Persistent Ideology and the Determination of Public Policy Over Time*

Zheng Song
Fudan University
May 2009

Abstract

This paper investigates how public policy responds to persistent ideological shifts in dynamic politico-economic equilibria. To this end, we develop a tractable model to analyze the dynamic interactions among public policy, individuals’ intertemporal choice and the evolution of political constituency. Analytical solutions are obtained to characterize Markov perfect equilibria. Our main finding is that a right-wing ideology may increase the size of government. Data from a panel of 18 OECD countries confirm that after controlling for the partisan effect, there is a positive relationship between the right-wing political constituency and government size. This is consistent with our theoretical prediction, but hard to explain by existing theories.

**JEL Classification:** D72 E62

**Key Words:** Markov Equilibrium, Persistent Ideology, Political Economy, Public Policy, Repeated Voting

---

*This paper is based on Chapter 2 of my dissertation at IIES, Stockholm University. I am deeply indebted to my advisor, Fabrizio Zilibotti, for his guidance and encouragement. I thank anonymous referees, John Hassler, Giovanni Favara, Jose V. Rodriguez Mora, Kjetil Storesletten and seminar participants at IIES and Oslo for helpful comments. I also thank Christina Lonnblad for editorial assistance. Any remaining errors are mine.*
1 Introduction

Modern political economy is designed to reveal the underlying mechanism of policy decision-making. A salient feature in real-world democracies is that policy attitudes are often driven by motives that seem hard to reconcile with mere economic factors. The empirical literature has long documented that ideology plays a key role in shaping policy preferences.\(^1\) Many theoretical frameworks, such as the probabilistic voting model (e.g., Lindbeck and Weibull, 1987), also incorporate ideology as an important factor for political decisions. Existing theory, however, ignores the persistence of ideological shifts. For example, pro-redistribution “leftist” policies were highly popular in the 1950s and 1960s, while a “rightist” mood appeared to dominate in the late 1970s and 1980s.\(^2\) The impacts of such persistent ideological waves are far from trivial. In particular, they lead to prospective changes in the type of government and associated policy outcomes, which may influence private intertemporal choices and even the distribution of future voters.\(^3\) Such variations in response to ideological shifts naturally affect the incumbent government’s choices, indicating a distinct role of ideology in the policy decision process.

This paper, therefore, aims to show explicitly how persistent ideology influences the determination of public policies. To this end, we construct a politico-economic model that has the ability to capture rich dynamic interactions among policies, private decisions, and the evolution of the distribution of voters. Our main finding is that a right-wing ideology may increase the size of government. The underlying mechanism is two-fold. First, a persistent ideological shift towards the right implies a higher probability that a right-wing government will be elected in the future. Since a right-wing government features lower taxes, on average, the ideological shock encourages investment by reducing expected future tax rates. This makes the investment less elastic and, hence, provides the incentive for the incumbent to increase taxes. Moreover, the shock generates a self-reinforcing process on the distribution of future voters.

---

\(^1\)For instance, Sears, Lau, Tyler and Allen (1980) show that symbolic attitude (mainly liberal-conservative ideology and party identification) far outstripped all self-interest variables in terms of predicting support for policies in the United States: The contribution of symbolic attitudes to \(R^2\) ranged between 10 percent and 17 percent, while the contribution of self-interests never exceeded 4 percent. In addition, Levitin and Miller (1979), Knight (1985), and Alvarez and Nagler (1995, 1998), among many others, show that ideology turns out to be a significant predictor for individuals’ voting choice in U.S. presidential elections.


\(^3\)Recent theoretical work has shown that the evolution of political constituency can be endogenously driven by private intertemporal choices in a full-fledged dynamic environment (e.g., Hassler, Rodriguez Mora, Storesletten and Zilibotti, 2003).
More investment results in more individuals in favor of the right-wing, which further increases the right-wing’s future election probability. The impact of ideology can, thus, be amplified by this endogenous response of probabilities over future government types.

The model is based primarily on a tractable framework recently developed by Hassler, Storesletten and Zilibotti (2007). There are two types of individual economic status, the rich and the poor. Individuals make human-capital investments that increase their likelihood of being rich. Two political parties run electoral competition. The right-wing and left-wing parties, modeled as citizen-candidates (Osborne and Slivinski, 1996, Besley and Coate, 1997), represent the rich and poor, respectively. To incorporate ideology, we assume that a proportion of the poor (rich) vote for the right-wing (left-wing) party. The discrepancy between individuals’ economic interests and their political preferences captures the impact of ideology on voting behavior. The election is, thus, codetermined by two fundamentals in the economy: the size of the rich (or the poor) and the ideological state.

A distinctive feature of our model is that public policies, private investment and the distribution of future voters are mutually affected over time. To show how policies are determined in this environment, we focus on Markov perfect equilibria, where the dynamic interactions are characterized by two fixed-points: the ideology-contingent distribution of future voters and the ideology-contingent policy rules. Under quasi-linear preferences and uniformly distributed ideological shocks, the equilibrium can be solved analytically.

The standard partisan model suggests that ideological shifts play no role in the policy decision process, as long as the current type of government remains unchanged. By contrast, our model implies a positive relationship between government size and the right-wing ideology within each political regime. It is then left for empirical study whether the positive relationship holds in real-world democracies or is just a counterfactual result. We provide evidence from an OECD panel that a more right-wing political constituency indeed leads to a larger government, which is consistent with our prediction but hard to explain by existing theories. Specifically, we find that one percentage point increase in the vote share of right-wingers is associated with an increase in the central government revenue GDP ratio of 0.17 percentage points. This result is statistically significant and quite stable to a number of control variables and estimation specifications.

There is a growing literature on the dynamics of government without commitment techniques (e.g., Besley and Coate, 1998; Hassler et al., 2003; Hassler et al., 2007). This strand of research, including the present paper, emphasizes the fact that in representative democ-

\footnote{See, also, Krusell, Quadrini and Rios-Rull (1997) and Krusell and Rios-Rull (1999).}
racies, the incumbent government has limited abilities to commit to policies after the next election. The effect of a change in the future government type on equilibrium outcomes has been studied in some recent work, such as Amador (2003) and Song, Storesletten and Zilibotti (2007). However, much of the literature ignores a potentially important channel that runs from current policies back to future election probabilities. Azzimonti Renzo (2005) extends the analysis by endogenizing the distribution of voters in a dynamic setup. Like Azzimonti Renzo, we also allow current political and private decisions to affect the evolution of political constituency. The focus of our paper, however, is fundamentally different. We are interested in how persistent ideological shifts change policies, while in Azzimonti Renzo (2005) ideological shocks are purely i.i.d., acting to endogenize election outcomes and, therefore, playing a role similar to that played in the partisan model.

Although this paper aims to understand the influence of persistent ideology on the determination of public policies, it is also relevant for a long-standing issue in political science and sociology concerning the cause of changes in political constituency. A sizable empirical literature shows that political identifications are related to lagged economic conditions. However, few works have formalized the dynamic interaction between macroeconomy and political cycles. Our model, based on rational choices of parties and individuals, contributes to the literature by building a theoretical framework of analyzing changes in political constituency in response to both exogenous ideological shocks and endogenous public policies.

The rest of the paper is organized as follows. Section 2 describes the model and solves a static example. Section 3 gives the conditions for the existence and uniqueness of the Markov perfect equilibrium. In Section 4, we provide a closed-form solution when the ideological shock follows uniform distribution. Section 5 discusses the robustness of our results under alternative political objectives. Section 6 shows empirical evidence, and Section 7 concludes.

2 The Model

2.1 The Model Economy

The model economy is based primarily on a tractable framework recently developed by Hassler et al. (2003) and Hassler et al. (2007). The economy is inhabited by an infinite sequence of overlapping generations. Each generation has a unit mass and lives two periods. There are two types of old individuals endowed with different productivity, referred to as the old poor

5 Earlier research includes Persson and Svensson (1989) and Alesina and Tabellini (1990), among others, providing examples of strategic policy decision-making under future electoral uncertainty.

and the old rich, respectively. The wage of the old rich is normalized to unity, and the poor earn zero. The benefits from public good consumption $g$ are identical across old individuals. The government imposes a proportional income tax rate $\tau^o$ on the old. Let $u^{ou}_t$ and $u^{os}_t$ be the utilities of the old poor and old rich, respectively. These are equal to

\begin{align}
    u^{ou}_t &= a^o g_t, \quad (1) \\
    u^{os}_t &= 1 - \tau^o_t + a^o g_t, \quad (2)
\end{align}

where $a^o \in (0, 2]$ is the constant marginal utility of a public good for the old.$^7$

Young individuals are ex-ante homogenous. They make a human-capital investment $h$ at birth, which will increase the probability $p$ of being rich over their lifetime.$^8$ Without loss of generality, let $p = h \in [0, 1]$. As with old individuals, the wage of the young rich equals unity, and the poor earn zero. Therefore, once being rich, an individual earns a high wage, normalized to unity, in both periods of her life. On the other hand, a poor individual receives zero earnings.$^9$ $\tau^y_t$ is the proportional income tax rate for young individuals. For analytical convenience, we assume a linear-quadratic preference over consumption and costs of human-capital investment. The expected utility of a young household is

\begin{equation}
    u^y_t = h_t (1 - \tau^y_t) + a^y g_t - h_t^2 + \beta E[u^o_{t+1}], \quad (3)
\end{equation}

where $E$ is the expectation operator and $\beta \in [0, 1]$ denotes the discount factor. $a^y$ is the marginal utility of a public good for the young. Since the probability of being rich when old is equal to $h$, we have

\begin{equation}
    E \left[ u^o_{t+1} \right] = h_t E \left[ u^{os}_{t+1} \right] + (1 - h_t) E \left[ u^{ou}_{t+1} \right]. \quad (4)
\end{equation}

Age-dependent taxation has its counterparts in the real world. Many public programs and tax policies have important age-dependent elements. In addition, the young and old may evaluate public goods, such as public health care, in quite different ways. Allowing for age-dependent taxation also simplifies the analytical characterization, without fundamentally changing the results.$^{10}$

$^7$Assuming equal marginal utility of public spending across households is for notational convenience. It can be argued that the poor care about public spending more than the rich. The following results carry over to the case in which $a^o$ is different between the poor and the rich.$^8$This implies that human capital investment increases productivity contemporaneously. The assumption simplifies the analysis substantially. Otherwise we would have to work on a multi-period model that captures conflicts of interests across generations.$^9$The assumption on the perfectly correlated earnings in both periods is not essential. Our results will be qualitatively unchanged as long as earnings in the two periods are positively correlated.$^{10}$See Section 4.5 for more details.
Through the wage structure, the old and young produce $h_{t-1}$ and $h_t$, respectively. Thus, the aggregate output $y_t$ equals

$$y_t = h_{t-1} + h_t.$$  \hspace{1cm}(5)$$

Total tax revenue and public spending amount to $\tau_0 h_{t-1} + \tau_y h_t$ and $2g_t$, respectively. We assume that the government budget must be balanced in each period, which implies

$$g_t = \frac{\tau_0 h_{t-1} + \tau_y h_t}{2}. \hspace{1cm}(6)$$

2.2 The Political Decision Process

The sequence of tax rates is set through a repeated political decision process. We assume that only old individuals vote. This captures, in an extreme fashion, the phenomenon that the old are more influential in determining public policies.\(^{11}\) It would be observationally equivalent to assume that voting occurs at the end of each period. Old individuals have no interests at stake and, thus, abstain from voting. For expositional ease, we keep the former interpretation throughout the paper. We will show in Section 5 that a relaxation of this assumption leads to no major changes in our main findings.

The left-wing and right-wing parties, modeled as citizen-candidates, represent the old poor and old rich, respectively. The party candidates cannot credibly commit to any policy other than that preferred by the group they represent. For simplicity, we assume zero entry cost.\(^{12}\) In the absence of ideology, the majority rule implies that election outcomes are deterministic and depend solely on the distribution of old individuals’ economic situation. However, the literature has provided convincing evidence that, besides economic reasons, the electorate’s ideological label also plays a significant role in policy preference and voting choice. To capture this phenomenon, we introduce another variable - namely, the ideological state - to reflect the discrepancy between the electorate’s economic interests and political preference. As a consequence, election outcomes become codetermined by the distribution of old individuals’ economic situation and the ideological state.\(^{13}\)

\(^{11}\)For instance, Mulligan and Sala-i-Martin (1999) argue that the old have more influence in the political decision process because they have a lower cost of time. Empirically, the voting turnout is, indeed, lower for younger households (e.g., Wollinger and Rosenstone, 1980). See Hassler et al. (2003) and Hassler et al. (2007) for more-detailed discussions.

\(^{12}\)For simplicity, we assume zero entry cost, which shuts down the entry game in the standard citizen candidate model. Consequently, both the candidate representing the rich and the one representing the poor will participate in the electoral competition. However, we still regard the two-party system as a simplified citizen-candidate model, since the party candidates cannot credibly commit to any policy platform other than their preferred policies, as in Osborne and Slivinski (1996) and Besley and Coate (1997).

\(^{13}\)s can also be considered the quality of the party leadership, or even the popularity of the leadership. More generally, “ideology” may capture any factor affecting vote shares unrelated to economic concerns. We thank a referee for this alternative interpretation.
Specifically, we assume that an ideological shock can switch a proportion of the poor (rich) to the right-wing (left-wing) side in terms of voting choice. Define the left-wingers (right-wingers) as old households voting for the left-wing (right-wing) party. The election outcome is determined by the proportion of right-wingers $e_t$:

$$
e_t = \begin{cases} 
1 & s_t \geq 1 - h_{t-1} \\
h_{t-1} + s_t & s_t \in (-h_{t-1}, 1 - h_{t-1}) \\
0 & s_t \leq -h_{t-1}
\end{cases}, \tag{7}
$$

where $s_t$ is the ideological state at time $t$ and $h_{t-1}$ is the population of the old rich or, equivalently, the human-capital investment at time $t-1$. A positive (negative) $s_t$ switches some of the poor (rich) to vote for the right-wing (left-wing) party. Thus, a high (low) $s_t$ refers to a more right-leaning (left-leaning) ideology. The right-wing party wins the election if $e_t > \frac{1}{2}$. Otherwise, the left-wing party is elected.\textsuperscript{14} Note that (7) ensures that $e_t \in [0, 1]$ always holds. When $s_t$ takes an extreme value (either very high or very low), the economic determinant $h_{t-1}$ is wiped out. Outside these "ages of extremes" (Hobsbawm, 1996), economic motives may sway voters. Since ideological movements tend to be persistent (e.g., Robinson and Fleishman, 1984), $s_t$ is assumed to follow a stationary AR(1) process whose properties will be defined and discussed below.\textsuperscript{15}

It has been a long tradition in the literature of political economy that poor (rich) is synonymous with the left (right). This receives some empirical support from the finding that increased employment raises the popularity of the left government, while inflation reduces the popularity of the right via the wealth effect (e.g., Haynes and Jacob, 1994). However, it is a far-fetched idea that political constituency is purely determined by economic factors.\textsuperscript{16} In fact, (7) can be thought of as a parsimonious way of capturing the influence of both economic and ideological factors on the formation of political constituency.

The timing of events in each period is described as follows. Citizen candidates announce their policy platforms at the beginning of each period. An ideological shock is realized afterwards. The elected party then implements its preferred tax rates and public spending. Given public policies, young individuals invest in human capital. Their being rich or poor is revealed after they invest.

\textsuperscript{14} We assume that the left-wing comes into power if the proportion of left-wingers and right-wingers is equal.

\textsuperscript{15} The existence and uniqueness of the dynamic politico-economic equilibrium can easily be extended to an AR($n$) process with $n > 1$.

\textsuperscript{16} Besides the extensive evidence provided by political scientists, it is worth mentioning a recent empirical study from Di Tella and MacCulloch (2005) suggesting the importance of ideology. Based on survey data from ten OECD countries for 1975-1992, they find that "... respondents declare themselves to be happier when the party in power has a similar ideological position to themselves, even after we control for key performance indicators such as unemployment, inflation and income." (Di Tella and MacCulloch, 2005, pp.378)
2.3 Policy Choices under Different Political Regimes

The right-wing party sets $\tau^o_t$ so as to maximize the utility of the rich $u^o_s$ in (2), subject to the balanced-budget constraint (6). The assumption that $a^o \leq 2$ is sufficient for the right-wing to set $\tau^o_t = 0$. The left-wing sets $\tau^o_t$ by maximizing $u^{ou}_t$ in (1), which is equivalent to maximizing fiscal revenues $\tau^o_th_{t-1} + \tau^o_yh_t$. Since $h_{t-1}$ is predetermined and $\tau^o$ does not distort young individuals’ human-capital investment, the left-wing will set $\tau^o_t = 1$. In other words, the left-wing government entirely eliminates the income inequality of old individuals by imposing a 100-percent tax rate. To conclude, $\tau^o_t$ follows a binary rule

$$\tau^o_t = \begin{cases} 1 & \text{if } e_t \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}.$$ 

(8)

The disagreement on tax $\tau^o_t$ exhibits the feature of a two-party system. Despite the conflict of interest between left-wingers and right-wingers in terms of $\tau^o_t$, their preferences on $\tau^o_t$ are perfectly aligned: attaining the top of the Laffer curve to maximize taxes from young individuals. This is because citizen candidates represent the interests of only the old rich and old poor. None of them care about the welfare of the young. So, $\tau^o_t$ solves

$$\tau^o_t = \arg\max T_t,$$ 

(9)

where $T_t \equiv \tau^y_th_t$. Four remarks are in order. First, the incumbent at time $t$ would be better off if she could promise $\tau^o_{t+1} = 0$ to encourage human-capital investment $h_t$. Without commitment techniques, however, the promise is not credible since future policies are repeatedly decided by the winners of future elections. So, $\tau^o_{t+1}$ must follow the binary rule (8). Second, the political parties would disagree on the tax rate imposed on the young if the government were not allowed to adopt age-dependent taxation. As will be discussed in Section 4.5, the disagreement will not qualitatively change the results. Third, we adopt majority voting as our benchmark political decision process. However, Section 5 shows that majority voting is essentially identical to probabilistic voting à la Lindbeck and Weibull (1987), once we assume politicians to be altruistic towards the young. Finally, in the present setup, both parties would like to see the right-wing elected in the next period since $\tau^o_{t+1} = 0$ in the right-wing regime encourages investment and, thus, enlarges the tax base. We shall see in Section 5.1 that including re-election concerns does not change our main findings.

2.4 Two Effects of Ideology with Exogenous Political Constituency

We now distinguish two channels for ideology to affect policies. For expositional ease, we assume (temporarily) political constituency to be exogenous; i.e., the proportion of right-wingers $e_t$ is determined purely by the ideological state $s_t$. Specifically, the right-wing party
wins the election if $s > 0$. The next section will analyze our benchmark setup, in which $e_t$ obeys (7). The exogenous political constituency shuts down the link between the distribution of future voters and the current private investment in human capital. This makes the analysis straightforward.

The effect of an ideological shock on policies via election (the first channel) is analogous to that in the standard partisan models, which will be referred to as the partisan effect. The second channel, a novel feature of our model, governs how policies respond to a change in the distribution of future voters driven by a persistent ideological shock. As a warm-up exercise to facilitate the intuition, let us first study a static example with no private intertemporal trade-off. This helps identify the partisan effect of ideology in the first channel.

2.4.1 The Partisan Effect of Ideology

In this static example, we assume the probability of being rich in old age, $p$, to be exogenous. We will continue to drop the time subscript when it does not create any confusion. The corresponding politico-economic equilibrium is straightforward. The policy rule of $\tau^o_t$ follows (8). According to (3), young individuals’ human-capital investment solves

$$h = \arg \max_{h \in [0, 1]} (1 - \tau^y) h - \tilde{h}^2,$$

which yields

$$h = \frac{1 - \tau^y}{2}.$$  \hspace{1cm} (11)

(11) shows that private choice is independent of ideology. Substituting (11) into (9), we obtain an equalized distorting tax rate across ideological states:

$$\tau^y = \frac{1}{2}.$$  \hspace{1cm} (12)

Now consider the policy rule (8) and (12). Assuming away intertemporal trade-offs shuts down the link between ideology and the distortionary tax rate $\tau^y$. Nevertheless, if the ideology shock $s \leq 0$, the left-wing party, representing the interests of the poor, will win the election and spend more for redistribution by letting $\tau^o = 1$. Therefore, an ideological shift may affect government size by changing the identity of the incumbent party. The implication from this partisan effect is, thus, in accordance with the standard prediction of partisan theory.

2.4.2 The Intertemporal Effect of Ideology

We now proceed by incorporating private intertemporal choices into the above static model. We maintain the exogeneity of the distribution of voters; i.e., the right-wing party wins the election if $s > 0$. 
Denote $x'$ as the variable $x$ in the next period. The expected utility $u^y$ in (3) implies that $h$ depends on $E[\tau^o]$. According to the binary tax rule (8), $E[\tau^o]$ is equal to $1 - \pi$, where

$$\pi \equiv \Pr(e > 1/2)$$

denotes the right-wing’s probability of being elected in the next period. Alternatively, $\pi$ can be considered a variable characterizing the distribution of future voters. $\pi = \Pr(s' > 0)$ under exogenous political constituency.

Plugging (8) into (3), young individuals solve

$$h = \arg\max_{\hat{h} \in [0,1]} (1 - \tau^y + \beta \pi) \hat{h} - \hat{h}^2.$$ \tag{13}

The utility from a public good is irrelevant for the decision $h$, due to the atomistic unit assumption on individuals. (13) yields

$$h = \frac{1 - \tau^y + \beta \pi}{2}.$$ \tag{14}

Compared with (11), a new feature of (14) is that $h$ increases in $\pi$, the probability that a right-wing government will be elected. The reason is simple: The right-wing, if elected, would adopt a tax-free policy for the current young when they become old. Moreover, the elasticity of tax base $h$ depends on the election probability:

$$\epsilon = \frac{\tau^y}{1 - \tau^y + \beta \pi},$$ \tag{15}

where $\epsilon$ stands for the absolute value of the elasticity of $h$ with respect to $\tau^y$. Clearly, a higher $\pi$ leads to a lower $\epsilon$. That is to say, the current tax base tends to be less elastic when the future election is more favorable to the right-wing.

Substituting (14) back into (9) solves the following first-order condition for $\tau^y$:

$$\tau^y = \frac{1 + \beta \pi}{2}.$$ \tag{16}

(16) shows that $\pi$ can affect $\tau^y$. Compared with the policy rule (12) in the static example, we find that the effect appears whenever the tax base $h$ is subject to private intertemporal choices. The intuition is straightforward. Since the right-wing regime features a lower $\tau^o$, a more rightist political constituency in the future will encourage private investment and, therefore, reduce the elasticity of the tax base. This provides the incentive for the incumbent to increase $\tau^y$.

If ideological movements are persistent, i.e., $d\pi/ds > 0$, (16) implies that $d\tau^y/ds > 0$. This illustrates the second channel for ideology to affect political decisions via private intertemporal choices. By contrast, ideological shifts play no role in the standard partisan model, as long as the current type of government remains unchanged. Moreover, through this second channel, a right-leaning ideology may actually encourage the government to impose a higher tax rate,
which is opposite to the partisan effect. This effect will be referred to as the intertemporal effect of ideology.

The above models are built upon an ad hoc assumption that the distribution of voters is driven entirely by ideology. However, being rich or poor indeed shapes individual policy references, as can be seen from (1) and (2). Moreover, the exogenous political constituency has no interactions with policy decision and private choices. These interactions are not only theoretically appealing, but reflect the essence of democracy. In the rest of the paper, we will remove this assumption and go back to our benchmark model, in which political constituency is codetermined by ideology and human-capital investment. Endogenizing political constituency results in a self-reinforcing process running from an ideological shock to the distribution of future voters. Nevertheless, our main finding on the intertemporal effect of ideology still holds true.

3 The Politico-Economic Equilibrium with Endogenous Political Constituency

The next-period election probability, \( \pi \), is governed by the stochastic process of \( s \) under exogenous political constituency. In our benchmark setup (7), \( \pi \) becomes an equilibrium outcome involving private investment. To see this, we substitute (14) into (7). The definition of \( \pi \) establishes

\[
\pi = \Pr \left( e' > \frac{1}{2} \right) = \Pr \left( \frac{1 - \tau^y + \beta \pi}{2} + s' > \frac{1}{2} \right). \tag{17}
\]

The fixed point of equation (17) solves the equilibrium probability \( \pi \). In particular, the link between \( h \) and \( \pi \) establishes a channel for an ideological shock to affect the distribution of future voters. An increase of \( s \) (a right-wing ideology) leads to a high \( \pi \) and, therefore, a high \( h \). More human-capital investment, in turn, increases \( \pi \) as more individuals will be rich and in favor of the right-wing in the next period. In other words, a right-wing ideological wave may move future political constituency further towards to the right through this self-reinforcing process. (17) also shows that \( \pi \) depends on \( \tau^y \) since \( \tau^y \) can affect \( h \) and, thus, \( e' \). This establishes a channel for \( \tau^y \) to influence \( h \) via \( \pi \), resulting in a more distortive \( \tau^y \) under the endogenous political constituency. We shall see explicitly in Section 4 how the endogeneity of political constituency affects political choices.

3.1 The Endogenous Distribution of Future Voters

Before characterizing the fixed-point problem, we first specify the properties of the stochastic process of \( s \) as follows. The density function is defined by \( f : R^2 \rightarrow [0, \infty) \) with \( \int f (s'|s) \; ds' = 1 \)
for any given $s$. By (17), we know that $\pi$ depends on $\tau^y$ and the probability of the future ideological state $s'$, which is, in turn, contingent on the current ideological state $s$. Hence, $\pi$ can be written as a function of $\tau^y$ and $s$, $\pi : [0,1] \times R \rightarrow [0,1]$, which solves the following functional equation implied by (17):

$$\pi (\tau^y, s) = \int_{s' \geq \tau^y - \beta \pi (\tau^y, s)} f (s'|s) ds'.$$

(18)

The existence of the ideology-contingent probability $\pi (\tau^y, s)$ can easily be obtained by assuming the following properties on $f (s'|s)$. Define $X \equiv [\underline{s}, \bar{s}]$, where $-\infty < \underline{s} < \bar{s} < \infty$.

Assume

\textbf{A1}: $s'$ and $s \in X$.

\textbf{A2}: $f (s'|s)$ is bounded and uniformly continuous.

\textbf{Lemma 1} Assume A1 and A2. Then, there exists a uniformly continuous function $\pi (\tau^y, s)$ that solves (18).

\textbf{Proof}: See the appendix.

A2 is a sufficient condition for the existence. $\pi (\tau^y, s)$ can exist under discontinuous distributions, as shall be seen in Section 4.

We can further establish the uniqueness of $\pi (\tau^y, s)$ by assuming

\textbf{A3}: $f (s'|s) < 2/\beta$ for all $s'$ and $s \in X$.

\textbf{Lemma 2} Assume A1 and A3. Then, there exists a unique $\pi (\tau^y, s)$ that solves (18).

\textbf{Proof}: See the appendix.

Again, A3 is a sufficient condition for the uniqueness. Lemma 2 implies that sufficient ideological uncertainty can rule out the indeterminacy of beliefs, which features a number of recent studies on dynamic politico-economic equilibrium with endogenous identity of the policymaker (e.g., Hassler et al., 2003).\footnote{In an earlier version of this paper (Song, 2005, Chapter 2), we relaxed assumption A3 and investigated the multiplicity of equilibria.}

Plugging the probability $\pi (\tau^y, s)$ into (14) gives

$$h (\tau^y, s) = \frac{1 - \tau^y + \beta \pi (\tau^y, s)}{2}.
\quad (19)$$

Then, (7) shows that the future political constituency $e'$ evolves according to

$$e' (s', \tau^y, s) = \begin{cases} 1 & \text{if } s' \geq 1 - h (\tau^y, s) \\ h (\tau^y, s) + s' & \text{if } s' \in (-h (\tau^y, s), 1 - h (\tau^y, s)) \\ 0 & \text{if } s' \leq -h (\tau^y, s) \end{cases}.\quad (20)$$
3.2 Markov Perfect Equilibrium

Given the ideology-contingent probability $\pi(\tau^y, s)$ solved from (18) and the individual investment decision (19), an incumbent government will choose $\tau^y$ according to (9). Let $T(\tau^y, s) \equiv h(\tau^y, s) \tau^y$; i.e., tax revenues from the young. The problem can be written as

$$\tau^y(s) = \arg \max_{\tau^y \in [0,1]} T(\tau^y, s). \tag{21}$$

Two remarks are in order. First, the theorem of maximum implies that $\tau^y: R \to [0, 1]$ be an upper hemi-continuous mapping. Second, $\tau^y$ depends only on the current ideological state $s$.

One may guess that $\tau^y$ should also depend on the state of political constituency $e = s + h - 1$, as the policy rule of $\tau^o$ in (8). In fact, $e$ or the identity of an incumbent has no influence on $\tau^y$ since the objectives of the two parties over $\tau^y$ are perfectly aligned: maximizing tax revenue from the young. We will study party-specific $\tau^y$ in Section 5.1.

This paper focuses on Markov perfect equilibria, in which private and public choices are conditioned to payoff-relevant state variables. There are two state variables in our model: the ideological state $s$ and the proportion of right-wingers $e = s + h - 1$. These two state variables are payoff-relevant since they determine the current election and, thus, policy outcomes. So, the Markovian political equilibrium can be defined as follows.

**Definition 1** A (Markov perfect) political equilibrium is a set of mappings $\tau^o(e)$, $\tau^y(s)$, $\pi(\tau^y(s), s)$, and $h(\tau^y(s), s)$ such that:

1. $\tau^o(e)$ follows (8);
2. given $\tau^y(s)$, the next-period election probability $\pi(\tau^y(s), s)$ solves (18);
3. given $\pi(\tau^y(s), s)$, the human-capital investment $h(\tau^y(s), s)$ follows (19);
4. given $h(\tau^y(s), s)$, the incumbent solves $\tau^y(s)$ by (21).

4 An Analytical Solution

In this section, we provide a closed-form solution of the Markov perfect equilibrium. The complete characterization of the equilibrium reveals dynamic interactions among political constituency, policy decision-making and private intertemporal choice. We assume that $s'$ follows an AR(1) process with a symmetric uniformly-distributed innovation:

$$s' = \rho s + \varepsilon'. \tag{22}$$

---

18 The dynamic game in our model also allows for equilibria with trigger strategies.

19 An analytical solution is also available, though much more tedious, under more general setups. For example, $s'$ follows an AR($n$) process with the innovation that has a piecewise linear cumulative distribution function.
The ideological shock is stationary and persistent, i.e., \( \rho \in (0,1) \). The density of \( \varepsilon \) equals \( 1/(2z) \) if \( \varepsilon \in (-z,z) \) and 0 otherwise. So, the conditional density function of \( s' \) is

\[
f(s'|s) = \begin{cases} \frac{1}{2z} & \text{if } s' \in (\rho s - z, \rho s + z) \\ 0 & \text{otherwise} \end{cases}
\]

(23)

4.1 Exogenous Political Constituency

Before proceeding, it is instructive to solve \( \pi \) and \( \tau^y \) when the distribution of voters is determined entirely by ideology, as in Section 2.4. There, \( \pi = \Pr(s' > 0) \) is exogenous and only depends on \( s \):

\[
\pi(s) = \begin{cases} \frac{1}{2z} & \text{if } s \geq \frac{z}{\rho} \\ \frac{\rho s + z}{2z} & \text{if } s \in \left(-\frac{z}{\rho}, \frac{z}{\rho}\right) \\ 0 & \text{if } s \leq -\frac{z}{\rho} \end{cases}
\]

(24)

Clearly, \( \pi \) increases in \( s \) as long as \( \rho > 0 \). The marginal effect of ideology on \( \pi \) is equal to \( \rho/4z \) for \( s \in (-z/\rho, z/\rho) \). Substituting (24) back into (16), we obtain the distortionary tax rule associated with exogenous political constituency:

\[
\tau^y(s) = \begin{cases} \frac{1}{1+\beta(\rho s + z)/(2z)} & \text{if } s \leq -\frac{z}{\rho} \\ \frac{1+\beta}{2} & \text{if } s \in \left(-\frac{z}{\rho}, \frac{z}{\rho}\right) \\ \frac{1}{1+\beta} & \text{if } s \geq \frac{z}{\rho} \end{cases}
\]

(25)

We see explicitly from this example that a persistent right-wing ideological wave may increase the distortionary tax rate. The intertemporal effect of ideology on \( \tau^y \) is equal to \( \beta \rho/4z \) for \( s \in (-z/\rho, z/\rho) \).

4.2 The Endogenous Distribution of Future Voters

Now we solve \( \pi \) when the distribution of future voters is endogenous and affected by private investment. Given (23), the functional equation (18) becomes

\[
\pi(\tau^y, s) = \begin{cases} \frac{1}{2z} \left(\rho s + z - \frac{\tau^y - \beta \pi(\tau^y, s)}{2}\right) & \text{if } \frac{\tau^y - \beta \pi(\tau^y, s)}{2} \leq \rho s - z \\ 0 & \text{if } \frac{\tau^y - \beta \pi(\tau^y, s)}{2} \in (\rho s - z, \rho s + z) \\ \frac{1}{2z} \left(\rho s + z - \frac{\tau^y - \beta \pi(\tau^y, s)}{2}\right) & \text{if } \frac{\tau^y - \beta \pi(\tau^y, s)}{2} \geq \rho s + z \end{cases}
\]

(26)

The linearity makes the analytical solution straightforward. Assumption A3 implies that \( z > \beta/4 \), which is sufficient for the uniqueness of \( \pi(\tau^y, s) \) under the uniform distribution (23). In this section, we assume that \( z > \beta/4 \). It can be shown that \( z > \beta/4 \) is also necessary.\(^{20}\)

\(^{20}\)The opposite case, \( z < \beta/4 \), which produces multiple equilibria, was studied in Song (2005, Chapter 2). \( \pi(\tau^y, s) \) does not exist if \( z = \beta/4 \). The non-existence of \( \pi(\tau^y, s) \) is due to the fact that the uniform distribution (23) is not continuous and, thus, does not satisfy assumption A2.
where $\lambda^- (s) \equiv 2 (ps - z) + \beta$ and $\lambda^+ (s) \equiv 2 (ps + z)$. Note that $\lambda^+ (s) > \lambda^- (s)$ as long as $z > \beta / 4$. For notational convenience, we refer to $\lambda^+ (s) \leq 0$ or, equivalently, $s \leq -z / \rho$ as the left-dominating region, where the left-wing will be elected with probability one in the next period, irrespective of $\tau^y$. Symmetrically, $\lambda^- (s) \geq 1$ or, equivalently, $s \geq ((1 - \beta) / 2 + z) / \rho$ is referred to as the right-dominating region, where the right-wing will be elected with probability one under any $\tau^y$.

It immediately follows that $\partial \pi (\tau^y, s) / \partial s \geq 0$. Such an effect was illustrated by the model with exogenous political constituency in Section 2.4 (e.g., $d \pi / ds$ in (24)). A novel feature of the endogenous political constituency is that there is a self-reinforcing process running from $s$ to $\pi$. A comparison between (27) and (24) shows that the marginal effect of ideology on $\pi$ increases from $\rho / (2z)$ to $2 \rho / (4z - \beta)$. The intuition is straightforward. A right-wing ideology increases $\pi$ and $h$. The higher $h$, in turn, leads to more rich voters. The future political constituency moves further towards the right. Since the right-wing features zero $\tau^o$, such a self-reinforcing process encourages more private investment, giving an extra incentive for the incumbent to increase $\tau^y$. This tends to amplify the intertemporal effect of ideology.

(27) shows that the endogenous political constituency also allows $\tau^y$ to affect $\pi$. Differentiating (27) w.r.t. $\tau^y$ gives $\partial \pi (\tau^y, s) / \partial \tau^y < 0$ for $\tau^y \in (\lambda^- (s), \lambda^+ (s))$. Intuitively, a higher $\tau^y$ discourages $h$. The smaller number of rich individuals in the next period yields a lower $\pi$, which further decreases $h$. Hence, the incumbent may have the incentive to cut $\tau^y$ due to the more distortive $\tau^y$. Opposite to the self-reinforcing process, the link between $\tau^y$ and $\pi$ tends to dampen the intertemporal effect of ideology.

In the left-dominating (right-dominating) region with $\lambda^+ (s) \leq 0$ ($\lambda^- (s) \geq 1$), $h$ and $\tau^y$ have no impact on the next-period government’s identity. Therefore, $\pi$ is independent of $h$ and $\tau^y$, as in the exogenous political constituency.

4.3 The Equilibrium Tax Rule

Now we are well-equipped to solve for $\tau^y (s)$. By (19) and (27), tax revenues from young individuals are

\[
T (\tau^y, s) = \begin{cases} 
T^R (\tau^y, s) \equiv \frac{1}{2} (1 - \tau^y + \beta) \tau^y & \text{if } \tau^y \in [0, \lambda^- (s)] \\
T^M (\tau^y, s) \equiv \frac{1}{2} (1 - \tau^y + \beta \frac{2(ps + z) - \tau^y}{4z - \beta}) \tau^y & \text{if } \tau^y \in (\lambda^- (s), \lambda^+ (s)) \\
T^L (\tau^y, s) \equiv \frac{1}{2} (1 - \tau^y) \tau^y & \text{if } \tau^y \in [\lambda^+ (s), 1]
\end{cases}
\]
Taking $s$ as the state variable, $\tau^y$ can be pinned down by maximizing the piecewise quadratic function $T(\tau^y, s)$. A characterization of the policy rule $\tau^y(s)$ is given by

**Proposition 1** Assume that (23) and $z \geq \hat{z}$, where

$$\hat{z} \equiv \frac{\beta}{8 \left[ - (\beta^2 + 2\beta) + (1 + \beta) \sqrt{\beta^2 + 2\beta} \right]}.$$  

Then, the Markov perfect equilibrium is such that

$$\tau^y(s) = \begin{cases} 
\frac{1}{2} \phi(s) & \text{if } s \leq s^1 \\
\lambda^-(s) & \text{if } s \in [s^1, s_M^H] \\
\frac{1 + \beta}{2} & \text{if } s > s_R \end{cases},  \tag{29}$$

where

$$\phi(s) \equiv \frac{(2\beta (\rho s + z) + 4z - \beta)}{8z}, \quad s^1 \equiv \frac{\sqrt{z (4z - \beta) - (4z - \beta) / 2 - \beta z}}{\beta \rho},$$

$$s_M^H \equiv \frac{16z^2 - 6\beta z + 4z - \beta}{\rho (16z - 2\beta)}, \quad s_R \equiv \frac{(1 - \beta) / 4 + z}{\rho}.$$

**Proof**: See the appendix.

To simplify the statement in the paper, we assume that $z \geq \hat{z}$. The other case, where $z \in (\beta/4, \hat{z})$, delivers qualitatively similar results and will be investigated in Appendix 8.4.21 Panels A and B in Figure 1 plot the policy rule $\tau^y(s)$ and the ideology-contingent probability $\pi(\tau^y(s), s)$ under the benchmark parameter values, which are set to $z = 0.22$, $\rho = 0.47$ and $\beta = 0.67$.22

[Insert Figure 1]

In the left-dominating region, $\partial \pi(\tau^y, s) / \partial \tau^y = 0$, $h$ and $\tau^y$ have no effect on future election outcomes. (28) reduces to a quadratic function $(1 - \tau^y) \tau^y/2$, and the incumbent sets $\tau^y = 1/2$.

For $\lambda^+(s) > 0$ or, equivalently, $s > -z/\rho$, ideology becomes less hostile for the right-wing. $h$ and $\tau^y$ start to affect the distribution of future voters. The corresponding objective function $T$ is composed of two different quadratic functions, $T = T^M$ for low $\tau^y$ and $T = T^L$ for high $\tau^y$ (see Panel A of Figure 2). Since $h$ and the tax base are increasing in $\pi$, the incumbent would like to see a higher $\pi$. However, when $\lambda^+(s)$ is close to zero, $\tau^y$ needs to be cut substantially.

---

21 However, there will be no electoral uncertainty if $z < \hat{z}$.

22 The parameter values are calibrated to match some long-run electoral patterns in OECD countries. See Appendix 8.5 for details.
from 1/2 to a level lower than \( \lambda^+ (s) \), in order to affect \( \pi \) in (27). Proposition 1 shows that for \( s < s^1 \), it is still optimal for the incumbent to set \( \tau^y = 1/2 \) (see Panel B of Figure 2). Consequently, \( \pi \) remains to be zero.

[Insert Figure 2]

As \( s \) moves rightward, it is less costly to influence \( \pi \) by adjusting \( \tau^y \), as the top of \( T^M \) gets closer to the top of \( T^L \). In particular, when \( s \) reaches the threshold point \( s^1 \), we find an equalized maximum of the two quadratic functions. In other words, the incumbent is indifferent between \( \tau^y = 1/2 \) and \( \tau^y = \phi (s^1) \), where

\[
\phi (s^1) = \frac{\sqrt{4 - \beta/z}}{4}.
\]

This can be seen directly in Panel B of Figure 2. The indifference produces a discontinuity of \( \tau^y \) at \( s^1 \).\(^{23}\) For a small increment \( \xi \) in \( s \), the incumbent will cut \( \tau^y \) from 1/2 to \( \phi (s^1 + \xi) \), to attain the top of the Laffer curve. \( \pi \) becomes positive accordingly.

Next, we investigate a more realistic region \([s^1, s^M_H] \), where both parties have positive probabilities to win the next election. This is referred to as “the competitive political region.” The first observation is that now \( s \) plays a role in \( \pi \) (see the middle panel of Figure 1), which gives rise to the intertemporal effect of ideology as illustrated in Section 2.4. Consequently, \( \tau^y \) is increasing in \( s \) (see the top panel of Figure 1). Moreover, Proposition 1 shows that in this region, the optimal \( \tau^y \) always attains at the top of \( T^M \) in (28), which is equal to \( \phi (s) \in (\lambda^- (s), \lambda^+ (s)) \). For \( \tau^y \in (\lambda^- (s), \lambda^+ (s)) \), both \( h \) and \( \tau^y \) can influence \( \pi \) according to (27). As discussed in Section 4.2, the link between \( h \) and \( \pi \) establishes a self-reinforcing process, which amplifies the intertemporal effect of ideology. Meanwhile, the link between \( \tau^y \) and \( \pi \) makes \( \tau^y \) more distortive and, thus, dampens the intertemporal effect of ideology. Interestingly, (29) shows that the marginal effect of ideology on \( \tau^y \), \( d\tau^y (s)/ds \), equals \( \beta \rho/4z \) and is identical to \( d\tau^y (s)/ds \) in (25). Therefore, the two opposite effects cancel each other out.

When \( s = s^M_H \), the optimal \( \tau^y \) for maximizing \( T^M \) coincides with the boundary \( \lambda^- (s) \) (see Panel C of Figure 2). Here, the right-wing will win the next election with probability one. When ideology moves further towards the right, the optimal \( \tau^y \) will be equal to \( \lambda^- (s) \); i.e., the incumbent will set \( \tau^y \) such that \( \pi = 1 \) (see Panel D of Figure 2). Notably, this region features a stronger intertemporal effect of ideology.\(^{24}\) The reason is simple. The fact that \( \pi = 1 \) breaks

\(^{23}\)More specifically, \( \tau^y (s) \) is not lower hemi-continuous. The theorem of maximum (e.g., Stokey and Lucas, 1989, pp. 62) ensures only that \( \tau^y (s) \) is upper hemi-continuous.

\(^{24}\)According to Proposition 1, \( d\tau^y (s)/ds = \beta \rho/4z \) for \( s \in [s^1, s^M_H] \) and \( d\tau^y (s)/ds = 2 \rho \) for \( s \in (s^M_H, s^R] \). The latter is greater than the former since \( z > \beta/4 \).
down the link between $\tau^y$ and $\pi$, resulting in a less elastic private investment. The incumbent, therefore, has the incentive to increase further $\tau^y$. This illustrates how the intertemporal effect of ideology is amplified under the endogenous political constituency.

Finally, in the right-dominating region when $s > s^R$, young individuals rationally expect the right-wing to win the next election under any $\tau^y$. Consequently, (28) reduces to a quadratic function $T^R$, which solves $\tau^y = (1 + \beta)/2$.

Given the equilibrium tax rule $\tau^y(s)$ and the ideology-contingent probability $\pi(\tau^y(s), s)$, human-capital investment $h$ follows

**Proposition 2** Assume that (23) and $z \geq \hat{z}$. Then, the Markov perfect equilibrium is such that

$$h = \begin{cases} 
\frac{1}{4} + \frac{\beta(\rho s + z)}{1+\beta} & \text{if } s \leq s^1 \\
\frac{1}{4} - \frac{\rho s - z}{1+\beta} & \text{if } s \in [s^1, s^M_H] \\
\frac{1}{4} & \text{if } s \in (s^M_H, s^R] \\
\frac{1}{4} & \text{if } s > s^R.
\end{cases}$$

The proof is straightforward and immediately follows from (19), (27) and (29). $h$ increases in $s$ for $s \in [s^1, s^M_H]$ and peaks at $s = s^M_H$. Then, $h$ decreases in $s$ for $s \in (s^M_H, s^R]$. Panel C in Figure 1 plots the inverted U-shaped $h$. The non-monotonicity of $h$ is due to the fact that a right-wing ideology has two opposite effects on $h$. First, it helps the right-wing win the next election and, thus, increases $\pi$. This has a positive effect on $h$. However, a high $\pi$ also induces the incumbent to increase $\tau^y$, which has a negative effect on $h$. For $s \in [s^1, s^M_H]$, the positive effect dominates the negative effect. For $s \in (s^M_H, s^R)$, the positive effect disappears since $\pi$ has already reached its upper boundary, while the negative effect is still prevalent due to the increasing $\tau^y(s)$ in this region. This results in a decreasing $h$.

To conclude, the intertemporal effect of ideology on $\tau^y$ under the endogenous political constituency turns out to be similar to that under the exogenous political constituency. Particularly, if we focus on the competitive political region $[s^1, s^M_H]$, there is a positive relationship between the distortionary tax rate and the right-wing ideology, which sharply contrasts with the partisan effect in the static example.

### 4.4 Vote Shares and the Size of Government

This subsection investigates the main empirical prediction of our model, which will be tested in Section 6. Although (29) has provided a prediction on the correlation between ideology and $\tau^y$, there are two major difficulties in testing the prediction. The first one is how to measure ideology. A commonly used measure of ideology in political science literature is self-placement
scores of the left-right position from opinion polls or survey data (Inglehart, 1990). This approach obviously suffers from limited comparative observations across countries and time. Second, it is equally hard to find an empirical counterpart of $\tau^y$, though age-dependent taxation contains some realistic components. Given these concerns, we adopt an alternative approach: looking at government size and vote shares, for which data can easily be obtained. Since $s$ and $\tau^y$ are positively related to the right-wing’s vote share and government size within each political regime, respectively, we expect the correlation between vote shares and government size qualitatively similar to that between $s$ and $\tau^y$.

To see this precisely, we aggregate the two types of taxes in (8) and (29), and then compute the size of government as a percentage of aggregate output:

$$\gamma = \frac{2g}{y} = \left\{ \begin{array}{ll} \frac{h(s-1) + \tau^y(s) h(s)}{h(s-1) + h(s)} & \text{if } s \leq \frac{1}{2} - h(s-1) \\ \frac{\tau^y(s) h(s)}{h(s-1) + h(s)} & \text{otherwise} \end{array} \right.,$$

(32)

where $\tau^y(s)$ and $h(s)$ follow (29) and (31), respectively. Note that $\gamma$ is not only affected by the current ideological state $s$, but also depends on the past ideological state $s-1$, which determines the current size of the old rich. Due to the predetermined size of the old rich, an analytical characterization of the correlation between $\gamma$ and $e$ (the right-wing’s vote share) is not applicable. So, we first simulate the model and use simulated data to estimate the following linear equation.

$$\gamma = b_0 + b_1 R + b_2 e + \varepsilon,$$

(33)

where $\varepsilon$ is an error term and $R$ is a dummy variable which equals zero and one for the left- and right-wing regime, respectively. We run 1100 simulations 50 times with the benchmark parameterization. The estimated results are $b_1 = -0.4950$ ($-410.43$) and $b_2 = 0.1010$ (55.16), with $t$ statistics in brackets. Consistent with our theory, $b_1$ is negative and $b_2$ is positive, suggesting a negative partisan effect and a positive intertemporal effect, respectively. In addition, $R^2$ is 0.98, indicating a high degree of fitness of the linear specification.

4.4.1 Sensitivity to Model Parameters

Now, we check the parameter sensitivity of the coefficient of interests, $b_2$. Specifically, we analyze sensitivity to two key model parameters: $z$ and $\rho$. Panel A of Figure 3 shows that an
increase in $z$, implying more volatile ideology, tends to reduce $b^2$. This is consistent with (29); the intertemporal effect of ideology, captured by $d\tau_y(s)/ds = \beta\rho/4z$, mitigates as $z$ increases. The effect of $z$ on $b^2$, however, is non-monotonic. When $z$ is sufficiently large, $b^2$ turns out to be increasing in $z$. A larger $z$ increases the likelihood for $s$ to fall into $(s_H^L, s_R]$, where the intertemporal effect gets stronger (recall that $d\tau_y(s)/ds = 2\rho$ for $s \in (s_H^L, s_R]$). This yields a larger $b^2$. The U-shaped estimates of $b^2$ in Panel A suggest that the second effect tends to dominate for large $z$.

[Insert Figure 3]

Panel B shows the estimates of $b^2$ w.r.t. $\rho$. When $\rho = 0$, the intertemporal effect of ideology goes away. This implies a zero $b^2$. Adding persistence into the stochastic process of ideology $\rho$ gives rise to the intertemporal effect and, therefore, a positive $b^2$. Similar to the effects of $z$ on $b^2$, $\rho$ also has two opposite effects on $b^2$. On the one hand, $\rho$ increases the magnitude of the intertemporal effect, $d\tau_y(s)/ds$ for $s \in (s^1_H, s^M_H] \cup (s^M_H, s^R]$. On the other hand, the region $(s^1_H, s^M_H] \cup (s^M_H, s^R]$ shrinks with a larger $\rho$, suggesting a lower likelihood for the intertemporal effect to be functioning. This reduces the estimate of $b^2$. Different from the U-shaped estimates of $b^2$ w.r.t. $z$, Panel B shows that the aggregate effect of $\rho$ seems monotonic; i.e., the positive effect always dominates the negative one.

To conclude, we find a positive $b^2$, implying that an increase in the right-wing voter share leads to a larger government within each political regime. The result is robust to a wide range of parameter values. This allows us to test our theory against the standard partisan theory predicting a zero $b^2$. The empirical analysis will be conducted in Section 6.

### 4.5 Age-Independent Taxation

Throughout the paper, we maintain the assumption that the government can condition taxes on age. Although age-dependent taxation has its realistic counterpart and substantially simplifies the analysis, this assumption is not innocuous. One may wonder whether binary taxation (8), which obviously overstates the partisan effect of ideology, is crucial for the positive relationship between $s$ and $\tau_y$. An earlier version of this paper (Song, 2005, chapter 2) assessed the robustness of the main result under age-independent taxation and found that imposing the weaker policy instrument does not lead to any major change. The intuition is simple. Age-independent tax rates in the right-wing regime are, on average, lower than those in the left-wing regime. A right-wing ideology, as in our benchmark case with age-dependent taxation, reduces the expected tax rate by increasing $\pi$. This encourages investment and induces the incumbent to behave in a way similar to that described above.
5 Alternative Political Objectives

We have characterized equilibria in which only the old vote. This assumption seems too extreme, as (old) politicians entirely ignore the welfare of the young in policy decision-making. A natural extension is to assume that politicians (party candidates) are altruistic towards the young. The political objective function can, therefore, be written as

\[ W^L = u^{o_{u}} + \tilde{\omega} u^{y}, \]  
\[ W^R = u^{o_{s}} + \tilde{\omega} u^{y}, \]

where \( \tilde{\omega} \geq 0 \) is the intensity of altruism. To avoid trivial results, we assume that \( a^{o} + \tilde{\omega} a^{y} < 2 \). Appendix 8.6 proves the following proposition.

Proposition 3 Assume political objective functions (34) and (35). Then,

(i) \( \tau^{o} \) follows (8).

(ii) \( h \) and \( \pi \) satisfy (14) and (17), respectively,

(iii) Define

\[ V \equiv \tau^{y} h + \omega \left( (1 - \tau^{y} + \beta \pi) h + \frac{a^{o}}{2} \beta (1 - \pi) h - \frac{h^{2}}{2 \text{ investment costs}} \right), \]

where \( \omega \equiv \frac{2 \tilde{\omega}}{a^{o} + \tilde{\omega} a^{y}} \). Then, \( \tau^{y} \) solves

\[ \tau^{y} = \arg \max_{\tau^{y} \in [0,1]} V. \]

(iv) The political outcomes associated with (34) and (35) are identical to those under probabilistic voting à la Lindbeck and Weibull (1987), in which politicians maximize a weighted average of individual utilities:

\[ (h_{-1} + \hat{s}) u^{o_{s}} + (1 - h_{-1} - \hat{s}) u^{o_{u}} + \tilde{\omega} u^{y}, \]

where \( \hat{s} \equiv s + (a^{o} + \tilde{\omega} a^{y} - 1)/2 \).

Three remarks are in order. First, due to age-dependent taxation, introducing altruism towards the young does not affect the decision on \( \tau^{o} \). Therefore, the binary rule (8) carries over to the model with altruism. It follows immediately that the equilibrium \( \pi \) still satisfies the functional equation (17). This property substantially simplifies the analysis below. Second, when choosing \( \tau^{y} \), politicians face a trade-off between tax revenues and the welfare of the young, shown as a sum of their lifetime earnings, redistributive benefits and costs of human-capital.
investment in \((36)\).\footnote{Finally, the last part of the proposition shows that the probabilistic voting à la Lindbeck and Weibull (1987) can be considered a micro-foundation for political objective functions \((34)\) and \((35)\). In particular, the political weight on the left-wing (right-wing) is equal to the population of the poor (rich) plus (minus) the ideological state (subject to a linear transformation).

The solid lines in Figure 4 depict the equilibrium policy rules of \(\tau^y\), \(\pi\) and \(h\) when \(\omega = 0.1\).\footnote{The dotted lines are those from the benchmark model in Figure 2. The results are qualitatively similar. This politico-economic equilibrium also features an increasing \(\tau^y\) as ideology leans towards the right in the competitive political region. The human-capital investment \(h\) is, again, non-monotonically related to \(s\). Nevertheless, two quantitative changes deserve mention. First, \(\tau^y\) is lower than that in the benchmark model. Second, the competitive political region moves towards the left, and both \(\pi\) and \(h\) become larger. In other words, the right-wing party will be more likely to win the election when politicians are altruistic towards the young. The intuition is straightforward: The young dislike distortionary taxes. Therefore, altruistic politicians have the incentive to cut \(\tau^y\), resulting higher \(h\) and \(\pi\).}

\[\text{[Insert Figure 4]}\]

5.1 Ideology-Dependent Altruism

In this subsection, we extend our theory further by introducing ideology-dependent altruism. Note that altruism in \((34)\) and \((35)\) is independent of ideology. Politicians care about the ex-ante welfare of the young, irrespective of their ex-post economic status of being rich or poor. Consequently, as in the benchmark model, both parties are perfectly aligned with \(\tau^y\) and the partisan effect works only for \(\tau^o\). Now, we modify the political objective functions as follows.

\[
\begin{align*}
W^L &= u^o + \omega u^y, \\
W^R &= u^o + \omega u^y,
\end{align*}
\]

\[
\begin{align*}
W^L &= u^o, \\
W^R &= u^o + \omega u^y,
\end{align*}
\]

where

\[
\begin{align*}
u^y &= a^y g + \beta a E \left[ g \right], \\
u^y &= 1 - \tau^y + a^y g + \beta E \left[ 1 - \tau^o + a^o g \right],
\end{align*}
\]

\footnote{\(1 - \pi\) \(h\) in the middle term of the bracket in \((36)\) is the next-period redistribution if the left-wing party wins the election.}

\footnote{We set a low \(\omega\) in order to facilitate the visual comparison with results from the benchmark model in Figure 4. The main feature that \(\tau^y\) increases in \(s\) holds true even if \(\omega = 1\).}
stand for the ex-post utilities of the young poor and the young rich, respectively. Altruism in (38) and (39) is ideology-dependent: The left-wing (right-wing) party candidate cares only about the ex-post welfare of the young poor (rich). Appendix 8.7 proves the following proposition.

**Proposition 4** Assume the political objective functions (38) and (39).

(i) $\tau^o$ still follows (8).

(ii) $h$ and $\pi$ satisfy (14) and (17), respectively.

(iii) Define

\begin{align*}
V^L &\equiv h\tau^y + \omega \left( \frac{a^o \beta (1 - \pi) h}{\text{ex-post welfare of the young poor}} \right), \\
V^R &\equiv h\tau^y + \omega \left( 1 - \tau^y + \beta \pi + \frac{a^o}{2} \beta (1 - \pi) h \right),
\end{align*}

where $\omega \equiv \frac{2 \hat{\omega} a^o}{\hat{\omega} + a^o \hat{\omega}_y}$. Then, $\tau^{yi}$ solves

\[ \tau^{yi} = \arg\max_{\tau^y \in [0,1]} V^i, \]

where $i \in \{L, R\}$.

The first and second parts of the proposition come directly from age-dependent taxation, as in Proposition 3. (40) illustrates two additional effects on the determination of $\tau^y$ when the left-wing party candidate has ideology-dependent altruism (38). On the one hand, given $\pi$, $V^L$ is increasing in $h$ since a larger $h$ increases the next-period redistributive benefits. Consequently, the left-wing would like to cut $\tau^y$ in a cynical way that contradicts their political colors. This is an analogue of “the strategic effect” in Persson and Svensson (1989). On the other hand, given $h$, $V^L$ is decreasing in $\pi$. Since the next-period redistributive policy benefits the current young poor, the left-wing have the incentive to increase their next-period election probability, $1 - \pi$. This is referred to as “the opportunistic effect,” reflecting an incumbent’s re-election concerns. Clearly, such an effect results in a higher $\tau^y$ and a lower $h$ than those in the benchmark model.\(^{31}\)

Now, let us look at the right-wing’s objective function (41). An immediate observation is that (41) differs from (40), implying a partisan effect on $\tau^y$. The partisan effect consists of two components. First, given $\pi$, $V^R$ is increasing in $h$ (or decreasing in $\tau^y$). This captures the direct effect of the ideology-dependent altruism, since the young rich are averse to $\tau^y$. On

\(^{31}\)The opposite strategic and opportunistic effects can also be found in Jonsson (1997).
top of that, (41) implies an opportunistic effect similar to that discussed above. Since the current young rich would suffer from the next-period redistributive policy, the right-wing have the incentive to increase $\pi$. Specifically, one can see that given $h$, $V^R$ is increasing in $\pi$ by the assumption that $a^o < 2$.

The equilibrium policy rules of $\tau^y$, $\pi$ and $h$ are plotted in Figure 5 (solid lines). As in Figure 4, the dotted lines are those from the benchmark model. It can be seen that in the left-wing region, the results associated with ideology-dependent altruism (38) are quantitatively close to those in the benchmark model. This suggests that the strategic and opportunistic effects largely cancel each other out. The right-wing regime features a lower $\tau^y$ since both the direct and opportunistic effects tend to reduce $\tau^y$ when the right-wing is in office. Despite this difference, our major finding is robust to the introduction of ideology-dependent altruism: The equilibrium $\tau^y$ is still increasing in $s$ within each political regime.

[Insert Figure 5]

6 Empirical Evidence

In this section, we test the main prediction of our theory regarding the intertemporal effect of ideology: After controlling for the partisan effect, a more right-leaning political constituency should lead to a larger government.

6.1 Data and Specification

We use the Comparative Welfare States Data Set assembled by Huber et al. (1997) and updated by Brady et al. (2004). The sample consists of, at most, 41 years of observations (1960-2000) from 18 democracies, including Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Sweden, Switzerland, the United Kingdom and the United States.32

Following Persson (2002) and Persson and Tabellini (2004), among many others, we use central government expenditure (or revenue) as a percentage of GDP, denoted by $CGEXP$ (or $CGREV$), to measure the size of government. We use the percentage of vote for the right-wing parties ($RVOT$) as the empirical counterpart of the proportion of right-wingers in the model. A dummy variable, $R$, is created for controlling the partisan effect. We let $R = 1$ if both shares of seats of the right and center parties in parliament and government are larger than 66.6 percent. $R = 0$ otherwise. As will be seen below, such a super-majority of the right-wing

\[32\text{All the data used in this subsection are from the Comparative Welfare States Data set, which is available at http://www.lisproject.org/publications/welfaredata/welfareaccess.htm.}\]
helps to empirically illustrate the partisan effect.\footnote{The partisan effect would be statistically insignificant if we use a simple majority as the criterion of setting the dummy variable \(R\).} Of 738 observations, 201 are associated with \(R = 1\). In words, about 27 percent governments in the sample are labeled as right-wing.\footnote{A similar criterion is adopted by Woldendorp \textit{et al.} (1993, 1998), where \(R = 1\) is referred to as the "right-wing dominance" regime.}

Based on (33), we estimate regressions in which policy outcomes \(Y_{it}\) are linear functions of \(R\) and \(RVOT\).

\[
Y_{it} = b_0^i + b_1^i R_{it} + b_2^i RVOT_{it} + \phi X_t + \varepsilon_{it}, \tag{43}
\]

where \(b_0^i\) is a country-specific effect, and \(X_t\) is a set of year dummies to control for the unobserved common shocks. We also run this regression with some additional explanatory variables containing country-specific factors. Specifically, we use the log of real GDP per capita (\(YPC\)) to control the potential impact of Wagner’s Law; i.e., the size of government will rise as income rises. Deviation of \(YPC\) from its trend (obtained by the HP filter), denoted by \(YGAP\), is added, as government tends to implement countercyclical policies to smooth economic fluctuations. Other control variables include the unemployment rate (\(UNEMPL\)), export plus import share of GDP (\(OPEN\)), proportion of population over 65 (\(POP65O\)) and below 14 (\(POP14U\)), which are constantly adopted in recent empirical studies (e.g., Razin \textit{et al.}, 2004; Persson and Tabellini, 2005). Finally, since the pre-determined debt level may affect government revenues, we include the debt/GDP ratio as a control variable in the regression when using government revenues as the dependent variable.\footnote{We thank a referee for raising this concern. Excluding the debt/GDP ratio yields essentially the same results.}

6.2 Results

Table 1 gives estimation results from fixed-effects regressions. We start with column 1, where the intertemporal effect of ideology is shut down by excluding \(RVOT\) in (43). This parallels the standard approach for testing the partisan effect. The baseline regression (column 1) shows that \(b_1^i\) is negative, but statistically insignificant. When additional explanatory variables are included (column 2), \(|b_1^i|\) increases from 0.82 to 1.05 and becomes significant at 10\% level. According to the point estimation, switching from a right-wing to a left-wing regime causes government expenditures to increase by about one percent of GDP. The magnitude of the partisan effect in OECD countries is roughly in line with previous findings.\footnote{In Blais \textit{et al.} (1993), such a shift from the right to the left will increase government spending by 0.7 percentage point. In Perotti and Kontopoulos (2002), the increase by a shift from a modest right government to a modest left government amounts to 1.6 percentage points in the steady state.} Column 3 and 4 display the same regressions on \(CGREV\). There is a much stronger partisan effect for
government revenue: The estimated $|b^0|$ amounts to 1.8 and is significant at 1% level. This result is rather stable, irrespective of whether additional explanatory variables are included.

[Insert Table 1]

Our main finding is in columns 5 to 8. The baseline regression (column 5) shows that $b^2$, the coefficient on RVOT, is positive and significant at the level of 1%. Including additional explanatory variables reduces the estimate of $b^2$ substantially (column 6). The statistical significance, however, is the same. The point estimation in column 6 implies that given the incumbent’s political identity, a one-percentage-point increase in the vote share for right parties will, on average, increase government expenditures by 0.06 percent of GDP. The intertemporal effect of ideology is more evident in columns 7 and 8, where we use central government revenue as an indicator of government size. A one-percentage-point increase in RVOT will lead to 0.17 percentage points increase in CGREV when additional explanatory variables are added. These results provide statistical evidence in favor of the intertemporal effect of ideology.

The positive $b^2$ does not suggest that right-wing governments will be larger. In fact, controlling for political constituency reveals a quantitatively larger partisan effect. The estimate of $b^1$ that is significant in columns 1 to 4 remains significant in columns 5 to 8. This suggests that left-wing governments are indeed more favorable towards public spending. The co-existence of the two effects is essential for our theory, since the intertemporal effect cannot exist without the partisan effect.

6.3 Robustness

A key issue is whether RVOT is an appropriate proxy of political constituency. Party ideology, on which RVOT is completely silent, is clearly an important dimension of ideology. Our model adopts a two-party system for simplicity. However, real democracies have much more complex political systems. One country may have several left (right) parties with differing ideological positions. Taking this issue into consideration, we use the index developed by Kim and Fording (1998, 2003) as an alternative proxy of political constituency. The advantage of the index is that it reflects not only changes in voting, but also changes in party ideology.\footnote{Kim and Fording first estimate party ideology, based on party manifesto statements, and then use the percentage of the vote received by each party to construct an adjusted index for the median ideological position.} Using the Kim-Fording index does not qualitatively change to the results. The coefficients that are significant remain significant.\footnote{The results are available upon request.}
Our panel regressions contain 18 countries. It is important to check the sensitivity of the results to individual countries. To this end, we run all the regressions in Table 1, excluding one country at a time. In every regression, the coefficients that are significant at the 5% level are still significant at the same level no matter which country is excluded.

Some evidence shows that fiscal policies legislated in the United States typically take one year to be effective (Poterba, 1994, Gilligan and Matsusaka, 1995). It is not clear whether the same mechanism carries over to the sample of OECD countries. Be it as it may, we replace $R_{it}$ in the regressions with one-year lagged variable $R_{it-1}$. The results change only marginally, and the statistical significance for the coefficients of interest is never affected.

We have not yet addressed that there might be omitted variables affecting political variables and fiscal outcomes simultaneously. Concerning this, we choose one-year lagged political variables $R_{it-1}$ and $RVOT_{it-1}$ to instrument the current political variables $R_{it}$ and $RVOT_{it}$. These lagged variables are highly correlated with the respective current values and are expected to be independent of current policy outcomes. Table 2 reports the results of reestimating (43) using Two-State-Least Squares (2SLS). First, note that the 2SLS estimates of $b^1$ are significant in all cases. There is some evidence for the endogeneity of $R$ since a Hausman test rejects the null hypothesis in columns 3 and 4 at the level of 10%. Second, though the endogeneity leads to a significant underestimation of the partisan effect, the endogeneity of $RVOT$ is much less severe. Application of the Hausman test cannot reject the null hypothesis. The 2SLS estimate of $b^2$ becomes statistically insignificant in column 2, but remains highly significant in all other cases.

[Insert Table 2]

Finally, our reduced-form empirical analysis relies on an important hypothesis that the fraction of right-wing voters helps to predict the identity of the next government. According to the dummy variable $LR$, our data set has only 24 political regime shifts in 18 countries over the sample period from 1960 to 2000. The limited variation does not allow us to run a panel regression to test the hypothesis. We then use the total shares of seats of the right and center parties in parliament and government, $LR$, as a proxy of the identity of the government. Note that, by construction, $LR$ and $LR$ are highly correlated: The correlation coefficient is equal to 0.84. We run the following regression:

$$LR_{it} = a^0 + a^1RVOT_{it-1} + \phi X_t + \epsilon_{it},$$

---

39 We thank a referee for pointing out this concern.
where $a_i^0$ is a country-specific effect, $RVOT_{it-1}$ is the lagged vote share for the right-wing parties, and $X_t$ is a set of year dummies to control for the unobserved common shocks. The estimated key coefficient, $a^1$, is equal to 0.199 and statistically significant at the level of 1%. The result is very robust to adding additional explanatory variables containing country-specific factors.

To conclude, we find a significantly positive relationship between government revenue and the right-wing vote share, after controlling for the partisan effect. The point estimate is quite stable to a number of control variables and specifications. The empirical finding is in line with our theoretical prediction, but hard to explain by the existing literature. An interesting remaining question is why the positive relationship is less evident for government expenditure? The less significant partisan effect for government expenditure might be an important reason. Moreover, recall that in the model, the intertemporal effect of ideology is driven mainly by the government’s target of tax-revenue maximization. Therefore, as an indicator of tax policy, $CGREV$ is perhaps a better measure of $\gamma$ for identifying the intertemporal effect of ideology.\footnote{Although government expenditures and revenues usually move in tandem, imbalances between expenditures and revenues occur occasionally since governments may run deficits, which is totally assumed away in the theory.}

7 Conclusion

In spite of the growing literature on public policy decision-making in dynamic politico-economic equilibria, most works are silent on the role of ideological shifts, which tend to be persistent and have a nontrivial influence on political outcomes. To explore the underlying mechanism of policy decision-making under stochastic ideological movements, we develop a tractable model to investigate the dynamic interactions among public policy, individuals’ intertemporal choice and the evolution of political constituency. Our main finding is that the relationship between right-wing ideology and the size of government is positive within each political regime. This distinguishes the literature of partisan politics predicting that ideology has no effect on public policies if the political regime remains unchanged. We document empirical evidence from an OECD panel that supports our theory.

Our analysis is subject to a number of caveats. For instance, our theory is completely silent on the determination of public policies under coalition government. Moreover, we abstract away public debt. When a government is allowed to borrow, however, public choices may appear to be strategic. In a related work, Song, Storesletten and Zilibotti (2007) analyze the determination of public debt in a stochastic ideological environment. In that model, however, private choices are irrelevant for the evolution of political constituency. It will be interesting
future research to incorporate public debt in the current setup, to see how public intertemporal trade-off interacts with private intertemporal choices.

References


8 Appendix

8.1 Proof of Lemma 1

Apply the Schauder fixed-point theorem. Let \( C \) be a set of bounded and uniformly continuous functions mapping from \([0,1] \times X\) to \([0,1]\). Define \( F = \int_{s' \geq \tau_y - \beta \alpha (\tau_y, s)} f(s') \, ds' \), where \( \alpha (\tau_y, s) \in C \). We need to prove that the mapping \( F \) has a fixed point.

Let \( \Omega = \{ F(\alpha), \alpha \in C \} \). We first claim that \( \Omega \) is equicontinuous; i.e., \( F(\alpha) \) is bounded and uniformly continuous for any \( \alpha \in C \). The boundedness is trivial since \( F(\alpha)(\tau_y, s) \leq \int f(s') \, ds' = 1 \). To prove that \( F(\alpha) \) is uniformly continuous, we pick up any two vectors \( x = (\tau_y, s_1) \) and \( y = (\tau_y, s_2) \) from \([0,1] \times X\). It is straightforward to show that

\[
|F(\alpha)(\tau_y, s_1) - F(\alpha)(\tau_y, s_2)| \\
= \left| F(\alpha)(\tau_y, s_1) - \int_{s' \geq \tau_y - \beta \alpha (\tau_y, s)} f(s') \, ds' \right| \\
\leq \left| \int_{s' \geq \tau_y - \beta \alpha (\tau_y, s)} f(s') \, ds' - \int_{s' \geq \tau_y - \beta \alpha (\tau_y, s)} f(s') \, ds' \right| \\
\quad \quad + \left| \int_{s' \geq \tau_y - \beta \alpha (\tau_y, s)} f(s') \, ds' - \int_{s' \geq \tau_y - \beta \alpha (\tau_y, s)} f(s') \, ds' \right| \\
\leq \left| \tau_y - \tau_y - \beta (\alpha (\tau_y, s_1) - \alpha (\tau_y, s_2)) \right| \|f(s')\|_{\sup} \\
\quad \quad + \sup s' - \frac{\tau_y - \beta \alpha (\tau_y, s_2)}{2} \|f(s')\|_{\sup}.
\]

As \( \|x - y\| \to 0 \), we have \( |(\tau_y - \tau_y - \beta (\alpha (\tau_y, s_1) - \alpha (\tau_y, s_2))) / 2| \|f(s')\|_{\sup} \to 0 \) by the uniform continuity of \( \alpha \) and \( \|f(s')\|_{\sup} < \infty \) (A2). Moreover, from A1 and the boundedness of \( \alpha \), it immediately follows that \( \sup s' - \frac{\tau_y - \beta \alpha (\tau_y, s_2)}{2} < \infty \). By A2, \( \|f(s') - f(s')\|_{\sup} \to 0 \) as \( \|x - y\| \to 0 \). Therefore, we have \( |F(\alpha)(x) - F(\alpha)(y)| \to 0 \) as \( \|x - y\| \to 0 \).

Next, we check the conditions of the Schauder fixed-point theorem (Theorem 17.4, Stokey and Lucas, 1989). \( \Omega \) has been proved equicontinuous. And it is easily shown that \( C \) is nonempty, closed and convex, and \( F \) is continuous. Thus, all conditions are satisfied. □

8.2 Proof of Lemma 2

We only need to prove that, given any \((\tau_y, s)\), the following equation has a unique solution

\[
x = G(x) \equiv \int_{s' \geq \tau_y - \beta \alpha (\tau_y, s)} f(s') \, ds'.
\]

A3 implies that \( dG(x) / dx = \beta f(s') / 2 < 1 \). The proof is complete by applying the contraction mapping theorem. □
8.3 Proof of Proposition 1

The proof is based on the following lemma.

**Lemma 3** Assume that (23) and \( z > \beta /4 \).

(i) If \( z \geq \hat{z} \),

\[
\tau^y(s) = \begin{cases} 
\frac{1}{2} & \text{if } s \leq s^1 \\
\phi(s) & \text{if } s \in [s^1, s^M_H] \\
\lambda^-(s) & \text{if } s \in (s^M_H, s^R] \\
\frac{1 + \beta}{2} & \text{if } s > s^R
\end{cases},
\]  

(ii) If \( z < \hat{z} \),

\[
\tau^y(s) = \begin{cases} 
1/2 & \text{if } s \leq s^2 \\
\lambda^-(s) & \text{if } s \in [s^2, s^R] \\
\frac{1 + \beta}{2} & \text{if } s > s^R
\end{cases}.
\]  

where

\[
\hat{z} = \frac{\beta}{8 \left(1 - (\beta^2 + 2\beta) + (1 + \beta) \sqrt{\beta^2 + 2\beta}\right)},
\]

\[
s^1 = \frac{\sqrt{z(4z - \beta) - (4z - \beta) / 2 - \beta z}}{\beta \rho},
\]

\[
s^2 = \frac{1 - \beta - \sqrt{\beta^2 + 2\beta + 4z}}{4\rho},
\]

\[
s^M_H = \frac{16z^2 - 6\beta z + 4z - \beta}{\rho (16z - 2\beta)}.
\]

\[
s^R = \frac{(1 - \beta) / 4 + z}{\rho}.
\]

8.3.1 Proof of Lemma 3

The solution of maximizing (28) is straightforward under two polarized cases, i.e., \( s \geq ((1 - \beta) / 2 + z) / \rho \) and \( s \leq -z / \rho \). Thus, we need only to focus on \( s \in (-z / \rho, ((1 - \beta) / 2 + z) / \rho) \).

For notational convenience, we define

\[
L(s) = \max_{\tau^y \in (\lambda^+(s), 1]} T^L(\tau^y, s) \quad \tau^L(s) = \arg\max_{\tau^y \in (\lambda^+(s), 1]} T^L(\tau^y, s)
\]

\[
M(s) = \max_{\tau^y \in (\lambda^-(s), \lambda^+(s))} T^M(\tau^y, s) \quad \tau^M(s) = \arg\max_{\tau^y \in (\lambda^-(s), \lambda^+(s))} T^M(\tau^y, s)
\]

\[
R(s) = \max_{\tau^y \in [0, \lambda^-(s)]} T^R(\tau^y, s) \quad \tau^R(s) = \arg\max_{\tau^y \in [0, \lambda^-(s)]} T^R(\tau^y, s)
\]
It is also convenient to classify the regions where interior solutions hold.

\[
\tau^L(s) = \begin{cases} 
\frac{1}{2} & \text{if } s \leq s^L, \\
\lambda^+(s) & \text{if } s > s^L,
\end{cases}
\]

(47)

\[
\tau^M(s) = \begin{cases} 
\phi(s) & \text{if } s \in [s^M_L, s^M_H], \\
\lambda^-(s) & \text{if } s > s^M_H, \\
\lambda^+(s) & \text{if } s < s^M_L
\end{cases},
\]

(48)

\[
\tau^R(s) = \begin{cases} 
\frac{1+\beta}{2} & \text{if } s \geq s^R, \\
\lambda^-(s) & \text{if } s < s^R,
\end{cases}
\]

(49)

where

\[
s^L \equiv \frac{1/4 - z}{\rho},
\]

\[
s^M_{L} = \frac{-16z^2 + 2\beta z + 4z - \beta}{\rho (16z - 2\beta)},
\]

and

\[
s^R > s^M_H > s^M_L,
\]

(50)

\[
s^R > s^L > s^M_L.
\]

(51)

We proceed by classifying the following six cases (see Table A-1), according to the conditions in (47) and (48). Some results are immediate. By (51), the sixth case is empty. In Cases 1-3, \(s \leq s^L\). So, \(\tau^L(s) = 1/2\). In Cases 4 and 5, \(\tau^M(s) = \lambda^+(s)\). The continuity of \(T(s)\) implies that \(M(s) \geq L(s)\). So we are left to compare \(M(s)\) and \(R(s)\) in Cases 4 and 5.

<table>
<thead>
<tr>
<th>Case 1: if (s \leq s^L) and (s \in [s^M_L, s^M_H])</th>
<th>Case 2: if (s \leq s^L) and (s &gt; s^L)</th>
<th>Case 3: if (s &gt; s^L) and (s &lt; s^L)</th>
<th>Case 4: if (s &gt; s^L) and (s \in [s^M_L, s^M_H])</th>
<th>Case 5: if (s &gt; s^L) and (s &gt; s^L)</th>
<th>Case 6: if (s &gt; s^L) and (s &lt; s^M_L)</th>
</tr>
</thead>
</table>

Now let us consider the five cases in order. In the first case, \(\tau^M(s) = \phi(s)\). Moreover, (50) and (49) give that \(s < s^R\) and \(\tau^R(s) = \lambda^-(s)\). The continuity of \(T(s)\) implies that \(M(s) > R(s)\). So, we only need to compare \(L(s)\) and \(M(s)\). It immediately follows that \(\tau^L(s) = 1/2\) if \(s \leq s^1\) and \(\tau^M(s) = \phi(s)\) if \(s \geq s^1\), where \(s^1\) solves

\[
T^M(\phi(s^1), s^1) = T^L(1/2, s^1).
\]

This yields

\[
s^1 = \frac{\sqrt{z(4z - \beta)} - (4z - \beta) / 2 - \beta z}{\beta \rho}.
\]

(52)

The other root is omitted since \(s^1 > -z/\rho\). It is easy to see that

\[
s^1 < s^L.
\]

(53)
Moreover, by the assumption that $4z > \beta$, one can show that

$$s^1 > -\frac{z}{\rho} + \frac{\beta}{4\rho} > s^M_L.$$ (54)

(53) and (54) will be used in obtaining (57).

Turn to the second case, where $\tau^M(s) = \lambda^-(s)$. Since $s^L < s^R$, $\tau^R(s) = \lambda^-(s)$ and $M(s) = R(s)$. So, we have that $\tau^y(s) = 1/2$ if $s \leq s^2$ and $\tau^y(s) = \lambda^-(s)$ if $s \geq s^2$, where $s^2$ solves

$$TR(\lambda^-(s^2), s^2) = TL(1/2, s^2).$$

This yields

$$s^2 = \frac{1 - \beta - \sqrt{\beta^2 + 4\beta + 4z}}{4\rho}.$$ (55)

Since $L(s^1) > R(s^1)$, $L(s^1) = L(s^2)$ and $R(s)$ is increasing in $s$, $R(s^2) = L(s^2)$ implies that

$$s^2 \geq s^1,$$ (56)

Similarly, since $s^M_L < s^R$, $\tau^R(s) = \tau^M(s) = \lambda^-(s)$ and $M(s) = R(s)$ in the third case. So $\tau^y(s)$ follows the same rule as in the second case.

In the fourth case, $\tau^M(s) = \phi(s)$. Moreover, (50) and (49) establish that $s < s^R$ and $\tau^R(s) = \lambda^-(s)$. Hence, $M(s) > R(s)$ and $\tau^y(s) = \phi(s)$.

Finally, the fifth case gives that $\tau^M(s) = \lambda^-(s)$. If $s \leq s^R$, $\tau^R(s)$ is also equal to $\lambda^-(s)$ and $R(s) = M(s)$. On the other hand, if $s > s^R$, $\tau^R(s) = (1 + \beta)/2$, $R(s) > M(s)$. So, $\tau^y(s) = \lambda^-(s)$ if $s \leq s^R$ and $\tau^y(s) = (1 + \beta)/2$ otherwise.

To conclude, we have

$$\tau^y = \begin{cases} 
\frac{1}{2} & \text{if } s \in (\lceil s, s^L \rceil \cap \lceil s, s^1 \rceil) \cup (\lceil s, \min \{s^L, s^2\} \rceil \cap (s^M_H, \bar{s}) ) \cup (\lceil s, \min \{s^L, s^2\} \rceil \cap \lceil s, s^M_L \rceil) \\
\phi(s) & \text{if } s \in (\lceil s, s^L \rceil \cap \lceil s^1, \bar{s} \rceil \cap [s^M_L, s^M_H] ) \cup (\lceil s^L, \bar{s} \rceil \cap [s^M_L, s^M_H] ) \\
\lambda^-(s) & \text{if } s \in (\lceil s^2, \bar{s} \rceil \cap \lceil s, s^L \rceil \cap (s^M_H, \bar{s}) ) \cup (\lceil s^2, \bar{s} \rceil \cap [s^L, s^M] ) \cup (\lceil s, s^R \rceil \cap (s^L, \bar{s}) ) \cup (\lceil s, s^L \rceil \cap (s^M_H, \bar{s}) ) \\
\frac{1+\beta}{2} & \text{if } s \in (s^L, \bar{s}] \cap (s^R, \bar{s}] \cap (s^M_H, \bar{s}) )
\end{cases}.$$ (57)

The first line on the RHS of (57) comes from the results in Cases 1, 2 and 3. The second line follows the results in Cases 1 and 4. Cases 2, 3, and 5 give the third line, and the last line collects the result from Case 5.
To simplify (57), now we need to further classify two cases: $s^M_H < s^1$ and $s^M_H \geq s^1$. $s^M_H < s^1$ is equivalent to

$$2 - \frac{1}{4} \beta/z > (1 + \beta) \sqrt{(4 - \beta/z)}$$

after some algebra. It follows that $s^M_H < s^1$ if and only if

$$z < \hat{z}.$$ 

When $z < \hat{z}$, (56) establishes that $s^M_H < s^1 \leq s^2$. Moreover, when $z < \hat{z}$, one can show that $s^L > s^2$ must hold. Together with (50), (51), (53), (54) and (56), (57) can be reduced to (46). Finally, since $-(\beta^2 + 2\beta) + (1 + \beta) \sqrt{\beta^2 + 2\beta} < 1/2$ always holds, $z \in (\beta/4, \hat{z})$ is not an empty set. Similarly, when $z \geq \hat{z}$, (56) establishes that $s^M_H \geq s^2 \geq s^1$. Then (57) can be written as (45). □

8.4 $z \in (\beta/4, \hat{z})$

**Proposition 5** Assume that (23) and $z \in (\beta/4, \hat{z})$. Then, the Markov perfect equilibrium is such that

$$\tau^y(s) = \begin{cases} \frac{1}{\pi} - (s) & \text{if } s \leq s^2 \\ \frac{1}{\lambda} - s & \text{if } s \in [s^2, s^R] \\ \frac{1}{\lambda} + s & \text{if } s > s^R \end{cases}.$$  

(58)

**Proposition 6** Assume that (23) and $z \in (\beta/4, \hat{z})$. Then, the Markov perfect equilibrium is such that

$$h = \begin{cases} \frac{1}{\pi} - (ps - z) & \text{if } s \leq s^2 \\ \frac{1}{\lambda} - s & \text{if } s \in [s^2, s^R] \end{cases}.$$  

(59)

The proof is straightforward and immediately follows from Lemma 3, (19), (27) and (29). One can see that the implications of Propositions 4 and 5 are qualitatively the same as those of Propositions 1 and 2. It is worth noting that $\pi(\tau^y(s), s) = 0$ or 1. That is to say, there is no electoral uncertainty under a small $z$.

8.5 Calibration

We first use the panel for OECD countries in Section 6.1 and run the following regression to estimate the persistence of $RVOT$ (the percentage of vote for the right-wing parties):

$$RVOT_{it} = \rho^0_t + \rho^1 RVOT_{it-1} + \varepsilon_{it},$$
where $\rho_i^0$ is a country-specific effect. The estimated $\rho^1$ equals 0.927.\footnote{We estimate the regression using a Least Square Dummy Variable (LSDV) estimator. For sample sizes of $T \geq 90$ and $N = 20$, the bias is small and the LSDV estimator generally performs better than the Arellano-Bond estimator or the Anderson-Hsiao estimator.} We then calibrate $\rho$ to the estimated $\rho^1$. Assume that one period in our model contains ten years. Hence, $\rho = 0.927^{10} = 0.469$. Moreover, the standard deviation of estimated $\varepsilon_{it}$ from the above regression is 2.67\%, implying a standard deviation of 6.14\% over ten years. Since we use $\varepsilon$ as a proxy for ideological shocks, we target the “competitive political region” to the 95\% confidence interval of $\varepsilon$, which is $[-12.03\%, 12.03\%]$. More precisely, we let $s^1 = -12.03\%$ and $s^M = 12.03\%$. This pins down two other parameters: $z = 0.2233$ and $\beta = 0.6689$. Note that the calibration implies a symmetric competitive political region around $s = 0$.

### 8.6 Proof of Proposition 3

Using (1) to (4), (34) can be written as

\[ W^L = a^o g + \hat{\omega} (h (1 - \tau^y) + a^y g - h^2 + \beta h E [1 - \tau^o + a^o g^1] + \beta (1 - h) a^o E [g^1]) \]

\[ = (a^o + \hat{\omega} a^y) g + \hat{\omega} (h (1 - \tau^y) - h^2 + a^o \beta E [g^1] + \beta h \pi) \]

\[ = \frac{(a^o + \hat{\omega} a^y)}{2} (h_{-1} \tau^o + h \tau^y) + \hat{\omega} \left( h (1 - \tau^y) - h^2 + \frac{a^o \beta h}{2} (h (1 - \pi)) + h \beta \pi \right). \]

(6) is used to replace $g$ and $g^1$ in the third line. We also drop the irrelevant term, $h^2 \tau^y$, as it is fully determined by the next-period young households and government and, thus, independent of $\tau^o$ and $\tau^y$. Similarly, (35) can be written as

\[ W^R = 1 - \tau^o + (a^o + \hat{\omega} a^y) \left( h_{-1} \tau^o + h \tau^y \right) + \hat{\omega} \left( h (1 - \tau^y) - h^2 + \frac{a^o \beta h}{2} (h (1 - \pi)) + h \beta \pi \right). \]

Recall that $a^o + \hat{\omega} a^y < 2$ and $h_{-1} \leq 1$. Hence, maximizing $\tau^o$ gives the binary rule (8). Since $h$ follows (14), $\pi$ satisfies the same functional equation (17) as in the benchmark model. Then, it is immediate that $\tau^y$ solves (37).

Now, we prove the third part of the proposition. Under probabilistic voting, the party candidates maximize

\[ W = (1 - h_{-1} - \hat{s}) (a^o g) + (h_{-1} + \hat{s}) (1 - \tau^o + a^o g) \]

\[ + \hat{\omega} \left( h (1 - \tau^y) + a^y g - h^2 + \beta h E [1 - \tau^o + a^o g^1] + \beta (1 - h) a^o E [g^1] \right) \]

\[ = (h_{-1} + \hat{s}) (1 - \tau^o) + \frac{(a^o + \hat{\omega} a^y)}{2} (h_{-1} \tau^o + h \tau^y) \]

\[ + \hat{\omega} \left( h (1 - \tau^y) - h^2 + \frac{a^o \beta h}{2} (h (1 - \pi)) + h \beta \pi \right). \]
Clearly, maximizing $W$ w.r.t. $\tau^o$ yields

$$\tau^o = \begin{cases} 1 & \text{if } \hat{s} + h_{-1} < \frac{a^o + \hat{\omega}a^y}{2} \\ 0 & \text{otherwise} \end{cases}.$$  

The definition of $\hat{s}$ implies that the above equation is identical to (8). Then, the problem for $\tau^y$ reduces to (37).

### 8.7 Proof of Proposition 4

The political objective functions (38) and (39) can be rewritten as

$$W^L = \frac{(a^o + \hat{\omega}a^y)}{2} (h_{-1} \tau^o + h \tau^y) + \hat{\omega} \beta \frac{a^o}{2} (h (1 - \pi)),$$

$$W^R = 1 - \tau^o + \frac{(a^o + \hat{\omega}a^y)}{2} (h_{-1} \tau^o + h \tau^y)$$

$$+ \hat{\omega} \left(1 - \tau^y + \beta \left( \pi + \frac{a^o}{2} h (1 - \pi) \right) \right).$$

Here we drop the irrelevant term, $h' \tau^y$. The rest of the proof is immediate.
Figure 1: Equilibrium Results

Figure 1: Panel A represents the equilibrium policy rule \( \tau^*(s) \). The probability for the right-wing to be elected, \( \pi(\tau^*(s), s) \), is plotted in Panel B. Panel C corresponds to the equilibrium investment rule \( h(\tau^*(s), s) \). The parameter values are set equal to the benchmark case: \( z = 0.22, \rho = 0.47, \beta = 0.67 \).
Figure 2: Laffer Curves

Figure 2: Panel A to D plot $T(\tau^y, s)$ with respect to $\tau^y$ under different ideological states. The parameter values are set equal to the benchmark case as in Figure 1.
Figure 3: Sensitivity Analysis

Figure 3: Panel A and B plot the estimated $b^2$ with respect to $z$ and $\rho$, respectively. When changing $z$ or $\rho$, we recalibrate $\beta$ accordingly so that the competitive political region remains symmetric at $s = 0$. 
Figure 4: Equilibrium Results with Ideology-Independent Altruism

Panel A

Panel B

Panel C

Figure 4: Solid and dotted lines stand for equilibrium results with ideology-independent altruism and those in the benchmark case, respectively. Panel A represents the equilibrium policy rule $\tau^y(s)$. The probability for the right-wing to be elected, $\pi(\tau^y(s),s)$, is plotted in Panel B. Panel C corresponds to the equilibrium investment rule $h(\tau^y(s),s)$. $\omega = 0.1$ and the other parameter values are held constant as in the benchmark case.
Figure 5: Equilibrium Results with Ideology-dependent Altruism \((h_{-1} = 0.5)\)

Figure 5: Solid and dotted lines stand for equilibrium results with ideology-dependent altruism and those in the benchmark case, respectively. The predetermined human capital \(h_{-1} = 0.5\). Panel A represents the equilibrium policy rule \(\tau^y(s)\). The probability for the right-wing to be elected, \(\pi(\tau^y(s), s)\), is plotted in Panel B. Panel C corresponds to the equilibrium investment rule \(h(\tau^y(s), s)\). \(\omega = 0.1\) and the other parameter values are held constant as in the benchmark case.
Table 1: OLS Estimation of the Determinants of Government Sizes

<table>
<thead>
<tr>
<th>Dep. Variable</th>
<th>CGEXP</th>
<th>CGREV</th>
<th>CGEXP</th>
<th>CGREV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>R</td>
<td>-0.8150</td>
<td>-1.0462*</td>
<td>-1.8006***</td>
<td>-1.8270***</td>
</tr>
<tr>
<td></td>
<td>(-0.84)</td>
<td>(-1.65)</td>
<td>(-2.78)</td>
<td>(-3.10)</td>
</tr>
<tr>
<td>RVOT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Variables</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Obs.</td>
<td>673</td>
<td>663</td>
<td>718</td>
<td>708</td>
</tr>
<tr>
<td>Adj. R²</td>
<td>0.832</td>
<td>0.881</td>
<td>0.812</td>
<td>0.833</td>
</tr>
</tbody>
</table>

Notes: Country dummies and year dummies are included to control for the fixed effects and time effects. Control variables are YPC (the log of real GDP per capita), YGAP (HP filtered YPC), openness, unemployment rate, and the sizes of population over 65 and below 14. We also include debt/GDP ratio as additional control variable when the dependent variable is CGREV (government revenue). Robust t statistics is in brackets. ***, ** and * is significant at 1%, 5% and 10%, respectively.
Table 2: 2SLS Estimation of the Determinants of Government Sizes

<table>
<thead>
<tr>
<th>Dep. Variable</th>
<th>CGEXP</th>
<th></th>
<th>CGREV</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>R</td>
<td>-2.4844*</td>
<td>-2.3420**</td>
<td>-3.8392***</td>
<td>-3.8592***</td>
</tr>
<tr>
<td></td>
<td>(-1.95)</td>
<td>(-2.13)</td>
<td>(-3.36)</td>
<td>(-3.45)</td>
</tr>
<tr>
<td>RVOT</td>
<td>0.1632***</td>
<td>0.0380</td>
<td>0.2156***</td>
<td>0.1761***</td>
</tr>
<tr>
<td></td>
<td>(5.55)</td>
<td>(1.34)</td>
<td>(7.64)</td>
<td>(5.86)</td>
</tr>
<tr>
<td>Control</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Control</td>
<td>Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>657</td>
<td>648</td>
<td>700</td>
<td>677</td>
</tr>
</tbody>
</table>

Notes: Country dummies and year dummies are included to control for the fixed effects and time effects. R and RVOT are identified as endogenous. The corresponding instruments are the same variables, but one-year lagged. Control variables are YPC (the log of real GDP per capita), YGAP (HP filtered YPC), openness, unemployment rate, and the sizes of population over 65 and below 14. We also include debt/GDP ratio as additional control variable when the dependent variable is CGREV (government revenue). t statistics is in brackets. ***, ** and * is significant at 1%, 5% and 10%, respectively.