

University of Illinois at Urbana-Champaign

From the Selected Works of Mattias K Polborn

2009

A political-economy model of taxation and government expenditures with differentiated candidates

Stefan Krasa, *University of Illinois at Urbana-Champaign*
Mattias K Polborn



Available at: <https://works.bepress.com/polborn/20/>

A political-economy model of taxation and government expenditures with differentiated candidates

Stefan Krasa*

Mattias K. Polborn†

August 13, 2011

Abstract

We develop a model of political competition between two candidates who choose which level of taxation (and implied government spending) to propose as their election platform. Candidates differ in the amount of public good they can produce for a given level of tax revenue. Voters differ in their incomes, as well as in their preference for the public good relative to private consumption.

Even though candidates are office-motivated, they propose strictly differentiated platforms that play to their strengths of providing lots of public good, or managing limited government, respectively. Equilibrium platforms depend on the candidates' production functions and on the preferences of (some) voter types, but – in contrast to the Downsian model – not on the distribution of voter types in the population. We also derive comparative statics and empirical predictions.

JEL Classification Numbers: D72, D60.

Keywords: Differentiated candidates, issue ownership, policy divergence, fiscal policy.

*Department of Economics, University of Illinois, 1407 W. Gregory Dr., Urbana, IL 61801 USA. E-mail: skrasa@uiuc.edu.

†Department of Economics and Department of Political Science, University of Illinois, 1407 W. Gregory Dr., Urbana, IL 61801 USA. E-mail: skrasa@uiuc.edu.

1 Introduction

A standard economic interpretation of the median voter model is that the policy space is the set of possible tax rates. Tax revenues are used to produce a public good, measured by how much money is spent on its provision (see, e.g., Persson and Tabellini (2000)). Voters differ either in their income, or in their preferences over private and public goods. It is well known that, in this setup, both candidates propose the level of taxation that corresponds to the ideal point of the “median voter” (for example, the agent with the median income, if voters have the same utility function and only incomes differ).

It has long been recognized in the political economy literature that candidates for political office differ in their “ability to deliver” public goods for citizens. However, the standard way of modeling these ability differences is as additive “valence” that is appreciated by all citizens, and is independent of the policy proposed by the candidate (see, e.g., Stokes (1963), Adams (1999), Ansolabehere and Snyder (2000), Groseclose (2001)). This modeling applies best if we think of valence as the charisma or physical appearance of the politician (and citizens prefer any policy to be delivered by a charismatic or good-looking candidate). However, it is less appropriate if we think of “valence” as the general competence of the candidate to manage the provision of public goods by the state, because how important a candidate’s ability to manage is for voters depends on the amount of resources used by the government for the production of public goods, and thus on the policy proposed by the candidate. Building on the framework introduced by Krasa and Polborn (2009), our model analyzes such interactions between candidate characteristics and platform choice.

We assume that the two candidates have access to different production technologies. Thus, the level of public goods provided by the government does not only depend on the amount of taxation, but also on the identity of the winning candidate. Specifically, we assume that one candidate has an advantage at running a small government, while the other one has a higher marginal productivity. The relevant parts of the platforms for voters are the taxation level and the amount of public good produced. Voters are differentiated by income and preferences for the public good (relative to private consumption), and the distribution of voter preferences is uncertain for candidates.

We identify a condition that guarantees the existence of a unique pure strategy equilibrium. In this equilibrium, both candidates focus on their respective strengths: The candidate who has an advantage at running a small government proposes a smaller tax rate than his competitor who has an advantage in producing more units of the public good. We also characterize how changes in the technologies of the two candidates affect the positions candidates take. Specifically, improvements in the technology of one candidate may increase or decrease this candidate’s equilibrium tax rate, but unambiguously induce his competitor to focus even more on his strong point.

The equilibrium of our model contrasts sharply with the equilibrium in a standard model in which both candidates have the same technology. In that model, both candidates converge to the ideal policy

of the “median median”.¹ For example, a shift in the likely position of the median voter to prefer more spending would mean that candidates of both parties choose a higher spending level, and the equilibrium winning probabilities remain the same. In contrast, candidates in our model are rigid and do not change their platform in the face of shifting voter preferences. The candidate with an advantage in running a small government sticks to a small government platform, not because he is convinced that a small government is “right for the country” (after all, he is, by assumption, purely office-motivated), but rather, because adjusting his platform to propose more spending is not going to help him electorally in those states of the world where a majority of voters indeed puts a larger emphasis on public goods. The reason is that his opponent has an insurmountable advantage in these states anyway.

The rigidity of candidates in our model and their focus on their respective “strong points” captures the notion of “issue ownership” by candidates who are endowed with advantages and disadvantages in different policy fields. The political science literature on issue ownership, starting with the seminal paper by Petrocik (1996), argues that the weak candidate in a particular policy area cannot benefit by simply copying the platform of the strong candidate in this area so that candidates remain differentiated; each candidate has strong and weak areas (from the perspective of all voters), but those voters who care primarily about a candidate’s strong area are likely to support him. In contrast to our model, though, the issue ownership literature focuses on parties choosing to emphasize different policy issues, rather than proposing different policies within the same policy space.

Our model provides an explanation for the puzzling empirical findings of Lowry, Alt, and Ferree (1998) who analyze the electoral effects of fiscal policy in the term before the election on gubernatorial elections between 1968 and 1992. They show that voters expect higher spending from Democratic rather than Republican governors, and that they appear to punish (reward) Republican governors for higher (lower) than expected spending, while they punish (reward) Democratic governors for lower (higher) than expected spending. Thus, a majority of voters appear to *appreciate* their politicians from both parties to be a bit more extreme than they expect them to be, and this result obtains in the equilibrium of our model.

Our model is built on the general framework provided by Krasa and Polborn (2009), in which candidates have exogenously differentiated characteristics and choose policy to maximize their winning probability. Voters have general preferences over candidate characteristics and policy. When voter preferences satisfy a property called *uniform candidate ranking* (UCR), policy convergence arises (generically) in equilibrium (even if candidates have differentiated characteristics). A voter has UCR preferences if and only if the voter’s preferred candidate when both candidates choose the same policy is independent of which particular policy the two candidates converge on. If voters have non-UCR preferences, then there can be policy divergence in the pure strategy equilibrium. While most models in the existing literature have voters with UCR preferences, voter preferences implied by the differentiated public good produc-

¹The state of the world affects the position of the median voter. Ordering the states by their induced median, the “median median” is the position of the median voter in the median state of the world.

tion possibilities of candidates in our model here do not satisfy UCR, as each candidate has a range of potential policies in which he is more productive than his opponent. This is the fundamental reason for equilibrium divergence in the present paper.

Our paper contributes to the large literature on candidate convergence and divergence in the spatial model of candidate competition. Some papers argue that candidates may diverge even though this decreases their winning probability, because they care about the implemented policy (see, e.g., Wittman (1983), Calvert (1985), Roemer (1994), Martinelli (2001)). In contrast, in our model, divergence may increase a candidate's probability of winning. Other models obtain policy divergence with office-motivated candidates with incomplete information among voters about candidate characteristics (e.g. Callander (2008)) or among candidates about the position of the median voter (Castanheira (2003), Bernhardt, Duggan, and Squintani (2006)). Another branch of literature on divergence with office motivation, which is less directly related to our paper, explains policy divergence as entry deterrence by two dominant parties (e.g., Palfrey (1984), Callander (2005)).

Our model analyzes the effects of competence differences between candidates, i.e. differentials in their ability to successfully implement certain policies. Most of the existing literature models this phenomenon through an additive "valence" term, giving one of the candidates an advantage over his opponent, no matter which policy is supposed to be implemented (e.g. Ansolabehere and Snyder (2000), Groseclose (2001)). In contrast, each candidate in our model has a specialization advantage – the Republican's advantage is to implement a small government while the Democrat is better-equipped to handle the provision of a large amount of public goods.

Soubeyran (2009) and Krasa and Polborn (2010) analyze settings in which candidates allocate a given budget between two different public goods such as domestic spending and military spending.² Candidates differ in how much of each public good they can produce for a given level of expenditures; while one candidate has an advantage in procuring domestic public goods, the other one has an advantage in producing national security. While some features of the equilibrium, notably the focus of the candidates on a certain cutoff voter type when choosing their platforms, are very similar, the substantive setting differs markedly: The present one features only one public good, but the size of the budget is endogenously determined by the candidates' platform. Tax policy is obviously one of the most important policy areas for economists, and our paper contributes to an important empirical literature (e.g., Lowry, Alt, and Ferree (1998)) that aims to understand the determinants and consequences of tax policy.

Finally, the probabilistic voting literature (e.g., Hinich (1978), Lindbeck and Weibull (1987), Lindbeck and Weibull (1993), Coughlin (1992), Dixit and Londregan (1995), Banks and Duggan (2005)) shares with our paper the feature that voters care both about candidates' unchangeable characteristics and their flexible policy positions. However, voter utility in probabilistic voting models is, like in va-

²In the same general area, there are some papers that analyze settings in which candidates have some (differentiated) fixed characteristics and choose a policy platform, and voters care about the combination of fixed characteristics and the candidates' policies (e.g., Adams and Merrill (2003), Callander (2008), Kartik and McAfee (2006) and Jensen (2009)).

lence models, additively separable, and so there is no interaction between a candidate's fixed characteristics and his policies. Consequently, candidates choose the same policy platform in the equilibrium of probabilistic voting models, and equilibrium voter choices are only determined by their ideological (i.e., non-policy) preferences.

We proceed as follows. The next section introduces the model. The main theoretical results are presented in Section 3, and we also discuss some empirical implications of the model there. Section 4 concludes. Some proofs are in the Appendix.

2 The Model

Individual voters have preferences over consumption bundles that consist of one private good x and one public good g .³ There is a continuum of citizens whose type is given by (θ, m) , where θ is the preference type and m is income. Preferences are given by utility functions of the form $u_\theta(x, g) = x + h(\theta)w(g)$, where $h(\cdot)$ is strictly increasing and positive, and $w(\cdot)$ is strictly increasing and strictly concave. The distribution of voters is given by the cumulative distribution function $\Psi_\omega(\theta, m)$, where $\omega \in \Omega$ is uncertainty that is revealed to all players ex-post. We denote the distribution of ω by μ , and assume that average income, \bar{m} is independent of ω , i.e., $\bar{m} = \int m d\Psi_\omega(\theta, m)$ for every ω .⁴

Two candidates, $j = 0, 1$, compete in an election. Candidates are office-motivated: They receive utility 1 if elected, and utility 0 otherwise, independent of the implemented policy. Candidate j 's policy platform is given by a tax rate t_j , so that total tax revenue is $t_j\bar{m}$, if candidate j is elected. Tax revenues are used to pay for government fixed cost and for the provision of a public good.

Candidates are differentiated in their ability to manage government by transforming tax revenue into services that voters ultimately care about. Specifically, we assume that candidate j has to cover fixed cost b_j , and is endowed with a technology $f_j: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ for public good production. By having two dimensions along which candidates differ, we avoid the uninteresting case where one candidate is strictly inferior to the other. In particular, we analyze situations in which candidate 0 has an advantage with respect to fixed costs, while his opponent has the advantage of having a higher marginal product in public good provision. Thus, candidate 0 is better at running a small government, while candidate 1 would be preferable given large government expenditures. Formally,

Assumption 1 *The candidates' production functions satisfy the following conditions:*

1. *Strictly increasing and concave: i.e., $f'_j > 0$, $f''_j \leq 0$, for $j = 0, 1$.*
2. *Single crossing property:*

³One can think of the private good x as a composite good. It is easy to generalize our model to one where individuals have (possibly different) utility functions over several private goods, as long as these private utility functions are homothetic.

⁴We discuss and relax this assumption in Section 3.4.

- (a) *Per-capita fixed costs satisfy $0 \leq b_0 < b_1$.*
- (b) *Candidate 1's marginal product strictly exceeds that of candidate 0: $f_1'(z) > f_0'(z)$ for all $z \geq 0$.*
- (c) *$f_0(\bar{m} - b_0) < f_1(\bar{m} - b_1)$.*

The timing of events is as follows:

Stage 1 Candidates $j = 0, 1$ simultaneously announce tax rates $t_j \in [0, 1]$.

Stage 2 Nature draws ω , which determines the distribution of voter preferences. Each citizen votes for his preferred candidate, or abstains when indifferent.⁵ The candidate with a majority of votes wins, collects taxes and provides the public good.

2.1 Discussion of modeling assumptions

The key innovation of our model is that the two candidates have different production possibility sets. While almost all of the existing literature either assumes that different candidates can choose the same policies (all models in the Downsian tradition) or that they can only propose their most preferred policy (the citizen candidate model), the assumption that different candidates have different production possibilities appears eminently reasonable. Economists agree that workers or firms differ in their productivities, and this fact is evident as output can easily be measured in many private sector occupations. In contrast, the “output” of politicians in terms of public good production is significantly more difficult to measure, and thus it is tempting to use expenditures on inputs as a proxy measure for the quantity of the public good supplied. However, in reality, citizens derive utility, for example, from the quality of education in state schools and not *per se* from the money spent on education. Thus, when two competing candidates propose to spend the same amount of money on schools, this does not mean that both of them would produce the same quality of service for citizens if elected. Our model formalizes this notion.

There are several different interpretations of the candidates’ differentiated production possibilities. First, there is a widespread notion that Republicans have an advantage when it comes to running a small government. For example, Egan (2008) demonstrates that Republicans have a long-run public opinion advantage over Democrats on the issue of “taxes”, while simultaneously a majority of people say that they trust Democrats more than Republicans on large expenditure issues such as education and health care. Of course, it is not straightforward to interpret what these opinion poll results actually mean, as revenues and expenditures are two sides of the same coin.⁶ Our preferred interpretation of these opinion

⁵If a voter has a strict preference, then it is a weakly dominant strategy to vote for the preferred candidate. If a voter is indifferent, he could in principle vote for any candidate or abstain, and we assume that he abstains.

⁶Possibly, (some) people just want to say that they would most like to have a Democratic (i.e., large) level of spending on issues such as education and health care, while being only lightly taxed (as under Republicans). Of course, such a “can I have my cake and eat it, too”-attitude would not be a meaningful political preference in a world of limited resources.

polls is therefore that (many) people think that the advantage of a Republican government is that it is better in taking care of taxpayer dollars by trimming government spending to a minimum, a task in which Democrats may be hampered, for example by their connections to unions of government workers. On the other hand, Democrats are preferable for delivering a high level of public good service (think of the Bush administration's handling of hurricane Katrina).

A difference between political parties can also arise as a consequence of specialization on different types of expenditure projects: Republicans may be specialized in the efficient provision of services such as law enforcement that are "basic" in the sense that every government – whether Democratic or Republican – has to provide, while the Democrats' efficiency advantage lies in the provision of "optional" services (i.e., services that could, but need not be provided by the government) such as, for example, government provision of health care.

An alternative interpretation that does not require any ex-ante differences between the political parties is that incumbency changes the production possibility set of a politician: It appears plausible that learning by doing is important in the political sector, and so the marginal productivity of the incumbent should, in general, be higher than that of the challenger. For example, the incumbent may have learned to better identify new worthwhile public good projects. However, incumbency may also lead to entrenchment, so if the next office holder were charged with reducing bureaucracy and government spending, it may well be the case that the challenger is better able to achieve this objective.

Finally, it is useful to interpret the specific way how we model the difference in production possibility sets. Fundamentally, we want our model to capture the notion that a low level of output is cheaper for society if produced by candidate 0 than by candidate 1, while the reverse holds for a high level of output. Our approach that distinguishes candidates by their "fixed costs" and "marginal productivities" is a simple way to achieve this objective. However, while marginal productivity is readily interpretable, it is probably less useful to think of "fixed cost" literally as fixed expenditures for government services that do not create any useful output, but rather as a shift parameter.⁷

While we do not detail in our model what ultimately causes the difference in the production possibilities of the two candidates, we discuss in the conclusion how our model can be embedded in a larger model in which parties choose candidates; we will argue that candidates with differentiated production possibilities are quite likely to arise endogenously in such a model.

⁷For an analogous interpretative problem, consider the macroeconomic consumption function $C = a + bY$. While b is naturally interpreted as the marginal propensity to consume, it is not useful to think of a as "the amount that society would consume if there was no economic activity."

3 Analysis of the Model

3.1 Characterization of Equilibria

We start by providing some intuition for the equilibrium of the game, and will then proceed to a more formal statement of our results. Tax revenue under candidate j 's plan is $t_j \bar{m}$. The net revenue after fixed costs will generate $f_j(t_j \bar{m} - b_j)$ units of the public good. For $t_j \geq b_j / \bar{m}$, define

$$W_j(t_j) = w(f_j(t_j \bar{m} - b_j)). \quad (1)$$

This is the common component of utility from the public good. Remember, however, that citizens also differ in the term $h(\theta)$ that multiplies $W_j(t_j)$.

Using the functions $W_0(t)$ and $W_1(t)$ in Figure 1, we can provide an intuition for the nature of the equilibrium. Remember that $b_0 < b_1$ so that $W_0(t) > W_1(t)$ for low levels of t . The larger marginal productivity of candidate 1 implies that W_1 increases more steeply than W_0 . Because of the single crossing property, Assumption 1, W_0 and W_1 intersect at a unique tax rate \bar{t} .

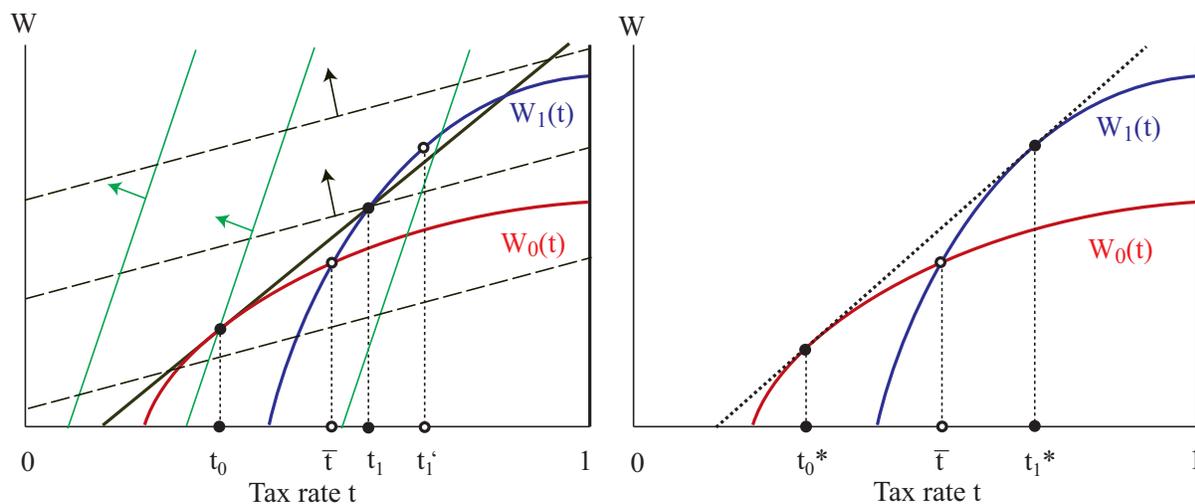


Figure 1: Characterization of Equilibrium Policies t_0^*, t_1^*

Consider first what would happen if both candidates were to propose the same tax rate t : If $t < \bar{t}$, then *all* voters would prefer candidate 0 over candidate 1 (as he produces more of the public good, with no difference in the level of taxation). Conversely, if both candidates propose the same tax rate $t > \bar{t}$, then all voters would prefer candidate 1. Thus, there cannot be an equilibrium in which both candidates propose the same tax rate. More generally, there also cannot be a pure strategy equilibrium in which both candidates propose different tax rates t_0, t_1 where either $t_0, t_1 < \bar{t}$ or $t_0, t_1 > \bar{t}$. For example, if $t_0, t_1 < \bar{t}$, then candidate 0 could win 100% of the votes (with certainty) by selecting $t_0 = t_1$. Thus, in any pure

strategy equilibrium, each candidate must locate in a region where he dominates his competitor.

To gain more insight into the properties of equilibrium, we need to consider the preferences of voters. While the type space of voters is two-dimensional, we now show that there is a one-dimensional sufficient statistic similar to a marginal rate of substitution that determines the voter's choice. The expected utility of voter type (θ, m) if candidate j is elected is

$$(1 - t_j)m + h(\theta)w(f_j(t_j\bar{m} - b_j)). \quad (2)$$

Thus, the voter's indifference curves in the (t, W) space are linear and upward-sloping with slope

$$\frac{dW}{dt} = \frac{m}{h(\theta)} \equiv \tau(m, \theta), \quad (3)$$

and the better-direction is to the Northwest in Figure 1. Voters with a higher income m , or those with a low preference parameter θ for the public good, have relatively steep indifference curves, while those with low m and/or high θ have relatively flat indifference curves.

Suppose that candidates propose tax rates $t_0 < \bar{t} < t_1$ as in the left panel of Figure 1. A voter with the solid (steep) indifference curves strictly prefers candidate 0, while a voter with the dashed and flatter indifference curves prefers candidate 1. The third line represents the indifference curves of a voter who is indifferent between the two candidates. Note that the situation depicted in the left panel is not an equilibrium: If candidate 1 deviates to t'_1 , then the previously indifferent voter is now solidly in the camp of candidate 1 as well as voters with slightly steeper preferences who previously preferred candidate 0.

Consider the line that connects $(t_0, W_0(t_0))$ and $(t_1, W_1(t_1))$. Given t_1 , the objective of candidate 0 is to choose t_0 such that this line is as flat as possible, as this maximizes the set of types who vote for him. Conversely, the objective of candidate 1 is to make the line as steep as possible. If the indifference curve of a voter who is indifferent between candidates 0 and 1 is tangent to $W_0(t_0^*)$ and $W_1(t_1^*)$ (as indicated in the right panel), then candidate 0 cannot make the line connecting $(t_0, W_0(t_0))$ and $(t_1, W_1(t_1))$ flatter, and candidate 1 cannot make it steeper. Thus, both candidates maximize the set of types who vote for them, respectively. We refer to the indifferent voter in the right panel as the "cutoff voter", and to the slope of his indifference curve as τ^* .

We still need to consider under which conditions (t_0^*, t_1^*) is in fact an equilibrium, and under which conditions it is the unique equilibrium. First, are there any deviations from (t_0^*, t_1^*) that would increase the winning probability of the deviating candidate? By the geometric argument above, it is clear that *small* deviations strictly decrease the set of voter types who prefer the deviating candidate. For example, a small deviation by candidate 0 (in either direction) will not change the fact that those voters with the flattest indifference curves prefer candidate 1, and leads to a cutoff voter with a steeper indifference curve so that the set of voters who support candidate 0 is a strict subset of the set of voters who previously supported candidate 0. However, it is in principle conceivable that a *large* deviation such that, after the deviation, voters with flat indifference curves vote for candidate 0 and those with steep indifference

curves vote for candidate 1, is profitable. For this to happen, it would have to be possible for one of the candidates to “outflank” his opponent on that side on which his opponent is stronger. In real life, it appears implausible that, for example, a Democrat can win in such a way that he attracts all right-wingers by “out-Republicaning” the Republican candidate. In Theorem 1 below, we provide a condition that prevents such outflanking, and we discuss this condition in more detail below.

As for the second question of uniqueness, it is clear that without uncertainty about the type distribution, there are generically multiple equilibria, because one of the candidates wins with certainty (and possibly by choosing any of some set of policies), and his opponent receives a payoff of 0 no matter which policy he chooses. Assumption 2 below is a sufficient condition that guarantees that there is enough uncertainty about the distribution of voters so that both candidates have a strictly positive winning probability, and that any policy change matters with positive probability for the election outcome. As we will show, this guarantees equilibrium uniqueness.

Assumption 2 *Let τ_ω be the median of the distribution of $m/h(\theta)$ in state ω and denote the cdf of this distribution by Φ . Then Φ is strictly monotone on $[W'_0(1), W'_1(0)]$.*⁸

With a slight abuse of terminology, we talk about a median voter τ_ω in each state ω , and a “median median”, τ_m , i.e. the median of the distribution of τ_ω . Assumption 2 ensures that any slope between $W'_0(1)$ and $W'_1(0)$ could be the median in some state ω . If a candidate deviates, he loses voter types τ^* and some voters close to τ^* . By Assumption 2, these types could be pivotal for the election outcome, so that his winning probability strictly decreases. If, in contrast, Assumption 2 is violated, then it could be the case that the realized median voter is never in a neighborhood of τ^* , so that candidates could change their tax rates without affecting their payoff. In such a case, we would have multiple equilibria.

Theorem 1 below shows formally that an equilibrium exists and is unique. In equilibrium, candidate 0 proposes a strictly lower tax rate and public good provision than candidate 1. Voter behavior is characterized through a set of cutoff voter types who are indifferent between both candidates. Both candidates try to appeal to cutoff voters by choosing their most preferred policy. Most voters have a strict preference for one of the two candidates. High voter types (i.e., those with high income and/or a low preference for public goods) strictly prefer candidate 0, while low voter types strictly prefer candidate 1.

Theorem 1 *Suppose that Assumptions 1 and 2 are satisfied. Let t_0^*, t_1^* solve*

$$W'_0(t_0^*) = W'_1(t_1^*) = \frac{W_1(t_1^*) - W_0(t_0^*)}{t_1^* - t_0^*}. \quad (4)$$

Suppose that $W_0(1) \leq W_1(t_1^)$ and that $t_0^* \bar{m} \leq b_1$.*

1. *(t_0^*, t_1^*) is the unique Nash equilibrium (pure or mixed) of the voting game.*

⁸Note that we could also assume that Φ has a density that is strictly positive on $[W'_0(1), W'_1(0)]$.

2. In equilibrium both candidates maximize the utility of voters (m, t) for whom $m/h(\theta) = \tau^*$, where $\tau^* = W'_0(t_0^*) = W'_1(t_1^*)$. All voters (m, t) with $m/h(\theta) > \tau^*$ strictly prefer candidate 0, while all voters (m, t) with $m/h(\theta) < \tau^*$ strictly prefer candidate 1.
3. In equilibrium candidate 0's proposed tax rate and public good provision is strictly less than that of candidate 1, i.e., $t_0^* < t_1^*$ and $f(\bar{m}t_0^* - b_0) < f(\bar{m}t_1^* - b_1)$.

It is useful to start with two comments about the assumptions in Theorem 1. First, note that t_0^* and t_1^* are *defined as* the unique solution of an equation system that depends only on exogenous parameters (and which is relatively simple to solve; see Appendix 5.2 for an example). Thus, the assumptions that guarantee equilibrium existence do not “rely on assuming that an equilibrium exists.” Instead, the main work in the proof of the theorem is to show that (t_0^*, t_1^*) constitutes an equilibrium.

Second, we need an assumption that captures the notion that it is not possible for the Democrat to “out-Republican” the Republican candidate by proposing to slash taxes and spending to gain the votes from the highest τ -types, and similarly for the Republican. A sufficient and simple condition for this notion is that $W_0(1) \leq W_1(t_1^*)$ and $t_0^* \bar{m} \leq b_1$. It can be interpreted as requiring that there are sufficiently large differences between the production possibilities of both candidates. However, it should be noted that these conditions are considerably stronger than necessary since they ensure existence of an equilibrium independently of the distribution of the voters. Theorem 2 below instead characterizes conditions on voter type distributions under which the same strategy pair characterized above in Theorem 1 continues to be an equilibrium.

The fact that both candidates maximize the utility of a cutoff voter τ^* is reminiscent of the equilibrium in the standard model where both candidates choose their policies to maximize the utility of the median voter. However, the similarity stops here: In the equilibrium of the standard model with policy convergence, *all* voters are indifferent between the candidates, while in the present model, only a small set of voters (of measure 0) is indifferent between both candidates. Thus, our model preserves the Downsian notion that candidates in an election campaign fiercely contest a relatively small set of swing voters, but avoids the rather counterfactual prediction that *all* voters are indifferent between candidates.⁹

Policy-motivation is probably the most widely-accepted explanation in the standard model for the widespread observation that policies diverge. While this model and ours both lead to a prediction of policy divergence, there is a subtle difference in the prediction of candidates' reaction to *changes* in the perceived distribution of voter preferences. Suppose, for example, that voters develop a stronger preference for public goods (i.e., there is a first-order stochastic dominance shift in the distribution of τ_ω). In our model, this change in the distribution has no effect on the candidates' equilibrium policies in our model, but changes the winning probability in favor of candidate 1. In contrast, a shift of the preference

⁹To avoid indifference of all voters in a Downsian model, some papers such as Erikson and Romero (1990) and Adams and Merrill (2003) assume that voters have exogenous “partisan” preferences for one of the candidates. However, since these partisan preferences for candidates are independent of the policies that candidates choose, they are difficult to interpret.

distribution has a significant effect on the policy platforms chosen by policy-motivated candidates in the standard model. For example, if the shift in preferences is due to a common shock to utility functions that affects both voters and candidates in the same way,¹⁰ then both candidates' platforms shift in parallel by the same amount. Even if only voter, but not candidate, preferences shift, both platforms generically shift in the same direction. Finally, with only office-motivated candidates, a change in the expected position of the median would translate only in a change of the equilibrium policies of candidates, without affecting their equilibrium winning probabilities – exactly the opposite from the effect in our model.

The rigidity of candidates is, at least in some cases, an appealing feature of our model. Consider, for example, the political positioning of the Republican party after suffering a severe defeat in the 2006 and 2008 elections. The distribution of U.S. voter preferences appeared to have shifted in a way that favored the Democrats. However, Congressional Republicans in early 2009 almost unanimously opposed President Obama's economic stimulus package that was initially supported by a clear supermajority of voters. As argued above, this contradicts the prediction of a standard model with identical party capabilities and policy motivated candidates. In contrast, in our model, Republicans cannot successfully imitate the Democrats' policy position; sticking with their previous platform and hoping for a reversal of the preference shift is the best (from an electoral perspective) that Republicans can do in our model, and they were actually successful with this strategy in the 2010 Congressional elections.

A significant difference between standard valence models and ours concerns the two candidates' equilibrium winning probabilities and expected vote shares. In a model with convergence, candidates receive either equal vote shares (if their valence is equal), or one of the candidates wins all votes (say, in a model with convergence and ex-ante uncertain valence). Moreover, the equilibrium in our model is an ex-post stable situation in the sense that, even knowing the realization of ω , no candidate could increase his vote share by choosing a different position. In standard models with divergence (say, with policy-motivated candidates), ex-post vote shares may differ substantially, but at least one candidate could increase his vote share ex-post, and possibly ex-ante.

In order to exclude “outflanking” deviations, Theorem 1 assumes that the two candidates' production possibilities are sufficiently differentiated. This guarantees equilibrium existence for all possible distributions of voter types, in particular, for all possible distributions of the median type τ_ω . Theorem 2 below takes a different approach, by identifying conditions on the distribution of voter types that guarantee that candidates do not have a profitable outflanking deviation. In these cases, the equilibrium pair of tax rates identified in Theorem 1 remains the unique equilibrium.

Theorem 2 *Suppose that Assumptions 1 and 2 are satisfied and that τ_ω is distributed symmetrically*

¹⁰For example, suppose that some external threat arises and increases the preferred amount of defense spending (interpreted as the policy variable in a standard one-dimensional model) for all voters and candidates.

around τ_m . Let t_0^* and t_1^* be defined by (4). Let $\tau^* = \frac{W_1(t_1^*) - W_0(t_0^*)}{t_1^* - t_0^*}$ be the cutoff voter. Let

$$\hat{\tau}_0 = \max_{t \geq t_1^*} \frac{W_0(t) - W_1(t_1^*)}{t - t_1^*}, \quad (5)$$

$$\hat{\tau}_1 = \min_{t \leq t_0^*} \frac{W_0(t_0^*) - W_1(t)}{t_0^* - t}. \quad (6)$$

1. Suppose that $W_0(1) \leq W_1(t_1^*)$ and $t_0^* \bar{m} > b_1$. Then $0 \leq \hat{\tau}_0 < \tau^*$ and (t_0^*, t_1^*) is an equilibrium if and only if $\tau_m \leq \frac{\hat{\tau}_0 + \tau^*}{2}$.
2. Suppose that $W_0(1) > W_1(t_1^*)$ and $t_0^* \bar{m} \leq b_1$. Then $\tau^* < \hat{\tau}_1 \leq 1$ and (t_0^*, t_1^*) is an equilibrium if and only if $\tau_m \geq \frac{\hat{\tau}_0 + \tau^*}{2}$.
3. Suppose that $W_0(1) > W_1(t_1^*)$ and $t_0^* \bar{m} > b_1$. Then $0 \leq \hat{\tau}_0 < \tau^* < \hat{\tau}_1 \leq 1$ and (t_0^*, t_1^*) is an equilibrium if and only if $\tau_m \in \left[\frac{\hat{\tau}_0 + \tau^*}{2}, \frac{\hat{\tau}_1 + \tau^*}{2} \right]$.

Theorem 2 considers whether deviations by a candidate that appeal to his opponent's "natural" supporters are profitable. Note that any such deviation can only appeal to a strict subset of his opponent's former equilibrium supporters, because the previous cutoff voter now strictly prefers the opponent.¹¹ The conditions in Theorem 2 guarantee that the support group after such an outflanking deviation is less likely to be a majority than a candidate's equilibrium supporters.

For example, if $W_0(1) > W_1(t_1^*)$, then candidate 0 could provide more public goods than candidate 1 does in equilibrium (provided that candidate 0 chooses a sufficiently high tax rate). The optimal such deviation by candidate 0 is the one that maximizes the slope of the indifference curve of the new cutoff voter $\hat{\tau}_0$, because candidate 0 will receive the votes of all voters who have a relatively flat indifference curve (i.e., who are either poor or very keen on public goods). Similarly, the optimal deviation by candidate 1 minimizes the slope of the new cutoff voter's indifference curve. There are three different scenarios regarding the possibility of successful deviations in Theorem 2. Case 1 refers to the case that only candidate 1 can appeal to candidate 0's natural supporters, but not vice versa.¹² If the candidates play (t_0^*, t_1^*) , candidate 0 receives the votes of all voters with $\tau > \tau^*$, and candidate 1 those of all voters with $\tau < \tau^*$. Consequently, candidate 1 wins with probability $\Phi(\tau^*)$ (remember that $\Phi(\cdot)$ denotes the cumulative distribution of τ_ω). If candidate 1 deviates to appeal to fiscal conservatives (and does this in an optimal way), then he receives the votes of all voters $\tau > \hat{\tau}_1$, and thus wins with probability $1 - \Phi(\hat{\tau}_1)$. Such a deviation is not profitable if and only if $\Phi(\tau^*) > 1 - \Phi(\hat{\tau}_1)$. Given that τ_ω is distributed symmetrically around τ_m , this condition is equivalent to $\tau_m \leq \frac{\hat{\tau}_0 + \tau^*}{2}$.

¹¹This follows from the facts that the cutoff voter given (t_0^*, t_1^*) is indifferent between the candidates in a situation where both of them choose the policy that is optimal for the cutoff voter, respectively. If one candidate deviates to a different policy while his opponent still offers the cutoff voter's ideal policy, the cutoff voter must strictly prefer the opponent.

¹²Note that $W_0(1) > W_1(t_1^*)$ means that candidate 0 cannot exceed candidate 1's public good provision, and therefore cannot appeal to those voter types with the highest taste for public good provision.

Similarly, the second case in Theorem 2 refers to the case that only candidate 0 can appeal to candidate 1's natural supporters, and in the third case, both candidates could, in principle, appeal to their opponent's natural supporters. The conditions for equilibrium existence are adjusted accordingly in these two cases. In summary, (t_0^*, t_1^*) remains an equilibrium if the “median median” τ_m of the type distribution is not too far from the cutoff voter, and consequently if the winning probabilities of the two candidates are sufficiently close. Of course, how asymmetric the winning probabilities can be for (t_0^*, t_1^*) to be an equilibrium depends on parameters.

Consider the following example, in which we choose utility and production functions so as to match the values reported by Lowry, Alt, and Ferree (1998) for state expenditures as a fraction of state income, of about 10% for Republicans and 12% for Democrats.¹³ This observed policy difference is quite substantial, and it might seem that one could only generate such differences if there are implausibly large productivity differences between candidates. However, we will see that this is not the case, as the production functions whose values are only slightly differentiated in the relevant range. We also use the example to show that the equilibrium continues to exist even in a relatively asymmetric situation.

Example 1 Suppose that candidate j 's production function is of the form $f_j(z) = a_j z$, with $a_0 = 1$ and $a_1 = 1.2$, and let $b_0 = 0.12$ and $b_1 = 0.15$. The voters' form of utility functions is given by $w(g) = -e^{-cg}$, with $c = 3$. Furthermore, assume that the average income is $\bar{m} = 2.7$.

It is easy to show that $t_0^* = 10.0\%$ and $t_1^* = 12.1\%$ (details of the calculations are in the Appendix). Around these equilibrium values, the difference between the two candidates' production possibilities is relatively small. For example, with a tax rate of $t = 0.10$, candidate 0 can produce $(0.1 \cdot 2.7 - 0.12) \cdot 1 = 0.15$ units of the public good, while candidate 1 could produce $(0.1 \cdot 2.7 - 0.15) \cdot 1.2 = 0.144$ (i.e., 4 percent less). With a tax rate of $t = 0.12$, candidate 0 could produce $(0.12 \cdot 2.7 - 0.12) \cdot 1 = 0.204$ units of the public good, while candidate 1 can produce $(0.12 \cdot 2.7 - 0.15) \cdot 1.2 = 0.2088$ (i.e., 2.3 percent more).

Given (t_0^*, t_1^*) , the cutoff type is $\tau^* = 5.156$. In the Appendix, we also calculate the optimal outflanking deviations as defined in Theorem 2 to be $\hat{t}_0 = 14.4\%$ and $\hat{t}_1 = 8.2\%$, so that $\hat{\tau}_0 = 3.618$ and $\hat{\tau}_1 = 7.513$.

Consider a deviation by candidate 0. In equilibrium, candidate 0 receives the votes of all types $\tau > 5.156$, while after deviating to \hat{t}_0 , all citizens $\tau < 3.618$ support the candidate. In order to understand the impact of the change of the cutoff type, recall from (3) that $\tau = m/h(\theta)$. Suppose that under (t_0^*, t_1^*) a citizen type (θ, m_θ) was indifferent between the candidates. Then candidate 0, the Republican, receives the support of everyone who is wealthier (i.e., all (θ, m) with $m > m_\theta$) while candidate 1, the Democrat, is supported by everyone who is poorer (i.e., all (θ, m) with $m < m_\theta$). If the Republican outflanks the Democrat, then he receives the votes of all sufficiently poor voters, i.e., all (θ, m) with $m < m'_\theta$, where $m'_\theta = (3.618/5.156)m = 0.702m$. That is, after the deviation the Republican receives the support of all citizens who have at most about 70% of the income of the previous cutoff voter. Under a uniform

¹³We should stress though that this is a numeric example and not a calibration exercise — there are more free parameters than targets to match.

distribution of (θ, m) ,¹⁴ the Republican would thus receive 30% fewer votes after a deviation than the Democrat did under (t_0^*, t_1^*) . Moreover, if the distribution of types is single peaked and has less mass in the tails, the reduction relative to the Democrat's vote under (t_0^*, t_1^*) is even larger. Thus, for outflanking to be successful in a particular state of the world, the Democrat would have to win with at least 65% of the votes, and for outflanking to be profitable in expectation, under (t_0^*, t_1^*) , those state in which the Democrat wins with more than this threshold number of votes would have to be more likely than those in which the Republican wins.

Similarly, outflanking is unlikely to be optimal for the Democrat. If the Democrat chooses \hat{t}_1 then he receives the voters of all types (θ, m) with $m \geq (7.513/5.156)m_\theta = 1.457m_\theta$, i.e., of everyone who is at least 46% wealthier than the previous cutoff voter. Again, under a uniform distribution of (θ, m) , the Democrat would receive 46% fewer votes than the Republican did in the original equilibrium, and this number is a lower bound if the distribution is single peaked. Thus, a deviation by the Democrat is not optimal unless the Republican wins under (t_0^*, t_1^*) sufficiently often with a vote share of over 70%. ■

An interesting question is whether ex-post a majority of voters would like the winning candidate to change his policy relative to his announced platform. If the Democrat wins, then the realized median voter has a higher demand for public goods than the cutoff voter. Thus, a majority of voters would benefit if the Democrat were to choose a higher tax rate than the one that maximizes the utility of cutoff voters. Conversely, if the Republican wins, then the median voters would like less government spending than the cutoff types, so that a majority prefers a strictly smaller tax rate than t_0^* . Unless the election happens to end in a tie (which is very unlikely from an ex-ante perspective), a majority of voters would therefore like the winning candidate to adopt a *more extreme* policy than he announced as the platform.¹⁵

Note that this result is exactly the opposite from what happens in standard one-dimensional models with policy divergence. Consider, for example, the citizen-candidate model of Osborne and Slivinski (1996). In their model, there exists (for large parameter sets) an equilibrium in which two candidates located symmetrically at opposite sides of the median voter run against each other, and each wins with probability 1/2. Independent of whether the right-wing or left-wing candidate wins the election, it is always true that a majority of voters would like the winning candidate to implement a more moderate policy (i.e., a policy that is closer to the median).¹⁶

It is interesting to consider the empirical findings of Lowry, Alt, and Ferree (1998) in the context of these different predictions. They analyze the electoral effects of fiscal policy (in the period before the election) on gubernatorial and legislative elections between 1968 and 1992. They find that voters expect

¹⁴I.e., m would be uniformly distributed between 0 and $2\bar{m}$.

¹⁵A formal proof of this claim can be made along the lines of Proposition 3 in Krasa and Polborn (2010) and is available from the authors upon request.

¹⁶The same result applies in models where policy divergence is due to entry deterrence (Palfrey (1984), Callander (2005)) or policy motivation with uncertainty about the position of the median (Calvert (1985)).

higher spending from Democratic rather than Republican governors, and that they appear to punish Republican governors for higher than expected spending, while they punish Democratic governors for lower than expected spending. For example, suppose that control of the state house and the governorship is unified (i.e., belongs to the same party) and that the government unexpectedly increases spending by 2.5%. The empirical results of Lowry, Alt, and Ferree (1998) imply that, if Democrats are in control, they will gain about 2.6% in the next gubernatorial election. In contrast, Republicans in the same situation would lose about 11.4%. Of course, an unanticipated decrease in spending has exactly the opposite effects on both parties, making Republicans more and Democrats less popular. In other words, voters appear to *appreciate* politicians from both parties to be a bit more extreme than they expect them to be, which is consistent with the prediction of our model, but inconsistent with the standard model discussed above.

3.2 Comparative Statics

We now discuss how changes of the technology parameters of the candidates affect equilibrium tax rates and public good provision levels. Theorem 3 considers changes in the fixed costs of government.

Theorem 3 *Suppose that (t_0^*, t_1^*) is an equilibrium.*

1. *If b_0 increases then:*

(a) *t_0 increases if $-W_0''(t_0^*)/W_0'(t_0^*) > 1/(t_1^* - t_1^0)$ and decreases otherwise. Candidate 0's public good production decreases.*

(b) *t_1 as well as candidate 1's provision of public good decrease.*

2. *If b_1 increases then:*

(a) *t_0 and candidate 0's public good provision increase.*

(b) *t_1 and candidate 1's public good provision increase.*

Intuitively, an increase in b_i weakens candidate i so that his opponent, candidate j , is now able to appeal to more voter types. As a consequence, the cutoff type shifts and both candidates change their equilibrium platform so as to appeal to the new cutoff type. For example, suppose that candidate 1's (the Democrat's) fixed cost increases. The Republican is now in a better position to compete for voters who care more about public good provision, i.e., the indifference curve of the new cutoff voter τ^* is flatter. Thus, it is optimal for the Republican to increase taxes and public good provision. The Democrat also has to compete for a cutoff voter with a stronger preference for public goods. Moreover, he needs to increase taxes in order to cover his increased fixed cost. Both effects increase the tax level proposed by the Democrat.

Suppose now that b_0 increases, i.e., the Republican is weakened. The new cutoff voter cares less about public goods than the previous one, and thus the Democrat lowers his proposed tax rate and public good provision. For the Republican, however, the effect on the tax rate is ambiguous: The shift in the cutoff voter alone would also induce the Republican to lower his proposed tax level. However, since the Republican needs more money to pay for fixed costs, his public good provision level is reduced, which increases the marginal benefit from higher taxes. These effects move in opposite directions, and which one dominates depends on the condition on the curvature of W detailed in the theorem. In contrast, the level of public goods supplied by the Republican always decreases when b_1 increases.

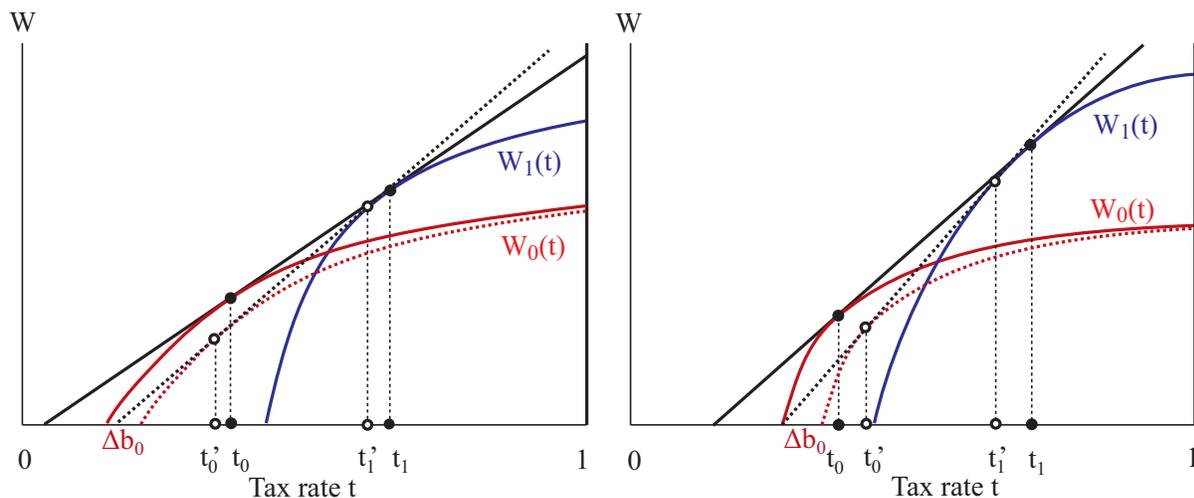


Figure 2: Illustration of Theorem 3

Figure 2 illustrates Theorem 3. When b_0 increases by Δb_0 , the curve $W_0(t)$ shifts everywhere to the right by Δb_0 (of course, the vertical distance decreases because the curves are concave). In both panels, the previous cutoff voter, represented by the solid indifference curve, now strictly prefers candidate 1. The new cutoff voter is represented by the dashed indifference curve. In the left panel, where $W_0(t)$ has a relatively low curvature at t_0 , candidate 0 adjusts by lowering taxes. In contrast, in the right panel, where $W_0(t)$ has higher curvature, candidate 0 responds by increasing taxes. In both panels, public good provision by both candidates strictly decreases.

It is easy to find numerical examples in which t_0^* increases and others where t_0^* decreases, as b_0 increases. For example, for $w(g) = \sqrt{g}$ and production functions $f_j(x) = a_j x$, one can show that $\frac{\partial t_0^*}{\partial b_0} = \frac{a_1 - 2a_0}{m(a_1 - a_0)}$. Thus, the sign of $\frac{\partial t_0^*}{\partial b_0} > 0$ depends on the relationship of a_0 and a_1 . In the Appendix, we compute an example with preferences $w(g) = -e^{-cg}$. Equation (45) shows that, in this case, t_0^* is always decreasing in b_0 .

Next, Theorem 4 analyzes the effect of changing a candidate's marginal productivity. To do this, we consider production functions with constant marginal productivity, $f_j(x) = a_j x$, and change a_j .

Theorem 4 *Suppose that (t_0^*, t_1^*) is an equilibrium and that $f_j(x) = a_j x$ for $j = 0, 1$.*

1. *If a_0 is increased then:*

- (a) t_0 *decreases if $-\frac{\bar{m}t_0 - b_0}{\bar{m}t_1 - b_0} \frac{W_0''(t_0^*)}{W_0'(t_0^*)} > 1/(t_1^* - t_1^0)$ and increases otherwise. Candidate 0's public good production increases.*
- (b) t_1 *as well as candidate 1's provision of public good increases.*

2. *If a_1 is increased then:*

- (a) t_0^* *and public good provision by candidate 0 decreases.*
- (b) *Public good provision by candidate 1 decreases if and only if $t_0^* \bar{m} > b_1$. t_1^* decreases if and only if $\frac{W_0''(t_0^*)}{W_0'(t_0^*)} < \frac{t_0^* \bar{m} - b_1}{t_1^* \bar{m} - b_1} \frac{1}{t_1^* - t_0^*}$. In particular, if $t_0^* \bar{m} > b_1$ then t_1^* always decreases.*

The results in Theorem 4 resemble those in Theorem 3. In particular, if candidate j 's productivity increases, then his opponent moves to a more extreme position in order to appeal to the new cutoff voter. Thus, if a_0 increases, then the Democrat increases the proposed tax level (as well as public good provision), and if a_1 increases, then the Republican decreases taxes and spending.

The effect on the candidate's own equilibrium tax level is ambiguous. If the Republican's productivity a_0 increases, then the new cutoff voters τ^* have a higher preference for public goods than before (i.e., flatter indifference curves). To appeal to the new cutoff type, the Republican would like to supply more of the public good, which would require a higher t_0^* at his old productivity level. Second, because the Republican is more productive, he can provide the same level of output with a lower t_0^* . Which effect dominates depends on the curvature of $W(\cdot)$, as well as the distance between t_0^* and t_1^* . This effect is similar to the ambiguous effect of a change in b_0 and t_0 , discussed in Theorem 3. Again, the condition for increase or decrease depends on the curvature of $W(\cdot)$. Note, however, that the additional term relative to the condition in Theorem 3, $\frac{\bar{m}t_0 - b_0}{\bar{m}t_1 - b_0}$, is smaller than 1. Thus, $\frac{\partial t_0^*}{\partial a_0} < 0$ implies that $\frac{\partial t_0^*}{\partial b_0} > 0$. Therefore, it is more likely that a decrease in fixed costs b_0 results in a lower tax rate t_0^* than that an increased productivity a_0 has this effect.

Now suppose that candidate 1's marginal productivity increases. A priori, one might have suspected that this induces candidate 1 to increase the activity level, as *every voter type* now has a higher preferred level of public good provision under candidate 1. However, Theorem 4 shows that candidate 1 lowers his tax rate and even provides *less* of the public good, whenever tax revenue by candidate 0 exceeds candidate 1's fixed costs. The Democrat lowers taxes, because the higher a_1 makes it possible to appeal to voters who have a lower preference for the public good. In principle, it could still be the case that the Democrat provides the same or a higher level of public goods (since he is more productive). However, as Theorem 3 indicates public good provision also decreases, provided that $t_0^* \bar{m} > b_1$.

3.3 Application of Comparative Static Results: The Marginality Hypothesis

There is a long-lasting discussion in political science about the extent to which candidates or parties in close races behave differently from those in safe races. The traditional view is that (in particular, incumbent) candidates in close elections choose positions that are more moderate and more responsive to their constituents than those whose (re)election is more secure. This view was dubbed the “marginality hypothesis” by MacRae (1952) (see also Huntington (1950)). Fiorina (1973) provides a discussion of this early literature, as well as an empirical analysis that refutes the marginality hypothesis.

More recently, Groseclose (2001) provides an influential theoretical model that provides an explanation for the contradictory results in empirical studies of the marginality hypothesis.¹⁷ In his model, candidates are policy-motivated and have differential “valence” that enters additively in all voters’ utility functions. The median voter’s ideal point is unknown, and the candidates’ ideal positions are located symmetrically around the expected median. In the benchmark case where no candidate has an advantage, the two candidates locate symmetrically around the expected median. When the valence of one of the candidates increases, Groseclose shows that his equilibrium position initially becomes more moderate before eventually (i.e., for sufficiently high valence advantage) becoming more extreme. In contrast, the disadvantaged candidate always becomes more extreme as his opponent’s valence increases.

While the parameters of the candidates’ production functions in our model are not directly observable for the empirical researcher, they correlate in obvious ways with election outcomes (or approval ratings for an incumbent): *Ceteris paribus*, a more productive candidate receives more votes in the election. An increase in a_1 or a decrease in b_1 moves the cutoff voter τ^* in candidate 1’s favor and therefore increases his (the Democrat’s) support and approval. Since Theorems 3 and 4 prove that spending decreases as b_1 decreases or a_1 increases, our model implies that a Democrat’s vote share (or approval rating) and spending are negatively correlated. Similarly a Democrat may also win with a high margin if he faces a relatively incompetent Republican (high b_0 and/or low a_0); again, this scenario shifts the cutoff voter in candidate 1’s favor, and the spending that is optimal for the new cutoff voter decreases. In contrast, for Republicans decreasing b_0 or increasing a_0 has an ambiguous effect. Therefore it appears plausible that the correlation between spending and approval ratings for Republicans are weaker than for Democrats.

Thus, our model has the following testable implications:

1. For a Democratic executive, spending is negatively correlated with the politician’s approval rating or vote share.
2. For a Republican executive, the correlation between spending and vote share/approval rating is weaker than for a Democrat, and its sign is theoretically ambiguous.

¹⁷Similar to our model, Groseclose (2001) is more applicable to an executive who has more direct control over spending rather than a legislator.

We do not know of any empirical study of these predictions, but such studies should be feasible. In particular, an advantage for empirical work in our model framework is that the policy choice variable t (or spending for public goods) is more easily observable than the candidates' "positions" in an ideological left-right spectrum, which are often more difficult to measure convincingly. The main complication of a serious empirical study in our framework relative to the work of Lowry, Alt, and Ferree (1998), in which causation only runs from fiscal policy to election outcomes, is that our model indicates that a candidate's unobserved competence affects both his electoral success and the fiscal policy he chooses. Another interesting aspect of our model is that the marginality hypothesis applies in different ways to Republican and Democratic candidates. In contrast, all models based on the standard setup with undifferentiated candidates are, by construction, symmetric.

3.4 Uncertainty about the distribution of income

In our model, the state of the world ω may affect the distribution of income, but average income is known to both candidates and voters. A fundamental change in the analysis would only be required if individual voters receive private signals on the average income \bar{m} and thus disagree on the precise way in which tax rates translate into public good provision by the two candidates. In such a world, voters would also have to consider that their vote is important only when they are pivotal, and being pivotal may contain essential information about the likely distribution of income in society and thus about the relation between tax rates and public good provision. Our assumption that \bar{m} is common knowledge is meant to exclude these complications in order to focus on the central novel features of the model.

It is easy to extend our model to a setting where candidates and voters are uncertain about the average income, as long as there is no asymmetry of information. For example, suppose that in addition to learning their own income, all citizens as well as candidates receive a signal about average income \bar{m} . Citizens and candidates update their prior using the signal to get a distribution ν of \bar{m} , which is the same for everyone (i.e., there is symmetric information in the posterior state). We can now define

$$W_j(t_j) = \int w(f_j(t_j\bar{m} - b_j)) d\nu(\bar{m}),$$

and all our results are sustained.

Also, it is possible to interpret uncertainty about the income distribution as uncertainty only about the distribution of income among *voters*. In this interpretation, the aggregate income in the economy (which determines how tax rates translate into government revenue and thus, into public good provision) is either known to both voters and candidates, or symmetrically unknown such as in the last paragraph. However, since a substantial proportion of the electorate abstains in many elections, it is quite conceivable that candidates are uncertain about the characteristics of those citizens who choose to vote: There may be substantial uncertainty about the income distribution among voters, while \bar{m} for all citizens, which determines tax revenue in the economy, is much better known. In such a scenario, our results are

sustained.

3.5 Relation between income and vote choice

While rich voters *ceteris paribus* have steeper indifference curves and are therefore more likely to vote for the candidate who proposes a lower tax rate, the voting decision also depends on the taste parameter for public goods, θ . Thus, voting is not characterized by an income threshold such that every voter with an income higher than the threshold necessarily votes for the low tax candidate. Moreover, our model admits any arbitrary joint distribution of m and θ . Thus, even if we were to observe that rich people were on average more likely to vote for the *high-tax* candidate than poor people, this would be consistent with our model. This is a desirable property of our model, because in reality, the empirically observed correlation between income and voting for Republicans is certainly not perfect.¹⁸

3.6 The role of ideology

Our model departs from some economists' favorite explanation of policy divergence, by assuming that it is driven by different *capabilities* of candidates rather than by different *ideologies*.¹⁹ We have several reactions to this.

First, note that the “evidence” for ideological differences driving policy divergence is only indirect. It is easy to believe that ideological differences exist between politicians, because they exist between voters. However, what is almost impossible to observe directly is whether and to what relative extent ideology and the desire to win office shape the positions that politicians take.²⁰ For all that we know, the “indirect evidence” for the proposition that ideological differences between candidates drive policy divergence is just that policy divergence exists and some form of policy motivation is considered necessary to generate this result in the standard model, and thus considered “implied” by the observation of divergent platforms.

Similarly, we would argue that the supposition that politicians differ in their ability is a priori not implausible, as we can observe that people in general differ in their capabilities. This said, just as with ideological policy differentiation models, our primitive assumption of differential capabilities to implement different policies, is extremely difficult or impossible to verify or falsify directly. As mentioned in the discussion of the model, the “output” of politicians is usually not easy to quantify. Moreover, we only

¹⁸See Ansolabehere, Rodden, and Snyder (2006), Gelman, Shor, Bafumi, and Park (2008) and Krassa and Polborn (2011) for an analysis of the relation between income and voting behavior.

¹⁹Indeed, some people are very dogmatic about this point, almost as if to demonstrate that *their* ideology matters, and so should politicians'. Of course, we prefer to interpret this behavior more as indicative of a difference between their capabilities and ours.

²⁰There is certainly anecdotal evidence that politicians sometimes change positions because they believe that their new position is electorally more successful. However, there is no way to tell how frequent this type of behavior is in the larger population of politicians. Moreover, there might be completely office-motivated politicians whose environment does not change; evidently, observing that they do not change positions does not prove that they are only ideologically motivated.

observe how one of the two candidates (namely the election winner) does in office, and we only observe one point on his production possibility set (his implemented policy), rather than the whole set of plans that are feasible for the election winner.

Second, we do see a role for policy motivation within more elaborate versions of our theory. Our model takes the production possibility sets of candidates as exogenously given. However, it is conceivable that our model can be expanded to develop a long run theory that explains *why* candidates differ in capabilities, and ideology may play an important role for that. A young aspiring politician who ideologically prefers a substantial role of government has potentially a stronger incentive to invest in human capital that is specifically useful for managing a larger role of government than another ideological type who prefers a more limited role of government.

Third, we should stress that the logical antithesis to our assumption that candidates have *different* production possibilities is that they have *exactly the same* production possibilities (and not that candidates have “more-or-less similar” possibilities). Even though this counter-position is difficult to falsify (as production possibility sets are unobservable), it does appear to be a rather precarious special case. As long as candidates have even slightly different production possibilities, our analysis applies, and the equilibrium does not converge to the standard median voter equilibrium. Moreover, Example 1 demonstrates that production functions whose values are relatively similar in the relevant range can lead to substantial differentiation. Thus, it is by no means necessary to assume unrealistically large differentiation between candidates to generate significant policy divergence.

4 Conclusion

In this paper, we have developed a simple model of political competition between two candidates who have access to different production technologies, so that the level of public goods provided by the government does not only depend on the tax rate, but also on the identity of the winning candidate. The model is very tractable and produces novel results that can shed light on some empirical regularities.

We identify conditions under which a unique pure strategy equilibrium exists where candidates adopt platforms that reinforce their respective strength: The candidate who has an advantage at running a small government proposes a smaller tax rate than his competitor who has an advantage in producing more units of the public good. This result connects to the literature on issue ownership that argues that candidates do not converge but rather choose to emphasize their strong points, because they cannot successfully replicate their opponent’s platform.

Citizens are characterized by two parameters, their income m and their preference for the public good, $h(\theta)$, but we show that their voting behavior is determined entirely by the ratio $m/h(\theta)$. In equilibrium, there exists a class of “cutoff” voters who are indifferent between the candidates, and whose utility both candidates maximize. All other voters have a strict preference for one of the candidates. Changes in

the distribution of voter preferences do not affect the positions candidates take. An improvement in the production technology of one candidate always induces his opponent to move to a more extreme position (away from his competitor). In contrast, the sign of the effect of a technological improvement on the candidate's own position often depends on parameters.

There are a number of interesting questions for further theoretical and empirical research. In terms of theory, a very interesting aspect would be to endogenize the production function of political candidates. Political parties, representing sets of voters could choose which one of a number of feasible candidates (differentiated according to their production functions) to nominate. Then, the respective nominees choose their platforms. There is good reason to presume that the two parties will not choose identical candidates: Party members are likely predominantly policy motivated. The members of the party composed of those voters who prefer low government spending are likely to choose someone whose expertise is in the provision of limited government, at least in part because such a candidate will not turn around in the general election campaign and propose large spending, as this would not be electorally successful for him. An analogous argument shows that the party of big spenders should choose a candidate whose advantage is in providing a large public good provision level.

Moreover, when individuals who want to become politicians choose which party to join, they will take into account in which party the career prospects are higher for someone with their own set of characteristics. Given the likely behavior of party members discussed in the last paragraph, individuals with high (low) a and b have better prospects in the Democratic (Republican) party. The resulting self-selection of potential candidates with different areas of expertise into parties is also likely to sustain an equilibrium nomination outcome in which the candidates of the two parties are differentiated.

Endogenizing the choice of candidates would also provide a venue through which long-run persistent changes in the distribution of voter preferences can eventually affect the policies chosen by the two parties' candidates. If a party notices that their current leadership has a low probability of winning the election given a persistent shift in voter preferences, then the party has an interest in exchanging the characteristics of their leaders (in the sense of choosing one with a more competitive production function), and the new leaders will then choose to run on a different platform than the old ones.

In terms of empirical research, several questions were already discussed in the previous section. There are also some additional questions that come to mind. As explained in the model section, learning by doing (i.e. increasing productivity during tenure in office) may be an important determinant of the incumbency advantage. If so, our model framework can be used to derive predictions regarding the relative positions of incumbents and challengers, as well as the intertemporal development of politicians' positions. A clear-cut prediction of our model is that, if one candidate's strength increases, his opponent chooses a more extreme platform. For example, if a competent Democrat runs against a Republican, then the Republican will shift to decrease proposed spending, relative to a race in which he faces a relatively incompetent or unpopular Democrat.

Finally, differences between party production functions may differ systematically across polities. For example, in most European countries, even very high-level positions in government bureaucracies are usually occupied by career civil servants who remain in charge of their tasks independent of which party is in government. Also, government workers are generally very well protected in Europe from reorganization or firing, so that one can argue that candidates' production functions are less differentiated in Europe than in the U.S. Consequently, the difference between policies enacted by left-wing and right-wing parties should be smaller in Europe than in the US.

5 Appendix

5.1 Proof of Theorems

Proof of Theorem 1. We first show that condition (4) is necessary. As argued in the text, an individual voter (θ, m) 's expected utility if candidate j is elected is given by (2); using this and (1), it follows that the citizen prefers candidate j to candidate k if

$$(t_k - t_j)m \geq h(\theta) [W(t_k) - W(t_j)] \quad (7)$$

If $t_0 = t_1 = t$ then candidate j wins if $f_j(t\bar{m} - b_j) > f_k(t\bar{m} - b_k)$ and there is a tie if $f_j(t\bar{m} - b_j) = f_k(t\bar{m} - b_k)$.

Now suppose that $t_0 \neq t_1$. Define

$$S(t_0, t_1) = \frac{W(t_1) - W(t_0)}{t_1 - t_0} \quad (8)$$

Then $S(t_0, t_1)$ is the cutoff type who is indifferent between the candidates. First, suppose that in an equilibrium $t_0 < t_1$. Then all types $\tau > S(t_0, t_1)$ vote for candidate 0. Since the distribution Φ of the median τ_ω is strictly monotone, it follows that candidate 0 must choose t_0 to maximize the set of all citizens (θ, m) that satisfy (7) for $j = 0, k = 1$. This is achieved by selecting t_0 to maximize $S(t_0, t_1)$ for given t_1 . Similarly, it follows that t_1 must minimize $S(t_0, t_1)$. In contrast, if $t_0 > t_1$ then t_0 must minimize $S(t_0, t_1)$ while t_1 maximizes $S(t_0, t_1)$.

Thus, a necessary condition for a Nash equilibrium with $0 < t_0 \neq t_1$ is that the first order condition, $\frac{\partial S(t_0, t_1)}{\partial t_j} = 0$ is satisfied for $j = 0, 1$. This, however, is exactly condition (4)

First, note that t_j with $W_j(t_j) < W_k(t_j)$ cannot be an equilibrium since candidate k could get 100 percent of the votes by deviating to $t_k = t_j$. Further, if $W_j(t_j) = W_k(t_k)$ then the fact that $W'_j(t) \neq W'_k(t)$ for any t implies that the first order condition is violated and hence at least one of the candidates does not maximize his vote share. Thus, it follows that $t_0 < t_1$.

We now take the second order condition:

$$\frac{\partial S(t_0, t_1)}{\partial t_0} = \frac{W''_0(t_0)}{t_1 - t_0} - \frac{2W'_0(t_0)}{(t_1 - t_0)^2} + \frac{2(W_1(t_1) - W_0(t_0))}{(t_1 - t_0)^3}; \quad (9)$$

$$\frac{\partial S(t_0, t_1)}{\partial t_1} = -\frac{W''_1(t_1)}{t_1 - t_0} - \frac{2W'_1(t_1)}{(t_1 - t_0)^2} + \frac{2(W_1(t_1) - W_0(t_0))}{(t_1 - t_0)^3}; \quad (10)$$

For any point choice of t_j that satisfies the first order condition, (4) must hold. Inserting (4) into (9) and (10) implies

$$\frac{\partial S(t_0, t_1)}{\partial t_0} = \frac{W'_0(t_0)}{t_1 - t_0}, \text{ and } \frac{\partial S(t_0, t_1)}{\partial t_1} = -\frac{W'_1(t_1)}{t_1 - t_0} \quad (11)$$

Since $t_0 < t_1$, (11) and $W'' < 0$ implies that $\frac{\partial S(t_0, t_1)}{\partial t_0} < 0$ for any t_0 that satisfies the first order condition. Thus, t_0 maximizes $S(t_0, t_1)$. Similarly, $W'' < 0$ implies $\frac{\partial S(t_0, t_1)}{\partial t_1} > 0$ and hence t_1 minimizes $S(t_0, t_1)$ as required.

It remains to exclude the case where (4) is satisfied but it is either optimal for candidate 0 to deviate to $t'_0 > t_1^*$ or for candidate 1 to deviate to $t'_1 < t_0^*$. However, for candidate 1 such a deviation is not possible since $t'_1 \bar{m} \geq b_1 \geq t_0^* \bar{m}$. Now suppose that candidate 0 deviates. Then $W_0(t'_0) \leq W_0(1) \leq W_1(t_1^*)$. This and $t'_0 > t_1^*$ implies that all citizens prefer candidate 1's t_1^* . Hence, candidate 0's deviation is not optimal.

Let $\tau^* = W'_0(t_0^*) = W'_1(t_1^*)$. Then it is immediate that all voters with $m/h(\theta) > \tau^*$ strictly prefer candidate 0, while all voters (m, t) with $m/h(\theta) < \tau^*$ strictly prefer candidate 1. Also, since $W'_0(\bar{t}) = W'_1(\bar{t})$ it follows that $t_0^* < \bar{t} < t_1^*$. Hence $W_0(t_0^*) < W_1(t_1^*)$ and strict monotonicity of w implies $f(\bar{m}t_0^* - b_0) < f(\bar{m}t_1^* - b_1)$, i.e., candidate 1's production exceeds that of candidate 0.

We now prove uniqueness. We have already shown that $t_0 = t_0^*$ and $t_1 = t_1^*$ is a necessary condition for pure strategy equilibria. Now suppose that (σ_0, σ_1) is a mixed strategy equilibrium where σ_i is candidate i 's probability distribution over tax rates. Suppose that one candidate, say candidate 1, strictly mixes in equilibrium. If candidate 0 plays t_0^* , then all types with $\tau \geq \tau^*$ vote for candidate 0 if candidate 1 chooses t_1^* , and the set of voters supporting candidate 0 is strictly larger when $t_1 \neq t_1^*$. Thus, by Assumption 2, candidate 0's winning probability is strictly larger than in equilibrium (t_0^*, t_1^*) . Since candidate 0's winning probability using σ_0 against σ_1 must be at least as large as that using t_0^* against σ_1 , it follows that candidate 0's winning probability in equilibrium (σ_0, σ_1) is strictly larger than in equilibrium (t_0^*, t_1^*) . By an analogous argument, candidate 1's could always deviate to t_1^* , so that his winning probability in the mixed strategy equilibrium must be at least as large as in the equilibrium (t_0^*, t_1^*) . However, this is a contradiction since the sum of the winning probabilities must be 1. ■

Proof of Theorem 2. In view of Theorem 1, it is sufficient to prove that it is not optimal for candidates to “ouflank” each other, i.e. for candidate 0 to choose a tax rate $t > t_0^*$ and for candidate 1 to choose a tax rate $t < t_1^*$.

Consider the first case, where $W_0(1) \leq W_1(t_1^*)$ and $t_0^* \bar{m} > b_1$. As indicated in Theorem 1, candidate 0 cannot outflank candidate 1 since any $t_0 > t_1^*$ results in a lower production compared to candidate 1 since $W_0(1) \leq W_1(t_1^*)$. Now suppose that candidate 1 chooses $t_1 < t_0^*$, while candidate 0 chooses t_0^* . Then there exists a cutoff type $\hat{\tau}_1$ such that all types $\tau > \hat{\tau}_1$ vote for candidate 1 and all types $\tau < \hat{\tau}_1$ vote for candidate 0. Clearly, it is optimal for candidate 0 to choose t_1 to minimize $\hat{\tau}_1$, i.e., $\hat{\tau}_1$ must solve (6). It remains to prove that (t_0^*, t_1^*) is an equilibrium if and only if $\tau_m \leq \frac{\hat{\tau}_1 + \tau^*}{2}$.

The symmetry of Φ implies that

$$1 - \Phi(\hat{\tau}_1) = 1 - \Phi(\tau_m + (\hat{\tau}_1 - \tau_m)) = \Phi(2\tau_m - \hat{\tau}_1). \quad (12)$$

It is immediate that

$$\Phi(\tau^*) = \Phi\left(2\frac{\hat{\tau}_1 + \tau^*}{2} - \hat{\tau}_1\right). \quad (13)$$

Candidate 1's winning probability given strategies (t_0^*, t_1^*) is $\Phi(\tau^*)$, since all citizens with $\tau < \tau^*$ vote for candidate 1. After the deviation, the winning probability is $1 - \Phi(\hat{\tau}_1)$. Thus, (12) and (13) imply that a

deviation by candidate 1 is not optimal if and only if

$$\Phi\left(2\frac{\hat{\tau}_1 + \tau^*}{2} - \hat{\tau}_1\right) \geq \Phi(2\tau_m - \hat{\tau}_1). \quad (14)$$

Since Φ is strictly monotone (14) is equivalent to $\frac{\hat{\tau}_1 + \tau^*}{2} \geq \tau_m$, which proves the first case.

Now consider the second case, where $W_0(1) > W_1(t_1^*)$ and $t_0^* \bar{m} \leq b_1$. Then (as indicated in Theorem 1) any deviation $t_1 \leq t_0^*$ would lead to zero output and higher taxes since b_1 exceeds the tax revenue $t_0^* \bar{m}$ if candidate 0 were elected. Thus, we can focus solely on deviations by candidate 0 to $t_0 \geq t_1^*$. The $\hat{\tau}_0$ be the cutoff voter after such a deviation, i.e., all citizens $\tau < \hat{\tau}_0$ vote for candidate 0 and all citizens with $\tau > \hat{\tau}_0$ vote for candidate 1. Thus, it is optimal for candidate 0 to choose $\hat{\tau}_0$ that solves (5). We now prove that (t_0^*, t_1^*) is an equilibrium if and only if $\tau_m \geq \frac{\hat{\tau}_0 + \tau^*}{2}$. The argument is similar to the one above. In particular, symmetry of Φ implies

$$\Phi(\hat{\tau}_0) = \Phi(\tau_m + (\hat{\tau}_0 - \tau_m)) = 1 - \Phi(2\tau_m - \hat{\tau}_0), \quad (15)$$

and it follows immediately that

$$1 - \Phi(\tau^*) = 1 - \Phi\left(2\frac{\hat{\tau}_0 + \tau^*}{2} - \hat{\tau}_0\right). \quad (16)$$

Candidate 1's winning probability given strategies (t_0^*, t_1^*) is $1 - \Phi(\tau^*)$. The winning probability after the deviation is $\Phi(\hat{\tau}_0)$. Thus, (15) and (16) imply that a deviation by candidate 0 is not optimal if and only if

$$1 - \Phi\left(2\frac{\hat{\tau}_0 + \tau^*}{2} - \hat{\tau}_0\right) \geq 1 - \Phi(2\tau_m - \hat{\tau}_0). \quad (17)$$

Since Φ is strictly monotone (14) is equivalent to $\frac{\hat{\tau}_0 + \tau^*}{2} \leq \tau_m$, which proves the second case.

The third case follows from combining the argument for the first two cases. ■

Proof of Theorem 3. Let $W_{j,b}(s) = w(f_j(s\bar{m} - b_j))$. In the following, we write $t_j^*(b_0, b_1)$ to denote the dependence of candidate j 's tax rate on the parameters that we change. Thus,

$$W'_{0,b_0}(t_0^*(b_0, b_1)) = W'_{1,b_1}(t_1^*(b_0, b_1)) = \frac{W_{1,b_1}(t_1^*(b_0, b_1)) - W_{0,b_0}(t_0^*(b_0, b_1))}{t_1^* - t_0^*}, \quad (18)$$

for all b_0, b_1 . Note that

$$W'_{j,b_j}(s) = w'(f_j(s\bar{m} - b_j))f'_j(s\bar{m} - b_j)\bar{m}. \quad (19)$$

Thus,

$$\begin{aligned} \frac{\partial W_{0,b_0}(t_0^*(b_0, b_1))}{\partial b_0} &= w'(f_0(t_0^*(b_0, b_1)\bar{m} - b_0))f'_0(t_0^*(b_0, b_1)\bar{m} - b_0) \left(\bar{m} \frac{\partial t_0^*(b_0, b_1)}{\partial b_0} - 1 \right) \\ &= W'_{0,b_0}(t_0^*(b_0, b_1)) \left(\frac{\partial t_0^*(b_0, b_1)}{\partial b_0} - \frac{1}{\bar{m}} \right), \end{aligned} \quad (20)$$

and

$$\begin{aligned}\frac{\partial W_{1,b_1}(t_1^*(b_0, b_1))}{\partial b_0} &= w'(f_1(t_1^*(b_0, b_1)\bar{m} - b_1))f_1'(t_1^*(b_0, b_1)\bar{m} - b_1)\bar{m}\frac{\partial t_1^*(b_0, b_1)}{\partial b_0} \\ &= W'_{1,b_1}(t_1^*(b_0, b_1))\frac{\partial t_1^*(b_0, b_1)}{\partial b_0}.\end{aligned}\quad (21)$$

Similarly, it follows that

$$\begin{aligned}\frac{\partial W'_{0,b_0}(t_0^*(b_0, b_1))}{\partial b_0} &= W''_{0,b_0}(t_0^*(b_0, b_1))\left(\frac{\partial t_0^*(b_0, b_1)}{\partial b_0} - \frac{1}{\bar{m}}\right), \\ \frac{\partial W'_{1,b_1}(t_1^*(b_0, b_1))}{\partial b_0} &= W''_{1,b_1}(t_1^*(b_0, b_1))\frac{\partial t_1^*(b_0, b_1)}{\partial b_0}.\end{aligned}\quad (22)$$

Using (18), (20), and (21) implies

$$\begin{aligned}\frac{\partial}{\partial b_0}\left(\frac{W_{1,b_1}(t_1^*(b_0, b_1)) - W_{0,b_0}(t_0^*(b_0, b_1))}{t_1^* - t_0^*}\right) &= \frac{W'_{1,b_1}(t_1^*(b_0, b_1))\left(\frac{\partial t_1^*(b_0, b_1)}{\partial b_0} - \frac{\partial t_0^*(b_0, b_1)}{\partial b_0} + \frac{1}{\bar{m}}\right)}{t_1^*(b_0, b_1) - t_0^*(b_0, b_1)} \\ &\quad - \frac{(W_{1,b_1}(t_1^*(b_0, b_1)) - W_{0,b_0}(t_0^*(b_0, b_1)))\left(\frac{\partial t_1^*(b_0, b_1)}{\partial b_0} - \frac{\partial t_0^*(b_0, b_1)}{\partial b_0}\right)}{(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))^2} \\ &= \frac{W'_{1,b_1}(t_1^*(b_0, b_1))\left(\frac{\partial t_1^*(b_0, b_1)}{\partial b_0} - \frac{\partial t_0^*(b_0, b_1)}{\partial b_0} + \frac{1}{\bar{m}}\right)}{t_1^*(b_0, b_1) - t_0^*(b_0, b_1)} - \frac{W'_{1,b_1}(t_1^*(b_0, b_1))\left(\frac{\partial t_1^*(b_0, b_1)}{\partial b_0} - \frac{\partial t_0^*(b_0, b_1)}{\partial b_0}\right)}{t_1^*(b_0, b_1) - t_0^*(b_0, b_1)} \\ &= \frac{W'_{1,b_1}(t_1^*(b_0, b_1))}{\bar{m}(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))} = \frac{W'_{0,b_0}(t_0^*(b_0, b_1))}{\bar{m}(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))}.\end{aligned}\quad (23)$$

Thus, taking the derivative of (18) with respect to b_0 implies

$$\frac{\partial t_0^*(b_0, b_1)}{\partial b_0} = \frac{1}{\bar{m}}\left(1 + \frac{W'_{0,b_0}(t_0^*(b_0, b_1))}{W''_{0,b_0}(t_0^*(b_0, b_1))(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))}\right)\quad (24)$$

$$\frac{\partial t_1^*(a_0, b_1)}{\partial b_0} = \frac{W'_{1,b_1}(t_1^*(b_0, b_1))}{W''_{1,b_1}(t_1^*(b_0, b_1))\bar{m}(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))}.\quad (25)$$

First, note that (25) implies that $\frac{\partial t_1^*(b_0, b_1)}{\partial b_0} < 0$. Note that $\frac{\partial t_0^*(b_0, b_1)}{\partial b_0} \geq 0$ if and only if $-W''_{0,b_0}(t_0^*)/W'_{0,b_0}(t_0^*) \geq 1/(t_1^* - t_0^*)$. The amount of public good produced by candidate 0 also decreases when $\frac{\partial t_0^*(b_0, b_1)}{\partial b_0} > 0$. In particular,

$$\frac{\partial(\bar{m}t_0^*(b_0, b_1) - b_0)}{\partial b_0} = \bar{m}\frac{\partial t_0^*(a_0, a_1)}{\partial b_0} - 1 = \frac{W'_{0,b_0}(t_0^*(b_0, b_1))}{W''_{0,b_0}(t_0^*(b_0, b_1))(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))},$$

which is strictly negative. Thus, input is decreased and therefore less public good is provided.

We next consider the derivatives with respect to b_1 . It follows immediately that

$$\begin{aligned}\frac{\partial W_{0,b_0}(t_0^*(b_0, b_1))}{\partial b_1} &= W'_{0,b_0}(t_0^*(b_0, b_1)) \frac{\partial t_0^*(b_0, b_1)}{\partial b_1} \\ \frac{\partial W_{1,b_1}(t_1^*(b_0, b_1))}{\partial b_1} &= W'_{1,b_1}(t_1^*(b_0, b_1)) \left(\frac{\partial t_1^*(b_0, b_1)}{\partial b_1} - \frac{1}{m} \right).\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{\partial W'_{0,b_0}(t_0^*(b_0, b_1))}{\partial b_1} &= W''_{0,b_0}(t_0^*(b_0, b_1)) \frac{\partial t_0^*(b_0, b_1)}{\partial b_1} \\ \frac{\partial W'_{1,b_1}(t_1^*(b_0, b_1))}{\partial b_1} &= W''_{1,b_1}(t_1^*(b_0, b_1)) \left(\frac{\partial t_1^*(b_0, b_1)}{\partial b_1} - \frac{1}{m} \right).\end{aligned}$$

Next, it follows that

$$\frac{\partial}{\partial b_1} \left(\frac{W_{1,b_1}(t_1^*(b_0, b_1)) - W_{0,b_0}(t_0^*(b_0, b_1))}{t_1^* - t_0^*} \right) = - \frac{W'_{0,b_0}(t_0^*(b_0, b_1))}{\bar{m}(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))} = - \frac{W'_{1,b_1}(t_1^*(b_0, b_1))}{\bar{m}(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))}$$

These equations imply

$$\frac{\partial t_0^*(b_0, b_1)}{\partial b_1} = - \frac{W'_{0,b_0}(t_0^*(b_0, b_1))}{W''_{0,b_0}(t_0^*(b_0, b_1)) \bar{m}(t_1^*(b_0, b_1) - t_0^*(b_0, b_1))} \quad (26)$$

$$\frac{\partial t_1^*(b_0, b_1)}{\partial b_1} = \frac{1}{\bar{m}} \left(1 - \frac{W'_{1,b_1}(t_1^*(b_0, b_1))}{W''_{1,b_1}(t_1^*(b_0, b_1)) (t_1^*(b_0, b_1) - t_0^*(b_0, b_1))} \right). \quad (27)$$

Note that (27), in combination with $W' > 0$, $W'' < 0$ and $t_1^* > t_0^*$, implies that $\frac{\partial t_1^*(b_0, b_1) \bar{m}}{\partial b_1} > 1$, which shows that candidate 1's public good provision increases as b_1 increases. ■

Proof of Theorem 4. The proof is similar to that of Theorem 3. Let $W_{j,a_j}(s) = w(a_j(s\bar{m} - b_j))$. Let t_j^* be the optimal tax rate chosen by candidate j . With a slight abuse of notation we now write $t_j^*(a_0, a_1)$ to denote the dependency of the solution on a_0 and a_1 . The equilibrium condition is again given by

$$W'_{0,a_0}(t_0^*(a_0, a_1)) = W'_{1,a_1}(t_1^*(a_0, a_1)) = \frac{W_{1,a_1}(t_1^*(a_0, a_1)) - W_{0,a_0}(t_0^*(a_0, a_1))}{t_1^* - t_0^*}, \quad (28)$$

for all b_0, b_1 . Note that

$$\frac{\partial W_{0,a_0}(t_0^*(a_0, a_1))}{\partial a_0} = W'_{0,a_0}(t_0^*(a_0, a_1)) \left(\frac{t_0^*(a_0, a_1) \bar{m} - b_0}{a_0 \bar{m}} + \frac{\partial t_0^*(a_0, a_1)}{\partial a_0} \right) \quad (29)$$

$$\frac{\partial W_{1,a_1}(t_1^*(a_0, a_1))}{\partial a_0} = W'_{1,a_1}(t_1^*(a_0, a_1)) \frac{\partial t_1^*(a_0, a_1)}{\partial a_0} \quad (30)$$

Similarly,

$$\frac{\partial W'_{0,a_0}(t_0^*(a_0, a_1))}{\partial a_0} = \frac{W'_{0,a_0}(t_0^*(a_0, a_1))}{a_0} + W''_{0,a_0}(t_0^*(a_0, a_1)) \left(\frac{t_0^*(a_0, a_1)\bar{m} - b_0}{a_0\bar{m}} + \frac{\partial t_0^*(a_0, a_1)}{\partial a_0} \right) \quad (31)$$

$$\frac{\partial W'_{1,a_1}(t_1^*(a_0, a_1))}{\partial a_0} = W''_{1,a_1}(t_1^*(a_0, a_1)) \frac{\partial t_1^*(a_0, a_1)}{\partial a_0} \quad (32)$$

Next, using (29), (30), (31), (32) and the equilibrium condition (28) imply

$$\frac{\partial}{\partial a_0} \left(\frac{W_{1,a_1}(t_1^*(a_0, a_1)) - W_{0,a_0}(t_0^*(a_0, a_1))}{t_1^* - t_0^*} \right) = - \frac{W'_{0,a_0}(t_0^*(a_0, a_1))(t_0^*(a_0, a_1)\bar{m} - b_0)}{a_0\bar{m}(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))}. \quad (33)$$

We now differentiate (33) and use (28) to get

$$\frac{\partial t_0^*(a_0, a_1)}{\partial a_0} = - \frac{t_0^*(a_0, a_1)\bar{m} - b_0}{a_0\bar{m}} \left(1 + \frac{t_1^*(a_0, a_1)\bar{m} - b_0}{t_0^*(a_0, a_1)\bar{m} - b_0} \frac{W'_{0,a_0}(t_0^*(a_0, a_1))}{W''_{0,a_0}(t_0^*(a_0, a_1))(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))} \right) \quad (34)$$

$$\frac{\partial t_1^*(a_0, a_1)}{\partial a_0} = - \frac{W'_{1,a_1}(t_1^*(a_0, a_1))(t_0^*(a_0, a_1)\bar{m} - b_0)}{W''_{1,a_1}(t_1^*(a_0, a_1))a_0\bar{m}(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))}. \quad (35)$$

Note that $\frac{\partial t_0^*(a_0, a_1)}{\partial a_0} < 0$ if and only if $-W''(t_0^*)/W'(t_0^*) > \frac{\bar{m}t_1^* - b_0}{\bar{m}t_0^* - b_0} \frac{1}{t_1^* - t_0^*}$. However, the amount produced always increases. In particular,

$$\frac{\partial (a_0\bar{m}t_0^*(a_0, a_1) - b_0)}{\partial a_0} = \bar{m}t_0^*(a_0, a_1) - b_0 + a_0\bar{m} \frac{\partial t_0^*(a_0, a_1)}{\partial a_0} \geq - \frac{W'_{0,a_0}(t_0^*(a_0, a_1))}{W''_{0,a_0}(t_0^*(a_0, a_1))(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))},$$

which is strictly positive. (35) immediately implies that t_1 increases. Thus, tax revenue and public good provision is increased.

Next, we consider changes in a_1 . We can again conclude that

$$\frac{\partial W_{0,a_0}(t_0^*(a_0, a_1))}{\partial a_1} = W'_{0,a_0}(t_0^*(a_0, a_1)) \frac{\partial t_0^*(a_0, a_1)}{\partial a_1}; \quad (36)$$

$$\frac{\partial W_{1,a_1}(t_1^*(a_0, a_1))}{\partial a_1} = W'_{1,a_1}(t_1^*(a_0, a_1)) \left(\frac{t_1^*(a_0, a_1)\bar{m} - b_1}{a_1\bar{m}} + \frac{\partial t_1^*(a_0, a_1)}{\partial a_1} \right); \quad (37)$$

and

$$\frac{\partial W'_{0,a_0}(t_0^*(a_0, a_1))}{\partial a_1} = W''_{0,a_0}(t_0^*(a_0, a_1)) \frac{\partial t_0^*(a_0, a_1)}{\partial a_1}; \quad (38)$$

$$\frac{\partial W'_{1,a_1}(t_1^*(a_0, a_1))}{\partial a_1} = \frac{W'_{1,a_1}(t_1^*(a_0, a_1))}{a_1} + W''_{1,a_1}(t_1^*(a_0, a_1)) \left(\frac{t_1^*(a_0, a_1)\bar{m} - b_1}{a_1\bar{m}} + \frac{\partial t_1^*(a_0, a_1)}{\partial a_1} \right) \quad (39)$$

Equations (36), (37), (38), (39) and the equilibrium condition (28) imply

$$\frac{\partial}{\partial a_1} \left(\frac{W_{1,a_1}(t_1^*(a_0, a_1)) - W_{0,a_0}(t_0^*(a_0, a_1))}{t_1^* - t_0^*} \right) = \frac{W'_{0,a_0}(t_0^*(a_0, a_1))(t_1^*(a_0, a_1)\bar{m} - b_1)}{a_1\bar{m}(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))}. \quad (40)$$

Differentiating (28) and using (40) implies

$$\frac{\partial t_0^*(a_0, a_1)}{\partial a_1} = \frac{W'_{0,a_0}(t_0^*(a_0, a_1))(t_1^*(a_0, a_1)\bar{m} - b_1)}{W''_{0,a_0}(t_0^*(a_0, a_1))a_1\bar{m}(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))} \quad (41)$$

$$\frac{\partial t_1^*(a_0, a_1)}{\partial a_1} = -\frac{t_1^*(b_0, b_1)\bar{m} - b_1}{a_1\bar{m}} \left(1 - \frac{t_0^*(a_0, a_1)\bar{m} - b_1}{t_1^*(a_0, a_1)\bar{m} - b_1} \frac{W'_{1,a_1}(t_1^*(a_0, a_1))}{W''_{1,a_1}(t_1^*(a_0, a_1))(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))} \right). \quad (42)$$

It follows immediately that both partial derivatives are negative, i.e., increasing a_1 decreases both t_0 and t_1 . This also implies that candidate 0's public good provision decreases. We now consider the change in input of candidate 1's public.

$$\begin{aligned} \frac{\partial (a_1\bar{m}t_1^*(a_0, a_1) - b_1)}{\partial a_1} &= \bar{m}t_0^*(a_0, a_1) - b_1 + a_1\bar{m} \frac{\partial t_1^*(a_0, a_1)}{\partial a_1} \\ &= (t_0^*(a_0, a_1)\bar{m} - b_1) \frac{W'_{1,a_1}(t_1^*(a_0, a_1))}{W''_{1,a_1}(t_1^*(a_0, a_1))(t_1^*(a_0, a_1) - t_0^*(a_0, a_1))}. \end{aligned}$$

Thus, inputs are decreased if and only if $\bar{m}t_0^* - b_1 > 0$. ■

5.2 Calculations for Example 1

We start by computing the equilibrium (t_0^*, t_1^*) . By Theorem 1, $W'_0(t_0^*) = W'_1(t_1^*) = [W_1(t_1^*) - W_0(t_0^*)]/(t_1^* - t_0^*)$, which implies

$$a_0c\bar{m}e^{-ca_0(\bar{m}t_0^* - b_0)} = a_1c\bar{m}e^{-ca_1(\bar{m}t_1^* - b_1)} \quad (43)$$

and

$$-a_1c\bar{m}e^{-ca_1(\bar{m}t_1^* - b_1)}(t_1^* - t_0^*) = e^{-ca_1(\bar{m}t_1^* - b_1)} \left(1 - \frac{a_1}{a_0} \right) \Rightarrow a_0a_1c\bar{m}(t_1^* - t_0^*) = a_1 - a_0. \quad (44)$$

Solving equations (43) and (44) for t_0^* and t_1^* yields

$$t_0^* = \frac{ca_0(a_1b_1 - a_0b_0) + a_0 \ln\left(\frac{a_1}{a_0}\right) - (a_1 - a_0)}{c\bar{m}a_0(a_1 - a_0)} \quad (45)$$

$$t_1^* = \frac{ca_1(a_1b_1 - a_0b_0) + a_1 \ln\left(\frac{a_1}{a_0}\right) - (a_1 - a_0)}{c\bar{m}a_1(a_1 - a_0)}, \quad (46)$$

and substitution of parameter values yields $t_0^* = 0.1$ and $t_1^* = 0.121$, as reported in the text.

Next, we compute \hat{t}_0 and \hat{t}_1 as defined in Theorem 2. The first-order conditions of (5) and (6) are

$$\hat{\tau}_0 = W'_0(\hat{t}_0) = \frac{W_0(\hat{t}_0) - W_1(t_1^*)}{\hat{t}_0 - t_1^*}, \quad (47)$$

$$\hat{\tau}_1 = W'_1(\hat{t}_1) = \frac{W_0(t_0^*) - W_1(\hat{t}_1)}{t_0^* - \hat{t}_1}, \quad (48)$$

It is easy to find numerically $\hat{t}_0 = 14.4\%$ and $\hat{t}_1 = 8.2\%$, so that $\hat{\tau}_0 = 3.618$ and $\hat{\tau}_1 = 7.513$.²¹

²¹The other solution of (47) and (48) is $\hat{t}_0 = t_0^*$ and $\hat{t}_1 = t_1^*$, i.e., the original profile in which $\hat{\tau}_0$ is minimized and $\hat{\tau}_1$ is maximized.

References

- Adams, J. (1999). Policy divergence in multicandidate probabilistic spatial voting. *Public Choice* 100(1), 103–122.
- Adams, J. and S. Merrill (2003). Voter turnout and candidate strategies in american elections. *Journal of Politics* 65, 161–189.
- Ansolabehere, S., J. Rodden, and J. Snyder (2006). Purple America. *Journal of Economic Perspectives* 20(2), 97–118.
- Ansolabehere, S. and J. M. Snyder (2000). Valence politics and equilibrium in spatial election models. *Public Choice* 103, 327–336.
- Banks, J. S. and J. Duggan (2005). Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. In *Social Choice and Strategic Decisions*, pp. 15–56. Springer: Berlin, Heidelberg.
- Bernhardt, D., J. Duggan, and F. Squintani (2006). Private polling in elections and voter welfare. mimeo, University of Illinois.
- Callander, S. (2005). Electoral competition in heterogeneous districts. *Journal of Political Economy* 113, 1116–1145.
- Callander, S. (2008). Political motivations. *Review of Economic Studies* 75, 671–697.
- Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science* 29, 69–95.
- Castanheira, M. (2003). Why vote for losers? *Journal of the European Economic Association* 1(5), 1207–1238.
- Coughlin, P. (1992). *Probabilistic Voting Theory*. Cambridge, Cambridge University Press.
- Dixit, A. and J. Londregan (1995). Redistributive politics and economic efficiency. *American Political Science Review* 89(4), 856–866.
- Egan, P. J. (2008). Issue ownership and representation: A theory of legislative responsiveness to constituency opinion. New York University.
- Erikson, R. and D. Romero (1990). Candidate equilibrium and the behavioral model of the vote. *American Political Science Review* 84, 1103–1126.
- Fiorina, M. (1973). Electoral margins, constituency influence, and policy moderation: A critical assessment. *American Politics Research* 1(4), 479–498.
- Gelman, A., B. Shor, J. Bafumi, and D. Park (2008). Rich state, poor state, red state, blue state: What? the matter with Connecticut? *Quarterly Journal of Political Science* 2, 345–367.
- Groseclose, T. (2001). A Model of Candidate Location when One Candidate Has a Valence Advantage. *American Journal of Political Science* 45(4), 862–886.
- Hinich, M. (1978). The mean versus the median in spatial voting games. In P. Ordeshook (Ed.), *Game Theory and Political Science New York*. NYU Press: New York.

- Huntington, S. (1950). A revised theory of American party politics. *The American Political Science Review* 44, 669–677.
- Jensen, T. (2009). Electoral Competition when Candidates are Better Informed than Voters. mimeo, University of Copenhagen.
- Kartik, N. and P. McAfee (2006). Signaling character in electoral competition. Forthcoming, American Economic Review.
- Krasa, S. and M. Polborn (2009). Political competition between differentiated candidates. University of Illinois.
- Krasa, S. and M. Polborn (2010). Competition between specialized candidates. *American Political Science Review* 104, 745–765.
- Krasa, S. and M. K. Polborn (2011). The relative importance of cultural and economic issues for the polarization of the u.s. electorate, 1972–2008. University of Illinois.
- Lindbeck, A. and J. Weibull (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice* 52, 273–297.
- Lindbeck, A. and J. Weibull (1993). A model of political equilibrium in a representative democracy. *Journal of Public Economics* 51, 195–209.
- Lowry, R., J. Alt, and K. Ferree (1998). Fiscal Policy Outcomes and Electoral Accountability in American States. *American Political Science Review* 92, 759–774.
- MacRae, D. (1952). The relation between roll call votes and constituencies in the massachusetts house of representatives. *American Political Science Review* 46, 1046–1055.
- Martinelli, C. (2001). Elections with privately informed parties and voters. *Public Choice* 108(1), 147–167.
- Osborne, M. and A. Slivinski (1996). A model of political competition with citizen candidates. *Quarterly Journal of Economics* 111, 65–96.
- Palfrey, T. R. (1984). Spatial equilibrium with entry. *Review of Economic Studies* 51, 139–156.
- Persson, T. and G. Tabellini (2000). *Political Economy: Explaining Economic Policy*. MIT Press.
- Petrocik, J. R. (1996). Issue ownership in presidential elections, with a 1980 case study. *American Journal of Political Science* 40(3), 825–850.
- Roemer, J. E. (1994). A theory of policy differentiation in single issue electoral politics. *Social Choice and Welfare* 11(4), 355–380.
- Soubeyran, R. (2009). Does a disadvantaged candidate choose an extremist position? *Annales d’Economie et de Statistique* 93-94, 328–348.
- Stokes, D. (1963). Spatial Models of Party Competition. *American Political Science Review* 57(2), 368–377.
- Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *American Political Science Review* 77, 142–157.