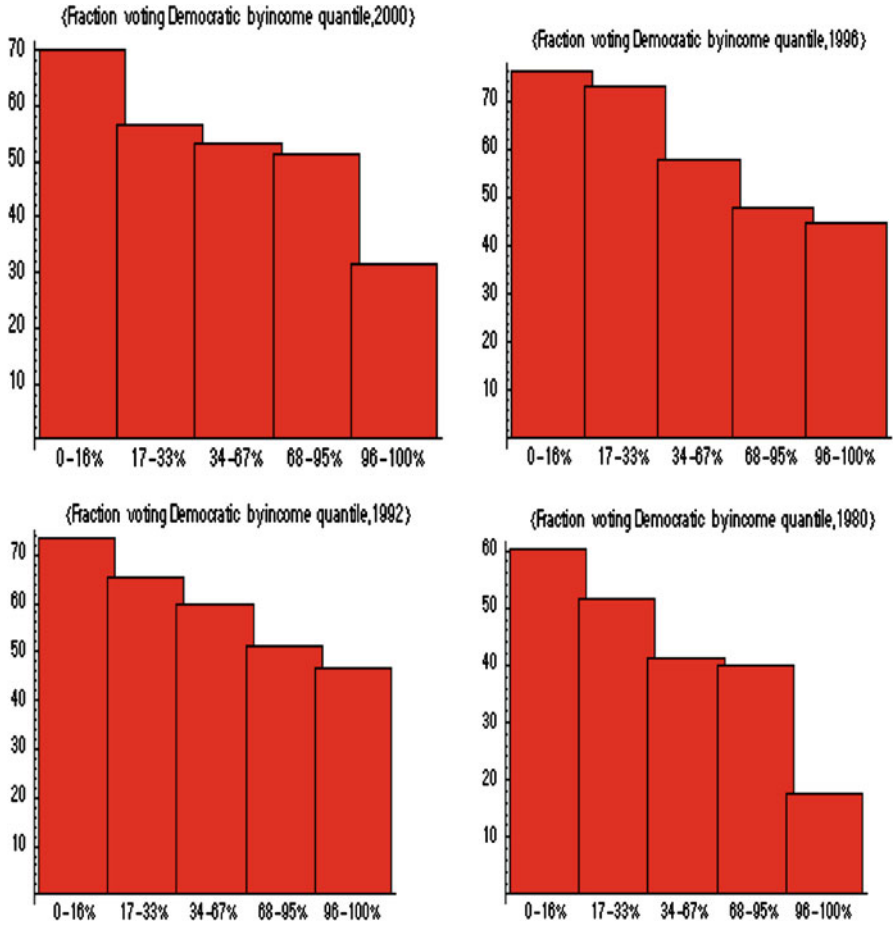

Fig. 8 The empirical function θ: Democratic presidential candidate vote share, by voter income quantile, various years

5 Discussion and conclusion

I have modeled political competition in a general election between two parties, incorporating two features of what appears to be American political reality: that parties compete on a very large policy space, and that their leaders or strategists conflict internally over whether to appeal to their core voters, or to appeal to voter types sitting on the fence, or to maximize the expected welfare of their constituents. I chose the policy space to consist of all continuous functions restricted only by a budget constraint, and by a requirement that marginal tax rates lie everywhere in an interval $[0, 1 - \alpha]$. Choosing $\alpha > 0$ was our simple strategy for capturing concerns with labor supply elasticity. In the simplest stationary equilibria of the model, the parties propose piecewise linear post-fisc distributions of income, with the same treatment for what may be a quite large interval of middle-income voters. The more "swing-voter" concerns

dominate in the parties, the larger will this interval be. But even on the lower envelope of the equilibrium manifold, where parties are most partisan, the policies will coincide for a non-negligible fraction of the income distribution.

We raise now the issue of multiple equilibria. A central problem in modeling political competition is conceiving of it in such a way that equilibria exist when policy spaces are multi-dimensional. We have solved that problem, but have instead a plethora of equilibria. Indeed, the bi-dimensionality of the equilibrium manifold of 1-equilibria here arises from their being, loosely speaking, two payoff functions in each party, associated with its two "factions." Roemer (2001, Chap. 8) shows that, in the related equilibrium concept of PUNE, the bi-dimensionality of the equilibrium manifold can be interpreted as due to there being two missing parameters in the model, which express the relative bargaining powers of the two factions in the two parties. (Every PUNE, in that model, can be generated as an outcome of a Nash-type bargaining game between factions, but with variable relative bargaining powers.) A similar interpretation may be possible for the equilibria in this paper.

Our most important result is that, in *all* equilibria, there is a substantial interval of "middle class" voters, endogenously determined, for which both Left and Right parties propose the same tax treatment. The empirical section intended to test this prediction.

We did so by using the income-tax data assembled by Piketty and Saez (2006), with an amendment to transform their *post-tax* data to *post-fisc* data. We believe that we have found at least mild support for the central prediction. Indeed, as this paper is being written, just prior to the November 2008 US elections, the non-partisan Tax Policy Center (a joint venture of the Urban Institute and the Brookings Institution) reports that the tax plans of Obama and McCain conform to the main prediction of our model (see the epigraph to Sect. 4).

It is perhaps appealing to view certain aspects of tax policy as being due to *simplicity* or *inertia*: thus, one might conjecture that piece-wise linear policies with a small number of pieces are adopted for reasons of simplicity, and that policies do not change much between Left and Right administrations for a large group of middle-income voters for reasons of both simplicity (costly to change the entire tax code) and inertia. We have shown, however, that these characteristics of policies derive from political competition—they are equilibrium characteristics. We need not appeal to simplicity and inertia.

One could point to many ways in which the model simplifies real politics. As just indicated, one of the most important is that detailed tax policy is not proposed by parties in general elections: it is the consequence of legislation, and in particular, of legislative bargaining between the parties, and between the Congress and the executive branch. Another is that taxes and transfer payments are typically dealt with under separate pieces of legislation. Modeling the problem of tax policy as a legislative bargaining problem is, of course, an entirely different undertaking. Nevertheless, if we take Schumpeter's dictum seriously, as stated in the paper's epigram, and also Riker's (1982) dictum, that the most important moment of democracy is the general election, then the investigation reported here should shed some light on the qualitative features of tax policy.

Acknowledgments I am grateful to Seok-ju Cho, Philippe De Donder, and Michael Graetz for helpful discussions. I thank Joseph Bafumi for sharing his data with me on vote shares, and Kenneth Couch for assembling some of the data on transfer payments used in Sect. 4. Emmanuel Saez also provided useful advice on the tax data. I thank seminar participants for their comments, when earlier versions of the paper were presented at the ESF workshop on public economics in Marseille-Luminy, the Paris School of Economics (PSE), Yale University, the University of Rochester, and the June 2008 Barcelona conference on political economy. I thank three referees for their useful comments.

Appendix 1

Proof of Proposition 2 1. Note that if $X_a \in M_a(h_*)$ and $X_b \in M_b(h_*)$ then $y = x_a + \alpha h_1 + h_* - h_1$ and $y = x_b + h_*$.

2. Write the budget constraint for a policy $X \in M_a(h_*)$:

$$x_a + \alpha \int_0^{h_1} h \mathrm{d}F(h) + \alpha h_1(1 - F(h_1)) + \int_{h_1}^{h_*} (h - h_1)\mathrm{d}F(h)$$
$$+(h_* - h_1)(1 - F(h_*)) + \alpha \int_{h_*}^{\infty} (h - h_*)\mathrm{d}F(h) = \mu.$$

We can rewrite this equation as:

$$x_a = (1 - \alpha)\left(\int_0^{h_1} h \mathrm{d}F(h) + \int_{h_*}^{\infty} (h - h_*)\mathrm{d}F(h) + h_1(1 - F(h_1))\right).$$

3. Viewing $M_a(h_*)$ as parameterized by h_1, and differentiating the first expression for y in step 1 w.r.t. h_1, we have

$$\frac{\mathrm{d}y}{\mathrm{d}h_1} = \frac{\mathrm{d}x_a}{\mathrm{d}h_1} - (1 - \alpha).$$

Now differentiating the expression derived in step 2 for x_a w.r.t. h_1 gives:

$$\frac{\mathrm{d}x_a}{\mathrm{d}h_1} = (1 - \alpha)(1 - F(h_1)).$$

These two equations together tell us that:

$$\frac{\mathrm{d}y}{\mathrm{d}h_1} = (\alpha - 1)F(h_1) < 0.$$

Therefore the smallest (largest) value of y compatible with a policy's being in $M_a(h_*)$ is associated with $h_1 = h_*$ (respectively, $h_1 = 0$). Using the equation

for x_a in step 2, we have

$$x_a(h_*) = (1-\alpha)\mu, \quad x_a(0) = (1-\alpha)\int_{h_*}^{\infty}(h-h_*)\mathrm{d}F(h),$$

and so these two values of y are given by:

$$y_a(h_*) = (1-\alpha)\mu + \alpha h_*, \quad y_a(0) = h_* + (1-\alpha)\int_{h_*}^{\infty}(h-h_*)\mathrm{d}F(h).$$

4. We perform a similar analysis of policies in $M_b(h_*)$. For any such policy, we may rewrite the budget constraint as:

$$x_b = (1-\alpha)\int_{h_*}^{h_2} h\ \mathrm{d}F(h) + (1-\alpha)h_2(1-F(h_2)) - (1-\alpha)h_*(1-F(h_*)).$$

Differentiating this equation w.r.t. the parameter h_2 gives:

$$\frac{\mathrm{d}x_b}{\mathrm{d}h_2} = (1-\alpha)(1-F(h_2)) > 0;$$

now using the expression for y in step 1, we have

$$\frac{\mathrm{d}y}{\mathrm{d}h_2} = \frac{\mathrm{d}x_b}{\mathrm{d}h_2} > 0.$$

Therefore, the smallest (largest) value of y compatible with a policy's being in $M_b(h_*)$ is associated with $h_2{=}h_*$ (respectively, $h_2{=}\infty$). These two values of y are

$$y_b(h_*) = h_*, \quad y_b(\infty) = (1-\alpha)\int_{h_*}^{\infty}(h-h_*)\mathrm{d}F(h) + h_*.$$

5. To summarize, the number y is associated with a policy in $M_a(h_*)$ if and only if

$$y_a(h_*) \le y \le y_a(0),$$

and y is associated with a policy in $M_b(h_*)$ if and only if

$$y_b(h_*) \le y \le y_b(\infty).$$

Notice that $y_a(0) = y_b(\infty)$; parts A and B of the proposition follow immediately.

6. We prove part C. We have shown that the smallest value of x_a is $(1{-}\alpha)\int_{h_*}^{\infty}(h{-}h_*)\mathrm{d}F(h)$ which is positive, as long as F has some support on (h_*, ∞). The argument in step 4 above shows that $x_b{>}0$ except in the

singular case that $h_*=h_2$. In that case, the policy X_b is the laissez-faire policy $X_b(h)=h$. □

Proof of Theorem 1 1. The theorem will be proved if we can show that X^a and X^b solve the programs in conditions ($\beta 2$) and ($\beta 3$) of Definition 2, respectively. We first address ($\beta 2$). Let the density function of F be denoted f. The numbers h_1 and $h_2$come with the functions X^a and X^b.

2. Define the number ρ, the functions $r(h)$ on $[0, h_1]$, $s(h)$ on $[h_1, h_*]$ and $t(h)$ on $[h_*, \infty)$, and the number λ as follows.
   (i) $\rho = \frac{\int_0^{h_1} \theta(h) dF(h)}{F(h_1)}$,
   (ii) $r(0) = 0$ and $r'(h) = (\theta(h) - \rho) f(h)$ on $[0, h_1]$,
   (iii) $s(h_1) = 0$ and $s'(h) = (\rho - \theta(h)) f(h)$ on $[h_1, h_*]$,
   (iv) $t(h_*) = \rho(1 - F(h_*)) - \int_{h_*}^{\infty} \theta(h) \mathrm{d}F(h)$ and
   $t'(h) = (\theta(h) - \rho) f(h)$ on $[h_*, \infty)$
   (v) $\lambda = s(h_*) + t(h_*)$.

   Note that $\rho > 0$. Note, from Proposition 3, that the function r is first increasing and then decreasing. Compute that $r(h_1) = r(0) + \int_0^{h_1} r'(h) \mathrm{d}h = 0$. Therefore r is a non-negative function on its domain. Note from Proposition 3 that s is an increasing function on its domain: since θ is constant on $[h_1, h_*]$, by Proposition 3, we know that $\rho > \theta(h)$ on this interval. Therefore s is a non-negative function on its domain, and $s(h_*) = \rho(F(h_*) - F(h_1)) - \int_{h_1}^{h_*} \theta(h) \mathrm{d}F(h) > 0$. Note that t is decreasing on its domain, and $t(\infty) = t(h_*) + \int_{h_*}^{\infty} t'(h) \mathrm{d}h = 0$. Therefore t is non-negative on its domain. Finally, note that $\lambda > 0$.

3. Suppose that X^a were not the solution to the program ($\beta 2$) of Definition 2, and that the true solution is some other policy X. Define the function g by the equation $X(h) = X^a(h) + g(h)$. Now define the function $\Delta : \mathbb{R} \to \mathbb{R}$ as follows.

$$\Delta(\varepsilon) = \int_0^{\infty} (X^a(h) + \varepsilon g(h))\theta(h) dF(h) + \int_0^{h_1} \left(X^{a\prime}(h) + \varepsilon g'(h) - \alpha\right) r(h) \mathrm{d}h$$
$$+ \int_{h_1}^{h_*} \left(1 - (X^{a\prime}(h) + \varepsilon g'(h))\right) s(h) \mathrm{d}h + \int_{h_*}^{\infty} \left(X^{a\prime}(h) + \varepsilon g'(h) - \alpha\right) t(h) \mathrm{d}h$$
$$+ \lambda \left(X^a(h_*) + \varepsilon g(h_*) - y\right) + \rho \left(\mu - \int_0^{\infty} (X^a(h) + \varepsilon g(h)) \mathrm{d}F(h)\right).$$

Note that Δ is a linear function, and that $\Delta(0) = \int_0^{\infty} X^a(h)\theta(h) \mathrm{d}F(h)$: this is the objective of program ($\beta 2$) evaluated at the policy X^a. Note as well that when $\varepsilon = 1$, all the terms in the expression defining Δ are non-negative: this follows from the fact that $r, s, t, \lambda,$ and ρ are all non-negative functions or numbers, and that $X \in \Im$. Suppose we can show that $\Delta'(0) = 0$: then Δ will be equal to a

constant, and consequently $\Delta(0) = \Delta(1)$. But this implies that the value of the objective function of ($\beta 2$) at X^a is at least as large as its value at X: a contradiction. Thus we will have proved that X^a solves program ($\beta 2$) if we can show that $\Delta'(0) = 0$.

4. Compute that

$$\Delta'(0) = \int_0^\infty \theta(h)g(h)\mathrm{d}F(h) + \int_0^{h_1} g'(h)r(h)\mathrm{d}h - \int_{h_1}^{h_*} g'(h)s(h)\mathrm{d}h$$
$$+ \int_{h_*}^\infty g'(h)t(h)\mathrm{d}h + \lambda g(h_*) - \rho \int_0^\infty g(h)\mathrm{d}F(h)$$

Hence, integrating three times by parts, we have

$$\Delta'(0) = \int_0^\infty \theta(h)g(h)\mathrm{d}F(h) + g(h)r(h)|_0^{h_1} - \int_0^{h_1} r'(h)g(h)\mathrm{d}h$$
$$- g(h)s(h)|_{h_1}^{h_*} + \int_{h_1}^{h_*} s'(h)g(h)\mathrm{d}h + g(h)t(h)|_{h_*}^\infty - \int_{h_*}^\infty t'(h)g(h)\mathrm{d}h$$
$$+ \lambda g(h_*) - \rho \int_0^\infty g(h)\mathrm{d}F(h).$$

We next re-group terms and write:

$$\Delta'(0) = \int_0^{h_1} \left((\theta(h) - \rho)f(h) - r'(h)\right) g(h)\mathrm{d}h + \int_{h_1}^{h_*} \left(s'(h)\right.$$
$$- \left.(\rho - \theta(h))\, f(h)\right)g(h)\mathrm{d}h + \int_{h_*}^\infty \left((\theta(h) - \rho)f(h) - t'(h)\right) g(h)\mathrm{d}h$$
$$+ g(h_*)(\lambda - s(h_*) - t(h_*)) - g(0)r(0) + g(h_1)(r(h_1) + s(h_1))$$
$$+ t(\infty)g(\infty).$$

Now check, by the definitions of r, s, t, and λ that every term on the r.h.s. of this equation vanishes, which proves that $\Delta'(0) = 0$.

5. We proceed to prove that X^b is the solution to program ($\beta 3$) of Definition 2. Suppose that the true solution is X and now define the function g by $X = X^b + g$. We now define functions R, S, and T, and numbers γ and δ as follows:

i. $\delta = \int_{h_2}^{\infty} (1-\theta(h))dF(h)/(1-F(h_2))$,
ii. $R(0) = 0$ and $R'(h) = (\delta - (1-\theta(h))f(h)$ on $[0, h_*]$,
iii. $S(h_*) = \int_{h_2}^{h_*}(\delta - (1-\theta(h))dF(h)$ and $S'(h) = (1-\theta(h)-\delta)f(h)$ on (h_*, h_2),
iv. $T(h_2) = 0$ and $T'(h) = (\delta - (1-\theta(h))f(h)$ on (h_2, ∞),
v. $\gamma = R(h_*) + S(h_*)$.

Since the function $1-\theta(h)$ is (weakly) increasing (see Proposition 3), it follows from the definition of δ that $R' \geq 0$, $S' \leq 0$, and that $S'(h_2) = 0$. The functions R, S, and T are non-negative on their domains. As well, $R(h_*)$, $S(h_*)$ and γ are positive.

6. We now define the function Φ by:

$$\begin{aligned}\Phi(\varepsilon) = &\int_0^{\infty} (1-\theta(h))(X^b(h) + \varepsilon g(h))\mathrm{d}F(h)\\ &+\int_0^{h_*}\Big(1-(X^{b\prime}(h)+\varepsilon g'(h)\Big)R(h)\mathrm{d}h\\ &+\int_{h_*}^{h_2}\Big(X^{b\prime}(h)+\varepsilon g'(h)-\alpha\Big)S(h)\mathrm{d}h\\ &+\int_{h_2}^{\infty}\Big(1-(X^{b\prime}(h)+\varepsilon g(h))\Big)T(h)\mathrm{d}h+\gamma\Big(X^b(h_*)+\varepsilon g(h_*)\Big)\\ &+\delta\left(\mu-\int_0^{\infty}(X^b(h)+\varepsilon g(h))\mathrm{d}F(h)\right).\end{aligned}$$

All the terms on the r.h.s. of this equation are non-negative, and so, as we argued above, if we can demonstrate that $\Phi'(0) = 0$, then we will have proved that X^b solves the program in condition ($\beta 3$) of Definition 2.

7. Compute that

$$\begin{aligned}\Phi'(0) = &\int_0^{\infty}(1-\theta(h))g(h)\mathrm{d}F(h) - g(h)R(h)|_0^{h_*} + \int_0^{h_*}R'(h)g(h)\mathrm{d}h\\ &+ g(h)S(h)|_{h_*}^{h_2} - \int_{h_*}^{h_2}S'(h)g(h) - g(h)T(h)|_{h_2}^{\infty} + \int_{h_2}^{\infty}T'(h)g(h)\mathrm{d}h\\ &+\gamma g(h_*) - \delta\int_0^{\infty}g(h)\mathrm{d}F(h).\end{aligned}$$

Re-grouping terms, we have

$$\Phi'(0) = \int_0^{h_*} ((1-\theta(h)-\delta)f(h) + R'(h))g(h)\mathrm{d}h + \int_{h_*}^{h_2} ((1-\theta(h)-\delta)f(h)$$
$$- S'(h))g(h)\mathrm{d}h + \int_{h_2}^{\infty} ((1-\theta(h)-\delta)f(h) + T'(h))g(h)\mathrm{d}h$$
$$+ (\gamma - R(h_*) - S(h_*)) - g(0)R(0) + g(h_2)(S(h_2)$$
$$+ T(h_2)) - g(\infty)T(\infty).$$

From the definitions of the functions R, S, T and the numbers δ and γ, we observe that all terms on the r.h.s. of this equation vanish, which proves the theorem.[17]

□

Proof of Theorem 2 1. It is clear that the ideal policy for a type h_*—the policy in $\Im$ that maximizes its (post-fisc) income—has some value y at h_*, increases as slowly as possible for $h > h_*$, and decreases from the (h_*, y) as rapidly as possible for $h < h_*$. This is the way to spend as few resources as possible on everyone other than h_*. Thus the ideal policy for h_* is defined by:

$$X^*(h) = \begin{cases} x_0 + h, & h \le h_* \\ x_0 + h_* + \alpha(h - h_*), & h > h_* \end{cases}$$

where x_0 is such that this policy integrates to μ. But this is precisely the policy in $M_a(h_*) \cap M_b(h_*)$ when $y = y_{\max}(h_*)$.

2. If $h_* \le \mu$ and $y = (1-\alpha)\mu + \alpha h_*$then the policy $X^a \in M_a(h_*)$ is a line of slope α such that $x_a = (1-\alpha)\mu$. We prove below (see the proof of Theorem 5), using the variational technique of the proof of Theorem 1, that this is the policy that maximizes $\int \theta(h)X(h)\mathrm{d}F(h)$ on $\Im$. Moreover, the fact is intuitively clear. Because θ is a decreasing function and the objective functional is linear in X, the objective wishes to push resources as much as possible to the poorest. The solution is to maximize what is given to $h = 0$, which means to increase as slowly as possible (that is, at rate α) on the whole positive line, subject to having given just enough to $h = 0$ so that the policy integrates to μ.
3. If $h_* \ge \mu$ and $y = y_{\min}(h_*) = h_*$ then the policy $X^b \in M_b(h_*)$ is the laissez-faire policy $X(h) = h$. It is also intuitively clear that this is the policy that maximizes $\int (1-\theta(h))X(h)\mathrm{d}F(h)$: for now, the objective wishes to push resources to the very rich. Once it is decided how much the very rich get, the strategy must be to decrease as fast as possible (i.e., at rate one) for h smaller. This yields in the limit the laissez-faire policy. Of course, this can also be proved using the variational technique of Theorem 1. □

[17] The convention that "$t(\infty) = 0 = T(\infty)$" is short-hand for a transversality condition. The proof can be rigorously completed by checking that $\lim_{h\to\infty} t(h)g(h) = 0$ and $\lim_{h\to\infty} T(h)g(h) = 0$: these claims are true.

Proof of Theorem 3 1. In the political contest at date 1, the L party (the one whose constituency is defined by the function θ_0) solves this program:

$$\max_{X \in \Im} \int \theta_0(h) X(h) \mathrm{d}F(h)$$
$$\text{s.t.}$$
$$X(h_*) \geq X^{\mathrm{R}}(h_*)$$

since the unique swing type at date 0 is h_*. There is a similar program for the R party. Thus the political equilibrium at date one is exactly a triple $(X^{\mathrm{L}}, X^{\mathrm{R}}, y)$ such that:

(1) X^{L} maximizes $\int \theta_0(h) X(h) \mathrm{d}F(h)$ subject to $X(h_*) \geq y$;
(2) X^{R} maximizes $\int (1 - \theta_0(h)) X(h) \mathrm{d}F(h)$ subject to $X(h_*) \geq y$;
(3) $X^{\mathrm{L}}(h_*) = y = X^{\mathrm{R}}(h_*)$.

This looks almost like a 1-stationary equilibrium, except that the function θ_0 is not related to the vote shares engendered by the policies X^{L} and X^{R}. Now suppose we choose $y \in [y_{\min}(h_*), y_{\max}(h_*)]$ so that $(h_*, y) \in \Gamma$. Note, by examining the proof of Theorem 1, that the only fact invoked about the function θ_* was that it was weakly monotone decreasing. So the argument of theorem will apply just as well if we substitute the decreasing function θ_0 for θ_*. Hence, by the same argument as in Theorem 1, the functions X^{L} and X^{R} which satisfy conditions (1)-(3) above are an ordered pair in the set $M_a(h_*) \times M_b(h_*)$. This proves the first statement in the theorem.

2. Conversely, let $(X^{\mathrm{L}}, X^{\mathrm{R}})$ be any equilibrium (stationary or not) at date 1 emanating from θ_0. The policies have a common value $y = X^{\mathrm{L}}(h_*) = X^{\mathrm{R}}(h_*)$ since h_* was the unique swing voter. If $(h_*, y) \in \Gamma$, then again the optimization of each party yields inexorably to the ordered pair of policies $M_a(h_*) \times M_b(h_*)$ associated with the income y at the pivot h_*.

Thus, what must be shown is that $y = X^{\mathrm{L}}(h_*) \in [y_{\min}(h_*), y_{\max}(h_*)]$. Define the following two functions:

$$X^{\max}(h) = \begin{cases} (y - \alpha h_*) + \alpha h, & 0 \leq h \leq h_* \\ y - h_* + h, & h > h_* \end{cases}$$
$$X^{\min}(h) = \begin{cases} y - h_* + h, & 0 \leq h \leq h_* \\ y + \alpha(h - h_*), & h > h_* \end{cases}$$

Subject to passing through the point $(h_*, y)$and satisfying conditions (P0) and (P1) of the definition of policies in $\Im$, $X^{\max}$ is the way of allocating income which consumes the maximal amount of resource and $X^{\min}$ is the way of allocating income which consumes the minimal amount of resource. It therefore follows that

$$\int X^{\max}(h) \mathrm{d}F(h) \geq \mu \quad \text{and} \quad \int X^{\min}(h) \mathrm{d}F(h) \leq \mu.$$

But this is precisely the condition for (h_*, y) to be in Γ. □

Proof of Theorem 4 1. We prove the converse part first. Let $(X^{\mathrm{L}}, X^{\mathrm{R}})$ be the date-1 equilibrium in a sequence of historical equilibrium beginning with the datum θ_0. Define for $h \in [h_*, h_{**}]$, $Z(h) \equiv X^{\mathrm{L}}(h)$. From the definition of equilibrium, it follows that for $h \in [h_*, h_{**}]$, $X^{\mathrm{R}}(h) = Z(h)$ as well. Suppose that the functions $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$ defined in the theorem's statement can be defined and integrate to μ. We show that X^{L} and X^{R} are precisely the functions $\hat{X}^{\mathrm{L}}$ and $\hat{X}^{\mathrm{R}}$. Suppose, to the contrary, that $X^{\mathrm{L}} \neq \hat{X}^{\mathrm{L}}$. Then define the (non-zero) function g by $X^{\mathrm{L}}(h) = \hat{X}^{\mathrm{L}}(h) + g(h)$.

Let the functions $r(\cdot), s(\cdot), t(\cdot),$ and $u(\cdot)$ be some non-negative functions, defined on the intervals given by the limits of the integrals in which they appear below, and let λ_1, λ_2 and ρ be arbitrary non-negative numbers. Now define the function:

$$\begin{aligned}\Delta(\varepsilon) = &\int_0^\infty \theta_0(h)(\hat{X}^{\mathrm{L}}(h) + \varepsilon g(h))\mathrm{d}F(h) + \int_0^{h_1} \varepsilon g'(h) r(h)\mathrm{d}h \\ &- \int_{h_1}^{h_*} \varepsilon g'(h) s(h)\mathrm{d}h + \lambda_1 \varepsilon g(h_*) + \int_{h_*}^{h_{**}} \varepsilon g(h) t(h)\mathrm{d}h + \lambda_2 \varepsilon g(h_{**}) \\ &+ \int_{h_{**}}^\infty \varepsilon g'(h) u(h)\mathrm{d}h - \rho \int_0^\infty \varepsilon g(h)\mathrm{d}F(h).\end{aligned}$$

Note that $\Delta(0)$ is the value of the L party's objective at date 1 evaluated at $\hat{X}^{\mathrm{L}}$ and $\Delta(1)$ is the value of the L party's objective at X^{L} plus a series of terms all of which must be non-negative. (For example, in the interval $[0, h_1]$, $g' \geq 0$ because $(\hat{X}^{\mathrm{L}})'(h) \equiv \alpha$ on this interval, and so on for all the other terms.)

Suppose we can choose $r, s, t, u, \lambda_1, \lambda_2,$ and ρ so that $\Delta'(0) = 0$. Since Δ is a linear function, it will follow that it is maximized at $\varepsilon = 0$; in particular, $\Delta(0) \geq \Delta(1)$. This implies, a fortiori, that the value of the L party's objective is at least as great at $\hat{X}^{\mathrm{L}}$ as at X^{L}, which will be the desired contradiction.

2. Calculate, using integration by parts, that:

$$\begin{aligned}\Delta'(0) = &\int_0^\infty \theta_0(h) g(h)\mathrm{d}F(h) + g(h) r(h)|_0^{h_1} - \int_0^{h_1} g(h) r'(h)\mathrm{d}h \\ &- g(h) s(h)|_{h_1}^{h_*} + \int_{h_1}^{h_*} g(h) s'(h)\mathrm{d}h + \lambda_1 g(h_*) + \int_{h_*}^{h_{**}} g(h) t(h)\mathrm{d}h \\ &+ \lambda_2 g(h_{**}) + g(h) u(h)|_{h_{**}}^\infty - \int_{h_{**}}^\infty g(h) u'(h)\mathrm{d}h - \rho \int_0^\infty g(h)\mathrm{d}F(h);\end{aligned}$$

now organize the terms above to express:

$$\Delta'(0) = \int_0^{h_1} [\theta_0(h)f(h) - \rho f(h) - r'(h)]g(h)\mathrm{d}h + \int_{h_1}^{h_*} [\theta_0(h)f(h) - \rho f(h)$$
$$+ s'(h)]g(h)\mathrm{d}h + \int_{h_*}^{h_{**}} [\theta_0(h)f(h) - \rho f(h) + t(h)]g(h)\mathrm{d}h$$
$$+ \int_{h_{**}}^{\infty} [\theta_0(h)f(h) - \rho f(h) - u'(h)]g(h)\mathrm{d}h + r(0)g(0) + g(h_1)(r(h_1)$$
$$+ s(h_1)) + g(h_*)(\lambda_1 - s(h_*)) + g(h_{**})(\lambda_2 - u(h_{**})) + g(\infty)u(\infty),$$

where $f(h) \equiv \mathrm{d}F(h)$.
Therefore, we can annihilate all these terms if the "Lagrangian functions and multipliers" are chosen to fulfill the following equations:
(a) $r'(h) = (\theta_0(h) - \rho)f(h)$ on $[0, h_1]$
(b) $s'(h) = (\rho - \theta_0(h))f(h)$ on $[h_1, h_*]$
(c) $t(h) = (\rho - \theta_0(h))f(h)$ on $[h_*, h_{**}]$
(d) $u'(h) = (\theta_0(h) - \rho)f(h)$ on $[h_*, h_{**}]$
(e) $r(0) = 0 = r(h_1) = s(h_1)$
(f) $\lambda_1 = s(h_*)$
(g) $\lambda_2 = u(h_{**})$
(h) $u(\infty) = 0$.

3. Since r must be zero at its endpoints (statement (e)), using (a), we define:

$$\rho = \theta^{av}[0, h_1] \equiv \frac{\int_0^{h_1} \theta_0(h)\mathrm{d}F(h)}{F(h_1)}.$$

Since θ_0 is a weakly decreasing function, it follows that r is non-negative on $[0, h_1]$. Obviously $\rho > 0$. It now follows from statement (b) that $s' \geq 0$ on $[h_1, h_*]$, again invoking the fact that θ_0 is decreasing. Hence, $s(h_*) = \int_0^{h_1} s'(h)\mathrm{d}h \geq 0$ (here we use the fact that we choose $s(h_1) = 0$). Hence from (f), $\lambda_1 \geq 0$ and s is a non-negative function on its domain. From (c), $t(h)$ is a non-negative function, again invoking the fact that θ_0 is decreasing. Now from (d), u must be a decreasing function on $[h_{**}, \infty)$ and must converge to zero at infinity, so we define:

$$u(h_{**}) = \int_{h_{**}}^{\infty} (\rho - \theta_0(h))\mathrm{d}h > 0.$$

Hence from (g), $\lambda_2 > 0$. Hence the Lagrangian functions and multipliers have been defined, to be non-negative, and to fulfill the conditions (a)–(h), proving the claim.[18]

A similar argument shows that $X^{\mathrm{R}} = \hat{X}^{\mathrm{R}}$.

4. Finally, we remark that indeed the functions $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$ *can* be defined and integrate to μ: this argument is just like the one presented in the proof of Theorem 3.
5. We have shown that any date-1 equilibrium with the historical vote-share function θ_0 is of the form $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$. The first statement in the theorem is clearly proved by the same technique. That is, if the functions $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$ can be defined (which is true if the integral of Z on its domain is not to small or too large) then they comprise a stationary equilibrium reached at date 1. To show stationarity, we need only observe that the vote share function $\hat{\theta}(\cdot)$ defined by $(\hat{X}^{\mathrm{L}}, \hat{X}^{\mathrm{R}})$ is itself monotone decreasing, and the same optimization proof works.

□

Appendix 2: Amending the Piketty–Saez (2006) data to include transfer payments

I explain how we amended the Piketty–Saez data to attain the average adjusted post-fisc tax rate for a quantile—here, for the bottom quintile. Let

$$\begin{aligned}
x_{20} &= \text{post-tax income of bottom quintile}\\
y_{20} &= \text{pre-tax income of bottom quintile}\\
\bar{x} &= \text{average post-tax income of whole sample}\\
\bar{y} &= \text{average pre-tax income of whole sample}\\
T_{20} &= \text{average taxes paid by bottom quintile}\\
t_{20} &= \text{average transfers received by bottom quintile}\\
\bar{T} &= \text{average taxes paid, whole sample}\\
\bar{t} &= \text{average transfers received, whole sample}
\end{aligned}$$

We wish to compute

$$q_{20} = \text{ average post-fisc tax rate of bottom quintile} = \frac{T_{20} - t_{20}}{y_{20}}.$$

Piketty–Saez (2006) give us the pre-tax income share

$$r_{20} = \frac{20 y_{20}}{\bar{y}}, \tag{1}$$

and the average tax rate for the bottom quintile

[18] As we remarked in the proof of Theorem 1, the statement "$u(\infty) = 0$" is short-hand for the statement $\lim_{h\to\infty} g(h)u(h) = 0$, which is true.

$$u_{20} = \frac{T_{20}}{y_{20}} \tag{2}$$

Table A0 of Piketty–Saez (2006) gives the average nominal income, which is $\bar{y}$.

Hence, from (1), we can compute y_{20}; from (3) we can compute T_{20}. We extract t_{20} from the PSID, and thus compute q_{20}.

For 1974, Piketty–Saez (2006) does not provide the necessary data for the quantiles in the bottom 90% of the income distribution. We proceeded as follows. First, we calculated the income shares from the Current Population Survey in 2006 for the quantiles in the bottom 90%, and also those shares from the same source in 2004. Then we applied the factors by which those shares changed to the income shares provided by Piketty–Saez (2006) for 2001. This gave us the necessary income shares for all quantiles in 2004, which permitted us to compute the values shown in Fig. 7c.

Finally, the average adjusted post-fisc income was calculated by charging taxpayers a tax rate proportional to income to finance the deficit in the year in question. The deficit used is that reported in the *Economic Report of the President.*

References

Austen-Smith D (2000) Redistributing income under proportional representation. J Political Econ 108:1235–1269

Besley T, Coate S (1997) An economic model of representative democracy. Q J Econ 112:85–114

Brownlee WE (2004) Federal taxation in America: a short history. Cambridge University Press, New York

Carbonell-Nicolau O, Ok E (2007) Voting over income taxation. J Econ Theory 134:249–286

Cox G (2006) Swing voters, core voters and distributive politics. Working paper, http://www.yale.edu/polisci/info/conferences/Representation/papers/Cox.pdf

Current Population Survey (2006) Annual Demographic survey: march supplement. http://pubdb3.census.gov/macro/032006/rdcall/1_001.htm

Dixit A, Londregan J (1998) Ideoglogy, tactics, and efficiency in redistributive politics. Q J Econ 113: 497–529

Levy G (2004) A model of political parties. J Econ Theory 115:250–277

Lindbeck A, Weibull J (1987) Balanced-budget redistribution as the outcome of political competition. Public Choice 52:273–297

McCarty N, Poole K, Rosenthal H (2006) Polarized America. MIT Press, Cambridge, MA

Osborne M, Slivinski A (1996) A model of political competition with citizen–candidates. Q J Econ 111: 65–96

Piketty T, Saez E (2006) How progressive is the US federal tax system? An historical and international perspective. http://elsa.berkeley.edu/~saez/piketty-saezNBER06taxprog.pdf (Published in J Econ Perspect 21:3–24 2007, but without the data appendices)

Riker W (1982) Liberalism against populism: a confrontation between the theory of democracy and the theory of social choice. WH Freeman, San Francisco, CA

Roemer JE (1999) The democratic political economy of progressive income taxation. Econometrica 67: 1–19

Roemer JE (2001) Political competition. Harvard University Press, Cambridge, MA

Roemer JE (2006) Democracy, education and equality. Cambridge University Press, New York

Roemer JE (2009) A problem in conceptualizing party competition over redistributive taxation. In: Aragonès E, Beviá C, Llavador H, Schofield N. (eds) The political economy of democracy, Chap 15. Fundación BBVA, Bilbao, Spain, pp 259–266

Schumpeter J (1954 [1918]) The crisis of the tax state. In: Peacock A (ed) International economic papers: translations prepared for the International Economic Association, no 4. Macmillan, London, pp 5–38

Soc Choice Welf
DOI 10.1007/s00355-010-0515-0

Bandwagon, underdog, and political competition: the uni-dimensional case

Woojin Lee

Abstract This article studies the effects of bandwagon and underdog on the political equilibrium of two-party competition models. We adapt for voter conformism the generalized Wittman–Roemer model of political competition, which views political competition as the one between parties with factions of the opportunists and the militants that Nash-bargain one another, and consider three special cases of the general model: the Hotelling–Downs model, the classical Wittman–Roemer model, and what we call the ideological-party model. We find that the presence of voter conformism significantly affects the nature of political competition, and its effect on political equilibrium is quite different depending on the model one uses. In the Hotelling–Downs model, political parties put forth an identical policy at the equilibrium, *regardless of* the type of voter conformism, and this is the only equilibrium. In both the ideological-party and classical Wittman–Roemer models, parties propose differentiated policies at the equilibrium, and the extent of policy differentiation depends on the degree of voter conformism.

1 Introduction

In almost all democracies, opinion polls have become an integral part of national elections. Public opinion polls provide aggregate information to the public about the views of their fellow citizens. By doing so, they may sometimes influence the behavior of voters and thus who will be elected. In turn, opinion polls may influence announced policies of candidates as well.

W. Lee (✉)
Department of Economics, Korea University, Anam-dong, Seongbuk-Gu, Seoul, Korea
e-mail: woojinlee@korea.ac.kr

Published online: 29 December 2010

The many theories about how this happens can be divided into three categories: 'voter conformism,' 'strategic voting,' and 'participation/abstention.'

A well-known example of voter conformism is the *bandwagon effect*. The bandwagon effect occurs when the opinion poll prompts voters to support the candidate shown to be winning in the poll, thus increasing his/her chances of being on the winner's side in the end. The idea that voters are susceptible to such an effect is old, and has remained persistent in spite of much debate on its empirical existence. Bartels (1985, 1988), for instance, shows that voters are motivated in part by a desire to vote for the winning candidate. The opposite of the bandwagon effect is the *underdog effect*; this happens when people support, out of sympathy, the candidate perceived to be 'losing' the elections. In a meta-study of research on this topic, Irwin and Van Holsteyn (2000) show that from the 1980s onward, empirical evidence for the existence of the bandwagon effect is found more often than for the underdog effect.[1]

The second category of theories on the effect of polls on voting concerns *strategic voting*. These theories are based on the idea that voters will sometimes not choose the candidate they prefer the most, but another, less-preferred, candidate from strategic considerations. An example can be found in the UK general election of 1997. Then Cabinet Minister, Michael Portillo's constituency of Enfield was believed to be a safe seat, but opinion polls showed that the Labour candidate Stephen Twigg was steadily gaining support, which may have prompted supporters of other parties to vote for Twigg in order to remove Portillo.[2]

The third category of theories concerns *voter participation/abstention*. It is often suggested that supporters of the candidate shown to be significantly lagging behind may give up casting their ballots, resulting in a landslide victory of another candidate. In the South Korean presidential election of 2007, when a conservative candidate, M.B. Lee, achieved a landslide victory over a liberal candidate, D.Y. Chung, with the vote shares of 48.7 vs. 26.1%, it was widely believed that anti-Lee voters had abstained significantly, concluding from several pre-election polls that Chung would have no chance of winning even if they would cast votes for him. (Indeed the voting rate was 63%, the lowest one since 1987.) But the opposite of this phenomenon may happen as well. A well-known example is the *boomerang effect* where the likely supporters of the candidate shown to be winning feel that they are 'home and dry' and that their vote is not required, thus allowing another candidate to win.

Since Leibenstein (1950)'s pioneering work on consumer conformism, there have been many studies on the effect of conformism on economic behavior; see Banerjee (1992), Bernheim (1994), Birkhchandani et al. (1992), and Schelling (1978). There

[1] There have been at least two explanations for the existence of voter conformism. The first consists in assuming that polls may exert a normative influence over voters; when voters perceive the existence of a *social norm*—defined by a majority preference expressed in polls in the case of a bandwagon effect—they may feel compelled to abandon their views and comply with such norms, to avoid perhaps cognitive dissonance. The second, which seems more compelling, consists in assuming that individuals may be influenced by polls because they use revealed public preferences as *information* about the correct option to take. Considering they have strong incentives to minimize the costs of acquiring the information necessary to make right choices (Downs 1957), voters may rely upon 'information shortcuts,' such as group references, party identification, or knowledge about where other voters stand on issues.

[2] See http://www.pollsoncall.com/opinion-polls.htm.

are also some political models incorporating the effect of opinion polls on voter conformism and its consequence for actual vote shares (Aldrich 1980; Baumol 1957; Callander 2007; Myerson and Weber 1993; Simon 1954)

To the best of our knowledge, no political models have been developed to study the effect of voter conformism on the *nature of political competition* in a large polity. This article aims at filling this gap in the literature. In particular, we are interested in the effect of voter conformism, in the form of bandwagon or underdog, on *equilibrium party platforms*.

We ask the following: does the presence of voter conformism affect the policy positions of candidates? If it does, does it mitigate policy differentiation among candidates, or exacerbate it?

To answer these questions, we adapt the generalized Wittman–Roemer model of two-party competition for voter conformism. Instead of viewing political competition as occurring between two parties each of which is a unitary actor that maximizes a certain payoff function, the generalized Wittman–Roemer model views political equilibrium as the one obtained from competition between parties with factions that have different goals and Nash-bargain with one another to set the policy. Following Roemer (2001), we assume that there are two factions in each party: the opportunists whose goal is to win the election and the militants whose objective is to maximize the average well-being of their party members.[3]

The generalized Wittman–Roemer model has one advantage for our study; it covers various models of political competition as its special cases. Thus it allows us to study the consequence of voter conformism on the equilibrium of various political models in a unified framework. We will study three special cases of the generalized Wittman–Roemer model of political competition, which have received much attention among students of political economy. One is the Hotelling–Downs model in which parties maximize their probabilities of victory, and another is the classical Wittman–Roemer model (Roemer 1997) in which parties maximize the expected utilities of their key constituents. The third is the one, which we call the ideological-party model, in which each party sets its policy equal to the ideal tax rate of its endogenously determined average member.

In defining voter conformism, we follow Simon (1954). Simon (1954) holds that voting behavior is a function of voters' *expectations* of the electoral outcome, and these expectations are influenced by published poll data. The bandwagon effect exists if voters are more likely to vote for a candidate when they expect him/her to win than when they expect him/her to lose; if the opposite holds, the underdog effect exists.

Present voter conformism, voters' voting behavior is influenced not only by the policy difference but also by the expected vote share difference between political parties. The fact that voters are partly influenced by the expected vote share difference may have a dramatically different implication for the nature of political competition.

[3] A generalized Wittman–Roemer equilibrium, where bargaining power is fixed, can be considered a special case of Roemer (2001)'s party unanimity Nash equilibrium, where bargaining power is not specified a priori.

Were there no voter conformism, a party changing its policy to the direction of the policy position of the opposition party would always increase its vote share; it would take some of the voters who would vote for the opposition party otherwise.

If voter conformism (say, bandwagon) is present, however, choosing a policy that is too close to the opposition party's position may not be desirable, even in terms of the vote share. This is because as the policy difference of the two parties becomes smaller, voters' voting behavior is influenced more by the expected vote share difference than by the policy difference. If the party is currently winning in opinion polls, it would reinforce the party's electoral strength. If the party is currently losing in opinion polls, on the other hand, choosing a policy that is too close to the opposition party's policy position would be *devastating* for the party. If, for example, a party slightly lagging behind in opinions polls takes the opposition party's position, voters do not see any policy difference between the parties and thus vote only according to what the current polls predict; this implies that the party turns itself into a sure loser. Thus if the party is currently losing in opinion polls, *moving away* from the opposition party's policy, within a certain limit, can be a better strategy for increasing the vote share.

We will make these statements precise in section 2, but this difference from the standard models without voter conformism is important for many of the results obtained in this article.

Although we will define the generalized Wittman–Roemer model's equilibrium as a static concept, the issues we are studying—bandwagon, underdog, and policy positions of candidates—are inherently dynamic. Thus we develop a dynamic version of the model, and study some of the model's dynamic properties as well.

Section 2 presents the generalized Wittman–Roemer model of political competition, which is adapted for voter conformism. This section provides a unified framework covering all the special cases that we study in the article. We also make it precise the implication of voter conformism on the nature of political competition that we discussed above. This section also presents a dynamic process, a stationary point of which is identical to the model's equilibrium.

Section 3 studies the effect of voter conformism on the Hotelling–Downs political equilibrium, where the militants have no bargaining power in both parties. We prove that voter conformism has no effect on the equilibrium policy in this case, for the unique equilibrium in this model is both parties' putting forth the same policy. At the same time, however, we argue that the equilibrium a knife-edge equilibrium; it is difficult to justify dynamically.

In section 4 we study another extreme case of the general model: the ideological-party model where the militants have no bargaining power in both parties. In contrast with the Hotelling–Downs model, equilibrium policies are differentiated in this model and the presence of multiple equilibria is generic when the bandwagon effect is sufficiently strong. In those equilibria that are dynamically stable and have the membership share of party L greater than 0.5 (less then 0.5), an increasing bandwagon effect decreases (increases) the equilibrium tax rates of both parties; the opposite holds for the underdog effect.

Section 5 studies the classical Wittman–Roemer model in which both the militants and the opportunists have equal bargaining power in both parties. As in the ideological-party model, political parties in the classical Wittman–Roemer model propose

differentiated equilibrium policies, and multiple equilibria typically exist when the bandwagon effect is sufficiently strong. In contrast with the purely ideological parties, however, the Wittman–Roemer parties move in an *opposite* direction as the parameter capturing the extent of voter conformism increases. In those Wittman–Roemer equilibria that are dynamically stable, an increasing bandwagon effect exacerbates the policy differentiation of the two parties; the opposite holds when the underdog effect is considered.

Section 6 concludes. We collect all the proofs in Appendix.

2 The model

Throughout the article, we will maintain that there are two political parties (or candidates representing them), L and R, and that the policy space is the unit interval: $T = [0, 1]$. A generic element of T will be denoted by t, which we call a tax rate, or simply a policy. We assume that the party that wins the election implements its announced policy. Because we study the models with two parties, the issue of strategic voting is not our concern. Also the potential issue of voter participation/abstention is not explicitly modeled.

There is a continuum of voters; we are modeling an election in large polities, where no individual voter is noticeable. Voters are distributed by one-dimensional characteristic, $w \in \mathbb{R}_+$, with its probability measure $P(.)$.[4] We assume that the associated distribution function is given by a strictly increasing and continuous distribution function $F(.)$. We call w an income. The mean of w, denoted by $\mu \in \mathbb{R}_+$, is assumed to exist.

We postulate a perfectly representative democracy where: (1) every voter belongs to one and only one party (thus there are no 'undecided' voters); (2) each party member receives an equal weight in the determination of the party's von Neumann–Morgenstern utility function; and (3) each voter votes for the party of which he/she is a member.

Suppose $(t_L, t_R) \in T \times T$ is a pair of policy positions of the two parties and $x \in [0, 1]$ is an expected membership share of party L, which is ascertained through opinion polls. (Because there are only two parties, the expected membership share for party R is $1 - x$.) Given (t_j, x_j), where $j = L, R$, we assume that voter preferences are given by

$$(1 - t_j)w + \alpha h(t_j \mu) + \theta \phi(x_j), \tag{1}$$

where α is a positive constant, and $h(.)$ and $\phi(.)$ are functions satisfying the following conditions.

Assumption 1 (1) $\phi : [0, 1] \to \mathbb{R}$ is strictly increasing and finite-valued on $[0, 1]$; and
(2) $h : T \to \mathbb{R}$ is strictly increasing, strictly concave, and finite-valued on T.

Some remarks are in order regarding voter preferences.

[4] $\mathbb{R}_+$ is defined to be the set of all non-negative real numbers, not just of positive numbers.

First, facing an election, voters care not only about the policy positions of political parties but also the membership shares of the parties. Note that the voter's utility function consists of two parts: a quasi-linear utility function that represents the economic interests of voters, $(1-t_j)w+\alpha h(t_j\mu)$, and a utility bonus/penalty from supporting the winning/losing party, $\theta\phi(x_j)$, where $j=L,R$. As $\phi(.)$ is increasing, we have $\theta\phi(x)>\theta\phi(1-x)$ if and only if $\theta(x-\frac{1}{2})>0$. Thus the bandwagon effect is captured by the assumption that θ is *positive*; other things being equal, voters prefer a party whose expected membership share is greater than $\frac{1}{2}$. The presence of the underdog effect would be equivalent to assuming that θ is *negative*. Finally, if $\theta=0$, there is no voter conformism.

Second, if we interpret $t\mu$ as the per capita amount of public goods, then α measures the extent to which voters value the consumption of public goods. The parameter θ, on the other hand, measures the relative salience of voter conformism. By letting α and θ vary across voters, one might allow a voter to be equipped with three characteristics: (w,α,θ). Our equilibrium will be well-defined even in that case.[5] For the sake of simplicity, we will maintain that the parameter values of α and θ are identical for all voters. This is, of course, a great simplification. If some voters are vulnerable to the bandwagon effect, others may be susceptible to the underdog effect; still others may receive no influence at all. As we do not know who are more susceptible to which effect, we study each case separately by assuming that all voters are susceptible to the same effect.

Facing (t_L,t_R,x), voter w (weakly) prefers L to R if

$$-(t_L-t_R)w+\alpha(h(t_L\mu)-h(t_R\mu))+\theta(\phi(x)-\phi(1-x))\geq 0. \tag{2}$$

If $t_L>t_R$, the left-hand side expression of (2) is decreasing in w and goes to $-\infty$ as $w\to\infty$, but may be negative at $w=0$ when $\theta(x-\frac{1}{2})<0$. If $t_L<t_R$, it is increasing in w and goes to ∞ as $w\to\infty$, but may be positive at $w=0$ if $\theta(x-\frac{1}{2})>0$.

Thus, for $t_L\neq t_R$, we define the cutoff level of income for inequality (2) as:

$$w(t_L,t_R,x)\equiv\max[\varpi(t_L,t_R,x),0], \tag{3}$$

where $\varpi(t_L,t_R,x)=\alpha\frac{h(t_L\mu)-h(t_R\mu)}{t_L-t_R}+\theta\frac{\phi(x)-\phi(1-x)}{t_L-t_R}$.

By Assumption 1, the first term inside the max expression of Eq. 3 is always finite (for $t_L\neq t_R$). It is not always positive; it can be negative if $\theta(x-\frac{1}{2})(t_L-t_R)<0$. To prevent the first term from being always negative, we make the following assumption; without this assumption, there may exist some $x\in[0,1]$ at which one party is preferred by 'all' voters for all distinct pairs of $t_L\neq t_R$.

Assumption 2 For any $x\in[0,1]$, there exists at least one pair of distinct policies (t_L,t_R), $t_L\neq t_R$, such that $w(t_L,t_R,x)>0$.

[5] What is essential in our model is the uni-dimensionality of the policy space, not the uni-dimensionality of the voter characteristic space.

The set of voters who prefer L to R is the set of w for whom inequality (2) is satisfied. Thus, given (t_L, t_R, x), the set of voters who prefer L to R is

$$\Omega(t_L, t_R, x) = \begin{cases} \{w \in \mathbb{R}_+ | w \leq w(t_L, t_R, x)\} & \text{if } t_L > t_R \\ \{w \in \mathbb{R}_+ | w \geq w(t_L, t_R, x)\} & \text{if } t_L < t_R \\ \mathbb{R}_+ & \text{if } t_L = t_R \quad \text{and} \quad \theta(x - \frac{1}{2}) > 0 \\ \text{a random half subset of } \mathbb{R}_+ & \text{if } t_L = t_R \quad \text{and} \quad \theta(x - \frac{1}{2}) = 0 \\ \emptyset & \text{if } t_L = t_R \quad \text{and} \quad \theta(x - \frac{1}{2}) < 0, \end{cases} \tag{4}$$

and the actual membership share of party L is given by

$$x' = P(\Omega(t_L, t_R, x)), \tag{5}$$

where $P(\Omega(t_L, t_R, x)) = \begin{cases} F(w(t_L, t_R, x)) & \text{if } t_L > t_R \\ 1 - F(w(t_L, t_R, x)) & \text{if } t_L < t_R \\ 1 & \text{if } t_L = t_R \quad \text{and} \quad \theta(x - \frac{1}{2}) > 0 \\ \frac{1}{2} & \text{if } t_L = t_R \quad \text{and} \quad \theta(x - \frac{1}{2}) = 0 \\ 0 & \text{if } t_L = t_R \quad \text{and} \quad \theta(x - \frac{1}{2}) < 0. \end{cases}$

Note a difference between the case in which $t_L = t_R$ and $\theta(x - \frac{1}{2}) > 0$ and the case in which $t_L = t_R$ and $\theta(x - \frac{1}{2}) = 0$. In the former case, all voters strictly prefer L to R (due to voters' conformist preference), while in the latter case, voters are indifferent between them. We assume that indifferent voters decide their party membership by flipping a fair coin. This means that the two random half subsets of $\mathbb{R}_+$ will have exactly the same distributions of voters as $F(.)$.

So far, we described the basic data of the model. We now introduce the two factions that Nash-bargain one another in setting the party policy.

We define the payoff function of the opportunists in party L to be

$$\pi(t_L, t_R; x) = \Phi\left(P(\Omega(t_L, t_R, x))\right), \tag{6}$$

where $\Phi : [0, 1] \to [0, 1]$ is an increasing function such that $\Phi(\frac{1}{2}) = \frac{1}{2}$ and $\Phi(x) = 1 - \Phi(1 - x)$.

In like manner, the payoff function of party R's opportunists is defined by:

$$1 - \pi(t_L, t_R; x) = 1 - \Phi\left(P(\Omega(t_L, t_R, x))\right) = \Phi\left(1 - P(\Omega(t_L, t_R, x))\right). \tag{7}$$

Our formulation of the objective functions of the opportunists covers a number of different specifications in the literature on political economy.

First, it covers the models with electoral uncertainty, where $\pi(t_L, t_R, x)$ and $1 - \pi(t_L, t_R, x)$ are interpreted as probabilities of victory. Suppose the actual vote share for L is given by $P(\Omega(t_L, t_R, x)) + \varepsilon$, where ε is a random variable distributed by a symmetric distribution function $G(.)$ such that $G(0) = \frac{1}{2}$. Then party L's probability of victory is $\Pr(P(\Omega(t_L, t_R, x)) + \varepsilon > 0.5) = 1 - G\left(0.5 - P(\Omega(t_L, t_R, x))\right)$, and thus $\Phi(x) = G(x - 0.5)$.

Second, if $\Phi(x) = 1_{(\frac{1}{2},1]}(x) + \frac{1}{2}1_{\{\frac{1}{2}\}}(x)$, where $1_A(x)$ is an indicator function that takes 1 if $x \in A$ and 0 otherwise, then $\pi(t_L, t_R, x)$ can be considered a probability of victory in the models with electoral certainty.

Third, if $\Phi(x) = x$, then $\pi(t_L, t_R, x)$ and $1 - \pi(t_L, t_R, x)$ are actual membership shares. This might be the case if there is no electoral uncertainty and the election is not of the winner-takes-all type.[6]

Although our formulation is flexible enough to cover various specifications, we will call $\pi(t_L, t_R, x)$ and $1 - \pi(t_L, t_R, x)$ probabilities of victory throughout the article. Also we will maintain the following assumption.

Assumption 3 $\Phi : [0, 1] \to [0, 1]$ is strictly increasing.[7]

The actual membership share, $P(\Omega(t_L, t_R, x))$, has a number of *distinct* features that do not exist in the standard models of political competition. Figure 1 illustrates some possible shapes that $P(\Omega(., t_R, x))$ can take. (In what follows, we will discuss in terms of the membership share of party L; symmetric statements hold for the R membership share.)

The case of $\theta(x - \frac{1}{2}) = 0$ corresponds to the models without voter conformism. In this case, $P(\Omega(t, t_R, x))$ is *monotonic* both on $[0, t_R)$ and on $(t_R, 1]$, and discontinuous at $t = t_R$ (unless t_R is the ideal tax rate of the voter with the median income). It is monotonic in the following sense; as t increases up to t_R on $[0, t_R)$, $P(\Omega(t, t_R, x))$ increases up to $\lim_{t \to t_R^-}(1 - F(w(t, t_R, x)))$, and as t decreases down to t_R on $(t_R, 1]$, it increases up to $\lim_{t \to t_R^+} F(w(t, t_R, x))$. At $t = t_R$, $P(\Omega(t, t_R, x)) = \frac{1}{2}$, which is, in general, not equal to the two limits.

When $\theta(x - \frac{1}{2}) \neq 0$, on the other hand, $P(\Omega(t, t_R, x))$ is inherently *non-monotonic* and *continuous* everywhere in t. We make the statement precise in the following lemma.

Lemma 1 (1) *Suppose* $\theta(x - \frac{1}{2}) > 0$. *For any* $t_R \in [0, 1)$, $P(\Omega(t, t_R, x))$ *strictly decreases on* $(t_R, 1]$ *with* $\lim_{t \to t_R^+} P(\Omega(t, t_R, x)) = 1$. *Also for any* $t_R \in (0, 1]$, *there exists unique* $a(t_R, x)$ *on* $[0, t_R)$ *such that* (i) $P(\Omega(t, t_R, x)) = 1$ *for all* $t \in [a(t_R, x), t_R]$; *and* (ii) *whenever* $a(t_R, x) > 0$, $P(\Omega(t, t_R, x))$ *strictly increases on* $[0, a(t_R, x))$ *with* $\lim_{t \to a(t_R,x)^-} P(\Omega(t, t_R, x)) = 1$.

(2) *Suppose* $\theta(x - \frac{1}{2}) < 0$. *For any* $t_R \in (0, 1]$, *there exists* $b(t_R, x)$ *on* $[0, t_R)$ *such that* $P(\Omega(t, t_R, x))$ *strictly decreases on* $[b(t_R, x), t_R)$ *with* $\lim_{t \to t_R^-} P(\Omega(t, t_R, x)) = 0$. *Also for any* $t_R \in [0, 1)$, *there exist* $c(t_R, x)$ *and* $d(t_R, x)$ *on* $(t_R, 1]$ *such that* (i) $c(t_R, x) \leq d(t_R, x)$; (ii) $P(\Omega(t, t_R, x)) = 0$ *for all* $t \in [t_R, c(t_R, x)]$; *and* (iii)

[6] For instance, under proportional representation, the opportunists may care more about vote shares than probabilities of victory. Baron (1993) and Ortuno-Ortin (1997) study models of political competition under proportional representation, in which the influence of the groups favoring a certain policy is proportional to the percentage votes favoring that policy. Alternatively, $\pi(t_L, t_R, x)$ and $1 - \pi(t_L, t_R, x)$ in this specification may well be interpreted as probabilities of victory in the models with very large electoral uncertainty. Suppose the error term ε in a model with electoral uncertainty is uniformly distributed over $[-0.5, 0.5]$; thus uncertainty is very large. Then $\Phi(x) = G(x - 0.5) = x$ for $x \in [0, 1]$.

[7] Thus Assumption 3 rules out the models with electoral certainty where $\pi(t_L, t_R, x)$ is interpreted as a probability of victory (the second specification in the above discussion).

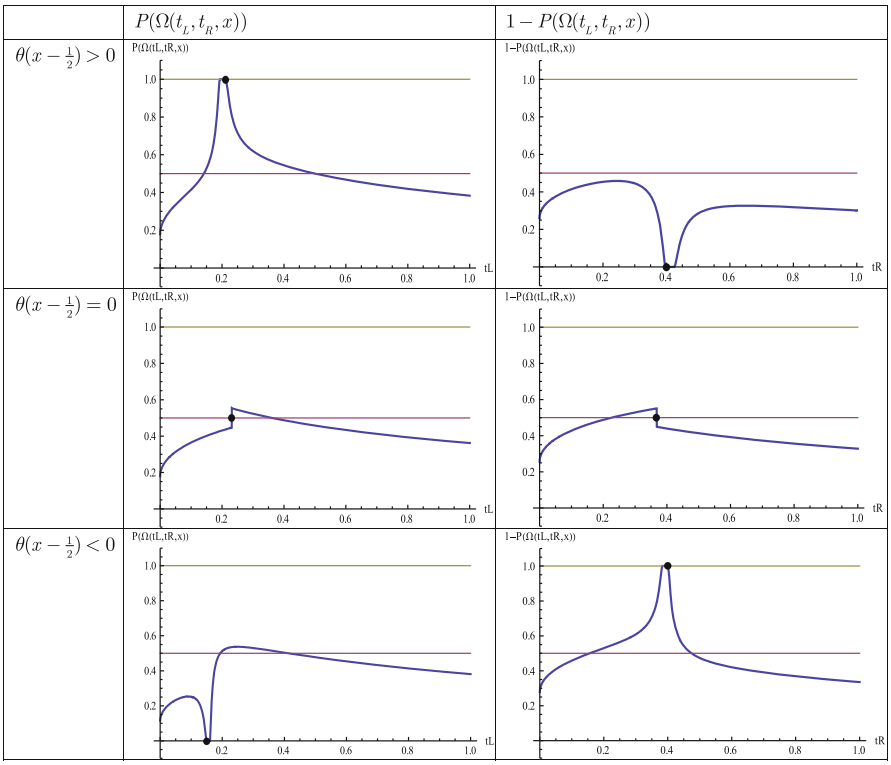

Fig. 1 Actual membership shares (*Note:* In each panel, the dot represents a policy position of a party's opponent *and* the membership share of the party should it choose the policy position of its opponent. These figures are drawn while holding constant the policy of the opposition party and the expected membership share of party L)

whenever $c(t_R, x) < d(t_R, x)$, $P(\Omega(t, t_R, x))$ *strictly increases on* $(c(t_R, x), d(t_R, x)]$ *with* $\lim_{t \to c(t_R,x)^+} P(\Omega(t, t_R, x)) = 0$.

We briefly remark on why $P(\Omega(t, t_R, x))$ can be *non-monotonic* and is *continuous* everywhere in t when $\theta(x - \frac{1}{2}) \neq 0$.

Mathematically, the possibility of non-monotonic $P(\Omega(t, t_R, x))$ is due to the presence of the second term in the definition of $\varpi(t, t_R, x)$ in Eq. 3, $\theta \frac{\phi(x) - \phi(1-x)}{t - t_R}$, which does not vanish unless $\theta(x - \frac{1}{2}) = 0$. For all $t \neq t_R$, the first term of $\varpi(t, t_R, x)$ is decreasing in t. The second term, on the other hand, is decreasing in t if $\theta(x - \frac{1}{2}) > 0$ but increasing in t if $\theta(x - \frac{1}{2}) < 0$. Therefore, in the case of $\theta(x - \frac{1}{2}) < 0$, two competing forces work in the determination of $P(\Omega(t, t_R, x))$.

The continuity of $P(\Omega(t, t_R, x))$, on the other hand, follows from two facts: first, although $\varpi(t, t_R, x)$ can be negative (its second term is negative if $\theta(x - \frac{1}{2})(t - t_R) < 0$), $w(t, t_R, x)$ is bounded below by 0; second, if $t_L = t_R$, voters vote only according to their conformist preference.

Intuitively, the non-monotonicity is due to the fact that voters' voting according to their conformist preferences may partially offset the effect of their voting according to their material interests. Consider the case in which θ is positive (i.e., bandwagon). As one party changes its policy to the direction of the policy position of the other party, it takes some of the voters who would vote for the opposition party otherwise. At the same time, however, voters are influenced more by the expected vote share difference than by the policy position difference as the policy difference of the two parties becomes smaller. If the party is currently winning in opinion polls, this second effect would be beneficial for the party. If the party is currently losing in opinion polls, on the other hand, the second effect is negative for the party, and would be devastating if the party chooses a policy that is too close to the opponent party's policy position.

Thus the presence of voter conformism has a dramatically different implication for the behavior of the opportunists. Part (1) of Lemma 1 states that the opportunists of the party with the expected membership share greater than $\frac{1}{2}$ can be better off by advocating a policy that is closer to its opponent's policy. But part (2) states that the opportunists in the party with the expected membership share less than $\frac{1}{2}$ become *worse off* if they advocate a policy that is too close to the other party's policy; by doing so, they *decrease* their party's actual membership share.

Thus the opportunists in the losing party can make themselves better off by *moving away*, within a certain limit, from the opposition party's policy. The higher the value of $\theta > 0$, the stronger the incentive of the opportunists in the losing party to move away from the policy of the opposition party. This implication is sharply in contrast with the one obtained in the models without voter conformism, where the opportunists of each party become better off as they move their policy to the direction of the other party's policy.

We now describe the objective function of the militants in each party. Consider an arbitrary partition of the polity into two sets of party members, H_L and H_R, such that $H_L \cup H_R = \mathbb{R}_+$ and $H_L \cap H_R = \emptyset$. Assume that a party's von Neumann–Morgenstern utility function is the average of its members' utility functions representing economic interests. Thus, for an arbitrary policy $t \in T$ and party memberships H_L and H_R, they are:

$$V(t; H_L) = \begin{cases} \frac{1}{P(H_L)} \int\limits_{w \in H_L} ((1-t)w + \alpha h(t\mu))dP(w) & \text{if } P(H_L) \neq 0 \\ 0 & \text{if } P(H_L) = 0 \end{cases}, \tag{8}$$

and

$$V(t; H_R) = \begin{cases} \frac{1}{P(H_R)} \int\limits_{w \in H_R} ((1-t)w + \alpha h(t\mu))dP(w) & \text{if } P(H_R) \neq 0 \\ 0 & \text{if } P(H_R) = 0 \end{cases} \tag{9}$$

In our model, these are the objective functions that the militants would like to maximize.[8]

[8] We are assuming that the militants care only about the economic well-being of their members, and not about the part due to voters' conformist preferences.

Because the utility function representing the economic interests is quasi-linear, each party's von Neumann–Morgenstern utility function, defined as the average well-being of its members, is identical to the utility function of the voter whose income equals the mean income of its members; for

$$\frac{1}{P(H_L)} \int\limits_{w \in H_L} ((1-t)w + \alpha h(t\mu))dP(w) = (1-t)w_L + \alpha h(t\mu), \tag{10}$$

and

$$\frac{1}{P(H_R)} \int\limits_{w \in H_R} ((1-t)w + \alpha h(t\mu))dP(w) = (1-t)w_R + \alpha h(t\mu), \tag{11}$$

where $w_L = \frac{1}{P(H_L)} \int\limits_{w \in H_L} w dP(w)$ and $w_R = \frac{1}{P(H_R)} \int\limits_{w \in H_R} w dP(w)$.

To model a within-party Nash-bargaining process between the factions, we need to specify the impasse payoffs (i.e., the payoffs of the factions should they fail to come to an agreement). If party L's factions fail to come to an agreement, party R wins the election by default; the probability of victory for party L is zero and party R's policy will be implemented. Thus, given (t_R, x, H_L), the Nash-bargaining solution between the two factions of party L is the policy t_L that maximizes a Nash product:

$$(\pi(t, t_R, x) - 0)^{\gamma_L}(V(t; H_L) - V(t_R; H_L))^{1-\gamma_L}, \tag{12}$$

for some $\gamma_L \in [0, 1]$. Similarly, given (t_L, x, H_R), party R's factions Nash-bargain to a policy t_R that maximizes:

$$(1 - \pi(t_L, t; x) - 0)^{\gamma_R}(V(t; H_R) - V(t_L; H_R))^{1-\gamma_R}, \tag{13}$$

for some $\gamma_R \in [0, 1]$.

We now define our equilibrium concept.

Definition 1 For given $\gamma_L, \gamma_R \in [0, 1]$, a generalized Wittman–Roemer political equilibrium with voter conformism is a partition of the polity into H_L^* and H_R^* and a triple (t_L^*, t_R^*, x^*) such that:

(1) $t_L^* \in \arg\max(\pi(t, t_R^*; x^*))^{\gamma_L}(V(t; H_L^*) - V(t_R^*; H_L^*))^{1-\gamma_L}$;
(2) $t_R^* \in \arg\max(1 - \pi(t_L^*, t; x^*))^{\gamma_R}(V(t; H_R^*) - V(t_L^*; H_R^*))^{1-\gamma_R}$;
(3) $w \in H_L^* \Rightarrow w \in \Omega(t_L^*, t_R^*, x^*)$,
$w \in H_R^* \Rightarrow w \in \mathbb{R}_+ \backslash \Omega(t_L^*, t_R^*, x^*)$;
(4) $x^* = P(H_L^*)$.

The first two conditions in Definition 1 require that given (x^*, H_L^*, H_R^*), (t_L^*, t_R^*) be a Nash equilibrium of the game in which each party's payoff function is a weighted Nash product of the payoff functions of its two factions. Thus a generalized Wittman–Roemer equilibrium is 'doubly Nash.' Each party plays a best response to the opponent

while holding (x^*, H_L^*, H_R^*) constant, and the best response is an outcome of a within-party Nash-bargaining process.

The third condition endogenizes party membership; it states that no member of either party is better represented by the other party at the equilibrium. Baron (1993) first uses the idea here ('voting with feet') in the context of political competition, although our formulation is closer to those of Ortuno-Ortin and Roemer (1998) and Roemer (2001, p. 92).

The fourth condition requires that the actual party membership shares be identical to the expected party membership shares at the equilibrium party platforms; polls are accurate in predicting party membership shares at the equilibrium party platforms. This condition is *weaker* than the requirement that the actual membership shares be identical to the expected membership shares for *all possible pairs of party platforms*.

It is difficult to characterize a generalized Wittman–Roemer equilibrium with voter conformism in its full generality. In the following sections, we will study the following three special cases.

First, if we set $\gamma_L = \gamma_R = 1$, we have the Hotelling–Downs model, adapted for voter conformism. In this model, the militants have no bargaining power in both parties.

Second, if we set $\gamma_L = \gamma_R = 0$, we have the model of political competition between two purely ideological parties in which the opportunists have no say in determining party policies. We call a political equilibrium in this case an ideological-party equilibrium with voter conformism.

Finally, if $\gamma_L = \gamma_R = \frac{1}{2}$, then we have the classical Wittman–Roemer model, adapted for endogenous party membership and voter conformism, where the two factions have equal bargaining power in both parties. [For details of the classical Wittman–Roemer model, see Roemer (2001, Chap. 3).]

Before proceeding to the formal analysis of these three models, we make a few remarks regarding Definition 1.

First, it would be useful to see how our equilibrium concept is different from those employed in the models without voter conformism. Take the classical Wittman–Roemer model. (A similar comparison can be made regarding the Hotelling–Downs model.) The original formulation of the classical Wittman–Roemer equilibrium requires only that (t_L^*, t_R^*) be mutual best responses of the two parties; party memberships and their shares are then automatically derived from (t_L^*, t_R^*). In contrast, for $(t_L^*, t_R^*, x^*, H_L^*, H_R^*)$ to be a Wittman–Roemer equilibrium *with* voter conformism, the following two conditions must be met simultaneously: given (x^*, H_L^*, H_R^*), (t_L^*, t_R^*) must be mutual best responses of the two Wittman–Roemer parties, and (t_L^*, t_R^*) must predict precisely (x^*, H_L^*, H_R^*). Put it mathematically, Definition 1 requires that (t_L^*, t_R^*, x^*) be a fixed point of

$$\beta_L(t_R; x, \Omega(t_L, t_R, x)) \times \beta_R(t_L; x, \Omega(t_L, t_R, x)) \times P(\Omega(t_L, t_R, x)), \qquad (14)$$

where β_i is the best response of party i derived while holding constant the membership of party i and the expected membership share of party L.

Second, condition (4) of Definition 1 shows another way of presenting the bandwagon and underdog effects. Condition (4) requires that given (t_L^*, t_R^*), the equilibrium

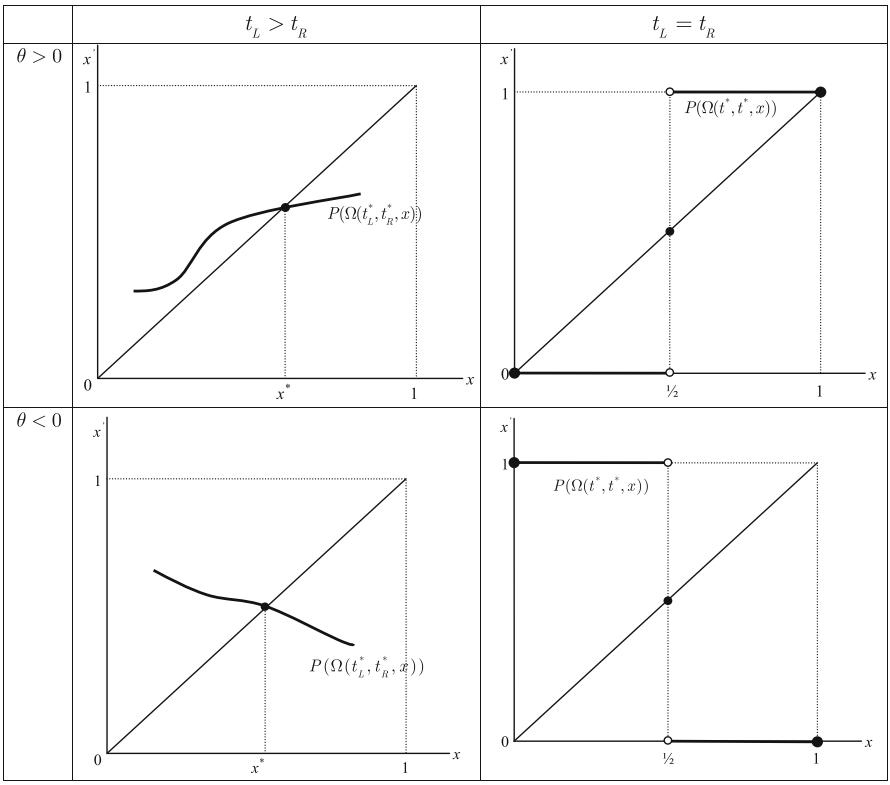

Fig. 2 Bandwagon and underdog (*Note:* In the case of $t_L < t_R$, we switch the top left and the bottom left figures of the case in which $t_L > t_R$)

membership share of party L be a fixed point of the map $P(\Omega(t_L^*, t_R^*, x))$. The presence of the bandwagon effect implies that the map is increasing, while the underdog effect corresponds to the case in which the map is decreasing.

Figure 2 illustrates. In Fig. 2, the horizontal axis measures the expected membership share of party L and the vertical axis measures its actual membership share. In Fig. 2, we draw four cases of the map $P(\Omega(t_L^*, t_R^*, x))$. In every case, the intersection of the map with the 45 degree line define fixed points of the map, given (t_L^*, t_R^*).

It is straightforward to prove that there always exists such a fixed point, given (t_L^*, t_R^*). If $t_L^* > t_R^*$, there is at least one fixed point when $\theta > 0$, and only one fixed point when $\theta \leq 0$. In particular, if $\theta > 0$, there may exist multiple fixed points; the exact number of fixed points is determined by the curvature of the map $P(\Omega(t_L^*, t_R^*, x))$. If $t_L^* = t_R^*$, on the other hand, there are three fixed points $(0, \frac{1}{2},$ and $1)$ when $\theta > 0$, and only one $(\frac{1}{2})$ when $\theta \leq 0$.[9]

[9] The fact that a membership share is a fixed point of the map at (t_L^*, t_R^*) does not guarantee that it is an equilibrium membership share; we repeat that for a membership share, as a fixed point, to be an 'equilibrium' membership share, the fixed point calculated at (t_L^*, t_R^*) must confirm the pair of policies as mutual best responses at the fixed point.

Third, at the generalized Wittman–Roemer equilibrium, if a move from t_i^* increases the payoff of party i's militants, then it must decrease the payoff of party i's opportunists. In other words, if a policy pair is a generalized Wittman–Roemer equilibrium, neither party's factions can unanimously agree to alter their proposal, given the policy played by the opposition party. (See Roemer 2001, pp. 159–163 for more details.)

Fourth, we briefly remark that there may exist 'trivial' non-differentiated equilibria in the generalized Wittman–Roemer model of political competition. If $F(.)$ is 'symmetric,' for instance, $(t^*, t^*, \frac{1}{2})$, where t^* is the ideal tax rate of the voter with the median (mean) income, is an equilibrium for all $\gamma_L, \gamma_R \in [0, 1]$.[10] Other trivial non-differentiated equilibria may also exist, depending on the functional forms of $F(.)$, $h(.)$, and $\phi(.)$. These equilibria are not of our interest; we are more interested in 'generic' equilibria.

Finally, although we presented the equilibrium as a static concept, it is possible to interpret it as a stationary point of the following dynamic process.

1. Suppose there is a sequence of decision making over time until party conventions, which are held simultaneously and, perhaps, some months prior to the election. Thus we are modeling a dynamic process of debate among citizens and politicians, which ultimately results in equilibrium party platforms and equilibrium party memberships.[11] Start with an arbitrary quintuple $(t_L^0, t_R^0, x^0, H_L^0, H_R^0)$ in the first period.
2. In each period after the first, each voter decides the party of which he/she will be a member in the current period, observing the two parties' previous policies and taking their past membership shares as the expected membership shares of the current period.
3. After observing the current party membership, each party chooses its current policy through a Nash-bargaining process between the factions, while assuming that the other party will choose the policy it chose in the previous period.
4. In the next period, voters revise their party membership according to rule 2, and parties revise the policies according to rule 3.
5. The process continues until the time of party conventions.

To see that a stationary point of this dynamic process is identical to a generalized Wittman–Roemer equilibrium defined in Definition 1, suppose (t_L, t_R) is a policy pair

[10] If both parties propose an identical policy at the expected membership share of $\frac{1}{2}$, the actual membership share is also $\frac{1}{2}$. At the membership share of $\frac{1}{2}$, no party would deviate unilaterally if the income distribution is symmetric because the policy is optimal for both the militants and the opportunists when the other party is holding that policy. Showing that it is optimal for the opportunists is immediate. To show that t^* is optimal for L's militants, we first note that $x^* = \frac{1}{2}$ implies that either (i) both H_L^* and H_R^* are random half subsets of $\mathbb{R}_+$, or (ii) $H_L^* = [0, \mu)$ and $H_R^* = [\mu, \infty)$. But the second case is impossible if both parties hold the same policy. Thus given that $x^* = 1/2$ and the other party is holding that policy, t^* is optimal for L's militants. (Were there no voter conformism, t^* would not have to be optimal for the militants. Here is another difference between the model *with* voter conformism and the one *without* it.)

[11] One might wish to model a dynamic process that continues until the election time, but such a modeling should assume that parties can costlessly change their announced policies, even after party conventions, up to the day of the election. We think it unrealistic; indeed changing its policy after the party convention would harm the party's credibility with the voters. Klingerman et al. (1994) show that parties adhere quite closely to the platforms they publish in the manifestoes written at party conventions.

of the previous period, (H_L, H_R) is a pair of the sets of party members in that period, and $x = P(H_L)$ is the membership share of party L in that period. If we follow the above dynamic process, the variables of interest in the current period (denoted with a prime) are updated from (t_L, t_R, x, H_L, H_R) to $(t'_L, t'_R, x', H'_L, H'_R)$ as follows:

$$H'_L = \Omega(t_L, t_R, x) = \Omega(t_L, t_R, P(H_L)), \quad H'_R = \mathbb{R}_+ \backslash H'_L, \tag{15}$$

$$x' = P(H'_L) = P(\Omega(t_L, t_R, x)), \tag{16}$$

$$t'_L \in \arg\max(\pi(t, t_R; x))^{\gamma_L}(V(t; H'_L) - V(t_R; H'_L))^{1-\gamma_L}, \tag{17}$$

$$t'_R \in \arg\max(1 - \pi(t_L, t; x))^{\gamma_R}(V(t; H'_R) - V(t_L; H'_R))^{1-\gamma_R}. \tag{18}$$

A stationary point of this dynamic process is clearly an equilibrium of Definition 1.

Note that a partition of the polity into the two parties is entirely determined by $w(t_L, t_R, x)$ defined in Eq. 3; i.e., $\overline{w}' = w(t_L, t_R, x)$. Thus the dynamic process justifying the Wittman–Roemer politics is reduced to the following system of three first-order difference equations:

$$t'_L \in \beta_L(t_R; x, w(t_L, t_R, x)), \tag{19}$$

$$t'_R \in \beta_R(t_L; x, w(t_L, t_R, x)), \tag{20}$$

$$x' = P(w(t_L, t_R, x)). \tag{21}$$

Note that the dynamic here is a 'best-response' dynamic. Best responses are well-defined for all $\gamma_L, \gamma_R \in [0, 1)$, but they may not be well-defined if $\gamma_L = 1$ and $\theta(x - \frac{1}{2}) = 0$.[12] Thus it is only the Hotelling–Downs parties that may not have a best response to some of its opponent's policies.

3 The Hotelling–Downs model of political competition when voter conformism is present

We first study the political equilibrium of the Hotelling–Downs model when voter conformism is present. (This is the case of $\gamma_L = \gamma_R = 1$ in our formulation.)

In the original Hotelling–Downs model without voter conformism, a pair of Condorcet winners constitutes a political equilibrium. In the model with voter conformism, which policy will be a Condorcet winner depends, in general, upon the expected membership share x. We define a strict x-Condorcet winner to be a policy that defeats all other policies in pairwise elections at the expected membership share of x. We now prove:

Theorem 1 *For any $F(.)$, the unique Hotelling–Downs equilibrium with voter conformism is $(t^*_L, t^*_R, x^*) = (t^*, t^*, \frac{1}{2})$, where t^* is a strict $\frac{1}{2}$-Condorcet winner. The equilibrium party membership is an arbitrary random half subset of R_+ for each party.*

[12] As we discussed earlier in this section, $\pi(t, t_R, x)$ is discontinuous at $t = t_R$ if $\theta(x - \frac{1}{2}) = 0$, and continuous everywhere if $\theta(x - \frac{1}{2}) \neq 0$. (See Fig. 1 again.) If $\gamma_L \neq 1$, $(\pi(t, t_R, x))^{\gamma_L}(V(t; H_L) - V(t_R; H_L))^{1-\gamma_L}$ is always continuous in t, even in the case of $\theta(x - \frac{1}{2}) = 0$. This is because as $t \to t_R$, $V(t; H_L) - V(t_R; H_L) \to 0$ while $\pi(t, t_R, x)$ approaches to a finite number.

Thus were real politics a Hotelling–Downs kind, there would be no differentiation of policies between political parties, and voter conformism would have no consequence on party platforms or policies; parties would propose the same policy whatever the types of voter conformism. Theorem 1 might be considered a generalization of the Condorcet winner theorem to political models with voter conformism.

One might wonder why there are no other equilibria than this. Figure 2 shows that there may exist other fixed points of the map $P(\Omega(t_L^*, t_R^*, x))$. For example, if $t_L^* = t_R^*$, the membership share of 0 or 1 is also a fixed point. Why don't these membership shares constitute an equilibrium? Lemma 1 provides an intuition for why there is no Hotelling–Downs equilibrium with $x^* \neq \frac{1}{2}$. If $\theta > 0$ and $x^* \neq \frac{1}{2}$, the opportunists in the 'winning' party would like to choose a policy that is as close as possible to the policy of the other party while the opportunists in the losing party would like to move away from the policy of the opponent party. A policy pair $t_L = t_R$ is not an equilibrium because a losing party will deviate from it. A policy pair $t_L \neq t_R$ is not an equilibrium because a winning party has an incentive to come closer to the policy of a losing party. (A similar intuition can be provided for the case in which $\theta < 0$.)

According to Theorem 1, the Hotelling–Downs model of political competition seems quite robust; the model's prediction *with* voter conformism is identical to its prediction *without* voter conformism. But we do not think the Hotelling–Downs equilibrium attractive as a description of real politics, in particular when voter conformism is present.

First, in the Hotelling–Downs world of politics, political parties and their candidates would have no concern about opinion polls and possible effects of voter conformism that opinion polls might generate, although their sole motivation is winning the election. We think it idiosyncratic.

Second, the Hotelling–Downs equilibrium with voter conformism is a knife-edge equilibrium; justifying the equilibrium from a dynamic perspective seems difficult.

Looking at best responses in the Hotelling–Downs model would be helpful. Referring to Fig. 1, one can easily verify the following:

(1) If $\theta(x - \frac{1}{2}) = 0$ and $t_j = t^*$, party i's best response to t_j is to choose the policy equal to $t_j = t^*$. If $\theta(x - \frac{1}{2}) = 0$ and $t_j \neq t^*$, on the other hand, party i has no best response. (As party i increases its tax rate on $[0, t_j)$, its probability of victory increases; it becomes $\frac{1}{2}$ at t_j, and decreases afterward.)

(2) If $\theta(x - \frac{1}{2}) > 0$, party i has a continuum of best responses, which always includes t_j. (If $\theta(x - \frac{1}{2}) > 0$, party i can have 'all' voters as its party members by choosing the same policy as the opponent's.) In this case, party i may use a random device to choose one among many.

(3) If $\theta(x - \frac{1}{2}) < 0$, party i has a best response that is not equal to t_j. (Choosing t_j is the worst option for party i in this case, because the entire polity will turn away from the party.)

Thus unless the initial expected fraction of voters who prefer L to R is precisely equal to $\frac{1}{2}$, and the initial pair of policies is exactly that of strict $\frac{1}{2}$-Condorcet winners, the dynamic process may stuck in the middle (due to the absence of a best response), or simply drift away from the equilibrium (due to the multiplicity of best responses).

We conjecture, without a proof, that the probability that the dynamic process the Hotelling–Downs politics converges to the equilibrium is zero.[13]

4 The ideological-party equilibrium with voter conformism

We now study the case in which both parties are purely ideological, consisting only of the militants. This is another extreme case in our formulation.

Political parties in this model are not strategic; each party simply chooses the ideal policy of its average member. Without endogenous party membership or voter conformism, this model would be trivial. With voter conformism and endogenous party membership, however, the model is no longer trivial. Although each party puts forth the ideal policy of its average member, the membership is endogenously determined and voter conformism affects the membership; this in turn changes the policy of the two parties. Our study in this section will also shed some lights on the classical Wittman–Roemer model of political competition that we will study in the next section.[14]

Let us denote the ideal tax rate of voter w by $\tilde{t}(w) \equiv \arg\max(1-t)w + \alpha h(t\mu)$. As T is compact and convex, and $(1-t)w + \alpha h(t\mu)$ is continuous in (t, w) and strictly quasi-concave in t, $\tilde{t}(w)$ is a well-defined continuous function of w. Also $(1-t)w + \alpha h(t\mu)$ has decreasing differences on $T \times \mathbb{R}_+$; thus $\tilde{t}(w)$ is non-increasing in w.[15]

In this model, Eqs. 19–21 can be merged into a single equation. We used the following equilibrium condition for the cutoff level that separates the party memberships:

$$\overline{w} = \max\left[\alpha \frac{h(\tilde{t}(w_L(\overline{w}))\mu) - h(\tilde{t}(w_R(\overline{w}))\mu)}{\tilde{t}(w_L(\overline{w})) - \tilde{t}(w_R(\overline{w}))} + \theta \frac{\phi(F(\overline{w})) - \phi(1-F(\overline{w}))}{\tilde{t}(w_L(\overline{w})) - \tilde{t}(w_R(\overline{w}))}, 0\right], \tag{22}$$

where $w_j(\overline{w})$ is the average income of the constituency of party $j = L, R$ at the cutoff level $\overline{w}$. Once the equilibrium cutoff level is determined by Eq. 22, the equilibrium tax rates and the equilibrium membership share follow immediately: $t_j^* = \tilde{t}(w_j(\overline{w}^*))$, $j = L, R$; and $x^* = F(\overline{w}^*)$.

We now prove the following:

Theorem 2 *Suppose* $\frac{\theta}{\alpha} < \frac{h(\tilde{t}(0)\mu) - h(\tilde{t}(\mu)\mu)}{\phi(1) - \phi(0)}$, *where* $\tilde{t}(w)$ *is the ideal tax rate of* w.

(1) *There exists a differentiated ideological-party equilibrium with voter conformism such that* $t_L^* > t_R^*$ *and* $x^* \in (0, 1)$.
(2) *Suppose* $h(.)$, $\phi(.)$ *and* $F(.)$ *are differentiable. At any asymptotically stable differentiated ideological-party equilibrium, the following holds locally:*

[13] One might suggest we consider *better-responses* as optimal responses when a best response does not exist. This does not solve the problem of multiplicity of optimal responses.

[14] Gomberg et al. (2004) study a similar model (without voter conformism) in a multi-dimensional policy space. Their model employs Aumann's Strong Nash Equilibrium as an equilibrium concept.

[15] A function $v(t, w)$ has decreasing differences on $T \times \mathbb{R}_+$ if $t^1 > t^2$ and $w^1 > w^2$ imply $v(t^1, w^1) - v(t^2, w^1) \leq v(t^1, w^2) - v(t^2, w^2)$.

$$x^* > \frac{1}{2} \Rightarrow \frac{\partial t_j^*}{\partial \theta} \leq 0 \textit{ and } x^* < \frac{1}{2} \Rightarrow \frac{\partial t_j^*}{\partial \theta} \geq 0 \textit{ for } j = L, R \textit{ and } \theta \neq 0.$$

Some remarks are in order regarding Theorem 2.

First, part (1) of Theorem 2 states that in contrast with the Hotelling-Down model, policies are generically differentiated at the equilibrium in this model. This is because parties propose the ideal policy of their average members and party memberships are sharply separated.

Second, due to the presence of voter conformism, multiple equilibria may exist, in particular, when the bandwagon effect is sufficiently strong. To get the intuition for this result, we first note that at the equilibrium such that $t_L^* > t_R^*$, Eq. 22 is equivalent to $x^* = F(w(t_L^*, t_R^*, x^*))$, for it implies

$$F^{-1}(x^*) = \max\left[\alpha \frac{h(t_L^*\mu) - h(t_R^*\mu)}{t_L^* - t_R^*} + \theta \frac{\phi(x^*) - \phi(1 - x^*)}{t_L^* - t_R^*}, 0\right], \tag{23}$$

In Fig. 1, we observed that multiple fixed points of x^* may exist if the bandwagon effect is sufficiently strong; a high value of $\theta > 0$ would make the map very curvy. As x^* and $\overline{w}^*$ are positively correlated and the relationship between them is one-to-one, multiplicity of x^* implies multiplicity of $\overline{w}^*$. (On the other hand, if $\theta < 0$, the map is downward sloping; there is only one fixed point.)

Of course, not all multiple equilibria are expected to be stable in the dynamic context. Part (2) of Theorem 2 establishes the comparative static result for dynamically stable equilibria. For any stable, differentiated, ideological-party equilibrium where party L's membership share is $>\frac{1}{2}$, an increasing bandwagon effect decreases the ideal tax rates of the two parties and an increasing underdog effect increases the ideal tax rates of the two parties; the opposite holds for any stable differentiated ideological-party equilibrium where party R's membership share is $>\frac{1}{2}$.

An intuition for this result is very simple. Take $\theta > 0$ and consider an equilibrium with $x^* > \frac{1}{2}$. As θ increases, some voters at the margin will switch from party R to party L. This conversion makes the average members of both parties richer than before, which in turn decreases the ideal tax rates of both parties. An intuition for the case in which $x^* < \frac{1}{2}$ is similar.

Figure 3 illustrates the ideological-party equilibrium with voter conformism; it shows the equilibrium cutoff level of income determined by Eq. 22.

For this numerical example, we chose for F a lognormal distribution derived from a normal distribution with mean m and standard deviation s. We estimate the parameter values of the lognormal distribution using the 2004 US Census Bureau data. Our estimated parameters are $m = 1.408$ and $s = 0.886$.[16] Finally, we chose $h(t\mu) = \sqrt{t\mu}$ and $\phi(x) = x^k$, where $k \geq 1$.

Figure 3 shows that there emerge multiple equilibria when the bandwagon effect is sufficiently strong. The bottom right panel of Fig. 3 shows that we have three equilibria. Among them, the middle one is unstable while the other two are stable.

[16] According to the US Census Bureau 2004 Economic Survey, the mean household income in the United States was \$60,528, and the Gini coefficient for household incomes in that year was 0.469.

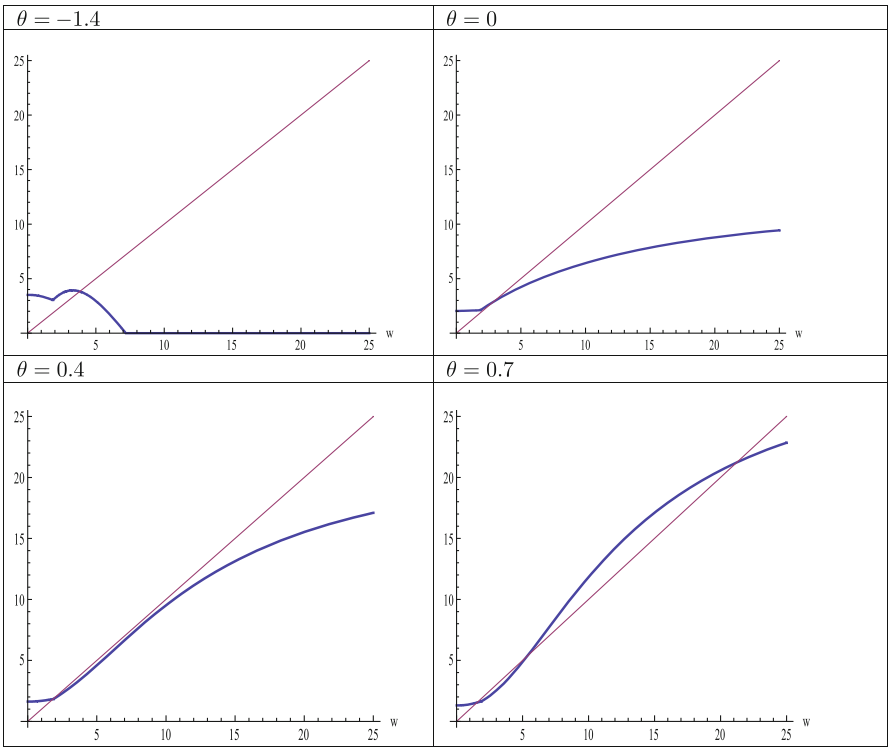

Fig. 3 The cutoff level of income that separates party membership in the ideological-party equilibrium (*Note:* Parameter values are: $m = 1.40804$; $s = 0.8860$; $k = 1.5$; $\alpha = 1.0$. The thick curve represents the function,$\max\left[\alpha\frac{h(\tilde{t}(w_L(\overline{w}))\mu)-h(\tilde{t}(w_R(\overline{w}))\mu)}{\tilde{t}(w_L(\overline{w}))-\tilde{t}(w_R(\overline{w}))}+\theta\frac{\phi(F(\overline{w}))-\phi(1-F(\overline{w}))}{\tilde{t}(w_L(\overline{w}))-\tilde{t}(w_R(\overline{w}))},0\right]$, and the thin line is the 45 degree line)

5 Voter conformism and the classical Wittman–Roemer model of political competition

So far we studied two extreme cases of a generalized Wittman–Roemer model of political competition with voter conformism. We now study the effect of voter conformism on the classical Wittman–Roemer political equilibrium, in which two factions have equal bargaining power in both parties. It appears that the classical Wittman–Roemer model shares at least two features of the ideological-party model of political competition. First, parties propose differentiated policies at the equilibrium. Second, the phenomenon of multiple equilibria seems typical when the bandwagon effect is sufficiently strong.

A general characterization of the classical Wittman–Roemer equilibrium with voter conformism is somewhat difficult to obtain. Thus we will calculate them numerically. In the numerical computation, we chose the same functions used in Sect. 4. For $\Phi(.)$, we set $\Phi(x) = G(x - 0.5)$, where $G(.)$ is a normal distribution with mean 0 and standard deviation 0.05.

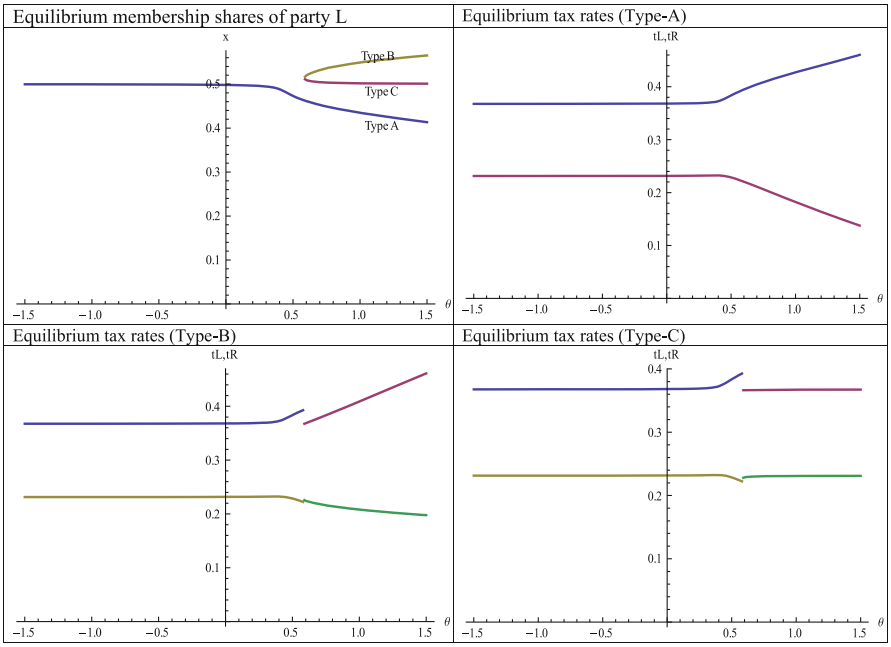

Fig. 4 Classical Wittman–Roemer equilibria with voter conformism (*Note:* Parameter values are: $m = 1.40804$; $s = 0.8860$; $a = 0$; $b = 0.05$; $k = 1.5$; $\alpha = 1.8$)

We varied the value of θ from -1.5 to 1.5, and found the classical Wittman–Roemer equilibrium with voter conformism for all values of θ in this range. As in the original Wittman–Roemer model without voter conformism, parties propose differentiated policies at the equilibrium. Figure 4 and Table 1 illustrate. (To save space, we do not report the equilibrium cutoff level of income: $\overline{w}^*$.)

We first note that the underdog effect has almost no significance on the classical Wittman–Roemer equilibrium, although it has a minor effect of mitigating the policy differentiation. We varied the value of θ from -0.01 to -1.5, but the tax rates and the vote shares are almost constant. (When we examine the equilibria carefully, we note that the difference of party platforms, $t_L^* - t_R^*$, keeps decreasing from 0.136541 at $\theta = -0.01$ to 0.136352 at $\theta = -1.5$.)

The bandwagon effect has, on the other hand, significantly different implications on the classical Wittman–Roemer equilibrium.

First, as in the ideological-party model, we observe multiple Wittman–Roemer equilibria when the bandwagon effect is sufficiently strong. In our numerical calculation, the branching point is at $\theta = 0.586$. If θ is less than it, there exists one equilibrium. At $\theta = 0.586$, there emerge two equilibria. After that, one of the two equilibria branches into two separate equilibria. Thus, if $\theta > 0.586$, there always exist three equilibria. When multiple equilibria exist, we call them type-A, type-B, and type-C equilibria. A type-A equilibrium is the one in which x^* is significantly <0.5, a type-B equilibrium is that in which x^* is significantly >0.5, and a type-C equilibrium is the one in which x^*

Table 1 Classical Wittman–Roemer equilibria with voter conformism

θ	x	t_L	t_R	$t_L - t_R$	x	t_L	t_R	$t_L - t_R$	x	t_L	t_R	$t_L - t_R$
−1.50	0.500	0.36770	0.23135	0.13635								
−1.00	0.499	0.36775	0.23138	0.13637								
−0.50	0.499	0.36785	0.23145	0.13640								
0	0.498	0.36817	0.23163	0.13655								
0.50	0.473	0.38358	0.22866	0.15493								
	Type A				Type B				Type C			
0.60	0.461	0.39487	0.22052	0.17435	0.520	0.36861	0.22451	0.14410	0.509	0.36631	0.22872	0.13759
0.70	0.453	0.40434	0.21114	0.19320	0.532	0.37820	0.21825	0.15995	0.504	0.36664	0.23022	0.13642
0.80	0.446	0.41261	0.20145	0.21116	0.539	0.38824	0.21415	0.17409	0.503	0.36688	0.23059	0.13629
0.90	0.440	0.42011	0.19176	0.22835	0.544	0.39853	0.21087	0.18766	0.502	0.36702	0.23076	0.13626
1.00	0.435	0.42712	0.18221	0.24491	0.549	0.40894	0.20809	0.20085	0.502	0.36711	0.23086	0.13625
1.10	0.430	0.43382	0.17284	0.26098	0.553	0.41938	0.20564	0.21374	0.502	0.36718	0.23093	0.13625
1.20	0.426	0.44037	0.16369	0.27668	0.556	0.42983	0.20343	0.22641	0.501	0.36723	0.23098	0.13625
1.30	0.421	0.44688	0.15476	0.29213	0.559	0.44027	0.20139	0.23888	0.501	0.36727	0.23101	0.13626
1.40	0.417	0.45350	0.14603	0.30747	0.562	0.45069	0.19948	0.25121	0.501	0.36730	0.23104	0.13626
1.50	0.413	0.46035	0.13750	0.32285	0.565	0.46110	0.19767	0.26343	0.501	0.36732	0.23106	0.13626

Note: Parameter values are: $m = 1.40804$; $s = 0.8860$; $a = 0$; $b = 0.05$; $k = 1.5$; $\alpha = 1.8$. In the numerical computation, we increased the value of θ from -1.5 to 1.5 by the size of 0.01, with the total of 301 separate calculations. Due to the space constraint, we report only some of them

is between them, which is nearby 0.5. We plot the three types of equilibria separately in Fig. 4.

It is not easy to merge Eqs. 19–21 into one in this model. We thus checked the stability of Wittman–Roemer equilibria by calculating the Jacobian of the system of Eqs. 19–21. If eigenvalues of the Jacobian are all less than 1 in their absolute terms, it would be asymptotically stable in a dynamic context. We find that all type-A and type-B equilibria are asymptotically stable while all type-C equilibria are not. Thus more meaningful equilibria in a dynamic context are those of type-A and type-B.

In type-A and type-B equilibria, political parties diverge more as the bandwagon effect becomes larger.[17] We thus observe that in dynamically stable classical Wittman–Roemer equilibria, an increasing bandwagon effect *exacerbates* the policy differentiation of the two parties.

Therefore, the classical Wittman–Roemer equilibrium is sharply in contrast with the ideological-party equilibrium, where parties move in the *same* direction as the degree of voter conformism changes; the classical Wittman–Roemer parties move in an *opposite* direction as the parameter that captures the degree of voter conformism increases.

To understand the intuition behind the result in this section, it would be useful to see how factions in each party bargain to an equilibrium policy, taking the equilibrium policy of the opposition party as given. Recall that at a generalized Wittman–Roemer equilibrium, neither party's factions can unanimously agree to alter their proposal. Any policy that increases both factions' payoffs cannot be a Wittman–Roemer equilibrium; given the policy played by the opposition party, if a move from t_i^* increases the payoff of party i's militants, then it must decrease the payoff of party i's opportunists. Thus the bargaining outcome lies always in the Pareto mini frontier defined by the payoff functions of the two factions calculated at the opposition party's policy.

We now provide an intuition while considering the bandwagon effect only; the underdog effect can be explained in similar ways.

Suppose party R is winning ($x^* < \frac{1}{2}$), as in a type-A equilibrium. Pick any level of $\theta > 0$. At the current level of θ, a proposal of R-militants to change the current R policy to the direction of the ideal tax rate of its average member would not be agreed upon within the party because it would require a sacrifice of R-opportunists. If the value of θ increases, however, it gives a windfall gain to R-opportunists, because more voters will lean toward R even at the same policy pairs. Thus, without sacrificing its opportunists, party R can change its policy slightly to the direction of the ideal tax rate of its average member.

The intuition for the direction of a policy change in the losing party, in this case party L, is different from that for the winning party. According to Lemma 1 of Sect. 2, the opportunists of a 'losing party' can be better off by moving away from the policy of the opposition party. The higher the value of θ, the stronger the incentive of the opportunists of a losing party to move away from the policy of the opponent. Such a

[17] In type-C equilibria, on the other hand, the difference in policies between the two parties is almost constant. When we examine the type-C equilibria carefully, we note that the difference of party platforms, $t_L^* - t_R^*$, keeps decreasing from 0.1383 at $\theta = 0.59$ to 0.13625 at $\theta = 0.98$, and afterward increases up to 0.13626 at $\theta = 1.5$. But such changes are almost unnoticeable.

move will be agreed upon by the militants of party L, because it implies that the party policy will be closer to the ideal policy of the militants.

An explanation for the case in which party L is winning, as in the type-B equilibrium, is similar. An increase in θ in this situation gives a windfall gain to L-opportunists; thus L-militants can call for a higher tax rate without scarifying L-opportunists. At the same time, party R's opportunists propose a policy which is further distant from the policy of party L.

The above explanations based on the Nash-bargaining perspective also provide an intuition for why neither bandwagon nor underdog has much impact on the policy differences when the vote share is nearby 0.5. If the vote share is close to 0.5, windfall gains to the opportunists of the winning party are very small; there is not much room for bargaining for policy changes. Also the change that the opportunists of a losing party demand will be small.

6 Conclusion

Using the framework of the generalized Wittman–Roemer model of political competition, we studied the potential effect that voter conformism might have on the political equilibrium of various models. This article shows that the effect of voter conformism on the nature of political equilibrium is quite different depending on the model one uses.

We find that voter conformism, both bandwagon and underdog, has no effect on the Hotelling–Downs political equilibrium. Even if voter conformism is present, the Hotelling–Downs parties propose an identical policy at the equilibrium, which is equal to a strict $\frac{1}{2}$ -Condorcet winner. But such an equilibrium seems difficult to justify in a dynamic context.

In the ideological-party model, political parties propose differentiated policies at the equilibrium and the presence of multiple equilibria is generic when the bandwagon effect is sufficiently strong. In those equilibria that are dynamically stable and have the membership share of party L greater than 0.5 (less than 0.5), an increasing bandwagon effect decreases (increases) the equilibrium tax rates of both parties; the opposite is true for the underdog effect.

The Wittman–Roemer parties behave differently not only from the Hotelling–Downs parties but also from the purely ideological ones. Unlike the Hotelling–Downs parties but like the purely ideological parties, the Wittman–Roemer parties propose differentiated equilibrium policies. Existence of multiple equilibria for a sufficiently strong bandwagon effect also seems typical. In contrast with the purely ideological parties which move in the same direction as the degree of voter conformism changes, the Wittman–Roemer parties move in an opposite direction as the parameter that captures the degree of voter conformism increases. In those equilibria that are dynamically stable, the stronger the bandwagon effect is, the more differentiated policies are. The opposite holds when the underdog effect is present.

This article studies the effect of voter conformism on political equilibrium in a *uni-dimensional* policy space. It is well known that both the Hotelling–Downs and Wittman–Roemer models of political competition do not possess generic equilibria

when the policy space is multi-dimensional. There are models that possess generic equilibria in a multi-dimensional policy space, such as the probabilistic voting model of Lindbeck and Weibull (1987), or the party unanimity Nash equilibrium model of Roemer (2001). We leave the study of the effect of voter conformism on these models for future research.

Acknowledgments I thank John Roemer for his comments on the first version of this article, which improved the article significantly. I also would like to thank three referees for many useful comments and suggestions.

Appendix

Proof of Lemma 1 We first note that $\varpi(t, t_R, x)$ consists of two terms.The first term, $\alpha\frac{h(t\mu)-h(t_R\mu)}{t-t_R}$, is finite- and positive-valued, and strictly decreases on $[0, t_R) \cup (t_R, 1]$.

(1) Suppose $\theta(x - \frac{1}{2}) > 0$.

Then for any $t_R \in [0, 1)$, $\theta\frac{\phi(x)-\phi(1-x)}{t-t_R}$ is strictly decreasing and positive-valued on $(t_R, 1]$ with $\lim_{t\to t_R^+}\theta\frac{\phi(x)-\phi(1-x)}{t-t_R} = \infty$, which implies that $\varpi(t, t_R, x)$ is positive-valued and strictly decreases on $(t_R, 1]$ with $\lim_{t\to t_R^+}\varpi(t, t_R, x) = \infty$. Because $w(t, t_R, x) = \max[\varpi(t, t_R, x), 0]$ and $P(\Omega(t, t_R, x)) = F(w(t, t_R, x))$ on $(t_R, 1]$, the first statement is proved.

On the other hand, for any $t_R \in (0, 1]$, $\theta\frac{\phi(x)-\phi(1-x)}{t-t_R}$ is strictly decreasing and negative-valued on $[0, t_R)$ with $\lim_{t\to t_R^-}\theta\frac{\phi(x)-\phi(1-x)}{t-t_R} = -\infty$, which implies that $\varpi(t, t_R, x)$ strictly decreases on $[0, t_R)$ with $\lim_{t\to t_R^-}\varpi(t, t_R, x) = -\infty$. Thus there exists unique $a(t_R, x)$ on $[0, t_R)$ such that (i) $\max[\varpi(t, t_R, x), 0] = 0$ for all $t \in [a(t_R, x), t_R)$; and (ii) whenever $a(t_R, x) > 0$, $\varpi(t, t_R, x)$ is positive and strictly decreases on $[0, a(t_R, x))$. Because $P(\Omega(t, t_R, x)) = 1 - F(w(t, t_R, x))$ on $[0, t_R)$ and $P(\Omega(t_R, t_R, x)) = 1$, the second statement is proved.

(2) Suppose $\theta(x - \frac{1}{2}) < 0$.

Then for any $t_R \in (0, 1]$, $\theta\frac{\phi(x)-\phi(1-x)}{t-t_R}$ is strictly increasing and positive-valued on $[0, t_R)$ with $\lim_{t\to t_R^-}\theta\frac{\phi(x)-\phi(1-x)}{t-t_R} = \infty$, which implies that $\varpi(t, t_R, x)$ is positive-valued and either strictly increasing or U-shaped on $[0, t_R)$ with $\lim_{t\to t_R^-}\varpi(t, t_R, x) = \infty$. Thus $P(\Omega(t, t_R, x))$ is either strictly decreasing or inverse U-shaped on $[0, t_R)$ with $\lim_{t\to t_R^-}P(\Omega(t, t_R, x)) = 0$. Therefore, there exists $b(t_R, x)$ on $[0, t_R)$ such that $P(\Omega(t, t_R, x))$ strictly decreases on $[b(t_R, x), t_R)$ with $\lim_{t\to t_R^-}P(\Omega(t, t_R, x)) = 0$. The first statement is proved.

For any $t_R \in [0, 1)$, $\theta\frac{\phi(x)-\phi(1-x)}{t-t_R}$ is strictly increasing and negative-valued on $(t_R, 1]$ with $\lim_{t\to t_R^+}\theta\frac{\phi(x)-\phi(1-x)}{t-t_R} = -\infty$, which implies that $\varpi(t, t_R, x)$ is either strictly increasing or inverse U-shaped on $(t_R, 1]$ with $\lim_{t\to t_R^+}\varpi(t, t_R, x) = -\infty$. Thus there exist $c(t_R, x)$ and $d(t_R, x)$ on $(t_R, 1]$ such that (i) $c(t_R, x) \le d(t_R, x)$; (ii) $\max[\varpi(t, t_R, x), 0] = 0$ for all $t \in (t_R, c(t_R, x)]$; and (iii) whenever $c(t_R, x) <$

$d(t_R, x)$, $\varpi(t, t_R, x)$ is positive and strictly increases on $(c(t_R, x), d(t_R, x)]$. Noting that $P(\Omega(t, t_R, x)) = F(w(t, t_R, x))$ on $(t_R, 1]$ and $P(t_R, t_R, x) = 0$, we complete the proof of the second statement. □

Proof of Theorem 1 We first prove that $(t_L^*, t_R^*, x^*) = (t^*, t^*, \frac{1}{2})$ is an equilibrium. At $x^* = \frac{1}{2}$, Eq. 1 becomes

$$\begin{cases} (1 - t_L)w + \alpha h(t_L \mu) + \theta\phi(\frac{1}{2}) \text{ from candidate L} \\ (1 - t_R)w + \alpha h(t_R \mu) + \theta\phi(\frac{1}{2}) \text{ from candidate R} \end{cases} \tag{24}$$

Because the same constant, $\theta\phi(\frac{1}{2})$, appears in both lines, voters decide only by looking at the policy positions of the two parties. Therefore, the standard theorems on the Hotelling–Downs model apply. (See pp. 21, 53 of Roemer 2001, for instance.) We thus have $t_L^*\big|_{x^*=\frac{1}{2}} = t_R^*\big|_{x^*=\frac{1}{2}} = t^*$, which in turn yields the vote share of $\frac{1}{2} = x^*$. Because voters are all indifferent at $(t_L^*, t_R^*, x^*) = (t^*, t^*, \frac{1}{2})$, the equilibrium party membership is an arbitrary random half subset of R_+ for each party.

It remains to prove that there are no other equilibria than this.

First, the above argument also proves that the case in which $t_L = t_R \neq t^*$ and $\theta(x - \frac{1}{2}) = 0$ and the case where $t_L \neq t_R$ and $\theta(x - \frac{1}{2}) = 0$ do not constitute an equilibrium.

Second, we prove that the case in which $t_L = t_R$ and $\theta(x - \frac{1}{2}) \neq 0$ does not constitute an equilibrium. Consider first the case of underdog: $\theta < 0$. In this case, if $t_L = t_R = t$, there will be no equilibrium membership share of party L other than $\frac{1}{2}$ (see Fig. 2), which completes the proof. Next we consider the case of bandwagon: $\theta > 0$. If $t_L = t_R = t$, the only candidates for the equilibrium membership share of party L not equal to $\frac{1}{2}$ are either 0 or 1. If $t_L = t_R = t$ and the expected membership share of party L is $x = 1$, the actual membership share for party R is $1 - P(\Omega(t, t, 1)) = 0$; party R has an incentive to deviate. (By Assumption 2, there is a profitable direction of deviation.) If $t_L = t_R = t$ and the expected membership share of party L is $x = 0$, its actual membership share is also 0; party L has an incentive to deviate.

Finally, we prove that any case in which $t_L \neq t_R$ and $\theta(x - \frac{1}{2}) \neq 0$ cannot be an equilibrium. Suppose the actual membership share for party L at the triple is $P(\Omega(t_L, t_R, x))$. There are three cases.

Case 1: Suppose $P(\Omega(t_L, t_R, x)) \in (0, 1)$ at the triple. This means that party R's membership share at the triple is $1 - P(\Omega(t_L, t_R, x)) \in (0, 1)$. If $\theta(x - \frac{1}{2}) < 0$, party R can increase its actual membership share to 1 by choosing the policy that party L chooses. If $\theta(x - \frac{1}{2}) > 0$, party L has an incentive to choose the same policy as party R's.

Case 2: Suppose $P(\Omega(t_L, t_R, x)) = 0$ at the triple. For all values of θ and x, party L has an incentive to deviate to a policy that gives it a positive membership share. (This is guaranteed by Assumption 2.)

Case 3: Suppose $P(\Omega(t_L, t_R, x)) = 1$ at the triple. For all values of θ and x, party R has an incentive to deviate to a policy that gives a positive membership share. □

Proof of Theorem 2 (1) Step 1: For any arbitrary $\overline{w} > 0$ such that $F(\overline{w}) \in (0, 1)$, define two functions, $w_L(.)$ and $w_R(.)$, as follows:

$w_L(\overline{w}) = \frac{1}{F(\overline{w})} \int_0^{\overline{w}} w dF$ and $w_R(\overline{w}) = \frac{1}{1-F(\overline{w})} \int_{\overline{w}}^{\infty} w dF$.

Note that $w_L(.)$ and $w_R(.)$ are increasing continuous functions of $\overline{w}$. Also note that $w_L(\overline{w}) < w_R(\overline{w})$ for all $\overline{w}$. (Recall that F is strictly increasing.)

Step 2: We now show that there is a positive-valued fixed point of the map:

$$\varpi(\overline{w}) \equiv \alpha \frac{h(\tilde{t}(w_L(\overline{w}))\mu) - h(\tilde{t}(w_R(\overline{w}))\mu)}{\tilde{t}(w_L(\overline{w})) - \tilde{t}(w_R(\overline{w}))} + \theta \frac{\phi(F(\overline{w})) - \phi(1 - F(\overline{w}))}{\tilde{t}(w_L(\overline{w})) - \tilde{t}(w_R(\overline{w}))}. \tag{25}$$

The map $\varpi(.)$is continuous: for $\tilde{t}(w_L(\overline{w})) > \tilde{t}(w_R(\overline{w}))$, $h(.)$ is strictly concave on T, and $\tilde{t}(w_j(.))$ is continuous.

As $\overline{w} \to 0, \tilde{t}(w_L(\overline{w})) \to \tilde{t}(0)$ and $\tilde{t}(w_R(\overline{w})) \to \tilde{t}(\mu)$. Likewise, As $\overline{w} \to \infty, \tilde{t}(w_L(\overline{w})) \to \tilde{t}(\infty)$ and $\tilde{t}(w_R(\overline{w})) \to \tilde{t}(\infty)$. Thus if

$$\varpi(0) = \alpha \frac{h(\tilde{t}(0)\mu) - h(\tilde{t}(\mu)\mu)}{\tilde{t}(0) - \tilde{t}(\mu)} + \theta \frac{\phi(0) - \phi(1)}{\tilde{t}(0) - \tilde{t}(\mu)} > 0, \tag{26}$$

and

$$\varpi(\infty) = \alpha \frac{h(\tilde{t}(\mu)\mu) - h(\tilde{t}(\infty)\mu)}{\tilde{t}(\mu) - \tilde{t}(\infty)} + \theta \frac{\phi(1) - \phi(0)}{\tilde{t}(\mu) - \tilde{t}(\infty)} < \infty, \tag{27}$$

the continuous map $\varpi(.)$ will have a positive-valued fixed point. The condition of $\varpi(\infty) < \infty$ holds, because $\phi(.)$ is finite-valued and $h(.)$ is strictly concave. The condition that $\varpi(0) > 0$ is ensured under the stated assumption.

Step 3: Denote the positive-valued fixed point by $\overline{w}^*$ and define: $H_L^* = [0, \overline{w}^*]$; $H_R^* = (\overline{w}^*, \infty)$; $t_j^* = \tilde{t}(w_j(\overline{w}^*))$, $j = L, R$; and $x^* = F(\overline{w}^*)$. Then they clearly constitute an ideological-party equilibrium with $t_L^* > t_R^*$ and $x^* \in (0, 1)$.

(2) The full dynamic system in this model consists of four equations: $t_j = \tilde{t}(w_j(\overline{w}))$, $j = L, R$; $x = F(\overline{w})$; and $\overline{w}' = \alpha \frac{h(t_L^\mu) - h(t_R^\mu)}{t_L - t_R} + \theta \frac{\phi(x) - \phi(1-x)}{t_L - t_R}$. As it can be reduced to a single difference equation, $\overline{w}' = \varpi(\overline{w})$, and the other three equations are not difference equations, the condition for the stability of the dynamic system is: $\left| \frac{\partial \varpi(\overline{w}^*)}{\partial \overline{w}} \right| < 1$.

Fix θ, and let $\overline{w}^*(\theta)$ be the cutoff level of income evaluated at a dynamically stable ideological-party equilibrium. Then we must have: $\overline{w}^*(\theta) = \varpi(\overline{w}^*(\theta))$, where $\varpi(\overline{w})$ is given by Eq. 25. Differentiating both sides, we obtain:

$$\frac{\partial \overline{w}^*}{\partial \theta} = \frac{\partial \varpi(\overline{w}^*)}{\partial \overline{w}} \frac{\partial \overline{w}^*}{\partial \theta} + \frac{\phi(F(\overline{w}^*)) - \phi(1 - F(\overline{w}^*))}{\tilde{t}(w_L(\overline{w}^*)) - \tilde{t}(w_R(\overline{w}^*))}. \tag{28}$$

Rearranging terms while using $x^* = F(\overline{w}^*)$ and $t_j^* = \tilde{t}(w_j(\overline{w}^*))$, we finally have:

$$\frac{\partial \overline{w}^*}{\partial \theta} = \frac{\phi(x^*) - \phi(1 - x^*)}{\left(1 - \frac{\partial \varpi(\overline{w}^*)}{\partial \overline{w}}\right)\left(t_L^* - t_R^*\right)}. \tag{29}$$

The denominator is positive if the equilibrium is stable. The numerator is positive if $x^* > \frac{1}{2}$ and negative if $x^* < \frac{1}{2}$. Thus $x^* - \frac{1}{2} > 0 \Rightarrow \frac{\partial \overline{w}^*}{\partial \theta} > 0$ and $x^* - \frac{1}{2} < 0 \Rightarrow \frac{\partial \overline{w}^*}{\partial \theta} < 0$. The proof is complete by noting that t_j^* is a non-increasing function of $\overline{w}^*$. □

References

Aldrich J (1980) A dynamic model of presidential nomination campaigns. Am Political Sci Rev 74(3):651–669

Banerjee A (1992) A simple model of herd behavior. Q J Econ 107(3):797–817

Baron D (1993) Government formation and endogenous parties. Am Political Sci Rev 87(1):34–47

Bartels L (1985) Expectations and preferences in presidential nominating campaigns. Am Political Sci Rev 79(3):804–815

Bartels L (1988) Presidential primaries and the dynamics of public choice. Princeton University Press, Princeton, NJ

Baumol W (1957) Interactions between successive polling results and voting intentions. Public Opin Q 21(2):318–323

Bernheim B (1994) A theory of conformity. J Political Econ 102(5):841–877

Birkhchandani B, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. J Political Econ 100(5):992–1026

Callander S (2007) Bandwagon and momentum in sequential voting. Rev Econ Stud 74:653–684

Downs A (1957) An economic theory of democracy. Harper, New York, NY

Gomberg A, Marhuenda F, Ortuno-Ortin I (2004) A model of endogenous political party platforms. Econ Theory 24:373–394

Irwin G, Van Holsteyn JJM (2000) Bandwagons, underdogs, the Titanic, and the Red Cross: the influence of public opinion polls on voters. In: Communication presented at the 18th Congress of the International Political Science Association, Québec, 1–5 August, 2000

Klingerman HD, Hofferbert R, Budge I (1994) Parties, policies, and democracy. Westview Press, Boulder, CO

Leibenstein H (1950) Bandwagon, snob, and Veblen effects in the theory of consumer demand. Q J Econ 64(2):183–207

Lindbeck A, Weibull J (1987) Balanced budget redistribution as the outcome of political competition. Public Choice 52(3):273–297

Myerson R, Weber R (1993) A theory of voting equilibria. Am Political Sci Rev 87(1):102–114

Ortuno-Ortin I (1997) A spatial model of political competition and proportional representation. Soc Choice Welf 14(3):427–438

Ortuno-Ortin I, Roemer JE (1998) Endogenous party formation and the effect of income distribution on policy. University of Alicante, Mimeo

Roemer JE (1997) Political economic equilibrium when parties represent constituents: The uni-dimensional case. Soc Choice Welf 14(4):479–520

Roemer JE (2001) Political competition: theory and applications. Harvard University Press, Cambridge, MA

Schelling T (1978) Micromotives and macrobehavior. Norton, New York, NY

Simon H (1954) Bandwagon and underdog effects and the possibility of election prediction. Public Opin Q 18(3):245–253

Soc Choice Welf
DOI 10.1007/s00355-010-0505-2

ORIGINAL PAPER

Competition for popular support: a valence model of elections in Turkey

Norman Schofield · Maria Gallego · Ugur Ozdemir · Alexei Zakharov

Abstract Models of elections tend to predict that parties will maximize votes by converging to an electoral center. There is no empirical support for this prediction. In order to account for the phenomenon of political divergence, this paper offers a stochastic electoral model where party leaders or candidates are differentiated by differing valences—the electoral perception of the quality of the party leader. If valence is simply intrinsic, then it can be shown that there is a "convergence coefficient", defined in terms of the empirical parameters, that must be bounded above by the dimension of the space, in order for the electoral mean to be a Nash equilibrium. This model is applied to elections in Turkey in 1999 and 2002. The idea of valence is then extended to include the possibility that activist groups contribute resources to their favored parties in response to policy concessions from the parties. The equilibrium result is that parties, in order to maximize vote share, must balance a centripetal electoral force against a centrifugal activist effect. We estimate pure spatial models and models with sociodemographic valences, and use simulations to compare the equilibrium predictions with the estimated party positions.

N. Schofield (✉) · U. Ozdemir
Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA
e-mail: schofield.norman@gmail.com

U. Ozdemir
e-mail: uozdemir@artsci.wustl.edu

M. Gallego
Department of Economics, Wilfrid Laurier University, Waterloo, Canada
e-mail: mgallego@wlu.ca

A. Zakharov
State University-Higher School of Economics, Moscow, Russia
e-mail: 3ax@mail.ru

Published online: 06 November 2010

1 Introduction: modeling popular support

The early work in modeling elections focused on two-party competition, and assumed a one-dimensional policy space, X, and "deterministic" voter choice. The models showed the existence of a "core" point, unbeaten under majority rule vote, at the median of the electoral distribution. Such models implied that there would be strong centripetal political forces causing parties to converge to the electoral center (Hotelling 1929; Downs 1957). In higher dimensions, such two-party "pure strategy Nash equilibria" (PNE) generally do not exist, so the theory did not cover empirical situations where two or more policy dimensions were relevant.[1] It has been shown, however, that there would exist *mixed* strategy Nash equilibria whose support lies within a subset of the policy space known as the "uncovered set."[2] "Attractors" of the political process, such as the "core", the "uncovered set" or the "heart" (Schofield 1999) are centrally located with respect to the distribution of voters' ideal points. The theoretical prediction that political candidates converge to the center is very much at odds with empirical evidence from U.S. Presidential elections that political candidates do not locate themselves close to the electoral center.[3]

The deterministic electoral model is also ill-suited to deal with the multiparty case. (Here multiparty refers to the situation where the number of candidates or parties, p, is at least three.) As a result, recent work has focused on "stochastic" models which are, in principle, compatible with empirical models of voter choice.[4] In such models, the behavior of each voter is modeled by a vector of choice probabilities. Various theoretical results for this class of models suggested that vote maximizing parties would converge to the mean of the electoral distribution of voter ideal points.[5]

Empirical estimates of party positions in European multiparty polities can be constructed on the basis of various techniques of content analysis of party manifestos.[6] More recent analyses have been based on factor analysis of electoral survey data to obtain a multidimensional description of the main political issues in various countries. All these empirical analyses have obtained policy spaces that are two dimensional. These techniques allow for the estimation of the positions of the parties in the empirically inferred policy space. These estimates have found no general tendency for parties to converge to the center.[7]

The various empirical electoral models can be combined with simulation techniques to determine how parties should respond to electoral incentives to maximize their vote shares. Schofield and Sened (2006), in their simulation of elections in Israel in the period 1988–1996, found that vote maximizing parties did not converge to the electoral origin. It may be objected that factor analysis of survey data gives

[1] See Saari (1997) and the survey in Austen-Smith and Banks (1999).

[2] Banks et al. (2006).

[3] Poole and Rosenthal (1984), Schofield et al. (2003). See also the empirical work in Schofield et al. (2011).

[4] Schofield and Sened (2006).

[5] Hinich (1977), Lin et al. (1999), Banks and Duggan (2005), McKelvey and Patty (2006).

[6] See Laver and Hunt (1992). Benoit and Laver (2006) use expert estimates.

[7] See Adams and Merrill (1999), for example.

only a crude estimate of the variation in voter preferences, while vote maximization disregards the complex incentives that parties face. Nonetheless, as a modeling exercise, the stochastic model for Israel seemed to provide a plausible account of the nature of individual choice[8] as well as the party positioning decision. Although the simulated equilibrium positions of the parties in Israel were not identical to the estimated positions, the positions were generally far from the origin, and for some of the parties very close to their estimated positions. The purpose of this paper is to attempt to extend the stochastic empirical model so as to close the apparent disparity between the simulated equilibrium positions of the parties, and the estimated positions.

The key to the contradiction between the non-convergence result of Schofield and Sened, and the convergence result in other work on the formal stochastic model was the incorporation of an asymmetry in the perception of the quality of the party leaders, expressed in terms of *valence* (Stokes 1963, 1992).

In the model presented here, the average weight, in the voter calculus, given to the perceived quality of the leader of the jth party is called the party's *intrinsic or exogenous valence.* In empirical models this valence is assumed to be exogenous, so it is independent of the party's position. The valence coefficients for each party are generated by the estimation of the stochastic model, based on the "multinomial logit" (MNL) assumption that the stochastic errors have a "Type I extreme value or Gumbel distribution" (Dow and Endersby 2004). These valence terms add to the statistical significance of the model. In general, valence reflects the overall degree to which the party is perceived to have shown itself able to govern effectively in the past, or is likely to be able to govern well in the future.[9]

The Appendix considers a pure spatial stochastic vote model,with party specific exogenous valences, based on the same distribution assumption, and on the assumption that each party leader attempts to locally maximize the party's vote share. Results from Schofield (2007a,b) give the necessary and sufficient conditions under which there is a "local pure strategy Nash equilibrium" (LNE) of this model at the *joint electoral mean* (that is, where each party adopts the same position, z_0, at the mean of the electoral distribution).[10] Theorem 2 in the Appendix shows that a "convergence coefficient", c, incorporating all the parameters of the model, can be defined. This coefficient, c, involves the differences in the valences of the party leaders, and the "spatial coefficient" β. When the policy space, X, is assumed to be of dimension w, then the necessary condition for existence of an LNE at the electoral center[11] is that the coefficient, c, is bounded above by w.

[8] Over 60% of the individual votes were correctly modeled.

[9] See Penn (2009). Notice, however, that valence refers to the *perception* by voters of the quality of political leaders. Recent work by Westen (2007), for the United States, suggests that voters' perceptions of the characteristics of political candidates are very important. Moreover, Schofield et al. (2011) shows that voter perception of character traits has a strong effect on candidate positions in the United States.

[10] A local Nash equilibrium under vote maximization is just a vector of positions such that no small unilateral move by a party can increase its vote. The usual notion of a pure stratgy Nash equilibrium (PNE) cannot be used because in the games we study there may exist no PNE.

[11] Again, the electoral center, or origin, is defined to be the mean of the distribution of voter ideal points.

When the necessary condition fails, then parties, in equilibrium, will adopt divergent positions. Because a pure strategy Nash equilibrium must be an LNE, the failure of existence of LNE when all parties are at the electoral mean implies non existence of such a centrist PNE. In this case, a party whose leader has the lowest valence will have the greatest electoral incentive to move away from the electoral mean. As the party moves away from the electoral mean, it increases the probability that voters on the electoral periphery will vote for it.[12] Other low valence parties will follow suit, and the local equilibrium will be one where parties are distributed along a "principal electoral axis."[13] The general conclusion is that, with all other parameters fixed, then a convergent LNE can be guaranteed only when β is "sufficiently" small. Thus, divergence away from the electoral mean becomes more likely the greater are β, the valence differences and the variance of the electoral distribution.[14]

The innovation of this paper is that in additional to exogenous valence, we also incorporate "sociodemographic valence." These party specific valence terms are associated with different groups in the society, and are defined by dichotomous or continuous characteristics of different subgroups in the population. This model is shown to be statistically superior to the spatial model with exogenous valence. This is the case because the exogenous valence model assumes that all voters have the same perception of the quality of the party leaders, whereas with the sociodemographic variables, these perceptions are allowed to vary across different subgroups.

We apply this valence model by considering in some detail a sequence of elections in Turkey from 1999 to 2007. The election results are given in Tables 1, 2, and 3, which also provide the acronyms for the various parties.

As in other related work, the empirical models were based on factor analyses of voter surveys.[15] Figures 1 and 2 show the electoral distributions (based on sample surveys of sizes 635 and 483, respectively) and estimates of party positions for 1999 and 2002.[16] The two dimensions in both years were a "left–right" religion axis and a "north–south" Nationalism axis, with secularism or "Kemalism" on the left and Turkish nationalism to the north. (See also Carkoğlu and Hinich (2006) for a spatial model of the 1999 election.)

Minor differences between these two figures include the disappearance of the Virtue Party (FP) which was banned by the Constitutional Court in 2001, and the change

[12] This follows for theoretical reasons as shown in Schofield (2007a). When $c > w$, at least one of the eigenvalues of the Hessian of the vote share function of a low valence party will be large and positive at the origin. As it moves from the origin, it will lose votes from centrist voters, but gain votes from more radical voters. Simulation of empirical models for Israel (Schofield and Sened 2006) has shown this to be the case.

[13] The principal electoral axis is defined to be the one dimensional subspace along which the variance of the distribution of voter ideal points is maximum.

[14] These results are presented for the reader's convenience in the context of the more general model described in the Appendix.

[15] The estimations presented below are based on factor analyses of sample surveys conducted by Veri Arastima for TUSES.

[16] The party positions were estimated using expert analyses, in the same way as the work by Benoit and Laver (2006).

Table 1 Turkish election results 1999

Party name		% Vote	Seats	% Seats
Democratic Left Party	DSP	22.19	136	25
Nationalist Action Party	MHP	17.98	129	23
Virtue Party	FP	15.41	111	20
Motherland Party	ANAP	13.22	86	16
True Path Party	DYP	12.01	85	15
Republican People's Party	CHP	8.71	–	–
People's Democracy Party	HADEP	4.75	–	–
Others	–	4.86	–	–
Independents	–	0.87	3	1
Total			550	

Table 2 Turkish election results 2002

Party name		% Vote	Seats	% Seats
Justice and Development Party	AKP	34.28	363	66
Republican People's Party	CHP	19.39	178	32
True Path Party	DYP	9.54	–	–
Nationalist Action Party	MHP	8.36	–	–
Young Party	GP	7.25	–	–
People's Democracy Party	HADEP	6.22	–	–
Motherland Party	ANAP	5.13	–	–
Felicity Party	SP	2.49	–	–
Democratic Left Party	DSP	1.22	–	–
Others and Independents	–	6.12	9	2
Total			550	

Table 3 Turkish election results 2007

Party name		% Vote	Seats	% Seats
Justice and Development Party	AKP	46.6	340	61.8
Republican People's Party	CHP	20.9	112	20.3
Nationalist Movement Party	MHP	14.3	71	12.9
Democrat Party	DP	5.4	–	–
Young Party	GP	3.0	–	–
Felicity Party	SP	2.3	–	–
Independents	–	5.2	27[a]	4.9
Others	–	2.3	–	–
Total		100	550	100

[a] Twenty-four of these "independents" were in fact members of the DTP—the Kurdish Freedom and Solidarity Party

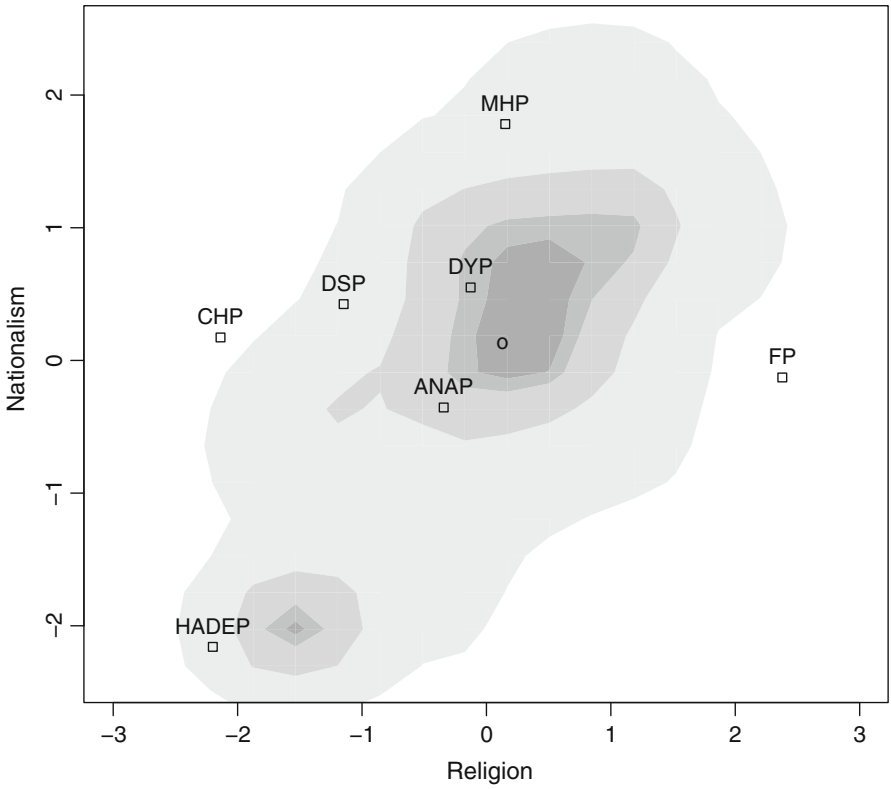

Fig. 1 Party positions and voter distribution in Turkey in 1999

of the name of the pro-Kurdish party from HADEP to DEHAP.[17] The most important change is the appearance of the new Justice and Development Party (AKP) in 2002, essentially substituting for the outlawed Virtue Party.

In 1999, a DSP minority government formed, supported by ANAP and DYP. This only lasted about 4 months, and was replaced by a DSP–ANAP–MHP coalition, indicating the difficulty of negotiating a coalition compromise across the disparate policy positions of the coalition members. During the period 1999–2002, Turkey experienced two severe economic crises. As Tables 1 and 2 show, the vote shares of the parties in the governing coalition went from about 53% in 1999 to less than 15% in 2002. In 2002, a 10% cut-off rule was instituted. As Table 2 makes clear, seven parties obtained less than 10% of the vote in 2002, and won no seats. The AKP won 34% of the vote, but because of the cut-off rule, it obtained a majority of the seats (363 out of 550). In 2007, the AKP did even better, taking about 46% of the vote, against 21% for the CHP. The Kurdish Freedom and Solidarity Party avoided the 10% cut-off rule, by contesting

[17] For simplicity, the pro-Kurdish party is denoted HADEP in the various Figures and Tables. Notice that the HADEP position in Figs 1 and 2 is interpreted as secular and non-nationalistic.

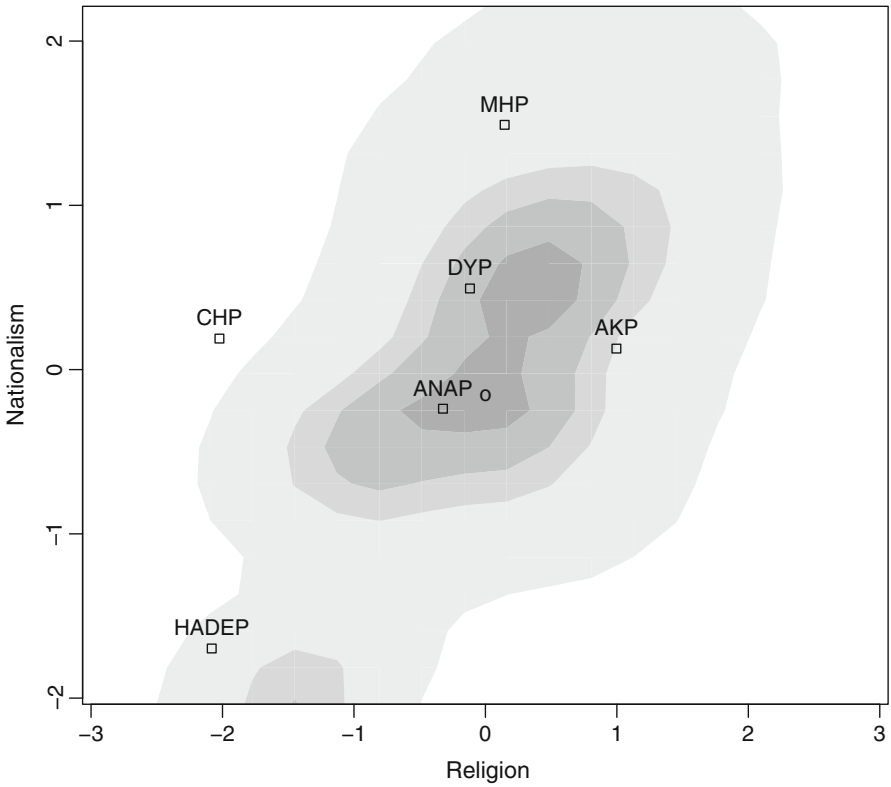

Fig. 2 Party positions and voter distribution in Turkey in 2002

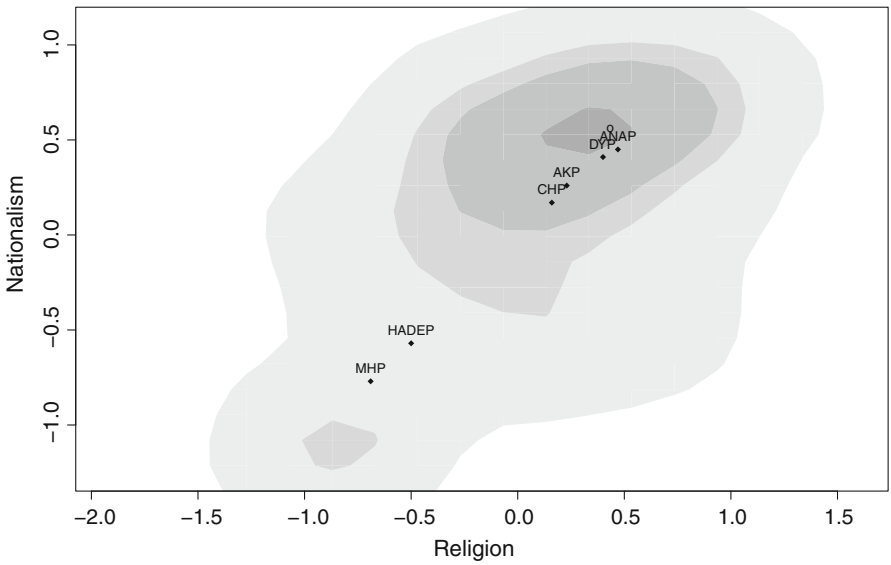

Fig. 3 A Local Nash Equilibrium for the pure spatial model in 2002

the elections as independent non-party candidates, winning 24 seats with less that 5% of the vote.

The point of this example is that a comparison of Figs 1 and 2 suggest that there was very little change in policy positions of the parties between 1999 and 2002. The basis of support for the AKP may be regarded as a similar to that of the banned FP, which suggests that the leader of this party changed the party's policy position on the religion axis, adopting a much less radical position.

In sum, the standard spatial model is unable to explain the change in the electoral outcome, taken together with the relative unchanged positioning of the parties between 1999 and 2002.

Section 2 of this paper considers the details of the MNL model for Turkey for 1999 and 2002. In particular, this section shows that the pure spatial model with exogenous valence predicts that the parties diverge away from the origin. To illustrate, Table 5 shows that the lowest valence party in 2002 was the Motherland Party (ANAP) while the Republican People's Party (CHP) had the highest valence. The convergence coefficient was computed to be 5.94, far greater than the upper bound of 2. Figure 3 presents an estimate of one of the LNE obtained from simulation of vote maximizing behavior of the parties, under the assumption of the pure spatial model with exogenous valence. As expected from the theoretical result, the LNE is non-centrist. Note, however, the the LNE positions for the pure spatial model given in Fig. 3 are quite different from the estimated positions in Fig. 2.

To improve the prediction of the model, we incorporated the sociodemographic variables. Estimating the LNE for this sociodemographic model gave a better prediction. To explain the difference between the estimated positions of the parties, and the LNE from the sociodemographic model, we then added the the influence of party activists to the model. Since sociodemographic variables can be interpreted as specific valences associated with different subgroups of the electorate, we can use these sociodemographic valences to estimate the influence of group-specific activists on party positions.

Theorem 1 in the Appendix[18] gives the first-order *balance* condition for local equilibrium in the stochastic electoral model involving sociodemographic valences and activists. The condition requires the balancing of a *centrifugal marginal activist pull (or gradient)* against a *marginal electoral pull*. In general, if the exogenous valence of a party leader falls, then the marginal electoral pull also falls, so balance requires that the leader adopt a position closer to the preferred position of the party activists.

The pure spatial model, with exogenous valences, and a joint model, with sociodemographic valences, but without activists, are compared using simulation to determine the LNE in these models. This allows us to determine which model better explains the party positions. For example, Fig. 5 shows the LNE based on a joint sociodemographic model for 2002. In this figure, the LNE position for the Kurdish party, HADEP, is a consequence of the high electoral pull by Kurdish voters located in the lower left of the figure. Similarly, the position of the CHP on the left of the figure is estimated to be due to the electoral pull by Alevi voters who are Shia, rather than Sunni and can be

[18] The results in the Appendix extend the version of the activist model originally proposed by Aldrich (1983) and developed in Schofield (2006).

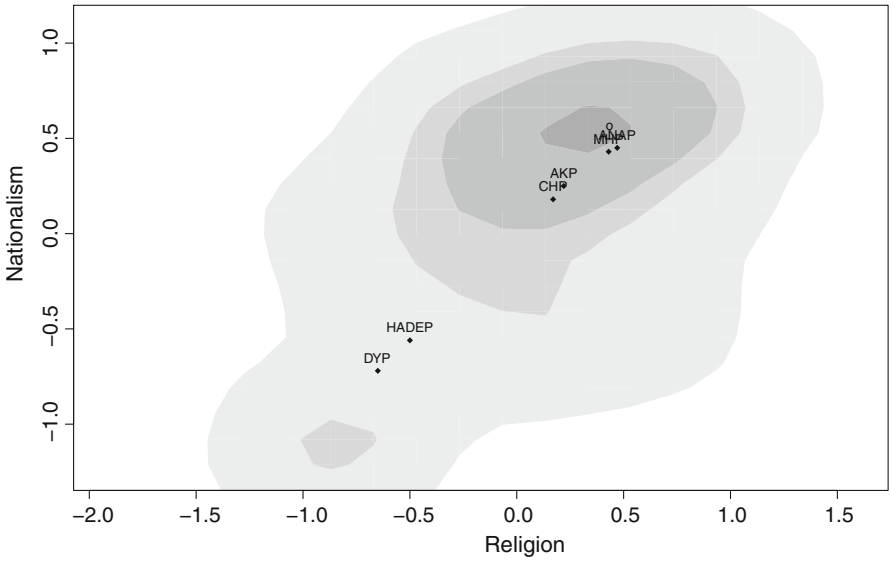

Fig. 4 A second local Nash Equilibrium for the pure spatial model in 2002

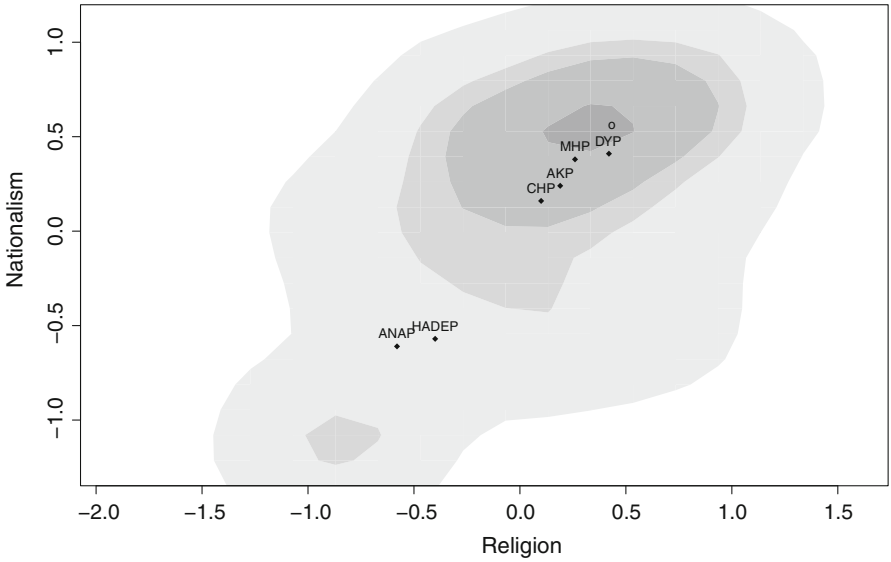

Fig. 5 A local Nash Equilibrium for the joint model in 2002

regarder as supporters of the secular state. Although Fig. 5 gives a superior prediction of the party positions than Fig. 3, there is still a discrepancy between the estimated positions of Fig. 2 and the LNE in Fig. 5. We argue that the difference between these two vectors of party positions, as presented in Figs. 2 and 5, can be used to provide an estimation of the marginal activist pulls influencing the parties.

More generally, we suggest that the combined model, with sociodemographic variables and activists, can be used as a tool with which to study the political configuration of such a complex society. In the conclusion we suggest that the full model involving activists may be applicable to the study of what Epstein et al. (2006) call "partial democracies", where a political leader must maintain popular support, not just by winning elections, but by maintaining the allegiance of powerful activist groups in the society.

2 Elections in Turkey 1999–2007

The Appendix defines an empirical electoral model, denoted $\mathbb{M}(\mathbf{\Lambda}, \boldsymbol{\theta}, \beta; \mathbf{V})$ which utilizes sociodemographic variables, denoted $\boldsymbol{\theta}$.

The symbol, $\mathbf{V}$, denotes a family of egalitarian vote functions, one for each party, and under which all voters are counted equally. The formal model of the Appendix considers a more general class of vote functions where the voters vary in their weights, thus allowing for complex electoral rules. In the Appendix, the egalitarian family is denoted $\mathbf{V}_{\mathrm{e}}$. The symbol, Ψ, denotes the Gumbel stochastic distribution on the errors. To simplify notation in the applications that follow we delete reference to $\mathbf{V}$ and Ψ.

This empirical model assumes that the utility function of voter i is given by the expression

$$u_{ij}(x_i, z_j) = \Lambda_j + (\theta_j \cdot \eta_i) - \beta \|x_i - z_j\|^2 + \varepsilon_j.$$

Here, the *spatial coefficient* is denoted β and $\mathbf{\Lambda} = \{\Lambda_j : j \in P\}$ are the *exogenous valences* (relative to a baseline party, k^*).[19] The relative *exogenous* valence, Λ_j, gives the average belief of the voters in the electorate concerning the quality of the leader of party j in comparison to the leader of the baseline party, k^*. The symbol, $\boldsymbol{\theta}$, denotes a set of m-vectors $\{\theta_j\}$ representing the effect of the m different sociodemographic parameters (class, domicile, education, income, religious orientation, etc.) on the beliefs of the various subgroups in the polity on the competence of party j. The symbol η_i is an m-vector denoting the ith individual's relevant "sociodemographic" characteristics. The composition $(\theta_j \cdot \eta_i)$ is the scalar product and can be interpreted as the group-specific valence ascribed to party j as a consequence of the various sociodemographic characteristics of voter i. Again, these sociodemographic variables will be normalized with respect to the baseline party k^*, essentially by estimating $((\theta_j - \theta_{k^*}) \cdot \eta_i)$. This scalar term is called the *total sociodemographic valence* of voter i for party j. The tth term in this scalar is called the *sociodemographic valence* of i as a result of membership by i of the tth group, or, more briefly, the tth group specific sociodemographic valence for the leader of party j.[20]

[19] Note that in the empirical models discussed below, these are specified relative to the baseline party, the DYP.

[20] For example, in Tables 6 and 7 there are six sociodemographic variables, so $m = 6$. An individual who is Alevi has $\eta_{i,\mathrm{Alevi}} = 1$. The coefficient for the CHP party for an Alevi is 3.089 in 1999, and this is the group-specific valence that a voter who is a member of the group of Alevi voters has for this party. Note again that this is specified relative to the baseline party, the DYP. These valences may be the result of the

The vector $\mathbf{z} = (z_1, \ldots, z_p) \in X^p$ is the set of party positions, while $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$ is the set of ideal points of the voters in N. When β is assumed zero then the model is called pure *sociodemographic* (SD), and denoted $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta})$. When $\{\theta_j\}$ are all assumed zero then the model is called *pure spatial*, and denoted $\mathbb{M}(\boldsymbol{\Lambda}, \beta)$. The pure spatial model implicitly assumes that the ranking over valence is identical among voters.The empirical model, $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \beta)$, including the sociodemographic terms is called *joint*. These sociodemographic variables allow us to incorporate characteristics common to specific groups of supporters of any party, and this permits the valence ranking to vary among voters in a way which depends on sociodemographics. Not accounting for these characteristics in the analysis will bias the estimates of the exogenous valences of the parties. Tables 4 and 5 give the details of the pure spatial MNL models for the elections of 1999 and 2002 in Turkey, while Tables 6 and 7 give the details of the joint MNL models. The differences in log marginal likelihoods for the three different models then gives the log Bayes' factor for the pairwise comparisons.[21] The log Bayes' factors show that the joint and pure spatial MNL models were clearly superior to the SD models. In addition, the joint models were superior to the pure spatial models.[22] We can infer that, though the sociodemographic variables are useful, by themselves they do not give an accurate model of voter choice.[23] It is necessary to combine the pure spatial model, including the valence terms, with the sociodemographic valences to obtain a superior estimation of voter choice.

Comparing Tables 4 and 5, it is clear that the relative valences of the ANAP and MHP, under the pure spatial model, dropped between 1999 and 2002. In 1999, the estimated Λ_{ANAP} was $+0.336$, while the confidence interval on Λ_{ANAP} for 1999 in Table 4 shows that the hypothesis that $\Lambda_{\text{ANAP}} = 0$ should be rejected. In contrast, the estimated value of Λ_{ANAP} for 2002 was -0.31, and the confidence interval on Λ_{ANAP} does not allow us to reject the hypothesis that $\Lambda_{\text{ANAP}} = 0$.[24] Similarly, Λ_{MHP} fell from a significant value of +0.666 in 1999 to -0.12 in 2002. The estimated relative valence, Λ_{AKP}, of the new Justice and Development Party (AKP) in 2002 was $+0.78$, in comparison to the valence of the FP of -0.159 in 1999. Since the AKP can be regarded as a transformed FP, under the leadership of Recep Tayyip Erdogan, we can infer from the confidence intervals on these two relative valences that this was a significant change due to Erdogan's leadership.[25]

Footnote 20 continued
perception of the leader's ability, as displayed in the past, or of the particular partiality of these voters to choose the party, independently of the party's policy position.

[21] Since the Bayes' factor (Kass and Raftery 1995) for a comparison of two models is simply the ratio of marginal likelihoods, the log Bayes' factor is the difference in log likelihoods.

[22] The log Bayes factors for the joint models over the sociodemographic models were highly significant at +31 in 1999 and +58 in 2002. The Bayes' factors for the joint over the spatial models were also significant, and estimated to be +6 and +5 in 1999 and 2002, respectively.

[23] Sociodemographic models are standard in the empirical voting literature.

[24] These tables show the standard errors of the coefficients, as well as the *t-values*, the ratios of the estimated coefficient to the standard error.

[25] Although Erdogan was the party leader, Abdullah Gul became Prime Minister after the November 2002 election because Erdogan was banned from holding office. Erdogan took over as Prime Minister after winning a by-election in March 2003.

Table 4 Pure Spatial Model of the Turkish election 1999

Party name		Λ_k	Std. error	$\lvert t\text{-Value}\rvert$
Democratic Left Party	DSP	0.724*	0.153	4.73
Nationalist Action Party	MHP	0.666*	0.147	4.53
Virtue Party	FP	−0.159	0.175	0.9
Motherland Party	ANAP	0.336	0.153	2.19
True Path Party	DYP	–	–	–
Republican People's Party	CHP	0.734*	0.178	4.12
People's Democracy Party	HADEP	−0.071	0.232	0.3
Normalized with respect to DYP				
Spatial coefficient β	0.375*		0.088	4.26
Convergence coefficient c	1.49*		0.22	6.77

$n = 635$; Log marginal likelihood (LML) $= -1183$
* Significant with probability <0.001

Table 5 Pure Spatial Model of the Turkish election 2002

Party Name		Λ_k	Std. error	$\lvert t\text{-Value}\rvert$
Justice and Development Party	AKP	0.78*	0.15	5.2
Republican People's Party	CHP	1.33*	0.18	7.4
True Path Party	DYP	–	–	–
Nationalist Action Party	MHP	−0.12	0.18	0.66
Young Party	GP	–	–	–
People's Democracy Party	HADEP	0.43	0.21	2.0
Motherland Party	ANAP	−0.31	0.19	1.63
Normalized with respect to DYP				
Spatial coefficient β	1.52*		0.12	12.66
Convergence coefficient c	5.94*		0.27	22.0

$n = 483$; Log marginal likelihood (LML) $= -737$
* Significant with probability <0.001

It should be noted that the β coefficients for the pure spatial models were 0.375 in 1999, and 1.52 in 2002. Both of these are estimated to be non-zero at the 0.001 level. Indeed, they are significantly different from each other,[26] suggesting that electoral preferences over policy had become more intense.

We first use the results of the formal pure spatial model given in the Appendix to compute estimates of the convergence coefficients. These computations suggest that convergence to an electoral center is not to be expected in these elections. We then use simulation to determine the LNE of the empirical joint models, again showing non-convergence. This allows us to obtain information about activist support for the parties.

[26] The 95% confidence interval for β_{1999} is [0.2,0.55] and for β_{2002} it is [1.28,1.76]

Table 6 Joint Model of the 1999 Election in Turkey (normalized with respect to DYP)

Variable	Party	Coefficient		95% Confidence interval	
		Est	Std Err	Lower bound	Upper bound
Spatial coeff. β		0.456*	0.104	0.243	0.648
Relative valence Λ_k	ANAP	−0.114	0.727	−1.513	1.227
	CHP	−0.673	0.770	−2.166	0.786
	DSP	0.463	0.720	−0.930	1.825
	FP	1.015	0.878	−0.709	2.755
	HADEP	−0.610	1.230	−3.004	1.803
	MHP	2.447*	0.669	1.167	3.664
Age	ANAP	0.001	0.012	−0.021	0.023
	CHP	−0.009	0.013	−0.033	0.016
	DSP	−0.008	0.012	−0.031	0.014
	FP	−0.023	0.014	−0.050	0.003
	HADEP	−0.053*	0.023	−0.103	−0.014
	MHP	−0.044*	0.012	−0.067	−0.022
Education	ANAP	0.006	0.065	−0.115	0.130
	CHP	0.106	0.063	−0.012	0.232
	DSP	0.077	0.058	−0.024	0.197
	FP	−0.129	0.081	−0.285	0.018
	HADEP	0.144	0.097	−0.038	0.335
	MHP	−0.060	0.061	−0.175	0.070
Urban	ANAP	0.531	0.367	−0.156	1.279
	CHP	0.354	0.395	−0.374	1.078
	DSP	0.582	0.359	−0.147	1.271
	FP	0.417	0.416	−0.418	1.183
	HADEP	0.264	0.634	−0.918	1.497
	MHP	−0.201	0.378	−0.922	0.593
Kurd	ANAP	1.132	0.924	−0.410	3.138
	CHP	1.715*	0.911	0.194	3.637
	DSP	−0.102	1.083	−2.650	2.098
	FP	1.116	0.972	−0.733	3.024
	HADEP	5.898*	0.926	4.290	7.904
	MHP	0.063	0.933	−1.751	2.148
Soc. Econ. Status	ANAP	0.080	0.165	−0.302	0.394
	CHP	0.163	0.176	−0.195	0.499
	DSP	−0.010	0.158	−0.322	0.333
	FP	0.120	0.179	−0.230	0.458
	HADEP	−0.119	0.264	−0.598	0.384
	MHP	0.168	0.159	−0.147	0.469
Alevi	ANAP	−0.697	0.972	−2.687	1.168

Table 6 continued

Variable	Party	Coefficient		95% Confidence interval	
		Est	Std Err	Lower bound	Upper bound
	CHP	3.089	0.693	1.965	4.715
	DSP	0.934	0.729	−0.383	2.423
	FP	0.346	0.939	−1.374	2.007
	HADEP	1.355	0.972	−0.332	3.605
	MHP	−0.873	0.925	−3.225	0.676

$n = 635$; Log marginal likelihood = −1178
* Prob <0.001

2.1 The 2002 election

Figure 3 shows the smoothed estimate of the voter ideal points in 2002. This distribution gives the 2×2 voter covariance matrix, with an electoral variance on the first axis (religion) estimated to be 1.18 while the electoral variance on the second axis (nationalism) was 1.15. The total electoral variance was $\sigma^2 = 2.33$, with an electoral standard deviation of $\sigma = 1.52$ The covariance between the two axes was equal to 0.74.

Thus, the voter covariance matrix is

$$\nabla_0 = \begin{bmatrix} 1.18 & 0.74 \\ 0.74 & 1.15 \end{bmatrix}$$

with $\text{trace}(\nabla_0) = 2.33$.

The eigenvalues of this matrix are 1.9, with major eigenvector $(+1.0, +0.97)$ and 0.43, with minor eigenvector $(-0.97, +1.0)$. The major eigenvector corresponds to the *principal electoral axis*, aligned at approximately 45° to the religion axis.

For the pure spatial model $\mathbb{M}(\mathbf{\Lambda}, \beta)$, the β coefficient was 1.52, The valence terms are estimated in contrast with the valence of the DYP, and the the party with the lowest relative valence is ANAP with $\Lambda_{\text{ANAP}} = -0.31$. By definition, $\Lambda_{\text{DYP}} = 0$. The vector of relative valences is then

$$\begin{aligned}&(\Lambda_{\text{ANAP}}, \Lambda_{\text{MHP}}, \Lambda_{\text{DYP}}, \Lambda_{\text{HADEP}}, \Lambda_{\text{AKP}}, \Lambda_{\text{CHP}}) \\ &= (-0.31, -0.12, 0.0, 0.43, 0.78, 1.33).\end{aligned}$$

When all parties are at the origin, the probability, ρ_{ANAP}, that a voter chooses ANAP, in the model $\mathbb{M}(\mathbf{\Lambda}, \beta)$, is independent of the voter. The Appendix, Eq. 7, shows that this is given by

$$\begin{aligned}&\frac{\exp(-0.31)}{\exp(-0.31) + \exp(-0.12) + \exp(0.0) + \exp(0.43) + \exp(0.78) + \exp(1.33)} \\ &= [1 + \exp(0.19) + \exp(0.31) + \exp(0.74) + \exp(1.09) + \exp(1.164)]^{-1}\end{aligned}$$

Table 7 Joint Model of the 2002 Election in Turkey (normalized with respect to DYP)

Variable	Party	Coefficient		95% Confidence interval	
		Est	Std Dev	Lower bound	Upper bound
Spatial coeff β		1.445*	0.143	1.180	1.723
Valence Λ_k	AKP	1.968*	0.667	0.708	3.432
	CHP	1.103	0.797	−0.579	2.615
	HADEP	2.596	1.246	−0.254	5.049
	MHP	1.714	0.889	−0.021	3.426
	ANAP	−0.567	0.880	−2.487	1.133
Age	AKP	−0.031	0.011	−0.052	−0.010
	CHP	−0.019	0.013	−0.045	0.005
	HADEP	−0.060	0.024	−0.110	−0.014
	MHP	−0.067	0.017	−0.103	−0.034
	ANAP	−0.004	0.014	−0.031	0.022
Education	AKP	−0.070	0.062	−0.185	0.045
	CHP	−0.007	0.068	−0.136	0.115
	HADEP	−0.142	0.108	−0.365	0.079
	MHP	−0.048	0.079	−0.202	0.106
	ANAP	−0.078	0.076	−0.237	0.064
Urban	DYP	0.050	0.406	−0.770	0.844
	CHP	0.121	0.443	−0.744	1.001
	HADEP	−1.138	0.688	−2.426	0.236
	MHP	−0.570	0.536	−1.649	0.504
	ANAP	0.661	0.479	−0.228	1.628
Kurd	AKP	2.086	1.105	0.203	4.596
	CHP	1.251	1.171	−0.891	3.839
	HADEP	5.996*	1.208	3.960	8.945
	MHP	1.595	1.312	−0.960	4.258
	ANAP	1.603	1.199	−0.535	4.358
Soc. Econ. Status	AKP	0.142	0.160	−0.160	0.457
	CHP	0.198	0.191	−0.196	0.560
	DEHAP	−0.217	0.281	−0.755	0.301
	MHP	0.317	0.204	−0.083	0.703
	ANAP	0.214	0.209	−0.182	0.613
Alevi	AKP	−0.249	0.983	−2.125	1.743
	CHP	2.567*	0.817	1.111	4.489
	DEHAP	0.377	1.045	−1.519	2.540
	MHP	−0.529	1.410	−3.565	2.292
	ANAP	1.392	0.931	−0.323	3.560

$n = 483$; Log marginal likelihood = −732
* Prob <0.001

$$= [1 + 1.2 + 1.36 + 2.09 + 2.97 + 3.2]^{-1}$$
$$= 0.08.$$

Below, we show that the 95% confidence interval on ρ_{ANAP} is [0, 05, 0.11], which includes the actual vote share (5.13%) in 2002.

Appendix shows that the Hessian of the vote share function of ANAP, when all parties are at the origin, is given by the characteristic matrix of ANAP:

$$\begin{aligned} C_{\text{ANAP}} &= 2\beta(1 - 2\rho_{\text{ANAP}})\nabla_0 - I \\ &= 2 \times (1.52) \times [(1 - (2 \times 0.08)]\nabla_0 - I \\ &= (2.55)\begin{bmatrix} 1.18 & 0.74 \\ 0.74 & 1.15 \end{bmatrix} - I \\ &= \begin{bmatrix} 2.01 & 1.88 \\ 1.88 & 1.93 \end{bmatrix}. \end{aligned}$$

Moreover, the convergence coefficient,

$$c = 2\beta(1 - 2\rho_{\text{ANAP}})\text{trace}(\nabla_0) = 2.55 \times 2.33 = 5.94.$$

This greatly exceeds the upper bound of +2.0 for convergence to the electoral origin. The major eigenvalue for the ANAP characteristic matrix is +3.85, with eigenvector (+1.0, +0.98), while the minor eigenvalue is +0.09, with orthogonal, minor eigenvector (−0.98, +1.0). The eigenvectors of this Hessian are almost perfectly aligned with the principal and minor components, or axes, of the electoral distribution.

Although the electoral origin satisfies the first order condition for local equilibrium, it follows from a standard result that the electoral origin is a *minimum* of the vote share function of ANAP, when the other parties are at the same position. On both principal and minor axes, the vote share of ANAP increases as it moves away from the electoral origin, but because the major eigenvalue is much larger than the minor one, we can expect that the AKP (as well the other parties) in equilibrium to adopt positions along a single eigenvector. Figures 3 and 4 present two LNE obtained from simulation of the pure spatial model. These are:

$$\mathbf{z}_1 = \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & 0.16 & -0.69 & 0.40 & -0.50 & 0.47 & 0.23 \\ y : \text{nat} & 0.17 & -0.77 & 0.41 & -0.57 & 0.45 & 0.26 \end{bmatrix}.$$

$$\mathbf{z}_2 = \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & 0.17 & 0.43 & -0.65 & -0.51 & 0.47 & 0.22 \\ y : \text{nat} & 0.18 & 0.43 & -0.72 & -0.56 & 0.45 & 0.25 \end{bmatrix}.$$

Note that all the positions in these two LNE lie close to the principal axis given by the eigenvector (1.0, 1.0). The higher valence parties, the AKP and CHP lie closer to the origin, while the lower valence parties tend to be further from the origin.

In contrast, the estimated positions of the parties for 2002 in Fig. 2 are:

$$\mathbf{z}^* = \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & -2.0 & 0.0 & 0.0 & -2.0 & -0.2 & 1.0 \\ y : \text{nat} & +0.1 & 1.5 & 0.5 & -1.5 & -0.1 & 0.1 \end{bmatrix}.$$

The equilibrium positions of the CHP and MHP, particularly, are very far from their estimated positions.

2.1.1 Errors in the models

The Appendix shows that the standard error on Λ_{ANAP} is $h = 0.19$, so

$$\begin{aligned} \rho_{\text{ANAP}}(\Lambda_{\text{ANAP}} + h) &= \rho_{\text{ANAP}}(\Lambda_{\text{ANAP}}) + h\frac{\mathrm{d}\rho_{\text{ANAP}}}{\mathrm{d}\Lambda} \\ &= \rho_{\text{ANAP}}(\Lambda_{\text{ANAP}}) + h\rho_{\text{ANAP}}(1 - \rho_{\text{ANAP}}). \end{aligned}$$

This gives a standard error of 0.014 and a 95% confidence interval on ρ_{ANAP} of [0.05, 0.11]. Since the standard error on β is 0.12, giving a confidence interval on β of approximately [1.28, 1.76], the standard error on c is 0.27. Using the lower bound on β and upper bound on ρ_{ANAP} gives an estimate for the 95% confidence interval on c of [4.65, 7.38], so we can assert that, with very high probability, the convergence coefficient exceeds 4.0. Another way of interpreting this observation is that even if we use the upper estimate of the relative valence for ANAP, and the lower bound on β, then the joint origin will still be a minimum of the vote share function for ANAP.

We now repeat the analysis for the election of 1999.

2.2 The 1999 election

The empirical model presented in Table 4 estimated the electoral variance on the first axis (religion) to be 1.20 while on the second axis (nationalism) the electoral variance, σ^2, was 1.14, giving a total electoral variance, σ^2, of 2.34, with the covariance between the two axes equal to +0.78.

The electoral covariance matrix is the 2×2 matrix

$$\nabla_0 = \begin{bmatrix} 1.20 & 0.78 \\ 0.78 & 1.14 \end{bmatrix}.$$

For the model ,the β coefficient was 0.375, while the party with the lowest valence was FP with $\Lambda_{\text{FP}} = -0.16$. The vector of valences is:

$$\begin{aligned} &(\Lambda_{\text{FP}}, \Lambda_{\text{MHP}}, \Lambda_{\text{DYP}}, \Lambda_{\text{HADEP}}, \Lambda_{\text{ANAP}}, \Lambda_{\text{CHP}}, \Lambda_{\text{DSP}}) \\ &\quad = (-0.16, +0.66, 0.0, -0.071, +0.34, +0.73, +0.72). \end{aligned}$$

When all parties are located at the origin, the probability, ρ_{FP}, that a voter chooses FP under $\mathbb{M}(\boldsymbol{\Lambda}, \beta)$ is equal to

$$\frac{1}{[1 + \exp(0.82) + \exp(0.16) + \exp(0.09) + \exp(0.5) + \exp(0.89) + \exp(0.88)]} = [11.27]^{-1} = 0.08.$$

The standard error on Λ_{FP} is 0.175, so the 95% confidence interval can be estimated to be [0.01, 0.15]. The FP vote share in 1999 was 15.41%, suggesting that the pure spatial model should be extended to include sociodemographic valences.

Now $2\beta(1 - 2\rho_{\mathrm{FP}}) = 2\beta \times (1 - 2 \times (0.08)) = 2 \times 0.38 \times 0.84 = 0.64$, so the characteristic matrix of the FP is

$$C_{\mathrm{FP}} = (0.64) \begin{bmatrix} 1.20 & 0.78 \\ 0.78 & 1.14 \end{bmatrix} - I = \begin{bmatrix} -0.24 & 0.448 \\ 0.448 & -0.27 \end{bmatrix}.$$

$$\text{and} \quad c = 0.64 \times 2.34 = 1.49.$$

Although $c < 2.0$, we can compute the eigenvalues of C_{FP} to be -0.74 with minor eigenvector $(+1, -1.116)$ and $+0.23$, with major eigenvector $(+1, +0.896)$, giving a saddlepoint for the FP Hessian at the joint origin. As with the 2002 election, on the basis of the pure spatial model, we again expect all parties to align along the major eigenvector, at approximately 45° to the religion axis. Note, however, that the standard error on c is of order 0.22, so unlike the result for the election of 2002, we cannot assert that there is a high probability that the convergence coefficient exceeds 2. However, there is a probability exceeding 0.95 that one of the eigenvalues is positive.

In comparing the pure spatial models of the elections of 1999 and 2002, we note there is very little difference between the model predictions.

2.3 Extension of the model for Turkey

We now use the empirical joint model, $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \beta)$, in order to better model party positioning. We use this model to estimate the influence of party activists in a more general activist model, denoted $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\mu}, \beta)$. In the activist model,the activist functions $\boldsymbol{\mu} = \{\mu_j : j \in P\}$ are presumed to be functions of party position, rather than exogenous constants. The idea behind this model is that activists provide campaign contributions to specific parties, and these contributions can be used by the parties to affect valence. For the game theoretic foundations of this model see Grossman and Helpman (1994, 2001). Grossman and Helpman (1996, p. 265) also define two distinct motives for these activists:

> Contributors with an *electoral* motive intend to promote the electoral prospects of preferred candidates. Those with an *influence motive* aim to influence the politicians' policy pronouncements.

Here we use a reduced form of the activist functions, based on Schofield (2006), since we only need the fact that the activist contribution to party j is a differentiable function of the party's position, and positively affects the parties valence. Galiani et al. (2010) provide a partial game theoretic model of this activist game that is consistent with Schofield (2006).

Theorem 1 of the Appendix shows that the first-order condition for a local equilibrium, $\mathbf{z}^* = (z_1^*, \ldots z_p^*)$, in the activist model is given by the set of *gradient balance conditions*:

$$\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j^*) + \frac{1}{2\beta}\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j^*) = 0. \tag{1}$$

Each term, $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j)$ is the *the marginal activist pull (or gradient) at* z_j, giving the marginal activist effects on party j, while the gradient term $\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j) = \left[z_j^{\mathrm{el}} - z_j\right]$ is the *gradient electoral pull on the party*, at z_j, pointing towards its weighted electoral mean, z_j^{el}, as defined for party j in (5) in the Appendix:

$$z_j^{\mathrm{el}} \equiv \sum_{i=1}^{n} \varpi_{ij} x_i, \quad \text{where } [\varpi_{ij}] = \left[\frac{[\rho_{ij} - \rho_{ij}^2]}{\sum_{k\in N}[\rho_{kj} - \rho_{kj}^2]}\right]. \tag{2}$$

The weighted electoral mean essentially weights voter policy preferences by the degree to which the sociodemographic valences influence the choice of the voter.

Note in particular that (2) gives the first-order condition for any of the various models considered here. In particular, if the sociodemographic and activist terms are zero, then (2) reduces to $[\varpi_{ij}] = \frac{1}{n}$, and, by the obvious coordinate transformation, we obtain $z_j = 0$, for all j, as the first-order condition.

The joint model, $\mathbb{M}(\mathbf{\Lambda}, \boldsymbol{\theta}, \beta)$, allows us to draw some inferences about equilibrium positions. First we note that in the joint model, the sociodemographic valences are substitutes for the relative valences. Table 7 shows that the only relative valence that is significantly non zero in 2002 is $\mathbf{\Lambda}_{\mathrm{AKP}}$. A number of the sociodemographic valences are, however, very significant.[27]

Figures 5 gives an LNE, $\mathbf{z}_3$, obtained by simulation of the joint model, $\mathbb{M}(\mathbf{\Lambda}, \boldsymbol{\theta}, \beta)$:

$$\mathbf{z}_3 = \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & 0.12 & 0.26 & 0.40 & -0.50 & -0.58 & 0.19 \\ y : \text{nat} & 0.16 & 0.38 & 0.41 & -0.51 & -0.61 & 0.24 \end{bmatrix}.$$

[27] The Bayes factors, or differences between the log marginal likelihoods of the joint models over the pure spatial models were +5 in both years.

Again the estimated positions are:

$$\mathbf{z}^* = \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & -2.0 & 0.0 & 0.0 & -2.0 & -0.2 & 1.0 \\ y : \text{nat} & +0.1 & 1.5 & 0.5 & -1.5 & -0.1 & 0.1 \end{bmatrix}.$$

Comparing the joint model with the pure spatial model, we see that the equilibrium positions are slightly better predictors for HADEP, MHP, and ANAP

For this joint model, Tables 6 and 7 show that the sociodemographic valences for HADEP (or DEHAP) by Kurdish voters were very high:

$$(\theta_{\text{HADEP}} \cdot \eta_{\text{Kurd}}) = 5.9 \text{ in } 1999$$
$$(\theta_{\text{HADEP}} \cdot \eta_{\text{Kurd}}) = 6.0 \text{ in } 2002.$$

Keeping the other variables at their means in 2002, then changing η_{Kurd} from non-Kurd to Kurd increases the probability of voting for HADEP from 0.013 to 0.45. The high significance level of the sociodemographic variables indicates that the joint electoral model would predict that HADEP would move close to Kurdish voters who tend to be located on the left of the religion axis, and are also anti-nationalistic. The position marked HADEP in Fig. 2 is consistent with this inference.

The joint model also shows that Alevi voters have very high sociodemographic valences for the CHP, with

$$(\theta_{\text{CHP}} \cdot \eta_{\text{Alevi}}) = 3.1 \quad \text{in } 1999$$
$$(\theta_{\text{CHP}} \cdot \eta_{\text{Alevi}}) = 2.6 \quad \text{in } 2002.$$

The Alevis are a non-Sunni religious community, who are adherents of Shia Islam rather than Sunni, and may be viewed as supporters of "Kemalism" or the secular state. Again, with other variables at their means, changing η_{Alevi} from non-Alevi to Alevi increases the probability of voting for CHP in 2002 from 0.16 to 0.63. Thus, the joint model indicates that the CHP will move to a vote maximizing position, on the left of the religious axis, again as indicated in Fig. 2.

Conversely, for Alevi voters $\theta_{\text{AKP}} \cdot \eta_{\text{Alevi}}) = -0.25$ in 2002, and we can infer that the AKP may move right to attract Sunni voters.

In the model $\mathbb{M}(\mathbf{\Lambda}, \boldsymbol{\theta}, \beta)$, we do not consider activist terms, so this is equivalent to setting $\{\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}\} = 0$. We can infer from (1) that the first-order balance condition will be satisfied at a vector $\mathbf{z} = (z_1, \ldots z_p)$ when

$$\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j} \equiv \left[z_j^{\text{el}} - z_j\right] = 0, \quad \text{for each } j.$$

Thus, we can use $\mathbf{z}_3$ as the estimator for the vector of weighted electoral means. We find that

$$
\begin{aligned}
\mathbf{z}^* - \mathbf{z}_3 = & \\
& \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & -2.0 & 0.0 & 0.0 & -2.0 & -0.2 & 1.0 \\ y : \text{nat} & +0.1 & 1.5 & 0.5 & -1.5 & -0.1 & 0.1 \end{bmatrix} \\
& - \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & 0.12 & 0.26 & 0.40 & -0.50 & -0.58 & 0.19 \\ y : \text{nat} & 0.16 & 0.38 & 0.41 & -0.51 & -0.61 & 0.24 \end{bmatrix}. \\
& = \begin{bmatrix} \text{Party} & \text{CHP} & \text{MHP} & \text{DYP} & \text{HADEP} & \text{ANAP} & \text{AKP} \\ x : \text{rel} & -3.2 & -0.26 & -0.40 & -1.50 & +0.38 & 0.81 \\ y : \text{nat} & -0.15 & +1.12 & 0.09 & -0.99 & +0.51 & -0.14 \end{bmatrix}
\end{aligned}
$$

Assuming that this vector is an LNE with respect to the full model, $\mathbb{M}(\mathbf{\Lambda}, \boldsymbol{\mu}, \beta)$ involving activists, then by (13) in the Appendix, we can make the identification:

$$
\frac{1}{2\beta}\left[\frac{\mathrm{d}\mu_1}{\mathrm{d}z_1}, \ldots, \frac{\mathrm{d}\mu_p}{\mathrm{d}z_p}\right] = \mathbf{z}^* - \mathbf{z}_3
$$

Here, $\{\frac{\mathrm{d}\mu_1}{\mathrm{d}z_1}, \ldots, \frac{\mathrm{d}\mu_p}{\mathrm{d}z_p}\}$ are the marginal activist pulls at the equilibrium vector $\mathbf{z}^*$. Under the hypothesis that the joint model with activists is valid, then the difference between these two vectors gives us an estimate of the vector of marginal pulls on the parties:

The estimated activist pull on HADEP is very high, pulling the party to the left on the religion axis, and in an anti-nationalist direction on the y-axis. Similarly,the estimated activist pull on the CHP is even higher on the religious axis, pulling the party in a secular direction, and we can infer that this is due to the influence of Alevi voters.

As a consequence, this asymmetry will cause Alevi activists to provide further differential support for the CHP. It is thus plausible that secular voters (on the left of the religious axis in Figs. 1, 2) would offer further support to the CHP, located close to them. This would affect the party's marginal activist pull, and induce the CHP leader to move even further left, towards its inferred equilibrium position in the full activist model.

We suggest that activist support for the AKP would move it slightly to the right on the religion axis, as well as in an anti-nationalism direction. This would result in its estimated position as in Fig. 2.

In contrast, we might conjecture that the military provides activist support for the MHP on the nationalism axis, and this will move the party to the left in a secular direction, and north on the nationalism axis, resulting in its position in Fig. 2.

Overall, we note that we can expect activist valence to strongly influence party positioning, and we can proxy this support to some degree using the sociodemographic variables. Notice that the sociodemographic variables are estimated at the vector $\mathbf{z}^*$, so the estimated sociodemographic valences have been influenced by activist support. The LNE obtained from the joint model is a hypothetical solution to the vote max-

imizing game involving the parties, based on some empirical assumptions about the underlying nature of the important sociodemographic groups in the polity.

2.4 General remarks on Turkish elections

Although we have not performed a MNL analysis of the 2007 election, it seems obvious that some of the changes in the nature of party strategies were due to changes in the electoral laws. The election results of 1999 were based on an electoral system that was quite proportional, whereas in 2002 and 2007, the electoral system was highly majoritarian. In 2002, for example, the AKP gained 66% of the seats with only 34% of the vote, while in 2007 it took 46.6% of the vote and 340 seats (or 61.8%), reflecting the continuing high valence of Erdogan. Similarly, the CHP went from about 9% of the vote in 1999 (and no seats) to 19% of the vote in 2002, and 32% of the seats. This is mirrored by the increase in the valence estimates of the joint model from $\Lambda_{\mathrm{CHP}} = -0.673$ in 1999 to $\Lambda_{\mathrm{CHP}} = 1.103$, in 2002. In contrast the MHP went from 18% of the vote in 1999 to 8% in 2002, while Λ_{MHP} for the joint model fell from 2.5 to 1.7. The turn around in the vote share of the MHP between 2002 and 2007 could be a result of increasing support for this party from nationalist activist groups in an attempt to offset the high valence and electoral support for the AKP in 2002. Indeed, the increased concentration of the vote share between 1999 and 2007 may be a consequence of the greater significance of activist influence as the electoral system became more majoritarian due to the nature of the electoral cut-off rule.[28]

In such a non-proportional electoral system there are incentives for members of different sociodemographic groups to to engage in strategic voting. There is some indication from the formal model that the intensity of the political contest between secularist, nationalistic, and religious activist groups had increased prior to 2007, and recent events suggest that this is continuing.

After the 2007 election, Abdullah Gul, Erdogan's ally in the AKP was elected as the country's 11th President, despite strong opposition from the army and many secular interests. In late February 2008, the Turkish military invaded the Kurdish controlled territory in north west Iraq in an attempt to destroy the bases of the P.K.K. (the Kurdistan Workers' Party). The secular Constitutional Court has also considered banning many members of the AKP. In September 2008, Turkey formed a Caucasus Stability and Cooperation Platform with five neighboring countries, in reponse to Russian aggression in Georgia, and President Gul visited Armenia, one of the countries in the Platform. On 30 January 2009, Erdogan returned home from the World Economic Forum in Davos after walking out of a televised debate with Shimon Peres, the Israeli President, over Israel's war on the Gaza Strip. The moderator had refused to allow Erdogan to rebut Peres' justification of the war. Erdogan was welcomed back in Turkey as a hero.

However, more secular voters have begun to worry that Erdogan had become more autocratic, and in the municipal elections in March, 2009, the vote for the AKP dropped

[28] The Herfindahl concentration measure of the vote shares went from 0.11 in 1999 to 0.16 in 2002 to 0.27 in 2007.

from 47 to 39%. It appears that the Turkish electorate had divided geographically into four different political regions: a liberal, secular litoral, a conservative interior, with a nationalistic center, and a Kurdish nationalistic southeast.[29] The conflicts between the secular military and the non-secular government have come to a head over the Ergenekon affair, which has involved the prosecution of more than 200 people, allegedly involved in plotting against the state. In February 2010, the government arrested a further 40 people, including three high-ranking ex-military officers, and in March the government proposed constitutional changes that would limit the power of the Constitutional Court, making it more difficult for the Court to ban parties, as it has in the past. The changes would also make it more difficult to restrict membership of the forces to those who had no allegiance to religious groups, and would also permit trials in civilian rather than military courts for officers who were accused of plotting against the government. These constitutional changes require 367 votes out of the 550 in the Parliament. Both opposition parties, the Nationalist Movement Party (MHP) and Republican's Peoples Party (CHP) oppose these changes in the constitution, while the AKP only has 340 votes.

In his visit to Turkey in April 2009, Barack Obama made it clear that in his view, Turkey should become a member of the European Union. At the same time, he urged Turkey to undertake more democratic reforms. Although Turkey has many of the characteristics of a full democracy, there does appear to be severe conflict between the government and secular activist groups such as the military and judiciary.

Although many business interests favor membership of the European Union, the opposition to this by President Sarkozy of France and Chancellor Merkel of Germany may cause Turkey to turn east. In October 2009, Erdogan visited Tehran and met with President Ahmadinejad of Iran, while Turkey and Russia are also discussing the possibility of having Russian gas supplies transit through Turkey. On May 31, there was an attack by Israeli commandoes against a boat traveling in international waters and carrying humanitarian supplies for Gaza. Nine people in the convey were killed. The convoy was partly organized by a Turkish organization, Insani Yardim Vakfi. The repercussions for Turkish Israel relations were likely to be extreme. On 8 June 2010, Erdogan met with President Ahmadinejad and Prime Minister Vladimir Putin of Russia at a regional security summit in Istanbul. Turkey may be shifting from its pro-western stance and seeking to be an independent power in the region.

3 Concluding remarks

Recent works by Acemoglu and Robinson (2006), Boix (2003), and Przeworski et al. (2000) have explored the transition from autocratic regimes to democracy. A recent contribution by Epstein et al. (2006) has emphasized the existence of the category of "partial democracies." These exhibit mixed characteristics of both democratic and autocratic regimes. In fact, Epstein et al. give Turkey as a prime illustration of the possible degree of democratic volatility of a regime. They observe that, in terms of Polity IV scores, Turkey fell from being a full democracy to an autocracy first in the

[29] Asli Aydintasbas in the *New York Times,* 7 April 2009.

mid 1960s and again in the early 1980s, and since then has hovered between partial and full democracy. Epstein et al. (2006, p. 564) also comment, on the basis of their empirical analysis, that "the determinants of the behavior of partial democracies elude our understanding." These models of democratic transitions have tended to consider a single economic axis, and to utilize the notion of a median citizen, or median kingmaker as the unique pivotal player.[30] While these models have been illuminating, we believe it necessary to consider policy spaces of higher dimension and to utilize a stochastic model so as to emphasize the aspect of uncertainty.

The analysis of Turkey in this paper indicates that both religion and nationalism define the political space.[31] The military in Turkey can be represented by a pro-nationalist, pro-secular position, far from the AKP, and it is this phenomenon which means that Turkish politics cannot be understood in terms of a median voter. Modeling partial democracies would seem to require a very explicit analysis of the power of activist groups.

This paper has applied a theoretical stochastic model to present an empirical analysis of elections in Turkey, and argues that there is no evidence of a centripetal tendency towards an electoral center. Instead it suggests that activist groups will tend to be located far from the electoral center. Once the sociodemographic valences have caused the parties to move away from the center to gain electoral support, the influence of activists will separate the parties even further, pulling them towards policy positions preferred by the activists. Thus, simulation of the joint model with sociodemographic valence can be used to infer aspects of this activist influence.

Acknowledgments This paper is based on work supported by NSF grant 0715929 and a Weidenbaum Center grant to Schofield. Zakharov thanks the DECAN laboratory, State University-Higher School of Economics, for support. This version was completed while Schofield was the Glenn Campbell and Rita Ricardo-Campbell National Fellow at the Hoover Institution, Stanford, 2010.

Appendix: Formal and empirical electoral models

The model with activists

The electoral model presented here is an extension of the multiparty stochastic model of McKelvey and Patty (2006), modified by inducing asymmetries in terms of valence. The justification for developing the model in this way is the empirical evidence that valence is a natural way to model the judgements made by voters of party leaders and candidates. There are a number of possible choices for the appropriate model for multiparty competition. The simplest one, which we first present, is that the utility function for leader j is proportional to the popular support, V_j, of the party in the election.[32]

[30] See Gallego and Pitchik (2004).

[31] Schofield and Sened (2006) found the electoral model for Israel to be very similar to Turkey, with two electoral axes, religion and security. Schofield and Zakharov (2010) found nationalism to be one of the principal axes in Russia, but the second axis was defined by attitudes to capitalism/communism, perhaps comparable to religion.

[32] The popular support may be identical to the vote share in a democratic election, or may be weighted by individual characteristics, such as domicile, income or ownership of land, in non-democratic polities.

With this assumption, we can examine the conditions on the parameters of the stochastic model which are necessary for the existence of a pure strategy Nash equilibrium (PNE). Because the vote share functions are differentiable, we use calculus techniques to obtain conditions for positions to be locally optimal. Thus, we examine what we call *local pure strategy Nash equilibria* (LNE). From the definitions of these equilibria it follows that a PNE must be a LNE, but not conversely. A necessary condition for an LNE is thus a necessary condition for a PNE. A sufficient condition for an LNE is not a sufficient condition for PNE. Indeed, additional conditions of concavity or quasi-concavity are required to guarantee existence of PNE.

The stochastic model essentially assumes that candidates cannot predict vote response precisely, but that they can estimate the effect of policy proposals on the expected vote share. In the model with valence, the stochastic element is associated with the weight given by each voter, i, to the average perceived quality or valence of each candidate. We also consider a formal model where the perceptions of the leader qualities vary across different sociodemographic groups in the society.

The data of the spatial model is a distribution, $\{x_i \in X\}_{i \in N}$, of voter ideal points for the members of the electorate, N, of size n. We assume that X is a subset of Euclidean space, of dimension w with w finite. Without loss of generality, we adopt coordinate axes so that $\frac{1}{n}\Sigma x_i = 0$. By assumption $0 \in X$, and this point is termed the *electoral mean*, or alternatively, the *electoral origin*. Each of the parties in the set $P = \{1, \ldots, j, \ldots, p\}$ chooses a policy, $z_j \in X$, to declare prior to the specific election to be modeled. Let $\mathbf{z} = (z_1, \ldots, z_p) \in X^p$ be a typical vector of party policy positions.

Given $\mathbf{z}$, each citizen, i, is described by a utility vector

$$\mathbf{u}_i(x_i, \mathbf{z}) = (u_{i1}(x_i, z_1), \ldots, u_{ip}(x_i, z_p))$$

where

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) - \beta||x_i - z_j||^2 + \varepsilon_{j.} = u^*_{ij}(x_i, z_j) + \varepsilon_j. \tag{3}$$

Here $u^*_{ij}(x_i, z_j)$ is the observable component of utility. The constant term, λ_j, is the fixed or *exogenous valence* of party j, The function $\mu_j(z_j)$ is the component of valence generated by activist contributions to agent j. We can also refer to this term as *endogenous valence*. The term β is a positive constant, called the *spatial parameter*, giving the importance of policy difference defined in terms of a metric induced from the Euclidean norm, $||\cdot||$, on X. The vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_j, \ldots, \varepsilon_p)$ is the stochastic error, whose multivariate cumulative distribution will be denoted by Ψ. The notation $\lambda_j + \mu_j(z_j)$ is intended to imply that this is the average valence for party j among the electorate, but the realized valence is a distributed by Ψ. The most common assumption in empirical analyses is that Ψ is the *Type I extreme value distribution* (sometimes called Gumbel). This cumulative distribution has the closed form

$$\Psi(x) = \exp\left[-\exp\left[-x\right]\right].$$

The theorems presented in this Appendix are based on this assumption. This distribution assumption is the basis for much empirical work based on MNL estimation (Dow and Endersby 2004). The variance of ε_j is fixed at $\frac{\pi^2}{6}$, so that by definition β has dimension L^{-2}, where L is whatever unit of measurement is used in X.

In empirical models, the exogenous valences are simply real numbers, estimated by the model. Since they are all finite, they can be ranked. We therefore assume that the exogenous valence vector is given by

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p) \text{ satisfies } \lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1.$$

This is a strong assumption, in that it assumes that every voter ranks the parties in this fashion. Adding sociodemographic valences, as in the body of the paper, means that this ranking over valences differs among the electorate.

Voter behavior is modeled by a probability vector. The probability that a voter i chooses party j at the vector $\mathbf{z}$ is

$$\rho_{ij}(\mathbf{z}) = \Pr[[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)], \quad \text{for all } l \neq j].$$

Here Pr stands for the probability operator generated by the distribution assumption on ε.

With this distribution assumption on Ψ, it follows, for each voter i, and leader j, that

$$\rho_{ij}(\mathbf{z}) = \frac{\exp[u^*_{ij}(x_i, z_j)]}{\sum_{k=1}^{p} \exp u^*_{ik}(x_i, z_k)}. \tag{4}$$

For any voting model the *likelihood* of a model is

$$\mathbb{L} = \prod_{i \in N,\ j_i \in P} \rho_{ij_i}(\mathbf{z}),$$

where j_i is the party that i chooses. The log likelihood of the model is $\log_e(\mathbb{L})$. Clearly as $\mathbb{L}$ approaches 0 then $\log_e(\mathbb{L})$ approaches $-\infty$.

To compare two models, $\mathbb{M}_1$ and $\mathbb{M}_2$, the *Bayes Factor* is $\mathbb{L}(\mathbb{M}_1)/\mathbb{L}(\mathbb{M}_2)$ and the *log Bayes factor* of $\mathbb{M}_1$against $\mathbb{M}_2$ is $\log_e(\mathbb{L}(\mathbb{M}_1) - \log_e(\mathbb{L}(\mathbb{M}_2)$. A log Bayes factor over 5.0 for $\mathbb{M}_1$ against $\mathbb{M}_2$ is considered strong support for $\mathbb{M}_1$ (Kass and Raftery 1995).

The *expected popular support* for leader j is

$$V_j(\mathbf{z}) \equiv \sum_{i \in N} s_{ij} \rho_{ij}(\mathbf{z}).$$

Here $\{s_{ij}\}$ are different weights that can be associated with different voters. In the case all weights are equal to $\frac{1}{n}$, we call the model *egalitarian.*

It is useful to have a formal model where voter weights differ. For example, in US Presidential elections, it is not the vote share *per se* but the *share of the electoral college total.* Voter weights in different States will therefor vary.

To present the model we now regard $\mathbf{V} = \{V_j : j \in P\}$ as a set of *vote share functions*, and identify $\mathbf{V}$ as a differentiable *profile function*, $\mathbf{V} : X^p \to \mathbb{R}^p$. We denote the egalitarian profile function as $\mathbf{V}_e$.

In this stochastic electoral model, it is assumed that each party j chooses z_j to maximize V_j, conditional on $\mathbf{z}_{-j} = (z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_p)$.

Thus a vector $\mathbf{z}^*=(z_1^*, \ldots, z_{j-1}^*, z_j^*, z_{j+1}^*, \ldots, z_p^*)$ is called a *local strict Nash equilibrium* (LSNE) if each z_j strictly locally maximizes V_j, conditional on $\mathbf{z}_{-j}$, while $\mathbf{z}^*$ is a *local weak Nash equilibrium* (LNE) if each z_j weakly locally maximizes V_j, conditional on $\mathbf{z}_{-j}$. The notion of LSNE is convenient so as to avoid degeneracy problems associated with the Hessians.

In the same way the vector $\mathbf{z}^*$ is a *strict (or weak) pure strategy Nash equilibrium* (PSNE or PNE) if each party j chooses z_j to strictly (or weakly) maximize V_j on X.

Now assume that the vector $\mathbf{z}$ is fixed, and let $\rho_{ij}(\mathbf{z}) = \rho_{ij}$ be the probability that i picks j. Define the p by n matrix array of weights by

$$[\varpi_{ij}] \equiv \left[\frac{s_{ij}[\rho_{ij} - \rho_{ij}^2]}{\sum_{k \in N} s_{kj}[\rho_{kj} - \rho_{kj}^2]} \right] \tag{5}$$

The vector $\sum_i \varpi_{ij} x_i$ is a convex combination of the set of voter ideal points and is called the *weighted electoral mean* for party j. Define

$$z_j^{\mathrm{el}} \equiv \sum_{i=1}^{n} \varpi_{ij} x_i \quad \text{and} \quad \frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j) \equiv \left[z_j^{\mathrm{el}} - z_j \right].$$

Then the *balance equation* for z_j^* is given by the expression

$$\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j^*) + \frac{1}{2\beta} \frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j^*) = 0. \tag{6}$$

The term $\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j)$ is the *marginal electoral pull of party* j at the point z_j and can be regarded as a gradient vector, at z_j, pointing towards the weighted electoral mean of the party. (Note that this electoral pull depends on the positions of all leaders.) When z_j is equal to the weighted electoral mean then the electoral pull is zero. The gradient vector $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j)$ is called *the marginal activist pull for party* j at z_j.

When $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j) = 0$, then then the balance equation reduces to setting $z_j = z_j^{\mathrm{el}}$.

If $\mathbf{z}^* = (z_1^*, \ldots z_j^*, \ldots z_p^*)$ is such that each z_j^* satisfies the balance equation then call $\mathbf{z}^*$ a *balance solution*. The balance solution requires that the electoral and activist gradients are directly opposed, for every party leader.

The model just presented is denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \beta; \mathbf{V})$. Schofield (2006) proves the following theorem for this model.

Theorem 1 *Consider the electoral model* $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \beta; \mathbf{V})$ *based on the distribution,* Ψ, *including both exogenous and activist valences, and defined by the family* $\mathbf{V}$ *of vote share functions.*

(i) *The first order condition for* $\mathbf{z}^*$ *to be an LSNE is that it is a balance solution.*
(ii) *If all activist valence functions are sufficiently concave,*[33] *then a balance solution will be a PNE.*

In the full activist model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \beta; \mathbf{V})$, with valence functions $\{\mu_j\}$ that are not identically zero or constant, then it is the case that generically $\mathbf{z}_0$ cannot satisfy the first order conditions for LNE even when $\mathbf{V}$ is egalitarian. Instead the vector $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}$ "points towards" the position at which the activist valence for leader j is maximized. When this marginal or gradient vector, $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}$, is increased (as activist groups become more willing to contribute to leader j) then the equilibrium position is pulled away from the weighted electoral mean of the leader, and we can say the "activist effect" for the leader is increased. In the case of two opposed leaders, j and k, if the activist valence functions are fixed, but the exogenous valence, λ_j, is increased, or λ_k, is decreased, then the weighted electoral mean, z_j^{el}, approaches the electoral origin. Thus, the local equilibrium of leader j is pulled towards the electoral origin. We can say the "electoral effect" is increased.

The egalitarian model without activists

In the case that the activist valence functions are identically zero, or constant, we denote the model by $\mathbb{M}(\boldsymbol{\lambda}, \beta; \mathbf{V})$. The key consideration for the egalitarian model, $\mathbb{M}(\boldsymbol{\lambda}, \beta; \mathbf{V}_e)$, when all voter weights are identical, is whether the electoral origin is a LSNE. For this model it can be shown that if all parties are at the same position, so $\mathbf{z}^* = (z^*, z^*, \ldots z^*)$ then every $\{\rho_{ij}(\mathbf{z}^*) : i \in N\}$ is independent of i, and can thus be written $\rho_j(\mathbf{z}^*)$. This implies that all α_{ij} in (5) are identical at $\mathbf{z}^*$ and equal to $\frac{1}{n}$. Thus, when there is only exogenous valence, the equation $z_j^* = \frac{1}{n}\Sigma x_i$ satisfies the balance solution for all j. By an appropriate coordinate change, we can assume $\frac{1}{n}\Sigma x_i = 0$. In this case, all marginal electoral pulls are zero at $\mathbf{z}_0 = (0, \ldots 0)$, so $\mathbf{z}_0$ satisfies the first-order conditions. However, to determine whether $\mathbf{z}_0$ is an LNE it is necessary to examine the Hessians of the vote share functions.

We first define the *electoral covariance matrix,* ∇_0, and then use ∇_0 to define the convergence coefficient of the model $\mathbb{M}(\boldsymbol{\lambda}, \beta; \mathbf{V}_e)$. Let $X = \mathbb{R}^w$ be endowed with a system of coordinate axes $r = 1, \ldots, w$. For each coordinate axis let $\xi_r = (x_{1r}, x_{2r}, \ldots, x_{nr})$ be the vector of the rth coordinates of the set of n voter ideal points. The scalar product of ξ_r and ξ_s is denoted $(\xi_r \cdot \xi_s)$. Let $(\sigma_r \cdot \sigma_s) = \frac{1}{n}(\xi_r \cdot \xi_s)$ be the electoral covariance between the r and s axes, and σ_s^2 be the variance on the s axis.

[33] By this we mean that the eigenvalues of the activist functions are negative and of sufficient magnitude everywhere. That is to say, there exists $\alpha < 0$, such that all eigenvalues $<\alpha$ is sufficient to guarantee existence of a PNE.

(i) The symmetric $w \times w$ *electoral covariance matrix* about the origin is denoted ∇_0 and is defined by

$$\nabla_0 \equiv [(\sigma_r \cdot \sigma_s)]_{s=1,\ldots,w}^{r=1,\ldots,w}.$$

(ii) The total electoral variance is $\sigma^2 \equiv \sum_{s=1}^{w} \sigma_s^2 = \text{trace}(\nabla_0)$.

(iii) At the vector $\mathbf{z}_0 = (0, \ldots, 0)$ the probability $\rho_{ij}(\mathbf{z}_0)$ that i votes for party j is independent of i, and is given by

$$\rho_j = \left[1 + \sum_{k \neq j} \exp\left[\lambda_k - \lambda_j\right] \right]^{-1}. \tag{7}$$

(iv) The Hessian of the egalitarian vote share function of party j at $\mathbf{z}_0$ is a positive multiple of the w by w *characteristic matrix*,

$$C_j \equiv 2\beta(1 - 2\rho_j)\nabla_0 - I. \tag{8}$$

(Here I is the identity matrix.)

The *convergence coefficient* of the egalitarian model, $\mathbb{M}(\boldsymbol{\lambda}, \beta; \mathbf{V}_e)$, is defined to be

$$c \equiv c(\boldsymbol{\lambda}, \beta; \mathbf{V}_e) \equiv 2\beta[1 - 2\rho_1]\sigma^2. \tag{9}$$

Note that the β-parameter has dimension L^{-2}, so that c is dimensionless, and therefore independent of the scale used to measure positions. We can therefore use c to compare different models.

Theorem 2 *Consider the electoral model $\mathbb{M}(\boldsymbol{\lambda}, \beta; \mathbf{V}_e)$ where all activist valence functions are zero (or constant) and $\mathbf{V}_e$ is the egalitarian party profile.*

(i) *The joint origin $\mathbf{z}_0 = (0, \ldots, 0)$ satisfies the first order condition to be a LSNE for this model.*

(ii) *In the case that X is w dimensional then the necessary condition for $\mathbf{z}_0$ to be a LNE for this model is that $c(\lambda, \beta; \mathbf{V}_e) \leq w$.*

(iii) *In the case that X is 2 dimensional, a sufficient condition for $\mathbf{z}_0$ to be a LSNE for this model is that $c(\boldsymbol{\lambda}, \beta; \mathbf{V}_e) < 1$.*

The proof and some applications of Theorem 2 are given in Schofield (2007a,b)

Empirical models

In empirical models with exogenous valence alone it is necessary to estimate the model with respect to the valence of a baseline party, say k^*. We set $\Lambda_j = \lambda_j - \lambda_{k^*}$, and call these the relative valences. We denote this egalitarian model by $\mathbb{M}(\boldsymbol{\Lambda}, \beta; \mathbf{V}_e)$.

At the joint origin,$\mathbf{z}_0$, we see that

$$\rho_{ij}(\mathbf{z}_0) = \frac{\exp(\lambda_j)}{\sum_{k=1}^{p}\exp(\lambda_k)} = \frac{\exp(\lambda_j - \lambda_{k^*})}{\sum_{k=1}^{p}\exp(\lambda_k - \lambda_{k^*})} = \frac{\exp(\Lambda_j)}{\sum_{k=1}^{p}\exp(\Lambda_k)} \tag{10}$$

is again independent of the individual, i, and can be written as ρ_j.

To estimate the standard error on ρ_j, we use Taylor's Theorem, which asserts that

$$\rho_j(\Lambda_j + h) = \rho_j(\Lambda_j) + h\frac{\mathrm{d}\rho_j}{\mathrm{d}\Lambda_j} = \rho_j(\Lambda_j) + h\rho_j(1-\rho_j). \tag{11}$$

Empirical models with sociodemographic valences

As described in the body of the paper, in empirical applications with sociodemographic variables, we typically assume that $\mathbf{V}$ is the egalitarian party profile function,$\mathbf{V}_e$, so the model $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \beta; \mathbf{V}_e)$ is based on the assumption that voter utility has the form

$$u_{ij}(x_i, z_j) = \Lambda_j + (\theta_j \cdot \eta_i) - \beta\|x_i - z_j\|^2 + \varepsilon_{j.}$$

The estimate of voter i's valence will then be $\Lambda_j + (\theta_j \cdot \eta_i)$, so this will vary from one voter to another. A consequence of this is that, in the expression (5) for the weighted electoral mean, even when all parties are at the origin, then the denominator term $\{\rho_{kj}(\mathbf{z}_0) : k \in N\}$ will depend on voter k. This implies that voters will be weighted differently, and generically, $\mathbf{z}_0$ will not satisfy the first order condition for LNE. However, the joint empirical model, $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \beta; \mathbf{V}_e)$, assumes that the sociodemographic effects are independent of party positions, and this implies $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j} = 0$, for all j. Using (6), we infer that the various LNE obtained by simulation of the joint model provides an estimate of a set of vectors of weighted electoral means: $\{\mathbf{z}^{\mathrm{el}} = (z_1^{\mathrm{el}}, \ldots, z_p^{\mathrm{el}})\}$.

Assuming that the estimated party positions are given by the vector $\mathbf{z}^* = (z_1^*, \ldots, z_p^*)$ and that this is in equilibrium with respect to the full activist model, then choosing one joint LNE, $\mathbf{z}^{\mathrm{el}}$, gives an estimate of

$$\left[z_j^{\mathrm{el}} - z_j^*\right] \equiv \frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j^*) = -\frac{1}{2\beta}\left[\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}\right]. \tag{12}$$

Thus,

$$\left[\mathbf{z}^* - \mathbf{z}^{\mathrm{el}}\right] = \frac{1}{2\beta}\left[\frac{\mathrm{d}\mu_1}{\mathrm{d}z_1}, \ldots, \frac{\mathrm{d}\mu_p}{\mathrm{d}z_p}\right]. \tag{13}$$

This observation suggests how the gradients of the activist valence functions may be inferred from a comparison of LNE of the joint empirical model with the estimated political configuration.

References

Acemoglu D, Robinson J (2006) Economic origins of dictatorship and democracy. Cambridge University Press, Cambridge

Adams J, Merrill S III (1999) Modelling party strategies and policy representation in multiparty elections: why are strategies so extreme. Am J Polit Sci 43:765–781

Aldrich J (1983) A Downsian spatial model with party activists. Am Polit Sci Rev 77:974–990

Austen-Smith D, Banks J (1999) Positive political theory I: collective preferences. The University of Michigan Press, Ann Arbor

Banks J, Duggan J (2005) The theory of probabilistic voting in the spatial model of elections. In: Austen-Smith D, Duggan J (eds) Social choice and strategic decisions. Springer, Berlin Heidelberg New York

Banks J, Duggan J, Le Breton M (2006) Social choice and electoral competition in the general spatial model. J Econ Theory 126:194–234

Benoit K, Laver M (2006) Party policy in modern democracies. Routledge, London

Boix C (2003) Democracy and redistribution. Cambridge University Press, Cambridge

Carkoğlu A, Hinich M (2006) A spatial analysis of Turkish party preferences. Elect Stud 25:369–392

Dow JK, Endersby J (2004) Multinomial probit and multinomial logit: a comparison of choice models for voting research. Elect Stud 23:107–122

Downs A (1957) An economic theory of democracy. Harper and Row, New York

Epstein D, Bates R, Goldstone J, Kristensen I, O'Halloran S (2006) Democratic transitions. Am J Polit Sci 50:551–568

Gallego M, Pitchik C (2004) An economic theory of leadership turnover. J Pub Econ 88:2361–2382

Galiani S, Schofield N, Torrens G (2010) Factor endowments, democracy and trade policy divergence. Working Paper, Washington University, Saint Louis

Grossman GM, Helpman E (1994) Protection for sale. Am Econ Rev 84:833–850

Grossman GM, Helpman E (1996) Electoral competition and special interest politics. Rev Econ Stud 63:265–286

Grossman GM, Helpman E (2001) Special interest groups. MIT Press, Cambridge, MA

Hinich M (1977) Equilibrium in spatial voting: the median voter result is an artifact. J Econ Theory 16:208–219

Hotelling H (1929) Stability in competition. Econ J 39:41–57

Kass R, Raftery A (1995) Bayes factors. J Am Stat Assoc 90:773–795

Laver M, Hunt WB (1992) Policy and party competition. Routledge, New York

Lin T, Enelow MJ, Dorussen H (1999) Equilibrium in multicandidate probabilistic spatial voting. Pub Choice 98:59–82

McKelvey RD, Patty JW (2006) A theory of voting in large elections. Games Econ Behav 57:155–180

Penn E (2009) A model of far-sighted voting. Am J Polit Sci 53:36–54

Poole K, Rosenthal H (1984) U.S. Presidential elections 1968–1980: a spatial analysis. Am J Polit Sci 28:283–312

Przeworski A, Alvarez ME, Cheibub JA, Limongi F (2000) Democracy and development: political institutions and well-being in the world, 1950–1990. Cambridge University Press, Cambridge

Saari D (1997) The generic existence of a core for q-rules. Econ Theory 9:219–260

Schofield N (1999) The heart and the uncovered set. J Econ Suppl 8:79–113

Schofield N (2006) Equilibria in the spatial stochastic model with party activists. Rev Econ Des 10:183–203

Schofield N (2007a) Political equilibrium with electoral uncertainty. Soc Choice Welf 28:461–490

Schofield N (2007b) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980

Schofield N, Milller G (2007) Elections and activist coalitions in the U.S.. Am J Polit Sci 51:518–531

Schofield N, Sened I (2006) Multiparty democracy: elections and legislative politics. Cambridge University Press, Cambridge

Schofield N, Zakharov A (2010) A stochastic model of the 2007 Russian Duma election. Public Choice 142:177–194

Schofield N, Miller G, Martin A (2003) Critical elections and political realignments in the U.S.: 1860–2000. Polit Stud 51:217–240

Schofield N, Claassen C, Ozdemir U, Zakharov A (2011) Estimating the effects of activists in two-party and multi-party systems: comparing the United States and Israel. Soc Choice Welf (this issue). doi:10.1007/s00355-010-0507-0

Stokes D (1963) Spatial models and party competition. Am Polit Sci Rev 57:368–377

Stokes D (1992) Valence politics. In: Kavanagh D (ed) Electoral politics. Clarendon Press, Oxford

Westen D (2007) The political brain. Perseus Books, New York

Soc Choice Welf
DOI 10.1007/s00355-010-0507-0

ORIGINAL PAPER

Estimating the effects of activists in two-party and multi-party systems: comparing the United States and Israel

Norman Schofield · Christopher Claassen · Ugur Ozdemir · Alexei Zakharov

Abstract This article presents an electoral model where activist groups contribute resources to their favored parties. These resources are then used by the party candidates to enhance the electoral perception of their quality or *valence*. We construct an empirical model of the United States presidential election of 2008 and employ the electoral perception of the character traits of the two candidates. We use a simulation technique to determine the local Nash equilibrium, under vote share maximization, of this model. The result shows that the unique vote-maximizing equilibrium is one where the two candidates adopt convergent positions, close to the electoral center. This result conflicts with the estimated positions of the candidates in opposed quadrants of the policy space. The difference between estimated positions and equilibrium positions allows us to estimate the influence of activist groups on the candidates. We compare this estimation with that of Israel for the election of 1996, and show that vote maximization leads low valence parties to position themselves far from the electoral origin. We argue that these low valence parties in Israel will be dependent on support of radical activist groups, resulting in a degree of political fragmentation.

N. Schofield (✉) · C. Claassen · U. Ozdemir
Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA
e-mail: schofield.norman@gmail.com

C. Claassen
e-mail: claassen@wustl.edu

U. Ozdemir
e-mail: uozdemir@artsci.wustl.edu

A. Zakharov
State University-Higher School of Economics Moscow, Moscow, Russia
e-mail: 3ax@mail.ru

Published online: 04 November 2010

1 Introduction

This article offers a unified model of the electoral process in order to account for a number of general empirical observations about the effects of political institutions. As Duverger (1954) and Riker (1953) have observed, there appears to be a relationship between the electoral rule in place, and the number of political parties in the polity. A highly majoritarian (or plurality) system tends to result in just two parties, while an electoral system based on proportional representation (PR) tends to give a fragmented political structure.[1] Many authors have also argued that there is a relationship between fragmention and the durability of government (Taylor and Herman 1971; Warwick 1994). Other authors have argued that these differing constitutional rules profoundly affect the nature of the policy process (Bawn and Rosenbluth 2005; Persson and Tabellini 2000, 2003).

It is possible that the degree of political fragmentation is a direct consequence of the details of the electoral rule, and the opportunities these provide for strategic voting in the electorate. However, the formal spatial electoral model has not, in our view, been able to offer a plausible account of this relationship. Indeed, as discussed in Schofield (2007a), the extensive literature on formal "deterministic" or "stochastic" vote models tend to suggest that all parties should adopt vote-maximizing positions at the center of the electoral distribution.[2] Such models assume an underlying symmetry in the motivations and dispositions of party leaders, and as a result they are unable to account for the extreme heterogeneity of political configurations observed by Benoit and Laver (2006), for example, in their analysis of party positions in European polities.

In this article, we offer a formal stochastic model of elections that emphasizes the importance of the idea of *valence*. In the standard spatial model, only candidate *positions* matter to voters. However, as Stokes (1963, 1992) has emphasized, the non-policy evaluations, or *valences*, of candidates by the electorate are equally important. Stokes (1963, p. 373) used the term *valence issues* to refer to those that "involve the linking of the parties with some condition that is positively or negatively valued by the electorate." As he observes, "in American presidential elections... it is remarkable how many valence issues have held the center of the stage." We use the stochastic valence electoral model to compare party strategies in Israel, where the electoral system is based on proportional rule, with that of the United States, where the electoral system is highly majoritarian.

We argue that, in the United States, the differences between the valences of the two major presidential candidates are insufficient to force them to adopt divergent positions. Instead, the logic of vote maximization should force convergence to the electoral origin.[3] Since candidates do not converge, we propose a model where activist groups provide the resources that are critical for political success. However, these activists require the candidates to adopt divergent positions in return for political support. In

[1] See Laakso and Taagepera (1979) for a formal definition of fragmentation.

[2] See Downs (1957), Riker and Ordeshook (1973), and McKelvey and Patty (2006).

[3] The *electoral origin* is simply that point which is at the mean on all dimensions of the distribution of voter preferred points.

essence, a small number of influential activist groups induce the polarization that has been noted in the U.S.polity.[4]

In contrast, in Israel, as in other polities based on proportional representation, there are significant valence differences between the parties. By themselves, these valence differences are sufficient to force the parties to diverge. Activists may well influence the parties, but this influence appears much less significant than in the United States. More importantly, since small parties may aspire to membership of coalition government, their activist coalitions have no incentive to coalesce. Thus, the relatively fragmented party structure is maintained. In the United States, the greater intensity of competition for activist support means that small activist groups, if they are to have any impact, must join one or other of the major party activist groups. This forces coalescence of the activist groups. We argue that the relationship between political fragmentation and the nature of the electoral system, noted by Riker (1953) and Duverger (1954), is the result of this logic of activist support.

There is a long tradition of argument that interest groups induce policy choices that are non-optimal for the society (Olson 1965; Keefer 2004; Acemoglu and Robinson 2006), but this article is the first, we believe, to develop a formal and empirical vote model that indicates how to estimate the impact of activist groups on the policy stances of political leaders in polities with different electoral systems.[5]

Electoral models involving this notion of valence has formed the basis for recent extensive analyses of British, Canadian and US electoral response by Clarke et al. (2009a) and Clarke et al. (2005, 2009b).[6]

For Britain, they argue that electoral responses

> were a reflection largely of [the] changing perceptions of the decision-making competence of the main political parties and their leaders. At any point in time, [the] preferences were strongly influenced by their perceptions of the capacity of the rival parties—the putative alternative governments of the day—to solve the major policy problems facing the country.

These works have shown that valence, as measured by the perceptions of the character traits of the candidates, or of party leaders, is a key element of election.

Here, we extend the usual spatial model by incorporating these electoral perceptions of candidate character traits in a stochastic model of the 2008 US election. Our purpose is different from the empirical work by Clarke et al. (2009b). Instead of focusing on the electoral response to candidates, we use this extended model to determine the response of candidates to the electoral situation: that is we compute the equilibrium candidate positions in the context of the chosen model.

[4] It is of interest that Bernhardt et al. (2009) use a one-dimensional vote model to argue that polarization of party positions, if not too extreme, is welfare enhancing because of the choice that it provides for the electorate.

[5] An early article by Enelow and Hinich (1982) presented a formal model involving valence, though they used the term "non-spatial characteristic" rather than valence, but their model was not related to the impact of interest groups.

[6] See Schofield et al. (2010c) for similar studies of Britain and Canada.

In the stochastic model, a voter's perception of each candidate's traits has a very significant impact on the probability that the voter chooses one candidate or the other. Since these voters are characterized by different preferred policy positions, a candidate's optimal policy position should be a function of the distribution of these correlated voter positions/perceptions.[7] Our simulation of the combined model, based on both position and valence, allows us to estimate what we call *Local Nash equilibria* (LNE) to the vote-maximizing game, as calibrated by the empirical model with the greatest statistical significance.[8]

We found, by simulation of the stochastic model of the US presidential election in 2008 involving perception of candidate traits, that there was a unique local Nash equilibrium very close to the electoral origin.

In contrast, simulation of the multiparty stochastic model of Israel in 1996 found that the local Nash equilibrium were characterized by divergence away from the electoral center.

In Sect. 2 of this article, we briefly sketch our argument about the fundamental differences in these two polities. Section 3 introduces the notion of the *convergence coefficient* which can be used to determine whether candidates or parties should converge to the electoral center when they attempt to maximize vote share. Sections 4 and 5 present the empirical analyses of these two polities. The concluding section emphasizes the differences between the majoritarian electoral system of the United States and the proportional system in Israel, as well as other fragmented polities such as Poland and Turkey, that are highlighted by the formal and empirical analyses. This conclusion mentions other work that has estimated the convergence coefficients for various polities, and suggests these coefficients are related to the degree of political fragmentation in these political systems. Formal definitions for the model are given in Appendix 1. See also the Appendix to Schofield et al. (2011) in this issue, for the definitions of the equilibrium concepts used here.

2 Comparison of the United States and Israel

To provide a brief sketch of the results on the 2008 U.S. election, consider Fig. 1, which presents the distribution of voter preferred points, as obtained from factor analysis of survey responses from the American National Election Study (ANES 2008). This survey allows us to estimate each respondent's *ideal point,* as a way of representing that citizen's responses. The estimated distribution of such points is the *electoral distribution.* We shall refer to the space in which the electoral distribution is embedded as the *factor space*. In formal spatial models, this space is usually known as the policy space. We use the term *factor space* to remind the reader that the basis for the construction of this space is the factor analysis of the survey. In particular, the electoral distribution is directly estimated from the factor analysis.

[7] Just as in Clarke et al. (2009b), we use factor analysis of the survey responses to obtain a two-dimensional representation of the voter preferred positions.

[8] Erikson and Romero (1990) used this procedure in an empirical model of the 1988 US Presidential election. However, they estimated local equilibrium positions on a number of separate policy dimensions, rather than in a single multidimensional policy space.

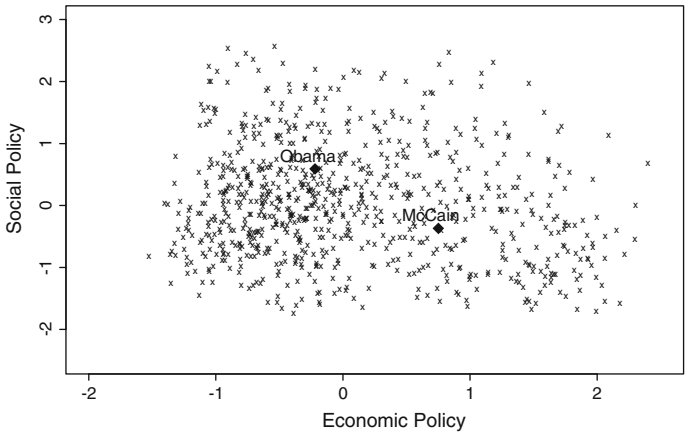

Fig. 1 Distribution of voter ideal points and candidate position in 2008

For any polity, we refer to that point in the factor space which is at the mean in all dimensions of the electoral distribution as *the electoral origin*. The electoral distribution is characterized by variances of the distribution on each axis, as well as the covariance between the various axes.[9]

Figure 1 represents the voter locations in a two-dimensional factor space, so the electoral distribution is given by a symmetric 2 by 2 *covariance matrix*. The x-axis involves economic or redistributive issues, and on this axis the variance of the electoral distribution is 0.80. The y-axis involves social issues, and the variance on this axis is 0.83.[10]

In an empirical model presented below, we estimate Obama's position in the factor space to be $z_{\text{Obama}} = (-0.22, +0.75)$, a distance of 0.77 units from the electoral origin.[11] McCain's position in the same factor space was estimated to be $z_{\text{McCain}} = (0.59, -0.37)$, a similar distance of 0.69 from the origin, but in a different quadrant of the policy space.[12]

Moreover, the average Democrat voter position was $z^{\text{vote}}_{\text{DEM}} = (-0.17, +0.36)$[13], while the average Republican voter position was $z^{\text{vote}}_{\text{REP}} = (0.72, -0.56)$.[14] The survey also gave information on activists of the parties (that is, individuals who contributed money to the parties). Figure 2 shows the distribution of activist positions, clearly very

[9] The factor analysis in w dimensions thus gives a symmetric w by w electoral covariance matrix.

[10] The covariance between the two axes is -0.127.

[11] Details of the estimation method are given below. These estimates are the average perceptions of the voters about the candidate positions.

[12] The distance in the factor space between the candidates was 1.38. Since the total electoral variance was 1.63, we term $\sqrt{1.63} = 1.27$ the electoral standard deviation (esd). The two candidates are thus located about 1.08 esd apart.

[13] The variances of the distribution of Democrat partisans ideal points on the two axes were (0.72, 0.75), giving standard errors of the means of (0.029, 0.03).

[14] The variances of the Republican partisans' ideal points were (0.38, 0.47), so the standard errors of these two means were (0.027, 0.03).

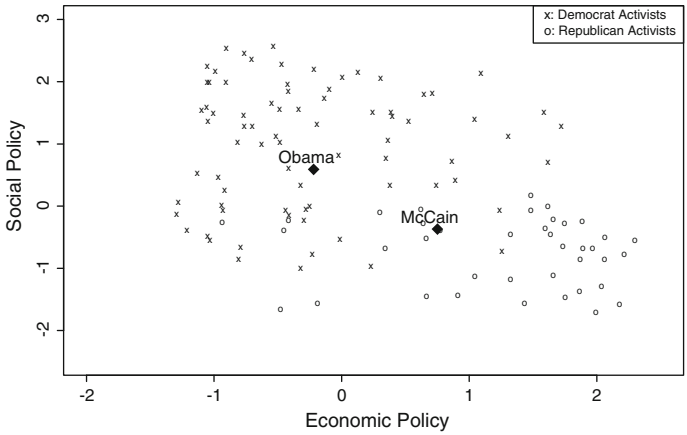

Fig. 2 Distribution of activist ideal points and candidate positions

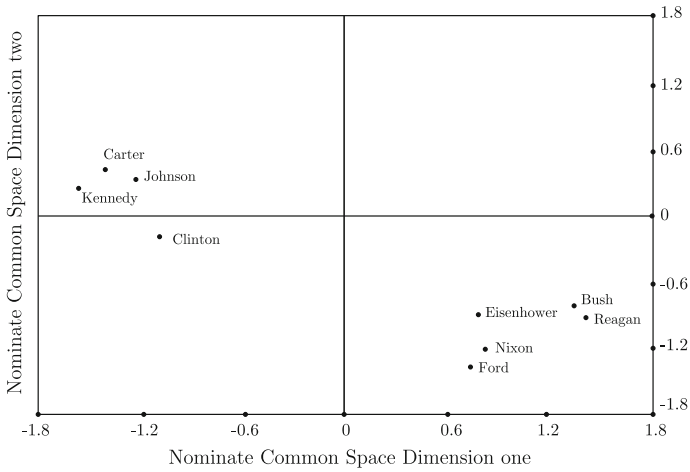

Fig. 3 Presidents' positions

different from the voter distribution given in Fig. 1. The mean activist positions for the two parties were estimated to be $z^{\text{act}}_{\text{DEM}} = (-0.2, +1.14)$ for the Democrats and $z^{\text{act}}_{\text{REP}} = (1.4, -0.82)$ for the Republicans.

The point to note about these estimates is that the Obama and McCain positions appear significantly different, and some distance from the electoral origin. Obama is located midway on the social axis between Democrat voters and activists, while McCain is more centrist than both Republican voters and activists on the two axes.

The positioning of Democrat presidential candidates in the upper left of the policy space, and Republican candidates in the lower right, has been noted in other empirical work, as suggested by Fig. 3.[15]

Related work (Schofield et al. 2003) has modeled the presidential elections of 1964 and 1980, and argued that such a configuration is a structural characteristic of the US

[15] This figure is taken from Schofield (2002). See other related work by Poole and Rosenthal (1984).

polity. Indeed, Miller and Schofield (2003, 2008) and Schofield and Milller (2007) suggest that the US polity is fundamentally two dimensional. They use this feature as the basis for a model of political realignment (Sundquist 1973), as one dimension becomes more important than the other. One purpose of this article is to provide a formal account of what seems to be opposed policy positions offered by the presidential candidates in elections in the United States.

We argue here that the shifts in candidate positions for the two parties over time are insufficient to account for the quite substantial changes in electoral support that occur. Instead our analysis suggests that electoral shifts are primarily the result of changes in the perceptions by the electorate of the candidates. These electoral transformation, in turn, are the consequence of the changing resources available to the candidates. Finally, these are due to the shifting coalition structures among the potential activist groups in the polity.

If this suggestion is correct, then it implies that formal models of elections based on position and valence alone are quite inadequate to account for candidate policy proposals. The following remarks and inferences suggest that any formal model of US elections must expliccitly include activist groups:

(i) The equilibrium analysis of spatial models of US presidential elections indicates that candidates should converge to positions very close to the electoral origin in order to maximize vote shares.[16]
(ii) However, estimates of candidate positions indicate that they are located in opposed quadrants of the policy space.
(iii) The incompatibility of the equilibrium locations and the estimated positions can be explained by the influence of activists in US elections.
(iv) Activist influence has increased over time.[17] The recent Supreme Court decision, *Citizens United v. Federal Election Commission*, on Thursday, January 21, 2010, has removed limits on campaign contributions and will likely increase the importance of activist contributions. Dworkin (2010) has called this decision "an unprincipled political act with terrible consequences for the nation." Obama, shortly after, in his State of the Union address declared

> the Supreme Court reversed a century of law [which] I believe will open the floodgates for special interests... to spend without limit in our elections.

(v) Although the distribution of voter positions may not change dramatically, so the distribution cannot be seen to be polarized, the positions of candidates for office have become more polarized.[18] The system of primaries in US elections is likely to further enhance the influence of activists on candidates.
(vi) Because of this polarization of candidate positions, a shift in the party controlling the presidency will have significant policy implications.

[16] This result has also been found by Enelow and Hinich (1989) for earlier U.S. elections.

[17] Indeed, Herrera et al. (2008) observe that spending by parties in federal campaigns went from 58 million dollars in 1976 to over 1 billion in 2004 in nominal terms.

[18] See for example McCarty et al. (2006), and Fiorina et al. (2005).

(vii) The same argument holds for members of Congress, and we would expect activist influence to increase the degree of polarization in Congress.[19]

(viii) The influence of activists in the strongly majoritarian polity of the United States is the fundamental cause of these policy shifts.

(ix) Because the winner of the presidential election will generally be located some distance from the electoral center, the policies supported by the President need not be supported by an electoral majority. This phenomenon can be seen with regard to the reform of health care, supported by Obama in 2009/2010. This policy is certainly located in the upper left quadrant of the policy space. As of January 22, 2010, about 39% of the electorate supported the health plan while 55% did not.[20]

(x) In between elections, diametrically different policy positions will be aggressively supported by opposed lobbying groups. For example, in 2009, health care, pharmaceutical and insurance lobbyists spent approximately $650 million on lobbying itself, and about $210 million on media advertising.[21] The oil and gas industry spent about $560 million.[22]

(xi) Actual policy choices will depend on complex bargaining between the President and Congress. As the health care issue illustrates, the supramajoritarian voting rule in the Senate will tend to favor the status quo.[23]

(xii) Activist-induced policy preferences in Congress is extremely heterogenous.[24] This, together with a non-centrist presidential policy position, can result in so-called "gridlock." [25]

(xiii) During 2010, there is an increasing perception in the electorate that Congress has become dysfunctional because of "strident partisanship, unyielding ideology and a corrosive system of campaign financing." [26]. For example, the CNN/Opinion Research Corp. poll, conducted on February 12–15,2010, with 1,023 respondents, found that 86% thought government was "broken." Of these, however, 81% felt it could be "fixed." In fact, gridlock can be overcome, as illustrated by the 62–30 vote in the Senate on February 22 to implement a multi-billion "jobs

[19] Conflict between the parties over health care in 2009 and 2010 is just one illustration of this phenomenon.

[20] The surprise victory by Republican Scott Brown over Democrat Martha Coakley in the special election for the Senate seat for Massachusetts, on January 19, 2010, may be indicative of this electoral response, as well as "Tea Party activism."

[21] The pharmaceutical industry was a strong supporter of reform of health care, because of an agreement with Obama to protect the industry's profits.

[22] Tomasky (2010) gives a figure of $3.47 billion for spending by lobbyists in the non election year of 2009, citing data from the Center for Responsive Politics.

[23] Scott Brown's victory in Massachusetts in January, 2010, deprived the Democrats of the 60 seat majority required to overcome the filibuster and push through legislation on health care and other policy issues such as financial reform.

[24] Work by Jeong et al. (2010) estimated the policy positions of US senators with regard to the 2006 immigration reform act and found the Republican senator positions to be very heterogenous, but all clearly in the lower right hand quadrant of the policy space.

[25] This of course contradicts the argument by Bernhardt et al. (2009) that divergence is welfare enhancing.

[26] Indeed, when Evan Bayh, Senator from Indiana, announced in February 2010, he would retire, these were the reasons he gave (Bayh 2010).

creation" program. Gridlock over health care was also broken on March 25, after strenuous efforts by President Barack Obama and House speaker, Nancy Pelosi, when the House voted 220–207 for the health care bill. Republicans had voted unanimously against the legislation, joined by 33 dissident Democrats. The President had signed a draft of the bill, the "Patient Protection and Affordable Care Act" on March 23, and the Senate passed the bill by simple majority of 56 to 43, as required for reconciliation. In July 15, the Senate voted 60–39 for the bill for Reform of Financial Regulation.[27] As of July 2010, there remain four major bills to put through Congress: A Deficit Reduction Act, an Energy Independence and Climate Change Act, an Expanded Trade and Export Act, and a Comprehensive Immigration Act. If these prove impossible to enact because of Republican opposition, the electorate may blame the GOP.

The success of the health care legislation and of reform of financial regulation, together with the signing of the new START arms reduction treaty in Prague on April 8 by Presidents Medvedev and Obama, has certainly increased Obama's international prestige.

However, given the uncertainty surrounding policy choice in the Legislature, it is hardly surprising that voters in the United States doubt that government can be effective. Part of the problem would appear to be the degree of political polarization resulting from the power of interest groups located in the opposed quadrants of the policy space.

We contrast these observations on the difficulties facing the U.S. government with inferences about a polity based on an electoral system using proportional representation, such as Israel. We argue that in such a polity, activist influence is weaker, and policy shifts between different governments will be significantly smaller.

Figure 4 presents a smoothed estimate of the voter distribution, as well as the estimated party positions in Israel in 1996.[28]

The x-axis is designated security, and is defined in terms of attitudes to the PLO. The factor model was normalized with respect to this factor, so the electoral distribution on this axis had a variance of 1.0. The y-axis involves religious attitudes and on this axis the variance was 0.732. Note that in Fig. 4, the estimated positions of the two major parties, Labor and Likud, in the factor space are $(-0.8, -0.3)$ and $(0.4, 0.2)$.[29] In the discussion of Israel elections in Sect. 3 of this article, we argue that the policy positions of coalition governments will depend on whether there is a core party, or one that is centrally located, and large enough to dominate coalition bargaining.

In the election of 2006, a centrist party, Kadima, initially under the leadership of Ariel Sharon, was able to position itself at the electoral origin, and form a coalition

[27] This complex bill was 2,300 pages long. Russ Feingold, a Democrat, voted against the bill, because it was not strong enough. Three moderate New England Republicans, Snowe and Collins of Maine, and Scott Brown of Massachusetts, voted for the bill.

[28] The party positions were obtained from expert estimates, and the voter distribution obtained from a survey by Arian and Shamir (1999). See Schofield and Sened (2006).

[29] The distance in the factor space between these two parties was 1.3. Since the total electoral variance was 1.732, the electoral standard deviation (esd) was $\sqrt{1.732} = 1.32$. The two parties were thus located 0.98 esd apart.

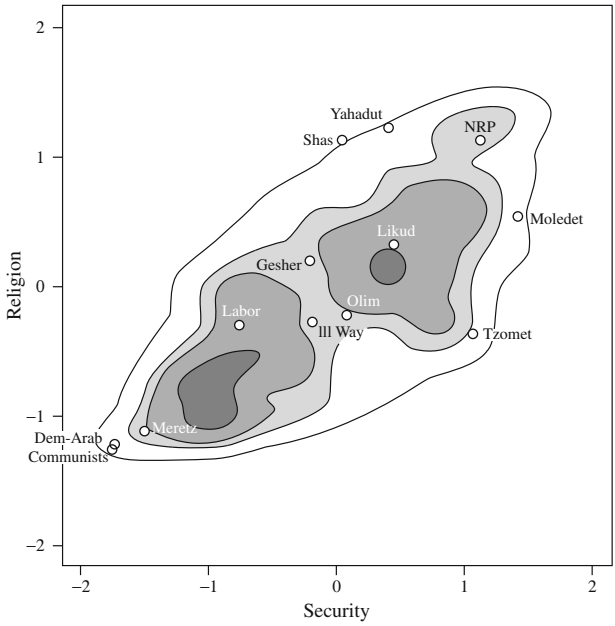

Fig. 4 Party positions and voter distribution in Israel in 1996

with other smaller parties. In the recent election of February 2009, the smaller parties, estimated to be located in the upper right quadrant of the factor space gained some further electoral support. Netanyahu, the leader of Likud, then constructed a winning coalition with the support of Labor and Israel Beitenu.

Our analysis suggests the following:

(i) In the US polity with a pronounced majoritarian electoral system, the two parties, or their candidates, adopt divergent positions that are symmetrically opposed.[30]

(ii) In Israel with a very proportional electoral system (but with a 2% cut-off) the major parties typically adopt positions relatively close to the electoral origin, while the smaller parties occupy quite divergent positions. The result of bargaining between a major party and smaller parties will tend to result in centrist outcomes.[31] The difficulty facing such a polity, especially when the political configuration is fragmented, is that agreement between the parties may be difficult to attain. We comment in the conclusion on other empirical work that has a bearing on this observation

Thus, a very rough interpretation of the significance of policy change, when normalized with respect to a natural characteristic of the electoral distribution, suggests

[30] In particular, an estimate of policy changes from one administration to another will be of the order of one esd.

[31] Policy changes from one government to another can be expected to be of the order of at most one esd. If the centrist Kadima party is included in the coalition, then policy switches will be less.

that volatility, in terms of change in the political outcome, is a function of the nature of the electoral system. In this article, we present a general electoral model to provide a formal account of this variation in political configuration.

3 Valence in the electoral model

The model we present is intended to cover both elections between *candidates* for office, such as president, from differing parties, as well as elections involving leaders of different parties. We shall use the term political leaders for both cases.

We assume first that leaders adopt positions to maximize their vote share in the context of a stochastic electoral model. Each leader, j is characterized by an *intrinsic (or exogenous) valence*, denoted by Λ_j.[32] This can be estimated as the intercept term in the pure spatial model, and can be interpreted as the average electorally perceived quality of each leader. Exogenous valence is estimated with respect to a baseline leader, so in the 2008 US election model we set $\Lambda_{\text{Obama}} = 0$, and estimate Λ_{McCain}. We use the notion of a *convergence* coefficient, presented in Schofield (2007a) to show that the difference between Λ_{Obama} and Λ_{McCain} is sufficiently small so that the equilibrium of the pure spatial model is one where the two candidates adopt identical positions at the electoral origin.

The notion of exogenous valence is then extended to include heterogenous, *sociodemographic* valence terms. These sociodemographic valences cause each party to seek out any group in the electorate which has a propensity to favor that leader independently of the leader's declared policy position. In our simulation of the joint sociodemographic model, we found that even with these sociodemographic valences the equilibrium positions in the 2008 US election were not perturbed from the electoral origin.

We then extended the model with exogenous and sociodemographic valences, by including the electoral perceptions of the candidate traits. These candidate traits add considerably to the significance of the valence model. Since these perceptions are individually based, and, therefore, determined by voter position, we can use simulation techniques to compute the equilibrium positions implied by the full valence model. We found that the local equilibrium of the full model with traits was one where candidates adopted positions slightly different positions at $z_{\text{Obama}} = (0.10, -0.07)$ and $z_{\text{McCain}} = (0.13, -0.12)$.

To account for the disparity between the simulated local equilibrium positions of the candidates and their estimated positions, we included *activist* valence in the formal model. We may regard activist valence as a kind of *endogenous* valence since it is the consequence of bargaining between party and activists.[33] A party that has in the past tended to adopt a policy position that favors a particular group may also benefit from the provision of resources, such as money and time, from activists belonging the group. The possibility of obtaining such resources, to enhance the electoral success of

[32] See for example the formal model in Serra (2010) and an empirical model of the 2008 election in Jessee (2010).

[33] See the models in Grossman and Helpman (1994, 1996, 2001) and Baron (1994).

a candidate, will exert a centrifugal force, drawing the candidate closer to the group. The marginal calculation by each candidate can be interpreted as a *balance condition*, which incorporates all the valence terms. The term balance is used because it involves equating the opposed centripetal attraction of the electoral center and the centrifugal activist force.

The empirical work presented here for the United States suggests that the endogenous valence terms for the candidates are very similar, so that, for the pure spatial model, the centripetal electoral force should dominate. Because the candidates diverge from the center, we infer that the centrifugal activist valence terms are very significant. We argue that since the electoral system is highly majoritarian, potential activist groups will tend to coalesce, so as to increase their influence on their chosen party. As we noted in the Sect. 1, increasing campaign expenditure by parties in the United States reflects the increase of activist influence.

Under proportional representation, as in Israel, major parties will be characterized by high exogenous valence, in comparison to the more peripheral, low valence parties. The balance condition will cause the high valence parties to pay less heed to sociodemographic valence, and they will be less dependent on activist valence. As a consequence, they will tend to be located near the electoral center. In contrast, small parties will tend to represent the interests of very specific groups in the society. Their exogenous valence, which is a measure of the perceived quality of the party leader in the whole electorate, will be very low. The centripetal electoral effect on small parties will be dominated by the centrifugal effect, and they will tend to adopt positions far from the electoral center.

In the application of the model to Israel in 1996, we show the intrinsic valence of such small parties is indeed very low. This is a general phenomenon, which holds true for models of elections in Poland and Turkey, as well as Israel.[34] The vote share of such parties will be very dependent on sociodemographic valence, as well as on the support of specific activist groups. Thus, the centrifugal force on such parties will be further enhanced.

As illustrated by the case of Israel, relatively small parties will be pivotal for the formation of coalition government. Such parties may expect to gain office, and bring important policy rewards to their activist supporters. Thus, these specific sociodemographic groups will, in expectation, gain from the support they provide to these parties. Both parties and activist groups will be motivated to maintain this mutually beneficial arrangement, and the leaders of such groups will have little motivation to coalesce with other groups. The high level of political fragmentation will be maintained unless a dominant center party can attract some of the relatively radical activists.

To develop this argument, we first consider the United States. Section 4.1 presents a standard binomial logit model for the 2008 presidential election. This model does not involve candidate positions, so we then develop a spatial mixed logit model that does involve candidate positions. Section 4.2 obtains the conditions that characterize the local Nash equilibria in the models with exogenous and sociodemographic valences

[34] See Schofield et al. (2010b) for Poland. Schofield et al. (2011) for Turkey.

as well as voter perceptions of the candidate traits. We then use these formal models to estimate the activist influences for this election on the candidates.

In Sect. 5, we perform the same analysis for the 1996 election in Israel, making use of the earlier analysis of Schofield and Sened (2006). Although we do have have available the voter perceptions of character traits of the party leaders, we find that the estimates of exogenous valence obtained from a joint spatial model are sufficiently different to account for the fact that parties do not converge to the electoral origin. We briefly comment on the recent election of 2009, and suggest that the loss of dominance by the centrist party Kadima was due to the decrease of sociodemographic and activist valence by the Labor party.

4 The election of 2008 in the United States

4.1 Empirical analysis

The 2008 American National Election Study (ANES 2008) introduced many new questions on political issues in addition to the existing set. Assignment of respondents into the "new" or "old" set was random, with 1,059 respondents assigned to the "new" condition and having completed the follow-up post-election interview.

The post-election interviews asked respondents whom they voted for, if at all. Since we use a conditional logit model, which requires data for both respondents and candidates (which we only have for the major party candidates) we removed observations where respondents claimed to have voted for a presidential candidate other than McCain or Obama, or not to have voted at all.

To create the two-dimensional policy space, 23 survey items were selected to broadly represent the economic and social policy dimensions of American political ideology (see Appendix 2 for question wording). There were multiple questions for abortion, gay and African American issues. These three sets of questions were combined using factor analysis to give three separate scales.

Factor analysis of the survey was then used to obtain measures of individual locations in the policy space (see Table 1 for factor loadings).

The ANES also includes questions on seven qualities or traits associated with Obama and McCain, asking respondents about the traits of the candidates, including the terms "moral, caring, knowledgable, strong, dishonest, intelligent, out of touch." Factor analysis of these questions gave two factors, and the resulting factor scores were used as estimates of voter perceptions of the candidate's personal traits.

To calculate the presidential candidate positions, we took advantage of new survey questions which asked respondents to locate the positions of Obama and McCain on seven distinct issues.

These seven questions (government spending, universal health care, citizenship for immigrants, abortion when non-fatal, abortion when gender incorrect, aid to blacks, and liberal-conservative) were otherwise worded the same as the corresponding items from the 23 policy issue questions.

To find McCain's ideal point, we simply took the average response for each of his seven candidate location questions. We then repeated the process using Obama's can-

Table 1 Factor loadings for economic and social policy

Question	Economic policy	Social policy
1. Government services	0.53	0.12
2. Universal health care	0.51	0.22
4. Government bigger	0.50	0.14
5. Government or market	0.56	
9. Welfare spending	0.24	
6. Less government	0.65	
7. Equality	0.14	0.37
8. Tax companies	0.28	0.10
12. Abortion scale		0.55
11. Immigrant scale	0.12	0.25
13–16. Gay scale		0.60
17. Traditional values		0.53
18. Gun access	0.36	
19–22. Afr. Amer. scale	0.14	0.45
23. Liberal v conservative	0.30	0.60
Eigenvalue	1.93	1.83

Table 2 Descriptive data

	Econ	Policy		Social	Policy		*n*
	Mean	s.e.	95% C.I	Mean	s.e	95% C.I	
Activists							
Democrats	−0.20	0.09	[−0.38, −0.02]	1.14	0.11	[0.92,1.37]	80
Republicans	1.41	0.13	[1.66, 1.16]	−0.82	0.09	[−0.99, −0.65]	40
Non-activists							
Democrats	−0.17	0.03	[−0.24, −0.11]	0.36	0.04	[0.29,0.44]	449
Republicans	0.72	0.06	[0.60,0.84]	−0.56	0.05	[−0.65, −0.46]	219
							788

didate location questions. See Tables 2 and 3 for the descriptive data and the estimated positions of the two candidates.

Respondents were coded as activists if they claimed to have donated money to a candidate or party. The survey data gave information on whether the respondent was African American, Hispanic, female, working class, from the South. Additional data on age, number of years of education, and level of income were used to construct eight different sociodemographic variables.

Figure 1, above, gave the voter distribution, while Fig. 2 gave the activist distribution.

Table 3 Obama and McCain perceived positions

Question	Obama	McCain
Estimated position on economic policy	−0.22	0.59
Estimated position on social policy	0.75	−0.37

As noted above, the positions of the major presidential candidates, McCain and Obama, in 2008 were estimated using the perceptions of the sampled individuals.

These positions were:

$z_{\text{Obama}} = (x_{\text{Obama}}, y_{\text{Obama}}) = (-0.22, 0.75)$,

$z_{\text{McCain}} = (x_{\text{McCain}}, y_{\text{McCain}}) = (0.59, -0.37)$.

We now use the formal model to analyze this election.

4.2 Estimation of political equilibria

Obama's victory on November 4, 2008 suggests that it was the result of an overall shift in the relative valences of the Democrat and Republican candidates from the election of 2004. In fact, since Obama took 52.3% of the vote, a simple estimate of the probability, ρ_{obama}, of voting for Obama is given by

$$\rho_{\text{Obama}} = [0.523] = \frac{\exp[\Lambda_{\text{Obama}}]}{1 + \exp[\Lambda_{\text{Obama}}]}$$

It immediately follows that an estimate of Λ_{Obama} relative to Λ_{McCain} is given by

$$\begin{aligned} \log_e \left[\frac{0.523}{0.477}\right] &= \log_e[1.096] \\ &\simeq 0.09. \end{aligned}$$

In fact there were differential shifts in different regions of the country. In a region of the country from West Virginia through Tennessee, Arkansas, and Oklahoma, there was a shift of 20% in the increase in the republican vote, suggesting a change of about 0.6 in McCain's valence advantage.

To model this election, we first constructed a *pure positional binomial logit model*.

According to this positional model, a voter i, with preferred position (x_i, y_i) is estimated to vote Republican with probability

$$\rho_{\text{rep}} = \frac{\exp(\Lambda_r + bx_i + cy_i)}{1 + \exp(\Lambda_r + bx_i + cy_i)}. \tag{1}$$

We estimated these coefficients to be $(\Lambda_r, b, c) = (-0.74, 1.49, -1.80)$, with standard errors $(0.11, 0.13, 0.15)$, respectively. All were significant at the 0.001 level.

This cleavage line derived from this the equation gives the locus of voting with equal probability for one or other of the candidates. This cleavage line is given by the equation

$$y = 0.82x - 0.4. \tag{2}$$

This cleavage line misses the origin, and goes through the point (0, −0.4), indicating the valence advantage of Obama. The coefficient Λ_r is a measure of the (negative) relative valance of McCain with respect to Obama for this positional model. This cleavage line is similar to those obtained by Schofield et al. (2003) for the presidential elections of 1964 and 1980. (One difference between this earlier estimate and the one presented here was that in 1980 they found that Reagan had a valence advantage over Carter.)

These positional models do not explicitly involve the candidate positions, and so cannot be used to determine political equilibria. We now discuss the spatial models, presented in Table 4.

The electoral covariance matrix for the sample is given by

$$\nabla_0 = \begin{bmatrix} 0.80 & -0.127 \\ -0.127 & 0.83 \end{bmatrix}.$$

The principal component of the electoral distribution is given by the vector (1.0, −1.8) with variance 1.02, while the minor component is given by the orthogonal eigenvector (1.8, 1.0) with variance 0.61.

Model (1) in Table 4 shows the coefficients in 2008 for the pure spatial model to be

$$(\Lambda_{\text{Obama}}, \Lambda_{\text{McCain}}, \beta) = (0, -0.84, 0.85).$$

Table 4 indicates, the loglikelihood, Akaike information criterion (AIC), and Bayesian information criterion (BIC) are all quite acceptable, and all coefficients are significant with $P < 0.01$.

Note that these parameters are estimated when the candidates are located at the estimated positions. Again, Λ_{McCain} is the relative negative exogenous valence of McCain, with respect to Obama, according to the model $\mathbb{M}(\boldsymbol{\Lambda}, \beta)$. We assume that the parameters of the model remain close to these values as we modify the candidates positions in order to determine the equilibria of the model.

According to the model $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, the probability that a voter chooses McCain, when the McCain and Obama positions are at the electoral origin, $\mathbf{z}_0 = ((0, 0), (0, 0))$ is

$\rho_{\text{McCain}} = [1 + \exp(0.84)]^{-1} = [1 + 2.31)]^{-1} = 0.3.$

Then $\beta(1 - 2\rho_{\text{McCain}}) = 0.85 \times 0.4 = 0.34.$

The characteristic matrix (essentially the Hessian of McCain's vote function at $\mathbf{z}_0$ is:

$$\begin{aligned} C_{\text{McCain}} &= [2\beta(1 - 2\rho_{\text{McCain}})\nabla_0 = [2 \times 0.34 \times \nabla_0] - I = (0.68)\nabla_0 - I \\ &= (0.68)\begin{bmatrix} 0.8 & -0.127 \\ -0.127 & 0.83 \end{bmatrix} - I = \begin{bmatrix} 0.54 & -0.086 \\ -0.086 & 0.56 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.46 & -0.086 \\ -0.086 & -0.44 \end{bmatrix} \end{aligned}$$

Table 4 β-Spatial conditional logit models for USA 2008

	Spatial	Sp. & Traits	Sp. & Demog	Full
McCain valence Λ	−0.84*** (0.11)	−1.08*** (0.13)	−2.60** (0.93)	−3.58*** (1.05)
Distance β	0.85*** (0.06)	0.78*** (0.07)	0.86*** (0.07)	0.83*** (0.08)
McCain traits		1.30*** (0.17)		1.36*** (0.19)
Obama traits		−1.02*** (0.15)		−1.16*** (0.18)
Age			−0.01 (0.01)	−0.01 (0.01)
Female			0.29 (0.23)	0.44 (0.26)
African American			−4.16*** (1.10)	−3.79*** (1.23)
Hispanic			−0.55 (0.41)	−0.23 (0.45)
Education			0.15* (0.06)	0.22*** (0.06)
Income			0.03 (0.02)	0.01 (0.02)
Working class			−0.54* (0.24)	−0.70** (0.27)
South			0.36 (0.24)	−0.02 (0.27)
Observations	788			
Log likelihood (LL)	−298.63	−243.14	−250.25	−206.88
AIC	601.27	494.28	520.50	437.77
BIC	610.59	512.92	567.11	493.69

Standard errors in parentheses. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$
Vote for Obama is the baseline outcome

The "convergence coefficient" is

$$c = 2\beta(1 - 2\rho_{\text{McCain}})trace\nabla_0 = 2(0.34)(1.63) = 1.1.$$

Schofield (2007a) shows that the necessary condition for convergence to $\mathbf{z}_0$ is that $c < 1$. Note that c is dimensionless, and therefore independent of the units of measurement.

The estimate for c exceeds this critical value for convergence. However, the determinant of C_{McCain} is positive and trace is negative so both the eigenvalues of C_{McCain} are negative. Standard results of calculus show that the origin is a maximum of McCain's vote share function. Simulation of the pure spatial model confirmed that $\mathbf{z}_0$ was an LNE. Indeed it was shown to be a Pure Strategy Nash equilibrium (PNE).

We also considered a spatial model where the x and y axes had different coefficients, $\beta_1 = 0.8$, $\beta_2 = 0.92$. The analysis showed the the Hessian for this case had negative eigenvalues, so again $\mathbf{z}_0$ is a LNE. This model is essentially the same as the model with a single β.

We now turn to the models with traits and sociodemographics.

Table 4, above, gave the various spatial models with these additional valences.

We found that the loglikelihoods of the pure sociodemographic model and pure traits models to be to be −427 and −356, respectively. Comparison of the loglikelihoods for the pure spatial model and the model with traits, as given in Table 4 shows that the perception of character traits is important for the statistical significance

Table 5 Comparison of LL for US spatial models in 2008

	JST	ST	S	T
JST	na	36	92	150
ST	−7	na	55	114
S	−92	−55	na	58
T	−150	−114	−58	na

JST Joint spatial with traits, *ST* spatial with traits, *S* pure spatial, *T* Pure traits

of the model. As Table 5 shows, the difference in the loglikelihoods of the spatial model with traits and the pure traits model is $-243 + 357 = 114$, while the difference between the full spatial model with traits and sociodemographics against the traits model is $-206 + 357 = 150$.

Simulation of the full spatial model with traits and sociodemographics showed that the LNE (and PNE) was one where the candidates adopted the positions $z_{\text{Obama}} = (+0.10, -0.07)$ and $z_{\text{McCain}} = (+0.13, -0.12)$.

We can, therefore, write

$$\mathbf{z}^{\text{el}} = (z^{\text{el}}_{\text{Obama}}, z^{\text{el}}_{\text{McCain}}) = ((+0.10, -0.07), (+0.13, -0.12))$$

since the joint model with traits has no activist valence terms.

This equilibrium is only a slight perturbation from the joint origin. We can infer that though the traits add to the statistical significance of the stochastic model they do not significantly affect the equilibrium. Analysis of the relationship between perceptions of candidate traits and vote choice showed that there were weak correlations and these had only a slight effect on the strong convergence induced by the electoral pull.

The results of the Appendix 1 show that $\mathbf{z}^{\text{el}}$ can be interpreted as the vector of "weighted electoral means" in a full model with activists. Assuming that the estimated candidate positions, $\mathbf{z}^*$, are in equilibrium with respect to the activist model, then by the balance condition, we obtain:

$$\begin{aligned}\mathbf{z}^* - \mathbf{z}^{\text{el}} &= \begin{bmatrix} & \text{McCain} & \text{Obama} \\ x & 0.59 & -0.22 \\ y & -0.37 & +0.75 \end{bmatrix} - \begin{bmatrix} & \text{McCain} & \text{Obama} \\ x & +0.13 & +0.10 \\ y & -0.12 & -0.07 \end{bmatrix} \\ &= \frac{1}{2\beta}\frac{\mathrm{d}\mu}{\mathrm{d}\mathbf{z}}(\mathbf{z}) = \begin{bmatrix} & \text{McCain} & \text{Obama} \\ x & 0.46 & -0.32 \\ y & -0.25 & 0.82 \end{bmatrix}.\end{aligned}$$

Here

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mathbf{z}}(\mathbf{z}) = \left(\frac{\mathrm{d}\mu_{\text{mc}}}{\mathrm{d}z_{\text{mc}}}(z_{\text{mc}}), \frac{\mathrm{d}\mu_{\text{ob}}}{\mathrm{d}z_{\text{ob}}}(z_{\text{ob}})\right)$$

is the pair of direction gradients, induced by activist preferences, acting on the two candidates.The difference between $\mathbf{z}^*$ and $\mathbf{z}^{\text{el}}$ thus provides an estimate of the activist pull on the two candidates. In this election, we estimate that activists pull the two candidates into opposed quadrants of the policy space. The estimated distributions

of activist positions for the two parties, in these two opposed quadrants (as given in Fig. 1) are compatible with this inference. The means of these activist positions are:

$$\begin{bmatrix} & \text{Rep Act} & \text{Dem Act} \\ x & 1.41 & -0.2 \\ y & -0.82 & 1.14 \end{bmatrix}.$$

Miller and Schofield (2003, 2008) propose a model where activists have eccentric or ellipsoidal utility functions. If we assume that the Democrat activists tend to be more concerned with social policy and Republican activists tend to be more concerned with economic policy, then we have an explanation for the candidate shifts from the estimated equilibrium. Note in particular that the distribution of activist positions for the two parties, given in Fig. 2, looks very different from the voter positions, given in Fig. 1. The latter is much more heavily concentrated near the electoral origin, while the former tends to be dispersed.

Miller and Schofield (2008) also emphasized the potential conflict between economically conservative and socially conservative Republican activists. In Indiana in February 2010, the incumbent Democrat Senator, Evan Bayh, announced that he would retire. This set off a contest by local "tea party" social conservatives against the Republican National Committee's support for Dan Coats, an economic conservative contender for the Senate seat. This example just illustrates the degree to which contenders for political office require support from activist groups with very different agendas.

When the candidates are at their estimated positions, the estimated vote shares, according to the traits model, are $(V_{\text{Obama}}, V_{\text{McCain}}) = (0.68, 0.32)$. Since the actual vote shares are $(0.52, 0.48)$, it appears that the trait model may give a statistically plausible account of voter choice, but it does not provide, by itself, a good model of how candidates obtain votes. We suggest that the missing characteristic of this model of the election is the effect on the vote by the contributions of party activists.

Indeed, we suggest that the addition of activists to the model can account for the difference between convergent, equilibrium positions and the divergent, estimated candidate positions, as obtained by Enelow and Hinich (1989) and Poole and Rosenthal (1984), respectively, in their various analyses of U.S. elections.

As we noted above, we could also interpret $\frac{\mathrm{d}\mu}{\mathrm{d}\mathbf{z}}(\mathbf{z})$ as the gradient obtained from a model where candidates have policy preferences derived from utility functions $(\mu_{\text{mc}}, \mu_{\text{ob}})$. Duggan and Fey (2005) have explored such a model for the case of a deterministic vote model, and obtained symmetry conditions for equilibrium similar to those obtained earlier by McKelvey and Schofield (1987). However, in such a model of policy seeking candidates, a candidate must be willing to adopt a losing position because of strong preferences for particular policies.

It is possible that our estimates of the positions, $z_{\text{Obama}} = (-0.22, 0.75)$ and $z_{\text{McCain}} = (0.59, -0.37)$, are incorrect. However, these estimated positions give us a statistically significant model of voter choice. We argue that the most plausible account for the difference in the estimated and equilibrium positions of the two candidates is the nature of activist competition.[35]

[35] See Schofield et al. (2010a) for the US election in 2000.

Table 6 Seats in the Knesset

Party	1988	1992	1996	1999	2003	2006	2009
Left (ADL, Arab, Hadash)	14	5	9	10	9	10	11
Meretz		12	9	10	6	5	3
Labor	39	44	34	28	21	19	13
Center (Olim, Gesher, Shinui)	2	8	11	18	15	7	–
Center (Kadima)						29	28
Likud	40	32	30	19	40	12	27
Shas, Yahadut	15	10	14	22	16	12+6	11+5
NRP, Mafdal	5	6	9	5	6	9	4+3
Moledat, Techiya, Beiteinu	5	3	2	8	7	11	15
Total	120	120	120	120	120	120	120

Note also that this model can be applied to the determination of policy positions of members of the House and Senate of the United States. In particular, we would expect local activist groups to be very heterogenous across states and House constituencies. As a result, policy positions of members of Congress can be expected to be very heterogenous, even within parties.

5 Elections in Israel

Schofield and Sened (2006) estimated various multinomial conditional logit models for the elections of 1988, 1992, and 1996 in Israel.[36] Table 6 gives the election results for 1988–2009, while Fig. 4, presented above, showed the electoral distribution in 1996, together with estimates of the party positions. Using the formal analysis, we can readily show that the *convergence coefficient* of the pure spatial model, $\mathbb{M}(\mathbf{\Lambda}, \beta)$ for 1996 greatly exceeds 2 (the dimension of the policy space). Indeed, one of the eigenvalues of the Hessian of the one of the low valence parties, Shas, can be shown to be positive. The principal electoral axis (or principal component of the electoral distribution) can be seen to be aligned at approximately 45° to the security axis. As we now show, this axis is the eigenspace of the positive eigenvalue. It follows from the computation of eigenvalues that low valence parties should position themselves close to this principal axis, as illustrated in the simulation of the model, given in Fig. 5.

The MNL estimation given in Table 7 presents the relative valences in the pure spatial model with respect to Meretz. The table shows that in 1996 Shas had a relative valence of $\Lambda_{\text{Shas}} = -2.02$, while Labor had the highest relative valence of 0.99, with Likud having a valence of 0.78. The spatial coefficient was $\beta = 1.21$, so to use the convergence theorem, we note that the valence difference between Shas and

[36] Schofield and Sened (2006) compared the joint MNL spatial model involving sociodemographic terms and valences with various less extensive models. The joint model correctly predicted 63.8% of the voter choices.

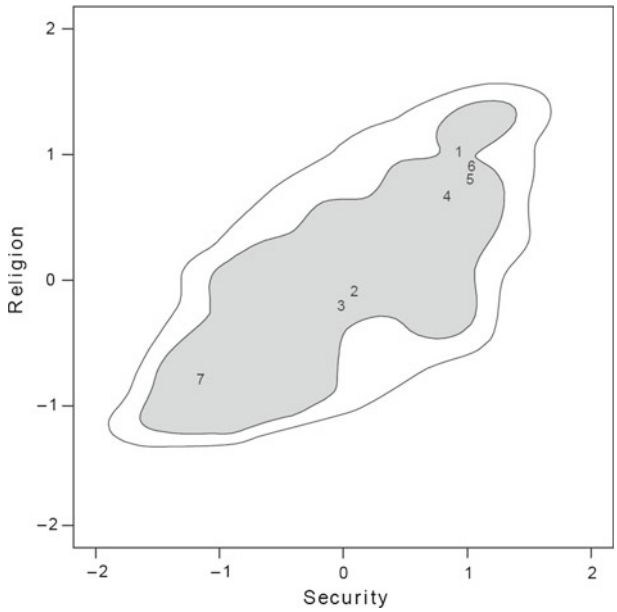

Fig. 5 Estimated local equilibrium positions in the Knesset in 1996. Key: *1* Shas, *2* Likud, *3* Labor, *4* NRP, *5* Moledat, *6* Third way, *7* Meretz

Labor was $0.99 - (-2.02) = 3.01$, while the difference between Shas and Likud was $0.78 - (-2.02) = 2.8$. The electoral covariance matrix is

$$\nabla_0 = \begin{bmatrix} 1.0 & 0.591 \\ 0.591 & 0.732 \end{bmatrix}$$

with trace $\sigma^2 = 1.732$. The principal component of this electoral distribution is given by the vector $(1.0, 0.80)$ with variance 1.47, while the minor component is given by $(1.0, -1.25)$ with variance 0.26. We can compute the characteristic matrix of Shas at the origin and the convergence coefficient as follows:

$$\begin{aligned} \rho_{\text{Shas}} &\simeq \frac{1}{1 + e^3 + e^{2.8} + e^{1.4} + e^{0.8}} \\ &\simeq 0.023. \\ 2\beta(1 - 2\rho_{\text{Shas}}) &= 2 \times 1.21 \times 0.95 = 2.30 \\ \text{so } C_{\text{Shas}} &= (2.3)\nabla_0 - I \\ &= \begin{bmatrix} 1.3 & 1.36 \\ 1.36 & 0.69 \end{bmatrix}. \\ \text{and } c &= 2.3 \times 1.732 = 3.98. \end{aligned}$$

From the estimate of C_{Shas}, it follows that the two eigenvalues are 2.39 and -0.39, giving a *saddlepoint*, and a value of 3.98 for the convergence coefficient. This exceeds

Table 7 Spatial model of the Israel election 1996, wrt Meretz

Variable	Party	Coefficient	Lower 95% bound	Upper 95% bound
β Spatial		1.207***	1.076	1.338
Λ_Valence	Likud	0.777***	0.400	1.154
	Labor	0.990***	0.663	1.316
	NRP	−0.626***	−1.121	−0.132
	Moledat	−1.259***	−1.858	−0.660
	Third way	−2.291***	−2.841	−1.741
	Shas	−2.023***	−2.655	−1.392
Convergence c		3.98	3.70	4.26
LL = − 777.0	$n = 922$	*** $P < 0.001$		

LL Log likelihood

Table 8 Comparison of LL for Israel models for 1996

	$\mathbb{M}_2$	Joint	Spatial	Socio-Dem.
	Joint	na	82	249
$\mathbb{M}_1$	Spatial	−82	na	167
	Socio-Dem.	−249	−167	na

the necessary upper bound of 2. The estimate for the standard error on ρ_{Shas} is 0.008, so the 95% confidence interval is [0.007, 0.02]. Note that this interval includes the actual sample vote share of 2% for Shas. The standard error on β is 0.065 so the standard error on c is of order 0.14, and we can infer that, with high probability, the convergence coefficient exceeds 2.0.

Using the above estimate for the major eigenvalue, we find that the major eigenvector for Shas is (1.0, 0.79), and along this axis the Shas vote share function increases as the party moves away from the origin. The minor, perpendicular axis associated with the negative eigenvalue is given by the vector (1, −1.26). Any LNE for the model $\mathbb{M}(\boldsymbol{\Lambda}, \beta)$ will be one where all parties are located on the major eigenvector.

We also constructed a joint MNL model, $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \beta)$, and a pure sociodemographic model of the election, $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta})$, details of which can be found in Schofield and Sened (2006). Table 8 reports the differences in the log likelihoods of the various models.

Figure 5 gives one of the local Nash equilibria, obtained by simulation of the model. Since this model does not involve activist terms, we can infer that this equilibrium gives an estimate of the weighted electoral means, $\mathbf{z}^{\text{el}}$, for the parties: This vector, $\mathbf{z}^{\text{el}}$, is given by:

$$\begin{bmatrix} \text{Party} & \text{Meretz} & \text{Moledat} & \text{IIIWay} & \text{Labor} & \text{Likud} & \text{NRP} & \text{Shas} \\ x & -1.1 & 1.0 & 1.0 & 0.0 & 0.2 & 0.9 & 1.0 \\ y & -0.8 & 0.8 & 0.8 & -0.2 & 0.0 & 0.6 & 1.0 \end{bmatrix}$$

All these equilibrium positions lie very close to an eigenvector (1.0, 0.85). It thus appears that the only effect of the inclusion of the sociodemographic variables is to

slightly rotate the principal eigenvector in an anticlockwise direction. In all, five different LNE were located. However, in every equilibrium, the two high valence parties, Labor and Likud, were located close to the simulated equilibrium positions shown in Fig. 5. The only difference between the various equilibria were slight differences in the positions of Shas, NRP, and Moledat.

It is evident that if the high valence party occupies the electoral origin, then each party with low valence can compute that its vote share will increase by moving up or down the principal electoral axis. In seeking local maxima of the vote shares, all parties other than the highest valence party should vacate the electoral center. Then, however, the first-order condition for the high valence party to occupy the electoral center would not be satisfied. Even though this party's vote share will be little affected by the other parties, it too should move from the center. The simulation for 1996 is compatible with the formal analysis: low valence parties, such as the NRP and Shas, in order to maximize vote shares must move far from the electoral center. As with the pure spatial model, their optimal positions will lie either in the "north-east" quadrant or the "south-west" quadrant. The vote-maximizing model, without any additional information, cannot determine which way the low valence parties should move.

The equilibrium position of Shas, by the joint model, will give greater weight to those voters who are observant. As Fig. 4 makes clear, Shas, Moledat, and NRP are located in the upper quadrant of the policy space. On the other hand, since the valence difference between Labor and Likud was relatively low, their local equilibrium positions will be close to, but not identical to, the electoral mean. Intuitively, it is clear that once the low valence parties vacate the origin, then high valence parties, like Likud and Labor, should position themselves almost symmetrically about the origin, and close to the principal axis.

We now compare the LNE obtained from the joint model with the vector, $\mathbf{z}^*$, of estimated positions given in Fig. 4:

$$\begin{bmatrix} \text{Party} & \text{Meretz} & \text{Moledat} & \text{IIIWay} & \text{Labor} & \text{Likud} & \text{NRP} & \text{Shas} \\ x & -1.5 & 1.4 & -0.2 & -0.8 & 0.6 & 1.0 & 0.0 \\ y & -1.0 & 0.5 & -0.4 & -0.2 & 0.2 & 1.1 & 1.1 \end{bmatrix}.$$

We hypothesize that $\mathbf{z}^*$ is a local equilibrium of the full activist model: the difference, $\mathbf{z}^* - \mathbf{z}^{\text{el}}$, between the vector of positions and the equilibrium of Fig. 5 is of order

$$\begin{bmatrix} \text{Party} & \text{Meretz} & \text{Moledat} & \text{IIIWay} & \text{Labor} & \text{Likud} & \text{NRP} & \text{Shas} \\ x & -0.4 & 0.4 & -1.2 & -0.8 & 0.4 & 0.1 & -1.0 \\ y & -0.2 & -0.3 & -1.2 & 0.0 & 0.2 & 0.5 & 0.1 \end{bmatrix}.$$

Thus, this vector gives an estimate of the influence of activist groups on the parties:

$$\mathbf{z}^* - \mathbf{z}^{\text{el}} = \frac{1}{2\beta}\left[\frac{\mathrm{d}\mu_1}{\mathrm{d}z_1}, \ldots, \frac{\mathrm{d}\mu_{\text{p}}}{\mathrm{d}z_{\text{p}}}\right].$$

Schofield and Sened estimate $\beta = 1.117$ for the joint model, so we obtain

$$\left[\frac{\mathrm{d}\mu_1}{\mathrm{d}z_1}, \ldots, \frac{\mathrm{d}\mu_p}{\mathrm{d}z_p}\right] = 2\beta(\mathbf{z}^* - \mathbf{z}^{\mathrm{el}})$$
$$= \begin{bmatrix} \text{Party} & \text{Meretz} & \text{Moledat} & \text{IIIWay} & \text{Labor} & \text{Likud} & \text{NRP} & \text{Shas} \\ x & -0.9 & 0.9 & -2.7 & -1.78 & 0.9 & 0.22 & -2.2 \\ y & -0.45 & -0.67 & -2.68 & 0.0 & 0.45 & 1.12 & 0.22 \end{bmatrix}$$

Although we have not performed the empirical analysis for the elections of 2003 and 2006, we can expect a similar result to hold. The analysis given in Schofield and Sened (2006) for the elections of 1992 and 1988 shows that in 1988 the two eigenvalues for Shas were $+2.0$ and -0.83, while in 1992 the eigenvalues for this party were $+2.12$ and -0.52. Just as in 1996, the theoretical model of vote maximization implies that all parties should be located on a principal electoral axis. The positioning of Shas off the principal electoral axis enables it to pivot between the two major parties, in the sense that it tended to be crucial for the formation of winning coalitions.

As Table 7 shows, after the elections of 1996, 1999, and 2003 any winning coalition based on either Labor or Likud needed additional support of Shas. In 1996, Netanyahu of Likud formed a government with Shas, but after Likud lost seats in 1999, it was the turn of Barak of Labor to form a government, again with Shas, followed in 2001 by Likud, led by Sharon, with Shas. In consequence, even though Shas controlled few seats in this period, it had significant bargaining power.

5.1 The elections of 2006 and 2009

This pattern of coalition government was transformed, to some degree, when Amir Peretz stood against Shimon Peres and won the election for leadership of Labor in November 2005.

Sharon then left the Likud Party and allied with Peres and other senior Labor Party members, to form the new party, *Kadima* ("Forward"). We can infer that the coalition of Sharon and Peres positioned Kadima at the center of the policy space. Because of Sharon's stroke in January 2006, Ehud Olmert took over as leader of Kadima, and in the election of March 2006, the new party was able to take 29 seats, while Likud only took 19 seats. One surprise of the election was the appearance of a Pensioners' party with 7 seats. A possible coalition of Likud and the religious parties, opposed to Kadima, did not have the required 61 seats for a majority (even with the Pensioners' Party). Schofield (2007b) discussed this election and argued that Kadima was at the core position, since no majority coalition could agree to overturn the Kadima position.[37] However, this "core property" was unstable, in the sense that it could be destroyed by small changes in positions or strengths of the parties.

[37] For a discussion of the core see Laver and Schofield (1990). For a spatial voting game in a legislature, the core is given by the intersection of legislative median lines between pairs of parties that pivot. If these lines do not intersect then the core is empty. In this case, the set bounded by these median lines is called the "heart". By definition, when the core is empty, then the heart is non empty. See Schofield (1999, 2007b)

As a result, Olmert needed the support of Labor to be able to deal with the complex issue of fixing a permanent border for Israel. The debacle in Lebanon severely weakened Olmert's popularity, and the 61 members of the Kadima-Labor coalition voted to bring Israel Beiteinu into the coalition. The report, in April 2007, on the failure of the government during the war with Lebanon in Summer 2006 seemed to threaten the Kadima-Labor–Israel Beiteinu coalition by bringing about a change in the Labor party leadership. Barak then won the election for the Labor Party leadership on June 12, 2007, and became Minister of Defense in the government on June 18, while Shimon Peres became President. In November 2007, Olmert proposed a land-for-peace proposal, possibly involving the separation of Jerusalem, and on January 15, 2008, Avigdor Lieberman, chairman of Israel Beiteinu announced that the party would quit the government because of disagreement over issues such as Jerusalem and negotiations with Hamas.

On February 3, 2008, Barak agreed to remain in the coalition, thus helping to sustain Kadima in power. However, in August 2008, Olmert faced charges of corruption, and formally resigned as leader of Kadima on September 21. He immediately gave an interview (Olmert 2008) in which he asserted that Israel would have to lose sovereignty over Jerusalem, and would have to come to an agreement with Syria by giving up the Golen Heights in return for Syrian forswearing their connections with Iran, Hezbollah, and Hamas.

The new leader of Kadima, and Prime Minister designate, Tzipi Livni, then had to face a revolt by Shas, over these security issues. On October 26, 2008, she announced, that she had failed to form a viable coalition, and an election would occur in February 2009. Even though the Kadima government was weakened, it responded to rocket attacks by Hamas from Gaza, and launched a 3 week attack on Gaza at the end of December 2008.

In the election of 2009, as Table 6 shows, the Pensioners' Party disappeared, and both Likud and Israel Beiteinu gained seats. Labor lost significantly, presumably because of the loss of valence by its leader, Ehud Barak. The core was destroyed, and it was unclear what government would form. Both Livni and Benjamin Netanyahu, of Likud, claimed the electoral mandate. However, on February 20, Avigdor Lieberman took the role of *formateur* of the coalition game, and offered his support to Netanyahu. On March 24, a majority of the Labor Party central committee voted to support Netanyahu, in return for four cabinet positions, and the retention of the defense portfolio by Barak. Tzipi Livni refused the offer to join this unity coalition government of Likud, Labor, Shas, and Israel Beiteinu, and will be in opposition. As prime minister designate, Netanyahu declared on March 26 that he would negotiate with the Palestinian Authority for peace. Five days later he was sworn in as Prime Minister, after a vote of 69 to 45, with the abstention of five Labor members (one Arab member of the Knesset was absent). Avigdor Lieberman became foreign minister. Although Netanyahu has tended to avoid mention of a sovereign Palestinian state, he declared in December 2009 that in order to proceed with this policy, he was willing to consider inviting Livni to join in a grand coalition.

In March 2010, during Vice President Biden's visit to Israel it was announced that Israel would add 1,600 housing units in eastern Jerusalem. Although the Obama administration was angered by the timing of the announcement, Netanyahu insisted

that Israel would go ahead with the construction. However, President Shimon Peres said: "We cannot afford to unravel the delicate fabric of friendship with the United States. Today we are also at a decisive moment and we must decide without the determination of external parties." However, on May 31,2010, there was an attack by Israeli commandoes against a boat traveling in international waters and carrying humanitarian supplies for Gaza. Nine people in the convoy were killed. The convoy was partly organized by a Turkish organization, Insani Yardim Vakfi. It is unclear who ordered the attack, but Netanyahu immediately cancelled a meeting that had been arranged with Obama.

Eventually, in September 2010, negotiations started in Washington, involving Netanyahu, Mahmoud Abbas (the President of the Palestinian Authority), King Abdullah II of Jordan and President Hosni Mubarak of Egypt.

5.2 Coalitions in Israel

We can see the nature of bargaining over this coalition government by joining the median lines between pairs of parties that pivot between majority coalitions after the 2009 election, as shown in Fig. 6. When these medians do not intersect, then they bound a finite, star shaped set known as the "heart."

Schofield (2007b) argues that the outcomes of coalition bargaining will be a constrained within this set. [38] The complex nature of this set suggests that there are many possible majority coalitions. In particular, small parties such as Shas, Yahadut, and Israel Beiteinu may join in government and may thus influence the outcome of coalition government. We have argued that the positions adopted by the parties are the result of activist choices to support particular parties. Thus, activist groups for these small parties may reason that the party they support has a good chance of taking part in government, thus bringing about policy changes that favor the activists. Consequently, there is little motivation for such activist groups to coalesce. As long as the logic of vote maximization maintains this policy divergence between the parties, then so will activist groups continue to provide support for these small parties. Thus, political fragmentation is preserved

These remarks about recent events in the Knesset are presented to illustrate the great difficulty of maintaining a stable government coalition, even when there is a large, centrally located party, such as Kadima. Such a party should, in principle, be able to dominate bargaining. However, it is only when the center party's leader has high valence is the party able to avoid threats to the government. Without such valence predominance, small parties, and their activist supporters have an incentive to act to maintain political fragmentation.

6 Concluding remarks

This article has argued, on the basis of an equilibrium analysis of elections, that the electoral pull on parties is very different in the United States and Israel, and this is a

[38] As mentioned above, when the medians intersect then the heart will collapse to the core, and we may assume that this will be the outcome.

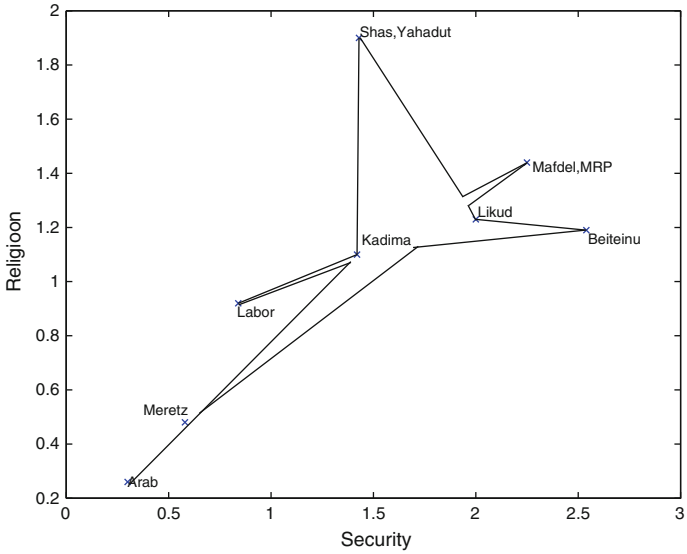

Fig. 6 The heart in Israel in 2008

consequence of the nature of the electoral system and the degree of political fragmentation.

We can express the difference between proportional representation and plurality rule as follows.

Analysis of the equilibria of the spatial valence model of a polity based on *proportional electoral methods*, such as Israel, indicates that only high valence parties will be positioned, in equilibrium, close to the electoral origin. Because of the wide variation in political valence, the low valence parties will move toward the electoral periphery. Activist groups, linked to small parties, may aspire to affect the policy choices of the chosen parties. Bargaining to create winning coalitions occurs *after* the election, and small parties may aspire to membership of government. As a consequence, there need to be no strong tendency forcing activist groups to coalesce, in order to concentrate their influence. If activist groups remain fragmented, then the outcome will be party fragmentation, and this will reinforce the tendency for variation in valence. Empirical analysis of the election in Israel for 1996 showed that the convergence coefficient had a high value of 3.98 in 1996. Other empirical work obtained a value for this convergence coefficient of 6.82 for 1997 in Poland (Schofield et al. 2010b) and of 5.94 for Turkey in 2002 (Schofield et al. 2011).

A standard way of estimating political fragmentation is in terms of the *effective number of party vote strength* (env) or *effective number of party seat strength* (ens).[39] The fragmentation in votes and seats is captured by the fact that in Israel in 1996 both

[39] Fragmentation can be identified with the *effective number* (Laakso and Taagepera 1979). That is, let H_v (the Herfindahl index) be the sum of the squares of the relative vote shares and $\text{env} = H_v^{-1}$ be the *effective number of party vote strength*. In the same way we can define ens as the effective number of party seat strength using shares of seats.

Table 9 Convergence coefficients (c) and fragmentation

	Country		
Variable	US	Britain	Canada
Convergence coeff. (c)	[0.40,1.1] (2000–2008)	[0.84, 0.98] (2005–2010)	2.0 (2004)
Political system	Pres. PL	Parl. PL	Parl. PL
env	2.0	3.2 (1997)	4.0 (2004)
env		2.7 (2005)	4.1 (2008)
ens	1.0	2.2 (1997)	3.1 (2004)
ens		2.5 (2005)	3.5 (2008)
	Poland	Turkey	Israel
Convergence coeff. (c)	6.82 (1997)	5.94 (2002)	3.98 (1996)
Political system	Frag. PR	Frag. PR, cut off	Frag PR
env	5.5 (1997)	7.7 (1999)	6.5 (1996)
env	7.7 (2005)	4.0 (2007)	10.0 (2009)
ens	3.1 (1997)	5.0 (1999)	6.5 (1996)
ens	5.0 (2005)	2.3 (2007)	10.0 (2009)

Parl parliamentary, *Pres.* presidential, *PL* plurality, *PR* proportional representation, *Frag.* fragmented

env and ens were equal to 6.5. As Table 9 shows, these had risen to 10.0 by 2009. The table shows that the effective numbers for vote shares in Poland and Turkey were also large, in the range [4.0, 7.7]. These three polities all have electoral systems that are approximately proportional, although Turkey introduced an electoral cut-off, which worked to the disadvantage of smaller parties, reducing the ens from 5.0 in 1999 to 2.3 in 2007.

The parliamentary polities of Canada and Britain, based on a plurality electoral system, have convergence coefficients that lie in the range [0.8, 2.0], and the fragmentation measures tend to be much lower than in the PR polities. For Canada, the env is about 4.0, and in Britain it lies in the range [2.7, 3.2]. In the very majoritarian polity of the United States the env is generally about 2.0. The equilibrium analysis of the spatial model presented here indicates that the convergence coefficient had a low value of 1.1 in 2008. In 2000 and 2004, Schofield et al. (2010a,b) estimate the coefficient to be 0.4.

Schofield and Zakharov (2010) suggest that Russia has an electoral system that is in between plurality and proportionality. It has a dominant party, United Russia, supportive of Putin, and the env was only 2.3 in the Duma election of 2007, while the convergence coefficient was estimated to be 1.7.

These comparisons of convergence coefficients are valid because c is dimensionless. The higher coefficients in polities with many parties is a result both of higher β-coefficients and greater electoral variance. In contrast, for US presidential elections, the convergence property still holds even when sociodemographic variables and individual perceptions of candidate traits are incorporated into the model. We argue that presidential candidates in the United States are pulled from these convergent equilibrium positions by the influence of interest groups. We suggest that in the election of

2008, social activists dominated in the Democrat party and economic activists in the Republican party. Activist resources are crucial for electoral success. If these interest groups do not coalesce *before* the election, then they will have little impact on political outcomes. Consequently, small parties, or activist groups, such as those led by independent candidates in recent elections, have little expectation of influencing government policy. Their valences will remain low, and they will have little impact, in general, on presidential elections. However, heterogenous local activist groups may influence the policy preferences of members of Congress, and this effect may induce conflict between the President of one party and Senators of the same party, as seems to be the case over the issue of health care in 2009 and 2010. Radical activist groups may also induce conflict within the support coalitions for one or other of the parties, particularly when the party is out of office. This phenomenon can be seen within the Republican party in the aftermath of the 2008 election. We may hypothesize that such conflict is the fundamental cause of realignment in the US polity.

Appendix 1: a formal stochastic model of elections

Details of the pure spatial stochastic electoral model are given in Schofield (2006, 2007a), to which the reader is referred. Here, we extend the presentation given in Schofield et al. (2011) by including electoral perceptions of candidate traits as well as allowing for multiple activist groups.

The aim of the model is to argue that the vote-maximizing equilibrium position of McCain will lie on what we shall call a *balance locus* in the lower right quadrant of the policy space, while Obama's position will lie in the opposite, upper left, quadrant.

The voter utility assumption for the stochastic vote model $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\mu}, \beta)$ is:

$$u_{ij}(x_i, z_j) = \Lambda_j + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) + \mu_j(z_j) - \beta\|x_i - z_j\|^2 + \varepsilon_j \tag{3}$$

$$\equiv u^*_{ij}(x_i, z_j) + \varepsilon_j. \tag{4}$$

Here, $u^*_{ij}(x_i, z_j)$ is the observable component of utility. The set of leaders is $P = \{1, \ldots j, \ldots p\}$. The term Λ_j is the exogenous valence of leader j, relative to the baseline leader. This is estimated from the intercept term of the model. The symbol $\boldsymbol{\theta}$ denotes a set of k -vectors $\{\theta_j : j \epsilon P\}$ representing the effect of the k different sociodemographic parameters[40] on voting for leader j while η_i is a k-vector denoting the ith individual's relevant "sociodemographic" characteristics. The compositions $\{(\theta_j \cdot \eta_i)\}$ are scalar products. We refer to the terms $\{(\theta_j \cdot \eta_i)\}$ as the total *sociodemographic valence* by i for leader j.

In similar fashion, the terms $\{(\alpha_j \cdot \tau_i)\}$ are scalar products, where τ_i is voter i's perception of the trait of leader j, with coefficient α_j. Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$.

The trait score for the election of 2008 was obtained by factor analysis from a set of survey questions, as mentioned above.

[40] These will depend on the survey but will include such characteristics as class, domicile, education, income, and religious orientation, etc.

The function $\mu_j(z_j)$ is the sum of other aspects of the valence of leader j. Note that we assume that this aspect of valence is dependent on the leader position, but not on the voter position. In the analysis below, we suggest that this term can be regarded as the effect of activist contributions to the leader. Let $\boldsymbol{\mu} = \{\mu_j\}$.

The term β is a positive constant, called the *spatial parameter*, giving the importance of policy difference defined in terms of a metric induced from the Euclidean distance, $||x_i - z_j||$, between the voter's ideal point,x_i, and the leader position z_j. The vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_j, \ldots, \varepsilon_p)$ is the stochastic error, with the multivariate Gumbel (Type I extreme value) distribution. The variance of ε_j is fixed at $\frac{\pi^2}{6}$, so that by definition β has dimension L^{-2}, where L is whatever unit of measurement is used in X.

Various submodels are *pure sociodemographic* (SD), denoted $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta})$, *pure spatial,* $\mathbb{M}(\boldsymbol{\Lambda}, \beta)$, *joint spatial,* $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \beta)$, and *joint spatial with traits*, $\mathbb{M}(\boldsymbol{\Lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$.

As shown in Schofield (2006), the *first-order condition for a local Nash equilibrium under vote maximization* is given by the *balance equation* for each z_j^*:

$$\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j^*) + \frac{1}{2\beta}\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j^*) = 0. \tag{5}$$

Here, the term

$$\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j) \equiv \left[z_j^{\mathrm{el}} - z_j\right].$$

is the *marginal electoral pull of leader* j at the point z_j and can be regarded as a gradient vector, at z_j, pointing toward the weighted electoral mean of the leader. The *weighted electoral mean* for leader j is given by

$$z_j^{\mathrm{el}} \equiv \sum_{i=1}^{n} \varpi_{ij} x_i$$

If $\rho_{ij}(\mathbf{z}) = \rho_{ij}$ is the probability voter i chooses leader j, at $\mathbf{z}$ then the weights are given by the p by n matrix array of weights

$$[\varpi_{ij}] \equiv \left[\frac{[\rho_{ij} - \rho_{ij}^2]}{\sum_{k\in N}[\rho_{kj} - \rho_{kj}^2]}\right] \tag{6}$$

When z_j is equal to the weighted electoral mean then the electoral pull is zero. The gradient vector $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j)$ is called *the marginal activist pull for leader* j at z_j.

If $\mathbf{z}^* = (z_1^*, \ldots z_j^*, \ldots z_p^*)$ is such that each z_j^* satisfies the balance equation then call $\mathbf{z}^*$ a *balance solution*. We may rewrite (5) as

$$\left[\mathbf{z}^* - \mathbf{z}^{\mathrm{el}}\right] = \frac{1}{2\beta}\frac{\mathrm{d}\mu}{\mathrm{d}\mathbf{z}}(\mathbf{z}) \tag{7}$$

For the pure spatial model, $\mathbb{M}(\mathbf{\Lambda},\beta)$, it follows from (6) that, when the leader positions are identical, then $\rho_{kj} = \rho_j$, is independent of the voter k. Thus $\varpi_{ij} = \frac{1}{n}$ for all i and j, gives the first-order condition for a LNE for $\mathbb{M}(\mathbf{\Lambda},\beta)$. By a change of coordinates, we can choose $z_j^* = \frac{1}{n}\sum_{i=1}^{n} x_i \equiv 0$. It follows that $\mathbf{z}_0 = (0,\ldots 0)$ is a candidate for a LNE for $\mathbb{M}(\mathbf{\Lambda},\beta)$. Schofield (2007a) shows that the Hessian of the vote share of leader j at $\mathbf{z}_0$ can be identified with the characteristic matrix

$$C_j \equiv 2\beta(1-2\rho_j)\nabla_0 - I. \tag{8}$$

Here, I is the identity matrix and ∇_0 is the electoral covariance matrix and $\rho_j \equiv \rho_j(\mathbf{z}_0)$.

The *convergence coefficient* of the model, $\mathbb{M}(\mathbf{\Lambda},\beta)$, is defined to be

$$c \equiv c((\mathbf{\Lambda},\beta) \equiv 2\beta[1-2\rho_1]\sigma^2 \tag{9}$$

where σ^2 is the total variance, the trace of ∇_0, and $\rho_1 \equiv \rho_1(\mathbf{z}_0)$ for the lowest valence leader.

Note, however, that this argument does not follow for the model $\mathbb{M}(\mathbf{\Lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$. Even when the leader positions are identical, the probabilities $\{\rho_{kj}\}$ will depend on k. It is necessary, therefore, to compute the vector $\mathbf{z}^{\mathrm{el}} = (z_1^{\mathrm{el}}, z_2^{\mathrm{el}}, \ldots, z_p^{\mathrm{el}})$ as a step to determine the LNE.[41]

The balance solution requires that the electoral and activist gradients are directly opposed, for every leader. If the various activist groups for leader j are given by a family $\{U_{jt} : t \in A_j\}$ of utility functions then we can represent their joint effect by some contract curve. This *contract curve,* generated by the family $\{U_{jt}\}$ of activist utilities, is the locus of points satisfying the gradient equation

$$\sum_{t\in A_j} a_{jt}\frac{dU_{jt}}{dz_j} = 0, \quad \text{where } \sum_{t\in A_j} a_{jt} = 1 \text{ and all } a_{jt} > 0. \tag{10}$$

This in turn implies that the optimal position of leader j will lie on the *balance locus*

$$\left[z_j^{\mathrm{el}} - z_j^*\right] + \frac{1}{2\beta}\left[\sum_{t\in A_j} a_t \frac{dU_t}{dz_j}\right] = 0. \tag{11}$$

The simplest case, discussed in Miller and Schofield (2003), is in two dimensions, where each leader has two activist groups. In this case, the contract curve for each leader will, generically, be a one-dimensional arc. Miller and Schofield (2003) also supposed that the activist utility functions were ellipsoidal, mirroring differing saliences on the two axes. In this case the contract curve for each leader would be a *catenary*, and the balance locus would be a one-dimensional arc. The balance solution for each leader naturally depends on the position(s) of opposed leader(s), and on the

[41] We did this using a MATLAB algorithm, based on the gradient of the vote share function.

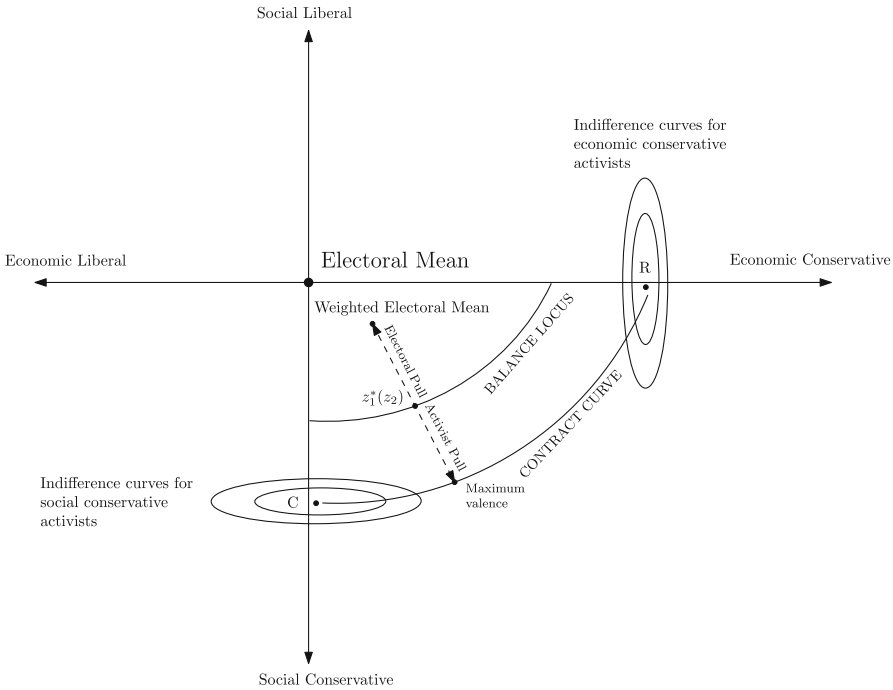

Fig. 7 Republican contract curve and balance locus

coefficients, as indicated above, of the various activists. These coefficients depend on the willingness of each activist group to supply resources in order to influence the political leader. The determination of the balance solution can be obtained by computing the vote share Hessian along the balance locus. Figure 7 illustrates the balance locus and contract curve for a Republican candidate.

Note that the combination

$$\sum_{t \in A_j} a_k \frac{\mathrm{d}U_t}{\mathrm{d}z_j}$$

may be interpreted as the marginal utility of the leader of party j, induced by the activist support.

To see this, suppose that each leader were to maximize the function

$$V_j(\mathbf{z}) = \delta \mu_j(z_j) + \frac{1}{n} \Sigma_i \rho_{ij}(\mathbf{z})$$

where μ_j is no longer an activist function, but a policy determined component of the leader's utility function, while δ is the weight given to the policy preference, Then the first-order condition is almost precisely as obtained above, namely

$$\frac{\mathrm{d}\mathcal{E}_j^*}{\mathrm{d}z_j}(z_j^*) + \frac{\delta'}{2\beta}\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j^*) = 0.$$

Here, $\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j^*)$ is a gradient pointing toward the policy preferred position of the leader and δ' is a product of n and δ, divided by a summation across the voter probabilities. Thus, we can make the identity

$$\delta'\frac{\mathrm{d}\mu_j}{\mathrm{d}z_j}(z_j^*) = \sum_{t\in A_j} a_t \frac{\mathrm{d}U_t}{\mathrm{d}z_j}.$$

This equation implies that the leader's marginal policy preference can be identifed with a combination of the marginal preferences of the party activists. To solve this equation in detail requires solving the game between activists and leaders, as outlined in Grossman and Helpman (1994, 1996, 2001). For our purposes, it is sufficient to use this reduced form, as we are interested in the difference $\left[z_j^{\mathrm{el}} - z_j^*\right]$ between the LNE position, z_j^{el}, of party j and its estimated position, z_j^*.

Appendix 2: question wording for the 2008 American national election study

1. Do you think the government should provide more services than it does now, fewer services than it does now, or about the same number of services as it does now?
2. Do you favor, oppose, or neither favor nor oppose the U.S. government paying for all necessary medical care for all Americans?
3. A proposal has been made that would allow people to put a portion of their Social Security payroll taxes into personal retirement accounts that would be invested in stocks and bonds. Do you favor this idea, oppose it, or neither favor nor oppose it?

   I am going to ask you three questions, and ask you to choose which of two statements in these questions comes closer to your own opinion.
4. One, the main reason government has become bigger over the years is because it has gotten involved in things that people should do for themselves. Two, government has become bigger because the problems we face have become bigger.
5. One we need a strong government to handle today's complex economic problems. Two, the free market can handle these problems without government being involved.
6. One, the less government, the better. Two there are more things that government should be doing.
7. This country would be better if we worried less about how equal people are. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?
8. Do you think that big companies should pay a larger percent of their profits in taxes than small businesses do, that big companies should pay a smaller percent

of their profits in taxes than small businesses do, or that big companies and small businesses should pay the same percent of their profits in taxes?

9. Should federal spending on welfare programs be increased, decreased, or kept about the same?
10. Do you favor, oppose, or neither favor nor oppose the U.S. government making it possible for illegal immigrants to become U.S. citizens?
11. Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?
12. I'd like to describe a series of circumstances in which a woman might want to have an abortion. For each one, please tell me whether you favor, oppose, or neither favor nor oppose it being legal for the woman to have an abortion in that circumstance.

    1. Staying pregnant would hurt the woman's health but is very unlikely to cause her to die.
    2. Staying pregnant could cause the woman to die.
    3. The pregnancy was caused by sex the woman chose to have with a blood relative.
    4. The pregnancy was caused by the woman being raped.
    5. The fetus will be born with a serious birth defect.
    6. Having the child would be extremely difficult for the woman financially.
    7. The child will not be the sex the woman wants it to be.

13. Do you favor or oppose laws to protect homosexuals against job discrimination?
14. Do you think homosexuals should be allowed to serve in the United States Armed Forces or don't you think so?
15. Do you think gay or lesbian couples, in other words, homosexual couples, should be legally permitted to adopt children?
16. Should same-sex couples be allowed to marry, or do you think they should not be allowed to marry?
17. This country would have many fewer problems if there were more emphasis on traditional family ties. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?
18. Do you think the federal government should make it more difficult for people to buy a gun than it is now, make it easier for people, or keep the rules the same?
19. Some people feel that the government in Washington should make every effort to improve the social and economic position of blacks. Others feel that the government should not make any special effort to help blacks because they should help themselves. Where would you place yourself on this scale, or haven't you thought much about this?
20. Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?
21. Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class. Do you agree

strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?

22. It's really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well off as whites. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?
23. We hear a lot of talk these days about liberals and conservatives. Where would you place yourself on a scale from liberal to conservative?

Acknowledgments This article is based on the work supported by NSF grant 0715929 and a Weidenbaum Center grant to Schofield. An earlier version was completed while Zakharov was a visitor at the Center. Zakharov thanks the DECAN laboratory, State University-Higher School of Economics, Moscow, for support. A version of the article was presented at the Southern Political Science Meeting, Atlanta, January 2010. This version was completed while Schofield was the Glenn Campbell and Rita Ricardo-Campbell National Fellow at the Hoover Institution, Stanford, 2010. Thanks are due to the helpful comments of two anonymous referees.

References

Acemoglu D, Robinson J (2006) Economic origins of dictatorship and democracy. Cambridge University Press, Cambridge

American National Election Study (2008)

Arian A, Shamir M (1999) The election in Israel: 1996. SUNY Press, Albany

Baron D (1994) Electoral competition with informed and uniformed voters. Am Political Sci Rev 88:33–47

Bawn K, Rosenbluth F (2005) Short versus long coalitions: electoral accountability and the size of the public sector. Am J Political Sci 50:251–265

Bayh E (2010) Why I am leaving the Senate. New York Times. February 21.

Benoit K, Laver M (2006) Party policy in modern democracies. Routledge, London

Bernhardt D, Duggan J, Squintani F (2009) The case for responsible parties. Am Political Sci Rev 103:570–587

Clarke HD, Sanders D, Stewart M, Whiteley P (2005) Political choice in Britain. Oxford University Press, Oxford

Clarke HD, Kornberg A, Scotto T (2009a) Making political choices. Toronto University Press, Toronto

Clarke H, Sanders D, Stewart M, Whiteley P (2009b) Performance politics and the British voter. Cambridge University Press, Cambridge

Downs A (1957) An economic theory of democracy. Harper and Row, New York

Duggan J, Fey M (2005) Electoral competition with policy-motivated candidates. Games Econ Behav 51:490–522

Duverger M (1954) Political parties: their organization and activity in the modern state. Wiley, New York

Dworkin R (2010) The "devastating" decision. New York Rev Books 57(3):39

Enelow JM, Hinich M (1982) Nonspatial candidate characteristics and electoral competition. J Politics 44:115–131

Enelow JM, Hinich M (1989) The location of American Presidential candidates. Math Comput Model 12:417–435

Erikson RS, Romero DW (1990) Candidate equilibrium and the behavioral model of the vote. Am Political Sci Rev 84:1103–1126

Fiorina M, Abrams SJ, Pope JC (2005) Culture war?: the myth of a polarized America. Pearson Longman, New York

Grossman GM, Helpman E (1994) Protection for sale. Am Econ Rev 84:833–850

Grossman GM, Helpman E (1996) Electoral competition and special interest politics. Rev Econ Stud 63:265–286

Grossman GM, Helpman E (2001) Special interest groups. MIT Press, Cambridge, MA

Herrera H, Levine D, Martinelli C (2008) Policy platforms, campaign spending and voter participation. J Publ Econ 92:501–513

Jeong G-H, Miller G, Schofield C, Sened I (2010) Cracks in the opposition: Partisan realigment in US Senate immigration negotiations. Working paper, Washington University
Jessee SA (2010) Partisan bias, political information and spatial voting in the 2008 Presidential Election. J Politics 72:327–340
Keefer P (2004) What does political economy tell us about economic development-and vice versa. Annu Rev Political Sci 7:247–272
Laakso M, Taagepera R (1979) Effective number of parties: a measure with applications to West Europe. Comp Political Sci 12:3–27
Laver M, Schofield N (1990) Multiparty Government: the politics of coalition in Europe. Oxford University Press, Oxford. Reprinted 1998, University of Michigan Press, Ann Arbor
McCarty N, Poole K, Rosenthal H (2006) Polarized America. MIT Press, Cambridge, MA
McKelvey RD, Patty JW (2006) A theory of voting in large elections. Games Econ Behav 57:155–180
McKelvey RD, Schofield N (1987) Generalized symmetry conditions at a core point. Econometrica 55: 923–933
Miller G, Schofield N (2003) Activists and partisan realignment in the U.S. Am Political Sci Rev 97: 245–260
Miller G, Schofield N (2008) The transformation of the Republican and Democratic coalitions in the U.S. Perspect Politics 6:433–450
Olmert E (2008) The time has come to say these things. New York Rev Books 55(19):6–8
Olson M (1965) The logic of collective action. Yale University Press, New Haven
Persson T, Tabellini G (2000) Political economics. MIT Press, Cambridge, MA
Persson T, Tabellini G (2003) The economic effect of constitutions. MIT Press, Cambridge, MA
Poole K, Rosenthal H (1984) U.S. Presidential elections 1968–1980: a spatial analysis. Am J Political Sci 28:283–312
Riker WH (1953) Democracy in the United States. Macmillan, New York
Riker WH, Ordeshook PC (1973) An introduction to positive political theory. Prentice-Hall, Englewood Cliffs, NJ
Schofield N (1999) The heart and the uncovered set. J Econ Suppl 8:79–113
Schofield N (2002) Representative democracy as social choice. In: Arrow K, Sen A, Suzumura K (eds) The handbook of social choice and welfare. North Holland, New York
Schofield N (2006) Equilibria in the spatial stochastic model with party activists. Rev Econ Des 10:183–203
Schofield N (2007a) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980
Schofield N (2007b) Political equilibrium with electoral uncertainty. Soc Choice Welf 28:461–490
Schofield N, Milller G (2007) Elections and activist coalitions in the U.S. Am J Political Sci 51:518–531
Schofield N, Sened I (2006) Multiparty democracy: elections and legislative politics. Cambridge University Press, Cambridge
Schofield N, Zakharov A (2010) A stochastic model of the 2007 Russian Duma election. Pub Choice 142:177–194
Schofield N, Miller G, Martin A (2003) Critical elections and political realignments in the U.S.: 1860–2000. Political Stud 51:217–240
Schofield N, Claassen C, Ozdemir U (2010a) Empirical and formal models of the United States presidential election in 2000. Working paper, Washington University
Schofield N, Claassen C, Ozdemir U, Tavits M (2010b) Modeling elections in the United States and Poland: voter perceptions, political leaders and activists. Working paper, Washington University
Schofield N, Gallego M, Ozdemir U, Jeon J (2010c) Modeling elections in Canada and Great Britain. Working Paper: Washington University in Saint Louis
Schofield N, Gallego M, Ozdemir U, Zakharov A (2011) Competition for popular support: a valence model of elections in Turkey. Soc Choice Welf (this issue). doi:10.1007/s00355-010-0505-2
Serra G (2010) Polarization of what? A model of elections with endogenous valence. J Politics 72:426–427
Stokes D (1963) Spatial models and party competition. Am Political Sci Rev 57:368–377
Stokes D (1992) Valence politics. In: Kavanagh D (ed) Electoral politics. Clarendon Press, Oxford
Sundquist JL (1973) Dynamics of the party system: alignment and realignment of political parties in the United States. The Brookings Institution, Washington, DC
Taylor M, Herman VM (1971) Party systems and government stability. Am Political Sci Rev 65:28–37
Tomasky M (2010) The money fighting health care reform. New York Rev Books 57(6):10–14
Warwick P (1994) Government survival in parliamentary democracies. Cambridge University Press, Cambridge

Soc Choice Welf
DOI 10.1007/s00355-010-0506-1

ORIGINAL PAPER

Secondary issues and party politics: an application to environmental policy

Vincent Anesi · Philippe De Donder

Abstract This article develops a political economy model to assess the interplay between party formation and an environmental policy dimension viewed as secondary to the redistributive dimension. We define being a secondary issue in terms of the intensity of preferences over this issue rather than in terms of the proportion of voters who care for the environment. Equilibrium policies are the outcome of an electoral competition game between endogenous parties. We obtain the following results: (i) The Pigouvian tax never emerges in an equilibrium; (ii) The equilibrium environmental tax is larger when there is a minority of green voters; (iii) Stable green parties exist only if there is a minority of green voters and income polarization is large enough relative to the saliency of the environmental issue. We also study the redistributive policies advocated by green parties.

1 Introduction

The objective of this article is to study electoral competition when the policy space is composed of a "frontline issue," redistribution, and of a "secondary issue," environmental policy. Frontline issues are those that are considered important enough that they drive the vote of a large fraction of the electorate. Prominent examples are the aggregate level of government spending or the degree of income redistribution. By contrast, secondary issues are not the main focus of a large fraction of the electorate. Such secondary issues include gun control, trade policy, foreign aid, or environmental

V. Anesi
University of Nottingham (CEDEX and CPE), Nottingham, UK

P. De Donder (✉)
Toulouse School of Economics (GREMAQ-CNRS and IDEI), Toulouse, France
e-mail: philippe.dedonder@tse-fr.eu

Published online: 21 December 2010

policy. Many authors then take the view that secondary issues are better studied in the context of special interest politics, and especially of lobbies. A recent article by List and Sturm (2006) argues to the contrary that electoral incentives constitute an important determinant of policy choices for secondary issues as well. While List and Sturm (2006) focus on political accountability (how incentives for being re-elected affect the incumbent's choices in both issues), our focus is on political compromise: how incentives for being re-elected affect all politicians' willingness to compromise on both issues.

We choose environmental policy as the secondary dimension because we are especially interested in understanding the role that the formation of political parties plays on that policy domain. Our main motivation is the emergence in the last decades of "green parties" which are mainly focused on the environment. We wish to better understand how these parties survive in a political system where environmental issues are not frontline for a majority of voters, and what type of redistributive policy green parties advocate at equilibrium. To be more precise, we wish to shed light on the following questions: Under which circumstances (if any) is the equilibrium environmental policy efficient? How is this policy affected by the proportion of voters who care about pollution? What are the necessary conditions to be satisfied for a green party to form at equilibrium? Can we have more than one green party at equilibrium? Who forms the constituency of a green party? Why is it that green parties are overwhelmingly associated with strong redistributive concerns?[1]

To the best of our knowledge, the political economy literature has not developed electoral competition models where the environment is secondary to another dimension. The majority of this literature assumes that the environmental policy is shaped by the action of lobbies and adopts mainly the menu auction approach first introduced by Bernheim and Whinston (1986) and popularized by Grossman and Helpman (2002). In this approach, elections are typically not explicitly modeled.[2] Recent surveys of this literature include Heyes and Dijkstra (2001) and Oates and Portney (2003).

A second branch of the political economy literature, beginning with Congleton (1992) and sometimes referred to as "majority voting models," applies variants and extensions of the median voter model to diverse economic settings. For instance, McAusland (2003) uses a majority voting model to analyze how inequality and openness to trade interact to determine voters' demand for environmental policy, and Jones and Manuelli (2001) and Kempf and Rossignol (2007) study voting over environmental policy in growth models. A small number of articles introduce, however, both the redistributive and the environmental dimensions, and use different political equilibrium concepts (sequential voting for Cremer et al. (2004), Party Unanimity Nash Equilibrium (PUNE) for Cremer et al. (2008)) but they do not model the environment as secondary.[3] As for List and Sturm (2006), we differ from them on two main accounts. First, they develop a political agency model with an incumbent, while we

[1] See Neumayer (2004) and the many references therein for empirical evidence.

[2] With the exception of Wilson and Damania (2005) who combine common agency and Downsian politics.

[3] Also related are the articles by Brett and Keen (2000) and Anesi (2006a), who study the earmarking of environmental taxes in different electoral competition models with two policy instruments.

focus on electoral competition between parties. Second, they introduce term limits to generate—and test—empirical predictions, while our article is exclusively theoretical.

We develop a two-dimensional model with endogenous parties based on Levy (2004). There is a continuum of voters who differ according to two traits: their income and their concern for the environment. Each trait can take two values, so that there are four groups of people. There are two goods in the economy, the *numeraire* and a polluting good. Public policy consists of two linear tax rates: one on income, and one on the consumption of the polluting good. Tax proceeds are rebated lump sum to all citizens. Public policy is the result of electoral competition between political parties. This can be viewed as a two-stage process. In the first stage, representatives of the different groups form political parties. In the second stage, these parties simultaneously propose political platforms, composed of an income tax rate and an environmental tax rate, to win the elections. The party that gets a plurality of the votes wins the election and implements its proposed policy. The crucial assumption is that the set of policies that a party can commit to is endogenous. If a party is made of a single type of citizens, then the only proposal it can commit to is their most-preferred policy. On the other hand, if a party is made of citizens of different types, then the party can commit to any policy that belongs to the Pareto set of its founders. An equilibrium political state is a partition of citizens into parties and a vector of electoral platforms such that (i) no citizen has an incentive to split up the party he/she belongs to, or to merge it with another party and (ii) no party can make its members better-off by choosing another electoral platform.

We obtain the following results. The Pigouvian level of the environmental tax rate is never an equilibrium of this game. Surprisingly, the equilibrium environmental tax is larger when there is a minority of green voters than when they form a majority. As a result, we find that a green party, defined as a party that proposes green voters' most-preferred environmental policy, can only be part of an equilibrium political state (i.e., be stable) if there is a minority of green voters. This result suggests that the main reason for the emergence of green parties is not simply to be found in an increase in the number of voters holding green views. Rather, we obtain that, for a green party to be stable, it is necessary for income polarization to be large enough, compared to the saliency of the environmental issue, for the non-green citizens. Finally, we show that a larger income polarization increases the minimum income tax rate proposed by the green party.

Before proceeding, the connections between the present article and our earlier study (Anesi and De Donder 2009) are noteworthy. In the latter, we studied the role of party formation in a similar model where the second dimension is the attitude toward racism instead of an environmental issue. We differ from that article in two main respects. First, the focus of Anesi and De Donder (2009) was to understand why racist policies may emerge when a minority of people hold racist views. We then made strong assumptions on the distribution of types but none on the relative saliency of the two issues. By contrast, in the current article, we make no assumption on the distribution of types (green voters may or may not form a majority) but rather assume that the environmental dimension is secondary to the redistributive one. Second, we adopt a different collective choice model and, more specifically, a different stability concept for political parties. In contrast to Anesi and De Donder (2009) who allowed for devi-

ations to smaller parties only and who considered purely policy-motivated politicians, we allow here for mergers between existing parties and we assume that politicians are both policy and office motivated.

The rest of the article is organized as follows: Section 2 presents the economic environment and the political equilibrium concept. Section 3 explains in what sense environmental taxation represents a secondary issue. Sections 4 and 5 analyze, respectively, equilibrium taxes and the parties formed at equilibrium. Final remarks are made in Sect. 6.

2 The Model

2.1 The economic environment

There is a large citizenry with total mass equal to one, in an economy with two goods: the *numeraire* and a polluting good, which both are produced at constant marginal cost, normalized to unity.[4] Citizens are differentiated by their exogenous income, $\omega \in \{\omega_\ell, \omega_h\}$, with $\omega_\ell < \omega_h$, and their concern about pollution, $j \in \{g, n\}$: the "green voters" ($j = g$) care for the pollution associated with aggregate consumption of the polluting good, while the others ($j = n$) do not. Following Fredriksson (1997), we assume that the preferences of green voters over the two consumption goods are given by

$$c + V(x) - \alpha\bar{x}, \tag{1}$$

where c and x are individual consumptions of the *numeraire* and the polluting good, respectively, $\bar{x}$ is the aggregate consumption of the polluting good, and $\alpha \in (0, 1)$ is a parameter that measures the intensity of the green voters' concerns about pollution. This intensity is assumed to be the same for all green voters. The utility of a non-green voter is simply given by

$$c + V(x). \tag{2}$$

All the individuals have the same taste for individual consumption of the polluting good, which is represented by the continuous function V with the following properties: $V(0) \geq 0$, $V' > 0$, $V'' < 0$, $\lim_{x\to 0} V'(x) = \infty$, and $\lim_{x\to\infty} V'(x) < 1$.

Let $\Theta \equiv \{\omega_\ell, \omega_h\} \times \{g, n\}$ be the type space, with generic element $\theta_i^j = (\omega_i, j)$. The fraction of the population that is of type θ_i^j is μ_i^j, where $\mu_i^j < 1/2$ for every $i = \ell, h$ and $j = g, n$. The proportion of voters with income level $\omega_i, i = \ell, h$ is denoted by $\mu_i = \mu_i^g + \mu_i^n$, whereas the proportions of green and non-green voters, are given by $\mu^g = \mu_\ell^g + \mu_h^g$ and $\mu^n = \mu_\ell^n + \mu_h^n$), respectively. Let $\bar{\omega} \equiv \mu_\ell\omega_\ell + \mu_h\omega_h$ be the aggregate income, and assume as usual that the median income is below the average ($\mu_\ell > 1/2$).

[4] This assumption is made to simplify the exposition. All the results carry through to the case of varying marginal costs.

The policy that voters must choose is composed of a proportional income tax, $t \in [0, 1]$, and an environmental tax on the consumption of the polluting good, $e \in [0, 1]$. Tax revenues are used to finance a lump sum transfer to all citizens, which is then determined as a residual: $T = t\bar{\omega} + e\bar{x}$. Once a public policy (t, e) has been decided, citizens choose the consumption level that maximizes their direct utility ((1) for green and (2) for non-green citizens) subject to the individual budget constraint

$$c + (1 + e)x \leq (1 - t)\omega + T.$$

Solving the consumers' problem leads to the following characterization of the demand for the polluting good, $x(e)$:

$$V'(x) \equiv 1 + e.$$

Each individual's choice is too small to affect the average quantity of the polluting good, $\bar{x}$, so that with quasi-linear preferences they all end up consuming the same amount of the good,[5] and $\bar{x}(e) = x(e)$. After appropriate rearrangements, the policy preferences of an individual of type (ω, j) can be represented by the following indirect utility function:

$$u(t, e, \omega, j) \equiv \begin{cases} \omega + t(\bar{\omega} - \omega) + V(x(e)) - (1 + \alpha)x(e) & \text{if } j = g \\ \omega + t(\bar{\omega} - \omega) + V(x(e)) - x(e) & \text{if } j = n. \end{cases} \tag{3}$$

It is easy to obtain individual θ_i^j's most-preferred policy. Obviously, in the absence of incentive effects from income taxation, poor voters favor income confiscation ($t = 1$), whereas rich voters prefer *laissez-faire* ($t = 0$). As for the environmental policy, non-green voters dislike any form of environmental taxation ($e = 0$), while green voters's most-preferred tax rate is equal to the intensity of their dislike of pollution ($e = \alpha$). Observe that, in our setting, the most-preferred environmental tax of an individual is independent of her income. This is due to the fact that all individuals consume the same quantity of the polluting good, so that environmental taxation is not redistributive.[6]

As nobody would prefer to increase e above α, without loss of generality, we restrict the policy space to be $P = [0, 1] \times [0, \alpha]$, with generic element (t, e). In this economy, the collective choice of a public policy (t, e) is made through electoral competition between endogenous political parties. We now turn to the description of the electoral competition side of the model.

[5] We assume that even poor individuals have income (or unmodeled wealth) large enough to consume that amount.

[6] Modeling a situation where richer people consume more of the polluting good would be more in line with reality (see Cremer et al. (2007), Sect. 4, for the case of energy consumption), but would complexify the analysis by adding a redistributive component to environmental taxation. Moreover, this added complexity would not generate additional insight about the political phenomena under study.

2.2 Political parties and elections

The electoral competition model is adapted from Levy (2004), to which the reader is referred for an in-depth discussion of the basic assumptions.

Each group of voters is represented by a single politician who is a perfect representative of her group, in that her policy preferences are given by (3). Politicians running alone are unable to commit to any proposal differing from their ideal policy. The key assumption of Levy (2004) is, however, that politicians can credibly commit to a larger set of policies by forming political parties (or *coalitions*, to use the language of game theory): the set of policies which a party can commit to is the Pareto set of its members. Formally, a *politician* is an element θ of Θ while a *party* is a non-empty subset S of Θ. A policy $(t, e) \in P$ is in the Pareto set of party S, denoted by P_S, if there is no other policy $\left(t', e'\right)$ such that $u\left(t', e'; \theta\right) \geq u(t, e; \theta)$ for all $\theta \in S$ and $u(t', e'; \hat{\theta}) > u(t, e; \hat{\theta})$ for some $\hat{\theta} \in S$.

The political game that we study has two steps: The first step is one of party formation, while the second step encompasses electoral competition, where all parties simultaneously choose a feasible policy and compete in a winner-takes-all election. Let us now describe how each step takes place, beginning with the electoral competition game.

Electoral competition

A party structure is a partition of Θ into parties (i.e., let us assume that all citizens belong to a political party). Let Π be the set of party structures. Let us further assume that the result of the party formation stage is some arbitrary party structure $\pi \in \Pi$. Elections then proceed as follows. Every party $S \in \pi$ chooses an *electoral strategy* (or *platform*), namely, a policy $(t_S, e_S) \in P_S \cup \{\varnothing\}$, where $\varnothing$ means that the party proposes no policy (we say that it does not run). In the case where no party runs for election, every politician receives a zero payoff. If at least one party runs, then we assume that voters record their preferences sincerely over any list of candidate platforms, $\mathbf{p} \equiv ((t_S, e_S))_{S \in \pi}$, and that the election is by plurality rule with no abstention.[7] The election outcome is then a fair lottery between the policies in

$$W(\mathbf{p}) \equiv \left\{ (t_S, e_S) : S \in \arg\max_{S' \in \pi} V_{S'}(\mathbf{p}) \right\},$$

where $V_{S'}(\mathbf{p})$ denotes party S' 's realized vote share. We assume that parties prefer not running to proposing a policy that will lose for sure.

Let $\psi_\theta(S)$ be the indicator function on 2^Θ taking on the value of 1 if $\theta \in S$ and 0 otherwise. Members of the winning party equally share an (arbitrarily small) non-policy benefit $\beta > 0$ (ego-rents, perks of office, etc.). As a consequence, the expected utility of politician θ resulting from a profile of electoral strategies $\mathbf{p}$ is given by

[7] Voters who are indifferent among several policies use a fair mixing device.

$$U(\mathbf{p},\theta) \equiv \frac{1}{|W(\mathbf{p})|} \sum_{(t_S,e_S)\in W(\mathbf{p})} \left[u\left(t_S, e_S, \theta\right) + \psi_\theta(S)\frac{\beta}{|S|} \right]$$

if there is at least one party $S \in \pi$ such that $(t_S, e_S) \neq \varnothing$, and $U(\mathbf{p},\theta) = 0$ otherwise.

Given a party structure $\pi \in \Pi$, a vector of electoral strategies $\mathbf{p} = ((t_S, e_S))_{S\in\pi}$ is a *π-equilibrium* of the electoral-competition game if, for all $S \in \pi$, there is no $\left(t'_S, e'_S\right) \in P_S \cup \{\varnothing\}$, $\left(t'_S, e'_S\right) \neq (t_S, e_S)$, that satisfies

$$U\left(\left(t'_S, e'_S\right), \mathbf{p}_{-S}; \theta\right) \geq U\left(\left(t_S, e_S\right), \mathbf{p}_{-S}; \theta\right)$$

for all $\theta \in S$, with at least one strict inequality. Let $\delta(\pi)$ be the set of π-equilibrium policy outcomes.[8]

Stability of party structures

Up to this point, we have taken the party structure π as given. We now turn to the party formation stage and ask whether π is a stable party structure. First of all, note that there may exist multiple π-equilibria, and therefore multiple equilibrium outcomes ($\delta(\pi)$ may not be a singleton). Thus, π may satisfy stability conditions for one electoral outcome but not for others. As a consequence, the stability of π will not be studied alone, but along with the stability of pairs $(\pi, \mathbf{p})$ where $\mathbf{p}$ is a π-equilibrium. They will be referred to as political states. Which of these should be considered as the set of equilibrium outcomes for the present model? The answer to this question depends on the stability requirements imposed on party structures. The stability condition that we adopt in this article is bi-core stability (Levy 2004).[9]

Let π and π' be two party structures. π' is said to be *induced* from π if π' is formed by either breaking a party in π into two, or by merging two existing parties in π into one. Forming a new party made up of subsets of current parties is excluded on the basis that nobody would trust a politician who is willing to betray her current partners. Mergers involving more than two existing parties are also excluded but, in the context of this article, this restriction does not affect the results and is only maintained for expositional simplicity.

A political state $(\pi, \mathbf{p})$ is *blocked* by another political state $\left(\pi', \mathbf{p}'\right)$ if there exists $S \subseteq \Theta$ such that (1) S can induce π' from π, and (2) for every $\theta \in S$, $U\left(\mathbf{p}', \theta\right) > U\left(\mathbf{p}, \theta\right)$. We are now ready to define equilibrium political states:

Definition 1 Let $\pi^* \in \Pi$ be a party structure and $\mathbf{p}^*$ a profile of electoral strategies for π^*. The pair $(\pi^*, \mathbf{p}^*)$ is an equilibrium political state (EPS) if (i) $\mathbf{p}^*$ is a π^*-equilibrium, and (ii) there is no political state $(\pi, \mathbf{p})$ that blocks $(\pi^*, \mathbf{p}^*)$.

[8] Any profile of electoral strategies induces an electoral outcome which, because of the possibility of a tie, may be a lottery between several policies. As a consequence, $\delta(\pi)$ is a subset of the family of fair lotteries over P. Throughout the article, we write $\langle x_1, \ldots, x_n\rangle$ for the random mixture between policies $x_1, \ldots, x_n$, but simply use x instead of $\langle x\rangle$.

[9] An extensive discussion of this solution concept is offered in the last section of the appendix.

Thus, an equilibrium situation is defined as one that meets two requirements: first, given the equilibrium party structure π^*, no party $S \in \pi^*$ can make all its members better-off by deviating from its equilibrium announcement to a different platform in $P_S \cup \{\varnothing\}$; second, given the equilibrium platform profile $\mathbf{p}^*$, parties in π^* are stable in the sense that no coalition of politicians can make all its members strictly better-off by inducing another political state.

We are now in a position to apply this political equilibrium concept to our economic environment.

3 Environmental policy as a secondary issue

Before we turn to the formal characterization of political equilibria, let us first define what we mean by environmental policy being a "secondary issue" compared to redistribution. Unlike List and Sturm (2006), our definition is not related to the number of people caring for the environment, but rather to the intensity of their preferences. To make this point more formally, some additional notation will prove handy. Since the indirect utility functions (1) and (2) are separable in t and e, let us denote by Δ^j the difference in utility level, for an individual of type θ_i^j, $j = g, n$, between her most-preferred and her least-preferred environmental policy in the policy space P—i.e.

$$\Delta^g \equiv u(t, \alpha, \omega, g) - u(t, 0, \omega, g) = V(x(\alpha)) - V(x(0)) - (1+\alpha)[x(\alpha) - x(0)],$$
$$\Delta^n \equiv u(t, 0, \omega, n) - u(t, \alpha, \omega, n) = V(x(0)) - V(x(\alpha)) - [x(0) - x(\alpha)].$$

Similarly, let us denote by Δ_i the difference in utility level, for an individual of type θ_i^j, $i = \ell, h$, between her most-preferred and her least-preferred income taxation policy in the policy space P—i.e.

$$\Delta_\ell \equiv u(1, e, \omega_\ell, j) - u(0, e, \omega_\ell, j) = \bar{\omega} - \omega_\ell = \mu_h(\omega_h - \omega_\ell),$$
$$\Delta_h \equiv u(0, e, \omega_h, j) - u(1, \alpha, \omega_h, j) = \omega_h - \bar{\omega} = \mu_\ell(\omega_h - \omega_\ell).$$

For future reference, note that Δ_ℓ (and similarly $\mu_\ell \Delta_\ell$) can be seen as a measure of income polarization, namely a measure of the saliency of the conflict between the rich and the poor. In the spirit of Esteban and Ray's (1994) original definition, polarization should indeed rise, as inequality $(\omega_h - \omega_\ell)$ increases and the sizes of the two groups become closer to each other ($\mu_h \to 1/2$).

We impose the following restriction on preferences: For all individuals θ_i^j, the difference in utility level from moving from the least-preferred to the most-preferred *taxation* policy is larger than the difference in utility from moving from the least-preferred to the most-preferred *environmental* policy. Formally, let us impose the following assumption:[10]

A1 $\max\{\Delta^g, \Delta^n\} < \Delta_\ell$

[10] Observe that $\Delta_l < \Delta_h$, because, by assumption, $\mu_l > 1/2$.

Assumption A1 is the precise statement that environmental policy is a secondary issue compared to redistributive policy. This assumption imposes restrictions on preferences over extreme policy bundles. It guarantees that every citizen prefers a policy bundle comprising her ideal redistributive and worst environmental policies to a bundle involving her worst redistributive and ideal environmental policies. For instance, the non-green rich prefers no redistribution accompanied with a high pollution tax to the total confiscation of their income without pollution tax. By this assumption, we do not deny that there may exist people who would be ready to give up all their resources for higher pollution taxes, but we assume their mass is not electorally significant.[11]

4 Environmental taxes

We start with the benchmark case where there is no party formation.[12]

Lemma 1 *Let* $\pi^0 \equiv \left\{\{\theta_h^g\}, \{\theta_\ell^g\}, \left\{\theta_\ell^n\right\}, \left\{\theta_h^n\right\}\right\}$. *Suppose A1 holds, and* $\mu^g \neq 1/2$.[13] *Then*

$$\delta\left(\pi^0\right) = \begin{cases} \{(1,\alpha)\} & \text{if } \mu^g > 1/2, \\ \{(1,0)\} & \text{if } \mu^g < 1/2. \end{cases}$$

In the absence of party formation, our model boils down to the standard citizen–candidate framework proposed by Osborne and Slivinski (1996),[14] where the only credible proposal by any citizen is her own most-preferred policy. In that case, the set of feasible policies is restricted to {(1,0),(1,α),(0,0),(0,α)}, and assumption A1 guarantees the existence of a transitive majority voting ordering over this set. Since poor outnumber rich citizens, any policy with income confiscation gets a majority compared to any policy with laissez-faire. If green voters outnumber non-green ($\mu^g > 1/2$), for any tax policy, a policy with $e = \alpha$ is favored by a majority to a policy without environmental tax ($e = 0$). The policy $(1, \alpha)$ is then a Condorcet winner among the four possible policies (i.e., it beats any other feasible option through pairwise majority comparisons). In the case where $\mu^g < 1/2$, the Condorcet winning policy is $(1, 0)$.

It is easy to see that the Condorcet winning policy is an equilibrium of the electoral competition game with partition π^0. More precisely, the candidate most preferring the Condorcet winner runs unopposed and obtains her most-preferred policy since, by definition, no other candidate can run with a different policy and defeat the Condorcet winner. The less easy part to prove in Lemma 1 is that there is no other equilibrium. To do that, cases where more than one candidate runs (i.e., proposes her most-preferred

[11] List and Sturm (2006) report that the number of members in the three largest environmental organizations (Greenpeace, the Sierra Club and the National Wildlife Federation) between 1987 and 2000 varies from a minimum of 0.25 percent of the population in Mississippi to a maximum of just over 2 percent in Vermont.

[12] All proofs are relegated to the Appendix.

[13] We assume away the case where $\mu^g = 1/2$ first because this is a knife-edge situation, but, mainly because considering this case increases considerably the length of the proofs without adding any new insight. For the interested reader, with $\mu^g = 1/2$ we obtain that $\delta(\pi^0) = \{\langle(1,\alpha),(1,0)\rangle\}$ if $\mu_\ell^g = \mu_\ell^n > \mu_h$, and $\emptyset$ otherwise. The proof is available upon request.

[14] The model proposed by Besley and Coate (1997) differs in that it assumes that voters behave strategically.

policy)are in turn considered, and we show that they can not constitute equilibria. We now briefly summarize how we proceed in the proof to give the reader a better feeling as to how the electoral competition stage gets solved in our model.

Since there is a strict transitive majority voting ordering over the four feasible policies, it is impossible for two candidates to run at equilibrium and to tie. Given our assumption that a losing candidate/party prefers no to run, we can rule out any situation with two candidates running. The same intuition carries through to the case where the four candidates run: given that poor outnumber rich voters, one rich candidate loses for sure if they run, and thus prefers not to run. This leaves only the possibility that three candidates run at equilibrium. Given that poor voters form a majority, it is impossible to have a three-way tie with two rich candidates running. With two poor and one rich candidates running, we show in the Appendix that one poor candidate has an incentive not to run to guarantee that the other poor candidate will win for sure. This is because, by assumption A1, a poor citizen prefers the policy favored by the poor citizen–candidate of the other environmental type to a random mixture between the three original policies. This proves that the only equilibrium under π^0 has one candidate running with the Condorcet winning policy. We then obtain the very intuitive result that, *in the absence of party formation, the pollution tax is larger when there is a majority of green voters in the electorate.*

We now turn to party formation. The main incentive to form a party is to enlarge the set of policies that may credibly be proposed to the voters. Figure 1 depicts the Pareto set of all potential parties.[15] Intuitively, parties made exclusively of rich and poor green (resp., non-green) citizens may credibly propose any income tax rate ($t \in [0, 1]$) provided that it is coupled with the maximum (resp. minimum) preferred environmental tax rate $e = \alpha$ (resp., $e = 0$). Similarly, a party made exclusively of green and non-green poor (resp., rich) citizens may credibly propose any environmental tax rate ($e \in [0, \alpha]$) provided they also propose full confiscation, $t = 1$ (resp., *laissez-faire*, $t = 0$).

As for parties with two opposite types ($\{\theta_\ell^n, \theta_h^g\}$ and $\{\theta_\ell^g, \theta_h^n\}$), observe that the environmental tax rates associated to interior income tax rates differ according to which opposite types compose the party. This is due to the fact that, with a majority of poor voters ($\mu_\ell > 1/2$), rich citizens care more about income tax policy than poor citizens (in the sense that $\Delta_h > \Delta_\ell$, as noted in footnote 10). A rich non-green citizen will then compromise more on environmental policy (i.e., accepts (t, e) with $e > \alpha/2$ and $0 < t < 1$ when forming a party with the poor green citizen) than a poor non-green citizen, who will insist on a low value of e ($e < \alpha/2$ for $0 < t < 1$) when joining forces with rich green citizens. Formally, when $0 < t < 1$, we have that $(t, \alpha\mu_\ell) \in P_{\{\theta_\ell^g, \theta_h^n\}}$, while $(t, \alpha\mu_h) \in P_{\{\theta_\ell^n, \theta_h^g\}}$. Finally, the Pareto set of parties composed of three types can easily be obtained from the previous case; and the Pareto set of the four-type party is equal to the feasible set P.

We now proceed to a comparison between EPS when green voters are a majority and when they are not. However, before we state those results, the following remark is in order. While EPS always exist in this model (see, for instance, the EPS described

[15] See Appendix for the notation used in figure 1.

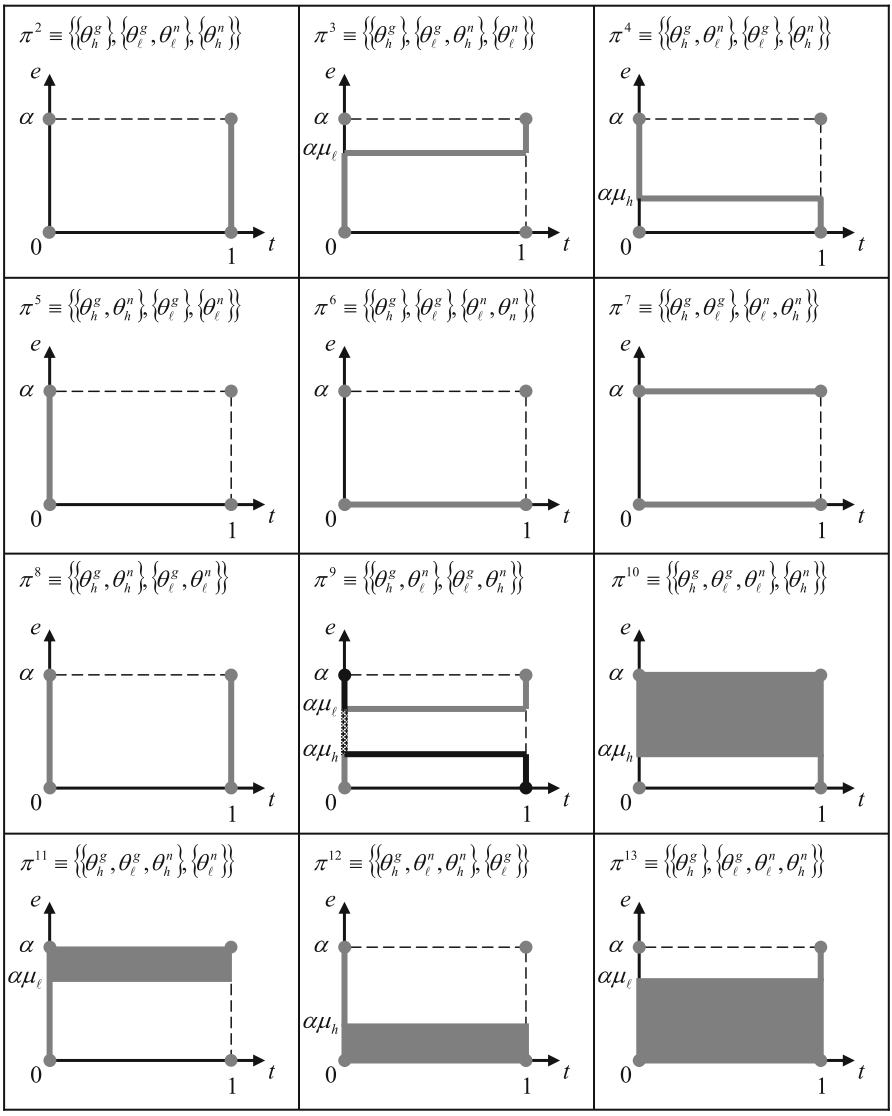

Fig. 1 Pareto sets of all potential parties

immediately after footnote 16), unicity is far from being guaranteed. Therefore, to make the equilibrium comparison meaningful, any statement about one or several equilibrium policies must be true for all the equilibria associated to the case under consideration. We start with the case where a majority of citizens are green.

Proposition 1 *Suppose A1 holds. If $\mu^g > 1/2$, then any environmental tax rate e^* that emerges in an EPS must satisfy $e^* \leq \alpha\mu_h$.*

The intuition for this result runs as follows. The poor green citizens are in a position of power since there are more poor than rich voters, and more green than non-green voters. As Lemma 1 shows, poor green citizens obtain their most favored policy when no party forms. This hinders the formation of any party containing poor green candidates. Take for instance a party composed of both green politicians. Poor green voters have a double incentive to disband such a party: they would not have to compromise on the income taxation issue and moreover they would not have to share the spoils of office (however small β is) with their partner.

On the other hand, poor green citizens are not powerful enough to win against all others because, by assumptions, $\mu_l^g < 1/2$. As a result, the citizen-candidate equilibrium depicted in Lemma 1 is not an EPS, since the rich green citizens have an incentive to form a party together with the poor non-green citizens to propose a compromise policy (a positive but not extreme income tax coupled with a low but positive environmental tax) that they both prefer to $(1, \alpha)$ and that obtains a majority of votes against $(1, \alpha)$.[16] Moreover, it can be shown that the party formed of θ_h^g and θ_l^n, proposing $(t, \alpha\mu_h)$ for some $0 < t < 1$, and running unopposed is an EPS. In a nutshell, the poor green voters are "too powerful" to form a stable party but "not powerful enough" to guarantee themselves against other parties. We then obtain that poor green voters can not obtain their most-preferred environmental policy ($e = \alpha$), and moreover that there is no EPS where the environmental tax is larger than $\alpha\mu_h$.

This reasoning does not carry through to the case where green citizens form a minority ($\mu^g < 1/2$). In that case, poor green voters have no incentive to break a party made of rich as well as poor green voters, but on the contrary have an incentive to join forces to increase environmental taxation. On the contrary, poor non-green voters become powerful enough that they are reluctant to form a party. This explains why the allocation described in the previous paragraph is not an EPS when $\mu^g < 1/2$, since the poor non-green voter would disband the party formed with the rich green candidate to win outright with her most-preferred policy (1,0). The next Proposition shows that all equilibrium political states exhibit a large environmental component ($e \geq \alpha\mu_\ell$) in that situation.

Proposition 2 *Suppose A1 holds. If $\mu^g < 1/2$, then any environmental tax rate $\bar{e}$ that emerges in an EPS must satisfy $\bar{e} \geq \alpha\mu_\ell$.*

Combining Propositions 1 and 2, we obtain the following surprising result: *Due to the party formation process, the environmental tax rate that emerges in a political equilibrium is larger when there is a minority of green voters.* This result illustrates very starkly that, given our modeling of party formation and electoral competition, an increase in the proportion of green voters need not result in more environmentally friendly policies. Another immediate consequence of Propositions 1 and 2 is that environmental quality is better when there is a minority of green voters.

Turning to the normative properties of EPS, observe first that, in our quasi-linear setting without income tax distortions, a utilitarian planner is indifferent between all values of the income tax rate. The optimal utilitarian environmental tax rate is given by its Pigouvian level, $e^* = \alpha\mu^g$. This Pigouvian level belongs to the Pareto sets of the

[16] This statement is made formally in Lemma 2, which is presented and proved in the Appendix.

grand four-member party, of several three-member parties, and also of two-member parties in the special case where $\mu^g \in \{\mu_h, \mu_\ell\}$.

We then obtain the following corollary to Propositions 1 and 2.

Corollary 1 *Every EPS is inefficient, in the sense that the Pigouvian tax rate $\alpha\mu^g$ is never implemented in equilibrium.*

The inefficiency of every EPS is driven by the link between the proportion of green voters and the equilibrium environmental tax rate. The presence of a majority of green citizens calls for a large Pigouvian tax ($e > \alpha/2$) but generates an equilibrium with a low tax rate ($e < \alpha/2$), and vice versa when green citizens form a minority.

To summarize, the main conclusion to draw from the discussion up to this point is the following: When party formation is taken into consideration, the explanation for the emergence of green policies is not to be found in an increase in the proportion of green voters. The next section will show that other factors, such as the saliency of the environmental issue and the income polarization may play an important role in explaining the emergence of green parties/policies.

5 Stable green parties

We now address the question of the existence of a green party, which is defined as a party offering the ideal environmental policy of green citizens. Formally, party $S \subseteq \Theta$ is a *stable green party* if there exists an EPS $(\pi, \mathbf{p})$ such that $S \in \pi$ and $e_S = \alpha$. We already know from the previous section that a stable green party exists only if there is a minority of green voters. The next proposition goes further.

Proposition 3 *A stable green party $S \subseteq \Theta$ exists only if $S = \left\{\theta_h^g, \theta_\ell^g\right\}$ and the following conditions hold: (a) $\mu^g < 1/2$, (b) $\mu_\ell \Delta_\ell \geq \Delta^n$, and (c) $\mu_\ell \Delta^g \geq \mu_h \Delta^n$. Furthermore, if (c) is replaced by $\mu_h \Delta^g > \mu_\ell \Delta^n$, then $\left\{\theta_h^g, \theta_\ell^g\right\}$ is a stable green party.*

The only class of green party that may exist at equilibrium (i.e., be stable) is composed of the two types of green voters. This is a consequence of the political power of the poor non-green candidate, who belongs simultaneously to the majority of poor voters ($\mu_\ell > 1/2$) and to the majority of non-green voters ($\mu^g < 1/2$). On one hand, the poor non-green candidate does not wish to constitute a party with a green candidate, since she is powerful enough alone. On the other, she has enough electoral power to defeat a party composed of green and non-green citizens that would run against him. The only stable green party must then be made of the two green voters (who would not want to share power with and accommodate a third type) running against the two separate non-green candidates.

In words, Proposition 3 establishes three necessary conditions for the green party $\left\{\theta_h^g, \theta_\ell^g\right\}$ to be stable:

(a) The green citizens form a minority;
(b) Income polarization, measured by $\mu_\ell \Delta_\ell$, is large enough compared to the saliency of the environmental issue for the non-green citizens, measured by Δ^n;

(c) The saliency of the environmental policy for the green citizens is sufficiently large compared to the saliency of this issue for the non-green citizens.

The intuition for condition (a) is familiar from the previous subsection: Proposition 1 establishes that equilibrium environmental taxes cannot exceed $\alpha\mu_h$ when the green citizens form a majority. However, even when the green citizens form a minority, the green party has still to guard itself against two dangers, one external and one internal to the party.

The external danger is the majority coalition formed by the non-green voters. We show that such a threat can only be countered if condition (b) holds. Suppose, to the contrary, that income polarization is relatively low. This weakens the redistributive conflict between rich and poor non-green politicians, thereby causing the defeat of the green party: a non-green candidate can run and win the elections against the green party by getting the votes of all non-green voters (who outnumber green voters), or the non-green politicians can compromise on the redistributive issue and form a party that defeats the green party.

The internal danger faced by the green party consists in one of its two members being wooed away by the policy of a non-green candidate. More precisely, if the poor green and non-green voters were to both prefer the policy $(1, 0)$ to the compromise policy (t, α) proposed by the green party, then the policy $(1, 0)$ would be proposed by the poor non-green candidate who would win the elections for sure since poor voters form a majority. We show in the proof of Proposition 3 that this threat does not materialize if condition (c) is satisfied.

By making it easier for the green politicians to compromise on the income tax rate than for the non-green politicians, condition (c) has another effect on the stability of the green party. It also ensures that, whenever it faces a party made of the two non-green politicians, the green party can always find a policy that attracts some non-green voters and then defeat its opponent. Combined with (b), condition (c) therefore guarantees that non-green cannot coalesce in a party to defeat the green party.

In the first part of Proposition 3, (a), (b), and (c) establish only necessary conditions for the existence of a stable green party. However, the second part of the Proposition reveals that reinforcing (c) suffices to obtain existence.

Proposition 3 has interesting implications in terms of the electoral alliances between green and non-green voters and in term of the policies proposed by green parties.

First, as we explained above, Proposition 3 shows that a large enough income polarization is necessary for the emergence of a stable green party. Moreover, increasing income polarization dampens the external threat to the existence of the green party and also increases the minimum tax rate that the green party needs to propose to fend off the internal threat—i.e., to prevent the poor green citizen from siding with the poor non-green citizen rather than supporting (t, α). This is illustrated in Fig. 2, where the set of policies (t, α) that are preferred by the poor green citizen to policy $(1, 0)$ shrinks (i.e., $1 - \Delta^g/\Delta_\ell$ increases) as income polarization—represented here by Δ_ℓ—increases relative to the saliency of the environmental issue for green voters. The role of income polarization is then summarized in the following

Remark 1 A large enough income polarization, namely $\mu_\ell \Delta_\ell \geq \Delta^n$, is necessary to have a stable green party. Furthermore, when this condition holds, the minimum

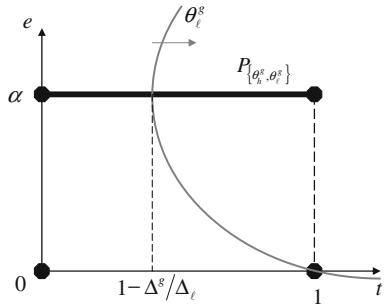

Fig. 2 Left-wing orientation of stable green parties

equilibrium income tax rate proposed by the green party increases and converges to one as income polarization becomes arbitrarily large.

Note that this result is in line with the fact that, according to the empirical evidence, green parties are overwhelming associated with strong redistributive concerns (see Neumayer 2004, and the many references therein).

A second implication of Proposition 3 is that a party involving a non-green politician never offers a green policy ($e = \alpha$) in equilibrium. Therefore, our model predicts that "Red-Green alliances" and (less common) "Blue-Green alliances" between green politicians and leftist/rightist non-green politicians typically fail to deliver the green type's most-preferred environmental policy.

A last implication is that a situation with two green parties is not stable. This is also in line with real world experience, where situations with several green parties coexisting (as in France in the 1990s) do not persist for long.

6 Conclusion

In this article, we have built a model of electoral competition to assess the interplay between political party formation and environmental taxation viewed as secondary to the income taxation. We have defined being a secondary issue in terms of the intensity of preferences over this issue rather than in terms of numbers of voters who care for it. We have built on Levy (2004) for the political equilibrium concept, defined as the solution to a two stage game where politicians first form parties and where parties then compete by choosing a policy bundle to win the elections.

The first two propositions together establish that the equilibrium environmental tax is larger when the green voters represent a minority of the electorate than when they form a majority. The main driving force behind this result is that, when green voters form a majority, they are electorally too powerful to compromise and form a party with other citizens, but not powerful enough to prevent other types from merging into a party and defeating them. Observe that this result is very different from what we would obtain with, for instance, a median voter approach applied sequentially to the two dimensions. In that case, a majority of poor and green voters would simply translate into a confiscatory policy coupled with a large environmental tax rate. Contrasting these results shows the importance of taking into account the endogeneity of

the political parties, both in terms of number of parties and of their constituency. It also shows very starkly that, at least within the confines of our model, the reason for the emergence of green parties and policies is not to be found in an increase in the proportion of voters who care for the environment. Rather, as Proposition 3 illustrates, the saliency of the environmental issue and income polarization play an important role in explaining why green parties and green policies emerge at equilibrium.

More precisely, our model suggests the existence of a positive relationship between income polarization and the existence of, and the degree of income redistribution proposed by, green parties. Although there exists a literature showing that higher income inequality is associated with worse environmental results (see, for instance, Boyce (1994); Heerink et al. (2001)), we are aware of no study linking income inequality and green parties' policies. We hope this article contributes to drawing the attention of applied researchers to this issue.

The remaining results of the article regarding the existence, the number and the policies of green parties, follow probably more closely our intuition. We obtain that there can only be one stable party, which is made of both types of green voters, who bargain over redistribution, but agree on environmental policy. This means that a situation with two green parties differing in redistributive policy, such as that experienced by France in the 1990s for instance, is not stable. We also obtain that green parties are associated with large redistribution, in the sense that there exists a lowerbound on the income tax rate proposed at equilibrium by any green party. This is in line with the numerous empirical evidence, surveyed by Neumayer (2004), which shows that green parties are located to the left on the redistributive dimension.

Acknowledgments This paper has been presented at the University of Illinois at Urbana-Champaign, Yale, Nottingham, Université de Montréal and at the 2008 Political Economy of Bargaining conference in Waterloo. The authors thank participants for their comments, as well as the three referees for very helpful suggestions. This article has been written while the second author visited Yale University. He thanks Yale Economics Department for its hospitality.

Appendix

Throughout this appendix, the following notation will be used:

$$\pi^1 \equiv \left\{\left\{\theta_h^g, \theta_\ell^g\right\}, \left\{\theta_\ell^n\right\}, \left\{\theta_h^n\right\}\right\}, \quad \pi^2 \equiv \left\{\left\{\theta_h^g\right\}, \left\{\theta_\ell^g, \theta_\ell^n\right\}, \left\{\theta_h^n\right\}\right\}$$
$$\pi^3 \equiv \left\{\left\{\theta_h^g\right\}, \left\{\theta_\ell^g, \theta_h^n\right\}, \left\{\theta_\ell^n\right\}\right\}, \quad \pi^4 \equiv \left\{\left\{\theta_h^g, \theta_\ell^n\right\}, \left\{\theta_\ell^g\right\}, \left\{\theta_h^n\right\}\right\}$$
$$\pi^5 \equiv \left\{\left\{\theta_h^g, \theta_h^n\right\}, \left\{\theta_\ell^g\right\}, \left\{\theta_\ell^n\right\}\right\}, \quad \pi^6 \equiv \left\{\left\{\theta_h^g\right\}, \left\{\theta_\ell^g\right\}, \left\{\theta_\ell^n, \theta_h^n\right\}\right\}$$
$$\pi^7 \equiv \left\{\left\{\theta_h^g, \theta_\ell^g\right\}, \left\{\theta_\ell^n, \theta_h^n\right\}\right\}, \quad \pi^8 \equiv \left\{\left\{\theta_h^g, \theta_h^n\right\}, \left\{\theta_\ell^g, \theta_\ell^n\right\}\right\}$$
$$\pi^9 \equiv \left\{\left\{\theta_h^g, \theta_\ell^n\right\}, \left\{\theta_\ell^g, \theta_h^n\right\}\right\}, \quad \pi^{10} \equiv \left\{\left\{\theta_h^g, \theta_\ell^g, \theta_\ell^n\right\}, \left\{\theta_h^n\right\}\right\}$$
$$\pi^{11} \equiv \left\{\left\{\theta_h^g, \theta_\ell^g, \theta_h^n\right\}, \left\{\theta_\ell^n\right\}\right\}, \quad \pi^{12} \equiv \left\{\left\{\theta_h^g, \theta_\ell^n, \theta_h^n\right\}, \left\{\theta_\ell^g\right\}\right\}$$
$$\pi^{13} \equiv \left\{\left\{\theta_h^g\right\}, \left\{\theta_\ell^g, \theta_\ell^n, \theta_h^n\right\}\right\}, \quad \pi^{14} \equiv \left\{\left\{\theta_h^g, \theta_\ell^g, \theta_\ell^n, \theta_h^n\right\}\right\}$$

Proof of Lemma 1

Let $\succeq^m$ stand for the majority preference relation, and let $\succ^m$ and $\sim^m$ be its asymmetric and symmetric parts, respectively. Under Assumption A1, this relation is a transitive linear order over the set of politicians' ideal policies whenever $\mu^g \neq 1/2$:

If $\mu^g > 1/2$: $(1, \alpha) \succ^m (1, 0) \succ^m (0, \alpha) \succ^m (0, 0)$,
If $\mu^g < 1/2$: $(1, 0) \succ^m (1, \alpha) \succ^m (0, 0) \succ^m (0, \alpha)$.

With these useful observations in mind, let us now turn to the determination of π^0-equilibria.

- *One-candidate equilibria*

Consider first π^0-equilibria in which a single party runs. An immediate consequence of the above observations is that $(\varnothing, (1, \alpha), \varnothing, \varnothing)$ (resp., $(\varnothing, \varnothing, (1, 0), \varnothing)$) is the unique π^0-equilibrium in which a single party runs whenever $\mu^g > 1/2$ (resp. $\mu^g < 1/2$). As a consequence, $(1, \alpha) \in \delta(\pi^0)$ when $\mu^g > 1/2$, and $(1, 0) \in \delta(\pi^0)$ when $\mu^g < 1/2$.

- *Two-candidate equilibria*

Suppose that $\mu^g > 1/2$. As $\geq^m$ is a transitive linear order, there is no π^0-equilibrium in which two candidates run against each other. Indeed, one of them would lose for sure in such a situation and, by assumption, would choose not to run. The same argument applies when $\mu^g < 1/2$.

- *Three-candidate equilibria*

Note first that $\mu_\ell > 1/2$ rules out ties between θ_h^g, θ_ℓ^n, and θ_h^n, and between θ_h^g, θ_ℓ^g, and θ_h^n.

Suppose now that the three running parties are $\{\theta_h^g\}$, $\{\theta_\ell^g\}$, and $\{\theta_\ell^n\}$. Such a situation cannot be a π^0-equilibrium. Indeed, party $\{\theta_\ell^n\}$ could deviate to $\varnothing$, thereby enforcing policy $(1, \alpha)$ she strictly prefers to the fair lottery between the policies offered by the three candidates under Assumption A1. A similar argument shows that $\{\theta_\ell^g\}$, $\{\theta_\ell^n\}$, and $\{\theta_h^n\}$ running against each other cannot be a π^0-equilibrium.

- *Four-candidate equilibria*

Our assumption on the distribution of types, namely $\mu_\ell > 1/2$, rules out the case where four candidates tie when running.

In summary, the θ_ℓ^g-politician (resp. θ_ℓ^n-politician) running alone and offering her ideal policy $(1, \alpha)$ (resp. $(1, 0)$) is the unique π^0-equilibrium when $\mu^g > 1/2$ (resp. $\mu^g < 1/2$). This proves the lemma.

Proof of Proposition 1

We start with a series of useful lemmas.

Lemma 2 *Suppose A1 holds. There exists* $t_1 \in (0, 1)$ *such that* $(t_1, \mu_h\alpha) \in \delta\left(\pi^4\right)$, *and*

$$u\left(t_1, \mu_h\alpha, \theta_h^g\right) > u\left(1, \alpha, \theta_h^g\right),$$
$$u\left(t_1, \mu_h\alpha, \theta_\ell^n\right) > u\left(1, \alpha, \theta_\ell^n\right).$$

Proof Note first that $(1, \alpha) \notin P_{\{\theta_h^g, \theta_\ell^n\}}$. From this (and the strict concavity of V), we can infer that there is $t_1 \in [0, 1]$ such that $(t_1, \mu_h\alpha) \in P_{\{\theta_h^g, \theta_\ell^n\}}$ and

$$u\left(t_1, \mu_h\alpha, \theta_h^g\right) > u\left(1, \alpha, \theta_h^g\right)$$
$$u\left(t_1, \mu_h\alpha, \theta_\ell^n\right) > u\left(1, \alpha, \theta_\ell^n\right).$$

Since $(t, \mu_h\alpha) \in P_{\{\theta_h^g, \theta_\ell^n\}}$, for any $t \in [0, 1]$, $(t_1, \mu_h\alpha) \in P_{\{\theta_h^g, \theta_\ell^n\}}$. Consider now party structure π^4 and suppose $\left\{\theta_h^g, \theta_\ell^n\right\}$ runs alone and offers $(t_1, \mu_h\alpha)$. Since the θ_ℓ^g- and θ_ℓ^n-politicians strictly prefer $(t_1, \mu_h\alpha)$ to $(0, 0)$ and $\mu_\ell > 1/2$, $\left\{\theta_h^n\right\}$ cannot profitably deviate by offering $(0, 0)$. Similarly, $\left\{\theta_\ell^g\right\}$ cannot profitably deviate by offering $(1, \alpha)$, for politicians of type θ_h^g, θ_ℓ^n, and θ_h^n all strictly prefer $(t_1, \mu_h\alpha)$ to $(1, \alpha)$. This proves that $(t_1, \mu_h\alpha) \in \delta\left(\pi^4\right)$.

Lemma 3 *Suppose A1 holds. If* $\mu^g > 1/2$, *then* $(1, \alpha) \in \delta\left(\pi^5\right)$ *and there is no EPS involving* π^5.

Proof Given that we assume that parties which are indifferent between running and not running do not run, the first part of the above statement means that there is a π^5-equilibrium which involves party $\left\{\theta_\ell^g\right\}$ running alone. Indeed, the Pareto sets of the other parties in π^5 do not contain $(1, \alpha)$.

To prove the lemma, note that for any $e \in [0, \alpha]$ the policy $(1, \alpha)$ defeats both $(0, e)$ and $(1, 0)$ in a pairwise vote ($\mu^g > 1/2$). As a result, if $\left\{\theta_h^g, \theta_h^n\right\}$ [resp. $\left\{\theta_\ell^n\right\}$] runs alone, and then offers $(0, e)$ [resp. $(1, 0)$], $\left\{\theta_\ell^g\right\}$ can profitably deviate by offering her ideal policy $(1, \alpha)$. Moreover, platform profiles of the form $((0, e), (1, \alpha), \varnothing)$ or $(\varnothing, (1, \alpha), (1, 0))$ cannot be π^5-equilibria since $\left\{\theta_\ell^g\right\}$ wins for sure. For the same reason, $\left\{\theta_\ell^g\right\}$ running alone is a π^5-equilibrium as no other potential candidate can defeat it.

However, party $\left\{\theta_\ell^g\right\}$ running alone in π^5 cannot be an EPS. To see this note that $(1, \alpha)$ is defeated by $(t_1, \alpha\mu_h) \in P_{\{\theta_h^g, \theta_\ell^n, \theta_h^n\}}$ in pairwise vote (see Lemma 2). Therefore, $\left\{\theta_h^g, \theta_h^n\right\}$ should coalesce with $\left\{\theta_\ell^n\right\}$ to induce π^{12}. Doing so, they could indeed implement $(t_1, \alpha\mu_h)$ which makes all of them strictly better-off and share the non-policy benefit β.

Consider now a profile of the form $((0, e), \varnothing, (1, 0))$. An immediate implication of Assumption A1 is that voters of type θ_ℓ^g strictly prefer $(1, 0)$ to $(0, e)$ for any $e \in [0, 1]$. As $\mu_\ell > 1/2$, this implies that $\left\{\theta_\ell^n\right\}$ wins for sure. This is then not an equilibrium situation.

To complete the proof of Lemma 3, it then remains to show that the three parties in π^5 running at the same time is not an EPS. To see this, consider a platform profile

$((0, e), (1, \alpha), (1, 0))$ with $e \in [0, 1]$. As $V(x(\alpha)) - x(\alpha) \leq V(x(e)) - x(e)$, we have

$$\begin{aligned}
&\frac{1}{3}\left[V(x(0)) - x(0) + V(x(\alpha)) - x(\alpha)\right] - \frac{2}{3}\left[V(x(e)) - x(e)\right] \\
&\quad \leq \frac{1}{3}\left[V(x(0)) - x(0) - (V(x(e)) - x(e))\right] \\
&\quad \leq \frac{1}{3}\Delta^n(\alpha) < \frac{1}{3}\mu_h(\omega_h - \omega_\ell) = \frac{1}{3}(\bar{\omega} - \omega_\ell) \qquad (4)
\end{aligned}$$

where the last inequality results from Assumption A1. Rearranging (4), we obtain

$$\begin{aligned}
&\frac{1}{3}\left[2(\bar{\omega} - \omega_\ell) + V(x(e)) - x(e) + V(x(0)) - x(0) + V(x(\alpha)) - x(\alpha)\right] \\
&\quad < \bar{\omega} - \omega_\ell + V(x(e)) - x(e)
\end{aligned}$$

or, equivalently,

$$\frac{1}{3}u\left(0, e, \theta_\ell^n\right) + \frac{1}{3}u\left(1, \alpha, \theta_\ell^n\right) + \frac{1}{3}u\left(1, 0, \theta_\ell^n\right) < u\left(1, e, \theta_\ell^n\right).$$

This means that the θ_ℓ^n-politician strictly prefers the policy $(1, e) \in P_{\{\theta_\ell^g, \theta_\ell^n\}}$ to the fair lottery between $(0, e)$, $(1, \alpha)$, and $(1, 0)$. Using a parallel argument, we can deduce from $\Delta^g(\alpha) < \mu_h(\omega_h - \omega_\ell)$ that the same is true for politician θ_ℓ^g.

As a consequence parties $\{\theta_\ell^n\}$ and $\{\theta_\ell^g\}$ can profitably merge with each other to induce π^8. Indeed, $\mu_\ell > 1/2$ ensures that $\{\theta_\ell^n, \theta_\ell^g\}$ offering $(1, e)$ and winning with probability 1 is a π^8-equilibrium. This proves that there is no ESP involving π^5 and ends the proof of Lemma 3.

Lemma 4 *Suppose A1 holds. Then* $\delta\left(\pi^8\right) = P_{\{\theta_\ell^g, \theta_\ell^n\}}$.

Proof By A1, all poor voters (and politicians) strictly prefer any policy in $P_{\{\theta_\ell^g, \theta_\ell^n\}}$ to any policy in $P_{\{\theta_h^g, \theta_h^n\}}$. As $\mu_\ell > 1/2$, this implies that any policy in $P_{\{\theta_\ell^g, \theta_\ell^n\}}$ beats any policy in $P_{\{\theta_h^g, \theta_h^n\}}$ in a pairwise vote. Thus, a strategy profile is a π^8-equilibrium if, and only if, it is of the form $(\varnothing, (t, e))$ with $(t, e) \in P_{\{\theta_\ell^g, \theta_\ell^n\}}$. This establishes Lemma 4. □

We now return to the main proposition. The idea is to check that, for every $j = 0, \ldots, 14$, the following statement is true:

($\mathbf{P}_k$) Suppose $\mu^g > 1/2$. If (t, e) is a policy that emerges with a positive probability in an EPS $\left(\pi^k, \mathbf{p}\right)$, then $e \leq \mu_h \alpha$.

($\mathbf{P}_k$) is evidently true for $k \in \{1, 2, 3\}$ since we know from Lemma 1 that politician θ_ℓ^g can always profitably induce π^0. Let us now turn to the other party structures.

- $k = 0$

From Lemmas 1 and 2, we immediately see that $\{\theta_h^g\}$ and $\{\theta_\ell^n\}$ can profitably induce π^4 from π^0. This proves ($\mathbf{P}_0$).

- $k = 4$

To show **(P$_4$)**, we have to check that $\{\theta_\ell^g\}$ can never win or tie for winning by offering $(1, \alpha)$, and that $\{\theta_h^g, \theta_\ell^n\}$ can never win or tie for winning by offering a policy of the form $(0, e)$, with $e \in (\mu_h\alpha, 1]$. Note first that a tie between the three parties in π^4 is not consistent with our assumptions on the distribution of voters' types ($\mu_\ell > 1/2$ and $\mu^g > 1/2$). A three-candidate equilibrium is therefore impossible. Moreover, the platform profile $(\varnothing, (1, \alpha), (0, 0))$ cannot be an equilibrium since $\{\theta_\ell^g\}$ wins for sure.

We know that $(t, \mu_h\alpha) \in P_{\{\theta_h^g, \theta_\ell^n\}}$ defeats $(1, \alpha)$ in a pairwise vote (recall the proof of Lemma 2). This guarantees that $\{\theta_\ell^g\}$ can never win with her ideal policy.

Let us now turn to party $\{\theta_h^g, \theta_\ell^n\}$. Under A1, the θ_ℓ^n- and θ_ℓ^g politicians strictly prefer $(1, \alpha)$ to any policy of the form $(0, e)$. This implies that $(1, \alpha)$ is preferred by a majority of voters to any policy $(0, e)$ with $e \in (\mu_h\alpha, 1]$ ($\mu_\ell > 1/2$), which in turn implies that $\{\theta_h^g, \theta_\ell^n\}$ can never win by offering such a policy.

Finally, $\{(1, e), (1, \alpha), \emptyset\}$ with $e \in (\mu_h\alpha, 1]$ cannot be an equilibrium, since a tie between $(1, e)$ and $(1, \alpha)$ would require that θ_h^g prefer the first to the latter, which is impossible.

- $k = 5$

(P$_5$) is a direct consequence of Lemma 3.

- $k = 6$

To show **(P$_6$)**, we have to check that neither $\{\theta_h^g\}$ nor $\{\theta_\ell^g\}$ can win with a positive probability in an EPS involving π^6. We distinguish among several cases:

(i) $\{\theta_\ell^g\}$ running alone and implementing $(1, \alpha)$ cannot be an equilibrium situation as policy $(t_1, \mu_h\alpha)$ (described in Lemma 2) makes politicians (and then voters) of types θ_ℓ^n, θ_h^n, and θ_h^g strictly better-off. Coalitions $\{\theta_\ell^n, \theta_h^n\}$ and $\{\theta_\ell^n\}$ can therefore profitably induce π^{12} to enforce that policy and grasp the non-policy benefit.

(ii) A strategy profile of the form $(\varnothing, (1, \alpha), (t, 0))$ is also impossible in an EPS. Indeed, for this to be possible there should be a tie between the running candidates, namely $\{\theta_\ell^n, \theta_h^n\}$ and $\{\theta_\ell^g\}$. As $\mu_\ell > 1/2$, this would imply that the voters of type θ_ℓ^n prefer $(t, 0)$ to $(1, \alpha)$, and then that $t > 0$. It would also imply that the θ_h^g-voters would be indifferent between $(1, \alpha)$ and $(t, 0)$. However, these last statements are in contradiction with $(\varnothing, (1, \alpha), (t, 0))$ being a π^6-equilibrium. Indeed, party $\{\theta_\ell^n, \theta_h^n\}$ could make all its members better-off by deviating to a platform $(t - \varepsilon, 0)$, with $\varepsilon > 0$ very small. Although the θ_ℓ^n politician would incur a small utility loss, she would be compensated by an increase in the non-policy benefit ($\beta/2$ instead of $\beta/4$) as, by continuity, the change in platform would attract θ_h^g-voters and ensure her party's victory.

(iii) As $(0, \alpha)$ is defeated by $(1, \alpha)$ in a pairwise vote, $\{\theta_h^g\}$ running alone or running against $\{\theta_\ell^g\}$ cannot be equilibrium situations.

(iv) Suppose now the strategy profile is $((0, \alpha), \varnothing, (t, 0))$. For this profile to be a π^6-equilibrium, voters of type θ_ℓ^g must be indifferent between $(0, \alpha)$ and $(t, 0)$. As $\beta > 0$, however, there exists $\epsilon > 0$ sufficiently small such that $\{\theta_\ell^n, \theta_h^n\}$ can

profitably deviate by offering $(t+\epsilon, 0)$. This would allow it to win and then to get β for sure, thus compensating its member of type θ_h^n for the small utility loss caused by the change in platform.

(v) Finally, the three parties in π^6 running at the same time cannot be an EPS. Indeed, coalition $\left\{\theta_\ell^g, \theta_\ell^n, \theta_h^n\right\}$ should deviate to π^{13}. To see this, define the policy (t_2, e_2) as follows:

$$t_2 \equiv \frac{1}{3}(1+t),$$

$$V(x(e_2)) - x(e_2) \equiv \frac{2}{3}[V(x(\alpha)) - x(\alpha)] + \frac{1}{3}[V(x(0)) - x(0)].$$

It is easy to see that (t_2, e_2) is a certainty equivalent of $\langle(0,\alpha), (1,\alpha), (t,0)\rangle$ for both non-green politicians. Now, define e_3 as follows

$$V(x(e_3)) - (1+\alpha)x(e_3) \equiv \frac{2}{3}[V(x(\alpha)) - (1+\alpha)x(\alpha)] + \frac{1}{3}[V(x(0)) - (1+\alpha)x(0)].$$

By definition, (t_2, e_3) is a certainty equivalent of $\langle(0,\alpha), (1,\alpha), (t,0)\rangle$ for the θ_ℓ^g-politician. Our curvature conditions further imply that $e_3 < \alpha/3 < \alpha\mu_\ell$ (the tie between the three candidates implies that $\mu_\ell^g = 1/3$, and then $\mu_\ell > 1/3$), and $e_3 < e_2$. This implies that $(t_2, e_3) \in P_{\{\theta_\ell^g, \theta_\ell^n, \theta_h^n\}}$, and that both non-green politicians strictly prefer (t_2, e_3) to $\langle(0,\alpha), (1,\alpha), (t,0)\rangle$. For the θ_ℓ^n-politician to accept the deviation towards π^{13}, just pick $\epsilon > 0$ sufficiently small so that $(t_2, e_3+\epsilon)$ belongs to $P_{\{\theta_\ell^g, \theta_\ell^n, \theta_h^n\}}$ and makes all members of $\left\{\theta_\ell^g, \theta_\ell^n, \theta_h^n\right\}$ strictly better-off (non-policy benefits remain unchanged for the green politician and increase for the non-green politicians).

- $k = 7$

First of all, note that there exists a sufficiently small $\varepsilon > 0$ such that $u\left(1-\varepsilon, \alpha, \theta_\ell^n\right) > u\left(0, 0, \theta_\ell^n\right)$ and $u\left(1-\varepsilon, \alpha, \theta_\ell^g\right) > u\left(1, 0, \theta_\ell^g\right)$, thus implying that $(1-\varepsilon, \alpha) \in \delta\left(\pi^1\right)$. Indeed, our assumptions on the distribution of types ($\mu_\ell > 1/2$ and $\mu^g > 1/2$) guarantee that party $\left\{\theta_h^g, \theta_\ell^g\right\}$ cannot be defeated in π^1 when it offers $(1-\varepsilon, \alpha)$.

Consider now party structure π^7. Since $\mu^g > 1/2$, party $\left\{\theta_h^g, \theta_\ell^g\right\}$ must win for sure in a π^7 equilibrium. Suppose first that it implements a policy $(t, \alpha) \in P_{\{\theta_h^g, \theta_\ell^g\}}$ such that $t < 1$. Then, θ_ℓ^n can profitably induce π^1 and then $(1, \alpha)$, which is her ideal policy in $P_{\{\theta_h^g, \theta_\ell^g\}}$. Suppose now that $\left\{\theta_h^g, \theta_\ell^g\right\}$ implements $(1, \alpha)$. Then, θ_h^n can profitably induce π^1 and then $(1-\varepsilon, \alpha) \in \delta(\pi^1)$. As a consequence, there is no EPS involving π^7 and **($\mathbf{P}_7$)** evidently holds.

- $k = 8$

As $\mu_\ell > 1/2$, $\left\{\theta_\ell^g, \theta_\ell^n\right\}$ wins with a probability of 1 in π^8-equilibrium. However, politician θ_ℓ^g can induce π^5, thereby enforcing her ideal policy and getting a benefit of β instead of $\beta/2$. Thus, there is no EPS involving π^8.

- $k = 9$

If condition **(P$_9$)** does not hold, then one of the following situations must arise.

(i) Suppose first that $\left\{\theta_\ell^g, \theta_h^n\right\}$ offers a policy $(0, e)$ with $e \in [\mu_h\alpha, \mu_\ell\alpha]$.
Then party $\left\{\theta_h^g, \theta_\ell^n\right\}$ can ensure its victory by offering $(0, e + \epsilon)$, with ϵ arbitrarily small. Both green politicians prefer this policy to $(0, e)$. Moreover, as ϵ is very small, the θ_ℓ^n-politician is compensated by an increase in her benefit of at least $\beta/4$:

$$u\left(0, e, \theta_\ell^n\right) - u\left(0, e + \epsilon, \theta_\ell^n\right) < \frac{\beta}{4}.$$

(ii) Suppose now that $\left\{\theta_\ell^g, \theta_h^n\right\}$ offers a policy (t, e) of the form $(t, \mu_\ell\alpha)$ with $t > 0$ or $(1, e)$ with $e > \mu_\ell\alpha$.
By the curvatures conditions imposed on V, $\left\{\theta_h^g, \theta_\ell^n\right\}$ has again a profitable deviation. To see this, take the indifference curves of politicians θ_h^g and θ_ℓ^n that pass through (t, e). These curves cross each other at another point, say (t', e'). It is easy to check that the unique intersection between the segment joining (t, e) to (t', e') and $P_{\left\{\theta_h^g, \theta_\ell^n\right\}}$ is a policy that enables $\left\{\theta_h^g, \theta_\ell^n\right\}$ to win for sure.

(iii) Finally, suppose $\left\{\theta_h^g, \theta_\ell^n\right\}$ offers a policy $(0, e)$ with $e > \alpha\mu_h$.
If $e \leq \alpha\mu_\ell$ then, by the same argument as in (i), $\left\{\theta_\ell^g, \theta_h^n\right\}$ has a profitable deviation. If $e > \alpha\mu_\ell$, then there exists a policy in $P_{\left\{\theta_\ell^g, \theta_h^n\right\}}$ which is strictly preferred to $(0, e)$ by the voters of type θ_ℓ^g, θ_h^n, and θ_ℓ^n. A deviation to this policy is therefore profitable to party $\left\{\theta_\ell^g, \theta_h^n\right\}$. As a result, $\left\{\theta_h^g, \theta_\ell^n\right\}$ cannot offer a pollution tax that exceeds $\alpha\mu_h$ in a π^9-equilibrium.

- $k = 10$

We first define the sets P_1, P_2, and P_3 as follows:

$$\begin{aligned}
P_1 &\equiv \left\{(t, e) \in P_{\left\{\theta_h^g, \theta_\ell^g, \theta_\ell^n\right\}} : u\left(t, e, \theta_\ell^n\right) \leq u\left(1, \alpha, \theta_\ell^n\right)\right\}, \\
P_2 &\equiv \left\{(t, e) \in P_{\left\{\theta_h^g, \theta_\ell^g, \theta_\ell^n\right\}} : u\left(t, e, \theta_h^g\right) \leq u\left(1, \alpha, \theta_h^g\right)\right\}, \\
P_3 &\equiv P_{\left\{\theta_h^g, \theta_\ell^g, \theta_\ell^n\right\}} \setminus (P_1 \cup P_2).
\end{aligned}$$

Under structure π^{10}, the three-member party must win for sure in an equilibrium, and then offer a policy in $P_{\left\{\theta_h^g, \theta_\ell^g, \theta_\ell^n\right\}} \equiv P_1 \cup P_2 \cup P_3$. We distinguish among the three different cases.

(i) It offers a policy in P_1. Then $\left\{\theta_\ell^g, \theta_\ell^n\right\}$ can induce π^2, thus enforcing $(1, \alpha)$ and obtaining a benefit of $\beta/2$ instead of $\beta/3$.

(ii) It offers a policy in P_2. By the same argument as previously, $\left\{\theta_h^g, \theta_\ell^g\right\}$ can profitably induce π^1.

(iii) It offers a policy (t, e) in $P_3 \setminus P_{\left\{\theta_h^g, \theta_\ell^n\right\}}$. Substituting (t, e) to $(1, \alpha)$ in the proof of Lemma 2, we obtain that there exists a policy $(t', \mu_h\alpha)$ such that $\left(t', \mu_h\alpha\right) \in \delta\left(\pi^4\right)$, and

$$u\left(t', \mu_h\alpha, \theta_h^g\right) > u\left(t, e, \theta_h^g\right),$$
$$u\left(t', \mu_h\alpha, \theta_\ell^n\right) > u\left(t, e, \theta_\ell^n\right).$$

This implies that $\left\{\theta_h^g, \theta_\ell^n\right\}$ can profitably induce π^4. Doing so, they indeed enforce a better policy and no longer share the non-policy benefit with θ_ℓ^g.
Suppose that $(t, e) \in P_3 \cap P_{\left\{\theta_h^g, \theta_\ell^n\right\}}$. This implies that (t, e) satisfies the conditions of Lemma 2, which in turn implies that $(t, e) \in \delta\left(\pi^4\right)$. Therefore, coalition $\left\{\theta_h^g, \theta_\ell^n\right\}$ can enforce the same policy without sharing the non-policy benefit with θ_ℓ^g.
This proves that there is no EPS involving party structure π^{10}.

- $k = 11$

From Lemma 3, we know that θ_ℓ^g's ideal policy $(1, \alpha) \in \delta\left(\pi^5\right)$. As $\beta > 0$, the θ_ℓ^g -politician has consequently a profitable deviation to π^5.

- $k = 12$

For $(\mathbf{P}_{12})$ to be true, it suffices to check that the big party in π^{12} never offers a policy $(0, e)$ with $e > \mu_h\alpha$, and that $\left\{\theta_\ell^g\right\}$ never wins in a π^{12} -equilibrium.

When $\left\{\theta_h^g, \theta_\ell^n, \theta_h^n\right\}$ offers a policy of the form $(0, e)$ with $e > \alpha\mu_h$, it is defeated with a probability of 1 by $\left\{\theta_\ell^g\right\}$ which offers $(1, \alpha)$. Indeed, under A1, voters of type θ_ℓ^n strictly prefer $(1, \alpha)$ to any policy $(0, e)$ with $e \in [0, 1]$, and $\mu_\ell > 1/2$. Therefore, if $\left\{\theta_h^g, \theta_\ell^n, \theta_h^n\right\}$ runs in a π^{12} -equilibrium, then it offers an environmental tax at most equal to $\alpha\mu_h$.

Let us now turn to party $\left\{\theta_\ell^g\right\}$. This party can only offer $(1, \alpha)$ which is defeated by $(t_1, \alpha\mu_h) \in P_{\left\{\theta_h^g, \theta_\ell^n, \theta_h^n\right\}}$ in pairwise vote (see Lemma 2). As a result, it can never win in a π^{12} -equilibrium.

- $k = 13$

Suppose that, contrary to $(\mathbf{P}_{13})$, a policy (t, e) with $e > \alpha\mu_h$ emerges in a π^{13}-equilibrium. This cannot be an EPS. To see this, suppose first that $t < 1$. Then (t, e) does not belong to the Pareto set of $\left\{\theta_\ell^g, \theta_\ell^n\right\}$. This implies that there exists a policy in $P_{\left\{\theta_\ell^g, \theta_\ell^n\right\}}$ that makes θ_ℓ^g and θ_ℓ^n strictly better-off and is a π^2-equilibrium policy. Indeed, it is easy to see that, under conditions A1 and $\mu_\ell > 1/2$, $(t, e) \in \delta\left(\pi^2\right)$ for every $(t, e) \in P_{\left\{\theta_\ell^g, \theta_\ell^n\right\}}$.

Suppose now that the policy (t, e) under consideration satisfies $t = 1$ and $e \geq \alpha\mu_\ell$. It is easy to see that any such a policy is a π^3-equilibrium policy (implemented by

party $\{\theta_\ell^g, \theta_h^n\}$). Therefore, coalition $\{\theta_\ell^g, \theta_h^n\}$ can profitably induce π^3, thus enforcing the same policy (t, e) without sharing the non-policy benefit with θ_ℓ^n. If $e < \alpha\mu_\ell$, then $(t, e) \notin P_{\{\theta_\ell^g, \theta_h^n\}}$. There consequently exists $(t', e') \in P_{\{\theta_\ell^g, \theta_h^n\}}$ (with $e' = \alpha\mu_\ell$) that makes θ_ℓ^g and θ_h^n strictly better off. Substituting (t', e') to (t, e) in the previous reasoning proves **($\mathbf{P}_{13}$)**.

- $k = 14$

Suppose first that the grand party offers a policy (t, e) outside the Pareto set of $\{\theta_\ell^g, \theta_\ell^n\}$. This implies that there exists $(t'', e'') \in P_{\{\theta_\ell^g, \theta_\ell^n\}}$ such that $u\left(t'', e'', \theta_\ell^j\right) > u\left(t, e, \theta_\ell^j\right)$ for every $j \in \{g, n\}$. By Lemma 4, $\{\theta_\ell^g, \theta_\ell^n\}$ should then induce π^8 so as to enforce (t'', e'') and raise the benefit of its members.

To show that **($\mathbf{P}_{14}$)** is true, we must therefore show that the grand party implementing a policy $(1, e) \in P_{\{\theta_\ell^g, \theta_\ell^n\}}$, with $e > \alpha\mu_h$, is not an EPS. For every $e > \alpha\mu_h > 0$, there exists by continuity an $\varepsilon > 0$ such that $e - \varepsilon \geq 0$ and

$$u\left(1, e, \theta_\ell^g\right) - u\left(1, e - \varepsilon, \theta_\ell^g\right) < \frac{\beta}{2}.$$

Moreover, Lemma 4 establishes that $(1, e-\varepsilon) \in \delta\left(\pi^8\right)$. This proves that a deviation to π^8 is again profitable to coalition $\{\theta_\ell^g, \theta_\ell^n\}$, thus completing the proof of Proposition 1.

Proof of Proposition 2

Substituting θ_i^n to θ_i^g, $i \in \{\ell, h\}$, and μ^n to μ^g in the proof of Proposition 1, we can prove Proposition 2 in like manner.

Proof of Proposition 3

Our proof of Proposition 3 will proceed in three short steps. Given our previous findings, we already know that there cannot be a stable green party when there is a majority of green voters. We consequently assume throughout that $\mu^g < 1/2$.

Step 1: If $S \subseteq \Theta$ is a stable green party, then $S = \{\theta_h^g, \theta_\ell^g\}$.

We can directly infer from Lemma 1 that there is no stable green party in π^0, π^2, and π^4. In π^3, if $\{\theta_\ell^g, \theta_h^n\}$ offers $(1, \alpha)$ than $\{\theta_\ell^n\}$ can win for sure by offering its ideal policy: θ_h^n-voters strictly prefer $(1, 0)$ to $(1, \alpha)$ and $\mu^n > 1/2$.

Policy $(0, \alpha)$ is defeated by both $(1, 0)$ and $(1, \alpha)$ in pairwise vote. Therefore, there is no π^5-equilibrium in which $\{\theta_h^g, \theta_h^n\}$ runs alone, or against a single opponent, and offers $(0, \alpha)$. Under assumption A1, the θ_h^n-politician strictly prefers $(1, 0)$ to $\langle(0, \alpha), (1, \alpha), (1, 0)\rangle$. A parallel argument to the one used to prove Lemma 3 would show that there is no three-candidate π^5-equilibrium.

The non-green party wins with a probability of 1 in any π^6- and π^7-equilibrium since $\mu^n > 1/2$.

There is no EPS involving π^8. Indeed, θ_ℓ^n can profitably induce the π^5-equilibrium in which she implements *alone* her/his ideal policy. In π^9, party $\{\theta_h^g, \theta_\ell^n\}$ [resp. $\{\theta_\ell^g, \theta_h^n\}$] can never win by offering $(0, \alpha)$ [resp. $(1, \alpha)$], for there is a policy in the Pareto set of $\{\theta_\ell^g, \theta_h^n\}$ [resp. $\{\theta_h^g, \theta_\ell^n\}$] that allows the latter to win for sure.

Consider π^{10} and π^{13} now. Under these party structures, the three-member party must win for sure. Suppose it offers a policy of the form (t, α). As $\beta > 0$, inducing π^2 and enforcing $(1, \alpha)$ is strictly profitable to coalition $\{\theta_\ell^g, \theta_\ell^n\}$.

For a policy $(t, \alpha) \in P_{\{\theta_h^g, \theta_\ell^g, \theta_h^n\}}$ to be a π^{11}-equilibrium policy, both θ_ℓ^g and θ_h^n must prefer (t, α) to $(1, 0)$ (otherwise, $\{\theta_\ell^n\}$ could win by offering $(1, 0)$). However, then, there exists a policy $(t', \alpha\mu_\ell)$ such that $\langle \varnothing, (t', \alpha\mu_\ell), \varnothing \rangle$ is a π^3-equilibrium, and both θ_ℓ^g and θ_h^n prefer $(t', \alpha\mu_\ell)$ to (t, α). Indeed, a brief inspection of the structure of preferences reveals that the analysis of EPS involving π^3 when $\mu^g < 1/2$ is symmetric to the analysis of EPS involving π^4 when $\mu^g > 1/2$. We can then deduce from Lemma 2 that such a policy exists. However, this implies that coalition $\{\theta_\ell^g, \theta_h^n\}$ can profitably deviate by inducing π^3 and enforcing $(t', \alpha\mu_\ell)$.

In a π^{12}-equilibrium, the bigger party never offers $(0, \alpha)$. Since θ_ℓ^n voters strictly prefer $(1, \alpha)$ to $(0, \alpha)$, $\{\theta_\ell^g\}$ could indeed win for sure by offering $(1, \alpha)$.

Finally, the grand coalition is not a stable green party. Suppose the unique party in π^{14} offers a policy of the form (t, α). As $\beta > 0$, inducing π^8 and enforcing $(1, \alpha)$ is strictly profitable to coalition $\{\theta_\ell^g, \theta_\ell^n\}$.

Step 2: A stable green party exists only if $\mu_h \mu_\ell (\omega_h - \omega_\ell) \geq \Delta^n$ and (c) hold.

An immediate consequence of Step 1 is that the only party structure in which there can be a stable green party is π^1. A little reflection suggests that the analysis of EPS involving π^1 when $\mu^g < 1/2$ is symmetric to the analysis of EPS involving π^6 when $\mu^g > 1/2$. Inspecting the case $k = 6$ in the proof of Proposition 1, thus, reveals that, when $\mu^g < 1/2$, there is no EPS in which the green party runs against one or two rival candidates.

Our focus is therefore on EPS of the form $\left(\pi^1, \langle (t, \alpha), \varnothing, \varnothing \rangle\right)$ where $t \in [0, 1]$. $((t, \alpha), \varnothing, \varnothing)$ cannot be a π^1-equilibrium if one of the following conditions hold:

(i) θ_ℓ^g-voters strictly prefer $(1, 0)$ to (t, α) (party $\{\theta_\ell^n\}$ can offer $(1, 0)$ and win for sure since $\mu_\ell > 1/2$) or, equivalently, $t < 1 - \Delta^g/\Delta_\ell$;
(ii) θ_ℓ^n-voters strictly prefer $(0, 0)$ to (t, α) (party $\{\theta_h^n\}$ can offer $(0, 0)$ and win for sure since $\mu^n > 1/2$) or, equivalently, $t < \Delta^n/\Delta_\ell$;
(iii) θ_h^n-voters strictly prefer $(1, 0)$ to (t, α) (party $\{\theta_\ell^n\}$ can offer $(1, 0)$ and win for sure since $\mu^n > 1/2$) or, equivalently $t > 1 - \Delta^n/\Delta_h$.

For none of these three conditions to hold, t must then belong to the interval

$$T \equiv \left[\max\left\{1 - \frac{\Delta^g}{\Delta_\ell}, \frac{\Delta^n}{\Delta_\ell}\right\}, 1 - \frac{\Delta^n}{\Delta_h}\right].$$

Therefore, a necessary condition for $((t, \alpha), \varnothing, \varnothing)$ to be a π^1-equilibrium is that T is nonempty. However, this is only the case if $\mu_\ell \Delta^g \geq \mu_h \Delta^n$ and $\mu_h \mu_\ell (\omega_h - \omega_\ell) \geq \Delta^n$.

Step 3: $\left\{\theta_h^g, \theta_\ell^g\right\}$ *is a stable green party whenever* $\mu_h \mu_\ell (\omega_h - \omega_\ell) > \Delta^n$ *and* $\mu_h \Delta^g > \mu_\ell \Delta^n$.

When $\mu_\ell \Delta^g > \mu_h \Delta^n$ and $\mu_h \mu_\ell (\omega_h - \omega_\ell) > \Delta^n$, there is a tax rate t that belongs to the interior of T. Then, it follows from the above argument that $((t, \alpha), \varnothing, \varnothing)$ is a π^1-equilibrium. What remains to be proved, therefore, is that $\left(\pi^1, ((t, \alpha), \varnothing, \varnothing)\right)$ is an EPS.

Note first that coalition $\left\{\theta_h^g, \theta_\ell^g\right\}$ cannot be part of a deviating coalition: (t, α) belongs to the Pareto set of that coalition and forming a larger party with another politician would make their non-policy benefit decrease. Moreover, we know from Lemma 1 that neither θ_h^g nor θ_ℓ^g have an interest in inducing π^0.

Politicians θ_h^n and θ_ℓ^n inducing π^7 is then the only possible deviation. As $\mu^g < 1/2$, $\left\{\theta_\ell^n, \theta_h^n\right\}$ must run alone in a π^7-equilibrium. Suppose first that it offers a policy $(t', 0) \in P_{\{\theta_\ell^n, \theta_h^n\}}$ such that $u\left(t', 0, \theta_h^n\right) < u\left(0, \alpha, \theta_h^n\right)$. Then, tedious computations reveal that the set of policies $\left(t'', \alpha\right) \in P_{\{\theta_h^g, \theta_\ell^g\}}$ such that politicians of types θ_h^g, θ_ℓ^g, and θ_h^n strictly prefer $\left(t'', \alpha\right)$ to $(t', 0)$ is nonempty whenever $\mu_\ell \Delta^g > \mu_h \Delta^n$. This implies that the green party can profitably deviate by offering $\left(t'', \alpha\right)$, and then $(t', 0) \notin \delta\left(\pi^7\right)$.

A parallel argument shows that if $\left\{\theta_\ell^n, \theta_h^n\right\}$ offers a policy $(t', 0) \in P_{\{\theta_\ell^n, \theta_h^n\}}$ such that $u\left(t', 0, \theta_\ell^n\right) < u\left(1, \alpha, \theta_\ell^n\right)$, then the green party can also profitably deviate whenever $\mu_h \Delta^g > \mu_\ell \Delta^n$. As

$$\left\{t' \in [0, 1] : u\left(t', 0, \theta_h^n\right) < u\left(0, \alpha, \theta_h^n\right)\right\} \cup \left\{t' \in [0, 1] : u\left(t', 0, \theta_\ell^n\right) < u\left(1, \alpha, \theta_\ell^n\right)\right\} = [0, 1]$$

whenever $\mu_h \mu_\ell (\omega_h - \omega_\ell) > \Delta^n$, this proves that there is no π^7-equilibrium, and then no possible deviation from π^1, when $\mu_h \mu_\ell (\omega_h - \omega_\ell) > \Delta^n$.

Combining Steps 1–3, we obtain the proposition.

On the choice of solution concept

As its definition reveals, the bi-core stability concept does not allow for deviating coalitions to be farsighted—to consider whether a blocking state is itself stable or not. For instance, any political state $\mathbf{s}$ of the form $\left(\pi^1, (t, \alpha)\right)$ is not bi-core stable because the θ_ℓ^g-politician can profitably induce π^0, and thus enforce her ideal policy. Nevertheless, one might argue that a farsighted θ_ℓ^g-politician would never do so, for she would anticipate a further deviation by $\left\{\theta_h^g, \theta_\ell^n\right\}$ to $\mathbf{s}' \equiv \left(\pi^4, (t_1, \mu_h \alpha)\right)$. Lemma 2 indeed shows that there exists a policy $(t_1, \mu_h \alpha)$ that makes both θ_h^g and θ_ℓ^n strictly better off than under the ideal policy of θ_ℓ^g.

Such an argument implicitly assumes, however, that state $\mathbf{s}'$ is itself immune to such deviations; otherwise those deviations would be anticipated by the farsighted θ_ℓ^g-politician. To see that this cannot be the case, assume for the moment that $\mathbf{s}'$ is deemed stable (i.e., immune to sequential deviations). It ensues that, by the same logic as that used to prove Lemma 2 and because of the continuity of preferences,

there exists t_1' in a small neighborhood of t_1 such that $\mathbf{s}'' \equiv \left(\pi^4, \left(t_1', \mu_k\alpha\right)\right)$ is also stable (and preferred by both θ_h^g and θ_ℓ^n to $\left(\pi^0, (1, \alpha)\right)$). However, *at least one member* of $\left\{\theta_h^g, \theta_\ell^n\right\}$ must strictly prefer $\left(t_1', \mu_h\alpha\right)$ to $(t_1, \mu_h\alpha)$, and therefore $\mathbf{s}''$ to $\mathbf{s}'$. This member of the green party should therefore initiate a sequence of deviations of the form: $\mathbf{s}' \to \left(\pi^0, (1, \alpha)\right) \to \mathbf{s}''$. That is, she induces π^0, anticipating that $\left\{\theta_h^g, \theta_\ell^n\right\}$ will then induce $\mathbf{s}''$. This is a contradiction with $\mathbf{s}'$ being stable. Besides, similar arguments can then be used to prove that $\mathbf{s}''$ is not stable, and so on and so forth.

This simple example is an illustration of the nonexistence problems that arise in voting models with infinite sets of states when one use "farsighted" stability concepts—such as the equilibrium processes of coalition formation (Konishi and Ray 2003; Anesi 2006b), or the extended equilibrium binding agreements (Diamantoudi and Xue 2007). Apart from rare exceptions, the (applied) literature using such farsighted solution concepts often assumes a finite number of states with a single equilibrium for each coalition structure. In the context of our article, the infiniteness of the set of possible states opens the room to numerous deviation opportunities, thereby causing existence problems and even cycles (in the sense of starting from a coalition structure, like π^4 in the above example, and ending up coming back to the original structure).

Furthermore, a large body of experimental literature rejects both individuals' and groups' farsightedness (the most recent contributions include Hey and Knoll 2007, and Bone et al. 2009).

References

Anesi V (2006a) Earmarked taxation and political competition. J Pub Econ 90:679–701

Anesi V (2006b) Committees with farsighted voters: a new interpretation of stable sets. Soc Choice Welf 27:595–610

Anesi V, De Donder P (2009) Party formation and minority ideological positions. Econ J 119:1303–1323

Bernheim B, Whinston M (1986) Menu auctions, resource allocation, and economic influence. Q J Econ 101:1–31

Besley T, Coate S (1997) An economic model of representative democracy. Q J Econ 112:85–114

Bone J, Hey JD, Suckling J (2009) Do people plan? Exp Econ 12:12–25

Boyce JK (1994) Inequality as a cause of environmental degradation. Ecol Econ 11:169–178

Brett C, Keen M (2000) Political uncertainty and the earmarking of environmental taxes. J Pub Econ 75:315–340

Congleton RD (1992) Political institutions and pollution control. Rev Econ Stat 74:412–421

Cremer H, De Donder P, Gahvari F (2004) Taxes, budgetary rule and majority voting. Pub Choice 119:335–358

Cremer H, De Donder Ph, Gahvari F (2007) Energy taxes in three political economy models. B.E. J Econ Anal Policy 7(1):Article 32

Cremer H, De Donder P, Gahvari F (2008) Political competition within and between parties: an application to environmental policy. J Pub Econ 92:532–547

Diamantoudi E, Xue L (2007) Coalitions, agreements and efficiency. J Eco Theory 136:105–125

Esteban JM, Ray D (1994) On the measurement of polarization. Econometrica 62:819–851

Fredriksson PG (1997) The political economy of pollution taxes in a small open economy. J Environ Econ Manag 33:44–58

Grossman G, Helpman E (2002) Interest groups and trade policy. Princeton University Press, Princeton

Heerink N, Mulatu A, Bulte E (2001) Income inequality and the environment: aggregation bias in environmental Kuznets curves. Ecol Econ 38:359–367

Hey JD, Knoll JA (2007) How far ahead do people plan? Econ Lett 96:8–13

Heyes A, Dijkstra B (2001) Interest groups and the demand for environmental protection. In: Tietenberg T, Folmer H (eds) International yearbook of environmental and resource economics 2001/2002, Edward Elgar, Cheltenham

Jones LR, Manuelli R (2001) Endogenous policy choice: the case of pollution and growth. Rev Econ Dyn 4:369–405

Kempf H, Rossignol S (2007) Is inequality harmful for the environment in a growing economy? Econ Polit 19:53–71

Konishi H, Ray D (2003) Coalition formation as a dynamic process. J Econ Theory 110:1–41

Levy G (2004) A model of political parties. J Econ Theory 115:250–277

List JA, Sturm DM (2006) How elections matter: theory and evidence from environmental policy. Q J Econ 121:1249–1281

McAusland C (2003) Voting for pollution policy: the importance of income inequality and openness to trade. J Int Econ 61:425–451

Neumayer E (2004) The environment, left-wing political orientation and ecological economics. Ecol Econ 51:167–175

Oates W, Portney P (2003) The political economy of environmental policy. In: Mahler KG, Vincent J (eds) Handbook of environmental economics. North-Holland/Elsevier Science, Amsterdam

Osborne M, Slivinski A (1996) A model of political competition with citizen-candidates. Q J Econ 111: 65–96

Wilson JK, Damania R (2005) Corruption, political competition and environmental policy. J Environ Econ Manag 49:516–535

Soc Choice Welf
DOI 10.1007/s00355-010-0508-z

Omnibus or not: package bills and single-issue bills in a legislative bargaining game

Johanna M. M. Goertz

Abstract Legislators sometimes pass unrelated issues in one bill with one vote (omnibus bills), but they often vote separately on different issues as well. In this article, a model is considered in which the first mover chooses whether to bargain over an omnibus bill or several separate bills. The main difference between the two types of bills is that trade-offs between issues are possible with omnibus bills, but not with separate bills. The underlying bargaining game is demand-bargaining (Morelli 1999). In this game, moderate legislators prefer to propose single-issue bills; extreme legislators prefer omnibus bills, if the asymmetry in policy ideal points in the legislature is large enough.

JEL Classification C78 · D02 · D72 · D78

1 Introduction

Legislators often package unrelated policy issues in one bill and pass them in the legislature with a single vote. Consider, for example, omnibus bills in the US federal legislation. Different issues are voted on with a single vote in such omnibus bills, instead of passing separate bills for each issue. Omnibus bills are typically considered to be implicit vote trades (Ferejohn 1986): Issues unfavorable to the majority are packaged with favorable issues to make sure they pass. Legislators can make trade-offs

J. M. M. Goertz (✉)
Department of Economics, University of Guelph, Guelph, ON N1G 2W1, Canada
e-mail: jgoertz@uoguelph.ca

J. M. M. Goertz
Center of Operations Research and Econometrics (CORE), Catholic University of Louvain, Louvain-la-Neuve, Belgium

Published online: 21 December 2010

between several policy issues, and such trade-offs usually increase efficiency. However, not all policy issues are approved in omnibus bills. Are such choices irrational? Not necessarily: It can be optimal for a proposal maker to propose separate bills if her/his expectation about the policy outcome in the subsequent legislative bargaining game warrants it. As it turns out, it depends on the underlying bargaining game whether a legislator finds omnibus or single-issue bills optimal, as well as on certain parameters of the legislature. Intuitively, the polarization of the legislature should play a role (because omnibus bills often contain at least some unfavorable issues or proposals), as well as the identity (or ideological position) of the proposal maker.

Here, legislative bargaining is modeled as a non-cooperative bargaining game between legislators. The proposal maker can decide to propose an omnibus bill containing all issues, or can decide to propose several single-issue bills instead. Omnibus bills allow legislators to make trade-offs between different policy issues, while single-issue bills do not. The two specific policy issues considered are purely distributive (sharing a fixed budget between the legislators) and ideological, with legislators having different policy ideal points in the ideological dimension. To make sure that omnibus bills are not optimal simply because of interdependence of preferences, it is assumed that legislators have separable preferences over the two dimensions. Trade-offs can still be optimal because of differences in the marginal rate of substitution.

The previous literature on non-cooperative legislative bargaining, both theoretical (Baron and Ferejohn 1989, and more recently Banks and Duggan 2000, 2006; Kalandrakis 2006; Battaglini and Coate 2007, 2008; Snyder et al. 2005), and empirical or experimental (for example, Diermeier et al. 2003, 2008; Frechette et al. 2005a,b), has not yet found a compelling argument for the proposal of single-issue bills.[1] Some previous models consider only distributive politics so that the question of packaging issues or not does not arise. Others provide models for government formation, so that the question of omnibus or single-issue bills is not suitable or relevant. In the remaining ones, omnibus bills are typically considered optimal. The main reason for this is that they assume the same underlying bargaining game—alternating-offers (or some variation of it), first proposed by Baron and Ferejohn (1989)—in which omnibus bills are indeed the only optimal choice for a proposal maker.

In the alternating-offers game, a generalization of Rubinstein (1982) to more than two players, a proposal maker is randomly chosen among the legislators in every round and makes a policy proposal. The legislature votes on this proposal, and if a majority rejects it, a new round starts with a new randomly chosen proposal maker. Because future proposal makers are unknown, legislators agree even to proposals not favorable for them because future proposals might be worse (but find majority support), and because there is discounting. The first proposal maker can take advantage of this and extract a large rent. To extract this rent in all dimensions, it is optimal to propose an omnibus bill.

Here, a different bargaining game is used with the following important characteristics: There is no uncertainty about the next proposal (or amendment) maker, legislators know that there will be a competing counterproposal in the same round for sure, and

[1] The list of articles is by far not exhaustive. The mentioned articles are the most recent, and/or the most closely related to this article.

coalitions between different pairs of legislators (even excluding the first proposal maker) are possible in the same round. Bernheim et al. (2006) show in a single-issue bargaining game, which shares some (but not all) of these features, that such characteristics effectively reduce the power of the first mover. It is important that the identity of at least some future proposers is known, that there is a sufficiently large number of counterproposals to be made for sure, as well as that there is no discounting between the possible implementation of the first proposal and a counterproposal.[2] Bernheim et al. (2006) show that the Condorcet winner alternative is implemented in their game; if there is no Condorcet winner, the last proposal maker can be a dictator.

The two most closely related models are those of Jackson and Moselle (2002) and Morelli (1999). Both models consider the same type of issues (purely distributive and ideological), but different bargaining games. Jackson and Moselle (2002) use alternating-offers and assume a more general utility function, but focus on a specific separable utility function to derive point predictions for several examples. They show precisely that with a concave utility function the first proposer of an alternating-offers game always proposes an omnibus bill. This is not the case for the bargaining game considered here. Instead, some legislators—depending on their ideological position—prefer single-issue bills, while others prefer omnibus bills.

Morelli (1999) assumes a bargaining game (demand-bargaining) very similar to the one used here but assumes a different utility function.[3] He does not consider the distinction between single-issue and omnibus bills and simply assumes that the first proposer proposes an omnibus bill. As it turns out, this is not a restriction in his model because of the utility function—quasilinear with the absolute distance function in the ideological dimension. However, there is a distinction between omnibus and single-issue bills if legislators have different marginal rates of substitution (Heifetz and Ponsati 2007, for example, reach the same conclusion in a slightly different context). This is true for the utility function used here, so that the decision of the first mover is non-degenerate.

The comparison with both Jackson and Moselle (2002) and Morelli (1999) shows that different bargaining games substantially affect the strategic considerations of the first mover. In a demand-bargaining-type game, the first mover has to fear that a coalition forms without her/him and approves a more unfavorable proposal. In fact, the first mover needs to obtain the support of the second mover by offering the second mover enough utility from the policy decisions. This turns out to be more expensive for omnibus bills than for single-issue bills, unless the proposal maker has an extreme ideological position. A first mover in an alternating-offers game does not face the

[2] There is an open-rule version of the alternating-offers game, which admits counterproposals. However, the identity of the amendment maker is unknown, counterproposals can only be made until one receives majority support (so that the game can still be over after the first proposal), and there is discounting before any counterproposal can be implemented (if the first proposal is not accepted).

[3] Montero and Vidal-Puga (2007) recently provided a critique of Morelli (1999), showing that one of the propositions (the distribution of benefits in a pure-benefits game is proportional to ex-ante bargaining power) does not generally hold if legislators have different voting weights. Morelli (2007) stresses in a response that one of the most important scenarios is the one in which legislators have only one vote (and equal weights); the critique does not apply in this case. This is also true for the model used here, in which legislators are symmetric and have one vote each.

same threat: The amendment maker is randomly chosen, and there is discounting if the counterproposal should be implemented, so that alternative coalitions are much harder to form.

To make the trade-offs faced by the first proposers in demand-bargaining and their influence on policy decisions as apparent as possible, a specific utility function is chosen. This makes it possible to point at those specific parameters of the game, which influence the decision of the first proposal maker. In addition, point predictions need to be derived to compare the exact utility from omnibus and single-issue bills. Assuming a specific utility function tremendously simplifies this exercise.

The fact that legislators prefer single-issue bills in a perfect- and complete-information bargaining game is quite different from other results in the literature on endogenous strategic agenda formation with perfect information.[4],[5] Inderst (2000), for example, argues for packaging of issues in two-player bargaining because this allows for mutually beneficial trade-offs and increases efficiency. In and Serrano (2003) extend the same framework to a larger class of utility functions and come to the same conclusion. And if agendas are restricted to issue-by-issue bargaining, then the outcome is inefficient (In and Serrano 2004). For trade negotiations, Horstmann et al. (2005) show that bargaining over several issues jointly generates gains for both parties, even if not both parties receive a positive gain from both issues. And Leblanc et al. (2000) find that legislators propose omnibus bills in an alternating-offers game over current spending and investment for the future, even if this yields inefficient policy outcomes.

In Sect. 2, the model is presented. In Sects. 3 and 4, omnibus bills and single-issue bills are discussed separately. The results are compared and, as a result, the choice of the first proposal maker can be characterized in Sect. 5. Section 6 concludes.

2 The model

Consider a legislature with three legislators $\{\mathrm{L}, \mathrm{C}, \mathrm{R}\} \in \mathbf{L}$. These legislators have to approve two different policies (ideological and purely distributive) with a majority rule and one vote per legislator. Let $Y = [\hat{y}_{\mathrm{L}}, \hat{y}_{\mathrm{R}}] \subset \mathbb{R}$ be the interval of possible ideological policies, with $\hat{y}_{\mathrm{L}}$ as legislator L's most preferred ideological policy (ideal point) and $\hat{y}_{\mathrm{R}}$ as legislator R's most preferred ideological policy.[6] Legislator C's ideal point $\hat{y}_{\mathrm{C}}$ is such that $\hat{y}_{\mathrm{L}} < \hat{y}_{\mathrm{C}} < \hat{y}_{\mathrm{R}}$ and, without loss of generality, $\hat{y}_{\mathrm{C}} \leq \frac{\hat{y}_{\mathrm{L}}+\hat{y}_{\mathrm{R}}}{2}$. In some cases, it is assumed in addition that $\hat{y}_{\mathrm{L}} = 0$ and $\hat{y}_{\mathrm{R}} = 1$ to simplify the analysis.

The legislators also have to divide a budget between them. Consider a budget of size X, and assume that X is large (a more precise definition is given in Sect. 3).

[4] Single-issue bargaining can be optimal in games with asymmetric information (e.g., Heifetz and Ponsati 2007): If some issues are harder to agree on, then negotiating over issues separately avoids a spillover of inefficiencies from harder issues to easier ones.

[5] Single-issue bargaining can be a preferred choice when players have convex utility functions, as in Lang and Rosenthal (2001). If players have a preference for risk and future proposal makers are uncertain, then single-issue bills are a way to settle issues through several lotteries.

[6] It is inconsequential to restrict the policy space to $[\hat{y}_{\mathrm{L}}, \hat{y}_{\mathrm{R}}]$ because no legislator would propose a policy outside this interval in equilibrium.

Legislator i's share of the budget is denoted by x_i, and it has to be true that $\sum_i x_i \leq X$ and $x_i \geq 0 \; \forall \; i =$L, C, R. The budget distribution is a vector $(x_{\mathrm{L}}, x_{\mathrm{C}}, x_{\mathrm{R}}) \in \mathbb{R}^3_+$.

Let $\mathbf{D} = Y \times \mathbb{R}^3_+$, and denote by $d = (y, x_{\mathrm{L}}, x_{\mathrm{C}}, x_{\mathrm{R}}) \in \mathbf{D}$ the approved policies. The utility of legislator $i \in \mathbf{L}$ from a particular decision d is given by

$$\begin{aligned} u_i(d) &= v_i(y) + x_i, \quad d \in \mathbf{D}, \\ v_i(y) &= k - b(y - \hat{y}_i)^2, \quad b > 0, \quad k \geq b(\hat{y}_{\mathrm{L}} - \hat{y}_{\mathrm{R}})^2. \end{aligned} \tag{1}$$

Given this utility function, legislators have different marginal rates of substitution. This implies that trade-offs between the two dimensions are mutually beneficial. By assumption, legislators cannot commit to trade-offs across different bills, so that trade-offs are only possible with omnibus bills (budget shares are used as sidepayments) and legislators are not indifferent between the two types of bills.

The parameter b captures the intensity of the legislators' utility from the ideological policy relative to the distributive policy. It is assumed that $b = 1$, but none of the results depends on the size of b. The parameter k ensures that even the most extreme legislators receive positive utility from implementing a policy in $[\hat{y}_{\mathrm{L}}, \hat{y}_{\mathrm{R}}]$, no matter how unfavorable to them. This assumption has no bearing on the results; it simply ensures that legislators cannot increase their utility by endlessly delaying the decision in the ideological dimension.

The legislators' preferences are separable: A legislator's preference over alternatives in one dimension is independent of the decision in the other dimension. This assumption eliminates one confounding factor for the choice of omnibus bills: interdependent preferences. The focus should be on the fact that the choice of omnibus bills or single-issue bills can be a consequence of the bargaining game and of parameters characterizing the legislature alone.

As mentioned above, there is no technology which allows legislators to make credible commitments to trade-offs between dimensions when they use single-issue bills. Suppose that legislators sequentially approve two single-issue bills. The assumption of no commitment to trade-offs is equivalent to assuming that after the first policy is settled, the game goes into a new bargaining round and all bets are off. As such, a bargaining game with two single-issue bills approved sequentially is equivalent to two independent simultaneous bargaining games, one for each policy. To avoid discounting of the future to be a confounding factor favoring omnibus bills (more rounds are needed to approve two bills than to approve one), it is assumed that legislators simultaneously play two independent bargaining games in the case of single-issue bills.

The legislative bargaining game is very similar to Morelli (1999)'s demand-bargaining, except that it now includes the choice over omnibus and single-issue bills. Legislators move sequentially in each round. To simplify the analysis, it is assumed that each bargaining round in each bargaining game is initiated by the same first mover. In Morelli (1999), the first mover is chosen by the head of state in each round, and the head of state chooses the same first mover each time. This stage is not included in the game: The outcome of the game would depend on the preferences of the head of state (through of the choice of the first mover), and the author does not want to make

assumptions about such preferences. Instead, the outcome of the game is characterized as a function of the first mover's identity. Assuming that the first mover is the same in each round makes the comparison with Morelli (1999) easy and tremendously simplifies the analysis.

Legislator i is chosen randomly by nature to be the first mover with probability π_i, with $\sum_i \pi_i = 1$ for $i \in \mathbf{L}$. The recognition probability π_i is exogenously fixed. If chosen as the first mover, legislator i then decides on the type of bill, one omnibus or two single-issue bills, and it remains fixed thereafter. It will become clear below that it is not restrictive to assume the same first mover to make initial proposals for both single-issue bills. The first mover also chooses the order of play of the remaining legislators.

Suppose legislator i is the first mover, fixes the order of play of the remaining legislators to be jk, $j, k \in \mathbf{L} \setminus \{i\}$, and proposes an omnibus bill. The policy proposal of legislator i contains both an ideological policy proposal y_i and a "demand" for a budget share x_i; it is denoted by by (y_i, x_i). In contrast to alternating-offers, there is no vote after the initial proposal of the first mover because the budget is not yet distributed. Instead, the second mover j can now make a policy "demand" as well, denoted by (y_j, x_j). If the proposal of the second mover is compatible with the proposal of the first mover, then the game is over. The two proposals are considered compatible if $x_j = X - x_i$ and $y_j = y_i$. The policy decision is then $d = (y_i, x_i, x_j, 0)$.[7]

Suppose instead that the second mover's demand is incompatible, so that there are two proposals on the table, (y_i, x_i) and (y_j, x_j). Legislator k, the third mover, can make a demand compatible with either one of the two standing proposals. If the demand is $(y_k = y_i, x_k = X - x_i)$, then the policy decision is $d = (y_i, x_i, x_k = X - x_i, x_j = 0)$, and the game is over. If the demand is $(y_k = y_j, x_k = X - x_j)$, then the policy decision is $d = (y_j, x_j, x_k = X - x_j, x_i = 0)$, and the game is over. If the third mover makes a proposal incompatible with either one of the two standing proposals, then the game goes into the next round and starts over with a new omnibus proposal by the same first mover.

Notice that the first mover's proposal in an alternating-offers game contains an entire distribution of the budget $(x_{\mathrm{L}}, x_{\mathrm{C}}, x_{\mathrm{R}})$, but specifies only the first mover's budget share in demand-bargaining. This difference is important because it implies that the first mover specifies the members of the winning coalition in alternating-offers (those with a positive budget share), but does not specify the coalition in demand-bargaining. Instead, there are three different potential coalitions which can approve a policy decision (depending on whose demand is compatible with whose). The equilibrium coalition will consist of the first and the second mover. But the possibility of a coalition between the second and the third mover reduces the rent of the first mover significantly. On the other hand, the possible coalition between the first and the third mover imposes an upper bound on the rent of the second mover as well.

Consider the case of single-issue bills. The first mover fixes the order of play to be jk and makes a policy proposal of (y_i) in the ideological bargaining game, and fixes

[7] Notice that a demand of $x_j < X - x_i$ is also compatible, but then $x_k > 0$. Since the first and second mover together are a majority in this game and can approve a policy without the approval of the third mover, it is never true that $x_j < X - x_i$ in equilibrium.

the order to be lm and makes a demand of (x_i) in the distributive bargaining game. Each of the two independent bargaining games follows the same rules as the game with omnibus bills. Legislator j makes a proposal of (y_j). If compatible $(y_j = y_i)$, then the policy is approved to be $d_{\mathrm{I}} = y_i$, and this particular bargaining game is over. If not compatible, then legislator k can make a proposal compatible with either one of the standing proposals y_i and y_j. If legislator k does so, then the game is over and either $d_{\mathrm{I}} = y_i$ or $d_{\mathrm{I}} = y_j$. Otherwise, the game goes into the next round and starts over (with a new ideological policy demand by the same first mover). Independently, legislator l, second mover in the distributive bargaining game, makes a demand of x_l. If compatible $(x_l = X - x_i)$, then the game is over and $d_{\mathrm{D}} = (x_i, x_l = X - x_i, x_m = 0)$. If not, then legislator m can make a proposal compatible with either one of the two standing proposals, in which case the game is over and the distributive policy is either $d_{\mathrm{D}} = (x_i, x_m = X - x_i, x_l = 0)$ or $d_{\mathrm{D}} = (x_l, x_m = X - x_l, x_i = 0)$. Otherwise, the game goes into the next round and starts over (with a new distributive policy demand by the same first mover).

Each bargaining game has an infinite horizon, i.e., $t \to \infty$. Legislators receive no utility in a particular dimension until a policy is approved in this dimension. Legislators are impatient: Their utility from a decision in a particular dimension decreases with the time it takes to reach the final decision. One can, for example, imagine that the legislators' constituencies do not appreciate the endless delay of policy decisions. To simplify matters, all legislators have the same time preferences, and the discount factor is the same for both issues. However, the model can easily be extended to allow for heterogeneous time preferences and different discount factors for the two issues. Denote the discount factor by $\delta \in [0, 1)$. The author assume that omnibus bills are necessarily approved with one vote, but single-issue bills are approved with two independent votes per legislator. Both issues are necessarily approved in the same round t with an omnibus bill. But they can be approved in different rounds with single-issue bills. Suppose that the ideological policy is approved in some round $t_{\mathrm{I}} \geq 1$, and that the distributive policy is approved in some round $t_{\mathrm{D}} \geq 1$. With omnibus bills, $t_{\mathrm{I}} = t_{\mathrm{D}}$. Taking time preferences into account, the utility of legislator i from decision $d = (y, x_{\mathrm{L}}, x_{\mathrm{C}}, x_{\mathrm{R}})$, approved as one omnibus or two single-issue bills, can be written as

$$U_i(d) = \begin{cases} \delta^{t_{\mathrm{I}}-1} v_i(y) + \delta^{t_{\mathrm{D}}-1} x_i, & t_{\mathrm{I}}, t_{\mathrm{D}} < \infty, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The utility from infinite disagreement is zero, so that legislators prefer implementing any policy decision to infinite disagreement. The rationale behind this assumption is, again, the fact that the legislators' constituencies elect politicians to implement policies and would not be happy with endless delay.

The preferences of the legislators, the rules of the bargaining game, the identity of the first mover, and the order of play (once determined by the first mover) are common knowledge. At any decision point, a legislator knows the history of previous proposals, including all those already made in the current round. Denote by $d_{i,t}$ and $d_{j,t}$ the proposals of the first mover i and the second mover j in the current round t. If no such proposal has yet been made, then $d_{i,t} = \varnothing$ and/or $d_{j,t} = \varnothing$. The strategy of legislator

l in round t with order of play ijk, denoted by $\sigma_l(ijk, d_{i,t}, d_{j,t}, h_t)$, is a function of the order of play and legislator l's position in it, all proposals made in the current round, and h_t, the history of all proposals in rounds $1, \ldots, t-1$.

The equilibrium concept is subgame-perfection. Typically, there are equilibria in legislative bargaining games which depend on punishment strategies. To eliminate such equilibria, the legislative bargaining literature focuses on equilibria in stationary strategies (e.g., Baron and Ferejohn 1989; Banks and Duggan 2000, 2006; Battaglini and Coate 2007; Jackson and Moselle 2002, among others). Stationary strategies are such that $\sigma_l(ijk, d_{i,t}, d_{j,t}, h_t) = \sigma_l(ijk, d_{i,t}, d_{j,t}) \ \forall \ h_t, l$, which means that they are history-independent: Only the information about the current round is relevant for a legislator's strategy. In this article, attention is restricted to pure stationary strategies.

A collection of strategies σ^* is an equilibrium if

1. Legislator i makes a proposal acceptable to either of the following movers if and only if the utility from the proposal to legislator i is at least as large as the utility from making an alternative proposal which will not be accepted;
2. Legislator i accepts a proposal if only if the utility from this proposal to legislator i is at least as large as a that from another standing proposal, if there is one, and also at least as large as legislator i's continuation value of the game (legislator i's expected utility from making a proposal incompatible with any of the standing proposals and inducing the game to continue).

A few tie-breaking assumptions are made : If the first mover is indifferent between omnibus and single-issue bills, then omnibus bills are chosen. If the second mover is indifferent between accepting the first mover's proposal and making a proposal incompatible with it, then the first mover's proposal is accepted. The third mover makes a proposal compatible with the second mover's proposal if it is not strictly worse than any other alternative. If it is, she/he makes a proposal compatible with the first mover's proposal unless this is also strictly worse than her/his continuation value of the game. If the first mover is indifferent with respect to the order of play, then the legislator closer to the first mover's ideal point on $[\hat{y}_{\mathrm{L}}, \hat{y}_{\mathrm{R}}]$ is chosen as the second mover. If both are equally close, then legislator L is chosen.[8]

3 Omnibus bills

In this section, the bargaining game with omnibus bills is considered. Proposition 1 shows that legislators make trade-offs between the two dimensions, using budget

[8] Two of the tie-breaking assumptions are necessary in the sense of Bennett and vanDamme (1991). They are 1) the first mover choosing an omnibus bill if indifferent, and 2) the first mover choosing the ideologically closer legislator as the second mover. If these were changed to the opposite, then for example, the equilibrium would be different in the following sense: Legislators C and R would now choose single-issue bills instead for one specific value of $\hat{y}_{\mathrm{C}}$ each, and the positive budget share would go to the ideologically farther legislator (but would not change in size). None of the qualitative predictions would change. The remaining tie-breaking assumptions are the only possible ones consistent with equilibrium. If they were changed, no indifference would occur in equilibrium. In addition, equilibrium predictions would stay within an ϵ-neighborhood of the present ones.

shares as sidepayments. To make the environment as favorable for omnibus bills as possible, these trade-offs are unconstrained in the sense that the budget is large enough to cover any sidepayment in a mutually beneficial trade.

Assumption 1 The budget X is large:

$$X > -\frac{1}{2}(\hat{y}_{\mathrm{C}} - \hat{y}_{\mathrm{L}})^2 + \frac{1}{16}\left[\left(2\hat{y}_{\mathrm{R}} + \hat{y}_{\mathrm{C}} - 3\hat{y}_{\mathrm{L}}\right)^2 + \left(\hat{y}_{\mathrm{R}} + \hat{y}_{\mathrm{L}} - 3\hat{y}_{\mathrm{C}}\right)^2\right].$$

Proposition 1 shows that the policy decision is approved in round 1. The ideological policy depends on all legislators' ideal points (more heavily on the first mover's ideal point than on the others). The first mover receives less than half of the budget, while the second mover receives more than half of the budget. The third mover receives no budget shares at all.

Proposition 1 (Omnibus Bills) *Suppose that legislator i is the first mover and has chosen the order of play of the remaining legislators to be jk. With omnibus bills and large budget X, the equilibrium policy outcome d^* is*

$$\begin{aligned} y^* &= \tfrac{2\hat{y}_i + \hat{y}_j + \hat{y}_k}{4}, \quad i, j, k \in \mathbf{L}, \\ x_i^* &= \tfrac{X}{2} + \tfrac{1}{2}(\nu_j(y_i^*) - \nu_j(y_j^*) + \nu_k(y_i^*) - \nu_k(y_j^*)), \\ x_j^* &= X - x_i^*, \\ x_k^* &= 0. \end{aligned} \tag{3}$$

The first mover is indifferent with respect to the possible orders of play. The proposal of the first mover is accepted immediately by the second mover, so that the game ends in $t = 1$.

All proofs can be found in the Appendix. An omnibus bill, approved by the first and second mover, makes it necessary for the first mover to allocate the larger share of the budget to the second mover to ensure the approval of the proposed bill. Otherwise, the second mover can form a coalition with the third mover in which no budget shares are allocated to the first mover and the ideological policy is $y_j^* = \frac{\hat{y}_j + \hat{y}_k}{2}$ (see proof of Proposition 1). To avoid this alternative coalition, the second mover has to be paid enough (in some cases, the second mover receives more utility than the first mover). Since coalitions between any two of the three legislators are possible in each round of demand-bargaining, the policy outcomes (including the choice of omnibus and single-issue bills) and the distribution of rents are quite different from alternating-offers.

The ideological policy outcome with omnibus bills is, to some extent, favorable to the first mover, but the budget distribution is not. If the combination of policy outcomes with single-issue bills gives more utility to the first mover, then the first mover will choose single-issue bills. The next section shows that this is indeed true for legislators L and C, and sometimes also for legislator R, depending on the degree of asymmetry in ideal points in the legislature.

4 Single-issue bills

In this section, the two independent bargaining games with single-issue bills are considered. Proposition 2 shows that the approved policies are different from omnibus bills. Legislators cannot make trade-offs between the policy issues, so that the ideological policy is the median ideal point, and the budget is evenly distributed between the first and second mover.

Proposition 2 (Single-Issue Bills) *Independent of the order of play, the ideological policy outcome is* $y^* = \hat{y}_C$. *If legislator* i *is the first mover and has chosen the order of play* jk *in the distributive bargaining game, then*

$$\begin{aligned} x_i^* &= x_j^* = \tfrac{X}{2}, \\ x_k^* &= 0. \end{aligned} \tag{4}$$

In each bargaining game, the first mover is indifferent with respect to the possible orders of play. In each game, the proposal of the first mover is accepted immediately by the second mover, so that each game ends in $t = 1$.

The policy outcomes with single-issue bills and omnibus bills are different, unless legislator C is the first mover and $\hat{y}_C = \frac{\hat{y}_L + \hat{y}_R}{2}$. Therefore, the first mover is typically not indifferent between the two types of bills. In the next section, the preference of the first mover over omnibus and single-issue bills is characterized.

5 Omnibus or single-issue? The choice of the first mover

Consider, for example, legislator R, the most extreme legislator. For this legislator, the implementation of $\hat{y}_C$ with a single-issue bill is more costly than the implementation of $\frac{2\hat{y}_R + \hat{y}_C + \hat{y}_L}{4}$ with an omnibus bill. At the same time, the legislator's budget share with omnibus bills is smaller than with single-issue bills. Proposition 3 shows how legislator R, as well as legislators L and C, resolve this type of trade-off as a first mover, and which type of bill they choose.

Assume that $\hat{y}_L = 0$ and $\hat{y}_R = 1$ and let $\overline{y}_C < 0.5$ define a threshold-level of asymmetry in the location of ideal points in the legislature (asymmetry is large if $\hat{y}_C < \overline{y}_C$, and small otherwise).

Proposition 3 *Legislator L always prefers single-issue bills; legislator C always weakly prefers single-issue bills (and is indifferent only if* $\hat{y}_C = \frac{\hat{y}_L + \hat{y}_R}{2}$). *Legislator R strictly prefers omnibus bills if* $\hat{y}_C < \overline{y}_C < 0.5$, *is indifferent only if* $\hat{y}_C = \overline{y}_C$, *and prefers single-issue bills otherwise.*

The assumption that $\hat{y}_L = 0$ and $\hat{y}_R = 1$ simplifies the analysis; but it is made mainly to pin down the threshold-level of asymmetry in ideal points, $\overline{y}_C$, for which legislator R strictly prefers single-issue bills. If $\hat{y}_C < \overline{y}_C$, then it is very costly for legislator R to implement $\hat{y}_C$ with single-issue bills. In fact, the cost of implementing the ideological policy with a single-issue bill rather than an omnibus bill is so large that it

outweighs the distributive benefit of a single-issue bill: Legislator R prefers omnibus bills. If legislator R is not so extreme, $\hat{y}_C > \overline{y}_C$, then the distributive advantage of single-issue bills prevails.

For legislators L and C, the trade-off is not so distinct. For them, the ideological outcome of a single-issue bill is sufficiently favorable not to forego the benefit in the distributive dimension. Indeed, legislator C can implement her/his policy ideal point $\hat{y}_C$ with single-issue bills. Legislator L prefers the ideological policy with omnibus bills if $\hat{y}_C > \frac{1}{3}$, but the distance to $\hat{y}_C$ is never large enough to outweigh the distributive advantage of a single-issue bill.

6 Conclusion

This article attempts to shed light on reasons for single-issue bills in legislative bargaining. As such, it should not be seen as an alternative to alternating-offers, but as a complement. It is found that single-issue bills can be optimal if the first-mover bargaining power is reduced by the underlying bargaining game (demand-bargaining, in this case). However, legislators prefer single-issue bills only if they are sufficiently moderate in their ideological policy preference. More extreme legislators prefer to propose omnibus bills. While these bills come with a distributive loss for the first mover, the implemented ideological policy is closer to the first mover's ideal point (i.e., sufficiently extreme).

These findings conform with the common sense that mostly unfavorable issues need to be packaged with other issues to make sure they pass. Rules in the US legislation (for example, the Germaneness Rule) forbid non-germane amendments to a bill; in particular, appropriation bills should not be combined with legislation (see, for example, Oleszek 2007). In the light of the findings in this article, this rule seems sensible because only extreme first proposers prefer omnibus bills. While the germaneness principle exists, rules are not always enforced by the Rules Committee and points of order are not always supported by the floor: Packaging of issues exists. Ferejohn (1986), for example, tells the tale of food-stamp legislation becoming part of an agricultural bill as an example for an implicit vote trade through omnibus bills.

The model delivers predictions that can be used for empirical testing of bargaining procedures. So far, empirical tests have not led to clear conclusions. Frechette et al. (2005a), for example, try to distinguish alternating-offers and demand-bargaining from experimental data using the power of (or payoff to) the first mover as the dependent variable. Theoretically, first-mover payoffs in purely distributive alternating-offers and demand-bargaining games are very different: The first mover in alternating-offers games receives a large payoff, while the first-mover payoff in demand-bargaining games is much smaller. Nevertheless, Frechette et al. (2005a) find it impossible to infer the underlying bargaining game from the data because empirically observed behavior is not significantly different in the two types of games: First movers get statistically similar payoffs in both types of games.

According to the theoretical predictions, the binary choice over omnibus and single-issue bills can be used to distinguish bargaining procedures (instead of using a continuous variable such as first-mover payoff). When there is strong proposer

power (alternating-offers), the choice of omnibus bills should be observed. With weak proposer power (demand-bargaining), single-issue bills should be observed (for all types of proposers if the asymmetry in ideal points is small enough, but for moderate ones only if the ideological asymmetry is large). A binary variable as this one is possibly more powerful in distinguishing the two types of games. It will be interesting to see the results of such an empirical test in the future.

Acknowledgements The author would like to thank participants of the 2005 Midwest Political Science Association Meeting, the 2004 Fall Midwest Economic Theory Meetings, and of the seminars at the University of Guelph, Ohio State University, and CORE, Catholic University of Louvain. The author is grateful to Massimo Morelli, Alan Wiseman, François Maniquet, Serkan Ozbeklik, Asha Sadanand, J. Atsu Amegashi, and Mike Hoy for their helpful comments and discussions, and to three anonymous referees for their detailed comments and suggestions. The usual disclaimer applies.

Appendix

Proof of Proposition 1

Equilibria in pure stationary strategies are considered. This implies that only the order of play and demands made in the current round are payoff-relevant. The game will necessarily end in some round t because the utility of eventual agreement is higher than the utility of infinite disagreement. Suppose that the game ends in some round $t > 1$. The order of play, demands made, and demands accepted in this round have to be utility-maximizing for all the legislators. Suppose that the order of play of the remaining legislators, chosen by first mover i, is jk. Denote the maximum utility achieved by legislators by ending the game in round t by $\overline{u}_{it}$, $\overline{u}_{jt}$, and $\overline{u}_{kt}$.

Let us now consider round $t-1$. The first mover is the same as in round t by assumption. As there is no restriction on demands and the order of play in round $t-1$ that does not also apply in round t, legislators (the first mover in particular) can achieve utilities of $\frac{\overline{u}_{it}}{\delta}$, $\frac{\overline{u}_{jt}}{\delta}$, and $\frac{\overline{u}_{kt}}{\delta}$ in round $t-1$. As $\delta < 1$, these are higher than the utilities achieved by ending the game in round t. Therefore, ending the game in round t is not consistent with utility maximization. The same argument applies to all rounds except round 1.

Notice that the utility achievable in round 1 is higher than in any other round in which the game could be ended, given that the first mover is the same in each round. Therefore, there is no scope for strategic delay by any of the legislators, and the utility achievable in later rounds does not impose binding constraints on demands made and accepted in round 1. Therefore, the game can be treated and solved as a one-shot game. This is true also for the two bargaining games with single-issue bills.

Suppose that legislator i is chosen as the first mover and decides that the order of play of the remaining legislators should be jk. Consider the decision of legislator k who can accept the proposal of the first mover, (y_i, x_i), or the proposal of the second mover, (y_j, x_j). For these proposals, $x_k = X - x_i$ or $x_k = X - x_j$, respectively. The third mover accepts the proposal that maximizes her/his utility, or, if indifferent, the proposal of the second mover according to the tie-breaking assumption.

Consider now the decision of legislator j. For a given (y_i, x_i), the maximum utility achieveable by not accepting the first mover's proposal, but by making a counterproposal accepted by third mover k is given by

$$\begin{aligned} &\max_{y_j, x_j} \ v_j(y_j) + x_j \\ &\text{s.t. } v_k(y_j) + X - x_j = v_k(y_i) + X - x_i. \end{aligned} \tag{5}$$

Notice that the second mover can always propose $x_j = x_i$ and $y_j = y_i$, which will be accepted by the third mover. The second mover's maximization problem leads to the following conditions for y_j^* and x_j^* (y_j^* and x_j^* are budget share and ideological policy in legislator j's counterproposal in equilibrium):

$$\frac{\partial v_j(y_j)}{\partial y_j} + \frac{\partial v_k(y_j)}{\partial y_j} = 0, \tag{6}$$

$$x_j^* = v_k(y_j^*) - v_k(y_i) + x_i. \tag{7}$$

Legislator i needs to make a proposal that is accepted by the second mover (it needs to make the second mover no worse off than a coalition between the second and the third mover). It is not possible for the first mover to increase utility by making a proposal that is not accepted by the second mover, but is accepted by the third mover: Given such a proposal, the second mover can always propose $x_j = x_i$ and $y_j = y_i$ and leave the first mover without budget shares. The maximization problem of the first mover is as follows:

$$\begin{aligned} &\max_{y_i, x_i} \ v_i(y_i) + x_i \\ &\text{s.t.} \quad v_j(y_j^*) + x_j^* = v_j(y_i) + X - x_i. \end{aligned} \tag{8}$$

This leads to the following conditions for y_i^* and x_i^*:

$$\frac{\partial v_i(y_i)}{\partial y_i} + \frac{1}{2}\left[\frac{\partial v_j(y_i)}{\partial y_i} + \frac{\partial v_k(y_i)}{\partial y_i}\right] = 0, \tag{9}$$

$$x_i^* = \frac{X}{2} + \frac{1}{2}\left[v_j(y_i^*) - v_j(y_j^*) + v_k(y_i^*) - v_k(y_j^*)\right]. \tag{10}$$

With $v_i(y) = k - b(y - \hat{y}_i)^2$, the solutions to Eqs. 6, 9, and 10 are unique and such that

$$y_i^* = \frac{2\hat{y}_i + \hat{y}_j + \hat{y}_k}{4}, \tag{11}$$

$$y_j^* = \frac{\hat{y}_j + \hat{y}_k}{2}, \tag{12}$$

$$x_i^* = \frac{X}{2} + \tfrac{1}{2}\left[v_j(y_i^*) - v_j(y_j^*) + v_k(y_i^*) - v_k(y_j^*)\right]. \tag{13}$$

The assumption of a large budget makes sure that $x_i^* > 0$. With the tie-breaking assumption, the second mover makes a demand compatible with the first mover's demand, so that the game ends immediately after the second mover's acceptance. The policy outcomes are y_i^* and $(x_i^*, x_j^* = X - x_i^*, x_k^* = 0)$. From Eqs. 11–13, it is obvious that the first mover is indifferent with respect to the order of play. With the tie-breaking assumption, the first mover chooses the ideologically closer legislator as the second mover. □

Proof of Proposition 2

In the following two lemmas, the author characterize the policy outcome of each of the two bargaining games for single-issue bills separately. The two lemmas prove Proposition 2.

Lemma 1 *In the ideological bargaining game, the policy outcome is* $d_{\mathrm{I}}^* = \hat{y}_{\mathrm{C}}$.

Proof The same arguments apply as in the first paragraphs of the proof of Proposition 1, so that it suffices to treat the game as a one-shot game ending in round 1. Suppose that legislator i is the first mover and has chosen order of play jk. There are two cases to consider:

Case 1: Legislator C is the first mover.

The proof is similar to the proof of Proposition 1. For a given y_i, the second mover maximizes

$$\begin{aligned} &\max_{y_j} \nu_j(y_j) \\ &\text{s.t.} \quad \nu_k(y_j) \geq \nu_k(y_i). \end{aligned} \tag{14}$$

The second mover's and the third mover's ideal points are on opposite sides of $\hat{y}_{\mathrm{C}}$. Each y_j which makes the second mover better off than y_i necessarily makes the third mover worse off. So, $y_j^* = y_i$ (y_j^* is the ideological policy in the second mover's counterproposal in equilibrium).

Given $y_i = y_j^*$, the first mover's maximization problem is

$$\begin{aligned} &\max_{y_i} \nu_i(y_i) \\ &\text{s.t.} \quad \nu_j(y_i) \geq \nu_j(y_j). \end{aligned} \tag{15}$$

The first mover, legislator C, implements her/his ideal point: $y_i^* = y^* = \hat{y}_{\mathrm{C}}$.

Case 2: The median legislator is not the first mover.

Suppose first that legislator C is the second mover, $j = \mathrm{C}$. In this case, the first mover's and the third mover's ideal points are on opposite sides of $\hat{y}_{\mathrm{C}}$. Suppose that y_i is between $\hat{y}_i$ and $\hat{y}_{\mathrm{C}}$, but not equal to $\hat{y}_{\mathrm{C}}$. The second mover proposes $\hat{y}_{\mathrm{C}}$, which makes both the second and the third mover better off. Suppose instead that the first mover proposes $\hat{y}_{\mathrm{C}}$. This is accepted by the second mover. Any policy between $\hat{y}_{\mathrm{C}}$

and $\hat{y}_k$, but unequal to $\hat{y}_C$, is worse for the first mover than $\hat{y}_C$ and would never be proposed. The policy outcome is $y_i^* = y^* = \hat{y}_C$.

Suppose instead that legislator C is the third mover, $k = C$. In this case, the first mover's and the second mover's ideal points are on opposite sides of $\hat{y}_C$. If the first mover proposes some policy y_i between $\hat{y}_i$ and $\hat{y}_C$, but different from $\hat{y}_C$, then the second mover j can get higher utility by proposing a policy between $\hat{y}_C$ and $\hat{y}_j$ which leaves the third mover, legislator C, indifferent. Such a policy necessarily makes the second mover better off. Only if $y_i = \hat{y}_C$, this is not possible. To maximize utility, the first mover proposes $\hat{y}_C$, and the second mover accepts. The policy outcome is $y_i^* = y^* = \hat{y}_C$.

The policy outcome does not depend on the order of play, and the first mover is indifferent with respect to the order of play. □

Lemma 2 *For the distributive bargaining game with legislator i as the first mover and order of play ijk, $d_D^* = (x_i^* = x_j^* = \frac{X}{2}, x_k^* = 0)$.*

Proof For the proof, refer to Morelli (1999)'s proof of Proposition 1(bargaining game over a single, purely distributive policy) because it can easily be adapted to the present game. □

Proof of Proposition 3

Assume that the first mover, legislator i, fixes the order of play of the remaining legislators to be jk. Below, the policy outcomes indexed by $*$ are the equilibrium policy outcomes of omnibus bills. With Propositions 1 and 2, legislator i prefers an omnibus bill if

$$\nu_i(y_i^*) + \frac{X}{2} + \frac{1}{2}\left(\nu_j(y_i^*) - \nu_j(y_j^*) + \nu_k(y_i^*) - \nu_k(y_j^*)\right) \geq \nu_i(\hat{y}_C) + \frac{X}{2},$$

or if

$$\frac{1}{2}\left[\nu_j(y_i^*) - \nu_j(y_j^*) + \nu_k(y_i^*) - \nu_k(y_j^*)\right] \geq \nu_i(\hat{y}_C) - \nu_i(y_i^*). \tag{16}$$

Denote the left-hand side of Eq. 16 by $D(i, y_i^*(ijk), y_j^*(ijk))$. It represents the first mover's loss in the distributive dimension from an omnibus bill relative to a single-issue bill. Denote the right-hand side of Eq. 16 by $I(i, \hat{y}_C, y_i^*(ijk))$: It represents the first mover's utility loss (or gain) in the ideological dimension from a single-issue bill relative to an omnibus bill. First mover i prefers an omnibus bill if $D(i, y_i^*(ijk), y_j^*(ijk)) \geq I(i, \hat{y}_C, y_i^*(ijk))$. Generally, $D(i, y_i^*(ijk), y_j^*(ijk)) \leq 0 \; \forall \, i$, while $I(i, \hat{y}_C, y_i^*(ijk)) \lesseqgtr 0$.

Case 1: Legislator L is the first mover.

Legislator L is indifferent with respect to the order of play, but chooses order of play CR according to the tie-breaking rule. Recall that $\hat{y}_L = 0$ and $\hat{y}_R = 1$, so that

$y_L^* = \frac{\hat{y}_C+1}{4}$ and $y_C^* = \frac{\hat{y}_C+1}{2}$ and $I(L, \hat{y}_C, y_L^*) = -\hat{y}_C^2 + \frac{1}{16}(\hat{y}_C+1)^2$, $D(L, y_L^*, y_C^*) = -\frac{1}{16}(\hat{y}_C+1)^2$, and $D(L, y_L^*, y_C^*) < I(L, \hat{y}_C, y_L^*)$ for all $\hat{y}_C \leq 0.5$. Legislator L always strictly prefers single-issue bills.

Case 2: Legislator C is the first mover.

Legislator C chooses order of play LR. In this case, $y_C^* = \frac{2\hat{y}_C+1}{4}$ and $y_L^* = \frac{1}{2}$. Therefore, $I(\mathrm{C}, \hat{y}_C, y_C^*) = (\frac{1-2\hat{y}_C}{4})^2 \geq 0$ and $D(\mathrm{C}, y_C^*, y_L^*) = -(\frac{2\hat{y}_C-1}{4})^2 \leq 0$: Legislator C strictly prefers single-issue bills unless $\hat{y}_C = \frac{1}{2}$, in which case legislator C is indifferent.

Case 3: Legislator R is the first mover.

Legislator R chooses order of play CL. In this case, $y_R^* = \frac{\hat{y}_C+2}{4}$ and $y_C^* = \frac{\hat{y}_C}{2}$. Therefore, $I(\mathrm{R}, \hat{y}_C, y_R^*) = -(\hat{y}-1)^2 + (\frac{\hat{y}_C-2}{4})^2 < 0$ and $D(\mathrm{R}, y_R^*, y_C^*) = -(\frac{\hat{y}_C-2}{4})^2 < 0$. There exists $\overline{y}_C$ such that $D(\mathrm{R}, y_R^*, y_C^*) > I(\mathrm{R}, \hat{y}_C, y_R^*)$ for $\hat{y}_C < \overline{y}_C$ and $D(\mathrm{R}, y_R^*, y_C^*) < I(\mathrm{R}, \hat{y}_C, y_R^*)$ for $\hat{y}_C > \overline{y}_C$. Solving $D(\mathrm{R}, y_R^*, y_C^*) - I(\mathrm{R}, \hat{y}_C, y_R^*) = 0$ yields $\overline{y}_C = \frac{6-\sqrt{8}}{7} \approx 0.453$.[9] □

References

Banks JS, Duggan J (2000) A bargaining model of collective choice. Am Polit Sci Rev 94(1):73–88

Banks JS, Duggan J (2006) A general bargaining model of legislative policy-making. Q J Polit Sci 1:49–85

Baron DP, Ferejohn JA (1989) Bargaining in legislatures. Am Polit Sci Rev 83(4):1181–1206

Battaglini M, Coate S (2007) Inefficiencies in legislative policy-making: a dynamic analysis. Am Econ Rev 97(1):118–149

Battaglini M, Coate S (2008) A dynamic theory of public spending, taxation, and debt. Am Econ Rev 98(1):201–236

Bennett E, vanDamme E (1991) Demand commitment bargaining, the case of apex games. In: Selten R (ed) Game equilibrium models III, strategic bargaining. Springer Verlag, Berlin

Bernheim BD, Rangel A, Rayo L (2006) The power of the last word in legislative policy making. Econometrica 74(5):1161–1190

Diermeier D, Eraslan H, Merlo A (2003) A structural model of government formation. Econometrica 71(1):27–70

Diermeier D, Swaab RI, Medvec VH, Kern MC (2008) The micro-dynamics of coalition formation. Polit Res Q 61(3):484–501

Ferejohn J (1986) Logrolling in an institutional context: a case study of food stamp legislation. In: Wright GCJr, Rieselbach LN, Dodd LC (eds) Congress and policy change. Agathon Press, Inc., New York

Frechette G, Kagel JH, Morelli M (2005a) Behavioral identification in coalitional bargaining: an experimental analysis of demand bargaining and alternating offers. Econometrica 73(6):1893–1938

Frechette G, Kagel JH, Morelli M (2005b) Nominal bargaining power, selection protocol, and discounting in legislative bargaining. J Public Econ 89(8):1497–1517

Heifetz A, Ponsati C (2007) All in good time. Int J Game Theory 35:521–538

Horstmann IJ, Markusen JR, Robles J (2005) Issue linking in trade negotiations: Ricardo revisited or no pain no gain. Rev Int Econ 13(2):185–204

In Y, Serrano R (2003) Agenda restrictions in multi-issue bargaining (ii): unrestricted agendas. Econ Lett 79:325–331

In Y, Serrano R (2004) Agenda restrictions in multi-issue bargaining. J Econ Behav Organ 53:385–399

Inderst R (2000) Multi-issue bargaining with endogenous agenda. Games Econ Behav 30:64–82

[9] The second solution, $\overline{y}_C = \frac{6+\sqrt{8}}{7}$, is outside of the range of ideal points.

Jackson MO, Moselle B (2002) Coalition and party formation in a legislative voting game. J Econ Theory 103:49–87
Kalandrakis T (2006) Proposal rights and political power. Am J Polit Sci 50(2):441–448
Lang K, Rosenthal RW (2001) Bargaining piecemeal or all at once?. Econ J 111:526–540
Leblanc W, Snyder JM, Tripathi M (2000) Majority-rule bargaining and the under-provision of public investment goods. J Public Econ 75:21–47
Montero M, Vidal-Puga JJ (2007) Demand commitment in legislative bargaining. Am Polit Sci Rev 101(4):847–850
Morelli M (1999) Demand competition and policy compromise in legislative bargaining. Am Polit Sci Rev 93(4):809–820
Morelli M (2007) Demand commitment and legislative bargaining: a response. Am Polit Sci Rev 101(4):851
Oleszek WJ (2007) Congressional procedures and the policy process. CQ Press, Washington, DC
Rubinstein A (1982) Perfect equilibrium in a bargaining model. Econometrica 50:97–109
Snyder JM, Ting MM, Ansolabehere S (2005) Legislative bargaining under weighted voting. Am Econ Rev 95(4):981–1004

Soc Choice Welf
DOI 10.1007/s00355-010-0510-5

Bargaining over the budget

Daniel Diermeier · Pohan Fong

Abstract This article presents a theory of government expenditure and identifies how an inefficient government budget is shaped by its initial size and allocation. Assuming that the parties in the legislative body agree with the optimal size of a government budget but have conflict of interests over its allocation, we show that, if the initial budget size is sufficiently large and the initial allocation is sufficiently unequal, in equilibrium the budget size is greater than what it would be had the initial budget size been sufficiently small.

1 Introduction

While the size and allocation of government budgets vary significantly over time and across countries, all contemporary governments routinely spend a large proportion of national income. Understanding the political and institutional determinants of government expenditures has consequently been one of the most important goals of research

D. Diermeier
Department of Managerial Economics and Decision Sciences (MEDS),
Kellogg School of Management, Northwestern University, Evanston, IL, USA

D. Diermeier
Ford Motor Company Center for Global Citizenship,
Kellogg School of Management, Northwestern University, Evanston, IL, USA

P. Fong (✉)
Department of Economics and Finance, City University of Hong Kong,
83 Tat Chee Avenue, Kowloon, Hong Kong
e-mail: phfong@cityu.edu.hk

Published online: 21 December 2010

in political economy.[1] One recurrent theme is whether and under what conditions governments overspend, i.e., choose a budget that is too large given the economic environment. Various mechanisms have been proposed to address this issue. In a classic article Meltzer and Richard (1981) show that if government expenditure is modeled as a one-dimensional collective choice variable financed by proportional income tax, the median voter theorem implies that with an income distribution skewed to the right, spending and taxation will be greater than the social optimum. In legislative procedures in which all spending decisions are decentralized but financing is centrally provided, a common pool problem arises and leads to inefficient overspending; each decision maker internalizes fully the benefit of its own spending category but only a fraction of the social cost due to taxation (Weingast et al. 1981; Tabellini 1987). In two-party politics a ruling party may spend and borrow too much because it enhances the incumbent's likelihood of reelection (Aghion and Bolton 1990), or because the future cost of debt repayment is not fully internalized by the incumbent due to the possibility of power turnover (Persson and Svensson 1989; Tabellini and Alesina 1990). Overspending may also result from strategic delegation, since voters in each electoral district may have incentives to elect a representative who spends aggressively as a tough bargaining agent in legislative bargaining over government expenditure (Chari et al. 1997).

In this article we propose a new mechanism of overspending and show how the choice of a government budget is shaped by the status quo, i.e., the budget currently in effect. We model a government budget as a multidimensional policy which specifies not only the size of total spending but also its allocation over various government programs that benefit different parties in the legislative body. We consider various institutions in which one party effectively controls the proposal power; yet, its budget proposal needs to be approved by some or all of the other parties. This general approach has been common in legislative models since it was first introduced by Romer and Rosenthal (1978). Such proposer models capture both authoritarian regimes characterized by "collective leadership", where a consensus is required among government officials, and majoritarian institutions, where a budget proposal by a committee or the executive needs to be approved in the legislature.

Like Romer and Rosenthal, we assume that any new policy must defeat the status quo in the legislative procedure. However, in contrast to their model we focus not only on the size of total spending, but also on the budget allocation. The existence of a pre-existing status quo may seem unusual in the context of legislative models as the "policy" considered in our theory is a government budget. After all, following Rubinstein (1982) and Baron and Ferejohn (1989), it has been a common assumption in legislative decision making that no private benefit is distributed or no public good is provided if no policy proposal is approved in the legislative process.[2] But one should note that in modern democratic countries a substantial proportion of total government spending is conducted in the form of entitlement programs. For instance, in fiscal

[1] Persson and Tabellini (2000) provide an extensive survey and discussion of the literature on the political economy of government expenditure. For recent developments, see Persson and Tabellini (2003, 2004), Hassler et al. (2003, 2005), Volden and Wiseman (2007), and Battaglini and Coate (2008).

[2] Noticeable examples include Morelli (1999), Persson et al. (2000) and Volden and Wiseman (2007).

year 2007 the United States federal government spent 1,604.9 trillion dollars (which composed more than half of the federal budget) on pensions, health care, welfare, and veterans.[3] A major part in these spending categories consists of entitlement programs that, once enacted, remain in effect until they are replaced by new legislations. In many cases, those programs do not have a predefined termination clause, a so-called "sunset provision", and the government is obliged by law to distribute the benefits. For example, based on the Social Security Act of 1935, beneficiaries can sue the government if their entitled benefits are withheld. These observations jointly justify our assumption that the status quo policy remains in effect if no new policy proposal is approved in the legislative procedure. This insight is reflected in a recent growing literature on dynamic legislative bargaining in which a new policy must defeat the status quo that was the policy chosen previously. Examples include Baron (1996), Kalandrakis (2004), Baron et al. (2008), and Duggan and Kalandrakis (2010), to name just a few.

We analyze various models that are intended to formally capture different institutions, including corporatist politics, parliamentary systems, and presidential systems. In each model we characterize the conditions for an inefficient budgetary decision. For each institution we show that if the status quo budget size is sufficiently small, then in equilibrium spending on each party in the voting coalition increases and total spending expands to the coalitionally efficient size. Whenever it is coalitionally welfare enhancing to tax more, the agenda setter raises taxes, spends part of the extra tax revenues on some of the other parties just sufficient to obtain the required voting support, and captures the rest of additional public resources. However, the situation is different when the status quo budget size is sufficiently large. In order for the agenda setter to enjoy the benefits of a coalitionally welfare-enhancing budget contraction, it needs to be able to cut down the government programs targeted at the other parties. If the status quo budget has already allocated few public resources to those parties from which the agenda setter seeks voting support, then there will not be sufficient room for the agenda setter to do so. In an extreme case the agenda setter must reduce total government expenditure down to the point that nothing is spent on its coalition partners. Beyond that the agenda setter has to bear any additional budget cut solely by itself. Since the agenda setter only partially internalizes the economic benefits from a coalitionally welfare-enhancing budget cut, it may not have the incentive to cut the budget down to the coalitionally efficient size. This implies that if the status quo budget size is sufficiently large and the status quo allocation is sufficiently unequal, then in equilibrium the budget shrinks but still stays greater than the budget size that would have been reached had the initial budget size been sufficiently small. The mapping from the initial budget size and initial allocation to an equilibrium budget size in a political bargaining environment is new to the literature and was not yet implied by the other existing mechanisms that account for inefficient government spending, such as the common pool arguments.

Overall, our theory implies that government spending is easy to expand but difficult to cut. This may provide an account for the so-called "ratchet effect", which states that if an external shock such as a war or a financial crisis raises government expenditures,

[3] The numbers are from the Budget of the United States Government: Fiscal Year 2007–2009.

expenditures after the shock will not fall all the way back to their pre-crisis level.[4] In addition, we show that there is a positive correlation between the size and allocative inequality of a budget proposed and approved in equilibrium. Moreover, *ceteris paribus* the more unequal the status quo allocation of the budget, the larger is the budget size proposed and approved in the subsequent legislative sessions. These theoretical results jointly suggest that budget size should not be isolated from budget allocation in an empirical investigation on the determinants of government expenditure.[5]

Whereas the results of ratchet effect hold across political systems, political institutions do matter. For the same status quo, it is more likely for the president in a presidential system to resist proposing a deep budget cut compared to a prime minister in a parliamentary system. These results are in line with an increasingly prominent literature that emphasizes the role of legislative procedures and government organizations in shaping fiscal policies and government expenditures. Various theories in this institutional approach have been developed based on variants of the agenda-setting model of Romer and Rosenthal (1978) and the legislative bargaining model of Baron and Ferejohn (1989), where legislative procedures are assumed to be a sequence of proposal making and voting. Persson et al. (1998, 2000) present comparative studies of electoral rules and political regimes, and conclude that parliamentary regimes and/or proportional electoral systems lead to a larger government size compared to the other constitutional arrangements. Milesi-Ferretti et al. (2002) extend the analytical framework of Persson and Tabellini to investigate the constitutional effects on the composition of government expenditure. Recently Battaglini and Coate (2007, 2008) incorporate the legislative institution into a full-fledged dynamic public finance environment, identify the source of policy inefficiency, and characterize the dynamics of taxation, public debt, and government spending.

The remainder of this article is organized as follows. Section 2 describes the policy environment. Sections 3 and 4 present the main findings for two-player and three-player models, respectively, and compare different institutional environments. Section 5 discusses possible extensions of our theory and concludes. All proofs can be found in Appendix A. Appendix B extends the three-player model to incorporate the possibility of reconsideration.

2 The policy environment

Consider a collective decision-making body that consists of $n \geq 2$ parties. Each party i is associated with a government program. A policy $x = (x_1, \ldots, x_n) \in \mathbf{R}^n_+$ specifies the budget allocation $x_i \geq 0$ for each program i and amounts to a total spending of $\Sigma_i x_i$. Policy choice involves both size and allocation of the government budget, so

[4] See Baumol (1967), Borcherding (1977) and Higgs (1987) for an extensive discussion of the ratchet effect, and Persson and Tabellini (2003, 2004) for a panel data analysis regarding how government expenditures respond to shocks under various constitutional arrangements.

[5] See Persson and Tabellini (2003, 2004) for cross-sectional and panel data analysis, as well as a survey of the related empirical literature.

the total spending is also endogenously determined. Each program i produces highly concentrated benefits to party i and diffuse costs to everyone. In order to finance a total spending of $g \geq 0$, each party incurs a cost given by

$$\kappa(g) = \frac{g^2}{4\Pi}$$

With any policy x, each party i thus derives a net utility of

$$u_i(x) = x_i - \kappa\left(\Sigma_i x_i\right).$$

The linear-quadratic setup admits a clean and intuitive closed-form solution.

This policy environment can be interpreted in a public finance context. Imagine that each party represents a particular socioeconomic group in the society, and the budget can be allocated for spending on various public programs, with each targeting at a different group. For example, the society could be divided into the retirees, the working poor, and the working rich; the government may enforce a social security policy that redistributes resources to the retired elderly, provide public education and health insurance that especially benefit low income families, and provide generous tax credits for owning a second home which benefit the wealthy group. The cost of these expenditures thus can be thought of as forgone private consumption and the deadweight efficiency loss due to distortionary taxation that is not explicitly modeled here. With such an interpretation, an equal cost-sharing arrangement assumed here can be justified by the fact that the government is unable, or constitutionally forbidden, to discriminate various socioeconomic groups through taxation based on their party affiliations or preferences over different government programs.

On the one hand, the policy space considered here generalizes the divide-the-dollar framework and allows the "size of the dollar" to be a choice variable as well. On the other hand, it generalizes the "pork barrel" policy domain modeled by Bernheim et al. (2006), with the additional feature that the budget for each program is endogenously determined instead of exogenously given.

In the next two sections we consider various legislative institutions in which policies are chosen through a sequence of proposal making and voting. We commonly refer to the proposing party as the agenda setter. For tractability we only consider models with two or three parties, although the intuitions developed can be carried over to cases with more parties.[6] In this article we also restrict our attention to single-period models and implicitly assume that any policy choice is irrevocable and not subject to reconsideration once approved. In Appendix B we show how the model could be extended to allow for the possibility of reconsideration during the bargaining process. A full investigation of dynamic bargaining environments is left for future work.

[6] As is evident later, the main results in this article rest in the corner solutions. With more parties in the model, there would arise a growing number of different types of corner solutions, in which different combinations of parties are at their constraints.

3 A two-player model

3.1 The political process

In this section we consider a legislative body with two parties, labeled $i \in \{a, r\}$, and model the political process as an ultimatum game. Specifically, the agenda setter (a) first makes a policy proposal $x \in \mathbf{R}^2_+$, which is then subject to approval by the opposition party (r). The proposal is implemented if party r approves. Otherwise a status quo $q \in \mathbf{R}^2_+$ is implemented. The status quo is exogenously given, and it can be any policy alternative that is technologically feasible. We interpret the status quo as the policy that has been in place prior to the beginning of the game considered here. We assume that party r always approves a proposal when indifferent.

The two-player model is the simplest case to illustrate the basic theoretical mechanism at work. At the same time it also captures political systems where legislative decisions effectively take place outside of a democratic, majoritarian institution. In corporatist politics, a budgetary decision may be effectively governed by bargaining between centrally organized interest groups, for example, a national union and a corresponding industry association (Katzenstein 1985; Schmitter and Lehmbruch 1979). In an authoritarian regime like China, the collective decision-making process is characterized by "collective leadership" and high officials in the Politburo's Standing Committee negotiate over any major policy choice without a meaningful formal vote (Xu 2010).

3.2 Analysis

We characterize the unique subgame perfect equilibrium for this game. Given any status quo q, the equilibrium policy $h(q) = (h_a(q), h_r(q)) \in \mathbf{R}^2_+$ is such that the total budget size, $H(q) \equiv \Sigma_i h_i(q)$, and party r's budget allocation, $h_r(q)$, jointly solve the agenda setter's maximization problem:

$$\begin{aligned} \max_{g, x_r} \ & (g - x_r) - \kappa(g) \\ \text{s.t.} \ & x_r - \kappa(g) \geq u_r(q), \\ & g - x_r \geq 0, \\ & x_r \geq 0. \end{aligned} \tag{1}$$

The agenda setter's own budget allocation, $h_a(q)$, is residually determined. The first inequality, an incentive-compatibility constraint, requires that in equilibrium party r be offered at least the utility derived from the status quo. The other two inequalities require that expenditures on parties a and r be nonnegative.

We say a policy is *efficient* if it maximizes the joint utility for the two parties. In the current setup with two parties, any efficient policy amounts to a total spending of Π, at which level the total cost incurred by the last dollar spent equals the additional utility contributed by that dollar through the government programs. Henceforth in this section we refer to Π as the efficient budget size.

In the rest of the analysis we discuss three cases based on how the status quo budget size is compared with the efficient budget size. First, suppose that the status quo

is already an efficient policy, i.e., $\Sigma_i q_i = \Pi$. In this case, any deviation from the efficient budget size would reduce the joint utility for the two parties. At the same time in order to pass any new policy the agenda setter needs to offer party r at least its reservation utility, $u_r(q)$. This implies that an attempt to deviate from the efficient budget size would make the agenda setter unambiguously worse off. Thus efficiency is sustainable.

Second, suppose that the status quo budget is below the efficient size, i.e., $\Sigma_i q_i < \Pi$. Then in equilibrium the budget must rise to the efficient level. As the budget increases from $\Sigma_i q_i$ to Π, the joint utility gets maximized. At the same time, the agenda setter can always allocate part of the additional budget to party r so as to compensate the addition cost that r needs to bear due to a rising budget and to make r indifferent. Thus the maximization of the joint utility is equivalent to the maximization of the agenda setter's own utility, as any additional utility created by the new policy would be captured entirely by the agenda setter. In equilibrium, party r has an allocation of $h_r(q) = q_r + [\kappa(\Pi) - \kappa(\Sigma_i q_i)]$, the agenda setter gets $\Pi - h_r(q)$, and efficiency is achieved. Since party a as the agenda setter also needs to bear the additional cost due to budget expansion, the fact that a is better off implies that the allocation for a must rise as well.

Third, suppose that the status quo budget is above the efficient size, i.e., $\Sigma_i q_i > \Pi$. The intuition developed previously may suggest that party a as the agenda setter would then cut down the budget to the efficient size. Potentially a could do so by offering party r just enough to break even and then taking the rest of the budget. However, this is true only if the status quo allocation is sufficiently egalitarian, in the sense that party r's share in the status quo, $q_r / \Sigma_i q_i$, is neither too small nor too large.

Let us hold constant a status quo budget size that is inefficiently too large, and begin with a situation in which the opposition party r is allocated a sufficiently small proportion of the budget in the status quo, or equivalently, the agenda setter is allocated a sufficiently large share. With a budget cut, both parties would be partially relieved from the high costs associated with the status quo. Thus in order to keep r at its reservation utility, the agenda setter needs to cut spending on r. However, given that the status quo allocation for r is already small, the agenda setter may not have room to cut party r's spending down to the level at which r just breaks even. This is because the budget allocation for r cannot be negative and would be constrained at zero in this case. As soon as the opposition party is at its constraint, any further budget contraction must solely rely on cutting spending for the agenda setter alone. This, however, may not be optimal for the agenda setter as it bears only half of the total costs of the budget and, therefore, only internalizes half of the gains from any deeper budget cut. Ironically, although the agenda setter could increase the joint utility by downsizing the budget, the additional surplus created by a smaller yet more efficient budget may mostly be enjoyed by the opposition party. At such a corner solution, the resistance to a deep budget cut comes from the party with proposal power.

Now consider an opposite situation in which the opposition party r is allocated a sufficiently large share of the budget in the status quo, or equivalently, the agenda setter is allocated a sufficiently small share. In this case it is easy for spending on the agenda setter to be at constraint in a fiscal contraction. Once the allocation for the agenda setter is down to zero, any further budget cut requires that this be achieved by

cutting spending on party r alone. At the same time party r internalizes only half of the gains from a further budget cut. As a consequence party r would reject any budget cut sufficiently deep and in equilibrium the budget ends up above the efficient size. In this situation a different corner solution arises, and the resistance to an efficient fiscal contraction comes from the opposition party.

Finally, if the status quo allocation is sufficiently egalitarian, then neither corner solutions occur. In equilibrium no party resists an efficient fiscal contraction and the budget shrinks to its efficient size. Our thought experiment thus implies that, holding constant a status quo budget that is inefficiently too large, the equilibrium budget size is a U-shaped function of the status quo allocation. In particular, inequality leads to inefficiency.

The preceding analysis suggests a general point that the equilibrium budget size depends on budget size as well as *allocation* of the status quo. Proposition 1 summarizes the results. The proof is presented in Appendix A.

Proposition 1 *Consider a two-player ultimatum game.*

1. *Suppose that the status quo budget size is inefficiently too small. Then in equilibrium the spending for each party increases and the budget expands to the efficient size. Specifically, for any $q \in \mathbf{R}^2_+$ such that $\Sigma_i q_i < \Pi$, the equilibrium policy $h(q)$ is such that*

$$h_a(q) = q_a + \frac{1}{4\Pi}\left(\Pi - \Sigma_i q_i\right)\left(3\Pi - \Sigma_i q_i\right) > q_a,$$
$$h_r(q) = q_r + \frac{1}{4\Pi}\left(\Pi - \Sigma_i q_i\right)\left(\Pi + \Sigma_i q_i\right) > q_r,$$

   and $H(q) = \Pi$.
2. *If the status quo is efficient, then it remains in equilibrium. In other words, for any $q \in \mathbf{R}^2_+$ such that $\Sigma_i q_i = \Pi$, the equilibrium policy is $h(q) = q$.*
3. *Suppose the status quo budget is inefficiently too large. If the status quo budget allocation is relatively egalitarian in the sense that $\Phi_1(q) < q_r/\Sigma_i q_i \leq \Phi_2(q)$, where*

$$\Phi_1(q) \equiv \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) - \frac{1}{4}\left(\frac{\Pi}{\Sigma_i q_i}\right), \tag{2}$$

   and

$$\Phi_2(q) \equiv \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) + \frac{3}{4}\left(\frac{\Pi}{\Sigma_i q_i}\right), \tag{3}$$

   then in equilibrium the spending for each party increases and the budget shrinks to the efficient size. On the other hand, if he status quo budget allocation is sufficiently unequal in the sense that either $q_r/\Sigma_i q_i \leq \Phi_1(q)$ or $q_r/\Sigma_i q_i > \Phi_2(q)$, then in equilibrium the budget shrinks but still ends up greater than the efficient size with one party capturing the entire budget and the other getting nothing. Moreover, a

more unequal status quo allocation leads to a weakly greater equilibrium budget size. Specifically, for any $q \in \mathbf{R}^2_+$ such that $\Sigma_i q_i > \Pi$, the equilibrium policy $h(q)$ satisfies all of the following properties:

A. *If $0 \leq q_r/\Sigma_i q_i \leq \Phi_1(q)$, then*

$$h_a(q) = \min\left\{2\Pi, (\Sigma_i q_i)\left[1 - 4\left(\frac{\Pi}{\Sigma_i q_i}\right)\left(\frac{q_r}{\Sigma_i q_i}\right)\right]^{\frac{1}{2}}\right\},$$

$h_r(q) = 0$, and $\Pi < H(q) = h_a(q) \leq 2\Pi$. Moreover, $H(q)$ is weakly decreasing in $q_r/\Sigma_i q_i$ holding $\Sigma_i q_i > \Pi$ constant.

B. *If $\Phi_1(q) < q_r/\Sigma_i q_i \leq \Phi_2(q)$, then*

$$h_a(q) = q_a - \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi} - 1\right)(3\Pi - \Sigma_i q_i) < q_a,$$
$$h_r(q) = q_r - \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi} - 1\right)(\Pi + \Sigma_i q_i) < q_r,$$

and $H(q) = \Pi$.

C. *If $\Phi_2(q) < q_r/\Sigma_i q_i \leq 1$, then $h_a(q) = 0$,*

$$h_r(q) = 2\Pi - 2(\Sigma_i q_i)\left[\left(\frac{\Pi}{\Sigma_i q_i}\right)^2 + \frac{1}{4} - \left(\frac{\Pi}{\Sigma_i q_i}\right)\left(\frac{q_r}{\Sigma_i q_i}\right)\right]^{\frac{1}{2}},$$

and $\Pi < H(q) = h_r(q) \leq 2\Pi$. Moreover, $H(q)$ is increasing in $q_r/\Sigma_i q_i$ holding $\Sigma_i q_i > \Pi$ constant.

Proposition 1 implies, roughly speaking, a positive correlation in equilibrium between the budget size and the inequality of budget allocation. The most unequal allocation of public resources is achieved only in cases in which the budget ends up above the efficient size.

Part 3 of Proposition 1 also suggests a novel relationship between the equilibrium budget size and the inequality of the status quo allocation. In particular, holding constant a status quo budget size $\Sigma_i q_i > \Pi$, in equilibrium, the total budget $H(q)$ is a U-shaped function of $q_r/\Sigma_i q_i$, the allocation $h_a(q)$ for the agenda setter is weakly decreasing in $q_r/\Sigma_i q_i$, and the allocation $h_r(q)$ for the opposition party is weakly increasing in $q_r/\Sigma_i q_i$, where $q_r/\Sigma_i q_i$ measures the opposition's share of the budget in the status quo.

4 Three-player models

4.1 The political process

In this section we assume that the legislative body consists of three parties, labeled $i \in \{a, b, c\}$ and consider a legislative procedure in line with the agenda-setting model

of Romer and Rosenthal (1978). Specifically, party a as the agenda setter first makes a policy proposal, which is then voted on by majority rule.[7] If at least two parties vote to approve the proposal, it is implemented. Otherwise a status quo is implemented. In the voting stage we assume that each party does not play any weakly dominated strategy and always vote for the proposal if indifferent.

In any subgame perfect equilibrium, the agenda setter forms a voting coalition with one of the other parties and maximizes its utility by selecting a proposal from those policy alternatives that make its coalition partner at least break even. In order to avoid indifference in coalition formation we only consider in this section a status quo $q \in \mathbf{R}^3_+$ where the budget allocations for the two nonproposing parties are different in size. In other words the status quo is taken from $Q = \left\{(q_a, q_b, q_c) \in \mathbf{R}^3_+ : q_b \neq q_c\right\}$. We will analyze two institutional variants.

4.1.1 Parliamentary institutions

In the first institution the agenda setter has exogenously formed a coalition with party $r \in \{b, c\}$ and is constrained to seek voting support from its coalition partner. This case approximates the collective choice situation in a multi-party parliamentary system, where a government must continuously maintain the confidence of a majority in parliament to stay in office.[8] The agenda setter can be thought of as the prime minister's party, and the exogenously given coalition can be thought of as a coalition government formed in a parliament where no party commands a majority of seats. This coalition may have been formed based on various historical and political factors, e.g., ideologies and electoral concerns, which do not necessarily reflect the parties' preferences over the choice of government budget. Intuitively, one may think of a case where a coalition has formed after an election, but now must decide on fiscal policy adjustments due to exogenous changes in the world economy. Diermeier and Feddersen (1998) show that once a coalition is in office, the agenda setter typically has a strong incentive to make a policy proposal that would allow all government members to approve unanimously. This holds because in case some government member does not agree with a proposed bill, a government crisis may be triggered due to a vote of no confidence or the resignation of some ministers in the cabinet. In that contingency the incumbent prime minister may lose its agenda control, in case the cabinet is reshuffled or an early election is called for, and the status quo would remain.

4.1.2 Presidential institutions

In the second institution, upon making a policy proposal, the agenda setter is free to seek voting support from any of the other parties. This case may be representative

[7] The insights obtained in the two-player model can be directly carried over to here if a unanimity rule is applied, so in this section we focus on the model with a simple majority rule.

[8] Baron and Diermeier (2001) and Baron et al. (2008) use this bargaining institution in their models of a parliamentary democracy. They, however, are mainly interested in parliamentary elections under proportional representation and the process of government formation.

of a multiparty presidential system.[9] Carey and Shugart (1992) and Cheibub (2006), among others, argue that in presidential countries it is common for presidents to hold the sole agenda-setting power with respect to government budgets. Therefore, the agenda setter in our model can be understood as the president's party. In a multiparty presidential regime, e.g., Argentina and Brazil, the president's party typically does not control a majority of seats in the chamber. Therefore, the proposal of government budget involves the selection of a voting coalition. Importantly, the composition of the coalition may change over time. That is, coalition choice is endogenous.

4.2 Analysis

We begin the analysis with some quick observations. First, the party excluded from the voting coalition will be fully expropriated and allocated no budget, as its vote is not needed for the passage of a new policy. This holds in both majoritarian institutions. Second, in the presidential institution, the agenda setter will form a voting coalition with the nonproposing party that has a smaller status quo allocation. This is because the party disadvantaged by the status quo is "cheaper to buy" and makes the agenda setter's incentive-compatibility constraint less stringent. In what follows let r denote the nonproposing party included in the voting coalition and s denote the out party. In a parliamentary institution $r, s \in \{b, c\}$ are exogenously given, whereas in a presidential system $r = \arg\min_{i \in \{b,c\}} q_i$ and $s = \arg\max_{i \in \{b,c\}} q_i$.

Given any status quo $q \in Q$, the agenda setter (a) and its coalition partner (r) thus effectively play a two-player ultimatum game with an equivalent status quo $\widehat{q} = (\widehat{q}_a, \widehat{q}_r)$ such that $\widehat{q}_a = q_a + q_s$ and $\widehat{q}_r = q_r$. In equilibrium, the agenda setter would optimally select a policy proposal $f(q, r) \in \mathbf{R}^3_+$ such that, the budget size $F(q, r) = \Sigma_i f_i(q, r)$ and party r's budget allocation $f_r(q, r)$ jointly solve problem (1) with the equivalent status quo $\widehat{q}$, the out party receives no allocation, and the agenda setter's budget allocation $f_a(q, r)$ is residually determined.

As a consequence, the intuition built up for the two-player model applies here with some modifications. We say a policy x is *coalitionally efficient* if it maximizes the joint utility for parties in the voting coalition, i.e., mathematically $x_s = 0$ and $x_a + x_r = \Pi$. In the context of a three-player model, we refer to Π as the coalitionally efficient budget size. As presented in the next proposition, the changes in budget size are asymmetric in cases in which the status quo budget is below or above the coalitionally efficient size.

Proposition 2 *Consider the two institutions in a three-player agenda-setting model.*

1. *The party whose voting support is not needed receives no budgetary allocation in equilibrium. In the presidential institution this is the nonproposing party that has a higher status quo allocation. As a consequence, with the same status quo, the president in a presidential institution will be allocated no less than the prime minister in a parliamentary institution.*

[9] For example, Robinson and Torvick (2008) assume a similar legislative procedure in their comparative model of a presidential democracy.

2. *Suppose that the status quo budget is below the coalitionally efficient size. Then in equilibrium the budget increases and reaches the coalitionally efficient size; the allocations for both the agenda setter and its coalitional partner expand.*
3. *Suppose that the status quo budget size is coalitionally efficient. Then in equilibrium the budget size remains coalitionally efficient; the agenda setter's budget allocation expands due to expropriation but the allocation for its coalition partner stays the same.*
4. *Suppose that the status quo budget size is above the coalitionally efficient size. Then in equilibrium the budget shrinks but always ends up greater than or equal to the coalitionally efficient size; whereas the allocation for the nonproposing party in the voting coalition always gets cut, the agenda setter's budget allocation may not decrease as it could always expropriate the out party. The two institutions are compared in the following with a status quo budget that is above the coalitionally efficient size:*
   A. *In a parliamentary institution, whenever the equilibrium budget ends up strictly above the coalitionally efficient size, one of the parties in the government coalition takes the entire budget and it may or may not be the agenda setter. In other words, any government member may resist a welfare-enhancing fiscal contraction. The equilibrium budget reaches the coalitionally efficient size if and only if the nonproposing party in the government coalition has a moderate share of the budget in the status quo, in the sense that* $\Phi_1(q) \leq q_r / \Sigma_i q_i \leq \Phi_2(q)$, *where* $\Phi_1(\cdot)$ *and* $\Phi_2(\cdot)$ *are defined in* (2) *and* (3). *Holding constant a status quo budget that is above the coalitionally efficient size, the equilibrium budget size has a U-shaped relationship with* $q_r / \Sigma_i q_i$.
   B. *In a presidential institution, whenever the equilibrium budget ends up strictly above the coalitionally efficient size, the agenda setter, i.e., the president, takes the entire budget in equilibrium whereas all the other parties are left with no budgetary allocation. In other words, the nonproposing party in the voting coalition formed by the president never resists a welfare-enhancing fiscal contract but the president may do. The equilibrium budget reaches the coalitionally efficient size if and only if the share of the budget for disadvantaged nonproposing party in the status quo is sufficiently large, i.e.,* $\min\{q_b, q_c\} / \Sigma_i q_i \geq \Phi_1(q)$. *Whenever this happens, the president's share of the budget in the status quo is sufficiently small, i.e.,* $q_a / \Sigma_i q_i \leq 1 - 2\Phi_1(q)$. *Holding constant a status quo budget that is above the coalitionally efficient size, the equilibrium budget size is weakly increasing in* $\min\{q_b, q_c\} / \Sigma_i q_i$.

Again, this proposition illustrates a general point that the budget size in equilibrium depends on its initial size and allocation in a majoritarian institution. If the initial budget size is sufficiently large and the initial allocation are sufficiently unequal, the equilibrium budget size can be greater than the coalitionally efficient amount. On the other hand, if the initial budget size is sufficiently small, regardless of the initial allocation the equilibrium budget size always expands and reaches the coalitionally efficient amount. Such asymmetric movement of the budget size in a majoritarian institution is reminiscent of the pattern in the two-player ultimatum game that captures authoritar-

ian regimes and corporatist politics; the ratchet effect of public spending is present in all forms of political institutions analyzed in this article.

Our theory also allows us to compare institutions. Given the same status quo, it is more likely for the president in a presidential system to resist and, therefore, do not propose, a coalitionally efficient fiscal contraction than for a prime minister in a parliamentary system to behave in the same way. This is jointly implied by Parts 1 and 4 of Proposition 2. In a presidential system, the president always forms a voting coalition with the opposition party that has a smaller status quo budget allocation. At the same time, the president opposes a coalitionally efficient budget cut under the condition that $\min\{q_b, q_c\}/\Sigma_i q_i < \Phi_1(q)$. Naturally this condition is satisfied whenever the prime minister in a parliamentary system resists a coalitionally efficient budget cut, i.e., $q_r/\Sigma_i q_i < \Phi_1(q)$ for some exogenously given $r \in \{b, c\}$. On the other hand, the fact that the president in a presidential system is resisting a coalitionally efficient budget cut does not necessarily imply that the prime minister in a parliamentary system must be doing the same.

5 Concluding remarks

In this article we present a theory of government expenditure. Our analysis demonstrates a general point that the size of government budget and its composition are inter-related and should not be analyzed in isolation. Specifically, the total budget size in equilibrium depends on both size and composition of the status quo budget. Moreover, budgetary change occurs in an asymmetric manner. Government spending is easy to rise but difficult to fall.

In our model, excessively large government budget occurs in corner solutions where either the agenda setter or its coalition partner is at constraint with a zero budget allocation. Of course, this finding should not be interpreted literally. The agenda setter may have to provide some minimum allocation to the other parties or itself for some reasons not explicitly modeled here. For example, with a large senior population with little personal savings it may be difficult or constitutionally prohibited to reduce social security benefits to zero. In other words our model only captures that part of the budget that realistically and legally can be cut. Moreover, our model does not capture positive externalities of government programs across different parties. If the agenda setter can partially benefit from spending programs targeted at the other parties, it may not want to reduce those programs as much as possible. Also, quasilinearity may oversimplify the policy environment, although it provides tractability necessary for a clean and intuitive closed-form solution. Under the assumption of diminishing marginal returns for spending programs, for instance, it follows that if the expenditure level for some party is sufficiently small, the marginal utility received by that party from its allocation may be sufficiently high. In this case, a corner solution will never be reached, yet the qualitative results of our model would still be obtained, as an interior solution. This intuition implies that, whenever the government size is not fully downward elastic, the allocation of public resources on various parties would be highly unequal.

The fact that the cost-sharing arrangement is exogenously given is crucial in the present model. Indeed, the ratchet effect arises in corner solutions, as the agenda set-

ter cannot propose policies in which the level of public spending for some group is negative. If the agenda setter could propose a distribution of both public spending and costs, then the problem of corner solutions might disappear. Indeed, the agenda setter could then propose zero public spending for one group together with a larger share of the cost borne by that group. In this case the ratchet effect would no longer be present. However, the equal cost-sharing arrangements may be a reasonable assumption if the government is unable to discriminate various socioeconomic group through taxation based on their party affiliations or preferences over different government programs. The asymmetric movements of government budget and the associated inefficiency thus result from the interaction between restriction on policy instruments and the legislative procedure of agenda setting.

Our model can be extended to allow the possibility of counteroffers upon rejection of a proposed budget as in Rubinstein (1982) or Baron and Ferejohn (1989). Consider the two-player bilateral bargaining game as an example. Imagine that the two parties make policy proposals in turn until a proposal is unanimously agreed by both. Suppose that upon a proposal is rejected, with probability β the other party gets a chance to make a counteroffer whereas with probability $1 - \beta$ the game ends with the status quo being implemented. Here the parameter β may be interpreted as dispersion of the proposal power. In the extreme case with $\beta = 0$ as analyzed in this article, one party has monopoly of the proposal power, whereas in the other extreme case with β sufficiently approaching to 1, the proposal power is almost equally shared and the first agenda setter barely has a first-mover advantage (Rubinstein 1982). As β gets larger, proposal power is shared more equally, which leads to a more egalitarian distribution of utilities between the two parties in equilibrium. As a consequence it is less likely to hit a corner solution with a larger β, the probability of making a counteroffer upon rejection of a proposed bill. This intuition suggests that concentration, or unequal allocation, of the proposal power is also a critical factor that leads to the ratchet effect identified in this article.

Another possible extension is to incorporate the bargaining model into a dynamic setup, in which government budget is chosen repeatedly and some of the government projects, once enacted, persist until new projects are legislated to replace them. Diermeier and Fong (2008) have presented a two-period model in this direction with temporary shocks on the marginal costs of government spending, but a fully developed dynamic model awaits future research.

Acknowledgments This article is based on our conference paper for the Political Economy of Bargaining (PEB) Conference at Wilfrid Laurier University. We especially thank an anonymous referee for insightful comments and suggestions. For helpful discussion we also thank seminar participants in the PEB Conference, the North American Summer Meeting of the Econometric Society, the Wallis Institute's Political Economy Working Group, and the CEANA-TEA Joint Conference. Part of this research project was conducted when Fong visited the Wallis Institute of Political Economy, University of Rochester, and the Ford Motor Company Center for Global Citizenship, Northwestern University. Fong thanks these institutes for hospitality and financial aids. We also acknowledge financial support from the Canadian Institute for Advanced Research (CIFAR). All errors are our own.

Appendix

A Mathematical proofs

Proof of Proposition 1 Let

$$L \equiv \left(G - x_r - \frac{G^2}{4\Pi}\right) + \eta\left(x_r - q_r + \frac{(\Sigma_i q_i)^2 - G^2}{4\Pi}\right) + \lambda_a (G - x_r) + \lambda_r x_r,$$

where η, λ_a and λ_r are the Lagrangian multipliers for the corresponding inequality constraints. By the Kuhn–Tucker Theorem, the first-order necessary conditions are:

$$\frac{\partial L}{\partial G} = 1 + \lambda_a - \frac{(1+\eta)\, H(q)}{2\Pi} = 0, \tag{4}$$

$$\frac{\partial L}{\partial x_r} = \lambda_r + \eta - (1 + \lambda_a) = 0, \tag{5}$$

$$\lambda_a \geq 0, \quad h_a(q) \geq 0, \quad \lambda_a h_a(q) = 0,$$

$$\lambda_r \geq 0, \quad h_r(q) \geq 0, \quad \lambda_r h_r(q) = 0,$$

$$\eta \geq 0, \quad h_r(q) \geq q_r + \frac{[H(q)]^2 - (\Sigma_i q_i)^2}{4\Pi}, \quad \text{and}$$

$$\eta\left(h_r(q) - q_r - \frac{[H(q)]^2 - (\Sigma_i q_i)^2}{4\Pi}\right) = 0.$$

Notice that the sufficient conditions obviously hold. We claim two useful properties of the Lagrangian multipliers. First, $\lambda_r + \eta > 0$. This is given by (5) together with the requirement that $\lambda_a \geq 0$. In words, either the incentive-compatibility constraint or the nonnegative constraint on the allocation for party r must be binding. Second, $\lambda_a \lambda_r = 0$. In order to understand this, suppose to the contrary that $\lambda_a > 0$ and $\lambda_r > 0$. This immediately implies $H(q) = h_r(q) = 0$. However, (4) implies $H(q) > 0$, which is a contradiction. Due to these two properties, in what follows it suffices to discuss four cases based on the values of the Lagrangian multipliers.

Case 1 Suppose that $\lambda_a = \lambda_r = 0$ and $\eta > 0$. This immediately implies that $h_r(q) > 0$. As a consequence, the incentive-compatibility constraint must be binding, so $h_r(q) = q_r + \left([H(q)]^2 - (\Sigma_i q_i)^2\right)/4\Pi$. By (5), $\eta = 1$. By (4), $H(q) = \Pi$. Putting all these together,

$$h_r(q) = q_r + \frac{1}{4}\left(1 - \frac{\Sigma_i q_i}{\Pi}\right)(\Pi + \Sigma_i q_i),$$

and

$$h_a(q) = \Pi - q_r - \frac{1}{4}\left(1 - \frac{\Sigma_i q_i}{\Pi}\right)(\Pi + \Sigma_i q_i)$$
$$= q_a + \frac{1}{4}\left(1 - \frac{\Sigma_i q_i}{\Pi}\right)(3\Pi - \Sigma_i q_i).$$

Since $x_r \geq 0$,

$$\frac{q_r}{\Sigma_i q_i} \geq \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) - \frac{1}{4}\left(\frac{\Pi}{\Sigma_i q_i}\right).$$

This right-hand side of the inequality above is smaller than or equal to 0 if $\Sigma_i q_i \leq \Pi$ whereas the left-hand side must be nonnegative. So this inequality holds whenever $\Sigma_i q_i \leq \Pi$. Also, since $x_a \geq 0$,

$$\frac{q_r}{\Sigma_i q_i} \leq \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) + \frac{3}{4}\left(\frac{\Pi}{\Sigma_i q_i}\right).$$

The right-hand side of the inequality above can be decomposed into $\frac{1}{4}\left[\left(\frac{\Sigma_i q_i}{\Pi}\right) + \left(\frac{\Pi}{\Sigma_i q_i}\right)\right] + \frac{1}{2}\left(\frac{\Pi}{\Sigma_i q_i}\right)$, where the minimum of the first term is $\frac{1}{2}$, attained at $\Sigma_i q_i = \Pi$, and the second term is greater than $\frac{1}{2}$ if $\Sigma_i q_i < \Pi$. As the left-hand side must be smaller than or equal to 1, this inequality holds as whenever $\Sigma_i q_i \leq \Pi$. Overall this case corresponds to Parts 1, 2, and 3-B in the proposition.

Case 2 Suppose that $\lambda_a = 0$, $\lambda_r > 0$ and $\eta = 0$. This immediately implies $h_r(q) = 0$. Since the incentive-compatibility constraint must be satisfied, i.e., $h_r(q) = 0 \geq q_r + \left([H(q)]^2 - (\Sigma_i q_i)^2\right)/4\Pi$, it follows that

$$\frac{q_r}{\Sigma_i q_i} \leq \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) - \left(\frac{\Pi}{\Sigma_i q_i}\right).$$

By (5), $\lambda_r = 1$. As a consequence, $H(q) = 2\Pi$ by (4).

Case 3 Suppose that $\lambda_a = 0$, $\lambda_r > 0$ and $\eta > 0$. This immediately implies that

$$h_r(q) = q_r + \frac{[H(q)]^2 - (\Sigma_i q_i)^2}{4\Pi} = 0.$$

Solving the second equality above,

$$H(q) = 2\,(\Sigma_i q_i)\left[\frac{1}{4} - \left(\frac{\Pi}{\Sigma_i q_i}\right)\left(\frac{q_r}{\Sigma_i q_i}\right)\right]^{\frac{1}{2}}.$$

By (5), $\eta + \lambda_r = 1$. This means $0 < \eta < 1$. By (4), we thus have $\Pi < H(q) = 2\Pi/(1+\eta) < 2\Pi$. Since $H(q) > \Pi$,

$$\frac{q_r}{\Sigma_i q_i} < \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) - \frac{1}{4}\left(\frac{\Pi}{\Sigma_i q_i}\right).$$

This inequality always fails if $\Sigma_i q_i \leq \Pi$, as in that case the right-hand side is smaller than or equal to 0 whereas $q_r / \Sigma_i q_i$ must be nonnegative. Since $H(q) < 2\Pi$,

$$\left(\frac{q_r}{\Sigma_i q_i}\right) > \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) - \left(\frac{\Pi}{\Sigma_i q_i}\right).$$

Overall, this case and Case 2 jointly correspond to Part 3-A in the proposition.

Case 4 Suppose that $\lambda_a > 0$, $\lambda_r = 0$ and $\eta > 0$. This immediately implies that $h_a(q) = H(q) - h_r(q) = 0$ and

$$H(q) = h_r(q) = q_r + \left(\frac{[H(q)]^2 - (\Sigma_i q_i)^2}{4\Pi}\right).$$

Solving the quadratic equation for $H(q)$, we have

$$H(q) = h_r(q) = 2\Pi - 2\left(\Sigma_i q_i\right)\left[\left(\frac{\Pi}{\Sigma_i q_i}\right)^2 + \frac{1}{4} - \left(\frac{q_r}{\Sigma_i q_i}\right)\left(\frac{\Pi}{\Sigma_i q_i}\right)\right]^{\frac{1}{2}}.$$

It is straightforward to verify that $H(q) \leq 2\Pi$ and the equality holds if $q_r = \Sigma_i q_i = 2\Pi$. For the square root in the expression above to be well defined, we must have

$$\frac{q_r}{\Sigma_i q_i} \leq \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) + \left(\frac{\Pi}{\Sigma_i q_i}\right).$$

Notice that this inequality unconditionally holds. This is because the minimum of the right-hand side is 1, attained at $\Sigma_i q_i = 2\Pi$, whereas the left-hand side must be no greater than 1. By (5), $\eta = 1 + \lambda_a$. By (4), $H(q) = 2\left(\frac{1+\lambda_a}{2+\lambda_a}\right)\Pi$. Since $\lambda_a > 0$, we have $H(q) > \Pi$ and consequently

$$\frac{q_r}{\Sigma_i q_i} > \frac{1}{4}\left(\frac{\Sigma_i q_i}{\Pi}\right) + \frac{3}{4}\left(\frac{\Pi}{\Sigma_i q_i}\right).$$

The last inequality always fails if $\Sigma_i q_i \leq \Pi$. Due to the discussion in Case 1, the right-hand side is greater than or equal to 1 provided $\Sigma_i q_i \leq \Pi$, which contradicts the fact that $q_r / \Sigma_i q_i \leq 1$. Overall this case corresponds to part 3-C in the proposition.

□

Proof of Proposition 2 This proposition directly follows the explanation in Sect. 4.2 except for the last part. Part 4-B is Implied by Proposition 1. If $\min\{q_b, q_c\} / \Sigma_i q_i > \Phi_2(q)$, then in equilibrium the nonproposing party in the voting coalition would take the entire budget whereas the president would get nothing. However, this would never happen since $\min\{q_b, q_c\} / \Sigma_i q_i \leq \frac{1}{2}$ and $\min_{q \in Q} \Phi_2(q) \geq \frac{1}{2}$. □

B A model with reconsideration

B1 The political process

In this section we extend the three-player model with endogenous coalitional choice to incorporate the possibility of reconsideration by a persistent agenda setter. As before such a model can be interpreted as a presidential system. But, in this extension the president is allowed to make new proposals to replace any previously approved bill.

We apply the analytical framework of Diermeier and Fong (2011) and assume that any allocation must be a multiple of some minimal spending unit $\epsilon > 0$, which can be as small as, for example, one cent. The discrete policy space is then represented by

$$\mathbf{R}^3_{\epsilon,+} = \left\{ x \in \mathbf{R}^3_+ : x = \epsilon z \text{ for some } z \in \mathbf{Z}^3_+ \right\}.$$

For technical convenience, we assume that ϵ is extremely small relative to Π and moreover $\frac{2}{3}\Pi$ is a multiple of ϵ, where $\frac{2}{3}\Pi$ is the efficient total spending level in a model with three players.

The legislature selects a policy over the course of potentially infinite rounds of proposal making and voting, where the number of rounds depends on exogenous factors and the decision made by the agenda setter. As the legislative session commences, a status quo policy $q \in \mathbf{R}^3_{\epsilon,+}$ is exogenously given. The status quo, for example, can be the policy that has been enacted prior to this legislative session. We define a "default" as the policy that would be implemented at the end of the legislative session if no new law is made in the rest of the session. Intuitively, the default in the first proposal round is simply the status quo; i.e., $x^0 = q$. Since then, activities prior to round $t > 1$ establish a default $x^{t-1} \in \mathbf{R}^3_{\epsilon,+}$. In any proposal round t, the agenda setter could either "pass" or make a policy proposal $y^t \in \mathbf{R}^3_{\epsilon,+}$. A "pass" in this context means inaction by the agenda setter and, for mathematical convenience, is modeled as a proposal $y^t = x^{t-1}$. The proposal y^t is then put to an immediate vote against the default x^{t-1} by simple majority rule. If the proposal y^t is approved, it replaces x^{t-1} and the new default is $x^t = y^t$. Otherwise the prevailing default remains and $x^t = x^{t-1}$. The policy that survives as default till the end of the legislative session is implemented.

The last proposal round is not predetermined so we have to specify how the legislative session may end. We say the legislative session ends *endogenously* after proposal round t, if the default x^t established by the first t rounds of proposal making and voting is such that the agenda setter will choose to pass any possible proposal round $t' > t$. In addition, after any proposal round the legislative session may be terminated *exogenously* with probability $1 - \delta$, where $\delta \in [0, 1)$ is the probability that the agenda setter will have an opportunity to reconsider the policy that emerges from the current round.

Here we assume that δ is sufficiently close to 1 given any policy space $\mathbf{R}^3_{\epsilon,+}$. This approximates the case in which policy-making in the legislative session proceeds until the agenda setter no longer wants to make a new proposal to replace a previously approved policy, i.e., the prevailing default.

B2 Equilibrium definition

Without loss of generality, we assume that party a is the agenda setter. As is customary in the legislative bargaining literature, for the model with the possibility of reconsideration we focus the analysis on stationary Markov perfect equilibria, in which the parties condition their strategies only on the prevailing default policy. We thus drop the superscript t for the proposal round from the notations.

Let $f(x)$ be the policy that the agenda setter proposes when the prevailing default is x. We refer to $f : \mathbf{R}^3_{\epsilon,+} \to \mathbf{R}^3_{\epsilon,+}$ as the policy rule. Let $U_\ell(x)$ be the expected utility of player ℓ if policy x is approved. With probability $1-\delta$ the legislative session is exogenously terminated after the current proposal round and this player receives a utility of $u_\ell(x)$. With probability δ the agenda setter has a chance to reconsider the approved policy and make a new proposal according to f. In this case, player ℓ receives a continuation utility of $U_\ell\left(f(x)\right)$. Thus,

$$U_\ell(x) = (1-\delta)\, u_\ell(x) + \delta U_\ell\left(f(x)\right). \tag{6}$$

We refer to $U_\ell : \mathbf{R}^3_{\epsilon,+} \to \mathbf{R}$ as the value function of player ℓ.

We make two technical assumptions regarding how the players break indifference. First, any party votes against a policy proposal if and only if passage of the proposal makes it strictly worse off. Second, the agenda setter never proposes any shift in policy that is destined to be vetoed by a majority. None of our qualitative results depend on the first assumption. The second one simplifies the notation, but is otherwise innocuous, since making a losing proposal is equivalent to remaining at the prevailing default.

In each proposal round, the agenda setter selects a policy proposal to maximize its expected utility. By the second assumption above, the agenda setter must be choosing from the set of policy alternatives which at least one player would vote for. By the first assumption above, this is equivalent to an incentive-compatibility constraint that requires that at least one player be weakly better off with the proposal than with the default. To sum up, given any default $x \in \mathbf{R}^3_{\epsilon,+}$, $f(x)$ must solve

$$\begin{aligned} \max_{y \in \mathbf{R}^3_{\epsilon,+}} \quad & U_a\left(y\right) \\ \text{s.t. } & U_i\left(y\right) \geq U_i(x) \quad \text{for some } i \in \{b, c\}\,. \end{aligned} \tag{7}$$

We are now ready to define the equilibrium.

Definition 1 A stationary Markov perfect equilibrium is a policy rule f such that $f(x)$ solves (7) for all $x \in \mathbf{R}^3_{\epsilon,+}$, where (U_a, U_b, U_c) solve the equation system defined by (6) given f.

In this article we characterize the class of equilibria where no reconsideration occurs, i.e., any equilibrium f such that $f\left(f(x)\right) = f(x)$ for all $x \in \mathbf{R}^3_{\epsilon,+}$. Diermeier and Fong (2011) prove that restricting the attention to equilibria where no reconsideration occurs does not reduce the set of possible policy outcomes, assuming that the legislative session may last for sufficiently many rounds. As our focus is on the policy

outcome and its relation to the status quo, we miss nothing by ignoring equilibria where reconsideration may actually occur.

B3 An equivalent problem

The possibility of reconsideration changes the nature of policy-making. This is true even if the agenda setter optimally decides not to revisit an approved policy in equilibrium. Our strategy to characterize an equilibrium is to answer the following question: If the agenda setter was restricted to making a policy proposal once and for all, i.e., reconsideration was not allowed, what additional constraints should be imposed so that the agenda setter's optimal proposal strategy in this hypothetical one-round game is identical to the equilibrium policy rule in the original game with the possibility of reconsideration? Our search for these additional constraints relies on a couple of necessary conditions summarized in Lemmas 1 and 2.

Lemma 1 *In any equilibrium f, $f_b(x) = f_c(x)$ for all $x \in \mathbf{R}^3_{\epsilon,+}$.*

Lemma 1 shows that government spending on both groups with no proposal power must be identical in any equilibrium. The proof simply applies the algorithm proposed by Diermeier and Fong (2011) and is omitted here. Intuitively, group b will not approve any proposal that offers group c strictly less than what it offers to b, say y_b. If, counterfactually, group b were to approve such a proposal, the agenda setter would then have an incentive to reconsider the policy and choose c as a cheaper coalition partner to pass a new proposal that further expropriates b, so that group b would end up with a spending level smaller than y_b. Therefore, by voting to protect c, group b can effectively safeguard its bargaining position in case the policy will be reconsidered. In equilibrium, the two groups with no proposal power are induced to "defend" the benefits for each other against further exploitation by the agenda setter. As a consequence, proposal power is endogenously limited and a more egalitarian allocation is reached.

Lemma 2 *In any equilibrium f and for all $x \in \mathbf{R}^3_{\epsilon,+}$, (A) $u_a(f(x)) \geq u_a(x)$, and (B) $u_b(f(x)) = u_c(f(x)) \geq \min\{u_b(x), u_c(x)\}$.*

Lemma 2 says that in equilibrium the policy outcome must deliver a weakly higher utility than the default policy would do to the agenda setter and at least one opposition party. Given our focus on equilibria where no reconsideration actually occurs, any equilibrium policy must persist as default. Therefore, any policy that makes the agenda setter strictly worse off than with default cannot be an equilibrium policy, as the agenda setter can always choose to keep the default. The same applies to any policy that makes both groups b and c strictly worse than with the default, as the two groups can jointly veto it. This lemma is a special case of the results in Diermeier and Fong (2011) so the proof is omitted here.

With Lemmas 1 and 2, we are ready to establish the equivalence between an equilibrium policy rule for the original game, where reconsideration is allowed, and the solution to a simple game where reconsideration is not allowed but additional constraints are imposed.

Proposition 3 *A policy rule f constitutes an equilibrium if and only if there exists an equivalent reservation value function $r : \mathbf{R}^3_{\epsilon,+} \to \mathbf{R}$ such that, for all $x \in \mathbf{R}^3_{\epsilon,+}$, (A) $f(x)$ solves*

$$\begin{aligned} \max_{y \in \mathbf{R}^3_{\epsilon,+}} \quad & u_a(y) \\ s.t. \quad & y_b = y_c, \\ & u_b(y) \geq r(x), \end{aligned} \tag{8}$$

(B) $u_a(f(x)) \geq u_a(x)$, and (C) $r(x) \geq \underline{r}(x) \equiv \min\{u_b(x), u_c(x)\}$.

The agenda setter faces three constraints in maximization problem (8). First, proposed spending on each group must be nonnegative, i.e., $y \in \mathbf{R}^3_{\epsilon,+}$, as is required in the original problem. Second, the spending levels for groups b and c must be the same, i.e., $y_b = y_c$, which is necessary in any equilibrium according to Lemma 1. The last constraint is more involved: the utility delivered by any proposed policy to group b (and c) must be greater than or equal to some cutoff $r(x)$, which is treated as exogenous in problem (8). We refer to $r(x)$ as an equivalent reservation value.[10]

Proposition 3 simply says that, with a suitably chosen function r, the policy rule that solves problem (8) constitutes an equilibrium for the original game with the possibility of reconsideration. Moreover, any equilibrium policy rule for the original game solves problem (8) for some suitably chosen function r.

An *admissible* function r that supports the equivalence of the two problems is bounded by Conditions (B) and (C) in Proposition 3. These conditions require that any proposed policy offers a weakly higher utility than the default policy to the agenda setter and at least one opposition party. Such requirement is consistent with the necessary equilibrium conditions in Lemma 2. Since $u_a(f(x))$ is an implicit, decreasing function of $r(x)$, a lower bound on $u_a(f(x))$ required by Condition (B) is equivalent to an upper bound on $r(x)$. We denote this upper bound by $\overline{r}(x)$.[11]

Proposition 3 implies not only the existence of an equilibrium, but also the possibility of multiple equilibria if $\overline{r}(x)$ is sufficiently greater than $\underline{r}(x)$ for some $x \in \mathbf{R}^3_{\epsilon,+}$.[12] Multiplicity arises due to self-fulfilling expectations of the players. In every proposal

[10] Recall that in the original game with the possibility of reconsideration, the reservation value for group i in any proposal round with default x is given by $U_i(x) = (1-\delta)\, u_i(x) + \delta U_i\left(f^e(x)\right)$, which depends on not only the current default policy x but also the players' expectation on the proposal strategy f^e to be played in the future if the agenda setter is given a chance to revisit the policy. With $\delta < 1$ sufficiently large, the expected future proposal strategy plays a disproportionally important role in the calculation of a reservation value. The search for a valid equivalent reservation value for problem (8) thus is a nontrivial task. Therefore, the equivalent reservation value for group b in problem (8) is not necessarily as simple as $\min\{u_b(x), u_c(x)\}$, which some readers may guess from Lemma 2.

[11] For all $x \in \mathbf{R}^3_{\epsilon,+}$, let $w(x, \rho)$ be the maximum of problem (8) associated with an equivalent reservation value ρ. Formally, the upper bound of a valid equivalent reservation value is defined as

$$\overline{r}(x) \equiv \max\{\rho \in \mathbf{R} : w(x, \rho) \geq u_a(x)\}. \tag{9}$$

It can be shown that $\overline{r}(x) \geq \underline{r}(x)$ for all $x \in \mathbf{R}^3_{\epsilon,+}$.

[12] For all $x \in \mathbf{R}^3_{\epsilon,+}$, $\overline{r}(x) \geq \underline{r}(x)$ and a solution to problem (8) obviously exists for any admissible equivalent reservation value. By equivalence, the existence of an equilibrium is ensured.

round all players anticipate the equilibrium policy rule to be played in all subsequent proposal rounds, and based on this common expectation they calculate their reservation values that determine the current policy rule. Stationarity requires that the anticipated future policy rule be consistent with the current equilibrium policy rule. Therefore, the players' expectation shapes their current play. Multiple equilibria thus result from the coexistence of multiple pairs of expectations and policy rule that are mutually consistent.

B4 Separation of choices over size and allocation

Given any status quo $q \in \mathbf{R}^3_{\epsilon,+}$ and any equivalent reservation value $r(q) \in \left[\underline{r}(q), \overline{r}(q)\right]$, the agenda setter's decision can be separated into two parts.

First, once the total spending level G is given, the agenda setter must offer players b and c the same minimal spending level that respects the constraints of incentive compatibility and nonnegativity. The minimal spending level that satisfies the incentive-compatibility constraint is given by

$$g_\epsilon(G, q; r) = r(q) + \tfrac{1}{6}\kappa G^2 + \sigma(G, q; r),$$

where $\sigma(G, q; r) \geq 0$ is a residual that rounds the spending level to a multiple of the minimal spending unit ϵ.[13] Since any spending level is required to be nonnegative, the agenda setter then must allocate

$$g^*_\epsilon(G, q; r) = \max\{0, g_\epsilon(G, q; r)\}$$

to both parties b and c, and leave $G - 2g^*_\epsilon(G, q; r)$ for itself.

Second, the agenda setter's problem then is reduced to the choice of the total spending level that maximizes its own utility. In particular, the agenda setter solves

$$\begin{aligned} \max_{G \in \mathbf{R}_\epsilon} \quad & G - 2g^*_\epsilon(G, q; r) - \tfrac{1}{3}\left(\tfrac{G^2}{2\Pi}\right) \\ \text{s.t.} \quad & G - 2g^*_\epsilon(G, q; r) \geq 0, \end{aligned}$$

where the inequality constraint requires that budget allocation for itself be nonnegative. Therefore, any equilibrium policy rule f associated with r must be such that $F(q) \equiv f_a(q) + f_b(q) + f_c(q)$ solves the above maximization problem and $f_b(q) = f_c(q) = g^*_\epsilon(G, q; r)$.

B5 Approximation

Given the assumed discrete policy space $\mathbf{R}^3_{\varepsilon,+}$, a search for an equilibrium requires pairwise comparison of utilities derived from different policy alternatives. A precise

[13] Formally, $\sigma(G, q; r) \equiv \left\lceil \left(r(q) + \frac{1}{6}\kappa G^2\right)/\epsilon \right\rceil \epsilon - \left(r(q) + \frac{1}{6}\kappa G^2\right)$, where $\lceil\cdot\rceil$ denotes the smallest integer that is greater than or equal to its argument. By definition, $0 \leq \sigma(G, q; r) < \epsilon$. This residual is negligible if the minimal spending unit is sufficiently small, i.e., $\lim_{\epsilon \to 0} \sigma(G, q; r) = 0$.

and complete description of an equilibrium thus is tedious and may blur the critical trade-offs of the policy choice.

In the rest of this section we present a "continuous" approximation of an equilibrium. The idea is to see what an equilibrium would approximate if the policy grid is sufficiently fine, i.e., $\epsilon > 0$ is sufficiently small.

Definition 2 A policy rule, $f : \mathbf{R}^3_{\epsilon,+} \to \mathbf{R}^3_+$, is an approximate equilibrium if for any $x \in \mathbf{R}^3_{\epsilon,+}$, there exists $r(x)$ such that: (A) $r(x) \geq \underline{r}(x)$; (B) $u_a(f(x)) \geq u_a(x)$; (C) $f_b(x) = f_c(x) = g^*(F(x), x; r)$, where

$$g^*(G, x; r) \equiv \max\left\{0, r(x) + \tfrac{1}{3}\left(\tfrac{G^2}{2\Pi}\right)\right\}$$

for any $G \geq 0$; and (D) $F(x)$ solves

$$\begin{aligned} \max_{G \in \mathbf{R}} \; & G - 2g^*(G, x; r) - \tfrac{1}{3}\left(\tfrac{G^2}{2\Pi}\right) \\ \text{s.t.} \; & G - 2g^*(G, x; r) \geq 0. \end{aligned}$$

B6 Results

We are now back to the theme of this article: How is the equilibrium budget size related to the initial size and allocation of the budget? To answer this question we characterize the approximate equilibrium associate with an arbitrary equivalent reservation value function $r \in [\underline{r}, \overline{r}]$.

Proposition 4 *Consider any approximate equilibrium f associated with some $r \in [\underline{r}, \overline{r}]$ and take any status quo $q \in \mathbf{R}^3_{\epsilon,+}$.*

1. *If the status quo budget size is either efficient or inefficiently too small, i.e., such that $q_a + q_b + q_c \leq \frac{2}{3}\Pi$, then the equilibrium budget (weakly) expands to the efficient size.*
2. *If the status quo budget size is inefficiently too large, i.e., such that $q_a + q_b + q_c > \frac{2}{3}\Pi$, then the equilibrium budget size may end up greater than the efficient level. Moreover, whenever the equilibrium budget size is strictly greater than the efficient level, either the agenda setter or the two opposition parties receive no budget allocation. Everything else equal, this happens only if the equivalent reservation value $r(q)$ is either sufficiently large or sufficiently small.*

The proof of this proposition resembles that of Proposition 1 and thus is omitted. Part 1 does not depend on the equivalent reservation value $r(q)$, so it holds for *any* approximate equilibrium. Therefore, in any approximate equilibrium if the status quo budget is associated with under-provision of government programs, then the total government spending would easily expand and reach the efficient level. The intuition is the same. Whenever it is social welfare enhancing to tax more, the agenda setter raises distortionary taxes, allocates part of the extra tax revenues to the opposition parties just enough to obtain a majority support, and leaves the rest of additional resources to itself.

Notice that the possibility of reconsideration plays a role to enhance policy efficiency whenever the status quo budget size is sufficiently small. This is because with the possibility of reconsideration, the agenda setter is endogenously constrained to propose an equal budget allocation to the opposition parties. This induces the agenda setter to internalize all costs of a budget expansion and, therefore, eliminates overspending.

Part 2 of the proposition says that, if the status quo budget is associated with overprovision of government programs, the total government spending may not drop to the efficient level due to corner solutions. For example, if the opposition parties are both at their constraint so that any fiscal reduction has to rely on the reduction of spending on the agenda setter, the agenda setter may find it too costly to reduce spending since it only internalizes part of the benefits from a welfare-enhancing budget cut.

Part 2 also points out that the resistance to downward adjustments is stronger only if the equivalent reservation value is sufficiently extreme, i.e., either sufficiently large or sufficiently small. At the same time the equivalent reservation value would be extreme only if the status quo allocation is sufficiently unequal. This implies that the persistence of an inefficiently large budget size must come with a sufficiently unequal allocation of the status quo. The reverse, however, is not true. With the possibility of reconsideration, a sufficiently unequal allocation of the status quo does not necessarily imply the persistence of inefficient overspending.

References

Aghion P, Bolton P (1990) Government domestic debt and the risk of a default: a political-economic model of a strategic role of debt. In: Dornbusch R, Draghi M (eds) Public debt management: theory and history. Cambridge University Press, New York

Baron DP (1996) A dynamic theory of collective goods programs. Am Polit Sci Rev 90:316–330

Baron DP, Diermeier D (2001) Elections, governments, and parliaments in proportional representation systems. Q J Econ 116:933–967

Baron DP, Ferejohn J (1989) Bargaining in legislatures. Am Polit Sci Rev 83:1181–1206

Baron DP, Diermeier D, Fong P (2008) A Dynamic theory of a parliamentary democracy. CMS-EMS Discussion Paper No. 1464

Battaglini M, Coate S (2007) Inefficiency in legislative policy-making: a dynamic analysis. Am Econ Rev 97:118–149

Battaglini M, Coate S (2008) A dynamic theory of public spending, taxation and debt. Am Econ Rev 98:201–236

Baumol WJ (1967) Macroeconomics of unbalanced growth: the anatomy of the urban crisis. Am Econ Rev 57:415–426

Bernheim D, Rangel A, Rayo L (2006) The power of the last word in legislative policy making. Econometrica 74:1161–1190

Borcherding T (1977) Budgets and bureaucrats: the sources of government growth. Duke University Press, Durham

Carey JM, Shugart MS (1992) Presidents and assemblies: constitutional design and electoral dynamics. Cambridge University Press, Cambridge

Chari VV, Jones LE, Marimon R (1997) The economics of split-ticket voting in representative democracies. Am Econ Rev 87:957–976

Cheibub JA (2006) Presidentialism, parliamentarism, and democracy. Cambridge University Press, Cambridge

Diermeier D, Feddersen T (1998) Cohesion in legislatures and the vote of confidence procedure. Am Polit Sci Rev 92:611–621

Diermeier D, Fong P (2008) Policy persistence in a multi-party parliamentary democracy. In: Elhanan H (ed) Institutions and economic performance. Harvard University Press, Cambridge
Diermeier D, Fong P (2011) Legislative bargaining with reconsideration. Q J Econ, Forthcoming
Duggan J, Kalandrakis T (2010) Dynamic legislative policy making. Mimeo
Hassler J, Mora JVR, Storesletten K, Zilibottiet F (2003) The survival of the welfare state. Am Econ Rev 93:87–112
Hassler J, Krusell P, Storesletten K, Zilibottiet F (2005) The dynamics of government. J Monet Econ 52:1331–1358
Higgs R (1987) Crisis and leviathan: critical episodes in the growth of American government. Oxford University Press, New York
Kalandrakis T (2004) A three-player dynamic majoritarian bargaining game. J Econ Theory 16:294–322
Katzenstein PJ (1985) Small states in world markets: industrial policy in Europe. Cornell University Press, Ithaca
Meltzer AH, Richard SF (1981) A rational theory of the size of government. J Polit Econ 89:914–927
Milesi-Ferretti GM, Perotti R, Rostagno M (2002) Electoral systems and public spending. Q J Econ 117:609–657
Morelli M (1999) Demand competition and policy compromise in legislative bargaining. Am Polit Sci Rev 93:809–820
Persson T, Svensson L (1989) Why a stubborn conservative would run a deficit: policy with time-inconsistency preferences. Q J Econ 104:325–345
Persson T, Tabellini G (2000) Political economics: explaining economic policy. MIT Press, Cambridge
Persson T, Tabellini G (2003) The economic effects of constitutions. MIT Press, Cambridge
Persson T, Tabellini G (2004) Constitutional rules and fiscal policy outcomes. Am Econ Rev 94:25–45
Persson T, Roland G, Tabellini G (1998) Towards micropolitical foundations of public finance. Eur Econ Rev 42:685–694
Persson T, Roland G, Tabellini G (2000) Comparative politics and public finance. J Polit Econ 108: 1121–1161
Robinson JA, Torvick R (2008) Endogenous presidentialism. NBER Working Paper No. 14603
Romer T, Rosenthal H (1978) Political resource allocation, controlled agendas and the status quo. Public Choice 33:27–43
Rubinstein A (1982) Perfect equilibrium in a bargaining model. Econometrica 50:97–109
Schmitter PC, Lehmbruch G (1979) Trends toward coporatist intermediation. Sage Publications, Beverly Hills
Tabellini G (1987) Money, debt and deficits in a dynamic game. J Econ Dyn Control 10:427–442
Tabellini G, Alesina A (1990) Voting on the budget deficit. Am Econ Rev 80:37–39
Volden C, Wiseman A (2007) Bargaining in legislatures over particularistic and collective goods. Am Polit Sci Rev 101:79–92
Weingast BR, Shepsle KA, Johnsen C (1981) The political economy of benefits and costs: a neoclassical approach to distributive politics. J Polit Econ 89:642–664
Xu C (2010) The fundamental institutions of China's reforms and development. J Econ Lit, Forthcoming

Soc Choice Welf
DOI 10.1007/s00355-010-0511-4

Intergovernmental negotiation, willingness to compromise, and voter preference reversals

Maria E. Gallego · David Scoones

Abstract A policy is the outcome of negotiations between two three-party parliamentary states. An election in jurisdiction A determines the composition of the legislature that selects a representative to negotiate an intergovernmental policy agreement with the representative from the legislature of jurisdiction B. Negotiations are modeled using Nash (Econometrica 18(2):155–162, 1950) bargaining framework. With heterogeneous parties, agreements and electoral outcomes depend on the concavity of the utility functions of negotiators and on the relative location of their ideal policies, i.e., depend on the negotiators relative willingness to compromise. Agreements between the bargainers may not follow the ordering of the parties' ideal policies. An electoral outcome where support for the center party comes from extreme voters may emerge.

1 Introduction

In an interdependent world, many policies are the outcome of negotiations[1] between different centers of power at national and international levels (see e.g., Breton 1996; Whalley 2008). In intergovernmental negotiations citizens are usually represented by

[1] Negotiated policies include situations of economic interdependence between jurisdictions (e.g., trans-border flows of goods and services) or areas where negative externalities arise when markets cannot be separated (e.g., pollution, trade barriers, taxation on mobile factors of production). As (Moravcsik, 1993, p. 485) argues "national governments have an incentive to co-operate where policy coordination increases their control over domestic policy outcomes, permitting them to achieve goals that would otherwise not be possible."

M. E. Gallego (✉)
Department of Economics, Wilfrid Laurier University, Waterloo, ON, N2L 3C5, Canada
e-mail: mgallego@wlu.ca

D. Scoones
Department of Economics, University of Victoria, Victoria, BC, V8W 2Y2, Canada
e-mail: scoones@uvic.ca

Published online: 21 December 2010

their jurisdiction's ruling party or its delegate. Elections or other political events that change the ruling party or governing coalition may change the government's policy preferences and may affect the outcome of intergovernmental bargaining (Moravcsik 1991). With heterogeneous parties and voters' preferences differing across jurisdictions, voters in one jurisdiction do not expect the ruling party of the other jurisdiction to represent their interests. Voters understand the interdependence of policy formation. Studying 14 multi-party parliamentary democracies, Kedar (2005) finds that citizens take into account the post-electoral legislative bargaining by voting for parties with positions different from their own. Rather than vote for the party whose most preferred policy closely matches their own, policy-oriented voters should choose a government that will negotiate the best possible policy.

We develop a model of elections where policy outcomes are the result of intergovernmental bargaining. We find that voters choices depend both on political parties' ideal policies—those they would implement were they to have total control—and their willingness to compromise. We show that in certain circumstances some voters will support a party whose ideal policy is more distant from their own than is that of an alternative party. When parties differ sufficiently in their willingness to compromise, groups of voters may swap allegiance: centrist voters support an "extreme" party, and extreme voters support the centre party.

Kedar (2005) finds that moderate voters in Norway and the Netherlands vote for a party whose position is more extreme than their own. Moravcsik (1991) argues that in 1983 Mitterrand realigned himself with the moderate wing of the Socialist Party as support for the Communist party declined allowing him to show greater public willingness to compromise on European integration. Moravcsik and Nicolaïdis (1999) argue that the hesitancy of European Union officials to publicize new policy initiatives during the negotiation of the Amsterdam Treaty in 1997 for fear of rallying support for the Eurosceptical Tories in the UK is evidence that parties are aware that voters understand party willingness to compromise. The 2006 Canadian Federal election campaign, widely expected to produce a minority government, featured much talk about willingness to compromise. The Conservative leader publicly stated that his government would find common ground with the left-leaning New Democratic Party (NDP) instead of the centre Liberal Party (Curry and Chase 2006). The possibility of the Conservative-NDP coalition influenced some voters at election time.

In the model two three-party parliamentary legislatures, A and B, jointly make policy decisions. With concurrent elections in any two jurisdictions being rare, we focus on the election in jurisdiction B and take as given the formateur in A and the existence of a status quo policy, and develop the simplest one-dimensional model that captures the point we wish to make: a single election, complete information, the same set of parties competing in the two jurisdictions and formateurs selected probabilistically by vote shares.[2] Intergovernmental negotiations are modeled as Nash (1950) bargaining between the formateurs. The resulting game has a simple and intuitive structure lead-

[2] Diermeier and Merlo (2004) find stronger support for probabilistic rather than deterministic selection rules. The probabilistic formateur selection rule allows us to focus on the effect on electoral and policy outcomes of parties' relative willingness to compromise.

ing to an intergovernmental political economy equilibrium for each possible formateur in jurisdiction B.

When choosing among the parties in jurisdiction B, policy-oriented voters must rank the agreements they anticipate will be reached by each pair of potential negotiators. Although, the agreement between formateurs A and B, the AB pair, always lies between their ideal policies, their agreement also depends on the bargainers' "intensity of preferences" (Peters 1992), which we interpret as willingness to compromise, represented by the interaction of the location of its ideal policy with the curvature of its utility function. The ranking of agreements reached by the three pairs of negotiators depends on the formateurs' relative willingness to compromise. Differences between the pairs' willingness to compromise generate different policy rankings each leading to different electoral equilibria. In one, the ranking of agreements and parties' ideal policies coincide. Since this is the ranking usually expected, we refer to it as the *typical ranking*. Associated with this set of agreements is an electoral equilibrium where voters partition themselves from left to right among the parties.

However, when one extreme party is more willing to compromise than the centre party, the anticipated policy agreements may not follow the ranking of the party's ideals. A left wing and centrist party, for example, could negotiate an agreement to the right of the agreement reached between the same left party and a right wing party.[3] The *policy reversal* leads to an *atypical* ranking which in turn leads to a voting equilibrium where support for one of the two extreme parties comes from moderate voters and support for the centre party comes from extreme voters. We extend our typical or atypical ranking labels to the corresponding voting equilibria.

To show these equilibria we use the results of Cressman and Gallego (2009). We examine the equilibrium and show how the relationship between parties' utility affects the ranking of agreements and the electoral outcome. When parties have twice continuously differentiable utility functions, the ranking depends on the interaction of the distance between the parties' ideal policies and the relative curvature of the parties' utility functions.[4] We examine the ideal point and curvature effects separately. First, we show the ranking of agreements when the utility functions of the centre and the right party are *translations* of one another, i.e., differ only in the location of their ideals. Then, we examine the ranking when the utility functions of the centre and right parties are *transformations* of one another at the same ideal point, so that they differ only in curvature. We can then examine, for example, what happens when the right party is more risk averse than the centre party by ruling out the influence that differences in ideal policy may have on the bargaining outcome (Kihlstrom et al. 1981). Finally, we examine the ranking of agreements and ideals for a broader set of utility functions by

[3] If given the chance the extreme parties themselves may prefer to abdicate their role as negotiator. We do not allow abdication. By assuming that parties also benefit from governing in ways additional to the power to implement a preferred policy, the incentive for abdication can be mitigated.

[4] For example when at least one party's preference is represented by a Euclidean utility, the ranking does not depend on the location of the party's ideal policy. See Sect. 3.

composing a right translation with a transformation.[5] In all three cases, we show that both a typical and an atypical electoral equilibrium may emerge.

Intergovernmental policy making has been studied in two-tier *two*-party systems. Alesina and Rosenthal (1995, 1996) and Fiorina (1996), among others, model policy making as an exogenous compromise between the preferred policies of the President and Congress. We show that this assumption is inappropriate when intergovernmental bargaining takes place between two *three*-party legislatures. Crémer and Palfrey (2000, 2002, 2006) endogenize intergovernmental negotiations over federal standards. Persson and Tabellini (1996) develop model in which two median regional voters engage in Nash bargaining over intergovernmental transfers in a Federation.

2 The model

Representatives of the governments of jurisdictions A and B negotiate a revision to an existing policy.[6] If negotiations fail, a status quo policy Q remains in force. In anticipation of the negotiations, an election determines the composition of the legislature in jurisdiction B and the identity of B's negotiator. The policy, $\theta \in [0, 1]$, is assumed to be unidimensional. The players are the (L)eft, (C)enter and (R)ight parties in each legislature, and a large number of heterogeneous voters. Players care only about policy outcomes, and preferences are single-peaked at an ideal policy. The voters' ideal policies are distributed on the policy space according to Γ. For a given policy θ that is implemented, player i whose ideal policy is $\hat{\theta}_i \in [0, 1]$ receives a payoff $u_i(\theta) = u_i(\theta|\hat{\theta}_i)$. We assume that if voters j and m prefer θ^1 to θ^0, so does every voter in-between: voter preferences satisfy Greenberg and Weber (1985) *Consecutiveness Condition* (CC): for any three voters j, k, and m and any two policies $\theta^0 < \theta^1$,

$$\begin{aligned} &\text{if } u_j(\theta^1) > u_j(\theta^0) \ \text{ and } \ u_m(\theta^1) > u_m(\theta^0) \\ &\text{then } u_k(\theta^1) > u_k(\theta^0) \text{ for all } k \text{ such that } \hat{\theta}_j < \hat{\theta}_k < \hat{\theta}_m. \end{aligned} \tag{CC}$$

Parties are labeled according to their ideal policies $0 \leq \hat{\theta}_{\mathrm{L}} < \hat{\theta}_{\mathrm{C}} < \hat{\theta}_{\mathrm{R}} \leq 1$, and may also differ in concavity of their utility functions. For simplicity, we assume that the same set of parties compete in both jurisdictions; if these sets differ, the notation becomes more cumbersome but our results remain. We assume that party discipline is strong enough to overcome any incentive problems that divide the interests of a party and its negotiators. The model divides into numerous cases depending on the identity of the jurisdiction A formateur and the location of the status quo. Many of these

[5] In Europe, parties have split or merged (Gallagher et al. 2005). When parties split, the two new parties can share the same ideal (e.g., the same position on European integration) yet differ in their willingness to compromise (e.g., on degree of European integration) as in our transformation case. Alternatively, the two parties could differ only in their ideal (e.g., a change of view with respect to European integration) as in our translation case. Of course, the new parties may have preferences that differ in both ideal and curvature.

[6] The model can be reinterpreted by identifying formateur A with the head of another government, an incumbent president during midterm elections, or an entrenched bureaucracy. We can also interpret the models as the election of a State formateur who engages in negotiations with its Federal counterpart.

are qualitatively similar, with one half being "reflections" of the other. To simplify, throughout we assume that $Q \leq \hat{\theta}_{\mathrm{C}}$.

2.1 Timing

Two subgames comprise the policy revision process: the election in jurisdiction B and the intergovernmental policy negotiation. All references to elections are implicitly those in jurisdiction B.

(i) *Election in jurisdiction B*: Taking as given the status quo policy Q and knowing the identity of A's formateur, citizens simultaneously vote to elect a government in jurisdiction B. Voters anticipate the behavior of all parties in post-election negotiations, and these depend on the parties' preferences *over all* outcomes, i.e., on their entire utility function. We assume that the only credible platforms are the parties' true preferences, $U = (u_{\mathrm{L}}, u_{\mathrm{C}}, u_{\mathrm{R}})$. The election determines each party's vote share $V^K \in [0, 1]$, $K \in \{\mathrm{L}, \mathrm{C}, \mathrm{R}\}$, which represents the weight of the party in the legislature. Following the election, B's formateur is selected according to vote shares. If party K wins a majority, $V^K \geq 1/2$, it becomes formateur. If no party wins a majority, a formateur is randomly chosen according to vote shares.

(ii) *Intergovernmental policy negotiation*: Next formateurs A and B, the AB pair, engage in negotiations, where the status quo Q is both the breakdown and deadlock outcome. We assume the intergovernmental agreement cannot be unilaterally altered by any formateur.[7] Based on the final policy θ^*, payoffs are realized, and the game ends.

2.2 A numerical example

At the heart of the model is a possibility result: voters may prefer to support a party whose ideal is further from their own than that of an alternative party, provided that the more distant party's *attitude toward compromise* leads to the negotiation of a more preferred outcome. Before deriving results formally to understand the mechanism behind these reversals, we set out some examples of the phenomenon.

In these examples, we assume that L is the formateur in jurisdiction A, so the possible pairs of AB formateurs are LL, LC, and LR, and that voter ideals are distributed uniformly on [0,1] with status quo $Q = 0$. The examples differ in what is assumed about the properties of the parties' utility functions. These are summarized in Table 1. We illustrate a case where the ordering of negotiated agreements coincides with the party ideals and cases where it does not. In the latter, atypical electoral outcomes emerges: some centre voters support an extreme party, and some extreme voters support the centre party.

Let $a \in [0, 1]$ denote the policy agreement. Clearly, the formateur pair $AB =$ LL agrees on $a^{\mathrm{LL}} = \hat{\theta}_{\mathrm{L}} = 0.25$. For the pairs LC and LR, we assume negotiations

[7] In a earlier version of the article, we assumed the intergovernmental agreement required simultaneous ratification by the two legislatures, and showed when and how ratification constrains negotiations.

Table 1 Example

	Formateur $A = L$ Status quo: $Q = 0$		
	Translation effect	Transformation effect	Combined effect
Preferences	$u_K(\theta) = -(\theta - \hat{\theta}_K)^2$ for K=L, C, R	$u_{\mathrm{K}}(\theta) = -(\theta - \hat{\theta}_K)^2$ for K=L, C and $u_{\mathrm{R}}(\theta) = -(\theta - \hat{\theta}_{\mathrm{R}})^4$	$u_K(\theta) = -(\theta - \hat{\theta}_K)^2$ for K=L, C and $u_{\mathrm{R}}(\theta) = -(\theta - \hat{\theta}_{\mathrm{R}})^4$
Ideal policies	$\hat{\theta}_{\mathrm{L}} = 0.25$, $\hat{\theta}_{\mathrm{C}}$=0.5, and $\hat{\theta}_{\mathrm{R}} = 0.75$	$\hat{\theta}_{\mathrm{L}} = 0.25$ and $\hat{\theta}_{\mathrm{C}} = \hat{\theta}_{\mathrm{R}} = 0.5$	$\hat{\theta}_{\mathrm{L}} = 0.25$, $\hat{\theta}_{\mathrm{C}}$=0.5 and $\hat{\theta}_{\mathrm{R}} = 0.75$
Formateur B			
(i) L	$a^{\mathrm{LL}} = \hat{\theta}_{\mathrm{L}} = 0.25 > Q = 0$	$a^{\mathrm{LL}} = \hat{\theta}_{\mathrm{L}} = 0.25 > Q = 0$	$a^{\mathrm{LL}} = \hat{\theta}_{\mathrm{L}} = 0.25 > Q = 0$
(ii) C	$a^{\mathrm{LC}} = 0.305$	$a^{\mathrm{LC}} = 0.305$	$a^{\mathrm{LC}} = 0.305$
(iii) R	$a^{\mathrm{LR}} = 0.317$	$a^{\mathrm{LR}} = 0.273$	$a^{\mathrm{LR}} = 0.293$
	Voting equilibrium	Voting equilibrium	Voting equilibrium
	(0, 0.277) vote for L (0.277, 0.311) vote for C (0.311, 1) vote for R A typical policy ranking and voting equilibrium emerge where R wins a majority	(0, 0.261) vote for L (0.261, 0.288) vote for R (0.288, 1) vote for C An ***atypical*** policy ranking and voting equilibrium emerge where right wing voters give C a majority	(0, 0.271) vote for L (0.271, 0.299) vote for R (0.299, 1) vote for C An ***atypical*** policy ranking and voting equilibrium emerge where right wing voters give C a majority

are carried out using Nash (1950) bargaining (discussed below). Table 1 shows the agreement for formateurs B = L, C, R and the electoral equilibrium. The agreement for each formateur is given in rows (i) to (iii) and the electoral equilibrium in the last row. In column 2, parties have quadratic preferences, $u_K(\theta) = -(\theta - \hat{\theta}_K)^2$ for K = L, C, and R where the parties ideals are $\hat{\theta}_{\mathrm{L}} = 0.25$, $\hat{\theta}_{\mathrm{C}} = 0.5$ and $\hat{\theta}_{\mathrm{R}} = 0.75$ (see Fig. 1). Under quadratic preferences the ranking of agreements and of parties' ideal policies coincide. Citizens have rational expectations regarding the behavior of other voters, and vote to realize their most preferred policy outcome. In this example, they vote for the agreement closest to their own ideal, and a *typical voting equilibrium* emerges where voters to the left of marginal voter 0.277 vote for L, those to the right for marginal voter 0.311 vote for R, and those in-between vote for C. Notice that vote balancing occurs as voters account for the influence of formateur A. Had they voted instead for the party whose *ideal* is closest to their own the partition would be such that those on the left of 0.375 vote for L, those on the right of 0.625 vote for R and those in-between vote for C.

To give a flavor of the formal results below, we separate the differences in party preferences into a "shift" and a "curvature" component. In column 3, L and C continue to have quadratic utilities, we assume C's and R's ideals *coincide*, $\hat{\theta}_{\mathrm{C}} = \hat{\theta}_{\mathrm{R}} = 0.5$ but change R's utility to $u_{\mathrm{R}}(\theta) = -(\theta - 0.5)^4$. The agreement of the LC pair (row ii) is to the ***right*** of that of LR's (row iii) showing that R who is more risk averse than C makes

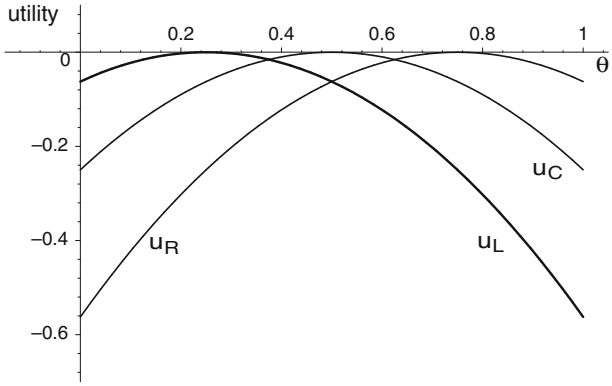

Fig. 1 The utility functions of the example in Sect. 4: L's utility is the left parabola $u_{\mathrm{L}}(\theta) = -(\theta - 0.25)^2$, C's the middle parabola $u_{\mathrm{C}}(\theta) = -(\theta - 0.5)^2$, and R's the right parabola $u_{\mathrm{R}}(\theta) = -(\theta - 0.75)^2$ for $\theta \in [0, 1]$

a bigger compromise when bargaining with L than C does to avoid the breakdown of negotiations (Kihlstrom et al. 1981). The ranking of agreements differs from that of parties' ideals leading to a *policy reversal* and to an *atypical voting equilibrium*. Voters to the left of 0.261 vote for L; those to the *right* of 0.288 vote for ***C***; and those *in-between* vote for R. In column 4, L and C maintain their quadratic utilities but R's changes to $u_{\mathrm{R}}(\theta) = -(\theta - 0.75)^4$, the reversal of policies leads to an *atypical* voting equilibrium.

We now develop the model in detail.

3 Equilibrium

An *intergovernmental political economy* (IPE) equilibrium is a set containing **(i)** a utility maximizing voting function $\pi_i^*(a^{AK}|Q, A, U)$ that maps voter types $\hat{\theta}_i$ to the probability of supporting each party, and **(ii)** a bargaining outcome function a^{AB} that maps each possible combination of formateurs A and B and the status quo into a policy agreement. These functions depend on the preferences of all parties, the status quo, the identity of formateur A, the distribution of voter preferences. We find the IPE equilibrium by solving backwards through the stages of the game.

3.1 The intergovernmental bargaining equilibrium

Let a^{AB} denote the policy agreement between formateurs A and B.[8] For any Q, let $\Phi_K(Q) = \{\theta | u_K(\theta) \geq u_K(Q)\}$, $K \in$ L, C, R denote the (convex) upper contour sets

[8] In a previous version intergovernmental negotiations were modeled using Binmore et al. (1986) alternating offers game where negotiations may exogenously breakdown after any rejection. In the limit as the time between periods and the risk of breakdown become small the agreement converges to the Nash solution. This introduces some more complications, but the primary results remain unchanged.

of party K, the set of policies that make party K no worse off than Q. Define the upper bound of party K's acceptable set $\Phi_K(Q)$ as $\overline{\theta}_K$. Two cases can be distinguished.

Case 1: When the *same* party holds power in both jurisdictions, there is no disagreement on the best policy, so this is not a bargaining problem in the sense of Nash (1950). The policy agreement is the party's ideal. $a^{KK} = \hat{\theta}_K$.

Case 2: When the two negotiators *differ*, there is disagreement over which feasible policy is best. Since jurisdictions have equal "bargaining power" we use JK to denote the pair of formateurs disregarding which jurisdiction they represent. The three JK, $J \neq K$, configurations are LC, LR, or CR. We model negotiations using Nash (1950) bargaining framework where Q represents the threat point with disagreement outcome $\mathrm{d}^{JK}(Q) = (u_J(Q), u_K(Q))$.

The feasible set, i.e., the set of utilities pairs over which the formateurs negotiate, is given by[9]

$$S^{JK}(Q) = \{u_J(a), u_K(a) | a \in \Phi_J(Q) \cap \Phi_K(Q)\}.$$

The feasible set $S^{JK}(Q)$ and the disagreement point $\mathrm{d}^{JK}(\mathrm{Q})$ define the bargaining problem. Among the many bargaining[10] solutions, we use that proposed by Nash (1950). For any $a \in \Phi_J(Q) \cap \Phi_K(Q)$, the Nash Solution (NS) to our bargaining problem maximizes the product of the utility gains from the disagreement point, i.e., solves

$$\max_a \; NP_{JK}(a|Q) = [u_J(a) - u_J(Q)][u_K(a) - u_K(Q)].$$

Since the Nash solution is Pareto optimal and Pareto dominates Q, not every point on the frontier of $S^{JK}(Q)$ is a candidate. First assume that $Q < \hat{\theta}_J$. For $a \in [Q, \hat{\theta}_J)$, both parties J and K prefer rightward changes in policy (e.g., the lower upward sloping segment in Fig. 2). When $a \in [Q, \hat{\theta}_J)$ and $\hat{\theta}_K \leq \bar{\theta}_J$, both prefer agreements to left of $\bar{\theta}_J$ (not shown in Fig. 2). In this case the set of potential agreements is $[\hat{\theta}_J, \hat{\theta}_K]$. When $a \in [Q, \hat{\theta}_J)$ and $\hat{\theta}_K > \bar{\theta}_J$, this set is smaller, restricted to $[\hat{\theta}_J, \bar{\theta}_J]$. Let the *upper* bound of the mutually acceptable policies be given by

$$\tilde{\theta}_J = \min\{\bar{\theta}_J, \hat{\theta}_K\} \tag{1}$$

Thus, the bargaining set of *mutually* acceptable policies is $\wp^{JK}(Q) = [\hat{\theta}_J, \tilde{\theta}_J]$. Instead $\wp^{JK}(Q) = Q$ when $Q \in [\hat{\theta}_J, \hat{\theta}_K]$ as no agreement is a Pareto improvement on the status quo. We thus conclude:

Result 1 (Nash 1950): *Given the status quo $Q \in [0, \hat{\theta}_{\mathrm{C}}]$, the party's preferences U, if different parties control the legislatures there is a unique solution to the intergovernmental Nash Bargaining problem.*

[9] With concave utilities, this set is convex, closed, and bounded.

[10] The results are qualitatively similar if we use instead the Kalai and Smorodinsky (1975), the Perles and Maschler (1981) or the Egalitarian (Kalai 1977) solution. See Cressman and Gallego (2009). On the solutions to Nash's bargaining problem see e.g., Peters (1992) and Thomson (1994, 2009).

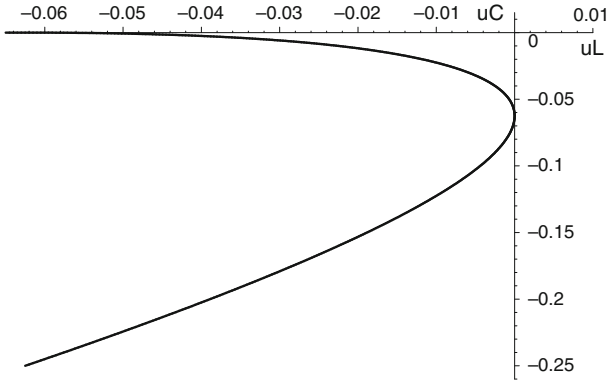

Fig. 2 The feasible set S between L (horizontal axis) and C (on the vertical axis). The downward sloping portion represents the Pareto set. The policy $\bar{\theta}_{\mathrm{L}}$ that keeps L indifferent to the status quo ($Q = 0$), does not constrain negotiations since $\hat{\theta}_{\mathrm{C}} = \bar{\theta}_{\mathrm{L}} = 0.5$. The upward sloping segment is for policies in $[Q, \hat{\theta}_{\mathrm{L}})$, a region in which both formateurs prefer rightward policy changes. The bargaining set is the set of mutually agreeable policies $\left[\hat{\theta}_{\mathrm{L}}, \bar{\theta}_{\mathrm{C}}\right]$

Case 1 (above) and Result 1 lead to the following proposition.

Proposition 1 *Given the status quo $Q \in [0, \hat{\theta}_{\mathrm{C}}]$, the party's preferences U and formateur A , intergovernmental negotiations lead to a unique intergovernmental agreement.*

When $\wp^{JK}(Q) = Q$, the LC ad LR pairs "agree" on Q, which is to the left of CR's agreement and agreements follow the typical ranking. Since we are interested in situations in which an atypical policy ranking may emerge, from now on we focus on situations in which $\wp^{JK}(Q) \neq Q$, i.e., $Q < \hat{\theta}_{\mathrm{L}}$ where both parties and voters want to change the status quo.

The set of possible equilibrium agreements Θ^{A*} for $A = \mathrm{L}$, C, and R respectively is

$$\Theta^{\mathrm{L}*} = \left\{a^{\mathrm{LL}} = \hat{\theta}_{\mathrm{L}}, a^{\mathrm{LC}}, a^{\mathrm{LR}}\right\} \tag{2}$$

$$\Theta^{\mathrm{C}*} = \left\{a^{\mathrm{CL}}, a^{\mathrm{CC}} = \hat{\theta}_{\mathrm{C}}, a^{\mathrm{CR}}\right\} \tag{3}$$

$$\Theta^{\mathrm{R}*} = \left\{a^{\mathrm{RL}}, a^{\mathrm{RC}}, a^{\mathrm{RR}} = \hat{\theta}_{\mathrm{R}}\right\} \tag{4}$$

Note that regardless who is formateur in jurisdiction A, there are only three possible agreements associated with each A formateur. This is important at the election stage.

3.2 The election in jurisdiction B

The formateur B is selected by vote share in the election. Party $K = \{\mathrm{L}, \mathrm{C}, \mathrm{R}\}$ forms government with probability one when $V^K \geq 1/2$. If no party wins a majority, K

forms government with probability V^K. The probability of K becoming the formateur in jurisdiction B is[11]

$$\mu^K = \begin{cases} 1 & \text{if } V^K \geq 1/2 \\ 0 & \text{if } \max\{V^{-K}\} \geq 1/2 \\ V^K & \text{otherwise} \end{cases} \tag{5}$$

Voters know the preferences of each party, the identity of formateur A and the location of the status quo. Under complete information and no uncertainty, voters rationally anticipate the final policy resulting from any particular election outcome.

Electoral games with a continuum of voters typically have many Nash equilibria since no single voter can affect the electoral outcome, i.e., no voter is pivotal in the election. Voting for any party is then a best response of any voter to any strategy used by all other voters. There is for example a voting equilibrium where everyone votes for the same party. In this equilibrium, some vote for a party that conditional on being chosen as formateur would deliver the voter's lowest ranked policy. We require voters use weakly undominated strategies, which rules out equilibria in which voters' vote less preferred parties (Patty et al. 2009).

Given formateur A and the associated intergovernmental agreements Θ^{A*}, voters *rank* parties by the agreements each can deliver (see additional discussion on policy rankings in Sect. 4). Using this ranking, we drive the voting equilibrium using Austen-Smith's (2000) voting model.

A voting equilibrium is a symmetric mixed voting strategy, π^* such that for all $\hat{\theta}_i \in [0, 1]$ given $Q, \wp, U$, and the anticipated agreements Θ^{A*}, $\pi_i^*(\theta^*(a^{AB*})|Q, A, U)$ for $B = \text{L, C, R}$, is weakly undominated and maximizes voter i's expected payoff.

Lemma 1 (Austen-Smith 2000) *Given the status quo $Q \in [0, \hat{\theta}_{\text{L}}]$, the party's preferences U, formateur A and the set of anticipated equilibrium agreements Θ^{A*}, if π^* is a voting equilibrium, voter of type i votes with positive probability only for the party that offers the highest payoff conditional on that party being selected as formateur in jurisdiction B. If two parties offer the same expected payoff, voters vote for the party with the ideal closest to their own.*

4 Equilibrium characterization when $A = \text{L}$

We now conduct comparative static exercises to demonstrate how the voting equilibrium depends on the parties' relative willingness to compromise. Given the assumption that the status quo is to the left of L's ideal, we can derive both typical and atypical voting equilibria in the case that $A = \text{L}$, so focus on this case.[12] To simplify notation, we

[11] This assumption implies that the marginal contribution of a voter is independent of the distribution of votes. Without it, many more equilibria would exist. We are grateful to a referee for emphasizing this point.

[12] A previous version of this article that included the requirement that the bargaining agreement be simultaneously ratified by each legislature, we needed to also consider the majority or minority status of the formateurs. Under the assumption that Q is left of party L's ideal, the results for $A = \text{L}$ were unaffected. However, this assumption does complicate the case where R is the formateur in jurisdiction A, but policy

use a^{AB*} to denote the equilibrium policy outcome $\theta^*(a^{AB*})$. The intergovernmental agreements are $\Theta^{L*} = \left\{\hat{\theta}_L, a^{LC*}, a^{LR*}\right\}$. Notice in particular that when $Q < \hat{\theta}_L$, if L controls both houses then $a^{LL*} = \hat{\theta}_L$ independently of L's willingness to compromise. Consequently, we need consider only the relationship between C's and R's preferences.

To add structure to the term "willingness to compromise," we consider two independent characteristics of the parties' preferences: its *ideal policy* and the *curvature* of its utility function, first separately and then in combination. We first examine what happens when R's utility is a *translation* of C's, so that the utilities differ only in the location of their ideals. This incorporates a common (though sometimes implicit) assumption of legislative bargaining models where parties have quadratic utility functions that translations of one another.[13] Below we give a more general treatment of translations than that given by quadratic functions.[14] Next we *transform* the curvature of C's utility to generate R's when C and R share the *same* ideal and use the well known result of Kihlstrom et al. (1981) to isolate the effect that changes in the curvature (or degree of risk aversion) of R's utility has on the equilibrium. Finally, following Cressman and Gallego (2009), we examine the more general case where R's utility can be obtained as a translation combined with a transformation of C's. Below we only summarize the effect of this combination; details are in the Appendix.

4.1 The examples again

First reconsider the examples in Table 1. Column 2 illustrates the translation effect since R's utility is a *right* translation of C's, i.e., $\hat{\theta}_R = \hat{\theta}_C + 0.25$ and $u_R(\theta) = u_C(\theta - 0.25)$. The ranking of agreements and of the parties' ideals coincide leading to a typical voting equilibrium (Lemma 1). In Column 3, R's utility is a *concave* transformation of C's, since $u_R(\theta) = -(u_C(\theta - 0.5))^2$ and $\hat{\theta}_C = \hat{\theta}_R = 0.5$. The agreement of the LR pair is to the *left* of LC's illustrating Kihlstrom et al. (1981) result: L prefers to bargain with R who is more risk averse and easier to satisfy than C in negotiations.

In column 4, R's utility is a combination of a right translation ($\hat{\theta}_R = \hat{\theta}_C + 0.25$) and a concave transformation of C's utility, since $u_R(\theta) = -(\theta - \hat{\theta}_R)^4 = -[u_C(\theta - 0.25)]^2$. Columns 2 and 3 illustrate that the combined effect (column 4) can be decomposed into the translation (column 2) and the transformation (column 3)

reversals between the RL and RC pairs may still occur. Among other changes, it need not be that voters vote with positive probability only for the party that offers the highest payoff conditional on that party being selected as formateur. Voting to maximize the probability that their most preferred agreement is implemented may require voters to support a party they do not prefer as formateur.

13 Quadratic utilities give tractable solutions (see e.g., Austen-Smith and Banks 1988; Coate and Knight 2007 and Cho and Duggan 2003) and satisfy the Single-Crossing (SC) condition given by Milgrom and Shannon (1994). Cho and Duggan (2003, p. 128) point out that under quadratic utilities "… individual preferences over lotteries on X are order restricted when individuals are ordered in the order of their ideal points." This important increasing monotonicity property is the basic building block of many proofs.

14 Cressman and Gallego (2009) point out that there are utility functions that satisfy Milgrom and Shannon's SC condition that do not satisfy the increasing monotonicity in ideals property. Implying that the SC condition is insufficient to generate agreements whose ranking coincides with the parties' ideals.

effects. In this example, the transformation effect shifts the LR agreement to the *left* of LC's by more than the rightward shift induced by the translation effect.

We now analyze for more general preferences the effect on the IPE equilibrium of a translation and a transformation of party preferences. In the appendix we show how these results combine, in some cases reinforcing one another and in others offsetting one another.

4.2 Translation

Assume that R's utility is a *translation* of C's to the right by $\alpha \geq 0$ units, so that, C's and R's utilities differ only in the location of their ideals, $u_{\mathrm{R}}(\theta) = T_\alpha(u_{\mathrm{C}}(\theta)) = u_{\mathrm{R}}(\theta - \alpha)$. Let $a_\alpha^{\mathrm{LR}*}$ denote the agreement of the LR pair and ν_α^{1*} and ν_α^{2*} the marginal voter types in the electoral equilibrium.

In spite of being translations of one another, C's and R's willingness to compromise differ. This is because, in general, the concavity of a utility function varies with the distance to the ideal, so at any particular agreement the *local* concavity of the parties' utilities *differ*. For the translation case, concavity and hence willingness to compromise in the neighborhood of a particular agreement is captured by the Arrow-Pratt coefficient of absolute risk aversion, ARA.[15] To make comparisons we examine classes of preferences for which the ARA coefficient either increase or decreases in policy outcomes. We examine preferences characterized by Increasing ARA (IARA, $\mathrm{d}(\mathrm{ARA}(\theta))/\mathrm{d}\theta > 0$) or Decreasing ARA (DARA, $\mathrm{d}(\mathrm{ARA}(\theta))/\mathrm{d}\theta < 0$).

Recall that the upper bound of the set of mutually agreeable policies is $\tilde{\theta}_{\mathrm{L}}$ (Eq. 1). The following lemma follows immediately from Cressman and Gallego (2009).

Lemma 2 (Cressman and Gallego 2009) *Suppose formateur $A = L$ and that u_{L} is a unimodal twice continuously differentiable utility function with ideal $\hat{\theta}_{\mathrm{L}}$ with bargaining set $\wp^{LK}(Q) = [\hat{\theta}_{\mathrm{L}}, \tilde{\theta}_{\mathrm{L}}]$ for $K = C,R$ and status quo $Q < \hat{\theta}_{\mathrm{L}}$. Suppose u_{C} and u_{R} are increasing utility functions for all $a \in [Q, \tilde{\theta}_{\mathrm{L}}]$ where u_{R} is the horizontal translation of u_{C} to the right by $\alpha \geq 0$ units (i.e., $u_{\mathrm{R}}(\theta) = T_\alpha(u_{\mathrm{C}}(\theta)) = u_{\mathrm{C}}(\theta - \alpha)$).*

(a) *If u_{C} has IARA for all $Q - \alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$, the agreements of the LR and LC pairs satisfy $a_\alpha^{\mathrm{LR}*} \geq a^{\mathrm{LC}*}$.*
(b) *If u_{C} has DARA for all $Q - \alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$, the agreements of the LR and LC pairs satisfy $a_\alpha^{\mathrm{LR}*} \leq a^{\mathrm{LC}*}$.*

When R's utility is a right translation of C's, the ranking of agreements depends on whether u_{C} exhibits IARA or DARA. When u_{C} has IARA for all $Q - \alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$, the agreement of the LC pair is to the *left* of LR's, so the ranking of agreements and party's ideals coincide (note: quadratic utilities exhibit IARA). The equilibrium agreements of the LL, LC and LR pairs follow the ordering of the party's ideals, so that a typical ranking emerges, i.e.,

$$a^{\mathrm{LL}*} = \hat{\theta}_{\mathrm{L}} < a^{\mathrm{LC}*} < a_\alpha^{\mathrm{LR}*} < \hat{\theta}_{\mathrm{R}}.$$

[15] $\mathrm{ARA}_K(\theta) = -u''_K(\theta)/u'_K(\theta)$ for $K = \mathrm{L}, \mathrm{C}, \mathrm{R}$.

Table 2 Summary of Propositions 2 and 3

Assume formateur $A = L$ and …	
Proposition 2: translation	
u_R a right translation of u_C by $\alpha > 0$ units and ….	
….u_C exhibits *IARA*	….u_C exhibits *DARA*
Typical policy ranking and voting equilibrium	Atypical policy ranking and voting equilibrium
Proposition 3: transformation	
u_R a …... transformation of u_C	
…. concave …	…. convex…
Atypical policy ranking and voting equilibrium emerge	Typical policy ranking and voting equilibrium emerge

Given that there are only three potential agreements, from Lemma 2, in the electoral equilibrium voters partition themselves such that those to the left of the marginal voter ν_α^{1*} vote for L, those to the right of the marginal voter ν_α^{2*} vote for R and those in between $(\nu_\alpha^{1*}, \nu_\alpha^{1*})$ vote for C. We call this the *typical* voting equilibrium.

However, when C's utility exhibits[16] DARA the agreement of the LR pair is to the *left* of LC's, so the equilibrium agreements of the LC and LR pairs lead to a *policy reversal* and thus to an atypical ranking of agreements, i.e.,

$$a^{\mathrm{LL}*} = \hat{\theta}_{\mathrm{L}} < a_\alpha^{\mathrm{LR}*} < a^{\mathrm{LC}*} < \hat{\theta}_{\mathrm{R}}.$$

Policy-oriented voters choose the party that conditional on being chosen formateur implements the agreement they most prefer, rather than vote for the party whose ideal is closest to their own. When C's utility exhibits DARA and R's is a translation of C's to right, in the electoral equilibrium voters to the *left* of ν_α^{1*} vote for L, voters to the *right* of ν_α^{2*} vote for C and *moderate* voters—those *in between* $(\nu_\alpha^{1*}, \nu_\alpha^{2*})$ — vote for R. Even in this simple case where utilities are translations of one another, the model produces an *atypical* voting equilibrium.

In the translation case, the parties' willingness to compromise depends only on the ARA coefficient of C's utility function which in turn determines R's willingness to compromise and effectively determines the voters' left–right ordering of parties.

We summarize the above results in the following Proposition (see summary in Table 2).

Proposition 2 *Suppose that the conditions of Lemma 2 hold. Assume voters know that formateur $A = L$, the status quo $Q \in [0, \hat{\theta}_{\mathrm{L}}]$, the party's preferences U, and the set of anticipated equilibrium agreements $\Theta^{\mathrm{L}*}$.*

[16] Cressman and Gallego (2009) show that Lemma 2 must be applied with care since a utility function that is always DARA increases for values greater than $\theta = 1$. This is important when we shift utilities to the right. They show however that C's utility function must be DARA only on the interval $[Q, \tilde{\theta}_{\mathrm{L}}]$ to conclude the agreement of the LR pair is to the *right* of LC's and thus further from L's ideal. Consequently, C's DARA utility function can be modified so that its ideal is in $(\tilde{\theta}_{\mathrm{L}}, 1]$. We assume C's unimodal utility is DARA in $\tilde{\theta}_{\mathrm{L}} < \theta < 1$ so that the statements of Lemma 2 and Proposition 2 below remain unaffected.

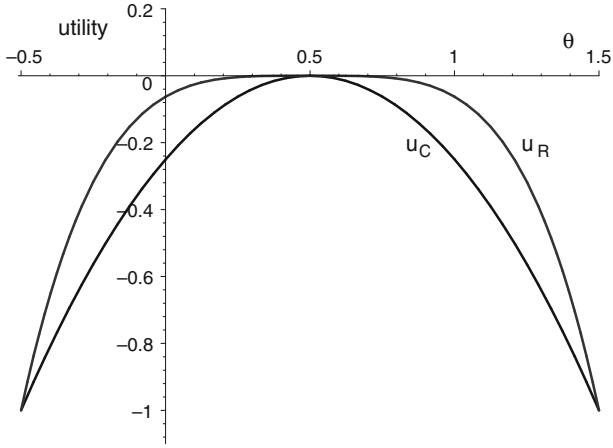

Fig. 3 The graph of R's utility function when it is a concave transformation of C's: $u_{\mathrm{C}}(x) = -(x-0.5)^2$ and $u_{\mathrm{R}}(x) = -(x-0.5)^4$

*(a) If u_{C} has **IARA** for all $Q-\alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$, the intergovernmental policy agreements follow the ranking of the party's ideal policies. There exists a typical electoral equilibrium where voters partition themselves along the left–right scale among the three parties.*

*(b) If u_{C} has **DARA** for all $Q-\alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$, the ranking of the intergovernmental agreements and of party's ideal policies differ. There exists an atypical electoral equilibrium where some extreme voters vote for the center party and some moderate voters vote for R, one of the extreme parties.*

4.3 Euclidean preferences

Proposition 2 should be applied with care as it requires that the parties' utilities be twice continuously differentiable. For instance, Proposition 2 does not apply when either C or R or both have Euclidean preferences. Euclidean preferences exhibit constant ARA everywhere except at the party's ideal policy. In this case, the LC and LR pairs reach the *same* agreement implying that the agreement is independent of the shift factor α.

4.4 Transformation when A = L

We next examine how differences in the *curvature* of C's and R's utility affect the equilibrium when C and R share the *same* ideal. Let $u_K(\theta) = T_\Psi(u_J(\theta))$ where T_Ψ denotes the transformation of $u_{\mathrm{C}}(\theta)$ and $\Psi : [u_J(Q), u_J(\hat{\theta}_J)] \to \Re$ is a continuously strictly increasing function such that $\Psi'(\theta) > 0$ and $\Psi''(\theta) < 0$ (Fig. 3). Let $a_\Psi^{\mathrm{AC}*}$ denote the agreements of the LR pair and let the marginal voter types of the electoral equilibrium be identified by ν_Ψ^{1*} and ν_Ψ^{2*}.

The bargaining outcome depends on whether R's utility is a *concave* transformation of C's, or vice versa. We treat these cases in turn.

Lemma 3 (Cressman and Gallego 2009) *Suppose formateur $A = L$ and that $u_{\mathrm{L}}(\theta)$ and $u_{\mathrm{C}}(\theta)$ are single-peaked twice continuously differentiable utility functions with ideals $\hat{\theta}_{\mathrm{L}} < \hat{\theta}_{\mathrm{C}}$, status quo $Q < \hat{\theta}_{\mathrm{L}}$ and bargaining set $\wp = [\hat{\theta}_{\mathrm{L}}, \tilde{\theta}_{\mathrm{L}}]$. If $u_{\mathrm{R}}(a)$ is a concave transformation of $u_{\mathrm{C}}(a)$, then $a_{\Psi}^{\mathrm{LR}*} < a^{\mathrm{LC}*}$.*

Intuitively, Lemma 3 says that when R's utility is a concave transformation of C's, R is easier to satisfy than C and thus, more willing to compromise than C (Peters 1992). The agreement of the LR pair is then to the *left* of LC's (Kihlstrom et al. 1981). This *policy reversal* leads to *an atypical* policy ranking,

$$a^{\mathrm{LL}*} = \hat{\theta}_{\mathrm{L}} < a_{\Psi}^{\mathrm{LR}*} < a^{\mathrm{LC}*} < \hat{\theta}_{\mathrm{R}}.$$

Lemma 3 states that there is an atypical voting equilibrium where those to the left of ν_{Ψ}^{1*} vote for L, those to the right of ν_{Ψ}^{2*} vote for C and those in between $(\nu_{\Psi}^{1*}, \nu_{\Psi}^{2*})$ vote for R. Obviously, the opposite ordering holds when C's utility is a *concave* transformation of R's. These results are summarized in Table 2.

Proposition 3 *Suppose that the conditions of Lemma 3 hold. Assume voters know formateur A = L, the status quo $Q \in [0, \hat{\theta}_{\mathrm{L}}]$, the party's preferences U, and rationally anticipate the set of equilibrium agreements $\Theta^{\mathrm{L}*}$ (Eq. 2).*

(*a*) *If u_{R} is a* ***concave*** *transformation of u_{C}, the intergovernmental policy agreements do not follow the ranking of the party's ideal policies. There exists an atypical electoral equilibrium.*
(*b*) *If u_{R} is a* ***convex*** *transformation of u_{C} (i.e., u_{C} is a* ***concave*** *transformation of u_{R}), the ranking of agreements and party's ideal policies coincide. There exists a typical electoral equilibrium.*

4.5 Combining a translation with a transformation when $A = L$

More general relationships between the parties' utilities can be captured by *combining* a translation with a transformation. Assume u_{R} is a transformation composed with a right translation by $\alpha > 0$ units of u_{C}. Given the results above, the translation and transformation effects may either reinforce or offset each other, depending on the form of preferences and the nature of the transformation. In fact, in the Appendix we show that it is possible to identify when the two effects *exactly* offset each other. This allows us to then determine the conditions under which one effect dominates the other and thus whether the agreements follow the typical or atypical ranking.

5 Conclusion

The implicit assumption made in many political economy models is that the more right leaning a party is, the more right leaning any policy agreement

(see e.g., Saporiti and Tohmé 2006, p. 364). The analysis here highlights the fact that there are many circumstances in which this assumption is not warranted. This is true in an otherwise simple environment—a single election, complete information and no uncertainty, a unidimensional policy, the same set of parties competing in both jurisdictions and formateurs selected by vote shares—the requirement for an intergovernmental agreement means that knowledge of the ideal policies of parties do not provide voters with sufficient information to cast their ballots.

In addition to knowing the parties' payoff functions, voters must also examine the relationship between the parties' utilities before making a decision. In particular, some (implicitly) argue that when the parties' preferences satisfy a single-crossing condition the policy ranking is well-behaved, in the sense of following the ranking of the parties' ideal policies. However, as Cressman and Gallego (2009) and our analysis show even when parties' preferences are translations of each other, the policy ranking may not coincide with the ranking of the parties' ideals. The ranking depends on the ARA coefficient of C's utility function. A more in depth analysis of the parties' utility functions is required where voters must rank the agreements they anticipate will be reached between parties.

Our results show that the IPE Equilibrium that we find depends on the relationship between parties' preferences. Two fundamental characteristics of the parties' utility functions—their ideal policies and the curvature of their utility functions—determine the parties' relative willingness to compromise. It is the relative willingness to compromise of the parties that determines whether a typical or an atypical voting equilibrium emerges in a multi-party unidimensional policy space.

In accordance with Kannai (1977); Kihlstrom et al. (1981) and Roth (1979), we find that intergovernmental negotiations favor the formateur least willing to compromise. However, in a three party setting where parties differ in their willingness to compromise, the ordering of intergovernmental agreements may not follow the ordering of the parties' ideal policies. We have shown that an atypical voting pattern can emerge where the center party's support comes mainly from voters who under conventional ranking align themselves with more extreme parties and where support for one of the extreme parties comes from moderate voters.

The model complements and adds insights to that of Romer and Rosenthal (1978) and Denzau and Mackay (1983). In our model, political competition forces voters to rank agreements. We show that since the ordering depends on the party's willingness to compromise, so do electoral outcomes.

Acknowledgments This project began when Gallego was visiting the Toulouse School of Economics. She thanks the institution for its hospitality and research support. Also acknowledged is the financial support from the Society of Management Accountants of Ontario. We thank J Crémer, T Romer, JF Wen and B Young for helpful comments. We are grateful to M Le Breton, P De Donder and R Cressman for reading an earlier draft and making many valuable suggestions. The usual caveat applies.

Appendix

Combining a translation with a transformation when $A = L$

More general relations between the parties' utilities are captured by *combining* a transformation with a translation. Assume u_{R} is a transformation *composed* with a right translation[17] by $\alpha > 0$ units of u_{C}.

Theorem 1 *Suppose that the conditions of Propositions 2 and 3 hold.*

(i) If u_{C} has IARA for all $Q - \alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$ ***and***
- *(a) $u_{\mathrm{R}}(\theta)$ is a* ***convex*** *transformation of $u_{\mathrm{C}}(\theta)$ (i.e., $u_{\mathrm{C}}(\theta)$ is a concave transformation of $u_{\mathrm{R}}(\theta)$) composed with a right translation, then $a^{\mathrm{LR}*} > a^{\mathrm{LC}*}$. The transformation and translation effects* ***reinforce*** *one another. A typical policy ranking emerges leading to a typical voting equilibrium.*
- *(b) $u_{\mathrm{R}}(\theta)$ is a* ***concave*** *transformation of $u_{\mathrm{C}}(a)$ composed with a right translation, then the transformation and translation effects* ***offset*** *one another so that the net effect on the ranking of agreements is* ***indeterminate*** *and* ***depends*** *on which effect dominates, i.e., depends α and Ψ.*

(ii) If u_{C} has DARA for all $Q - \alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$ ***and***
- (a) *$u_{\mathrm{R}}(\theta)$ is a* ***convex*** *transformation of $u_{\mathrm{C}}(\theta)$ followed by a right translation, then the transformation and translation effects* ***offset*** *one another. The policy agreements may not follow the ranking of the party's ideal policies since the ranking depends on which effect dominates.*
- (b) *$u_{\mathrm{R}}(\theta)$ is a* ***concave*** *transformation of $u_{\mathrm{C}}(\theta)$ followed by a right translation, then $a^{\mathrm{LR}*} < a^{\mathrm{LC}*}$. The transformation and translation effects* ***reinforce*** *one another. Consequently, the atypical ranking of agreements lead to an* ***atypical*** *voting equilibrium.*

We summarize the results of Theorem 1 in Table 3. For the indeterminate case of Theorem 1 (Table 3), we can find when the two effects *exactly* offset each other and when one dominates and determine if the agreements follow the typical or atypical ranking. Since R's utility depends on α and Ψ, to simply notation let R_α and R_Ψ represent R's type when either α or Ψ change in Corollaries 1 and 2 that follow from Theorem 1.

Corollary 1 *Suppose that the conditions of Theorem 1 hold and that u_{C} exhibits IARA on $Q - \alpha \leq \theta \leq \tilde{\theta}_{\mathrm{L}}$.*

(a) For a fixed ***concave*** *transformation T_Ψ there is a unique* ***right*** *translation $T_{\alpha*}$ of $T_\Psi(u_{\mathrm{C}})$ such that the agreement of the $LR_{\alpha*}$ pair coincides with LC's, $a^{\mathrm{LR}*}_{\alpha*} = a^{\mathrm{LC}*}$ for $u_{\mathrm{R}} = T_\Psi \circ T_{\alpha*}(u_{\mathrm{C}})$. Then for any other translation factor α such that $u_{\mathrm{R}} = T_\Psi \circ T_\alpha(u_{\mathrm{C}})$ two voting equilibria are possible. When C's and R's ideal policies are not too far apart, $\alpha < \alpha^*$, the concave transformation dominates the translation effect so that $a^{\mathrm{LR}*}_{\alpha} < a^{\mathrm{LC}*}$. The atypical ranking of agreements leads to an atypical voting equilibrium. When instead their ideals are very distant, $\alpha > \alpha^*$, the translation effect dominates the transformation effect, a typical ranking of agreements and voting equilibrium emerge since $a^{\mathrm{LR}*}_{\alpha} > a^{\mathrm{LC}*}$.*

[17] These operators commute, i.e., $T_\Psi \circ T_\alpha = T_\alpha \circ T_\Psi$, for any transformation T_Ψ and any translation T_α.

Table 3 Summary of Theorem 1: Combining a right translation with a transformation

Assume u_R a …. transformation composed with a right translation of u_C by $\alpha > 0$ and …			
Effects reinforce one another			
…. convex…. and….u_C exhibits *IARA*		…. concave…. and ….u_C exhibits *DARA*	
Typical policy ranking and voting equilibrium emerge		Atypical policy ranking and voting equilibrium emerge	
Effects offset one another			
…. concave…. / ….u_C exhibits *IARA*		…. convex…. / ….u_C exhibits *DARA*	
Translation and transformation effects **offset** one another ⇒ net effect on policies and on the voting equilibrium indeterminate		Translation and transformation effects **offset** one another ⇒ net effect on policies and on the voting equilibrium indeterminate	
Corollary 1a[a]		**Corollary 2a**[a]	
For α^* s.t. $a^{LR*}_{\alpha^*} = a^{LC*}$		For α^* s.t. $a^{LR}_{\alpha^*}* = a^{LC*}$	
$\alpha^* > \alpha$ (small)	$\alpha^* < \alpha$ (large)	$\alpha^* > \alpha$ (small)	$\alpha^* < \alpha$ (large)
Transformation dominates translation	Translation dominates transformation	Transformation dominates translation	Translation dominates transformation
Atypical policy ranking and voting equilibrium	Typical policy ranking and voting equilibrium	Typical policy ranking and voting equilibrium	Atypical policy ranking voting equilibrium
Corollary 1b[b]		**Corollary 2b**[b]	
For Ψ^* s.t. $a^{LR*}_{\Psi^*} = a^{LC*}$		For Ψ^* s.t. $a^{LR*}_{\Psi^*} = a^{LC*}$	
Ψ convex	Ψ concave	Ψ convex	Ψ concave
Transformation dominates translation	Translation dominates transformation	Translation dominates transformation	Transformation dominates translation
Atypical policy ranking and voting equilibrium	Typical policy ranking and voting equilibrium	Typical policy ranking and voting equilibrium	Atypical policy ranking and voting equilibrium

[a] α^* is given in the statement of the Corollaries
[b] Ψ^* is given in the statement of the Corollaries

(b) *For a fixed* ***right*** *translation T_α of u_C, there is a* ***concave*** *transformation $T_{\Psi*}$ of $T_\alpha(u_C)$ such that the agreement of the $LR_{\Psi*}$ pair coincides with LC's, $a^{LR*}_{\Psi*} = a^{LC*}$ for $u_R = T_{\Psi*} \circ T_\alpha(u_C)$. Then, for any other concave transformation T_Ψ where Ψ has the appropriate domain, the agreement a^{LR*}_{Ψ} for $u_R = T_\Psi \circ T_{\Psi*} \circ T_\alpha(u_C)$ satisfies $a^{LR*}_{\Psi} < a^{LC*}$. In this case, the transformation effect dominates the translation effect and an atypical policy ranking and voting equilibrium emerge.*

Corollary 2 *Suppose that the conditions of Theorem 1 hold and that u_C exhibits DARA on $Q - \alpha \leq \theta \leq \tilde{\theta}_L$.*

(a) *For a fixed* ***concave*** *transformation* T_Ψ *there is a unique* ***right*** *translation* $T_{\alpha*}$ *of* $T_\Psi(u_C)$ *such that the agreement of the* $LR_{\alpha*}$ *pair coincides with LC's,* $a_{\alpha*}^{LR*} = a^{LC*}$ *for* $u_R = T_\Psi \circ T_{\alpha*}(u_C)$. *Then for* $\alpha < \alpha*$ *and* $u_R = T_\Psi \circ T_\alpha(u_C)$, *the agreements follow the typical ranking since* $a_\alpha^{LR*} > a^{LC*}$ *and a typical voting equilibrium emerges. Similarly, for* $\alpha > \alpha*$ *and* $u_R = T_\Psi \circ T_\alpha(u_C)$ *we have* $a_\alpha^{LR*} < a^{LC*}$ *implying that an atypical ranking and voting equilibrium materializes.*

(b) *For a fixed* ***right*** *translation* T_α *of* u_C, *there is a* ***concave*** *transformation* T_Ψ *of* $T_\alpha(u_C)$ *such that the* LR_Ψ *and LC pairs agree on the same outcome,* $a_{\Psi*}^{LR*} = a^{LC*}$ *for* $u_R = T_{\Psi*} \circ T_\alpha(u_C)$. *Then, for any other concave transformation* T_Ψ *where* Ψ *has the appropriate domain, for* $u_R = T_\Psi \circ T_{\Psi*} \circ T_\alpha(u_C)$ *the agreement* a_Ψ^{LR*} *satisfies* $a_\Psi^{LR*} < a^{LC*}$. *An atypical policy ranking and voting equilibrium emerge.*

References

Alesina A, Rosenthal H (1995) Partisan politics, divided government, and the economy. Cambridge University Press, Cambridge

Alesina A, Rosenthal H (1996) A theory of divided government. Econometrica 64:1311–1341

Austen-Smith D (2000) Redistributing income under proportional representation. J Polit Econ 108(6): 1235–1269

Austen-Smith D, Banks J (1988) Elections, coalitions, and legislative outcomes. Am Polit Sci Rev 82(2):405–422

Breton A (1996) Competitive governments: an economic theory of politics and public finance. Cambridge University Press, Cambridge

Binmore K, Rubinstein A, Wolinsky A (1986) The nash bargaining solution in economic modeling. Rand J Econ 17:176–188

Cho S, Duggan J (2003) Uniqueness of stationary equilibria in a one-dimensional model of bargaining. J Econ Theory 113(1):118–130

Coate S, Knight B (2007) Socially optimal districting: a theoretical and empirical exploration. Q J Econ 122(4):1409–1471

Crémer J, Palfrey TR (2000) Federal mandates by popular demand. J Polit Econ 108(5):905–927

Crémer J, Palfrey TR (2002) Federal mandates with local agenda setters. Rev Econ Des 7(3):279–296

Crémer J, Palfrey TR (2006) An equilibrium voting model of federal standards with externalities. J Public Econ 90:2091–2106

Cressman R, Gallego M (2009) On the ranking of bilateral bargaining opponents. Math Soc Sci 58(1):64–83

Curry Bill, Chase Steven (2006) Willing to work with NDP, Harper says. Globe and Mail A4, Toronto

Denzau AT, Mackay RJ (1983) Gatekeeping and monopoly power of committees: an analysis of sincere and sophisticated behavior. Am J Polit Sci 27:740–761

Diermeier D, Merlo A (2004) An empirical investigation of coalitional bargaining procedures. J Public Econ 88:783–797

Fiorina M (1996) Divided Government, 2nd Edition. Allyn and Bacon, Boston

Gallagher M, Laver M, Mair P (2005) Representative government in modern europe, 4th edition. McGraw Hill, New York

Greenberg J, Weber S (1985) Multiparty equilibria under proportional representation. Am Polit Sci Rev 79(3):693–703

Kalai E (1977) Proportional solutions to bargaining situations: interpersonal utility comparisons. Econometrica 45:1623–1630

Kalai E, Smorodinsky M (1975) Other solutions to Nash's bargaining problem. Econometrica 43:513–518

Kannai Y (1977) Concavifiability and construction of concave utility functions. J Math Econ 4:1–56

Kedar O (2005) When moderate voters prefer extreme parties: policy balancing in parliamentary elections. Am Polit Sci Rev 99(2):185–199

Kihlstrom RE, Roth A, Schmeidler D (1981) Rish Aversion and Nash's solution to the bargaining problem. In: Moeschlin O, Pallaschke D (eds) Game theory and mathematical economics. Amsterdam, North Holland pp 65–71

Milgrom P, Shannon C (1994) Monotone comparative statics. Econometrica 62(1):157–180

Moravcsik A (1991) Negotiating the single European act: national interests and conventional statecraft in the European community. Int Organ 45:19–56

Moravcsik A (1993) Preferences and power in the European community: a liberal intergovernmentalist approach. J Common Mark Stud 31(4):473–524

Moravcsik A, Nicolaïdis K (1999) Explaining the treaty of Amsterdam: interest, influence and institutions. J Common Mark Stud 37(1):59–85

Nash JF (1950) The bargaining problem. Econometrica 18(2):155–162

Patty JW, Snyder JM Jr, Ting MM (2009) Two's company, three's an equilibrium: strategic voting and multicandidate elections quarterly. J Polit Sci 4(3):251–278

Perles MA, Maschler M (1981) A supper additive solution for the Nash bargaining game. Int J Game Theory 10:163–193

Persson T, Tabellini G (1996) Federal fiscal constitutions: risk sharing and redistribution. J Polit Econ 104(5):979–1009

Peters H (1992) A criterion for comparing strength of preferences with an application to bargaining. Oper Res 40(5):1018–1022

Roth A (1979) Axiomatic models of bargaining. Springer, Berlin and New York

Romer T, Rosenthal H (1978) Political resource allocation, controlled agenda, and the status quo. Public Choice 33:27–44

Saporiti A, Tohmé F (2006) Single-crossing, strategic voting and the median choice rule. Soc Choice Welf 26(2):363–383

Thomson W (1994) Cooperative models of bargaining. In: Auman RJ, Hart S (eds) Handbook of game theory. Elsevier Science, Amsterdam, pp 1237–1284

Thomson W (2009) Bargaining and the theory of cooperative games: John Nash and beyond. Rochester Center for Economic Research Working Paper No. 554, September 2009

Whalley J (2008) Recent regional agreements: why so many, why so much variance in form, why coming so fast, and where are they headed? World Econ 31(4):517–532

Soc Choice Welf
DOI 10.1007/s00355-010-0513-2

ORIGINAL PAPER

A Newton collocation method for solving dynamic bargaining games

John Duggan · Tasos Kalandrakis

Abstract We develop and implement a collocation method to solve for an equilibrium in the dynamic legislative bargaining game of Duggan and Kalandrakis 2008, unpublished manuscript. We formulate the collocation equations in a quasi-discrete version of the model, and we show that they are locally Lipchitz continuous and directionally differentiable. In numerical experiments, we successfully implement a globally convergent variant of Broyden's method on a preconditioned version of the collocation equations, and the method economizes on computation cost by more than 50% compared to the value iteration method. We rely on a continuity property of the equilibrium set to obtain increasingly precise approximations of solutions to the continuum model. We showcase these techniques with an illustration of the dynamic core convergence theorem of Duggan and Kalandrakis 2008, unpublished manuscript in a nine-player, two-dimensional model with negative quadratic preferences.

1 Introduction

To examine strategic incentives in ongoing collective choice problems, we consider a class of dynamic bargaining games in which a sequence of proposals and votes generates policy outcomes over time. The status quo policy evolves endogenously, with today's policy determining tomorrow's status quo and forward-looking players anticipating the future consequences of their decisions. Specifically, we take up the legislative bargaining framework of Duggan and Kalandrakis (2008). In the general

J. Duggan (✉) · T. Kalandrakis
Department of Political Science, University of Rochester, Rochester, NY, USA
e-mail: dugg@troi.cc.rochester.edu

J. Duggan
Department of Economics, University of Rochester, Rochester, NY, USA

Published online: 16 December 2010

model of that article, players possess some (perhaps small) degree of uncertainty about the future state of the game: there is noise in the transition from today's outcome to tomorrow's status quo, and players' preferences are subject to transitory preference shocks each period. The analytic derivation of equilibrium solutions is prohibitive in this framework, owing to the complexity of strategies (which are conditioned on the realized status quo and preference shocks), but the ability to compute equilibria is nevertheless essential for the development of our understanding of this type of dynamic social interaction. For example, dynamic bargaining of the type we analyze can be fruitfully used to model the formation of governments in parliamentary systems, or legislative policy making in Congressional systems of government. In such applications, computed equilibria can provide possibility theorems, and computational results, when used systematically, can also provide the germs of general theorems. Finally, in empirical work, the structural estimation of model parameters or the calibration of a model to observed data, as the case may be, both rely on the ability to compute equilibria.

We propose a method to compute equilibria in the dynamic bargaining framework. We formulate stationary equilibria as solutions to a system of functional equations, the unknowns of which are essentially the future expected utilities (or "dynamic utilities") of the players. We solve these functional equations using a collocation method, a method for solving functional equations that belongs in the general family of projection methods (Judd 1998, Chap. 11). In particular, we posit a finite-dimensional representation of the unknown equilibrium dynamic utility functions as linear combinations of Chebyshev polynomials, and we seek coefficients for these representations that solve the equilibrium functional equations exactly at a finite number of points (that coincide with the roots of the Chebyshev polynomials). This version of the collocation method is theoretically justified by the Chebyshev Interpolation Theorem (Judd 1998; Rivlin 1990) and by the theoretical results of Duggan and Kalandrakis (2008), which guarantee smoothness properties and a priori bounds for the equilibrium expected utilities and their derivatives: we can approximate the unknown functions to an a priori specified, but arbitrary, level of precision by adding higher degree Chebyshev polynomials to represent these functions.

Having transformed the original equilibrium functional equations to a system of a finite number of equations in an equal number of unknown coefficients, we face the practical problem of solving this system of equations. Our computational analysis then focuses on the "quasi-discrete" model, in which the uncertainty in the model is continuous (and so the state space of the bargaining game is infinite), but given the status quo, only a finite number of alternatives may be proposed. This allows us to exactly solve the optimization problem of a proposer by an exhaustive grid search. Since the collocation equations are intermediated by the players' best response behavior, and because they involve the integration over future uncertainty in the model, those equations are nonlinear and potentially ill-behaved. We establish, however, that they are locally Lipschitz continuous and directionally differentiable, i.e., they belong in a class of nonlinear equations for which various generalizations of Newton's method and the convergence properties thereof have been studied in recent years (Pang 1990, 1991; Ip and Kyparisis 1992; Qi 1993; Martinez and Qi 1995; Xu and Chang 1997). Moreover, we show that we can obtain an equilibrium of the model with a continuum

of feasible alternatives by taking the limit of equilibria of a sequence of quasi-discrete approximations.

We employ a version of Broyden's method (Broyden 1965) to solve for an equilibrium. Our implementation of Broyden's method "preconditions" the collocation equations to obtain faster rates of convergence, and we show that this algorithm outperforms two alternatives, a pseudo-Newton method and simple value iteration, in a series of experiments. Finally, we apply the method to illustrate the core convergence result of Duggan and Kalandrakis (2008) by specifying configurations of ideal points approaching the canonical setting in which one alternative belongs to the majority core; for each configuration, we compute the invariant distribution (representing the long run distribution over alternatives) generated by a stationary equilibrium; and we show that these invariant distributions pile mass near the limiting core point. Interestingly, convergence appears to be faster when the players are more patient.

Many situations of interest in political economy possess the structure of a dynamic bargaining game: some player proposes an alternative, that proposal is considered by other players and possibly agreed to, and the game possibly continues into future periods. One branch of this literature considers environments in which bargaining ends once agreement is reached, with play continuing into the future only if a proposal is rejected.[1] We focus, instead, on a class of models in which bargaining continues ad infinitum, whether there is agreement in a period or not. Baron (1996) analyzes the one-dimensional version of the model with single-peaked utilities, Kalandrakis (2004, 2010) studies the canonical divide-the-dollar environment, Cho (2005) considers policy making in a stage game that emulates aspects of parliamentary government, and Battaglini and Coate (2007) characterize stationary equilibria in a model of public good provision and taxation with identical legislators and a stock of public goods that evolves over time. With general stage payoffs and feasible set of alternatives, Duggan and Kalandrakis (2008) assume that stage payoffs and the transition to next period's status quo are subject to (arbitrarily small) shocks, adding uncertainty about the future state of the game. In that article, we establish existence and a number of desirable technical properties of stationary equilibria; we also examine the ergodic properties of equilibria, and we provide a core convergence result for long run equilibrium policy outcomes as the noise in the model goes to zero and the model becomes close to admitting a core alternative.[2]

Little work has been done on computation of equilibrium in this class of bargaining games. Baron and Herron (2003) give a numerical calculation of equilibrium in a three-player, finite-horizon version of the model, and Penn (2009) provides numerical illustrations of her model. Closest to the current article is the work of Duggan et al. (2008), who consider a special case of the model of Duggan and Kalandrakis (2008) to examine the effect of the presidential veto in a US-like political system. But the approach of the current article differs from the former in several respects. First, Duggan et al. (2008) use function approximation instead of function interpolation,

[1] See Rubinstein (1982), Binmore (1987), Baron and Ferejohn (1989), Banks and Duggan (2000), Banks and Duggan (2006), and others.

[2] At a further distance from our article is work on finite-state dynamic voting games, such as Acemoglu et al. (2008) and Diermeier and Fong (2008).

so that in their case equilibrium is not obtained as a solution to a system of collocation equations. Second, they work with a continuous proposal space for the legislators and use continuous optimization methods to solve for legislators' optimal proposals, whereas we implement our techniques in a model where the space of possible proposals at each status quo is finite. Third, those authors use a version of value iteration to obtain an equilibrium. In this article, we consider value iteration as one possible solution method, but we implement and provide theoretical justification for the use of Newton and Newton-like methods. These differences, especially the last two, amount to significant gains on computation time.

In what follows, we first present the model, define our equilibrium concept, and provide background results on the model in Sect. 2. In Sect. 3, we formally describe the collocation method and define the collocation equations. In Sect. 4, we provide theoretical results for the quasi-discrete model, establishing smoothness properties of the collocation equations and our approximation result for the continuum model. In Sect. 5, we describe in detail our implementation of Broyden's method for solving the collocation equations. In Sect. 6, we provide the results of our numerical experiments and our illustration of core convergence. Section 7 concludes, and the appendix contains the proof of our smoothness result.

2 Dynamic bargaining framework

In this section, we first present the bargaining model, and we then define our equilibrium concept, a refinement of stationary Markov perfect equilibrium, and review the known foundational results for the model.

2.1 Bargaining model

We consider a finite set N of players, $i = 1, \ldots, n$, who determine policy over an infinite horizon, with periods indexed $t = 1, 2, \ldots$. Interaction proceeds as follows in each period. A status quo policy $q \in \mathbb{R}^d$ and a vector $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^{nd}$ of preference parameters are realized and publicly observed. A player $i \in N$ is drawn at random, with probabilities $p_1, \ldots, p_n$, and proposes any policy $y \in X \cup \{q\}$, where $X \subseteq \mathbb{R}^d$ represents the set of feasible policies. The players vote simultaneously to accept y or reject it in favor of the status quo q. The proposal passes if a coalition $C \in \mathscr{D}$ of players vote to accept, and it fails otherwise, where $\mathscr{D} \subseteq 2^N \setminus \{\emptyset\}$ is a non-empty collection of *decisive coalitions* satisfying only the minimal monotonicity requirement that if one coalition is decisive, and we add players to that coalition, then the larger coalition is also decisive. Formally, we assume that if $C \in \mathscr{D}$ and $C \subseteq C' \subseteq N$, then $C' \in \mathscr{D}$. The policy outcome for period t, denoted x_t, is y if the proposal passes and is q otherwise. Each player j receives stage utility $u_j(x_t) + \theta_j \cdot x_t$, where $\theta_j \in \mathbb{R}^d$ is a utility shock for the player. Finally, the status quo q' for period $t+1$ is drawn from the density $g(\cdot|x_t)$, a new vector $\theta' = (\theta'_1, \ldots, \theta'_n)$ of preference shocks is drawn from the density $f(\cdot)$, and the above procedure is repeated in period $t+1$. Payoffs in the dynamic game are given by the expected discounted sum of stage utilities, as is standard, and we denote the discount factor of player i by $\delta_i \in [0, 1)$.

We impose a number of regularity conditions on the policy space. We assume that the set of feasible policies, X, is cut out by a finite number of functions $h_\ell : \mathbb{R}^d \to \mathbb{R}$, indexed by $\ell \in K$. We partition K into inequality constraints, K^{in}, and equality constraints, K^{eq}, and we assume that

$$X = \{x \in \mathbb{R}^d : h_\ell(x) \geq 0, \ell \in K^{in}, h_\ell(x) = 0, \ell \in K^{eq}\}.$$

We further assume that X is compact, and that h_ℓ is r-times continuously differentiable for all $\ell \in K$, where $r \geq \max\{2, d\}$.[3] For technical reasons, we impose the weak condition that for all $x \in X$, $\{Dh_\ell(x) : \ell \in \overline{K}(x)\}$ is linearly independent, where $\overline{K}(x)$ is the subset of $\ell \in K$, including equality constraints, such that $h_\ell(x) = 0$. These assumptions allow us to capture quite general manifolds. An important special case is the *quasi-discrete model,* in which the policy space $X \subseteq \mathbb{R}^d$ is finite.[4] Even if the space of interest is a continuum, this special class of model plays an important role in our computational analysis, where we make use of limits of equilibria of quasi-discrete models.

We assume $u_i : \mathbb{R}^d \to \mathbb{R}$ is r-times continuously differentiable. The presence of preference shocks in the model captures uncertainty about the players' future policy preferences. For example, in the important special case of negative quadratic stage utility, where $u_i(x) = -||x_j - x_t||^2$ and x_j is player i's underlying ideal point, the preference shock θ_i is equivalent to a perturbation of the player's ideal point x_i. We assume that the vector $\theta = (\theta_1, \ldots, \theta_n)$ is distributed according to a density f with support contained in the set $\Theta = [\underline{\theta}, \overline{\theta}]^{nd} \subseteq \mathbb{R}^{nd}$, and we let $\tilde{X} \subseteq [\underline{x}, \overline{x}]^d \subseteq \mathbb{R}^d$ be a compact set with $X \subseteq \tilde{X}$ and b_f be a bound such that for all $i \in N$, all $\theta \in \Theta$, and all $x \in \tilde{X}$, we have $|u_i(x) + \theta_i \cdot x| f(\theta) \leq b_f$. The noise on the status quo captures the idea that players are uncertain about the way policy decisions today will be implemented in the future. We assume that the density $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, with values $g(q|x)$, is jointly measurable in (q, x), and that for all x, the support of the density $g(\cdot|x)$ lies in $\tilde{X}$. We do not assume that the support of $g(\cdot|x)$ lies in X, though of course we allow it. Furthermore, we assume a bound b_g such that for all q, we have: $g(q|x)$ is r times continuously differentiable in x; if $r < \infty$, then all derivatives of order $1, \ldots, r$ are bounded in norm by b_f, and the rth derivative of $g(q|x)$ with respect to x is Lipschitz continuous with modulus b_f; and if $r = \infty$, then derivatives of all orders $1, 2 \ldots$ are bounded in norm by b_f.

A strategy in the game consists of two components, one giving the proposals of a player when recognized to propose and the other giving the votes of the player after a proposal is made. While these choices can in principle depend arbitrarily on histories, we seek subgame perfect equilibria in which players use stationary Markov strategies, which we denote $\sigma_i = (\pi_i, \alpha_i)$. Our main focus will be on pure

[3] Of course, we allow $r = \infty$.

[4] We obtain a finite X using suitably "oscillating" equality constraints. We can, for example, isolate a grid on $[0, 1]^d$ by using trigonometric functions, as in $\{x \in \mathbb{R}^d : \sin(2\pi x_i \alpha) = 0, i = 1, \ldots, d\}$, for appropriate α.

strategies.[5] Thus, player i's proposal strategy is a measurable mapping $\pi_i : \mathbb{R}^d \times \Theta \to \mathbb{R}^d$, where $\pi_i(q, \theta)$ is the policy proposed by i given status quo q and utility shocks θ; and player i's voting strategy is a measurable mapping $\alpha_i : \mathbb{R}^d \times \mathbb{R}^d \times \Theta \to \{0, 1\}$, where $\alpha_i(y, q, \theta) = 1$ if i accepts proposal y given status quo q and utility shocks θ and $\alpha_i(y, q, \theta) = 0$ if i rejects. We let $\sigma = (\sigma_1, \ldots, \sigma_n)$ denote a stationary strategy profile. We may equivalently represent voting strategies by the set of feasible proposals a player would vote to accept. We define this *acceptance set* for i as $A_i(q, \theta; \sigma) = \{y \in X \cup \{q\} : \alpha_i(y, q, \theta_i) = 1\}$. Letting C denote a coalition of players, we then define

$$A_C(q, \theta; \sigma) = \bigcap_{i \in C} A_i(q, \theta; \sigma) \quad \text{and} \quad A(q, \theta; \sigma) = \bigcup_{C \in \mathscr{D}} A_C(q, \theta; \sigma)$$

as the *coalitional acceptance set* for C and the *collective acceptance set,* respectively. The latter consists of all policies that would receive the votes of all members of at least one decisive coalition and would, therefore, pass if proposed.

Given strategy profile σ, we define $v_i(x; \sigma)$ as player i's discounted continuation value at the beginning of period $t + 1$ from policy outcome x in period t. We then define i's "policy-specific" dynamic payoff as

$$U_i(x; \sigma) = u_i(x) + \delta_i v_i(x; \sigma). \tag{1}$$

Then, the discounted payoff to player i from implementing policy x in the current period given preference shock θ_i is $U_i(x; \sigma) + \theta_i \cdot x$. We focus on voting strategies that are "deferential," i.e., players vote to accept when indifferent between a proposed policy and the status quo, which allows us to then consider only no-delay equilibria, meaning no player ever proposes a policy that is rejected. (In lieu of that, the player can just as well propose the status quo.) Our measurability assumptions on strategies imply that continuation values are also measurable, and therefore they satisfy

$$v_i(x; \sigma) = \int_q \int_\theta \sum_j p_j [U_i(\pi_j(q, \theta); \sigma) + \theta_i \cdot \pi_j(q, \theta)] f(\theta) g(q|x) \mathrm{d}\theta \mathrm{d}q \tag{2}$$

for all policies x.[6] Note that we can restrict the domain of U_i to the compact set $\tilde{X} \subset \mathbb{R}^d$, as players are restricted to propose in $X \cup \{q\} \subseteq \tilde{X}$ in any period with status quo q, and the distribution of the status quo has support in $\tilde{X}$.

[5] Duggan and Kalandrakis's (2008) Theorem 2 establishes that this is without loss of generality, as any equilibrium in stationary mixed strategies is essentially equivalent to some equilibrium in pure strategies. See Sect. 2.2 for further explanation.

[6] Note that continuation values v_i are "ex ante," in the sense that they are calculated by integrating over q and θ. The dynamic utilities U_i differ from those of Duggan and Kalandrakis (2008) by subtracting out the current period's preference shock. Also note that we do not normalize dynamic payoffs by $(1 - \delta_i)$ as we do in our earlier article.

2.2 Stationary bargaining equilibrium

With this formalism established, we can now define a class of stationary Markov perfect equilibria of special interest. Intuitively, we require that players always propose optimally and that they always vote in their best interest. It is well-known that the latter requirement is unrestrictive in simultaneous voting games, however, as arbitrary outcomes can be supported by Nash equilibria in which no voter is pivotal. To address this difficulty, we follow the standard approach of refining the set of Nash equilibria in voting subgames by requiring that players delete votes that are dominated in the stage game. Thus, we say a strategy profile σ is a *stationary bargaining equilibrium* if the following conditions hold:

- for every status quo q, every shock θ, and every player i, $\pi_i(q, \theta)$ solves

$$\begin{aligned} &\max_y U_i(y; \sigma) + \theta_i \cdot y \\ &\text{s.t. } y \in A(q, \theta; \sigma), \end{aligned} \tag{3}$$

- for every status quo q, every shock θ, every proposal y, and every player i,

$$\alpha_i(y, q, \theta_i) = \begin{cases} 1 & \text{if } U_i(y; \sigma) + \theta_i \cdot y \geq U_i(q; \sigma) + \theta_i \cdot q \\ 0 & \text{else.} \end{cases} \tag{4}$$

Thus, as required by subgame perfection, proposers choose optimally after all histories and the votes of players are, furthermore, consistent with the usual dominance criterion and deferential voting.

Duggan and Kalandrakis (2008) provide foundational results on stationary bargaining equilibria. Chief among those results is the following theorem, which establishes existence and a number of desirable regularity properties of equilibria.[7]

Theorem 1 *There exists a stationary bargaining equilibrium, σ, possessing the following properties.*

1. *Continuation values are differentiable: for every player i, $v_i(x; \sigma)$ is r-times continuously differentiable as a function of x.*
2. *Proposals are almost always strictly best: for every status quo q, almost all shocks θ, every player i, and every $y \in A(q, \theta; \sigma)$ distinct from the proposal $\pi_i(q, \theta)$, we have $U_i(\pi_i(q, \theta); \sigma) + \theta_i \cdot \pi_i(q, \theta) > U_i(y; \sigma) + \theta_i \cdot y$.*
3. *Proposal strategies are almost always continuously differentiable: for every status quo q, almost all shocks θ, and every player i such that $\pi_i(q, \theta) \neq q$, $\pi_i(q, \theta)$ is continuously differentiable in an open set around (q, θ).*
4. *Binding voters, if any, are almost never redundant: for every status quo q, almost all shocks θ, and every player i, if $\pi_i(q, \theta) \neq q$ and there exists j such that*

[7] Duggan and Kalandrakis (2008) use the term "pure stationary legislative equilibrium" for the concept we consider here. They state only continuity of equilibrium proposal strategies, rather than the condition of continuous differentiability stated in part 3. See the working paper version, Duggan and Kalandrakis (2007), for the statement and proof of the stronger result we give here.

$$U_j(\pi_i(q,\theta);\sigma) + \theta_j \cdot \pi_i(q,\theta) = U_j(q;\sigma) + \theta_j \cdot q, \textit{ then}$$

$$\{\ell \in N : U_\ell(\pi_i(q,\theta);\sigma) + \theta_\ell \cdot \pi_i(q,\theta) \geq U_\ell(q;\sigma) + \theta_\ell \cdot q\} \backslash \{j\} \notin \mathscr{D}.$$

Part 1 of Theorem 1 implies that dynamic utilities, $U_i(x;\sigma) = u_i(x) + \delta_i v_i(x;\sigma)$, are also r times differentiable, allowing us in principle to employ first order conditions to characterize optimal proposals. Optimal proposals are essentially strict, by part 2, and equilibrium proposal strategies are continuously differentiable almost everywhere, by part 3, permitting, in principle, the application of calculus techniques in computing comparative statics. Finally, part 4 of the theorem informs us that, generically, if a voter is indifferent between a proposal and the status quo, then the player is pivotal, in the sense that the remaining players willing to vote for the proposal are not decisive, i.e., the coalition of players who accept the proposal is minimally decisive. While Theorem 1 holds in the general model, its implications for the quasi-discrete model, in which the set X of alternatives is finite, take a simple form. There, dynamic utilities are almost always injective: for every status quo q, almost all shocks θ, every player i, and all distinct $x, y \in X \cup \{q\}$, we have $U_i(x;\sigma) + \theta_i \cdot x \neq U_i(y;\sigma) + \theta_i \cdot y$. Thus, the conditions in parts 2 and 4 are trivially satisfied. Finally, part 3 can be strengthened: equilibrium proposal strategies are such that for every q, almost all θ, and all i with $\pi_i(q,\theta) \neq q$, π_i is in fact constant in an open set around (q,θ).

Theorem 2 of Duggan and Kalandrakis (2008) considers the possibility that proposers mix over optimal proposals and that voters mix when indifferent between a proposal and the status quo. The result establishes that when we broaden our definition of equilibrium in this way, we do not introduce new equilibrium behavior in any meaningful sense: every mixed strategy equilibrium is equivalent (up to a measure-zero set of status quos and preference shocks) to a stationary bargaining equilibrium as defined above. Moreover, *every* stationary bargaining equilibrium satisfies the properties in parts 1–4 of Theorem 1. This allows us to focus on pure strategies without loss of generality and increases the scope for computation of equilibrium. In the earlier article, we also show that the correspondence of stationary bargaining equilibria has closed graph, a desirable continuity property that facilitates our analysis of quasi-discrete approximations of the continuum model in Sect. 4.2; and we provide general conditions under which equilibrium strategies induce an invariant distribution over the policy space that represents the long run policy outcomes of the system.

Finally, Duggan and Kalandrakis (2008) prove a core convergence theorem for models "close" to the canonical social choice model, in which players have quadratic stage utilities, the voting rule is strong (if a coalition is not decisive, then its complement is), there is a unique core policy, and the player located at the core has positive probability of proposing. Theorem 6 of that article establishes that in this case, the equilibrium invariant distributions collapse to the point mass on the limiting core policy. We take up this result in Sect. 6.3, where we demonstrate the numerical methods developed in this article and depict convergence to the core in a two-dimensional, majority-rule version of the bargaining model.

3 The collocation method

In this section, we develop a collocation method to solve for an equilibrium of the bargaining game. In particular, in Sect. 3.1, we formulate the problem of finding an equilibrium as the solution of a system of functional equations. In Sect. 3.2, we describe how we approximate these functional equations with a finite-dimensional system of equations, the collocation equations. Before we attempt to solve the collocation equations, we must be able to numerically evaluate them, and we discuss issues related to the numerical evaluation of the collocation equations in Sect. 3.3.

3.1 Equilibrium functional equations

We start with the observation that an equilibrium is fully characterized by the corresponding policy-specific dynamic payoffs $U_1, \ldots, U_n$ of the n players. Indeed, given $U = (U_1, \ldots, U_n)$, define

$$A(q, \theta; U) = \bigcup_{C \in \mathscr{D}} \left[\bigcap_{i \in C} A_i(q, \theta; U_i) \right],$$

where

$$A_i(q, \theta; U_i) = \{y \in X \cup \{q\} : U_i(y) + \theta_i \cdot y \geq U_i(q) + \theta_i \cdot q\} \tag{5}$$

gives the acceptance set of player i after eliminating stage-dominated voting strategies given dynamic payoffs U_i. Furthermore, define the policy $\pi_i(q, \theta; U)$ to be the player's optimal proposal given dynamic payoff U_i and voting behavior in (5), i.e., it solves

$$\begin{aligned} &\max_y U_i(y) + \theta_i \cdot y \\ &\text{s.t. } y \in A(q, \theta; U). \end{aligned} \tag{6}$$

By an application of Theorem I.3.1 in Mas-Colell (1985), the above optimization problem has a unique solution for almost all shocks θ_i, pinning down the optimal proposals $\pi_i(q, \theta; U)$ almost everywhere. Then, we can express equilibrium dynamic payoffs as solutions to the functional equation

$$\begin{aligned} U_i(x) = u_i(x) + \delta_i \int_q \int_\theta \sum_{h \in N} p_h \left[U_i(\pi_h(q, \theta; U)) \right. \\ \left. + \theta_i \cdot \pi_h(q, \theta; U) \right] f(\theta) g(q|x) \mathrm{d}\theta \mathrm{d}q, \end{aligned} \tag{7}$$

and we can focus our search for an equilibrium on computing equilibrium dynamic payoff functions $U = (U_1, \ldots, U_n)$ solving the functional Eq. 7.

Note that since both the stage utility functions u_i and, by property 1 of Theorem 1, the continuation values v_i are r times continuously differentiable with derivatives that

are uniformly bounded in $\tilde{X}$, we conclude that any solution U to the functional Eq. 7 must belong in the space of r times continuously differentiable functions $C^r(\tilde{X}, \mathbb{R}^n)$. Furthermore, Lemma 5 of Duggan and Kalandrakis (2008) establishes that the solutions to (7) must lie in an a priori specified compact space with derivatives satisfying a uniform bound across $\tilde{X}$.

3.2 Collocation equations

The numerical solution of the functional Eq. 7 on a computer must necessarily involve a representation of the infinite-dimensional objects U in finite dimensions. A standard approach is to proceed by choosing a finite-dimensional subspace of the function space of the candidate solutions U, and then restricting the search for approximate solutions to this subspace. We choose a subspace generated by a finite set of m Chebyshev polynomials,

$$\{T_1, \ldots, T_m\}.$$

Specifically, we specify a number m_ℓ of the univariate Chebyshev polynomials of degree 0 through $m_\ell - 1$ for each dimension $\ell = 1, \ldots, d$, and then we obtain the basis $\{T_1, \ldots, T_m\}$ of $m = \prod_{\ell=1}^{d} m_\ell$ polynomials using tensor products of these univariate polynomials. Once the basis $\{T_1, \ldots, T_m\}$ is fixed, we seek solutions for the expected payoff functions U_i that take the form

$$U(x; c_i) = \sum_{j=1}^{m} c_{i,j} T_j(x), \tag{8}$$

where $c_i = (c_{i,1,}, \ldots, c_{i,m}) \in \mathbb{R}^m$ is a vector of *collocation coefficients* corresponding to player i. We write $c = (c_1, \ldots, c_n) \in \mathbb{R}^{nm}$ for a vector specifying the coefficients of all players.

The choice of the Chebyshev polynomial basis is appealing for a number of reasons, including the fact, noted above, that any solutions U to (7) are continuously differentiable with bounded derivatives. Nevertheless, it is unlikely that the actual solutions reside in the subspace spanned by the choosen basis for any finite m. Thus, instead of satisfying Eq. 7 for all $x \in \tilde{X}$, the collocation method ensures that these equations are satisfied finitely many times, specifically at a finite number of judiciously chosen points in the domain of U_i. Thus, we choose m *collocation nodes*

$$\{\tilde{x}_1, \ldots, \tilde{x}_m\} \subset \tilde{X},$$

and we seek to find collocation coefficients $c = (c_1, \ldots, c_n) \in \mathbb{R}^{nm}$ for the n players so that for every player $i = 1, \ldots, n$, the following collocation equations are satisfied at each of the m collocation nodes $\tilde{x}_k$, $k = 1, \ldots, m$:

$$\begin{aligned} U(\tilde{x}_k; c_i) = u_i(\tilde{x}_k) + \delta_i \int_q \int_\theta \sum_{h \in N} p_h \left[U(\pi_h(q, \theta; c); c_i) \right. \\ \left. + \theta_i \cdot \pi_h(q, \theta; c) \right] f(\theta) g(q|\tilde{x}_k) \mathrm{d}\theta \mathrm{d}q. \end{aligned} \tag{9}$$

The function $\pi_h(q, \theta; c)$ in (9) is identical to the solution $\pi_h(q, \theta; U)$ of the optimization problem in (6) when $U = (U(\cdot; c_1), \ldots, U(\cdot; c_n))$. Given our choice of the Chebyshev polynomial basis, there is an elegant theory that dictates the location of the collocation nodes $\tilde{x}_k$ at the roots of the Chebyshev polynomials.[8] Since any solution U belongs in $C^r(\tilde{X}, \mathbb{R}^n)$, and since we have an a priori established bound on the derivatives of U, the Chebyshev Interpolation Theorem guarantees that we can approximate the function U up to arbitrary precision, using the combination of the Chebyshev polynomial basis and the roots of the corresponding Chebyshev polynomials as collocation nodes by increasing the degree of approximation and the corresponding number of collocation nodes Judd (1998).

We have thus reduced the problem of solving the functional Eq. 7 for equilibrium expected payoffs U to solving the nm collocation equations in (9) for the nm collocation coefficients $c \in \mathbb{R}^{nm}$. More compactly, we seek a solution to the equations $F(c) = 0$, where $F : \mathbb{R}^{nm} \to \mathbb{R}^{nm}$ is a function defined from the collocation Eq. 9 as

$$F_{i,k}(c) = U(\tilde{x}_k; c_i) - \Bigg[u_i(\tilde{x}_k) + \delta_i \int_q \int_\theta \sum_{h \in N} p_h \left[U(\pi_h(q, \theta; c); c_i) + \theta_i \cdot \pi_h(q, \theta; c) \right] f(\theta) g(q|\tilde{x}_k) \mathrm{d}\theta \mathrm{d}q \Bigg], \tag{10}$$

and where the index i, k corresponds to the kth collocation node, $\tilde{x}_k$, and player i. In the next subsection, we turn to the practical question of evaluating this collocation function.

3.3 Collocation function evaluation

The evaluation of the collocation function F requires us to tackle two computational issues. First, we must be able to evaluate the integrals with respect to the status quo q and the preference shocks θ. The integral with respect to the status quo q is d-dimensional. Thus, when the dimension of the policy space d is of small or moderate size, we can extend unidimensional quadrature techniques to perform this integration. In particular, in the applications we consider, we assume that each coordinate of the status quo is drawn independently, so that the density $g(q|x)$ is a product of densities. Thus, we can use Gaussian quadrature along each dimension with weight function given by the density of the coordinate of the status quo that corresponds to this dimension. The required d-dimensional nodes and weights are easily obtained from the unidimensional ones using tensor products (Judd 1998; Miranda and Fackler 2002). In practice, for each collocation node $\tilde{x}_k$, we specify a total of α quadrature nodes $q_{k,j} \in Q_k$ and corresponding weights $\omega_{k,j}$, $j = 1, \ldots, \alpha$. These are obtained

[8] We obtain these nodes by tensor products of the roots of the univariate bases. Judd (1998) refers to this version of the collocation method in which the collocation nodes $\tilde{x}_k$ coincide with the roots of the polynomial basis as *orthogonal collocation*.

using Gaussian quadrature with $g(q|\tilde{x}_k)$ as the weight function. Then we compute

$$\int_q \Phi(q,c)g(q|\tilde{x}_k)dq \approx \sum_{q_{k,j}\in Q_k} \Phi(q_{k,j},c)\omega_{k,j},$$

where $\Phi(q,c) = \int_\theta \sum_{h\in N} p_h\left[U(\pi_h(q,\theta;c);c_i) + \theta_i \cdot \pi_h(q,\theta;c)\right] f(\theta)\mathrm{d}\theta$.

Integration with respect to the preference shocks θ is more challenging, as the associated integral is nd-dimensional. The Guassian quadrature approach we described above would be impractical in this case, as the required number of quadrature nodes becomes prohibitive. Instead, we switch to a quasi-Monte Carlo integration method using a Sobol sequence of β quasi-random numbers (Press et al. 1992), $\theta^\ell \in \Theta$, $\ell = 1,\ldots,\beta$, so that for every quadrature node $q_{k,j}$, we compute

$$\int_\theta \phi(\theta,q_{k,j},c)d\theta \approx \beta^{-1}\sum_{\ell=1}^{\beta}\phi(\theta^\ell,q_{k,j},c),$$

where $\phi(\theta,q,c) = \sum_{h\in N} p_h\left[U(\pi_h(q,\theta;c);c_i) + \theta_i \cdot \pi_h(q,\theta;c)\right]$.

In addition to the above integrations, the second major numerical issue involved in evaluating the collocation function F is the computation of the optimal proposals $\pi_h(q,\theta;c)$. This entails solving the optimization problem of player h for each collocation node $\tilde{x}_k$, each quadrature node $q_{k,j}$, and each θ^ℓ in the Sobol sequence, i.e., we must solve a total of $n\times m\times\alpha\times\beta$ optimization problems in order to evaluate the collocation function F once. Note that for arbitrary values of the collocation coefficients c, the players' dynamic expected utility functions need not be concave, nor is there any guarantee that these functions would be concave in equilibrium. Even if these functions were concave, the feasible set of proposals available to player h is not convex, as is obvious from (5), since it is the union of sets of proposals acceptable to each coalition. Thus, these optimization problems are quite challenging and entail the possibility of significant numerical error in the evaluation of F if, for example, a local maximum is computed instead of a global maximum. When the space of policy proposals X is a continuum, Duggan et al. (2008) employ a *Nelder–Mead* maximization algorithm to compute $\pi_h(q,\theta;c)$ after initially approximating the solution via a grid search to safeguard against the possibility of locating a local, as opposed to global, maximizer. In this article, we take an alternative approach by assuming that the set of feasible proposals X is finite, i.e., we work in the quasi-discrete model, where we can compute optimal proposals without error by straightforward exhaustive search. Furthermore, the computation times reported in Sect. 6.2 suggest we can perform the players' maximization relatively efficiently, even when the feasible set X comprises a large number of points.

In the next section, we show that our use of the quasi-discrete model is justified on at least two more grounds. First, the collocation function in the quasi-discrete model is sufficiently smooth to allow us to use Newton-like methods in order to compute

equilibria. Second, we show that we can recover equilibria of the continuous model by computing equilibria for a sequence of quasi-discrete models.

4 The quasi-discrete model

In this section, we specialize the collocation method to the quasi-discrete model. In particular, we assume that the set of feasible proposals is a finite set, $|X| < \infty$, and we index a typical element as x_p. As discussed in the previous section, the main advantage of this formulation is that we are able to accurately evaluate the optimization problem of proposers, thus providing a realistic setting for the numerical evaluation of the collocation equations. It is key for the tractability of this problem, however, that the collocation equations be reasonably smooth. This precondition is no mere formality, for the definition of the collocation equations in (10) involves integration over optimal policy choices, $\pi_h(q, \theta; c)$, which are not even continuous in c. Nevertheless, we establish a sufficient level of smoothness for the applicability of Newton-like or inexact Newton methods: in particular, the collocation functions are locally Lipschitz and directionally differentiable. We then show that the solution of the quasi-discrete model allows us to compute equilibria of the continuous model as the limit of equilibria of a sequence of quasi-discrete approximations.

4.1 Smoothness of the collocation equations

Before proceeding to the result of this subsection, we provide some elementary definitions and facts used in the proof. We say that a function $g : \mathbb{R}^n \to \mathbb{R}^n$ is *locally Lipschitz continuous* if for every $x \in \mathbb{R}^n$, there exist $\epsilon > 0$ and a constant $M \in \mathbb{R}$ such that for all $y, z \in B_\epsilon(x)$,

$$||g(y) - g(z)|| \leq M||y - z||.$$

A locally Lipschitz function g is differentiable almost everywhere (by Rademacher's theorem). The *directional derivative* of g in direction s is

$$g'(x|s) = \lim_{\alpha \to 0^+} \frac{g(x + \alpha s) - g(x)}{\alpha},$$

provided this limit is well-defined. The smoothness properties of the collocation function rely on the properties of functions that can be represented as a continuous splicing of a finite number of continuously differentiable functions. A function $g : \mathbb{R}^n \to \mathbb{R}^n$ is *piecewise smooth* if it is continuous and there exists a finite collection of continuously differentiable functions $\{g_1, \ldots, g_m\}$ such that each g_j is defined on an open domain and for all $x \in \mathbb{R}^n$, there is a $j = 1, \ldots, m$ such that $g(x) = g_j(x)$.[9] It is known that piecewise smooth functions are locally Lipschitz and have directional derivatives (Kuntz and Scholtes 1994).

[9] Here, we use the generalized definition of piecewise smoothness suggested by Kuntz and Scholtes (1995).

The following theorem establishes that the collocation function F is sufficiently smooth in c in order to allow us to pursue Newton-like methods for the solution of nonlinear systems of equations. If further properties hold in a neighborhood of a solution to the equations, then convergence will also be fast.[10]

Theorem 2 *Assume X is finite and f is C^1. The collocation function $F : \mathbb{R}^{nm} \to \mathbb{R}^{nm}$ is locally Lipschitz continuous and directionally differentiable.*

The proof, located in the appendix, establishes that the collocation equations are integrals of piecewise smooth functions and that the properties of these piecewise smooth functions carry over to F. Specifically, given any status quo q, we break the integral

$$\int_{\theta} \sum_{h \in N} p_h [U(\pi_h(q, \theta; c); c_i) + \theta_i \cdot \pi_h(q, \theta; c)] f(\theta) \mathrm{d}\theta \tag{11}$$

into a finite number of integrals over polyhedral subsets of preference shocks. Each subset $\Theta(q, y, h, A_{-h}; c)$ corresponds to the set of preference shocks $\theta = (\theta_1, \ldots, \theta_n)$ such that player h's optimal proposal is y and the acceptance sets of the other players are given by A_{-h} when the status quo is q. As such, each "cell" is defined by a finite number of linear inequalities in which the collocation coefficients c enter the righthand side. The integral

$$\int_{\Theta(q, y, h, A_{-h}; c)} [U(y; c_i) + \theta_i \cdot y] f(\theta) \mathrm{d}\theta \tag{12}$$

over a particular cell is clearly continuous, but non-differentiabilities may arise because as we vary c, different sets of inequalities may become binding. In Fig. 1, we depict a change from c to c' to c'', supposing for simplicity that this variation only affects the upper most linear constraint. This change has a smooth effect on (12) until c', where the upper most constraint becomes binding, and the effect is smooth thereafter.

While our decomposition of (11) suppresses the dependence of optimal proposals on the collocation coefficients (through the term $\pi_h(q, \theta; c)$), we cannot avoid non-differentiabilities that may be inherent in the structure of equilibrium. Nevertheless, we show that the integral (12) over any cell is piecewise smooth, so that (11), rewritten as the sum

$$\sum_h p_h \sum_{A_{-h}} \sum_{y \in X \cup \{q\}} \int_{\Theta(q, y, h, A_{-h}; c)} [U(y; c_i) + \theta_i \cdot y] f(\theta) \mathrm{d}\theta$$

over all possible proposers h, acceptance sets of other players A_{-h}, and possible proposals y, is also piecewise smooth. Piecewise smoothness of (11) implies that for a

[10] If c^* solves $F(c) = 0$ and F has continuous directional derivatives in an open set around c^* with non-singular Jacobian, J^*, at c^*, it then follows from Theorem 4.1 of Ip and Kyparisis (1992) that Broyden's method (see Sect. 5.1) converges super-linearly to a solution in a neighborhood of (c^*, J^*).

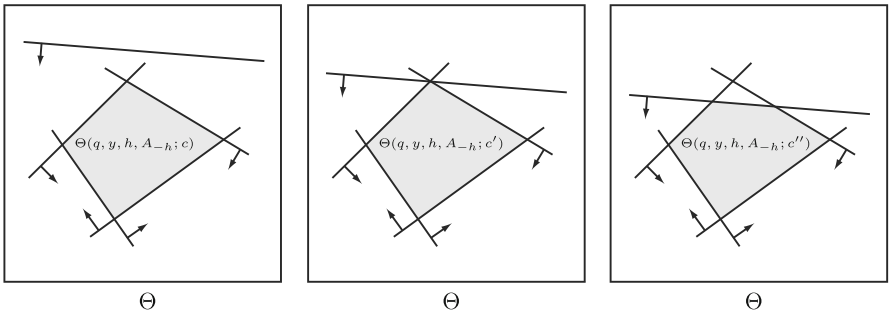

Fig. 1 Non-differentiability in collocation coefficients

given status quo q, it is locally Lipschitz in c, but in order to integrate this term over q, we need some bound on its behavior across status quos. The final hurdle in the proof is to construct, for every c, a local Lipschitz constant for (11) that is uniform over q, allowing us to conclude that the integral of (11) over status quos q, and, therefore, the collocation equations F, is locally Lipschitz and directionally differentiable.

4.2 Quasi-discrete approximation

In this section, we show that while the quasi-discrete model affords computational tractability, it also provides a means for computation of equilibrium in the continuum model. In particular, we can recover an equilibrium of the continuum model via a sequence of computed equilibria for quasi-discrete models. Given a model with a continuum X of feasible policies, we consider an algorithm for computing stationary bargaining equilibria by means of an increasing sequence of finite grids on the policy space. To be concrete, let $\{X^\ell\}$ be a sequence of finite approximations converging to X in the Hausdorff metric. For each ℓ, define a corresponding "quasi-discrete" model that is identical to the original model except for the fact that feasible proposals (but not the status quo) are now constrained to lie in X^ℓ. The quasi-discrete model is a special case of our bargaining model, for as mentioned above, we can obtain the finite set X^ℓ of feasible policies by appropriately specifying equality constraints. Therefore, Theorem 1 yields at least one stationary bargaining equilibrium, with policy-specific dynamic payoffs $U^\ell = (U_1^\ell, \ldots, U_n^\ell)$, in each quasi-discrete model.

We now establish that the sequence $\{U^\ell\}$ necessarily admits a convergent subsequence, and that the limit of any such subsequence corresponds to a stationary bargaining equilibrium of the continuum model. The notion of convergence we use is the topology of C^r-uniform convergence on compacta, a fairly strong topology (which entails a correspondingly strong convergence result).[11]

[11] To describe this topology, let $\hat{r}$ be a natural number and $Y \subseteq \mathbb{R}^d$, and define the norm $||f||_{\hat{r},Y}$ on $C^{\hat{r}}(\mathbb{R}^d, \mathbb{R}^n)$ as $\sup\{||\partial f(x)|| : x \in Y\}$, where ∂f is the $\hat{r}$th derivative of f. Then a sequence $\{f^m\}$ of functions converges to f in $C^r(\mathbb{R}^d, \mathbb{R}^n)$ if and only if for every $\hat{r} = 0, 1, \ldots, r$ and every compact set $Y \subseteq \mathbb{R}^d$, we have $||f^m - f||_{\hat{r},Y} \to 0$. We say $f^m \to f$ in $C^\infty(\mathbb{R}^d, \mathbb{R}^n)$ if and only if it converges in $C^r(\mathbb{R}^d, \mathbb{R}^n)$ for all $r = 0, 1, 2, \ldots$

Theorem 3 *Given a model with set X of alternatives, let $\{X^\ell\}$ be a sequence of quasi-discrete models $X^\ell \subseteq \tilde{X}$ converging to X in the Hausdorff metric. Then in each quasi-discrete model X^ℓ, there exists a stationary bargaining equilibrium with policy-specific dynamic payoffs $U^\ell = (U_1^\ell, \ldots, U_n^\ell)$; and for every sequence $\{U^\ell\}$ of equilibrium dynamic payoffs of quasi-discrete models, there is an accumulation point U of $\{U^\ell\}$ and a stationary bargaining equilibrium of the model with set X alternatives with dynamic payoffs U.*

Proof For each quasi-discrete model X^ℓ, existence of a stationary bargaining equilibrium σ^ℓ follows directly from Theorem 2, and we let U^ℓ be the policy-specific dynamic payoff generated by σ^ℓ as in (1) and (2). Now consider any such sequence $\{U^\ell\}$ of equilibrium dynamic payoffs corresponding to the sequence $\{X^\ell\}$ of quasi-discrete models. For each l, let v^ℓ be as in (1) and (2). Letting b_h denote the Lebesgue measure of $\tilde{X}$, Duggan and Kalandrakis (2008) define a subset $\mathcal{V} \subseteq C^r(\mathbb{R}^d, \mathbb{R}^n)$ that consists of $v\colon \mathbb{R}^d \to \mathbb{R}^n$ satisfying the following: (i) if $r < \infty$, then the derivatives of v of order $0, 1, 2, \ldots$ are bounded in norm by $\sqrt{n} b_f b_g b_h$, and the rth derivative of v is Lipschitz continuous with modulus $\sqrt{n} b_f b_g b_h$; and (ii) if $r = \infty$, then the derivatives of v of all orders are bounded in norm by $\sqrt{n} b_f b_g b_h$. Their Lemma 5 establishes that $\mathcal{V}$ is non-empty and compact in the topology of C^r-uniform convergence on compacta, and it shows that the equilibrium continuation values v^ℓ belong to $\mathcal{V}$ for all l. Thus, there is a convergent subsequence, still indexed by l for simplicity, such that $v^\ell \to v$ with limit $v \in \mathcal{V}$. Theorem 3 of Duggan and Kalandrakis (2008) establishes closed graph of the stationary bargaining equilibrium correspondence, and it follows that there is a stationary bargaining equilibrium, say σ, with continuation value $v(\cdot; \sigma) = v$. Defining U by $U_i(x; \sigma) = u_i(x) + \delta_i v_i(x)$, we have

$$U_i^\ell = u_i + \delta_i v_i^\ell \;\; \to \;\; u_i + \delta_i v_i \;\; = \;\; U_i,$$

as required. □

We have framed Theorem 3 in relatively simple terms for computational purposes, fixing stage utilities, discount factors, and other parameters as the policy space becomes finer. Since the upper hemicontinuity result of Duggan and Kalandrakis (2008) is general in these respects, however, it is straightforward to extend the proof of the above result to allow these other parameters to vary.

5 Solving the collocation equations

The collocation equations $F(c) = 0$ are nonlinear in the unknown collocation coefficients c, and hence a natural strategy to solve these equations is to apply a variant of Newton's method. Due to the smoothness properties of the collocation function established in Theorem 2, several generalizations of Newton's method are in principle applicable in our problem. In Sect. 5.1, we discuss one such method, Broyden's method, that offers several advantages compared to other alternatives. In Sect. 5.2, we discuss implementation issues and present a transformation of the collocation equations that significantly improves the performance of the resulting algorithm. In

Sect. 5.3, we briefly discuss two additional algorithms for the solution of the collocation equations.

5.1 Broyden's method

Recall that in Newton's method we start with an initial candidate c^0 for the solution and then at the $(\tau + 1)$th iteration, we obtain a new candidate solution according to the formula

$$c^{\tau+1} = c^{\tau} - [F'(c^{\tau})]^{-1} F(c^{\tau}),$$

where $F'(c^{\tau})$ is the Jacobian of F evaluated at the current iterate c^{τ}. Recently, a number of authors have studied modifications of Newton's method when F is not everywhere differentiable but satisfies weaker smoothness properties such as local Lipschitz continuity and directional differentiability, as is the case in our problem. One class of these methods directly generalizes Newton's method by using a generalized Jacobian in lieu of $F'(c^{\tau})$ when F is not differentiable at c^{τ} (e.g., Pang (1990); Qi and Sun (1993)). A second alternative falls within the broad class of inexact or quasi-Newton methods (e.g., Ip and Kyparisis 1992; Martinez and Qi 1995; Qi 1997). Primarily motivated by the observation that the evaluation of the Jacobian is typically very costly (even when it exists), these methods operate by providing an initial estimate of the Jacobian and then ensuring an inexact (but easy to compute) update to the Jacobian with each iteration. These methods behave quite well as long as the sequence of the updated Jacobians stays within a certain distance from the Jacobian at the solution. Any method that circumvents the need to compute the Jacobian is particularly appealing in our problem as the Jacobian of the collocation equations is very costly to compute analytically or numerically. We choose to work with a particularly robust version in the class of inexact Newton methods, namely Broyden's method (Broyden 1965).

Broyden's method requires an initial guess of the solution c^0, just as Newton's method does. In addition, at the beginning of the algorithm we must also supply an initial guess, B^0, of the Jacobian at the solution. At the τ-th iteration, we obtain a new update of the unknown collocation coefficients as in Newton's method,

$$c^{\tau+1} = c^{\tau} - [B^{\tau}]^{-1} F(c^{\tau}), \tag{13}$$

and a new update of the Jacobian according to the formula

$$\begin{aligned} B^{\tau+1} &= B^{\tau} + \frac{(F(c^{\tau+1}) - F(c^{\tau}) - B^{\tau}(c^{\tau+1} - c^{\tau}))(c^{\tau+1} - c^{\tau})^T}{(c^{\tau+1} - c^{\tau})^T (c^{\tau+1} - c^{\tau})} \\ &= B^{\tau} + \frac{F(c^{\tau+1})(c^{\tau+1} - c^{\tau})^T}{(c^{\tau+1} - c^{\tau})^T (c^{\tau+1} - c^{\tau})}, \end{aligned} \tag{14}$$

where $(c^{\tau+1} - c^{\tau})^T$ denotes the transpose of $c^{\tau+1} - c^{\tau}$. The conditions for fast convergence of Broyden's method, described following Theorem 2, are analogous to those needed for convergence of Newton's method, but in addition to requiring that

the initial guess c^0 must be sufficiently close to the solution, Broyden's method also requires that the initial approximation to the Jacobian, B^0, is sufficiently close to the Jacobian at the solution. We address the non-trivial problem of supplying an accurate initial approximation B^0 for our implementation of Broyden's method in the next subsection.

In practice, both Newton and quasi-Newton algorithms are modified to prevent oscillatory behavior or divergence of the iteration sequence. A focal point of intervention in these algorithms is a modification of the Newton step in cases when that step does not lead to a decrease in the residual, $||F(c^{\tau+1})|| \leq ||F(c^{\tau})||$. If the default (quasi-) Newton step does not produce a sufficient decrease in the residual norm, then the step is adjusted by performing a line search along the originally suggested direction until a decrease is achieved. A popular globalization strategy of this form is the Armijo rule (Armijo 1966), which we use in our implementation. In the particular version we use, the required line search along the direction suggested by (13) is performed using optimization techniques on a parabolic approximation of the residual function. Details of the Armijo rule and the particular implementation can be found in Kelley (2003).

5.2 Implementation and preconditioning

As we already discussed, Broyden's method requires the analyst to supply an initial estimate B^0 of the Jacobian, and the performance or even eventual convergence of the algorithm hinges on the quality of this initial estimate. Assuming the initial iterate c^0 is close to the solution, a good initial value for B^0 can be obtained by computing the actual Jacobian at c^0.[12] In principle, this can be done numerically, but aside from the possibility of numerical error, the numerical evaluation of the Jacobian is impractical in our case as it requires multiple evaluations of F, which are very expensive to perform. A superior alternative is the analytic evaluation of the Jacobian, but this is also prohibitively costly in our problem. In particular, in order to compute the partial derivative $\frac{\partial F_{i,k}(c)}{\partial c_{j,\ell}}$, we must account for the indirect effect of a change in the collocation coefficient $c_{j,\ell}$ on the proposal strategies $\pi_h(\theta, q; c)$.

Nevertheless, a strategy for a second-best approximation to the Jacobian is available analytically. This approximation evaluates the derivative by ignoring any effect of changes in the collocation coefficients that is channeled through changes in proposal strategies. Evaluating only the direct effects, it is immediate from (8) and (10) that the pseudo-derivative resulting from this approach, denoted $\psi_{(i,k),(j,\ell)}(c) \approx \frac{\partial F_{i,k}(c)}{\partial c_{j,\ell}}$, takes the form

$$\psi_{(i,k),(j,\ell)}(c) = \begin{cases} T_\ell(\tilde{x}_k) - \delta_i \int\limits_q \int\limits_\theta \sum\limits_{h \in N} p_h T_\ell(\pi_h(q, \theta; c)) & \text{if} \quad i = j \\ \quad f(\theta) \mathrm{d}\theta g(q|\tilde{x}_k) \mathrm{d}q & \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

[12] Recall that the collocation function F of the quasi-discrete model is differentiable almost everywhere by Theorem 2.

Let $\Psi(c)$ denote the $mn \times mn$ matrix with entries corresponding to the pseudo-derivatives computed above and evaluated at c, so $\Psi(c)$ is a block-diagonal matrix with the size of each block equal to $m \times m$. As a consequence, the evaluation of the inverse of $\Psi(c)$ is relatively inexpensive.[13] Thus, we can choose c^0, set $B^0 = \Psi(c^0)$, and apply Broyden's method to the collocation function F. As discussed by Kelley (2003), this is equivalent to setting $B^0 = I_{nm}$, where I_{nm} is the $nm \times nm$ identity matrix, and then solving a *left-preconditioned* version of the collocation equations using the function

$$\Psi F(c) = [\Psi(c^0)]^{-1} F(c)$$

instead of the function F.

While the above left-preconditioned version of Broyden's method offers one feasible route for the implementation of the method, the availability of a relatively cheap approximation to the Jacobian $\Psi(c)$ of F for every c suggests an even more appealing alternative. Observe that if $\Psi(c)$ is indeed a good approximation of the Jacobian of F and is invertible, then the Jacobian of the function $[\Psi(c)]^{-1} F(c)$ is close to the identity matrix I_{nm}. The low-cost of computing $[\Psi(c)]^{-1}$ suggests that we can consider applying Broyden's algorithm to a modified function $\hat{F}$ given by

$$\hat{F}(c) = [\Psi(c)]^{-1} F(c) = c - [\Psi(c)]^{-1} S(c), \tag{15}$$

where $S(c)$ is a $nm \times 1$ column vector whose entry corresponding to i's coefficient on T_k is

$$S_{i,k}(c) = u_i(\tilde{x}_k) + \delta_i \int_q \int_\theta \sum_{h \in N} p_h(\theta_i \cdot \pi_h(q, \theta; c)) f(\theta) \mathrm{d}\theta g(q|\tilde{x}_k) \mathrm{d}q. \tag{16}$$

Assuming $\Psi(c^*)$ is non-singular,[14] the collocation coefficients c^* solve the collocation equations (i.e., $F(c^*) = 0$) if and only if they solve $\hat{F}(c^*) = 0$. As a result, we have transformed the system of equations $F(c) = 0$ to a system $\hat{F}(c) = 0$ that has a lower condition number (see Judd 1998, pp. 67–70, Sect. 5.7) and is much closer to being linear. We thus proceed to apply Broyden's method to the function $\hat{F}$ instead of the collocation function F, using initial iterate c^0 and initial approximate Jacobian $B^0 = I_{nm}$. It turns out that this implementation of Broyden's method yields far superior performance in the numerical experiments we consider.

[13] In fact, in the case in which players share the same discount factor, the blocks forming the block-diagonal matrix $\Psi(c)$ are identical for each of the n players, and computation of the inverse of $\Psi(c)$ reduces to the computation of the inverse of one $m \times m$ matrix.

[14] It is straightforward to show that for all c, $\Psi(c)$ can be singular for at most a finite number (possibly zero) of vectors of discount factors $(\delta_1, \ldots, \delta_n)$.

5.3 Two alternative solution methods

In this section, we describe two additional methods that can be used as alternatives to Broyden's method in order to solve the collocation equations. Outside the class of Newton-like methods for solving nonlinear equations, a standard approach for the solution of the collocation equations is a form of *value iteration.* This method starts with an initial guess c^0, and in the τ-th iteration it generates a new set of collocation coefficients $c^{\tau+1}$ by performing Chebyshev function interpolation on the dynamic expected utility functions evaluated at the collocation nodes, where the values of these functions are computed using the collocation coefficients, c^τ. Specifically, at the τ-th iteration, the collocation coefficients c^τ imply a vector $\hat{U}^\tau = (\hat{U}_1^\tau, \ldots, \hat{U}_n^\tau)$ of nm values for the n players' dynamic expected utilities at each of the m collocation nodes according to

$$\hat{U}_i^\tau(\tilde{x}_k) = u_i(\tilde{x}_k) + \delta_i \int_q \int_\theta \sum_{h \in N} p_h \left[U(\pi_h(q, \theta; c^\tau); c_i^\tau) \right.$$
$$\left. + \theta_i \cdot \pi_h(q, \theta; c^\tau) \right] f(\theta) \mathrm{d}\theta g(q|\tilde{x}_k) \mathrm{d}q,$$

which updates $\hat{U}^\tau$ by computing "best response" proposals, $\pi_h(q, \theta; c^\tau)$. We then use the values $\hat{U}_i^\tau(\tilde{x}_k)$ to obtain $c^{\tau+1}$ by performing interpolation, i.e., by solving the linear system of equations

$$\hat{U}_i^\tau(\tilde{x}_k) = \sum_{\ell=1}^{m} c_{i,\ell}^{\tau+1} T_\ell(\tilde{x}_k), \tag{17}$$

$i = 1, \ldots, n, k = 1, \ldots, m$, for the unknown $c^{\tau+1}$. A form of value iteration is used by Duggan et al. (2008) to solve a variant of the model that we study. They also use a Chebyshev polynomial representation of the unknown functions to be solved for, but in their solution method the corresponding Eq. 17 are not solved exactly, as they use more nodes at which to evaluate the unknown functions than coefficients, i.e., they perform function approximation instead of interpolation.

A second alternative is motivated by the form of the transformed collocation function $\hat{F}$ in (15). In particular, it is apparent from the definition of the function $\hat{F}$ that collocation coefficients solving the equations $\hat{F}(c) = 0$ constitute a fixed point of the function

$$\hat{\hat{F}}(c) = -[\Psi(c)]^{-1} S(c),$$

where $S(c)$ is defined in (16). Thus, when viewed as a fixed point $c = \hat{\hat{F}}(c)$, a solution to the original collocation equations $F(c) = 0$ can then be obtained by *fixed point iteration* on the auxiliary function $\hat{\hat{F}}$, so that at the τ-th iteration we obtain

$$c^{\tau+1} = \hat{\hat{F}}(c^\tau). \tag{18}$$

An alternative way to motivate this iterative method is to construe it as a *pseudo-Newton* method. Again guided by (15), we note that the updating step described in (18) is equivalent to a Newton step for the collocation function F, since

$$c^{\tau+1} = \hat{\hat{F}}(c^{\tau}) = c^{\tau} - [\Psi(c^{\tau})]^{-1} F(c^{\tau}),$$

where instead of the Jacobian $F'(c^{\tau})$ used in the conventional Newton method, we have substituted the approximation $\Psi(c^{\tau})$. Viewed as a pseudo-Newton method, the iteration (18) then admits all the globalization strategies employed for correcting Newton iterations. In particular, if at the τ-th iteration the step suggested by the pseudo-Jacobian $\Psi(c^{\tau})$ does not lead to a sufficient decrease in the norm of the residual $||F(c^{\tau+1})||$ compared to the residual $||F(c^{\tau})||$, then we can adjust the length of the step along the same direction. In the numerical experiments we report, we implement both Broyden's method and this pseudo-Newton method using the same Armijo rule and a parabolic line search for the optimal step size at each iteration.

6 Numerical experiments and core convergence

In this section, we conduct a number of numerical experiments designed to evaluate the performance of the collocation method and the techniques for solving the collocation equations for dynamic bargaining games we developed in Sects. 3–5. We begin in Sect. 6.1 with a description of the numerical specification of the model parameters and other numerical specifications required to implement the algorithms. In Sect. 6.2, we discuss and compare the performance of the three methods for solving the collocation equations, and we provide an application of Theorem 3 for the purposes of approximating equilibria in a model with a continuous space of policies. As an application of these techniques, we conclude in Sect. 6.3 with an illustration of the core convergence result of Duggan and Kalandrakis (2008).

6.1 Specifications

Throughout this section, we specify models with a two-dimensional policy space ($d = 2$),[15] and we assume that the set of feasible policies is a finite grid contained in the square

$$\tilde{X} = [-1, 1]^2.$$

Given $x = (x_1, x_2)$, we set the support of the density of the status quo, $g(q|x)$, to

$$\left[\frac{9}{10}x_1 - \frac{1}{10}, \quad \frac{9}{10}x_1 + \frac{1}{10}\right] \times \left[\frac{9}{10}x_2 - \frac{1}{10}, \quad \frac{9}{10}x_2 + \frac{1}{10}\right] \subset \tilde{X}.$$

[15] We have written software that allows us (at increasing cost) to solve for equilibria in models with higher-dimensional policy spaces.

Each coordinate $q_i, i = 1, 2$, of the status quo is independently Beta distributed in $\left[\frac{9}{10}x_i - \frac{1}{10}, \frac{9}{10}x_i + \frac{1}{10}\right]$ with parameters $a = b = 5$.[16] This specification ensures that $g(q|x)$ is twice continuously differentiable with respect to x with Lipschitz bounded second derivative, as assumed in Sect. 2.1. We fix the distribution of the preference shocks, $f(\theta)$, to be uniform in $(-\frac{1}{20}, \frac{1}{20})^{2n}$. We solve models with $n = 9$ players and simple majority rule, so that the set of winning coalitions is given by

$$\mathcal{D} = \{C \in N : |C| \geq 5\}.$$

We specify uniform recognition probabilities across players, so that $p_i = \frac{1}{9}$, and set common discount factors $\delta_i = \delta$. With the exception of one set of computations in the core convergence subsection, we use a value of $\delta = 0.7$ for the discount factor. The stage utilities of all players are negative quadratic with the ideal point of player i denoted $\hat{x}_i$, so that

$$u_i(x) = -(x_1 - \hat{x}_{i,1})^2 - (x_2 - \hat{x}_{i,2})^2.$$

We specify a number $m = 31 \times 31 = 961$ of collocation nodes $\tilde{x}_k$, which are located at the roots of the Chebyshev polynomials as described in Sect. 3. Based on the experimentation with alternative numbers of collocation nodes, this choice seems to strike a good compromise between achieving sufficient precision in the approximation of the dynamic expected utility functions U and computational cost. We borrow routines from the MATLAB *compecon* toolbox of Miranda and Fackler (2002) to generate the collocation nodes as well as the values of the Chebyshev polynomials at these nodes. For each collocation node $\tilde{x}_k$, we specify $\alpha = 25$ Gaussian quadrature nodes in order to perform the integration with respect to the status quo q. These nodes and the corresponding weights are specified using the MATLAB routine *qnwbeta* from the *compecon* toolbox of Miranda and Fackler (2002). We use a Sobol sequence of $\beta = 128$ quasi-random numbers as implemented by Burkardt (2007) in the MATLAB environment. We vary the size of the policy grid X, which is in all cases uniform in $[-1, 1]^2$ with sizes ranging from $7 \times 7 = 49$ to $51 \times 51 = 2{,}601$ points, although most of the computations are performed with a grid of size $35 \times 35 = 1{,}225$. Given the above, each function evaluation requires us to solve $n \times m \times \alpha \times \beta = 27{,}676{,}800$ optimization problems.

We have written MATLAB routines to perform these optimizations and the integrations required in order to evaluate the collocation functions F and $\hat{F}$ or $\hat{\hat{F}}$. These routines take advantage of MATLAB's *Parallel Computing Toolbox,* so that the players'

[16] Thus, $g(q|x)$ takes the form

$$g(q|x) = \begin{cases} \prod_{i=1}^{2} \frac{(q_i - \frac{9}{10}x_i + \frac{1}{10})^4 (\frac{1}{10} + \frac{9}{10}x_i - q_i)^4}{B(5,5)(\frac{2}{10})^9} & \text{if } \; q_i \in \left[\frac{9}{10}x_i - \frac{1}{10}, \frac{9}{10}x_i + \frac{1}{10}\right], i = 1, 2, \\ 0 & \text{otherwise.} \end{cases}$$

Table 1 Performance of three solution methods

	Broyden	Pseudo-Newton	Value iteration
Number that converged to 10^{-5}	10	8	10
Median function evaluations needed	8.5	9.5	19
Average function evaluations needed	10.4	8.875	18.8
Min function evaluations needed	6	6	17
Max function evaluations needed	22	12	21
Number that converged to 5×10^{-5}	10	10	10
Median function evaluations needed	5	5	14
Average function evaluations needed	5.7	5	13.5
Min function evaluations needed	5	4	12
Max function evaluations needed	9	6	14

optimization problems are executed in parallel in eight processors.[17] We used Kelley (2003)'s MATLAB routine *brsola* to implement the Broyden–Armijo method with a parabolic line search to adjust the Newton step. We adapted the same code in order to implement the same line search procedure for the pseudo-Newton method described in Sect. 5.3.

6.2 Numerical experiments

To compare the performance of the three solution methods discussed in Sect. 5, we randomly drew ten sets of ideal points $\hat{x}_i$ for the nine players from the uniform distribution in $[-1, 1]^2$. We then applied each of the three procedures (Broyden's method, the pseudo-Newton method, and value iteration) to compute an equilibrium for each of the models specified by the ten sets of ideal points. In these computations, the size of the policy grid is given by $35 \times 35 = 1{,}225$ points. All iterations were initiated with a value $c^0 = 0$ for the collocation coefficients. We monitored convergence using the averaged L_2 norm of the residual of the transformed collocation function, i.e., $\frac{||\hat{F}(c)||_2}{\sqrt{nm}}$. The function $\hat{F}$ is evaluated in our implementation of Broyden's method, and for the purposes of comparison we evaluate the residual of that function for both the value iteration and the pseudo-Newton methods.[18] We required a residual of 10^{-5} for convergence of the iterations. In Table 1, we report on the performance of the three methods.

As is evident from the first row of Table 1, both Broyden's method and value iteration converged to the target tolerance level in all cases, while the pseudo-Newton method failed to converge in two cases.[19] This occurred when the line search proce-

[17] All computations reported in this article were executed on a 3.2 GHz dual Quad-Core Intel Xeon Mac-Pro machine with 8 GB of memory operating under Microsoft's 64-bit Windows Vista Professional operating system.

[18] These extra computations constitute a negligible fraction of the overall computation. Note that the cost of a function evaluation in Broyden's method is identical to the cost of evaluating the expected payoff values $\hat{U}$ in value iteration, so it is appropriate to gauge the cost of these methods in terms of function evaluations. The bulk of this computational cost arises from solving the optimization problems of the proposers.

[19] We discarded those two cases in calculating average function evaluations needed by pseudo-Newton.

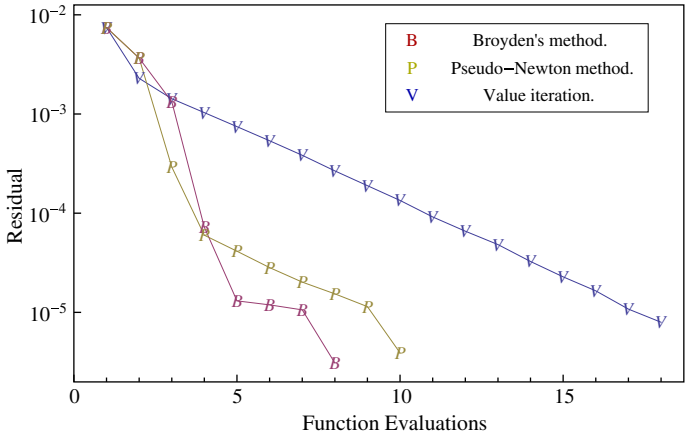

Fig. 2 Convergence paths for three solution methods

dure failed to find a step size that produced a sufficient reduction in the norm of the residual. According to Table 1, our implementation of Broyden's method outperforms value iteration in terms of the required function evaluations, as it required only 10.4 function evaluations on average compared to 18.8 for value iteration, thus economizing on function evaluations by approximately 50%. Furthermore, in the typical case, Broyden's method converged in 8 to 9 function evaluations, while value iteration consistently required roughly double that number in the range from 17 to 21 function evaluations. Table 1 also reveals that our pseudo-Newton method is quite competitive in the initial iterations. In particular, the pseudo-Newton method outperforms the other two methods when the required tolerance level is relaxed to 5×10^{-5}, in which case it converged in all ten specifications.

In Fig. 2, we display the convergence path of the three algorithms in one of the ten specifications in which the performance of the three methods (measured in terms of the number of function evaluations required for convergence) is roughly at the medians reported in Table 1. This figure is representative of the performance of the three methods. As expected, value iteration exhibits a consistent linear decrease in the residual norm. On the other hand, the other two methods exhibit super-linear convergence initially, but then slow down near the tolerance level. Although Newton-like methods perform better near the solution in the absence of numerical error in function evaluation, the slower convergence exhibited in the later iterations of Broyden's method in Fig. 2 is not surprising given the numerical precision that we have built into the problem. Interestingly, it is exactly near that tolerance level that the pseudo-Newton method loses its initial advantage over Broyden's method. The advantage of Broyden's method at these later iterations is due to the fact that it accumulates information across iterations that results in a more accurate representation of the Jacobian of the function $\hat{F}$ in later iterations. Our pseudo-Newton method, on the other hand, relies on the pseudo-Jacobian $\Psi(c^\tau)$ at each iteration, and that contains no information on the indirect effect of changes in the collocation coefficients c^τ on the likelihood of different values for the optimal proposals $\pi_h(q, \theta; c^\tau)$.

We performed a number of checks to test the quality of the solution produced at these tolerance levels for each method. Given solution c^*, we evaluated the dynamic expected payoffs of players at a large number of 50×50 test points on a uniform grid in $\tilde{X}$ that do not coincide with the 31×31 collocation nodes. We then computed the residual difference between these function values and those obtained from the collocation coefficients; namely, for each test point $z \in \tilde{X}$ and each player, we computed

$$R_i(z) = U(z; c_i^*) - \Bigg[u_i(z) + \delta_i \int_q \int_\theta \sum_{h \in N} p_h \left[U(\pi_h(q, \theta; c^*); c_i^*) + \theta_i \cdot \pi_h(q, \theta; c^*) \right] f(\theta) g(q|z) \mathrm{d}\theta \mathrm{d}q \Bigg].$$

The average L_2 norm of these residuals is roughly 7×10^{-5}, and it is less than 5×10^{-4} in the L_∞ norm. These numbers are consistent with the tolerance level of 10^{-5} used to gauge convergence of the algorithms.

It should be noted that we also attempted alternative implementations of Broyden's method, using the unconditioned function F or the left-preconditioned version ΨF discussed in Sect. 5.2, instead of the dynamically preconditioned version $\hat{F}$, but these versions of Broyden's method did not perform as well compared to either of the alternatives that we present. The conclusion we reach from this discussion and from Table 1 is that Broyden's method constitutes a viable solution method that can lead to a significant economization in computing time over value iteration when solving the quasi-discrete dynamic bargaining games that we study. Translated in terms of computer time, value iteration, with a median number of 19 function evaluations, required roughly in excess of 54 min to solve a typical instance of our nine-player model, while the corresponding cost for Broyden's method drops to just 23 min. If only a rough approximation of the solution is required, then our pseudo-Newton method is a competitive alternative.

We conclude this section with an illustration of Theorem 3, which allows us to obtain an increasingly precise approximation of the equilibrium of the continuum model from a sequence of equilibrium computations of quasi-discrete models. In particular, we chose the typical specification of ideal points among the ten used to produce Table 1, on which we reported in Fig. 2, and we used Broyden's method to compute an equilibrium for a sequence of six quasi-discrete models with grid sizes ranging from 7×7 to 51×51, roughly doubling the number of points in the policy grid X with each successive model. In order to ensure independence of the solutions for each model, we started Broyden's method from $c^0 = 0$ for each of the six models in the sequence.[20]

[20] This is not the approach we would use if our goal were to economize on computation time. In that case, we would use the solution output (and possibly the approximated Jacobian) from application of Broyden's method in each lower grid size model in order to provide a 'hot' starting point for the next model with finer policy grid, thus significantly economizing on the number of function evaluations needed for the models with finer policy grids.

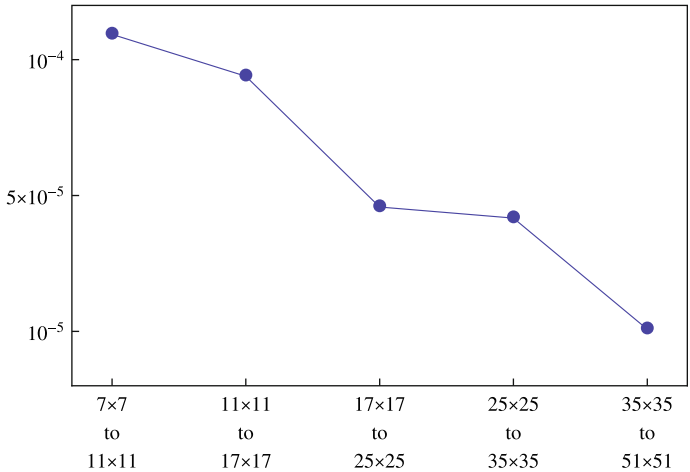

Fig. 3 Change in solution with finer policy grids

For each of the six models, we attained convergence to within the tolerance level of 10^{-5}.[21] In Fig. 3, we depict the change in the L_2 norm of the solution obtained for each successive pair of models. As is evident from the decreasing line in Fig. 3, these solutions grow successively closer to each other with each increase on the size of the policy grid. In fact, the equilibrium of the model with 35×35 grid differs from that obtained for the 51×51 grid by roughly 10^{-5}, suggesting that (at least for the specifications we use) we reasonably approximate equilibria of the continuum model by looking at these grid sizes.

6.3 Core convergence

Every equilibrium of the dynamic bargaining game we consider admits an invariant distribution (at least one) over implemented policies, an implication of Theorem 4 of Duggan and Kalandrakis (2008). In Theorem 6 of that article, we also show that when the noise on the status quo and the preference shocks are small, and when the stage utilities are close to quadratic and are close to admitting a core policy, and the player located at the core has positive probability of recognition, then the invariant distributions generated by equilibria of the model must be close (in the sense of weak convergence) to the unit mass on the limiting core policy. In this section, we implement the Broyden collocation method in the two-dimensional, nine-player model specified in Sect. 6.1 to provide a numerical illustration of this convergence result.

[21] Convergence to this tolerance level is generally harder in the models with coarser policy spaces, and Broyden's method converged to within 2×10^{-5} in the model with an 11×11 policy grid. We then ran a small number of value iterations to get the solution of that model to within the required tolerance level.

Table 2 Location of ideal points of players 1–8 in core convergence experiment

	$\hat{x}_1$	$\hat{x}_2$	$\hat{x}_3$	$\hat{x}_4$	$\hat{x}_5$	$\hat{x}_6$	$\hat{x}_7$	$\hat{x}_8$
$\hat{x}_{i,1}$	−0.8	0.3	−0.2	0.9	0.1	−0.15	0.3	−0.9
$\hat{x}_{i,2}$	0	0	0.2	−0.9	0.6	−0.9	0.2	−0.6

As a first step, we address the requirements of Duggan and Kalandrakis's (2008) dynamic core theorem. The specification we have chosen meets the requirements regarding the curvature of the players' stage utilities, as we assume they are quadratic. Duggan and Kalandrakis's (2008) theorem assumes a sequence of models that exhibit increasingly smaller levels of noise on preferences and the status quo. In our specification, we have assumed a relatively small level of noise, which we keep fixed in the computations that follow. These computations demonstrate that the equilibrium forces that yield the theorem take effect well before the noise on preferences and the status quo become negligible, i.e., even if fixed at the levels we have already specified. In addition to these requirements, the theorem assumes a configuration of stage preferences that becomes closer to admitting a core. For that purpose, we fix the ideal points of players 1 through 8 at the values reported in Table 2. It is straightforward to verify that, given negative quadratic stage utilities and absent preference shocks, these ideal points satisfy Plott's (1967) pairwise-symmetry conditions for the existence of a core point at the origin of the space $(0, 0)$. With the ideal points of the remaining players thus fixed, we can move the ideal point of player 9 from an arbitrary position toward the origin and monitor the effect of this move on the invariant distribution over policies induced in equilibrium.

In our numerical experiments, we varied the ideal point of player 9 from the location $\hat{x}_9 = (-0.4, -0.4)$, at which a core point does not exist, to an intermediate location $\hat{x}_9 = (-0.2, -0.2)$, and finally to the point $\hat{x}_9 = (0, 0)$, at which player 9 is located at the core (absent preference shocks). We considered two possible values for the common discount factor, a low value of $\delta = 0.3$ and the value $\delta = 0.7$ used for the computations we have already reported. We computed equilibria for each of these six configurations of ideal points and discount factors using a space of policies X given by a 51×51 uniform grid in $[-1, 1]^2$. To compute an equilibrium for that grid size, we first computed equilibria using coarser grids at a much lower computation cost, and then we gradually increased the size of the grid, using the solutions from smaller grids in order to initiate Broyden's algorithm for finer grids. At the 51×51 grid size, we required a more stringent convergence tolerance of 5×10^{-6} for termination of the algorithm, and convergence typically required only one or two iterations at this grid level, as the initial values were already quite close to the solution.[22]

Upon obtaining an equilibrium in this fashion, we simulated a long sequence of equilibrium play over 20, 000 periods, retaining the policy implemented in each period, and then used the last 15, 000 periods as a sample from an invariant distribution induced by the equilibrium. We used this sample in order to depict this invariant distribution over

[22] Thus, consistent with Theorem 3 and the conclusion drawn from Fig. 3, we can view the computed equilibria at this grid size as a good approximation of equilibrium in the continuous model.

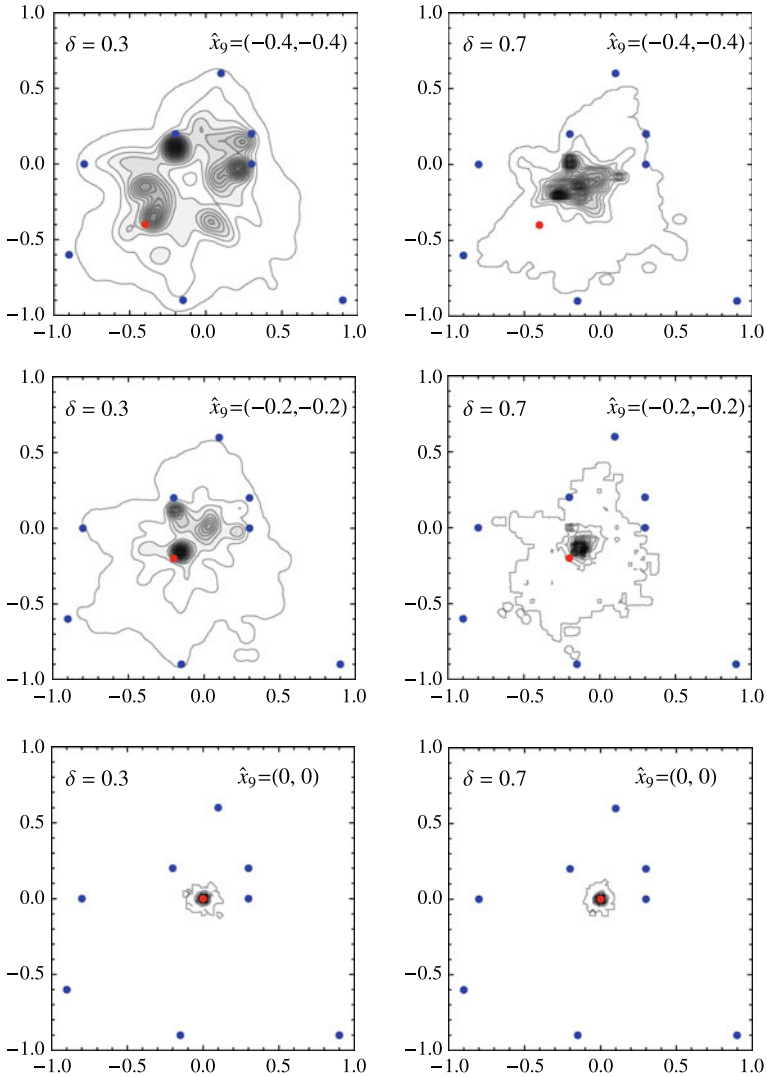

Fig. 4 Core convergence. Models with $\delta = .3$ appear on the left, those with $\delta = .7$ on the right. Long-term policies when $\hat{x}_9 \neq (0, 0)$ tend to be more moderate when players are more patient. Long-term policies pile mass at the core point when $\hat{x}_9 = (0, 0)$

policies in Fig. 4 using a kernel density estimator. The first column of Fig. 4 displays the invariant distribution over policies for the three locations for the ideal point of player 9 and a low discount factor $\delta = 0.3$, and the second column corresponds to a discount factor $\delta = 0.7$. The last row of these graphs corresponds to the case where player 9's location is at the core point (0, 0). Moving from the first row to the last row of these graphs, we find that the equilibrium invariant distribution piles more mass near the core point for both discount factors. In the last row, the invariant distribution

is obviously concentrated near the origin with virtually no policies occurring away from the limiting core point in accordance with Theorem 6 of Duggan and Kalandrakis (2008).

Observe that the invariant distribution when $\delta = 0.3$ and $\hat{x}_9 \neq (0, 0)$ is markedly more dispersed than the corresponding distribution when $\delta = 0.7$. Evidently, the strategic incentives of the players induce them to be more moderate and seek policies closer to the center of the policy space when the discount factor is larger. Such moderate policies insulate players from the risk of a large change in policy when a player with whom they disagree significantly is the proposer and can leverage the policy outcome in her favor due to the fact that a status quo is located too far away from the remaining players' preferences. The moderating effect of players' patience was also identified in the computations of Baron and Herron (2003) in a two-dimensional setting with three players symmetrically located in an equilateral triangle. Our analysis suggests that the observations of those authors are robust to the horizon of the game, and they raise the question of whether a theoretical explanation for this regularity is possible—a question we echo but leave open.

In order to illustrate the players' voting and proposal strategies, we depict the collective acceptance set and optimal proposals for various status quos in Fig. 5. The three figures on the right column of Fig. 5 correspond to the equilibrium for the specification of ideal points that is the furthest from satisfying Plott's conditions so that the ideal points of players 1 to 8 are as specified in Table 2, that of player 9 is located at $(-0.4, -0.4)$, and the discount factor is set to $\delta = 0.7$. For the purposes of comparison, we also depict the corresponding acceptance sets and proposal strategies when players are impatient ($\delta = 0$) in the left column of Fig. 5. A number of observations emerge from this comparison. First, the collective acceptance sets when players are patient tend to include more alternatives in the center of the policy space and fewer at the extremes. Second, many players strategically compromise their proposals by offering a moderate proposal instead of choosing the feasible policy that is closest to their ideal point, even in cases when it is feasible. Third, the collective acceptance set grows small as the status quo moves to a more central location at $q = (-0.1, -0.1)$.

In Fig. 6, we display the equilibrium preferences of one of the players in each of the three models with discount factor $\delta = 0.7$ to elucidate the nature of the equilibrium and the strategic incentives of the players. The first row of Fig. 6 depicts the continuation value of player 7 and is in some sense a representation of the effect of future equilibrium play on that player's incentives. Note that for all three models (player 9 located at $(-0.4, -0.4)$, $(-0.2, -0.2)$, and $(0, 0)$, respectively), the policy outcomes that engender the best future distribution of policies for player 7 are those that are close to the center of the policy space, near the ideal policy of player 9. This area of the policy space concentrates most of the mass of the equilibrium invariant distribution, as is evident by the second column of Fig. 4, and player 7's continuation value is maximized when the current policy is near that area of the policy space. Nevertheless, player 7 has a stage ideal point fixed at $\hat{x}_7 = (0.3, 0.2)$, which is removed from the policies that generate higher future utility for that player. The net effect of these two incentives, i.e., those concerning player 7's future expected utility versus the utility derived from present policy, is that the policy generating the highest expected discounted payoff for player 7 is a compromise between her stage ideal point and the policies that are near the center

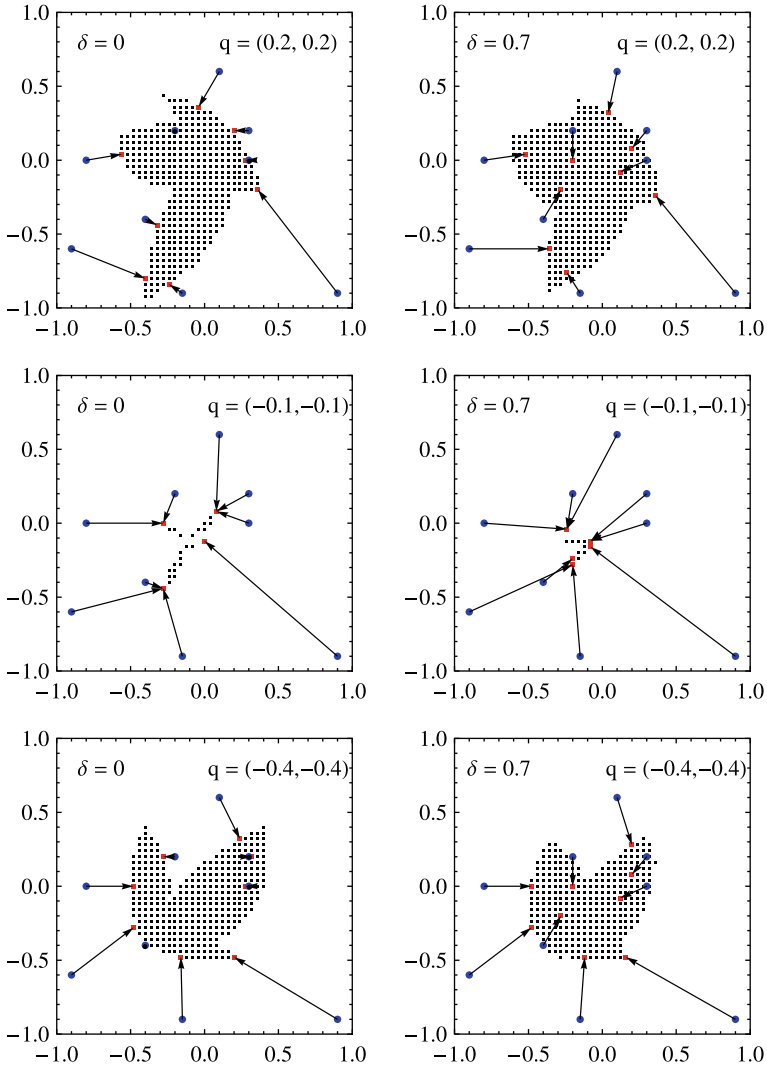

Fig. 5 Social acceptance sets and proposal strategies. Collectively acceptable alternatives are displayed using small black points, acceptable alternatives using larger points. *Arrows* originate from ideal points and point to the proposals of the corresponding player. Individual preference shocks θ_i are set to zero for all i

of the policy space. This is evident from the second row of Fig. 6. This row displays player 7's dynamic utility $U_7(x)$, which in all cases appears to have a maximizer closer to the origin compared to the stage ideal point of the player. The dynamic utility also appears to be non-concave in the equilibrium of the model with $\hat{x}_9 = (-0.4, -0.4)$, although this non-concavity is not very pronounced, despite the fact that player 7's stage utility is strictly concave. Of course, the dynamic utility U_7 is derived endogenously as the convex combination of the stage utility and player 7's continuation

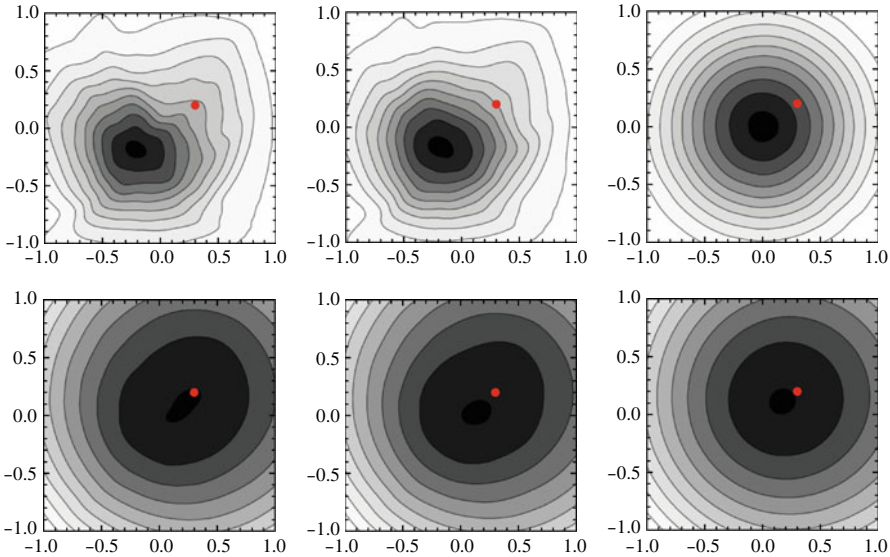

Fig. 6 Continuation value and dynamic utility. Continuation value $v_7(x)$ (*first row*) and dynamic utility $U_7(x)$ (*second row*) of player 7, for the models with player 9's stage utility ideal point located at $(-0.4, -0.4)$, $(-0.2, -0.2)$, and $(0, 0)$, respectively. The highlighted point indicates player 7's stage ideal point. *Darker areas* indicate higher utility

value function, so the non-concavity is not surprising given the fact that player 7's continuation value is shaped by the complicated nature of future equilibrium play.

7 Conclusion

We have proposed one approach to computation of equilibrium in a general class of dynamic bargaining games, and we have provided theoretical support for this approach. Despite the fact that proposals are endogenous in our model, we have proven that we can approximate equilibrium dynamic utilities as the solution to a sufficiently smooth system of collocation equations. We have shown that a preconditioned version of Broyden's method constitutes a viable solution method for the quasi-discrete model, and that it can lead to a significant economization in computing time over value iteration. Furthermore, taking the limit of equilibria of a sequence of increasingly finer quasi-discrete models, we can approximate equilibria of the continuum model. We employ these techniques in an illustration of core convergence in the model. Our computational results suggest that greater patience on the part of the players may induce greater policy moderation, opening the question of a theoretical explanation for this regularity. These contributions, taken together, should complement the future analysis of such theoretical questions as the effect of proposal power or the distribution of voting rights on equilibrium policy outcomes, and they may facilitate the empirical estimation of unobservable parameters, such as proposal probabilities or discount factors, in dynamic bargaining games.

Appendix: proof of Theorem 2

Theorem 2 *Assume X is finite and f is C^1. The collocation function $F : \mathbb{R}^{nm} \to \mathbb{R}^{nm}$ is locally Lipschitz continuous and directionally differentiable.*

Proof To any $c \in \mathbb{R}^{nm}$ we can associate policy-specific dynamic payoffs $U(y; c_i)$ as in (9), and we can consider the optimality conditions (5) and (6). We define $\Theta(q, y, h, a_{-h}; c)$ as the closure of the subset of preference shocks $\theta \in \mathbb{R}^{nd}$ for which the acceptance sets of the players other than h satisfying (5) are summarized by the vector $a_{-h} \in A = \{0, 1\}^{(n-1)|X|}$, and the proposal $y \in X \cup \{q\}$ solves (6) for proposer h. Here, $a_{-h} = (a_1, \ldots, a_{h-1}, a_{h+1}, \ldots, a_n)$, and $a_{j,p} = 1$ indicates that j will vote to accept x_p if proposed, and $a_{j,p} = 0$ indicates rejection. To be more precise, the shocks $\theta \in \Theta(q, y, h, a_{-h}; c)$ are characterized by two sets of inequalities. The first set of inequalities relates to the voting incentives of players other than h. For all $x_p \in X$ and all $j \in N \backslash \{h\}$, θ must satisfy

$$U(q; c_j) + \theta_j \cdot q \leq U(x_p; c_j) + \theta_j \cdot x_p \tag{19}$$

if $a_{j,p} = 1$, and it must satisfy

$$U(q; c_j) + \theta_j \cdot q \geq U(x_p; c_j) + \theta_j \cdot x_p \tag{20}$$

if $a_{j,p} = 0$. The second set of inequalities concerns the optimality of proposing y for player h. Let $C(x_p, h, a_{-h}) = \{i \in N \setminus \{h\} : a_{i,p} = 1\}$ be the coalition of players other than h who accept proposal x_p when acceptance sets of players other than h are given by a_{-h}. Given a status quo q and potential proposal $y \in X \cup \{q\}$, let $Y(y, h, a_{-h}) = (\{x_p \in X : C(x_p, h, a_{-h}) \cup \{h\} \in \mathcal{D}\} \cup \{q\}) \backslash \{y\}$ be the subset of policies other than y that can pass by the vote of h and those of other players, given acceptance sets a_{-h}. For proposing $y \in X \cup \{q\}$ to be optimal for player h given a_{-h}, θ must be such that for all $z \in Y(y, h, a_{-h})$,

$$U(z; c_h) + \theta_h \cdot z \leq U(y; c_h) + \theta_h \cdot y. \tag{21}$$

Thus, the closure of the subset of preference shocks such that h proposes y when the status quo is q and the acceptance sets of other players are summarized by a_{-h} is defined as follows:

$$\Theta(q, y, h, a_{-h}; c) = \{\theta \in \Theta : (19), (20), \text{and } (21) \text{ hold}\}.$$

It follows that $\Theta(y, q, h, a_{-h}; c)$ is a convex polytope defined by a finite number of inequalities that are linear in θ. Note that when $x_p \neq q$, the inequalities (19–21) are satisfied with equality only for a lower dimensional set of θ's, and furthermore, the intersection

$$\Theta(q, y, h, a_{-h}; c) \cap \Theta(q, y', h, a_{-h}; c)$$

is a measure-zero subset of $\mathbb{R}^{nd}$ for all distinct $y, y' \in \tilde{X}$. Hence, the probability that player h proposes y when the status quo is q is given by

$$\sum_{a_{-h} \in A} \int_{\Theta(q,y,h,a_{-h};c)} f(\theta)\mathrm{d}\theta.$$

Based on the above, we rewrite the collocation equations as follows:

$$F_{i,k}(c) = U(\hat{x}_k; c_i) - \Bigg[u_i(\hat{x}_k) + \delta_i \sum_{h \in N} p_h \sum_{a_{-h} \in A} \times \int_q \sum_{y \in X \cup \{q\}} \int_{\Theta(q,y,h,a_{-h};c)} [U(y; c_i) + \theta_i \cdot y]\, f(\theta)\mathrm{d}\theta g(q|\hat{x}_k)\mathrm{d}q \Bigg].$$

Note that in the above integral, the status quo q ranges over $\mathbb{R}^d$, and the probability that the realized status quo lies in the finite set X is zero; thus, we can neglect the measure-zero event that $q \in X$.

For all $i \in N$, all $q \in \tilde{X}$, all $y \in X \cup \{q\}$, all $h \in N$, and all $a_{-h} \in A$, we define a function $G_i(\cdot; q, y, h, a_{-h}) : \mathbb{R}^{nm} \to \mathbb{R}$ by

$$G_i(c; q, y, h, a_{-h}) = \int_{\Theta(q,y,h,a_{-h};c)} [U(y; c_i) + \theta_i \cdot y]\, f(\theta)\mathrm{d}\theta,$$

so that $F_{i,k}$ can be expressed equivalently as

$$F_{i,k}(c) = U(\hat{x}_k; c_i) - \Bigg[u_i(\hat{x}_k) + \delta_i \sum_{h \in N} p_h \sum_{a_{-h} \in A} \int_q \sum_{y \in X \cup \{q\}} G_i(c; q, y, h, a_{-h}) g(q|\hat{x}_k)\mathrm{d}q \Bigg]. \tag{22}$$

The properties of F of interest will follow from the properties of G.

We prove the theorem in three steps.

Step 1 *For all $i \in N$, all $q \in \tilde{X} \backslash X$, all $y \in X \cup \{q\}$, all $h \in N$, and all $a_{-h} \in A$, there exist functions $\hat{G}_j : C_j \to \mathbb{R}$, $j = 0, \ldots, K$, such that $\{C_j\}_{j=0}^K$ is an open covering of $\mathbb{R}^{nm}$; for all $j = 0, \ldots, K$, $\hat{G}_j$ is continuously differentiable; and $G_i(\cdot; q, y, h, a_{-h})$ is piecewise smooth with representation $\{\hat{G}_0, \hat{G}_1, \ldots, \hat{G}_K\}$.*

Fix $i \in N$, $q \in \tilde{X} \backslash X$, $y \in X \cup \{q\}$, $h \in N$, and $a_{-h} \in A$. To conserve notation, we will suppress the parameters i, y, h, a_{-h} in our construction of the representation; given q, these parameters range over a finite set, and so this convenience is harmless.

Because q ranges over the infinite set $\tilde{X}$, however, we will make the dependence of K, C_j, and $\hat{G}_j$ on q explicit. Define $\hat{G}_0 \equiv 0$, with $C_0 = \mathbb{R}^{nm}$ independently of q. If $\Theta(q, y, h, a_{-h}; c)$ has measure zero in $\mathbb{R}^{nm}$ for all c, then we let $K(q) = 0$, which completes the step. Otherwise, let $C^*(q) \subseteq \mathbb{R}^{nm}$ denote the open set of $c \in \mathbb{R}^{nm}$ such that $\Theta(q, y, h, a_{-h}; c)$ has positive measure. We then rewrite (19–21) as follows: for all $x_p \in X$, all $j \in N\backslash\{h\}$, and all $z \in Y(y, h, a_{-h})$,

$$\theta_j \cdot (q - x_p) \leq U(x_p; c_j) - U(q; c_j) \quad \text{if } a_{j,p} = 1 \tag{23}$$

$$\theta_j \cdot (x_p - q) \leq U(q; c_j) - U(x_p; c_j) \quad \text{if } a_{j,p} = 0 \tag{24}$$

$$\theta_h \cdot (z - y) \leq U(y; c_h) - U(z; c_h). \tag{25}$$

We can write these inequalities in the form $\alpha(q)\theta \leq \beta(c; q)$, where the first $(n-1)|X|$ rows of the matrix $\alpha(q)$ and the column vector $\beta(c; q)$ correspond to inequalities (23) or (24) (as appropriate), and the last $|Y(y, h, a_{-h})|$ rows correspond to inequalities (25). Note that the rows of $\alpha(q)$ are non-zero, as $q \notin X$. Of course, the constraint that θ belongs to Θ can also be formalized in terms of linear inequalities: for all $h = 1, \ldots, n$ and all $\ell = 1, \ldots, d$, $\theta_{h,\ell} \leq \overline{\theta}$ and $\theta_{h,\ell} \leq -\overline{\theta}$. Since these inequalities are fixed, we do not include them in the matrix representation of $\Theta(q, y, h, a_{-h}; c)$.

Given any $[(n-1)|X| + |Y(y, h, a_{-h})|] \times nd$ matrix A and any column vector b of dimension $(n-1)|X| + |Y(y, h, a_{-h})|$, define

$$\Theta(A, b) = \{\theta \in \Theta : A\theta \leq b\},$$

and note that the identity $\Theta(q, y, h, a_{-h}; c) \equiv \Theta(\alpha(q), \beta(c; q))$ holds on $C^*(q)$. Letting R index a subset of rows of A, we define $\Theta^R(A, b)$ as the polyhedral set of θ's satisfying the inequalities of R and the constraint that θ belong to Θ, so that

$$\Theta^R(A, b) = \{\theta \in \Theta : \text{ for all } r \in R,\ A_r\theta \leq b_r\},$$

where A_r is the rth row of the matrix A and b_r is the rth row of b. We say a set R is "minimal at (A, b)" if there is no set R' such that R' is a proper subset of R and $\Theta^R(A, b) = \Theta^{R'}(A, b)$. We let $\mathcal{R}(q)$ denote the collection of all sets of inequalities that are minimal at $(\alpha(q), \beta(c; q))$ for some $c \in C^*(q)$, i.e.,

$$\mathcal{R}(q) = \{R : \text{ there exists } c \in C^*(q) \text{ such that } R \text{ is minimal at } (\alpha(q), \beta(c; q))\},$$

and we set $K(q) = |\mathcal{R}(q)|$ and enumerate this collection as $R_1, \ldots, R_{K(q)}$. The domain of the function $\hat{G}_j(\cdot; q)$ to be defined will be

$$C_j(q) = \{c \in C^*(q) : R_j \text{ is minimal at } (\alpha(q), \beta(c; q))\},$$

an open set. For each $j = 1, \ldots, K(q)$, we define the functions Γ_j and Δ_j by

$$\Gamma_j(A, b) = \int_{\Theta^j(A,b)} f(\theta)\mathrm{d}\theta \quad \text{and} \quad \Delta_j(A, b) = \int_{\Theta^j(A,b)} (\theta_i \cdot y) f(\theta)\mathrm{d}\theta$$

for all (A, b), where we use $\Theta^j(A, b)$ for $\Theta^{R_j}(A, b)$. Finally, for each $j = 1, \ldots, K(q)$, we define the function $\hat{G}_j(\cdot; q)$ by

$$\hat{G}_j(c; q) = U(y; c_i)\Gamma_j(\alpha(q), \beta(c; q)) + \Delta_j(\alpha(q), \beta(c; q))$$

for all $c \in C_j(q)$.

Next, we claim that the functions $\hat{G}_j(\cdot; q)$, $j = 1, 2, \ldots, K(q)$, are continuously differentiable at each $c \in C_j(q)$. Given $r \in R_j$, define

$$\Theta_r^j(A, b) = \{\theta \in \Theta^j(A, b) : A_r\theta = b_r\}$$

as the $(nd-1)$-dimensional face of $\Theta^j(A, b)$ determined by the hyperplane $A_r\theta = b_r$. We claim that when $\Theta^j(A, b)$ has positive measure in $\mathbb{R}^{nd}$ and R_j is minimal at (A, b), the solution set $\Theta_r^j(A, b)$ has positive volume in the $(nd-1)$-dimensional hyperplane spanned by $\Theta_r^j(A, b)$. To see this, note that since $\Theta^j(A, b)$ contains an open set, there is some $\tilde{\theta}$ that satisfies the inequalities corresponding to R_j strictly. Since R_j is minimal, the rth inequality is not redundant: there exists $\hat{\theta} \in \Theta^{R_j\setminus\{r\}}(A, b)$ such that $A_r\hat{\theta} > b_r$. Then there exists $\alpha \in (0, 1)$ such that $A_r\theta' = b_r$, where $\theta' = (1-\alpha)\hat{\theta} + \alpha\tilde{\theta}$ satisfies the inequalities in $R_j\setminus\{r\}$ strictly. Letting $\tilde{\theta}$ vary while satisfying the inequalities of R_j, this implies that $\Theta_r^j(A, b)$ has positive $(nd-1)$-dimensional volume, as claimed. Therefore, Lasserre's (1998) Lemma 2.2 establishes that $\Gamma_j(A, b)$ and $\Delta_j(A, b)$ are continuously differentiable at such (A, b) with partial derivatives of the form

$$\begin{aligned}
\frac{\partial \Gamma_j}{\partial b_r}(A, b) &= \frac{1}{||A_r||} \int\limits_{\Theta_r^j(A,b)} f(\theta)\mathrm{d}\theta \quad \text{and} \\
\frac{\partial \Delta_j}{\partial b_r}(A, b) &= \frac{1}{||A_r||} \int\limits_{\Theta_r^j(A,b)} (\theta_i \cdot y) f(\theta)\mathrm{d}\theta,
\end{aligned} \tag{26}$$

where $||\cdot||$ denotes the Euclidean norm and integrals are with respect to Lebesgue measure in $(nd-1)$-dimensional space.[23] Given $c \in C_j(q)$, note that $\Theta(\alpha(q), \beta(c; q))$ has positive measure in $\mathbb{R}^{nd}$ by construction, and therefore $\Theta^j(\alpha(q), \beta(c; q)) \supseteq \Theta(\alpha(q), \beta(c; q))$ does as well; furthermore, R_j is minimal at $(\alpha(q), \beta(c; q))$. Thus, Γ_j and Δ_j are continuously differentiable in b at $(\alpha(q), \beta(c; q))$, and the chain rule implies $\hat{G}_j$ is continuously differentiable in c with partials

[23] Lasserre implicitly assumes that no rows of A are equal to zero. This is true for $\alpha(q)$ because we only consider $q \notin X$.

$$\frac{\partial \hat{G}_j}{\partial c_{i,k}}(c;q) = \frac{\partial U}{\partial c_{i,k}}(y;c_i)\Gamma_j(\alpha(q),\beta(c;q)) + \sum_{r\in R_j}\left[U(y;c_i)\frac{\partial \Gamma_j}{\partial b_r}(\alpha(q),\beta(c;q))\frac{\partial \beta_r}{\partial c_{i,k}}(c;q) + \frac{\partial \Delta_j}{\partial b_r}(\alpha(q),\beta(c;q))\frac{\partial \beta_r}{\partial c_{i,k}}(c;q)\right], \tag{27}$$

as claimed.

To show that $G_i(\cdot;q,y,h,a_{-h})$ is piecewise smooth with representation $\{\hat{G}_0,\hat{G}_1,\ldots,\hat{G}_K\}$, note that $G_i(c;q,y,h,a_{-h})$ takes values in the set $\{\hat{G}_0(c;q),\hat{G}_1(c;q),\ldots,\hat{G}_K(c;q)\}$. Indeed, if $c\notin C^*(q)$, then $G_i(c;q,y,h,a_{-h}) = 0 = G_0(c;q)$. Otherwise, consider any $c\in C^*(q)$, and let R_j be any set of inequalities that is minimal at $(\alpha(q),\beta(c;q))$ and such that $\Theta^j(\alpha(q),\beta(c;q)) = \Theta(q,y,h,a_{-h};c)$; then $G_i(c;q,y,h,a_{-h}) = \hat{G}_j(c;q)$, as claimed. Finally, we must show that $G_i(c;q,y,h,a_{-h})$ is continuous in c. Indeed, consider a sequence $c^\ell \to c\in\mathbb{R}^{nm}$. For each ℓ, there exists R_{j_ℓ} such that $c^\ell \in C_{j_\ell}(q)$ and $\Theta^{j_\ell}(\alpha(q),\beta(c^\ell,q)) = \Theta(q,y,h,a_{-h};c^\ell)$, which implies $G_i(c^\ell,q;y,h,a_{-h}) = \hat{G}_{j_\ell}(c^\ell,q)$. Without loss of generality (since the collection of subsets of inequalities is finite), we may suppose that j_ℓ is constant in ℓ, i.e., $j_\ell = j$. By continuity, we have $\Theta^j(\alpha(q),\beta(c;q)) = \Theta(q,y,h,a_{-h};c)$, though R_j may not be minimal at $(\alpha(q),\beta(c;q))$. It may be that c lies outside $C_j(q)$, but we nevertheless have

$$\begin{aligned} G_i(c^\ell,q;y,h,a_{-h}) &= \hat{G}_j(c^\ell,q) = \Gamma_j(\alpha(q),\beta(c^\ell,q)) \\ &\to \Gamma_j(\alpha(q),\beta(c;q)) = G_i(c;q,y,h,a_h), \end{aligned}$$

where the limit follows from a straightforward dominated convergence argument. Indeed, we define $\phi^\ell(\theta)$ to be $[U(y;c_i^\ell)+\theta_i\cdot y]f(\theta)$ times the indicator function of $\Theta^j(\alpha(q),\beta(c^\ell,q))$, and we define $\phi(\theta)$ as $[U(y;c_i)+\theta_i\cdot y]f(\theta)$ times the indicator function of $\Theta^j(\alpha(q),\beta(c;q))$. The sequence $\{\phi^\ell\}$ is dominated by an integrable function and converges pointwise almost everywhere to ϕ, verifying the limit. This establishes continuity, and we conclude that $G_i(c;q,y,h,a_{-h})$ is piecewise smooth in c.

Step 2 *For all $i\in N$, all $q\in\tilde{X}\backslash X$, all $y\in X\cup\{q\}$, all $h\in N$, and all $a_{-h}\in A$, the function $G_i(\cdot;q,y,h,a_{-h})$ is directionally differentiable in c; furthermore, $G_i(\cdot;q,y,h,a_{-h})$ is locally Lipschitz continuous with uniform Lipschitz constant, i.e., for all $c\in\mathbb{R}^{nm}$, there is a constant M and an open set $\tilde{C}$ containing c such that for all $q\in\tilde{X}\backslash X$, all $y\in X\cup\{q\}$, all $h\in N$, all $a_{-h}\in A$, and all $c',c''\in\tilde{C}$,*

$$|G_i(c'';q,y,h,a_{-h}) - G_i(c';q,y,h,a_{-h})| \le M||c''-c'||.$$

Fix $i\in N$, $q\in\tilde{X}\backslash X$, $y\in X\cup\{q\}$, $h\in N$, and $a_{-h}\in A$. Since $G_i(\cdot;q,y,h,a_{-h})$ is piecewise smooth with representation $\{\hat{G}_0(\cdot;q),\ldots,\hat{G}_{K(q)}(\cdot;q)\}$, by Step 1, Proposition 2.1 of Kuntz and Scholtes (1994) establishes that $G_i(\cdot;q,y,h,a_{-h})$ has directional derivatives at all $c\in\mathbb{R}^{nm}$, completing the first part of the step. To prove local

Lipschitz continuity with a uniform Lipschitz constant, consider any $c \in \mathbb{R}^{nm}$, and let $\tilde{C}$ be any open ball of finite radius containing c. We first claim that for every subset $R \subseteq \{1, \ldots, (n-1)|X|+|X|-1\}$ (corresponding to possible inequalities in (23–25)), there exists a bound M_R such that for all $q \in \tilde{X} \backslash X$, all $y \in X \cup \{q\}$, all $h \in N$, all $a_{-h} \in A$, all j with $R_j = R$, and all $\tilde{c} \in \tilde{C} \cap C_j(q)$, we have

$$\left| \frac{\partial \hat{G}_j}{\partial c_{i,k}}(\tilde{c}; q) \right| \leq M_R.$$

Referring to (27), the term $\frac{\partial U}{\partial c_{i,k}}(y; \tilde{c}_i)\Gamma_j(\alpha(q), \beta(\tilde{c}; q))$ is bounded over such q, y, h, a_{-h}, j, and $\tilde{c}$. The summand includes the term $U(y; \tilde{c}_i)$, which is also bounded over such q, y, h, a_{-h}, j, and $\tilde{c}$, and the terms

$$\frac{\partial \Gamma_j}{\partial b_r}(\alpha(q), \beta(\tilde{c}; q)) \frac{\partial \beta_r}{\partial c_{i,k}}(\tilde{c}; q) \quad \text{and} \quad \frac{\partial \Delta_j}{\partial b_r}(\alpha(q), \beta(\tilde{c}; q)) \frac{\partial \beta_r}{\partial c_{i,k}}(\tilde{c}; q).$$

We focus on the former, as the argument for the latter is similar, and rewrite it as

$$\frac{\partial \Gamma_j}{\partial b_r}(\alpha(q), \beta(\tilde{c}; q)) \frac{\partial \beta_r}{\partial c_{i,k}}(\tilde{c}; q) = \left(\int_{\Theta_r^j(\alpha(q), \beta(\tilde{c}; q))} f(\theta) \mathrm{d}\theta \right) \left(\frac{1}{||\alpha_r(q)||} \frac{\partial \beta_r}{\partial c_{i,k}}(\tilde{c}; q) \right).$$

The first term in the product is clearly bounded across such q, y, h, a_{-h}, j, and $\tilde{c}$; but the analysis of the second term in the product is complicated by the fact that $||\alpha_r(q)||$ can approach zero for some status quos $q \in \tilde{X} \backslash X$.

In the expression $\frac{1}{||\alpha_r(q)||} \frac{\partial \beta_r}{\partial c_{i,k}}(\tilde{c}; q)$, row r corresponds to one of two cases: either to an inequality in (23) or (24) with alternative x_{p_r} and voter $j_r \neq h$, or to an inequality in (25) with alternative z_r and proposer $i_r = h$. In the first case, the norm of $\alpha_r(q) = q - x_{p_r}$ becomes arbitrarily small when $||q - x_{p_r}||$ is small. The partial derivative $\frac{\partial \beta_r}{\partial c_{j,k}}(\tilde{c}; q)$ is equal to zero if $j_r \neq j$, and it is equal to $\pm T_k(q) - T_k(x_{p_r})$ otherwise. When $j_r = j$, note that as q approaches x_{p_r}, we have

$$\lim_{q \to x_{pr}} \left| \frac{1}{||\alpha_r(q)||} \frac{\partial \beta_r}{\partial c_{i,k}}(\tilde{c}; q) \right| = \lim_{q \to x_{pr}} \frac{|T_k(q) - T_k(x_{p_r})|}{||q - x_{p_r}||} \leq \|T_k'(x_{p_r})\| < \infty.$$

This limit holds independently of $q \in \tilde{X}$, $y \in X \cup \{q\}$, and $\tilde{c} \in \tilde{C}$, and since h, a_{-h}, and j belong to finite sets, we conclude that the expression $\frac{1}{||\alpha_r(q)||} \frac{\partial \beta_r}{\partial c_{i,k}}(\tilde{c}; q)$ is uniformly bounded. In the second case, we have $\alpha_r(q) = z_r - y$, which is constant in q unless $y \in X$ and $z_r = q$. Then the norm of $\alpha_r(q) = q - y$ becomes small when $||q - y||$ is small, and an analogous argument shows that the expression is again uniformly bounded. We conclude that the partial derivative $\frac{\partial \hat{G}_j}{\partial c_{i,k}}(\tilde{c}; q)$ is bounded over $q \in \tilde{X} \backslash X$, $y \in X \cup \{q\}$, $h \in N$, $a_{-h} \in A$, j with $R_j = R$, and $\tilde{c} \in \tilde{C} \cap C_j(q)$. If there do not exist such q, y, h, a_{-h}, j, and $\tilde{c}$, we set $M_R = 0$, delivering the claim. Using

the claim, we conclude that the directional derivatives $\hat{G}'_j(\cdot; q|s)$ are also bounded over directions s with $||s|| = 1, q \in \tilde{X}\backslash X, y \in X \cup \{q\}, h \in N, a_{-h} \in A, j$ with $R_j = R$, and $\tilde{c} \in \tilde{C} \cap C_j(q)$. Let $\tilde{M}_R$ be such a bound.

Finally, we deduce that $G_i(q, y, h, a_{-h}; \cdot)$ is Lipschitz on $\tilde{C}$ with uniform constant $M = \sum_R \tilde{M}_R$, the sum being over all subsets of possible inequalities. Consider any $c', c'' \in \tilde{C}$, and let $q \in \tilde{X}\backslash X, y \in X \cup \{q\}, h \in N$, and $a_{-h} \in A$ be arbitrary. Let $[c', c'']$ be the convex hull of $\{c', c''\}$. For each $\tilde{c} \in [c', c'']$, let $J(\tilde{c}) = \{j : \tilde{c} \in C_j(q)\}$ index the functions $\hat{G}_j(\cdot; q)$ that contain $\tilde{c}$ in their domain, and let $B(\tilde{c})$ be any open ball in $\mathbb{R}^{nm}$ around $\tilde{c}$ such that for all $j \in J(\tilde{c})$, we have $B(\tilde{c}) \subseteq C_j(q)$. We can cover $[c', c'']$ with such open balls, and by compactness there is a finite subcover. We focus on one such ball, say $\tilde{B}$, and we show that $G_i(q, y, h, a_{-h}; \cdot)$ is Lipschitz on $[c', c''] \cap \tilde{B}$ with uniform constant M. To this end, consider any $\tilde{c}', \tilde{c}'' \in [c', c''] \cap \tilde{B}$, and let $\tilde{J}$ be the set of j such that $\tilde{B} \subseteq C_j(q)$ and such that for some $\tilde{c} \in [\tilde{c}', \tilde{c}'']$, we have $G_i(\tilde{c}; q, y, h, a_{-h}) = \hat{G}_j(\tilde{c}, q)$. For each $j \in \tilde{J}$, let $I_j = \hat{G}_j([\tilde{c}', \tilde{c}''] ; q)$, and note that since $[\tilde{c}', \tilde{c}'']$ is compact and convex and $\hat{G}_j(\cdot; q)$ is continuous, the image of this set is a closed interval, say $I_j = [s_j, t_j]$. The image $G_i([\tilde{c}', \tilde{c}''] ; q, y, h, a_{-h})$ is also a closed interval contained in the union $\bigcup_{j \in \tilde{J}} I_j$, and by construction we have $G_i([\tilde{c}', \tilde{c}''] ; q, y, h, a_{-h}) \cap I_j \neq \emptyset$ for all $j \in \tilde{J}$. We, therefore, have

$$\begin{aligned} |G_i(\tilde{c}''; q, y, h, a_{-h}) - G_i(\tilde{c}'; q, y, h, a_{-h})| &\leq (\max_{j \in \tilde{J}} t_j) - (\min_{j \in \tilde{J}} s_j) \\ &\leq \sum_{j \in \tilde{J}} \tilde{M}_{R_j} ||\tilde{c}'' - \tilde{c}'|| \\ &\leq M ||\tilde{c}'' - \tilde{c}'||. \end{aligned}$$

Since M is independent of the ball $\tilde{B}$, this completes the step.

Step 3 *The collocation function F is locally Lipschitz continuous and directionally differentiable.*

We consider each coordinate function $F_{i,k}$, the result then following by a straightforward argument. Referring to (22), the term of interest is the integral

$$\int_q \sum_{y \in X \cup \{q\}} G_i(c; q, y, h, a_{-h}) g(q|\hat{x}_k) \mathrm{d}q. \tag{28}$$

For each q, we have shown that $G_i(c; q, y, h, a_{-h})$ is locally Lipschitz continuous and directionally differentiable, and that as a consequence, the integrand in (28) is locally Lipschitz continuous and directionally differentiable. In fact, we showed in Step 2 that for all $c \in \mathbb{R}^{nm}$, there is an open set $\tilde{C}$ containing c such that $G_i(\cdot; q, y, h, a_h)$ is Lipschitz continuous on $\tilde{C}$ with uniform constant. Then Proposition 1 of Qi et al. (2005) establishes that the expression in (28), and therefore $F_{i,k}$, is directionally differentiable and locally Lipschitz continuous. □

References

Acemoglu D, Egorov G, Sonin K (2008) Dynamics and stability of constitutions, coalitions, and clubs. Unpublished manuscript

Armijo L (1966) Minimization of functions having Lipschitz-continuous first patial derivatives. Pac J Math 16:1–3

Banks J, Duggan J (2000) A bargaining model of collective choice. Am Political Sci Rev 94:73–88

Banks J, Duggan J (2006) A general bargaining model of legislative policy-making. Q J Political Sci 1:49–85

Baron D (1996) A dynamic theory of collective goods programs. Am Political Sci Rev 90:316–330

Baron D, Ferejohn J (1989) Bargaining in legislatures. Am Political Sci Rev 83:1181–1206

Baron D, Herron M (2003) A dynamic model of multidimensional collective choice. In: Kollman K, Miller J, Page S (eds) Computational models in political economy. MIT Press, Cambridge

Battaglini M, Coate S (2007) Inefficiency in legislative policy-making: a dynamic analysis. Am Econ Rev 97:118–149

Binmore K (1987) Perfect equilibria in bargaining models. In: Binmore K, Dasgupta P (eds) The economics of bargaining. Basil Blackwell, New York

Broyden CG (1965) A class of methods for solving nonlinear simultaneous equations. Math Comput 19:577–593

Burkardt J (2007) The sobol quasirandom sequence. http://people.sc.fsu.edu/~burkardt/m_src/sobol/sobol.html

Cho SJ (2005) A dynamic model of parliamentary democracy. Unpublished manuscript

Diermeier D, Fong P (2008) Endogenous limits on proposal power. Unpublished manuscript

Duggan J, Kalandrakis T (2007) Dynamic legislative bargaining. W Allen Wallis Institute Working Paper Series 45

Duggan J, Kalandrakis T (2008) Dynamic legislative policy making. Unpublished manuscript

Duggan J, Kalandrakis T, Manjunath V (2008) Dynamics of the presidential veto: a computational analysis. Math Comput Model 48:1570–1589

Ip CM, Kyparisis J (1992) Local convergence of quasi-newton methods for b-differentiable equations. Math Program 56:71–89

Judd KL (1998) Numerical methods in economics. 1. The MIT Press, Cambridge

Kalandrakis T (2004) A three-player dynamic majoritarian bargaining game. J Econ Theory 116:294–322

Kalandrakis T (2010) Minimum winning coalitions and endogenous status quo. Int J Game Theory. 39(4):617–643

Kelley T (2003) Solving nonlinear equations with Newton's method. SIAM, Philadelphia

Kuntz L, Scholtes S (1994) Structural analysis of nonsmooth mappings, inverse functions, and metric projections. J Math Anal Appl 188:346–386

Kuntz L, Scholtes S (1995) Qualitative aspects of the local approximation of a piecewise differentiable function. Nonlinear Anal Theory Methods Appl 25:197–215

Lasserre JB (1998) Integration on a convex polytope. Proc Am Math Soc 126:2433–2441

Martinez JM, Qi L (1995) Inexact newton methods for solving nonsmooth equations. J Comput Appl Math 60:127–145

Mas-Colell A (1985) The theory of general economic equilibrium: a differentiable approach, no. 9, Econometric Society Monographs. Cambridge University Press, Cambridge

Miranda MJ, Fackler PL (2002) Applied computational economics and finance. The MIT Press, Cambridge

Pang JS (1990) Newton's method for b-differentiable equations. Math Oper Res 15:311–341

Pang JS (1991) A b-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems. Math Program 51:101–131

Penn EM (2009) A model of farsighted voting. Am J Political Sci 53:36–54

Plott CR (1967) A notion of equilibrium and its possibility under majority rule. Am Econ Rev 57:787–806

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes. In: The art of scientific computing, 2nd edn. Cambridge University Press, Cambridge

Qi L (1993) Convergence analysis of some algorithms for solving nonsmooth equations. Math Oper Res 18:227–244

Qi L (1997) On superlinear convergence of quasi-newton methods for nonsmooth equations. Oper Res Lett 20:223–228

Qi L, Sun J (1993) A nonsmooth version of newton's method. Math Program 58:353–367
Qi L, Shapiro A, Ling C (2005) Differentiability and semismoothness properties of integral functions and their applications. Math Program Ser A 102:223–248
Rivlin TJ (1990) Chebyshev polynomials: from approximation theory to algebra and number theory. Wiley Interscience, NewYork
Rubinstein A (1982) Perfect equilibrium in a bargaining model. Econometrica 50:97–110
Xu H, Chang XW (1997) Approximate newton methods for nonsmooth equations. J Optim Theory Appl 93:373–394