

INFORMED POLITICIANS AND INSTITUTIONAL STABILITY

A Dissertation  
submitted to the Faculty of the  
Graduate School of Arts and Sciences  
of Georgetown University  
in partial fulfillment of the requirements for the  
degree of  
Doctor of Philosophy  
in Economics

By

Helena P. Aten, M.Sc.

Washington, DC  
June 2, 2010

Copyright 2010 by Helena P. Aten  
All Rights Reserved

# INFORMED POLITICIANS AND INSTITUTIONAL STABILITY

Helena P. Aten, M.Sc.

Thesis Advisors: Luca Anderlini, Ph.D. and Roger Lagunoff, Ph.D.

## ABSTRACT

This dissertation examines two aspects of democratic institutions.

The first chapter studies the welfare implications of politicians who assume either the role of delegates or trustees in a representative democracy. I identify conditions under which the latter is preferable to the former. In this model, voters are uninformed about the value of a policy-relevant state. Two informed politicians compete for votes by committing to state-contingent policy platforms that may or may not reveal information about the underlying state. After the election, the winning politician announces the state and implements the relevant policy.

I find that if voters' policy preferences are not too sensitive to changes in the state, then the two politicians offer divergent policy platforms. In addition, the main result characterizes Perfect Bayesian Equilibria in which the offered platforms are non-revealing menu contracts, and the resulting welfare is higher than in any separating equilibrium. Such is the case when voters are sufficiently valence-driven and direct benefits to politicians are sufficiently important. The result provides a welfare explanation for why voters may defer policy choices to an elected representative, rather than select a politician that reflects their policy preferences based on information revealed in political competition.

The second chapter explores whether there are systematic differences in institutional stability between democracies and non-democracies. It exploits data on

56 countries that have experienced institutional change between 1980-2007. Since the institution variable is correlated with unobservables in the determination of institutional change, a maximum likelihood estimation that does not control for this correlation will yield biased estimates. Assuming that the endogeneity operates solely through country fixed effects, I estimate the likelihood of institutional change using fixed effects probit with bias correction.

I find that the consistently significant factor affecting institutional change is the interaction between democracy and the percentage of democracies in the world. The coefficient is negative and significant, which suggests that being a democracy has a positive externality on the stability of other democracies. Further classification of political institutions into democracy, autocracy, and intermediate ranges yields stronger results confirming this argument. Whether or not being a democratic institution directly affects a country's likelihood of institutional change depends on the assumption about the source of the endogeneity of the institution variable.

To My Grandmother

## TABLE OF CONTENTS

Introduction . . . . .	1
Chapter 1: Competing Informed Principals and Representative Democracy . . . . .	4
1 Introduction . . . . .	4
2 The Model . . . . .	10
3 Characterization . . . . .	15
4 Welfare-Dominant Pooling Equilibria . . . . .	22
4.1 Purely Valence-Driven Voter . . . . .	25
4.2 A Voter Driven by Both Valence and Policy . . . . .	28
4.3 Equilibrium Refinement . . . . .	32
5 Conclusion . . . . .	33
Chapter 2: Are Democracies more Stable than Non-Democracies? Evidence from Cross-Country Panel Data . . . . .	37
1 Introduction . . . . .	37
2 Econometric Estimation . . . . .	43
2.1 Econometric Model . . . . .	44
2.2 Data . . . . .	47
3 Results . . . . .	52
3.1 Fixed Effects Probit . . . . .	52

3.2	Marginal Effects . . . . .	55
4	Robustness and Source of Endogeneity . . . . .	60
4.1	Time Trend . . . . .	60
4.2	Time-variant Endogeneity . . . . .	61
5	Conclusion . . . . .	67
	Appendix . . . . .	70
A	Proof of Lemma 1 . . . . .	70
B	Strategy and Outcome Representation . . . . .	73
C	Aggregate Welfare . . . . .	74
D	The Truth-Telling Constraint . . . . .	75
E	Proof of Proposition 2 . . . . .	78
F	Proof of Lemma 2 . . . . .	86
G	Proof of Lemma 3 . . . . .	88
H	Proof of Proposition 3 . . . . .	90
I	Description of Data . . . . .	93
J	Tables of Results . . . . .	103
	References . . . . .	111

## Introduction

This dissertation is a study of the value and welfare implications of the rules that characterize democratic institutions. On a microeconomic level, I compare the welfare outcomes when politicians assume either the role of delegates or trustees in a representative democracy, and ask whether entrusting legislative decisions to elected representatives could be justified. On a macroeconomic level, I explore the value of democracy from the point of view of stability, and identify the channels through which democracy may lead to an increase or decrease in a country's institutional stability.

In the first chapter of the dissertation, I introduce a theoretical model of political competition between two privately-informed politicians. There are two dimensions of strategic interactions in play in this model. First, through the offer of a policy platform, each politician may be giving out the private information that she possesses; therefore she has to optimally devise her platform taking into account how the voters will react based on the information they have. Second, the two politicians are competing against each other in an election. As each politician decides what platform to offer, she must account for her opponent solving the exact same problem, and only one of them will be elected to office. The need to balance both strategic interactions is crucial for both politicians, and I characterize the set of equilibria in this game.

An important link between the model and the motivating question is that the two roles of politicians in a representative democracy are analogous to the two kinds of equilibria that emerge in the model – when information is and is not revealed through the policy platforms offered. When information is revealed through political



competition, voters will know the state of nature at the time of voting, therefore the politician elected will be the one whose policy platform most closely resembles the preferences of voters. The winning politician's role will only be to implement the policy after the election, and her role will be analogous to that of a delegate. On the other hand, when information is not revealed through political competition, the winning politician withholds the information she has until the policy-making stage. The voters must therefore rely on the elected representative to announce the state ex post, giving the winning politician the authority to use discretion on the choice of policy, while leaving the voters with no ability to influence policy after the election. The role of the elected representative resembles that of a trustee in this kind of equilibrium.

I find that if voters' policy preferences are sufficiently insensitive to changes in the state, then the two politicians offer divergent policy platforms. In addition, the main result characterizes Perfect Bayesian Equilibria in which the offered platforms are non-revealing menu contracts, and the resulting welfare is higher than in any separating equilibrium. Such is the case when voters are sufficiently valence-driven and direct benefits to politicians are sufficiently important. The result provides a welfare explanation for why voters may defer policy choices to an elected representative, rather than select a politician that reflects their policy preferences based on information revealed in political competition.

The second chapter is an empirical model estimating the likelihood of institutional change between democracies and non-democracies. The focus is to separate the channels through which democracy effects institutional change – whether there is something intrinsic about democratic institutions that leads to stability or instability, or democracy is an externality that influences the likelihood of change in

other countries.

The data is based on 56 countries that have experienced institutional change between 1980-2007. Since the institution variable is correlated with unobservables in the determination of institutional change, a maximum likelihood estimation that does not control for this correlation will yield biased estimates. Assuming that the endogeneity operates solely through country fixed effects, I estimate the likelihood of institutional change using fixed effects probit with bias correction.

I find that the consistently significant factor affecting institutional change is the interaction between democracy and the percentage of democracies in the world. The coefficient is negative and significant, which suggests that being a democracy has a positive externality on the stability of other democracies. Further classification of political institutions into democracy, autocracy, and intermediate ranges yields stronger results confirming this argument.

In addition, there is evidence suggesting that the endogeneity of the institution variable is attributable to both country-specific fixed effects and time-variant unobservables. Whether or not being a democratic institution directly affects a country's likelihood of institutional change depends on the assumption about the source of the endogeneity of the institution variable.

# Chapter 1: Competing Informed Principals and Representative Democracy

## 1 Introduction

Most democracies today involve elected representatives legislating on behalf of their constituents. However, what role politicians should play in a representative democracy is still a largely unsettled debate<sup>1</sup>. In order to analyze how political institutions should be designed, it is crucial to understand how the incentives of political actors and the outcomes of the political process vary across different configurations of the institution. Broadly speaking, there are two schools of thought on what the roles of politicians may be (Burke, 1854).

The first is commonly referred to as the *delegate model of representation*. The idea is that politicians should act as the mouthpiece of their constituents and make policy in accordance with the wishes of the electorate. This model is motivated by the view that politicians mainly exist to directly represent voters, who cannot realistically attend to all legislative procedures and decisions.

The second is referred to as the *trustee model of representation*. In this model politicians are given discretion over policy; even though their policy decisions may not always coincide with the view of the electorate, they are entrusted to implement policies that are geared towards the long-run good of a society. This model is motivated by the view that politicians should be more than simply representation for voters in absentia, but that they should be relied on for their competence, judgment, and leadership.

---

<sup>1</sup>The origins of the debate trace back to Burke (1854) and Mill (1882). For more recent work, see Fox and Shotts (2009).

This paper compares the welfare implications when politicians assume either of these roles in the political system. Intuitively, it seems preferable to treat politicians as delegates in the interest of aggregating voter preferences and accountability; we typically think politicians are entrusted with discretion only because it is too costly to extract the information they have or there is too much variation in what the right action should be given the information<sup>2</sup>. Somewhat surprisingly, this paper shows that aggregate welfare can actually be higher when politicians are given legislative discretion, and characterize conditions under which this is true. Provided that these conditions are met, even though politicians possess policy-relevant information, it is in fact welfare enhancing when this information is not revealed to voters, so that the elected representative can exercise judgment in implementing the appropriate policy.

This paper examines a simple but illustrative model of electoral competition between two politicians. There are two states of the world (good times and bad) and two public goods that can be produced (guns and butter). Politicians care first and foremost about being elected, but they also each prefers one of the two public goods. Voters have state-dependent preferences over the public goods and preference predispositions for the politicians that are not policy-related (valence). Politicians know the state of the world; voters do not<sup>3</sup>. However, during the elec-

---

<sup>2</sup>This will be discussed in greater detail when I review the existing literature.

<sup>3</sup>It is sufficient to think that politicians know something about which voters have no information. There are several ways to motivate why politicians are more informed than voters. There is a sizable literature that argues or assumes politicians are more informed because they have a team of experts working to provide them with higher quality information, and they have less of a free-rider problem in acquiring costly information relative to each member of a large electorate (for instance, see Caillaud and Tirole (2002), Maskin and Tirole (2004), Kessler (2005), and Fox and Shotts (2009)). It is also reasonable to think that politicians seeking higher office are already in a political position in which they have access to classified information unavailable to voters – think of John Kerry being in the

toral competition politicians announce policy platforms that may or may not reveal information about the state. After observing the candidates' platforms, voters can update their beliefs and vote accordingly. The winning politician then announces the state and implements the policy corresponding to it.

How is this model linked to the question of delegate versus trustee representation? In the process of competing in an election, the information possessed by the politicians – or more generally, political competitors<sup>4</sup> – may or may not be revealed to voters.

In an equilibrium in which the electoral process induces the politicians to reveal the state, the winning politician's only role is to implement the part of the policy platform that is consistent with the state. In fact, the winning politician's choice of policy in the true state is the one preferred by the pivotal voter. In this sense, the fully revealing equilibrium resembles the delegate model of representation.

In an equilibrium in which neither politician reveals information in the electoral process, the winning politician withholds the information she has until the policy-making stage. The elected representative is entrusted with the authority to use discretion on the choice of policy, thus eliminating voters' ability to influence policy ex post. In this sense, a non-revealing equilibrium corresponds to the trustee model of representation.

I find that if voters' policy preferences are sufficiently insensitive to the policy-relevant state, then policy platforms offered diverge and differ from the voters' ideal point. The sets of separating and pooling equilibria always coexist in this

---

Senate Armed Services Committee when he was running for president in 2004.

<sup>4</sup>While the term "politician" is used throughout, the application could be any general form of "political competition", e.g. two competing experts or informed lobbying groups that care about having their advice adopted and their suggested policies implemented.

game. More importantly, if the voters are sufficiently valence-driven, and direct benefits to politicians are sufficiently important, then welfare is higher when the state is *not* revealed to voters in the process of political competition, and the elected representative is given legislative discretion. Trustee representation may then be interpreted as a welfare improvement over delegate representation, in which the chosen politician reflects directly the policy preference of her constituents.

The intuition of the main result is as follows. In most political systems politicians compete against one another and are bound by institutional constraints. These are known in the principal-agent literature as individual rationality and incentive compatibility constraints, or, in the context of this application, constitutional guarantees or checks and balances<sup>5</sup>. If the state is revealed in the electoral process, these constraints have to be satisfied appropriately given each state. On the other hand, if the relevant information is not revealed to voters through political competition, politicians have some flexibility in choosing how to balance the various constraints to implement the policies that they deem the most appropriate, even as they compete with each other. The present paper identifies conditions under which it is welfare-enhancing to allow politicians the added flexibility, so that they may act as trustees rather than delegates.

The model in this paper is built on the standard Hotelling-Downsian (Hotelling, 1929; Downs, 1957) multi-dimensional spatial competition model, and includes the probabilistic voting element of Lindbeck and Weibull (1987) to represent voters' preferences along the non-policy (valence) dimension. The main departure from these models is in the payoff-relevant information possessed by politicians, and thus

---

<sup>5</sup>It is important to note, however, that these constraints arise as part of the equilibrium, rather than as assumptions of the model.

the need for them to decide whether and how to use the information in an electoral competition.

It is standard in the political economy literature to model the incentives of politicians with private information as an agency problem<sup>6</sup>, where the politician is the agent. In these models, the sets of separating and pooling equilibria generally exist for different parameter configurations, and pooling equilibrium arises because it is too costly to separate the different types of the agent or fix a narrow band of actions permissible for the agent. Although not directly comparable (because the parameters differ across the comparison), there is a sense in which the principal and at least one type of the agent are “worse off” when the relevant information is not fully revealed. A novelty of this paper is that politicians assume the role of *informed principals* while the voters represent “the agent.” This is more in line with the spatial political competition models, and highlights the strategic decisions made by competing politicians with payoff-relevant information<sup>7</sup>. Schultz (2008) has a model similar to that of this paper, in which politicians are informed and that information may be revealed in the process of political competition. The focus of Schultz (2008), however, is to evaluate whether term lengths should be long or short in a representative democracy. The current paper extends the informed principal framework in the existing literature by having two informed principals compete to contract with a single agent<sup>8</sup>. In addition to applying this to a model

---

<sup>6</sup>For instance, see Mirrlees (1976), Holmstrom (1979), Grossman and Hart (1983), Athey et al. (2005), and Vlaicu (2008).

<sup>7</sup>While voters may also have private information, this paper abstracts from it to highlight the results driven by the principals being the ones informed. Since voters’ only action is to vote, there is no screening or moral hazard component that utilizes voters’ private information.

<sup>8</sup>Besides Schultz (2008), the only other paper with more than one informed principal to my knowledge is Martimort and Moreira (2004). The setting and focus of Martimort and Moreira (2004) is quite different from the current paper – there are two principals with

of electoral competition, the extension that this paper provides may also be useful for understanding other applications with multiple competing principals<sup>9</sup>.

The signaling value of contracts offered by an informed principal was first pointed out in the seminal works of Myerson (1983) and Maskin and Tirole (1990, 1992). In this paper we exploit the well-known Inscrutability Principle of Myerson (1983), which says that the informed principal can without loss of generality offer pooling menu contracts, and reveal the information she has only *after* the contract is signed.

In their two related papers, Maskin and Tirole (1990, 1992) provide equilibrium characterization for the informed principal problem. Equilibrium characterization differs depending on whether we are in the case of private or common values<sup>10</sup>. Maskin and Tirole (1990) show that in the case of private values, there exist Pareto superior pooling menu contracts relative to the entire set of separating equilibria. Because the relevant information does not pose as a direct conflict between the different types of the principal, a pooling equilibrium simultaneously relaxes the individual rationality (IR) and incentive compatibility (IC) constraints, since these constraints need only to hold in expectation rather than state by state. This allows all types of the principal to be better off compared to any fully revealing equilibrium.

Although the premise of the current paper is different from Maskin and Tirole (1990) in a few crucial ways (for example, there are two competing principals and

---

private information about their preferences in a common agency game. The key insight is that it introduces additional incentive compatibility constraints (requiring the principals' truthfulness) in the common agency environment.

<sup>9</sup>For example, there is usually more than one contractor competing for a contract to carry out construction work; the agent can choose to sign a franchising agreement with either McDonald's or Burger King; or when purchasing a car, the agent has multiple brands to choose from, each with uncertain quality and purchasing plans.

<sup>10</sup>We are in the case of private values if, holding the contract offered by the principal constant, the information is not an argument in the agent's utility function. Common values refers to the case in which the agent's utility function is still a function of the principal's information even after fixing the contract offered.



risk neutral preferences for all parties), the intuition for the existence of welfare dominant pooling equilibria (over the set of separating equilibria) resonates with Proposition 1 of Maskin and Tirole (1990). The key in the present result is that each politician must maintain a balance between the gains to simultaneously relaxing the IR and IC constraints, and being competitive in the policy dimension against her opponent. Therefore, the gains that the set of pooling equilibria has over the separating equilibria only exist when competition is not too intense along the policy dimension.

The rest of the paper is organized as follows: Section 2.1 describes the model; section 3 provides basic results to consider a reduced strategy space; section 4 states and discusses details of the main welfare result; and finally section 5 concludes. All proofs in this paper, unless otherwise noted, can be found in the appendix.

## 2 The Model

Consider a spatial competition model in which two informed politicians – the principals – compete for votes. Voters care about the provision of two public goods and have preference predispositions for each political candidate’s valence. Because the details of the distribution of voters’ preferences do not play an important role, I simplify the analysis and suppose that there exists a pivotal voter – the agent – whose preferences ultimately determine who is elected<sup>11</sup>.

The pivotal voter’s policy preference depends on the value of the state  $\theta \in$

---

<sup>11</sup>See Rothstein (1990) for a general result on the existence of a median voter in models with multidimensional policy space, or Caplin and Nalebuff (1988) for the use of supermajority to resolve the problem of electoral cycles. In the most simple form, one can think of a large but finite number of voters having the same policy preferences described below, but different preference predispositions for each politician.

$\{H, L\}$ , which is observed only by the politicians (indexed  $i = 1, 2$ ). Each politician is primarily office-motivated, but if elected she cares about the resulting policy. A policy platform is a state-contingent policy, where each policy is comprised of four elements:  $(b, g, p, t)$ . There are two public goods –  $b$  and  $g$ ; think of them as “butter” and “guns”. The variable  $p$  is the direct transfer that the politician receives or gives. Finally,  $t$  is the total tax revenue from voters to finance both the public goods and direct transfer. Throughout the paper, the terms “policy platform” and “contract” are used interchangeably.

The timing of the game is as follows:

1. Nature draws the state of the world (observed only by the politicians) and the pivotal voter’s preference predisposition for the politicians’ valences (observed only by the voter).
2. The two politicians simultaneously announce their policy platforms.
3. Upon observing the platforms announced, the pivotal voter updates his beliefs about the state and votes for at most one politician.
4. If a politician is elected, the winning politician will announce the state, and then carry out the part of her policy platform that corresponds to the announced state<sup>12</sup>.

Next we describe the preferences of all involved in the game.

---

<sup>12</sup>We assume that the policy platforms are binding for the politicians. This means that while the winning politician can announce the false state, she is still bound to implement a policy that is part of her platform.

### Politicians' Preferences (Principals)

$$\begin{aligned} V_1 &= b + \lambda p \\ V_2 &= g + \lambda p \end{aligned} \quad ; \quad \lambda > 1$$

The politicians' preferences are state-independent and commonly known.  $V_i$  is the utility of politician  $i$  if she is elected and the policy  $(b, g, p, t)$  is implemented, otherwise her payoff is normalized to zero. We assume  $\lambda > 1$ <sup>13</sup>, denoting the politician's relative preference of direct transfer over the public good. Note that each politician prefers a different public good (politician 1 likes public good  $b$ , while politician 2 likes  $g$ ), but that preference does not vary by state.

### Pivotal Voter's Preferences (Agent)

$$S_i^\theta = U^\theta + c_i \quad ; \quad \text{where} \quad U^\theta = \begin{cases} \gamma [\eta b + (1 - \eta) g] - t & \text{if } \theta = H \\ \gamma [\eta g + (1 - \eta) b] - t & \text{if } \theta = L \end{cases} \quad ; \quad \text{and}$$

- $0 \leq t \leq 1$  ;
- $\eta \in [\frac{1}{2}, 1]$  ; and
- $c_1 \equiv 0$  and  $c_2 \equiv \psi c$ , where  $\psi \in \mathbb{R}_+$ ,  $c \sim F(c)$ , and  $F(\cdot)$  is a distribution with compact support, density  $f(\cdot)$ , and expected value 0.

The utility of the pivotal voter if politician  $i$  is elected and the state is  $\theta$  is given by  $S_i^\theta$ . There are two components in  $S_i^\theta$ : a policy-specific component  $U^\theta$ ,<sup>14</sup> and a

<sup>13</sup> $\lambda > 1$  is needed since we allow  $p$  to be positive or negative (See discussion on page 14). If  $\lambda \leq 1$  and  $p \in \mathbb{R}$ , the politician can offer an unboundedly high level of the public good that she prefers while still satisfying her individual rationality. In reality, there are often natural bounds for the public goods owing to resource or other constraints.

<sup>14</sup>From section 3 onward  $U_i^\theta$  is used to denote the offer of politician  $i$  to the pivotal voter if the state is  $\theta$ .

valence component  $c_i$ . The voter’s preference for the public goods is parameterized by  $\eta$ . Regardless of the value of  $\eta$ , the pivotal voter (at least weakly) prefers  $b$  in the high state, and  $g$  in the low state. He dislikes taxes, but in relative terms he always prefers to pay taxes in order to acquire his preferred public good (see Feasibility Constraint below). The valence component derived from choosing politician  $i$  is given by  $c_i$ ; it is sometimes referred to in the literature as the voter’s “ideology”.<sup>15</sup> The pivotal voter also has a normalized reservation utility of zero, reflecting the basic rights that are guaranteed to voters in the political system.

The variables that make the informed principal framework vary in two dimensions are  $\eta$  and  $\psi$ . Along the dimension of public good specificity,  $\eta$  represents the intensity of the voter’s relative public good preference across the two states. If  $\eta = \frac{1}{2}$ , the voter likes both public goods equally regardless of the state. If  $\eta = 1$ , the voter *only* wants  $b$  (respectively  $g$ ) when  $\theta = H$  (respectively  $\theta = L$ ). Along the dimension of political competition,  $c_i$  is the valence component in the pivotal voter’s preference, as used in a probabilistic voting model<sup>16</sup>. We normalize  $c_1$  to zero, and  $c_2 \equiv c$ , where  $c$  is a random variable representing the voter’s preference for politician 2’s valence relative to politician 1. The parameter  $\psi$  is a non-negative number denoting the relative importance of the valence and policy components to the voter. If  $\psi = 0$ , then the pivotal voter cares only about policy; a large  $\psi$  indicates a higher level of importance placed on valence compared to policy.

As mentioned above, our goal is to use a simple but illustrative model to provide

---

<sup>15</sup>This component broadly captures the non-policy-related preference that the voter has for each politician (hair color, height, personality). It can also be interpreted as capturing the overall preference that the voter has for each politician outside of the policy dimensions analyzed in the model.

<sup>16</sup>The purpose of this valence component is the standard one: to smooth out discontinuous jumps in the probability of winning when one politician offers a contract just infinitesimally better than that of her opponent.

a possible welfare explanation for trustee representation. The linearity of preferences serves two purposes. First, it greatly simplifies the characterization of the equilibrium set (see Lemma 1 below). More importantly, this strengthens our results because intuitively welfare gains from pooling across states may be higher under any form of risk aversion, given voters' desire to smooth their payoffs across the different states.

Finally, an economy-wide feasibility constraint is given by:

### Feasibility

$$\begin{array}{lcl} b + g + p \leq t & \text{if } \theta = H & \\ \alpha(b + g) + p \leq t & \text{if } \theta = L & \end{array} ; \quad \text{where} \quad \begin{array}{l} \gamma\eta > \alpha > 1; \\ b \geq 0 ; g \geq 0 \end{array}$$

The feasibility constraint describes the technology with which the politicians convert taxes into public goods and direct transfer. It is identical across politicians, but differs by state. The assumption is that public goods are more costly to produce in the low state. However, since  $\gamma\eta > \alpha$ , the voter's preference is such that public good production will not be shut down even in the low state.

Note the parameter restrictions imposed on the policies  $(b, g, p, t)$ . Clearly the amount of public good ( $b$  and  $g$ ) offered must be non-negative. There is also a natural limit as to how much voters can be taxed; recall it is set between zero and one. The "net transfer"  $p$  that each politician receives can be positive or negative<sup>17</sup>: implicit or indirect transfers are often used in politics to elicit votes from a subset

---

<sup>17</sup>From a technical standpoint, the advantage that each principal has in the model only exists if negative transfers are allowed. If we impose the restriction of  $p \geq 0$ , then the principals are equally competitive in both states, even when one principal's preference aligns with that of the agent's. By allowing possibly  $p < 0$ , each principal can use the advantage that she has in a particular state to increase her probability of winning by offering  $p < 0$ .

of the electorate, which can be either beneficial or detrimental for the politician. A negative  $p$  can be interpreted as the politician taking an otherwise disadvantageous move from her point of view in order to provide higher public good levels.

A contract offered by each politician describes what  $(b, g, p, t)$  will be implemented in each of the two states. Following Myerson (1983), we consider menu contracts without loss of generality<sup>18</sup>, meaning that each contract lays out what policies will be implemented in *both* states, with the relevant part of the menu pointed out and implemented if and when a contract is accepted. Of course, in a separating equilibrium, the pivotal voter will know the true state with certainty on equilibrium path. In terms of notation,  $b_i^{\theta\hat{\theta}}$  denotes the level of  $b$  that type  $\theta$  of principal  $i$  (henceforth  $Pi\theta$ ) promises to implement in state  $\hat{\theta}$ . Other policy variables follow this convention as well. A policy platform for  $Pi\theta$  is therefore  $(c_i^{\theta\hat{\theta}})_{\hat{\theta}\in\{H,L\}} \equiv (b_i^{\theta\hat{\theta}}, g_i^{\theta\hat{\theta}}, p_i^{\theta\hat{\theta}}, t_i^{\theta\hat{\theta}})_{\hat{\theta}\in\{H,L\}}$ .

### 3 Characterization

An equilibrium in this model refers to a pure strategy, weakly undominated Perfect Bayesian Equilibrium. A policy platform for each  $Pi\theta$  is an eight-dimensional object.<sup>19</sup> Fortunately, given the linear structure of the model, Lemma 1 below demonstrates that we can restrict the strategy space to only two dimensions without loss of generality. The difficulty is in ruling out all other possible policies for *any* beliefs that the pivotal voter may have. The result combines the politicians'

<sup>18</sup>In fact, the Inscrutability Principle that Myerson (1983) establishes is stronger than what is used here: he proves that *pooling* menu contracts are without loss of generality.

<sup>19</sup>Formally, the strategy of each principal includes the platforms for both  $PiH$  and  $PiL$ . However, in describing offers and deviations, it is often easier to consider a platform by  $Pi\theta$  – politician  $i$  in one of the states.

incentive constraints and the structure of policy platforms, so we can argue that the policy variables chosen *within each part*<sup>20</sup> of the platform must be as described in the lemma, regardless of any on- or off-equilibrium-path beliefs that the voter may have.

**Lemma 1.** *In any equilibrium (on or off path), it must be the case that  $t_i^{\theta\hat{\theta}} = 1 \forall \theta, \hat{\theta}, i$ . Furthermore,  $\exists \bar{\eta}_i, i = \{1, 2\}$  such that*

$$\begin{cases} g_1^{\theta\hat{\theta}} = 0 \\ b_2^{\theta\hat{\theta}} = 0 \end{cases} \quad \forall \theta, \hat{\theta} \quad \text{if} \quad \begin{cases} \eta \leq \bar{\eta}_1 \\ \eta \leq \bar{\eta}_2 \end{cases}$$

$$\begin{cases} g_i^{\theta H} = 0 \\ b_i^{\theta L} = 0 \end{cases} \quad \forall \theta, i \quad \text{if} \quad \eta > \bar{\eta}_i$$

It is important to note that Lemma 1 does not just apply to policies on equilibrium path; deviation policies must also have the features described.

Lemma 1 points to a few intuitive features of the policies considered. First, the “bang-bang” result of either  $g = 0$  or  $b = 0$  in all policies comes from the linear structure of the model: both public goods exhibit constant but different returns to the principal.

A politician is said to be “preference-aligned” in a given state if the public good that gives the politician positive utility coincides with that which gives the voter higher utility in that state; otherwise the politician is “preference-misaligned.” For instance, the pivotal voter prefers butter in  $\theta = H$ , so does politician 1; therefore, politician 1 is “preference-aligned” in  $\theta = H$ .

---

<sup>20</sup>There are two “parts” in each menu contract. One refers to what will be implemented in  $\theta = H$  (i.e.  $C_i^{\theta H}$ ), and the other refers to what will be implemented in  $\theta = L$  (i.e.  $C_i^{\theta L}$ ).

A preference-aligned politician ( $P1H$  and  $P2L$ ) will always offer the public good of her choice, exactly because there is no conflict between her preference and the pivotal voter's preference for the public goods. A preference-misaligned politician ( $P1L$  and  $P2H$ ), on the other hand, offers either the good of her choice or that of the pivotal voter, depending on the relative intensity of the voter's preference for the two public goods, parameterized by  $\eta$ . When  $\eta$  is small (close to  $1/2$ ), i.e. when the voter's public good preference is not too state-specific, the preference-misaligned politician could offer the public good that she prefers. However, when the pivotal voter has a marked preference for one of the public goods given the state ( $\eta$  large), then the preference-misaligned politician will offer the public good that the voter prefers<sup>21</sup>.

The case of  $\eta$  small is one where the politicians' policy platforms diverge (with one offer deviating from the pivotal voter's most preferred point). Intuitively, each politician offers public goods to achieve one or two goals: (1) to increase the pivotal voter's utility, which in turn increases her probability of being elected; and (2) to increase her own utility conditional on being elected. Ideally the politicians would like to offer a public good that attains both of these goals; this is possible for all politicians when  $\eta$  is relatively small. However, when  $\eta$  is large, having to fulfill both objectives for the preference-misaligned politicians also means that the former objective will be fulfilled rather ineffectively. The misaligned type will find herself better off by offering the public good that the voter prefers, and the policy convergence obtains as in most models of political competition.

---

<sup>21</sup>If  $\alpha = 1$ , i.e. there is no difference in feasibility across the two states, then the intermediate case of  $\eta$  (in which one preference-misaligned principal offers the public good she prefers, while the other preference-misaligned principal offers what the pivotal voter prefers) does not exist – both politicians either offer the public good that they want, or that which the pivotal voter wants.



The bounds of  $\eta$  for each politician ( $\bar{\eta}_i$ ), given in the appendix, are decreasing in  $\lambda$  and  $\alpha$ . This means that divergent policies are “less likely” (over the space of  $\eta$ ) to be offered the more important direct transfer is to the politicians, and the more expensive public goods are in the low state. Intuitively, policies offered are less likely to be divergent because the use of public goods to attain goal (2) above is less effective the higher  $\lambda$  and  $\eta$  are. There will be a greater range of  $\eta$  in which the politicians find it more cost-effective to offer the public good of the pivotal voter’s choice, and convert any residual from taxes (if available) into  $p$ .

Lemma 1 simplifies the structure of the game tremendously, in that we have two out of four of the elements in  $\mathcal{C}_i^{\theta\hat{\theta}}$  pinned down, and the final two are restricted by the feasibility constraint. This reduces our problem from a 32-dimensional strategy profile to just eight. Lemma 1 and the feasibility constraint can be combined to describe the relationship between any policy that the politician may offer and her utility level should she be elected (details in Appendix B). Provided that the winning politician will truthfully reveal the state ex post, we can also compute the voter’s utility from policy for a given platform promised by the winning candidate, and hence the relationship between the payoffs of the winning politician and the pivotal voter.<sup>22</sup> Hence we can without loss of generality think of the game as the politicians competing in the  $\left(U_i^{\theta\hat{\theta}}\right)_{\hat{\theta}\in\{H,L\}}$  dimension, where  $U_i^{\theta\hat{\theta}}$  is the utility that  $Pi\theta$  promises the pivotal voter if elected and state  $\hat{\theta}$  is implemented.

Before proceeding to the main welfare result, a few remarks are in order.

First, in the interest of brevity, two components of the equilibrium description

---

<sup>22</sup>For the purpose of what will follow in this paper, whenever the payoff of the voter is mentioned, we refer to the voter’s payoff *from policy* only. This is because while the valence component affects the probability of winning and hence indirectly affects the strategies of the politicians, it does not directly enter into the politicians’ strategies.

are omitted throughout the paper:

- (1) The pivotal voter's strategy – the voter's action in the game is trivial; his beliefs, on the other hand, are extremely important. Though omitted, it is understood that the pivotal voter's choice of politician is sequentially rational given his beliefs: he elects the candidate that gives him higher expected utility given his updated beliefs, provided that doing so also yields him a utility level that is above his reservation.
- (2) On-path beliefs – using Bayes' Rule, the pivotal voter's beliefs are degenerate and correct in a separating equilibrium; in a pooling equilibrium, the posterior belief will be the same as the prior.

Second, given Lemma 1, the strategy of  $Pi\theta$  can be summarized by  $U_i^\theta \equiv (U_i^{\theta H}, U_i^{\theta L})$ . For each state  $\theta$ , fix an equilibrium strategy profile  $\mathcal{C}^\theta \equiv (\mathcal{C}_1^\theta, \mathcal{C}_2^\theta)$ . Given  $\mathcal{C}^\theta$ , let  $V_i^{\theta\hat{\theta}}$  denote the utility of  $Pi\theta$  if she announces  $\hat{\theta}$  after she is elected, and  $U_i^{\theta\hat{\theta}}$  is the corresponding utility of the pivotal voter derived from policy.

Finally, equilibrium contracts will satisfy the following standard properties:

1. For all  $i$  and  $\theta$ ,  $\exists \hat{\theta}$  such that  $V_i^{\theta\hat{\theta}} \geq 0$  **(IR-P)**

IR-P is the individual rationality of the politicians. The individual rationality constraint says that no politician can offer a policy platform that gives her negative payoffs for *both* parts of the platform if she was elected. Given the politician's reservation utility, any such policy platform will be dominated by another in which the politician breaks even for at least one part of the platform.

2. For all  $i$  and  $\theta$ ,  $\pi U_i^{\theta H} + (1 - \pi)U_i^{\theta L} \geq 0$  **(IR-A)**

Suppose the pivotal voter’s posterior belief of  $pr(\theta = H)$  is given by  $\pi$ . Individual rationality must also hold for the voter, but since he does not know the state of the world, he calculates his expected utility given his beliefs (hence IR-A is a function of  $\pi$ ). In our application, IR-A can be viewed as constitutional guarantees for the voters’ basic rights. Given the voter’s reservation utility, any policy platforms in which IR-A is not satisfied is weakly dominated by another in which IR-A is weakly satisfied, since the former is chosen with probability zero and hence the politician obtains zero utility for sure.

**3.** For all  $i$ ,  $\theta$ , and  $\hat{\theta}$ ,  $V_i^{\theta\theta} \geq V_i^{\theta\hat{\theta}}$  (TC)

TC is the truth-telling constraint, which says that in each state any elected politician will implement the part of the policy platform that corresponds to the true state. TC often results in informational rent to be given to the party with private information. Interestingly, in our model *the voter* may also benefit from the politician’s binding TC to get above-reservation utility. Such is the case when the winning politician gets most of the surplus. Details of TC are given in Appendix D. Using similar arguments as those used for the Revelation Principle, we can without loss of generality consider only contracts that satisfy TC (see Appendix A for formal proof).

On the subject of the Revelation Principle, it must be noted that more complex contract structures such as escalation clauses<sup>23</sup> are simply ruled out in this paper. The main reason is that policy platforms of this nature are hardly observed and do not seem applicable to political competition. Escalation clauses require levels

---

<sup>23</sup>See Epstein and Peters (1999), Dasgupta and Maskin (2000), Peters (2001), and Martimort and Stole (2002) for examples of more complex contracts and the problem of the Revelation Principle with competing principals.

of commitment and detail that are unnatural in this context. Moreover, results of the current model will be qualitatively the same if the model is extended as follows: before the votes are cast, politicians are allowed a known, finite number of alternating sequential (counter) offers, and each politician's last offer supersedes all her earlier ones<sup>24</sup>.

Henceforth we will denote the set of policy platforms that satisfies IR-P, IR-A, and TC by  $\mathcal{U}_i^\theta$  (with its dependence on the pivotal voter's beliefs  $\pi$  suppressed). We can graphically represent the strategy space in which the two politicians compete. Recall that the strategy of  $Pi\theta$  can be summarized by the voter's utility from policy conditional on  $Pi\theta$  being elected (i.e.  $U_i^\theta$ ). Since state-contingent policy platforms are used,  $Pi\theta$ 's strategy is given by the pair  $(U_i^{\theta H}, U_i^{\theta L})$ . A generic policy platform will therefore span a two-dimensional space. Qualitative representations of each politician's strategy space are given in Figure 1.

IR-P puts maxima on the voter's utility for a given platform, with the bounds given by *either* of the dotted lines. Recall from our earlier discussion that IR-P rules out policy platforms with which the politician gets negative utility for both parts of the platform. Graphically, IR-P rules out the northeast quadrant of the dotted lines. IR-A is given by the negatively-sloped line. The voter's individual rationality depends on his beliefs, therefore this line can take on any slope between 0 (if he believes  $\pi = 0$ ) and  $-\infty$  (if he believes  $\pi = 1$ ). A feasible platform must lie northeast of IR-A. Finally, TC is given by the space between the two positively-sloped lines. It limits how different  $U_i^{\theta H}$  and  $U_i^{\theta L}$  can be to ensure incentive compatibility. If  $|U_i^{\theta H} - U_i^{\theta L}|$  is too large,  $Pi\theta$  will have an incentive to implement the part of the

---

<sup>24</sup>The results also do not depend on which principal being the first or last to announce her platform.

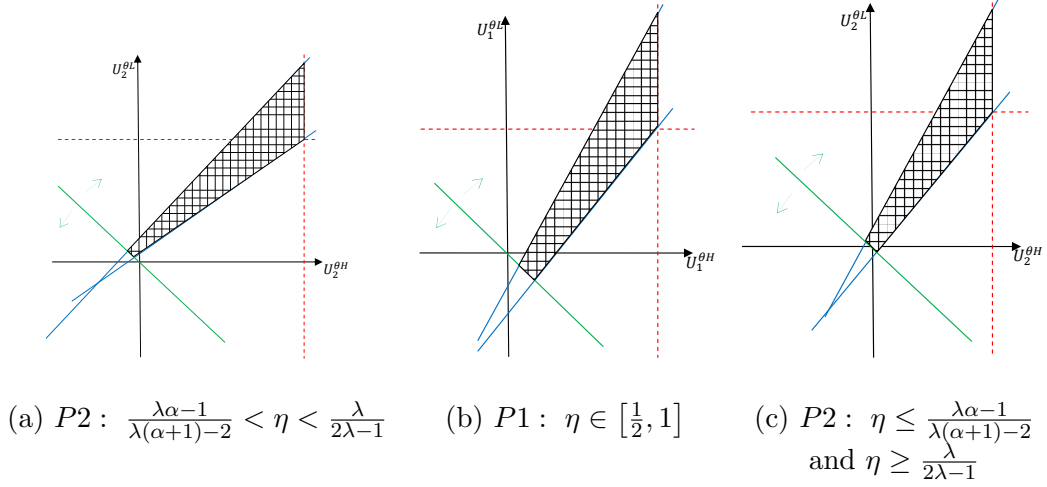


Figure 1: **[Space of Menu Contracts]**. Dotted lines: IR-P. Negatively-sloped line: IR-A. Positively-sloped lines: TC.

platform that yields her the higher utility regardless of the true state.

## 4 Welfare-Dominant Pooling Equilibria

Recall that the main goal of the paper is to provide a welfare explanation for politicians to play the role of trustees in a representative democracy. Let  $\mu \in [0, 1]$  and  $(1 - \mu)$  be the welfare weights society places on the winning politician and the voters respectively<sup>25</sup>. Fixing the equilibrium strategy profile as before and given a welfare weight  $\mu$ , let the ex post weighted aggregate welfare in state  $\theta \in \{H, L\}$  in which politician  $i$  wins as  $\mathcal{W}_i^\theta(\mu) = \mu V_i^{\theta\theta} + (1 - \mu)U_i^{\theta\theta}$ .<sup>26</sup> The welfare function  $\mathcal{W}_i^\theta(\mu)$  can

<sup>25</sup>We do not put explicit weight on the losing politician's welfare, since her payoff is normalized to zero if she loses.

<sup>26</sup>Here the pivotal voter's utility includes only the policy component; since the main goal is to compare across the sets of separating and pooling equilibria, we would fix a realization of  $c$  for both, and therefore the valence component would not affect the comparison.

also be derived from Lemma 1 and the feasibility constraint; details are in Appendix C.  $\mathcal{W}^{S\theta}$  and  $\mathcal{W}^{P\theta}$  (with dependence on  $\mu$  suppressed for brevity) refer to welfare *ranges* for the sets of separating and pooling equilibria, respectively<sup>27</sup>.

We will proceed to show the existence of and identify conditions for which pooling equilibria result in higher welfare relative to the entire set of separating equilibria. Provided that these conditions are met, entrusting politicians with the role of trustees yield higher welfare, and hence may emerge as the preferred institutional arrangement through time (Arrow, 1963).

**Proposition 1 (Welfare-Dominant Pooling Equilibria).** *For all  $\mu \in (0, 1]$ , there exist  $(F(\cdot), \psi, \gamma, \lambda, \alpha, \eta)$  such that*

$$\inf \mathcal{W}^{P\theta} \geq \sup \mathcal{W}^{S\theta} \quad \forall \theta \in \{H, L\}$$

Proposition 1 says provided that society places *some* weight on the winning politician's payoff ( $\mu$  being bounded away from 0), there exist parameter configurations such that being in an equilibrium in which information is *not* revealed prior to voting yields higher welfare than any equilibrium in which information is revealed. The result may be interpreted as a welfare justification for trustee representation: if welfare is higher in a system in which voters must select a representative before the relevant information is known, then trustee representation may have arisen as a result of these welfare gains, so that the elected official is given discretion in her choice of policy.

Intuitively, we can think of politicians as being bound by institutional constraints

---

<sup>27</sup>There is no subscript denoting the principal because fixing a set of parameters, we can calculate the set of equilibria corresponding to each principal winning, and the associated aggregate welfare.

(in equilibrium) as they compete with each other for office. For some of these equilibria, information is revealed to voters through the political process, and hence politicians are restricted in the sense that these constraints have to be satisfied in every state. The policy platform offered by the winning politician is the one that is preferred by the pivotal voter *given each state*, and therefore the role of politicians most closely resembles that of delegates. There are other equilibria in which information is not revealed to voters in the process of electoral competition, and politicians assume the role that resembles that of trustees. In these equilibria, politicians have the added flexibility of balancing these constraints so that they hold in expectation but not necessarily state-by-state, and the winning politician is entrusted to carry out the policy that is appropriate given the state. Having the ability to simultaneously relax the otherwise binding constraints is what allows us to achieve higher aggregate welfare.

To understand how Proposition 1 is obtained, let us first think along the dimension of policy competitiveness between the two politicians. By “policy competitiveness” we mean: how effective is an increase in  $U_i^{\theta\hat{\theta}}$  towards raising  $Pi\theta$ 's probability of winning?<sup>28</sup> In this model, the level of policy competition is given by the importance of the valence component in the voter's preference, parameterized by  $\psi$ . If policy competition between the politicians is intense, most of the surplus should be given to voters in the process; IR-P is much more likely going to bind than IR-A. The opposite is true if there is little policy competition between the politicians.

We will first consider the case in which the voter is purely valence-driven and explain how it is related to the model parameters, then prove upper hemicontinuity

---

<sup>28</sup>Of course, for a higher  $U_i^{\theta\hat{\theta}}$  to have any effect on  $Pi\theta$ 's probability of winning, we are considering  $\hat{\theta}$  where the pivotal voter's posterior belief is  $pr(\hat{\theta}) > 0$ .

of the equilibrium correspondence for the more general case.

#### 4.1 Purely Valence-Driven Voter

Suppose the distribution of valence is such that regardless of the platforms offered<sup>29</sup>, the pivotal voter will elect politician 1 (politician 2 resp.) if the  $c$  drawn lies in the negative (positive resp.) range of its support. This will be referred to as the case of purely valence-driven voter. The requirement of purely valence-driven voter is different from the specifications of  $F(\cdot)$  laid out in section 2.1. Continuity of the distribution of valence is not needed in this part; what we need is the support of the distribution to not contain  $U_i^{\theta\hat{\theta}}$  ranges in which the politicians can still compete in the policy dimension; for instance, we can exclude the range of  $U_i^{\theta\hat{\theta}}$  in which IR-P does not bind for all principals in every state<sup>30</sup>. We will analyze this case and show how it relates to the parameters in the general model.

The case of purely valence-driven voter is similar to a single-informed-principal model in MT90: when the pivotal voter's preference is overwhelmingly dominated by his preference for the politicians' valences, there is no competition along the policy dimension for either politician. With no real strategic interaction between the two principals, each principal maximizes her own utility subject to the agent's individual rationality, so that the contract will be accepted if the realized value of  $c$  is favorable to the principal. From Maskin and Tirole (1990) (henceforth MT90), we know that

<sup>29</sup>Of course, the platforms must still satisfy IR-P, IR-A, and TC, as argued in the previous section.

<sup>30</sup>A sufficient condition is  $F(\cdot)$  such that

$$\psi|c| > \max_{(i,\theta,\hat{\theta})} \left\{ U_i^{\theta\hat{\theta}} : (U_i^{\theta H}, U_i^{\theta L}) \in \mathcal{U}_i^\theta \right\} \text{ for all } c \in \text{supp } F(c)$$



at least two opposing binding constraints (with respect to the states) are needed in order for there to be gains in the set of pooling equilibria over its separating counterpart. Although there is not an incentive compatibility constraint for the agent in the present paper, the fact that the winning principal has to announce the state after a contract is signed gives rise to the truth-telling constraint. This structure allows us to obtain higher surplus in the set of pooling equilibria relative to the set of separating equilibria, and to do so with only risk neutral preferences. The proposition below states the result.

**Proposition 2 (Welfare with Purely Valence-Driven Voter).** *In the case of purely valence-driven voter,  $\forall \mu \in (0, 1]$  and  $(\gamma, \alpha, \eta)$ , there exists  $\bar{\lambda}(\mu, \gamma, \alpha, \eta) \in (1, \infty)$  such that for all  $\lambda \geq \bar{\lambda}$*

$$\inf \mathcal{W}^{P\theta} \geq \sup \mathcal{W}^{S\theta} \quad \forall \theta \in \{H, L\}$$

Though the proof is relegated to Appendix E, it is worthwhile to elaborate on the results of Proposition 2. There are essentially two ways to interpret the details of this proposition. The first is to follow MT90 and look only at the principals' payoffs: after all, politicians are the ones in power who shape the political institution. From the politicians' standpoint, the set of pooling equilibria is preferred over any separating equilibrium. Therefore, if politicians are the ones selecting the political institution in place ( $\mu = 1$ ), then trustee representation may be adopted on welfare grounds (from the point of view of the mechanism designer), so that payoff-relevant information is not revealed to voters even when politicians compete.

The principal-centric interpretation is admittedly quite restrictive, since voters could exert their power by voting politicians out of office (if there are repeated

interactions). A second approach is to consider more generally any weighted sum of the winning politician’s and the pivotal voter’s utilities. This follows the basic idea of Arrow (1963), which proposes that welfare dominant systems may emerge and persist because economic agents seek to benefit from the gains in these systems<sup>31</sup>. Taking this approach, we ask: under what conditions may politicians assume the roles of trustees because of welfare considerations? We find that in the case of pure policy, the requirement is that the welfare importance of the politicians’ direct transfers ( $\mu\lambda$ ) cannot be too small. Intuitively, this is because the opposing interest between politicians and voters is underscored in the variable  $p$ . For the gains that the politicians obtain in the set of pooling equilibria to translate into overall welfare gains, we need  $p$  to have a minimum level of significance in the welfare criterion.

It is important to note that (i) having a separating equilibrium in which TC binds for  $C_i^{\theta\theta}$  for at least one  $\theta$ , and (ii) having a range of TC contracts<sup>32</sup>, are key to obtaining the result. These two conditions together yield a range of pooling equilibria that is strictly welfare dominant over the set of separating equilibria. If either of the above does not hold, then the separating and pooling equilibria are outcome equivalent, like the quasi-linear case in MT90. See Appendix E for the graphical illustration and more details.

There are a few key differences between the case of purely valence-driven voters and MT90. First, there are two competing informed principals instead of just one. It is not clear whether the gains from pooling will be “competed” away, and whether there will be incentives to pool or information will necessarily be revealed in equilib-

---

<sup>31</sup>The notion Arrow (1963) uses is actually Pareto dominance, which is impossible in our application given the opposing preferences between the politicians and voters, and a fixed economic feasibility set. Therefore we use the weighted aggregated welfare or surplus of the political process instead.

<sup>32</sup>That is,  $U_i^\theta$  in Figure 1 is an area and not just a line.

rium. Second, there is no screening or moral hazard component to the agent (since the only action the voters take is to vote), so the incentive compatibility constraint for the agent is absent in this model. Third, the principals and the agent all have *linear* preferences and feasibility; the conditions laid out in MT90 require concavity in the utility functions. Proposition 2 suggests that the insight of MT90 still applies in this model for an appropriate range of preference parameters.

As mentioned earlier in the description of parameters,  $\psi$  represents the intensity of voter's preference for valence relative to policy. The case of purely valence-driven voter, therefore, should correspond to  $\psi \rightarrow \infty$ . Intuitively, this is because as  $\psi \rightarrow \infty$ , the likelihood of having a draw of the valence component lying in a range such that there is policy competition approaches zero. Unfortunately, the distribution for  $\psi c$  is not well-defined for  $\psi = \infty$ , but we can show that the set of equilibria for the case of purely valence-driven voter coincides with the set of equilibria in the limit as  $\psi \rightarrow \infty$ .

**Lemma 2.** *Let the set of equilibria in the case of purely valence-driven voter be  $\mathcal{E}^{PI}$ , and the set of equilibria for a given  $\psi$  (and other parameters fixed) be  $\mathcal{E}(\psi)$ .*

$$\lim_{\psi \rightarrow \infty} \mathcal{E}(\psi) = \mathcal{E}^{PI}$$

## 4.2 A Voter Driven by Both Valence and Policy

The welfare result obtained in Proposition 2 would not be very useful if it held only in the limit. Therefore, the next step is to show that the set of welfare dominant pooling equilibria (over the set of separating equilibria) exists over an open set of parameter values. Consider the intermediate case in which the pivotal voter cares

about a mixture of valence and policy. That is, while politicians' policy platforms alone do not determine who wins the election, they have a positive probability of affecting the pivotal voter's choice over the two politicians.

Since politicians only know the distribution of the valence component, and  $F(\cdot)$  is a continuous function with density  $f(\cdot)$ , offering a contract that yields the pivotal voter infinitesimally higher payoff than one's opponent still results in an increase in the probability of one's contract being accepted, but this increase is now a smooth function of one's offer.

Below we establish upper hemicontinuity of the equilibrium correspondence, which completes the proof of Proposition 1 when combined with Proposition 2. While it is proven only with respect to  $\psi$ , it can easily be extended to show that the equilibrium correspondence is also upper hemicontinuous with respect to  $\gamma, \lambda, \alpha$ , and  $\eta$ . The upper hemicontinuity of the Perfect Bayesian Equilibrium correspondence follows similar lines as existing arguments for upper hemicontinuity of the Nash correspondence (see, for instance, Fudenberg and Tirole (1991)).

**Lemma 3.** *The set of Perfect Bayesian Equilibria is upper hemicontinuous in  $\psi$ .*

Recall the upper hemicontinuity of the equilibrium correspondence implies that at points of  $\psi$  where continuity might fail, it could only be that the set of equilibria at that point is larger but not smaller – there might be equilibria at point  $\check{\psi}$  that cannot be reached by any sequence of equilibria given any  $\psi^n \rightarrow \check{\psi}$ . Since the welfare comparison in the case of purely valence-driven voter is  $\inf \mathcal{W}^{P\theta} \geq \sup \mathcal{W}^{S\theta}$  for all  $\theta$ , even if lower hemicontinuity fails and the set of equilibria “very close to” the case of purely valence-driven voter is smaller, the welfare comparison between the sets of separating and pooling equilibria will still hold. Therefore, provided that the

pivotal voter is sufficiently valence-driven, there are parameter configurations such that aggregate welfare is higher when the relevant information is not revealed to voters in the process of political competition.

A natural question that follows is whether this result applies for all levels of policy competitiveness, that is, for all  $\psi \in \mathbb{R}_+$ . The answer turns out to be no. The easiest way to understand this is by looking at the case of pure policy, or when  $\psi = 0$ . Equilibrium characterization for this case is omitted<sup>33</sup> since it is not central to proving Proposition 1; however, the intuition is as follows.

As discussed earlier, in the case of pure policy the pivotal voter will claim most of the surplus because of the competition between the politicians along the policy dimension. In a separating equilibrium, the state of the world is known with certainty on the equilibrium path, so the politicians must race to the bottom *state-by-state* in order to compete. Therefore, the preference-aligned politician, who has a competitive edge along the policy dimension, almost always wins (the only exceptions are ties), though the multiplicity of equilibria ranges from the preference-misaligned politician's break-even point to where the preference-aligned politician breaks even.

In a pooling equilibrium, however, since each politician must offer the same policy platform across the two states, the same politician must win in both states. It also means that there is always one preference-misaligned politician that is winning with positive probability in each state. This intuition suggests that the pivotal voter's utility is lower in the set pooling equilibria relative to the separating equilibria, although the multiplicity of equilibria described above makes a direct comparison inconclusive. The calculation and comparison of aggregate welfare are further complicated by the fact that different politicians win in each class of equilibria. A lower

---

<sup>33</sup>It is available from the author upon request.

equilibrium payoff for the voter does not necessarily imply a higher equilibrium payoff for the winning politician, since the identity of the winner may be different in each class. This problem does not arise in the case of purely valence-driven voter, because once we fix a realized  $c$ , we will have the same winner for both states and both classes of equilibria. Aggregate welfare across the sets of separating and pooling equilibria generally overlaps in the case of pure policy.

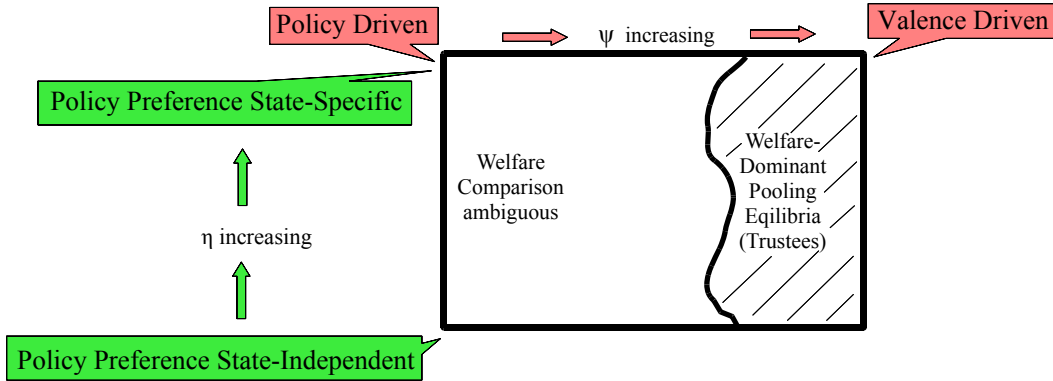


Figure 2: Summary of the Welfare Implications of the Model.

Figure 2 lays out the welfare results of the model. The model spans two dimensions parameterized by  $\eta$  and  $\psi$ :  $\eta$  describes the level of state-specificity for the voter's public good preference, while  $\psi$  denotes the intensity of policy competition. Proposition 2, Lemma 2, and Lemma 3 together imply that the set of pooling equilibria is welfare dominant for  $\psi$  sufficiently large. However, from the case of pure policy, we know that the area in which the set of pooling equilibria welfare-dominates cannot possibly span the entire box. Where the welfare ranking between the sets of separating and pooling equilibria changes depends on the specification of  $F(\cdot)$  and the values of other parameters.

### 4.3 Equilibrium Refinement

The multiplicity of equilibria is a well-known problem for signaling games. In particular, it is often possible to sustain intuitively “unreasonable” equilibria using “punishment by beliefs”, therefore it is important to discuss refinement of the set of Perfect Bayesian (or Sequential) Equilibria. Since our main result establishes that  $\inf \mathcal{W}^{P\theta} \geq \sup \mathcal{W}^{S\theta}$  for all  $\theta$  for an appropriate set of parameter configurations, the qualitative feature of this result will survive essentially all types of equilibrium refinement<sup>34</sup>. Our concern is the non-emptiness of each class of equilibria. In this section we consider the refinement used widely in the signaling literature: the Intuitive Criterion of Cho and Kreps (1987).

The Intuitive Criterion posits the following question about the equilibrium<sup>35</sup>. Upon observing an out-of-equilibrium action by a principal, the agent asks, “Is there a type of principal that will never find it profitable to take this out-of-equilibrium action *regardless* of beliefs that I may have, knowing that I will best respond given these beliefs?” If so, this type must be eliminated from the support of the agent’s beliefs following this off-path action.

In our model, the Intuitive Criterion limits the set of equilibria by “skimming the top” – it rules out equilibrium offers that give the agent higher payoffs from the set of Perfect Bayesian Equilibria. The intuition is not difficult to understand. Effectively what the Intuitive Criterion allows the preference-aligned type to do is to “signal her type” in a way that cannot possibly be profitable for the preference-

---

<sup>34</sup>In fact, equilibrium refinement may *strengthen* our results by possibly widening the gap between  $\inf \mathcal{W}^{P\theta}$  and  $\sup \mathcal{W}^{S\theta}$  for some  $\theta$ .

<sup>35</sup>The Intuitive Criterion is defined for the signaling of one party, while our model describes two informed principals competing. Since we are checking the condition equilibrium by equilibrium, we fix one principal to be offering her equilibrium contract, and impose the Intuitive Criterion on the other principal.

misaligned type. In a Perfect Bayesian (or Sequential) Equilibrium, there are many offers that are sustainable because of the relative freedom in assigning the agent's beliefs off the equilibrium path; in particular, we are able to sustain a range of  $U_i^{\theta\theta}$  up to where even the preference-*aligned* principal yields zero utility when her contract is accepted. The Intuitive Criterion will rule out a vast majority of such  $U_i^{\theta\theta}$  ranges, since the preference-aligned principal can always “break the equilibrium” by proposing an alternative contract that yields negative utility if implemented by the misaligned type. The agent, on observing such an offer, must assign probability zero to the misaligned type.

Proposition 3 formally states our discussion above. The full description of the sets of separating and pooling equilibria given the Intuitive Criterion and the proofs are relegated to Appendix H.

**Proposition 3 (Robustness with respect to the Intuitive Criterion).** *The sets of separating and pooling equilibria satisfying the Intuitive Criterion are non-empty and the welfare ranking preserves.*

## 5 Conclusion

The idea that the observed adoption of institutions may be explained by welfare comparisons goes back to at least Arrow (1963). A goal of this paper is to study the welfare implications of politicians assuming either the role of delegates or trustees in a representative democracy. I model competition between two politicians, who are informed about a payoff-relevant state. The pivotal voter has state-dependent policy preferences and preference for the two politicians' valences. In this environment, will competition between the politicians always result in the revelation of information?



Under what conditions will welfare be higher if politicians are treated as trustees and given legislative discretion?

This model of information asymmetry allows us to compare welfare across two classes of equilibria: separating and pooling. I argue that these two classes of equilibria parallel the models of delegate and trustee representation respectively. In a separating equilibrium, having the information fully revealed to them, voters elect the politician whose policy most closely reflect their preference. In a pooling equilibrium, voters must defer policy-making authority to the elected official, with no ability to influence policy ex post. Politicians are given discretion to implement the appropriate policy, and their role as the information keeper remains integral throughout the political process.

This paper finds that the policies offered by the politicians converge or diverge depending on the state specificity of the pivotal voter's preference for the public goods. Politicians offer divergent policy platforms if the pivotal voter's public good preference is sufficiently state-independent. Interestingly, this is true regardless of the intensity of policy competition between the two principals. The model in this paper distinguishes between getting a public good that the voter prefers and the voter getting a high payoff. The former depends on the pivotal voter's public good specificity ( $\eta$ ), while the latter depends on the intensity of policy competition between the two politicians ( $\psi$ ). If the pivotal voter is sufficiently policy driven but his public good preference is not very state-specific, he can have a relatively high equilibrium payoff but only getting the public good that he less prefers.

Equilibrium characterization shows that information is often withheld from voters even when politicians compete – the sets of separating and pooling equilibria always coexist. In fact, the main result of this paper identifies the existence of wel-

fare dominant *pooling* equilibria, and is interpreted as a possible welfare justification for elected representatives to be treated as trustees. For the set of pooling equilibria to be welfare dominant, we need the pivotal voter to be sufficiently valence-driven, and direct transfers to and from the politicians to have sufficient welfare significance. In this case, aggregate welfare is higher if voters select a politician prior to knowing the relevant information, so that the elected official can exercise discretion in choosing policy. The welfare explanation is robust to any equilibrium refinement with which the sets of separating and pooling equilibria remain non-empty.

There is a rough sense in which the welfare comparison made in this paper may be extended to the systems of direct versus representative democracy. If we take the existence of politicians as given due to transaction costs or frictions<sup>36</sup>, then delegate representation is simply a pragmatic solution to the practical difficulty of issue-by-issue voting in a direct democracy. Only when we adopt the model of trustee representation would representative democracy serve a substantial purpose beyond reasons of transaction costs. With this connection, the main question that this paper poses may be interpreted as an attempt to identify conditions under which “true” representative democracy – where politicians are given legislative discretion – may be justified on welfare grounds.

A few simplifying modeling assumptions were made in this paper. First, the utility functions and feasibility constraints are of specific parameterized forms; second, the politicians are assumed to have perfect information about the state; third, full commitment on the policy platform is assumed for both the politicians and voters. My conjecture is that qualitative features of the results will likely generalize if ei-

---

<sup>36</sup>Even Switzerland, a country whose political system is often described as the closest to a direct democracy, has a parliament (the Federal Assembly) and elected representatives.

ther of the first two assumptions is relaxed, though possibly at the expense of a less precise equilibrium characterization and welfare comparison. Whether or not the third assumption can be relaxed will depend on the structure of non-commitment and renegotiations.

# Chapter 2: Are Democracies more Stable than Non-Democracies? Evidence from Cross-Country Panel Data

## 1 Introduction

Though much studied throughout history, the value of democracy has been of especially great interest to social scientists since the twentieth century, which witnessed two world wars, the Cold War, and more instability across nations spurred by domestic and international conflicts. In search for what may lead to stability, prosperity, equality, and growth, many argue that democracy is a salient element, or even a universal criterion (Sen, 1999) towards achieving these ends. This paper is an empirical study focusing on one specific aspect of comparative political institutions: are democracies systematically more stable than non-democracies? I explore what contributes to the stability of a political institution, and whether the institution in place has an effect on stability, once other relevant factors are controlled for. Does democracy (the institution) itself contribute to stability, or is democracy an externality impacting institutional change through interactions among countries?

Before proceeding further, it is important to consider what one might mean by “political institution” and “stability.” Political institution can broadly refer to any rules of the political system; throughout this paper it is characterized by the level of openness in the government, the competitiveness in the selection of government leadership, the checks and balances constraining the power of the executive, and the citizens’ ability to participate in the political process (Marshall and Jaggers, 2005). These characteristics are summarized in a one-dimensional spectrum in which

at one end there is (strong) democracy, and at the other end there is (strong) autocracy. While there are other features of government that may be considered when characterizing political institutions (for instance, the number of competing parties in government, whether the government is ruled by one or several coalition parties, or the level of compromise among parties in the legislature), the definition used includes a broad spectrum of characteristics that are applicable to all countries, rather than only relatively established democracies.

Stability, in the context of the current paper, refers to a lack of significant change<sup>37</sup> in the characteristics of a political institution. In reality a change in political institution can often be thought of as a change in the country's constitution that alters its political structure in aspects described above. Changes in leadership within a political party, the number of parties sharing power, or the distribution of power across parties do not constitute a change in political institution, in so far as the rules of government remain unchanged.

Why is stability relevant or important? First, it is likely that frequent changes in the rules of government amplify the uncertainties faced by risk-averse individuals, leading them to choose actions that they may consider inferior in the absence of these added uncertainties. In the context of trade and international relations, instability of a political institution limits the enforcement and expected impact of any agreements ratified between firms or countries, which in turn stifles economic and political cooperation that may otherwise be mutually beneficial. Second, instability adversely affects the incentives of the individuals or parties with political power. The fact that the political institution may change overnight and the leader might be stripped of

---

<sup>37</sup>What constitutes a "significant change" will be formally defined when the econometric model is introduced.

all his powers (or in many cases imprisoned or killed) may lead him to discount the future heavily and take myopic actions that hurt the country and its people. Third, the change in institution defined in this paper is often associated with periods of violence in which the destruction of physical and human capital is commonplace<sup>38</sup>. These outbreaks inflict massive and horrific costs on the citizens, and often have a long-lasting effect on their economic and social development.

It is important to note at the outset that this paper does not take a stance on the value of stability in a democracy versus an autocracy – in fact, the main goal of this paper is to shed light on the value of a democracy purely from a stability standpoint. Therefore, I consider institutional stability across both democracies and non-democracies, rather than only look at changes *towards* democracy.

The idea of this paper was incepted as a test of the theoretical literature on democracy and stability. Beginning from the seminal works of Douglass North<sup>39</sup>, a strand of theoretical political economy studies why and how institutions persist. In recent literature, Acemoglu et al. (2008), Acemoglu and Robinson (2008), Jordan (2006), and Lagunoff (2009) have focused on characterizing rules or institutions that are stable. Democracy, or features and rules resembling what this paper defines as democracy, emerge from a number of these papers as a (more) stable institution. However, to the best of my knowledge there has not been an empirical study that formally identifies factors that affect institutional stability, and whether the status quo political institution has a significant effect on it. This paper tests the hypothesis that democratic institutions are more stable than non-democracies.

---

<sup>38</sup>Of course, violence may occur even in the absence of institutional change (for instance, when one autocratic leader overthrows another), but one can think of institutional change as an often sufficient but not necessary condition for violence.

<sup>39</sup>A majority of North's work (such as North (1990) or Alston et al. (1996)) includes some empirical elements, although most of the statistics are descriptive in nature.

In the field of political science many have argued that democracy should be promoted because (among other benefits) it leads to greater stability within the country and across all nations. Examples include Diamond (1992), Boutros-Ghali (1995), Sen (1999), and Ikenberry (1999). The reasons range from an idealistic or moral argument for democracy, to a more pragmatic strategic objective that spreading democracy minimizes conflicts and aggression among fellow democracies. In addition to testing whether democracy has a significant effect on institutional stability, I attempt to separate the channels through which democracy may affect stability; does a country's adoption of democracy lead to higher stability for itself, or does the benefit of a democracy come from its positive externality of higher stability on other democracies? This will allow us to identify what aspects of democratization are important towards promoting institutional stability, and may have foreign policy implications for governments that seek stability for themselves or other countries.

In the empirical literature on political institutions, there is an extensive study on the relationship between institutions and economic growth. Whether democracy is good for growth seems to depend on the methodology employed and the data or instruments used<sup>40</sup>. In the other direction of causality, the effect of economic factors on the likelihood of being a democracy is also far from settled<sup>41</sup>. Little has been done establishing the relationship between political institutions and institutional stability. The closest to the current paper is Persson and Tabellini (2009); however, their primary interest is to establish the three-way link between democracy, stability, and economic growth, and the focus of institutional instability is exclusively on the

---

<sup>40</sup>Examples include Barro (1996), Feng (1997), Rodrik et al. (2002), and Dollar and Kraay (2003).

<sup>41</sup>See Persson and Tabellini (2009) for a brief review of the literature on the relationship between institution and economic growth.

hazard rate out of autocracies.

A major problem in evaluating the effect of institutions on institutional change is that the institution variable is correlated with many unobservables (culture, history, characteristics of the people) in determining the likelihood of institutional change. Without any controls to restore the orthogonality between the error and the institution variable, the estimates obtained in a maximum likelihood estimation will be biased. Controlling for each country's fixed effects is a reasonable first start to eliminate the endogeneity described. Unfortunately, fixed effects probit could not have been estimated correctly until relatively recently because of the incidental parameters problem<sup>42</sup>, which results in the estimates being biased. In this paper I use the correction method proposed by Hahn and Newey (2004) to reduce this bias. Beyond identifying which factors have a significant effect on institutional change, it is also instructive to estimate the magnitude of the effects to understand their economic significance. For non-linear models, the marginal effect of the regressor is not immediate from the estimated coefficient; to get a sense of the magnitude, the average marginal effect of each factor on institutional change is calculated.

This paper uses data on 56 countries between 1980-2007. The institution variables are obtained or generated from the Polity IV dataset, and the economic variables are from the World Bank World Development Indicators. The results identified are only for countries that are in existence for the entire 28-year period and have experienced institutional change during this time. This is because the identification of the fixed effects estimates requires that the status of the variable of interest (institutional change) changes at least once. For that reason, this paper does not explain why some political institutions have persisted for over a hundred years (for example,

---

<sup>42</sup>See Greene (2002) for a more detailed discussion.



the United States and some European countries), but it helps us understand some of the factors that contribute to institutional change for countries that are either more prone to institutional change, or have experienced institutional change in the recent past.

Assuming that the endogeneity of the institution variable operates solely through country fixed effects (time-invariant, country-specific characteristics such as culture, main ethnicities, or language), I estimate the likelihood of institutional change using fixed effects probit with bias correction. I find that having a democratic institution per se does not lead to a decrease in the likelihood of institutional change. The consistently significant factor is the interaction between democracy and the percentage of democracies in the world. This effect is significantly negative, which suggests that the interactions between democracies and non-democracies play a more important role in affecting institutional change than the institution itself. Increasing the percentage of democracies in the world, for instance, *increases* the stability of democracies. In addition, a country with a higher urban population is also associated with a lower probability of institutional change. Surprisingly, while institutional persistence, GDP, and levels of trade are all positively correlated with stability, they do not have a significant effect on the likelihood of institutional change.

Further classification of political institutions into democracy, autocracy, and intermediate ranges yields stronger results confirming the conclusions above. In particular, the presence of and interaction with democratic countries have a significant impact on the likelihood of institutional change, in that having a higher percentage of democratic institutions in the world increases the stability of democracies and reduces the stability of autocracies. Interestingly, having a higher percentage of autocracies in the world also reduces the stability of current autocracies. On

the criterion of status quo institution alone, the data suggests that autocratic and democratic institutions are both associated with a lower probability of institutional change than institutions in the intermediate range of the democracy-autocracy spectrum.

Finally, I check the robustness of the results by including time fixed effects and testing whether the endogeneity of the institution variable is operating through time-variant and/or time-invariant unobservables. Controlling for the time trend in the data yields very similar results as the original specification; however, the endogeneity of the institution variable operates through time-invariant fixed effects as well as time-variant unobservables. The main difference in estimation results when controlling for this additional source of endogeneity is that being a democracy is associated with a lower likelihood of institutional change, and the point estimate is significant under two-stage least squares estimation.

The rest of the paper is organized as follows: section 2 describes the econometric model estimated and the data; section 3 provides results of the fixed effects probit estimation; section 4 discusses alternative specifications and tests the robustness of the results; and finally section 5 concludes.

## **2 Econometric Estimation**

The main objective of this paper is to identify factors that contribute to the likelihood of institutional change. In this section, I will first describe the econometric model used and the assumptions needed, and then discuss the variables that are included and the data sets from which they are obtained.

## 2.1 Econometric Model

Let  $i = 1, \dots, n$  denote the countries, and  $t = 1, \dots, T$  denote the time horizon. The primary equation to be estimated is

$$\Delta_{it} = I(\beta'x_{it} + \delta'D_{it} + e_{it} > 0) \quad (1)$$

where  $\Delta_{it}$  is a binary variable = 1 if institutional change takes place, = 0 otherwise. The vector of explanatory variables that affects the likelihood of institutional change is given by  $x'_{it}$ ;  $D'_{it}$  is a (vector of) binary variable(s) indicating the institution in place in country  $i$  at time  $t$ ; and  $e_{it}$  is a random variable capturing the error and unobservables in (1) that is assumed to be *i.i.d.* and normally distributed. Provided that

$$E[e_{it}|D_{it}, x_{it}] = 0 \quad (A1)$$

(1) can be estimated by probit.

There is legitimate concern that the error is correlated with the regressors in (1). In particular, country-specific characteristics such as the culture, history, and the citizens' ability to mobilize will likely affect both the likelihood of institutional change and the institution variable  $D_{it}$  (which means (A1) is violated). This leads to biased estimates of all index coefficients unless the endogeneity of  $D_{it}$  is addressed.

Many of the unobservable factors correlated with both the institution and institutional change are country-specific characteristics. In fact, some might argue that these characteristics that are often difficult to capture in data are more important in affecting the status quo institution and institutional change, which explains why policy advice that works for one country might not be good for another. For

most of this paper, it is assumed that the endogeneity of the institution variable operates solely through time-invariant characteristics that are specific to country  $i$  (the exception is section 4, in which the validity of this assumption is tested). This assumption implies that once we control for country fixed effects ( $\alpha_i$ ), the error is no longer correlated with the regressors<sup>43</sup>:

$$E[u_{it}|D_{it}, x_{it}, \alpha_i] = 0 \quad \forall i, t \quad (\text{A2})$$

Therefore, the primary equation can be estimated by fixed effects probit:

$$\Delta_{it} = I(\beta'x_{it} + \delta'D_{it} + \alpha_i + u_{it} > 0) \quad (2)$$

The fixed effects probit estimator maximizes the following log likelihood function by jointly choosing  $\beta$ ,  $\delta$ , and  $(\alpha_i)_{i=1}^n$ :

$$\begin{aligned} \mathcal{L}(\beta, \delta, \alpha) = & \sum_{i=1}^n \sum_{t=1}^T \{ \Delta_{it} \ln [\Phi(\beta'x_{it} + \delta'D_{it} + \alpha_i)] \\ & + (1 - \Delta_{it}) \ln [1 - \Phi(\beta'x_{it} + \delta'D_{it} + \alpha_i)] \} \end{aligned}$$

where  $\Phi(\cdot)$  is the cumulative distribution function of a normal distribution. It is important to note that in order for  $\alpha_i$  to be identified, for each  $i$ , there needs to be some  $t$  in which  $\Delta_{it} = 0$ , and some  $\hat{t}$  in which  $\Delta_{i\hat{t}} = 1$ . The reason for this requirement is that  $\alpha_i$  only appears in country  $i$ 's likelihood function and not any other country's. Suppose that some country  $i$  has  $\Delta_{it} = 1$  for all  $t$ , then the

---

<sup>43</sup>The corresponding error is relabeled as  $u_{it}$  to denote the inclusion of fixed effects in the main equation

fixed effect probit estimator chooses  $\alpha_i$  to maximize country  $i$ 's contribution to the likelihood,

$$\max \sum_{t=1}^T \ln [\Phi (\beta' x_{it} + \delta' D_{it} + \alpha_i)]$$

which does not have a well-defined solution for  $\alpha_i$ . This is why the sample used in this paper consists only of countries that have experienced institutional change during the period of study<sup>44</sup>.

Another issue with fixed effects probit that is well noted in the econometrics literature is the incidental parameters problem first pointed out by Neyman and Scott (1948). The basic problem is that in panel data, it is much easier to increase the number of individual observations, but the number of periods in which they are observed is relatively fixed. Because the increase in sample size is mostly due to having more individuals, as we increase the sample size there is also a need to estimate an increasing number of fixed effects, which leads to bias in the fixed effects estimates. The fact that the index coefficients are estimated jointly with the fixed effects means the estimation of the former will be affected by the bias of the latter. The incidental parameter problem is most prominent with data that has a very short time horizon; the period studied in this paper is 1980-2007, which is relatively long from the perspective of the problem. Nonetheless I employ the Hahn and Newey (2004) analytical bias correction to ensure that the estimates obtained are consistent; the standard errors will also need to be adjusted after the bias-corrected estimates are obtained<sup>45</sup>.

---

<sup>44</sup>Technically the requirement is that each country in the sample must have the institutional change variable being both 0 and 1; however, the latter constraint is the one that binds, while the former is not restrictive in the data.

<sup>45</sup>The basic steps are to first bias-correct the index coefficients, then re-maximize each individual's log likelihood choosing only their individual fixed effects, fixing the bias-corrected estimates. The variance-covariance matrix can be obtained next by evaluating the Hessian

## 2.2 Data

What are the factors that may be argued intuitively to have an impact on institutional stability? This paper explores five broad categories:

- Political institution –

This is the key variable that motivates this paper: are democracies more stable, once all relevant factors are controlled for?

- Persistence –

It has been argued that institutional inertia (for instance, see Coate and Morris (1999)) is an important factor in the stability of an institution – the longer an institution has been in place, the more costly it is to change it. While the connotation is mostly negative, persistence could be interpreted in a positive sense as well, as illustrated in Persson and Tabellini (2009).

- Economics –

GDP is certainly a factor to be included given the extensive literature on growth and stability (see section 1); the interdependence of countries through trade may also increase (increased resource conflicts) or decrease (increased reliance on and interest in the stability of trade partners) institutional stability.

- Demographics –

Demographic information is useful in understanding part of a country's characteristics, its policy emphasis, and the ease of information transfer among citizens. I use the level of urbanization of each country as a proxy for the ease of information exchange and the costs of communication and organization

---

of the log likelihood at the bias-corrected estimates and new fixed effects.

among the citizens.

- Foreign influence –

Throughout history, there have been many instances in which foreign nations directly or indirectly interfere and cause institutional change within a country. It can also be argued that being exposed to more of one type of institution leads to higher acceptance, or greater realization of the benefits and costs of that institution.

The data used in this paper is obtained from two main sources: the Polity IV data set, compiled by the Center for Systemic Peace; and the World Development Indicators (WDI), compiled by the World Bank.

The Polity IV data set has information on 187 countries from 1800-2007, but most countries only have data available from 1960 onward, with missing values scattered frequently in between. The WDI, on the other hand, has economic and demographic data on 209 countries between the years of 1960 and 2008 only. Unfortunately, a significant percentage of countries that this paper would like to capture (having experienced institutional change) do not have WDI data until after 1980; therefore, the sample constructed merges the two data sets from 1980-2007.

The variables of interest from the two data sets are listed in Table 1. We obtain two key variables from Polity IV: *polity2* and *durable*. The institution variable is *polity2*, which ranges from -10 to 10 and describes the institution that is in place; -10 denotes the highest level of autocracy, 10 denotes the highest level of democracy. It is a composite index that evaluates five elements of a political system: the competitiveness of executive recruitment; the openness of executive recruitment; constraints on the chief executive; the competitiveness of political participation; and the regula-

tion of participation (Marshall and Jaggers, 2005, p.14-15). This variable is also an improvement from the variable *polity* for the purpose of time-series data analysis, by smoothing the composite index across transition or interruption periods<sup>46</sup>. The variable *durable* captures the number of years since there is a three-point or more change in *polity2* over a period of three years or less.

All variables related to a country's political system used in this paper are generated from *polity2* and *durable*. First, *change* is a binary variable defined to take on value 1 if there is a change in *polity2* score of three or more from the last period; otherwise it is zero; the level chosen is to be consistent with the definition of the *durable* variable in Polity IV. The variable *persist* is a one-period lag of *durable*. If *durable* were used as a regressor, then whenever institutional change occurs, by definition *durable* would always be reset to zero, meaning that there would be colinearity issues in estimation. Therefore the one-period lag of *durable* is used to identify the impact of persistence on institutional stability – up until last year ( $t - 1$ ), how long it has been since the political institution last had a three point or more change.

The three institution variables are *d5*, *a5*, and *d0*. The main democracy variable is *d5*, defined to be 1 if *polity2* is strictly greater than 5; otherwise it is zero. The autocracy variable *a5* is a mirror image of *d5*: it is equal to 1 if *polity2* is strictly smaller than -5; zero otherwise. As an alternative (and broadened) definition of a “democracy”, *d0* is constructed to include countries that has a *polity2* score of strictly above zero.

Note that as defined in this paper, having the variable  $change = 1$  does not necessarily imply a change from democracy to non-democracy (or vice versa). For

---

<sup>46</sup>In the *polity* variable, a country going through a transition or interruption period often does not have data, or is given separate categories denoting transition or interruption.



instance, if a country were to move to  $-9$  to  $-4$ , it still remains an autocracy but has  $change = 1$ . On the other hand, if a country went from  $4$  to  $6$ , using the  $d5$  definition of democracy it went from a non-democracy to a democracy, but  $change = 0$ .

With the institution variables, we can calculate the percentage of democracies (based on either definition) and autocracies in the full set of data<sup>47</sup>; they are given by  $pcd5$ ,  $pca5$ , and  $pcd0$ . To study the full extent of foreign influence on the likelihood of institutional change, it is useful to know how the impact of the percentage of democracies or autocracies differs across countries with different status quo institutions. To do so, I generate interactions between the institution variables and the percentages of democracies and autocracies.

There are three economic or demographic variables obtained from the WDI. The economic variables are  $gdp05$ , which describes the GDP per capita in 2005 international dollars, adjusting for purchasing power parity across nations; and  $trade$ , which describes trade as a percentage of GDP. To capture the demographic characteristics,  $urban$  is the percentage of urban population in a country.

As mentioned in the description of the econometric model (subsection 2.1), the current model can only identify fixed effects for countries that have both 0 and 1 for the institutional change variable in the period studied. In addition, the sample is limited to countries that are in existence for the entire duration studied. This means that this paper does not capture, for instance, the wave of institutional changes seen after the fall of the Soviet bloc, leading to the independence of countries that were

---

<sup>47</sup>By “full set” I mean the original data before a sample with only countries that have complete data and experienced institutional change is extracted.

previously under Soviet control<sup>48</sup>. There is a total of 56 countries in the sample; the list of countries included are given in Table 2.

The summary statistics of the original data sets and the sample used in this paper are given in Table 3. As one can imagine, there is a greater percent of countries with  $change = 1$  in the sample (relative to the original, full data set), and hence the average number of years in the *durable* or *persist* variable is also lower in the sample. The other significant difference to note is that the sample has on average much lower per capita income than the original data set.

Correlation among the key variables are given in Tables 4, 5, and 6. Democracies (defined by  $d5$  or  $d0$ ) appear to have a higher GDP, lower *persist*, and higher urban population than non-democracies and autocracies<sup>49</sup>. Interestingly, there is no marked difference in the percentage of trade across different political institutions. Among the continuous variables,  $gdp05$  and  $urban$  have a high correlation at 0.77.

Trends of political institutions and institutional changes across time give an interesting description of the political climate in the world in this period. In Figure 6, we can see that there is a steady wave of democratization between 1980-2007. Instances of institutional changes (shown in Figure 7), however, are scattered in the period with no immediately obvious trend, and do not occur frequently. If we separate the instances of institutional changes among democracies and autocracies (Figures 8 and 9), we can see from casual observation that the fraction of countries with institutional changes among democracies seems to decrease over this time pe-

---

<sup>48</sup>This is done because the institutional change designations of these countries at the beginning of their independence are problematic; also, it will skew the sample to have the fall of a single country leading to the creation of and immediate institutional changes in so many countries, given the relatively rare occurrence of institutional change among others in the sample.

<sup>49</sup>As defined in this paper, autocracies are a subset of non-democracies.

riod, while the fraction of countries with institutional changes among autocracies appears to increase. These trends may be due to systematic differences in the likelihood of institutional change between democracies and autocracies, or may simply reflect a higher number of democracies and lower number of autocracies over time for a relatively stable likelihood of institutional change.

### 3 Results

This section describes the results obtained using fixed effects probit, discusses the differences across different specifications of the model and definitions of the institution variable(s), and analyzes the average marginal effects of the variables of interest on the likelihood of institutional change.

#### 3.1 Fixed Effects Probit

Results of the basic model is given in Table 7. The first thing to notice that is being a democracy (having  $d5 = 1$ ) has a negative but insignificant effect on the probability of institutional change. Persistence, though often hypothesized to play an important role, surprisingly also does not have a significant effect on the likelihood of institutional change. These results hold true regardless of the inclusion or exclusion of some of the relevant variables. Of course, as noted before the results are specific to a sample of countries that have all experienced institutional change at some point in the period studied, and hence are not meant to be representative of the full set of countries in the world.

The institution-based variable that seems significant in its impact on the likelihood of institutional change is the interaction between  $d5$  and  $pcd5$ . The effect

is negative, which means that having a higher percentage of democracies reduces the likelihood of institutional change among democracies. For non-democracies, the effect of having a higher percentage of democracies is not significantly different from zero. However, the point estimates suggest that having a higher percentage of democracies is associated with a higher probability of institutional change for non-democracies.

Among economic and demographic variables, *urban* is the only variable that has a significant effect on the likelihood of institutional change. A higher level of urban population is associated with a lower probability of institutional change. GDP and trade, while both having expected negative signs – meaning that a higher GDP or level of trade is associated with a lower likelihood of institutional change – are not significant regardless of the model specified.

To illustrate the importance of including country fixed effects, the result of probit estimations *without* fixed effects is shown in Table 8. We can see that when time-invariant, country-specific characteristics are not controlled for, the coefficient on democracy is significant and negative, while most of the other variables are not (the only exceptions were the trade and interaction variables, which were significant in one model specification each). The difference in results between probit and fixed effects probit suggests that these time-invariant country-specific characteristics are important in affecting the likelihood of institutional change, and therefore cannot be ignored. Though not shown in the paper<sup>50</sup>, the fixed effects obtained in the estimations are in fact mostly significant.

It is important to check whether the results obtained above is specific to the range of *policy2* considered as a democracy in this paper. To do so, I widen the

---

<sup>50</sup>Available from the author upon request.

definition of democracy to include all countries that have a *polity2* score above 0, by using the *d0* variable instead of *d5*. While the general features of the results are the same whether *d5* or *d0* is used as the institution variable – the only variables that are significant are *urban* and the interaction between *d0* and *pcd0*, both having negative signs as before – the significance of the interaction between institution (*d0*) and percentage and democracy (*pcd0*) is much weakened when we take a more inclusive definition of democracy.

Next, I classify the institution variable into three categories – democracy ( $d5 = 1$  as before), autocracy ( $a5 = 1$ ), and the intermediate range ( $d5 = a5 = 0$ ). The idea is to obtain a more detailed understanding of how the institution and interaction variables affect the likelihood of institutional change. Accordingly, I also include the percentage of autocracies, and allow for both own and cross interactions between the institutions and the percentages of different institutions in the estimations. The results are given in Table 9.

With a slightly finer classification of political institutions, we can see that the two ends of the *polity2* spectrum ( $d5 = 1$  and  $a5 = 1$ ) are both associated with having a lower likelihood of institutional change relative to the intermediate range. Though the point estimates of *a5* are always above those of *d5*, they are not statistically different from each other. The coefficient of *a5* is always significant, while the coefficient of *d5* is significant only for some model specifications. The more interesting results come from the added interaction variables, which gives a richer description of the interactions among *and across* political institutions. For instance, increasing the percentage of democracies on average lowers the likelihood of institutional change among democracies; however, it also leads to an increase in the likelihood of institutional change among autocracies. Increasing the percentage of autocracies,

on the other hand, has no statistically significant impact on the probability of institutional change among democracies, but is associated with an increased probability of institutional change among current autocracies. Higher percentages of democracies ( $pcd5$ ) and autocracies ( $pca5$ ) are both associated with a higher likelihood of institutional change among institutions in the intermediate range ( $d5 = a5 = 0$ ), though the effect is also not significant.

Results on the other variables are largely consistent with what is found earlier – *urban* is the only other variable with a significant point estimate (a higher level of urbanization is associated with a lower likelihood of institutional change), while the estimates of GDP, persistence, and trade are not significant.

### 3.2 Marginal Effects

In this subsection I look at the marginal effect of each of the variables on the likelihood of institutional change. In most econometric estimations, it is useful (and in fact often of primary interest) to evaluate the effect of an increase or decrease in the explanatory variable on the response variable. However, the point estimates in non-linear models are not informative about the effect of the regressor on the response variable beyond their signs (positive or negative). The marginal effect analysis in non-linear models requires the specification of the levels of the explanatory variables and fixed effects at which the marginal effect is evaluated.

For the purpose of this subsection, the vector of regressors ( $x'_{it}$ ) includes the institution variable(s) ( $D'_{it}$ ) that was separately listed before.

There are two ways to evaluate the marginal effect  $m(\cdot)$ . The first is to look at

the average marginal effect of an observation randomly drawn:

$$\frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n m(x'_{it}\beta + \alpha_i)$$

The second is to look at the marginal effect of an observation with average characteristics:

$$m\left(\frac{1}{nT} \sum_{t=1}^T \sum_{i=1}^n [x'_{it}\beta + \alpha_i]\right)$$

In panel data, especially when binary variables and fixed effects are used, it is unclear how an observation with average characteristics (including “average fixed effects”) should be interpreted. Therefore, I adopt the approach standardly used and evaluate the average marginal effect over the sample.

The evaluation of marginal effect depends on the nature of the variable of interest. Fernández-Val (2009) has a helpful discussion on how to evaluate the marginal effect based on the variable or situation, below I include some of it for the sake of completeness. Suppose we are interested in the marginal effect of variable  $x^k$  on the likelihood of institutional change ( $\Delta_{it}$ ). If  $x^k$  is continuous, then the marginal effect is typically the derivative of the error distribution  $\Phi(\cdot)$  with respect to  $x^k$ , evaluated at some  $\bar{x}_{it} = (\bar{x}_{it}^k, \bar{x}_{it}^{(-k)'})$  of interest:

$$m(\bar{x}_{it}, \beta, \alpha_i) \equiv \frac{\partial \Phi\left(\bar{x}_{it}^k \beta^k + \bar{x}_{it}^{(-k)'} \beta^{(-k)} + \alpha_i\right)}{\partial x^k} = \beta^k \phi\left(\bar{x}_{it}^k \beta^k + \bar{x}_{it}^{(-k)'} \beta^{(-k)} + \alpha_i\right)$$

where  $\psi(\cdot)$  is the *pdf* of the error. If the variable is binary or discrete, then the marginal effect is the difference in  $\Phi(\cdot)$  with a one unit change in  $x^k$  at  $\bar{x}_{it}$ :

$$\tilde{m}(\bar{x}_{it}, \beta, \alpha_i) \equiv \Phi\left((\bar{x}_{it}^k + 1)\beta^k + \bar{x}_{it}^{(-k)'} \beta^{(-k)} + \alpha_i\right) - \Phi\left(\bar{x}_{it}^k \beta^k + \bar{x}_{it}^{(-k)'} \beta^{(-k)} + \alpha_i\right)$$

To calculate the average marginal effect of a binary variable, we only sum over the subsample with the value of the binary variable at which we are evaluating. For instance, the “average treatment effect on the treated” (Fernández-Val, 2009, p.76) is given by

$$\frac{1}{N_1} \sum_{t=1}^T \sum_{i=1}^n \tilde{m}((0, \bar{x}_{it}^{(-k)})', \beta, \alpha_i) \mathbb{1} [x_{it}^k = 1]$$

where  $N_1 = \sum_{t=1}^T \sum_{i=1}^n \mathbb{1} [x_{it}^k = 1]$ .

There are additional issues with evaluating the marginal effect of a binary or continuous variable that has an associated interaction with another variable. They are not discussed in Fernández-Val (2009); however, given the importance of interactions between variables in this paper, I will elaborate on how these marginal effects should be interpreted. For this purpose, let  $x_{it} = (x_{it}^k, (x_{it}^\ell)', (x_{it}^L)')$ , where  $x_{it}^k$  is the variable of interest as before,  $(x_{it}^\ell)'$  is a vector of interaction variables associated with  $x_{it}^k$ , and  $(x_{it}^L)'$  is a vector of the remaining regressors. If  $x_{it}^k$  is a binary variable, then the overall marginal effect for the treated (the set  $\{i : x_{it}^k = 1\}$ ) evaluated at  $\bar{x}_{it}$  is

$$\frac{1}{N_1} \sum_{t=1}^T \sum_{i=1}^n \left[ \Phi \left( \beta^k + (\bar{x}_{it}^\ell)' \beta^\ell + (\bar{x}_{it}^L)' \beta^L + \alpha_i \right) - \Phi \left( (\bar{x}_{it}^L)' \beta^L + \alpha_i \right) \right] \mathbb{1} [x_{it}^k = 1]$$

If the variable of interest  $x_{it}^k$  is a continuous variable interacted with other binary variables, we will have to evaluate the marginal effect of the continuous variable for each possible realization of the binary (or discrete) variables, and as before only sum over the subsample whose effect we are capturing. Suppose  $x_{it}^k$  is interacted with a vector of binary variables  $(x_{it}^q)'$  to form interactions  $(x_{it}^\ell)'$  (so now  $x_{it} = (x_{it}^k, (x_{it}^q)', (x_{it}^\ell)', (x_{it}^L)')$ ). Then the marginal effect of  $x_{it}^k$  for subsample  $\{i : x_{it}^q =$



$1 \forall q$  evaluated at  $\bar{x}_{it}$ , for instance, will be given by

$$\frac{1}{N_q} \sum_{t=1}^T \sum_{i=1}^n \left( \beta^k + \sum_{\ell} \beta^{\ell} \right) \phi \left( \sum_q \beta^q + \bar{x}_{it}^k \beta^k + (\bar{x}_{it}^{\ell})' \beta^{\ell} + (\bar{x}_{it}^L)' \beta^L + \alpha_i \right) \mathbb{1} [x_{it}^q = 1 \forall q]$$

where  $N_q = \sum_{t=1}^T \sum_{i=1}^n \mathbb{1} [x_{it}^q = 1 \forall q]$ ; whereas the marginal effect of  $x_{it}^k$  for sub-sample  $\{i : x_{it}^q = 0 \forall q\}$  evaluated at  $\bar{x}_{it}$  is

$$\frac{1}{N_{q'}} \sum_{t=1}^T \sum_{i=1}^n \left[ \beta^k \phi \left( \bar{x}_{it}^k \beta^k + (\bar{x}_{it}^L)' \beta^L + \alpha_i \right) \right] \mathbb{1} [x_{it}^q = 0 \forall q]$$

where  $N_{q'} = \sum_{t=1}^T \sum_{i=1}^n \mathbb{1} [x_{it}^q = 0 \forall q]$ . The case where the binary variables assume different values can be evaluated similarly.

Since the fixed effects have to be re-estimated after the index coefficients are bias-corrected,  $\alpha_i$  depends on  $\beta$  even if evaluated at the true value of  $\beta$ . The slow convergence of the index coefficients also adds to the bias of the asymptotic distribution of the marginal effects (Fernández-Val, 2009, p.76). Therefore, the marginal effects need to have corrections made in addition to using the bias-corrected estimates for  $\beta$  and  $\alpha_i$ . For consistency with subsection 3.1, I use the bias correction of Hahn and Newey (2004) to calculate the bias-corrected marginal effect of each variable.

The average marginal effects are shown in Tables 10 and 11. The ranges described below denote the marginal effect across different specifications of the model. When political institution is defined by democracy versus non-democracy (i.e. only  $d5$  is used), increasing the percentage of democracies by one standard deviation (10.3%) leads to a 1.6 – 2.6% decrease in the probability of institutional change among democracies on average. Increasing the percentage of urban population by

one standard deviation (20.2%) is expected to reduce the probability of institutional change by 12.6 – 16.5%. Being a democracy ( $d5 = 1$ ) on average results in a 6.8 – 9.8% decrease in the likelihood of institutional change, though the effect is not significant.

When political institution is defined by democracy ( $d5 = 1$ ), autocracy ( $a5 = 1$ ), and the intermediate range ( $d5 = a5 = 0$ ), the results are similar to only having  $d5$ . Increasing the percentage of democracies by one standard deviation is expected to yield a 3.7 – 5.2% decrease and a 12.0 – 14.5% increase in the likelihood of institutional change in democracies and autocracies respectively. Interestingly, increasing the percentage of autocracies (by one standard deviation, or 16.5%) does not significantly impact democracies, but results in a 7.4 – 9.2% *increase* in the probability of institutional change among autocracies on average. Increasing the percentage of urban population by one standard deviation yields a similar effect as before, an average 16.2 – 19.5% decrease in the probability of institutional change. Finally, being a democracy and autocracy on average leads to an 8.1 – 10.0% and 23.9 – 32.8% decrease in the likelihood of change over the intermediate group respectively.

While the magnitudes of marginal effect described above do not appear to be too large, bear in mind that instances of institutional change are not frequent in this time period. Of all countries in the world with data between 1980-2007, the instances of institutional change constitute only 4.6% of total observations; even for the sample analyzed in this paper, in which each country has experienced at least one instance of institutional change, instances of institutional change are still only 8.0% of total observations. For some of the variables, the marginal effects could imply possibly impacting the probability of institutional change a few times over the current level.

## 4 Robustness and Source of Endogeneity

In this section I will address two issues about the robustness of the results. The first is related to the possible time trend observed in the data (recall Figures 6 to 9) that had not been controlled for in the estimations, while the second concerns the assumption about the source of endogeneity in the previous sections.

### 4.1 Time Trend

For time-series data and estimations, it is common to either de-trend the data or control for the time trend in estimations. The most natural way to control for the time trend in a fixed effects probit estimation is to include time fixed effects in the same way country fixed effects are used. The difficulty in adding annual time fixed effects is that it puts additional requirements on the data – as the use of country fixed effects requires *each country* to have both realizations 0 and 1 for its institutional change variable during the period studied, having annual time fixed effects would similarly require that for *each year*, there are realizations 0 and 1 across the countries studied. It also exacerbates the incidental parameters problem, since for each year that is added to the sample, there is an additional parameter to be estimated.

For comparison, the concerns for time trend in the data are addressed in three ways: (i) with annual time fixed effects; (ii) with period dummies; and (iii) with a linear time trend. The results are shown in Table 12, and are compared to column [1] of Tables 7 and 9 depending on whether the institution variable is divided into two or three categories. Columns [1] and [4] are estimated with annual time fixed effects. The estimates are qualitatively similar to the original specification (without

controlling for time trend) – being a democracy ( $d5$ ) does not significantly decrease the likelihood of institutional change relative to a non-democracy ( $d5 = 0$ ) or autocracy ( $a5 = 1$ ), while *urban* and the interactions between the institution variables and the percentage of institutions remain significant (except for  $d5 * pca5$ , which was not significant in the original specification either).

Columns [2] and [5] are estimated with period dummies, with seven periods and four years in each period. While the estimates of the variable  $d5$  are slightly larger when the period dummies are included, their statistical significance are still as in their previous levels. The interaction variables that are significant in the original specification remain so with the addition of period dummies. The results are qualitatively the same when the number of periods are lowered to four (not shown in table), the only difference is that the period dummies are less significant when there are fewer periods.

Finally, columns [3] and [6] are estimated with a linear time trend; we can see that the estimates obtained are again very similar to those of the original fixed effects probit estimations, and the linear time trend is not at all significant in either case.

I conclude that the results obtained in the section 3 is robust to the inclusion of controls for possible time trends.

## 4.2 Time-variant Endogeneity

An important assumption used thus far is that the endogeneity of the institution variable  $D_{it}$  operates solely through time-invariant, country-specific fixed effects. While the issue of time trend is addressed, meaning that the endogeneity of the institution variable may also operate through country-invariant, time-specific fixed

effects without qualitatively affecting the results, it may be argued that there are still remaining unobservables that are both country- and time-variant that may affect both the status quo institution and the likelihood of institutional change. Examples of these unobservables include financial crises or significant social or political events within a country or region. If the endogeneity of  $D_{it}$  does in fact operate through unobservables that are country- and time-variant as well, then assumption (A2) is violated, and the estimates obtained in fixed effects probit will be biased. In this section, I investigate whether the remaining country- and time-variant unobservables constitute part of the endogeneity of the institution variable.

Consider, in addition to the primary equation (2), the following equation that determines the institution in place:

$$D_{it} = I(\psi D_{i(t-1)} + \gamma' x_{it} + \eta_i + v_{it} > 0) \quad (3)$$

where  $D_{it}$  now is just a single binary variable (instead of possibly a vector as before) with  $D_{it} = 1$  indicating a democracy and  $D_{it} = 0$  indicating a non-democracy. We assume that  $u_{it}$  and  $v_{it}$  are jointly normally distributed, with a possibly non-zero covariance, and

$$E[u_{it}|x_{it}, D_{it}, D_{i(t-1)}, \alpha_i, \eta_i] = E[v_{it}|x_{it}, D_{it}, D_{i(t-1)}, \alpha_i, \eta_i] = 0 \quad \forall i, t \quad (A3)$$

The endogeneity of  $D_{it}$  is therefore due to the correlation of the errors and the correlation of fixed effects in (2) and (3) (Fernández-Val and Vella, 2007, p.3).

Given the environment described above, the idea is that while the fixed effects estimation alone yields biased estimates because assumption (A2) does not hold,

the orthogonality condition for maximum likelihood estimation can be restored by controlling for the part of  $u_{it}$  that is correlated with  $v_{it}$ . To do so, first estimate (3), then use the generalized residual of (3) as an additional regressor in (2) as a proxy for the time-varying element of the endogeneity. An important identification assumption required for this two-step estimation is the exclusion restriction in (3) – there must be at least one variable that is *only* in (3) and *not* in (2). The reason for this requirement is that without exclusion restriction, the residual obtained in the first stage will be a combination of only  $D_{it}$  and  $x_{it}$ , and therefore we will not be able to separately identify the coefficient for the residual. So for the two-step estimation to be valid, we need  $\psi \neq 0$ . I use the one-period lag of the institution variable as the exclusion restriction; this means that while the political institution yesterday has a significant effect on the political institution today, it does not have a significant effect on the likelihood of institutional change, once we control for the political institution today.

When performing the two-step estimation by maximum likelihood, a t-test of the generalized residual (obtained from the first step and included in the second step) is valid for testing whether the endogeneity of  $D_{it}$  operates through time-variant components (Rivers and Vuong, 1988), but the other point estimates obtained in the second step are biased and cannot be used for inference. The reason is that the generalized residual is non-linear coming from the first step (probit), and since it is only estimated, the error in the second step contains the difference between the estimated and true generalized residuals and will be non-linear as well. The non-linearity of the error term in the second stage implies that the normality assumption will no longer hold. As is well-known, maximum likelihood methods are very sensitive to whether the model is correctly specified, and therefore the violation

of normality will result in biased point estimates when using fixed effects probit in the second stage.

In order to have a valid t-test for the generalized residual, as in the case of fixed effects probit we need to do bias correction in each of the two steps to reduce possible bias stemming from the incidental parameters problem. In the first step, I use the Hahn and Newey (2004) bias correction as in subsection 2.1. In the second step, however, there are additional sources of bias that is not accounted for in the Hahn and Newey (2004) bias correction. These sources include the correlation of the fixed effects in the two steps, the asymptotic bias of the generalized residual, and the non-linearity of the second step (Fernández-Val and Vella, 2007, p.11). Fernández-Val and Vella (2007) proposes an analytical bias correction specific to two-step estimators for models with both time-variant and time-invariant heterogeneity. This is used in the second step of the estimation.

The procedure of the two-step control function estimator is as follows:

1. Estimate (3) using fixed effects probit, then use the Hahn and Newey (2004) bias correction to obtain  $\hat{\psi}$  and  $\hat{\gamma}$ .
2. Fixing  $\hat{\psi}$  and  $\hat{\gamma}$ , maximize each individual's log likelihood  $\mathcal{L}_i(\eta_i, \hat{\psi}, \hat{\gamma})$  choosing  $\eta_i$ . This yields a vector  $\hat{\eta}(\hat{\psi}, \hat{\gamma})$ .
3. Compute the generalized residual of (3),  $\hat{\lambda}(\hat{\eta}, \hat{\psi}, \hat{\gamma})$ , using  $\hat{\eta}$ ,  $\hat{\psi}$ , and  $\hat{\gamma}$ .
4. Estimate

$$\Delta_{it} = I\left(\beta'x_{it} + \delta D_{it} + \alpha_i + \theta\hat{\lambda}_{it} + \tilde{u}_{it} > 0\right)$$

with fixed effects probit, and use the Fernández-Val and Vella (2007) bias correction to obtain  $\hat{\beta}$ ,  $\hat{\delta}$ , and  $\hat{\theta}$ .

5. Repeat<sup>51</sup> step 2 with the corresponding parameters to obtain  $\hat{\alpha}(\hat{\beta}, \hat{\delta}, \hat{\theta})$ .

The results of the two-step fixed effects probit are listed in Table 13. The main interest is the coefficient of the generalized residual obtained in the first step and included in the second as a regressor. If the coefficient is significant, it means that country-specific fixed effects alone are not sufficient in solving the endogeneity problem of the institution variable, therefore time-variant unobservables must be accounted for as well. T-tests for the variable across all specifications are significant at the 5% level, suggesting that the endogeneity of the institution variable likely operates through time-variant unobservables, in addition to country-specific fixed effects. We can also see from the first-step estimation that the key determinant of the political institution today is the political institution yesterday; conditional on it, other variables does not have a significant impact on the status quo institution.

As explained above, the econometric model in this subsection cannot be estimated correctly in two steps using maximum likelihood estimation (with the assumption that the error in the second step follows a normal distribution). To estimate this model non-linearly while relaxing distributional assumptions on the errors, one could estimate the model semi-parametrically using methods such as Klein and Spady (1993). Although the use of semi-parametric estimation methods is beyond the scope of this paper, we can obtain consistent estimates in both steps using two-stage least squares (2SLS). The only concern for using 2SLS in this application is that for observations in the tails of the regressors, the implied estimated likelihood may not be as accurate as in a non-linear model and may result in the estimated probability lying outside the range of 0 and 1. Nonetheless, the 2SLS estimation

---

<sup>51</sup>The standard errors in each step are obtained in the same way as described in footnote 45.



will give us an idea about how the inclusion of the residual from the first step will impact the estimation of the main equation of interest (2).

Results of the 2SLS estimations are given in Table 14. Columns [2] and [5] are ordinary least squares (OLS) estimation of equation (2) – that is, the linear counterparts to the fixed effects probit estimation of Table 7 columns [1] and [5] – to facilitate comparison. As in the case of fixed effects probit, the institution variable ( $d5$ ) is not significant in its effect on institutional change in the OLS estimation. The interaction between institution and percentage of institution in the world ( $d5*pcd5$ ), and the variable *urban*, remain significant and with their original signs; however, the variable *pcd5* is positive and significant in the OLS estimations, meaning that increasing the percentage of democracies in the world is expected to increase the likelihood of institutional change for non-democracies.

In the 2SLS estimation, the residual obtained from the first stage (*resid*) is significant across various specifications, confirming the hypothesis that the institution variable is correlated with time-variant unobservables, even after fixed effects are included. More importantly, the institution variable ( $d5$ ) is significant and negative in the 2SLS estimation – being a democracy is associated with a 15.7 – 16.8% lower likelihood of institutional change relative to non-democracies. While the magnitude of the interaction variable ( $d5 * pcd5$ ) is smaller in 2SLS than OLS for each comparable specification, the point estimate remains significant and negative; the variable *pcd5*, on the other hand, remains positive and significant but is larger in 2SLS than OLS. In terms of stability, this means that without controlling for the endogeneity of institution through time-variant unobservables, the positive externality of democracies on other democracies (through increased stability) may be overstated, while the negative externality of democracies on non-democracies (through decreased sta-

bility) may be understated.

Combining the results of 2SLS with the fixed effects probit estimations in section 3, I find that the interaction between institution and percentage of institution is significant across different model specifications and estimation methods, highlighting the effect of democracies on institutional change as an externality. Whether or not democracy as an institution also has an impact on the likelihood of institutional change depends on the model assumption about the institution variable – whether the endogeneity of the institution variable operates through fixed effects only, or other time-varying unobservables.

## 5 Conclusion

Institutional stability is undoubtedly an important topic for social scientists. On the empirical front, there has been extensive research on the effect of political stability on economic variables, and the causes and effects of different political institutions on growth. However, there has not been an empirical study formally linking the effect of political institutions on institutional stability, even though the relationship has been explored widely in political science and theoretical political economy. This paper attempts to fill this gap in the literature, and explores whether there are systematic differences in institutional stability between democracies and non-democracies.

Because the institution variable is likely correlated with unobservables in the equation determining the likelihood of institutional change, a standard probit estimation (without time-variant or time-invariant controls) will yield biased estimates. For most of the paper, I assume that this endogeneity is driven solely by time-invariant country fixed effects. Therefore, I estimate the model using fixed effects

probit, and use the bias correction of Hahn and Newey (2004) to mitigate the incidental parameters problem. Using data from Polity IV and the World Bank WDI, I find that having a democratic institution does not lead to a decrease in the probability of institutional change, once other relevant parameters are controlled for. A more consistent factor affecting the likelihood of institutional change is the interaction between the institution variable and the percentage of each political institution in the world. The finding suggests that being a democracy increases the stability of *other democracies* rather than has a direct positive effect on its own stability.

This paper finds that increasing the percentage of democracies by one standard deviation (10.3%) is expected to result in a 3.7 – 5.2% decrease in the likelihood of institutional change among fellow democracies. However, it is also expected to increase the likelihood of institutional change among autocracies by 12.0 – 14.5%. Surprisingly, increasing the percentage of autocracies by one standard deviation (16.5%) leads to a 7.4 – 9.2% *increase* in the probability of institutional change among autocracies on average, but has no significant impact among democracies. A higher percentage of urban population is associated with a lower likelihood of institutional change; however, economic variables such as GDP and the level of trade do not appear to have a significant impact on stability.

Finally, a test on the source of endogeneity of the institution variable suggests that even after controlling for time-invariant fixed effects, there are still time-variant unobservables correlated with the institution variable. Using two-stage least squares to obtain consistent estimates given the additional source of endogeneity, I find that being a democracy is associated with a lower likelihood of institutional change, in addition to its positive externality on the stability of other democracies as identified before. Given the difference in the significance of the institution variable across

model specifications, it would be helpful to verify the results in this paper using semi-parametric methods in a two-step estimation.

This paper is a first attempt at evaluating the effect of political institution on institutional stability. Admittedly the proxy for institution is crude due to data limitations; however, I believe this paper opens up a number of interesting avenues that can be explored in future work. First, it will be interesting to study whether there are specific aspects within the democratic or autocratic institution that impacts institutional stability, and compare the effect of each. Second, it will be useful to distinguish between the types of institutional change – whether it be due a coup or revolution, civil or foreign-based wars, conflict or peaceful transitions – and whether the change was dramatic or gradual. The causes of the different kinds of change are likely going to be different; separating these effects will give us a better understanding of the mechanics of the various kinds of institutional change. Third, it is instructive to fine-tune the impact of a country’s neighbors using geographic distances, and identify the factors captured by the percentages of democracies and autocracies. Is it the strategic decisions by foreign nations to interfere that cause institutional change, or is it the mere presence of more democracies that leads to higher acceptance of democratic values, and appreciation of the benefits and costs of the various systems?

## A Proof of Lemma 1

We will first prove that the truth-telling constraint (TC) is without loss of generality, then proceed to argue what each component of the menu contract must entail. TC ensures that the principal will implement the part of the menu that correctly correspond to the state ex post. For more details, see Appendix D.

When the agent sees the menu contract offered, for any  $\theta$ , one of the following four conditions must hold regarding  $(\mathcal{C}_i^{\theta H}, \mathcal{C}_i^{\theta L})$ :

- (i) TC is satisfied;
- (ii) a type  $\theta$  principal would prefer to implement  $\mathcal{C}_i^{\theta(-\theta)}$ ,  $\forall \theta \in \{H, L\}$ ;
- (iii) both types would prefer to implement  $\mathcal{C}_i^{\theta H}$ ; or
- (iv) both types would prefer to implement  $\mathcal{C}_i^{\theta L}$

The condition we want is (i), for (ii) the agent can simply switch the labels of the two components such that TC holds. We will discuss (iii), and (iv) will be analogous.

Since the agent already knows that only  $\mathcal{C}_i^{\theta H}$  will be implemented for all  $\theta$ , a menu contract with the implemented terms above duplicated for both states – i.e.  $(\mathcal{C}_i^{\theta H}, \mathcal{C}_i^{\theta H})$  will yield the agent the same payoff<sup>52</sup>. The principal will also obtain

---

<sup>52</sup>Since the feasibility constraints differ across the two states, there might be a problem if we need to replicate the  $\theta = H$  part of the menu for the  $\theta = L$  part of the alternative contract. This issue will arise when implementing contracts that do not satisfy TC as well. We assume if any principal promises a contract that turns out to violate FC (e.g. if the true state is  $L$  but the principal implements the  $\theta = H$  part of the menu), the principal's direct transfer ( $p$ ) will be adjusted such that FC holds. This is often the case in reality: if a company is in a binding contract to sell a product whose cost turns out to exceed the agreed-upon trading price, the company will have to take a loss and deliver as promised. In this sense, the alternative contract is still payoff-equivalent.

the same payoff under the duplicated menu contract that satisfies TC. The only condition that we must add is a consistency in beliefs requirement – that the agent should place the same beliefs on the original non-TC contract as the TC contract with duplicated terms. Since the same contract terms are implemented and both the principal and the agent get the same payoffs across these two contracts, there should not be any signaling value to using a TC versus a non-TC contract.

Since TC holds without loss of generality, our argument from here onward refers to specific contract terms for each part of the menu contract (i.e. what will be implemented in a given state).

## I. Tax Always at Upper Bound

Suppose for some  $i$ ,  $\theta$ ,  $\hat{\theta}$ , we have  $t_i^{\theta\hat{\theta}} < 1$ , and that fixing all other parts of this contract, the principal and the agent get  $\hat{U}_i^{\theta\hat{\theta}}$  and  $\hat{V}_i^{\theta}(U_i^{\theta\hat{\theta}})$  respectively, should this contract be accepted and  $C_i^{\theta\hat{\theta}}$  is implemented. Now consider an alternative contract in which all other contract terms are the same as the above, except now  $\tilde{t}_i^{\theta\hat{\theta}} = 1$ . Denote  $\delta = \tilde{t}_i^{\theta\hat{\theta}} - t_i^{\theta\hat{\theta}}$ . Depending on the state, use  $\delta$  to acquire additional amounts of the public good that the agent prefers. Since  $\gamma\eta > \alpha > 1$ , this leads to a higher  $\tilde{U}_i^{\theta\hat{\theta}}$  compared to  $\hat{U}_i^{\theta\hat{\theta}}$ , meaning that the probability of winning increases for  $Pi\theta$ .

In addition, if  $Pi\theta$ 's preference is aligned with the agent's for this state  $\hat{\theta}$ , then  $Pi\theta$ 's utility conditional on the contract being accepted is also higher than if  $t_i^{\theta\hat{\theta}} < 1$ . If  $Pi\theta$ 's preference is not aligned with the agent's, then her utility remains the same. This is true for any  $t_i^{\theta\hat{\theta}} < 1$ , for all  $\hat{\theta}$ ,  $\theta$ .

## II. All Other Contract Terms

$p_i^{\theta\hat{\theta}}$  is omitted because it is always chosen such that the feasibility constraint holds with equality. Standard arguments (similar to the one for  $t$ ) can be used to show that any contract in which the feasibility constraint has slack cannot be optimal, and can be improved by using an alternative contract in which feasibility binds.

Below we present the argument for the  $\theta = H$  part of the menu contract for P1, and similar arguments can be applied to the other cases.

Since the feasibility constraint must bind, we have  $p_1^{\theta H} = 1 - b_1^{\theta H} - g_1^{\theta H}$ . Substitute that into the principal's utility function,  $V_1^H = b_1^{\theta H} + \lambda p_1^{\theta H}$ , we get  $V_1^H = \lambda - (\lambda - 1)b_1^{\theta H} - \lambda g_1^{\theta H}$ . Also, the agent's utility is given by  $U_1^{\theta H} = \gamma [\eta b_1^{\theta H} + (1 - \eta)g_1^{\theta H}] - 1$ .

Suppose we need to attain  $\bar{U}$  for the agent. P1H can do so via one of three ways: (i) offer  $b_1^{\theta H}$  only; (ii) offer  $g_1^{\theta H}$  only; (iii) offer a mixture of the two public goods<sup>53</sup>. Given the linear structure of the model,  $b$  and  $g$  are perfect substitutes, and we know that generically (except at a unique point of indifference) method (iii) will not be optimal.

If P1H offers  $b_1^{\theta H}$  only, she needs to set  $b_1^{\theta H} = \frac{1}{\gamma\eta} (\bar{U} + 1)$ . This gives

$$V_1^H (\bar{U}) = \lambda - \frac{\lambda - 1}{\gamma\eta} (\bar{U} + 1)$$

If P1H offers  $g_1^{\theta H}$  only, she needs to set  $g_1^{\theta H} = \frac{1}{\gamma(1-\eta)} (\bar{U} + 1)$ . This gives

$$V_1^H (\bar{U}) = \lambda - \frac{\lambda}{\gamma(1-\eta)} (\bar{U} + 1)$$

---

<sup>53</sup>This is the standard constrained maximization problem:

$$\max_{(b_1^{\theta H}, g_1^{\theta H}) \in \mathbb{R}^2} V_1^{\theta H} \text{ s.t. } U_1^{\theta H} \geq \bar{U}$$

$\frac{\lambda}{\gamma(1-\eta)} > \frac{\lambda-1}{\gamma\eta}$  since  $\eta > \frac{1}{2}$ , and so  $P1H$  is always strictly better off offering  $b_1^{\theta H}$  only and set  $g_1^{\theta H} = 0$ .

Intuitively, since the preference of  $P1H$  aligns with that of the agent, there is no reason why she should offer a public good that neither she nor the agent prefers. This will of course be different if the principal's preference does not align with that of the agent, but the calculation of which public good to offer is similar. ■

## B Strategy and Outcome Representation

We use  $V_i^\theta$  to denote the utility of principal  $i$  whose contract is accepted, and who is implementing the  $\theta$  part of her menu contract.

The 1-to-1 functions are obtained by straightforward algebra using Lemma 1. The details are as follows:

- $\eta \in \left[ \frac{1}{2}, \frac{\lambda\alpha}{2\lambda\alpha-1} \right)$

$$\left\{ \begin{array}{l} V_1^H = \lambda - (\lambda - 1)b_1^{\theta H} \\ V_2^H = \lambda - (\lambda - 1)g_2^{\theta H} \\ V_1^L = \lambda - (\lambda\alpha - 1)b_1^{\theta L} \\ V_2^L = \lambda - (\lambda\alpha - 1)g_2^{\theta L} \end{array} \right. \quad \left\{ \begin{array}{l} V_1^H = \lambda - \frac{\lambda-1}{\gamma\eta} (1 + U_1^{\theta H}) \\ V_2^H = \lambda - \frac{\lambda-1}{\gamma(1-\eta)} (1 + U_2^{\theta H}) \\ V_1^L = \lambda - \frac{\lambda\alpha-1}{\gamma(1-\eta)} (1 + U_1^{\theta L}) \\ V_2^L = \lambda - \frac{\lambda\alpha-1}{\gamma\eta} (1 + U_2^{\theta L}) \end{array} \right.$$



- $\eta \in \left[ \frac{\lambda\alpha}{2\lambda\alpha-1}, \frac{\lambda}{2\lambda-1} \right)$

$$\left\{ \begin{array}{l} V_1^H = \lambda - (\lambda - 1)b_1^{\theta H} \\ V_2^H = \lambda - (\lambda - 1)g_2^{\theta H} \\ V_1^L = \lambda - \lambda\alpha g_1^{\theta L} \\ V_2^L = \lambda - (\lambda\alpha - 1)g_2^{\theta L} \end{array} \right. \quad \left\{ \begin{array}{l} V_1^H = \lambda - \frac{\lambda-1}{\gamma\eta} (1 + U_1^{\theta H}) \\ V_2^H = \lambda - \frac{\lambda-1}{\gamma(1-\eta)} (1 + U_2^{\theta H}) \\ V_1^L = \lambda - \frac{\lambda\alpha}{\gamma\eta} (1 + U_1^{\theta L}) \\ V_2^L = \lambda - \frac{\lambda\alpha-1}{\gamma\eta} (1 + U_2^{\theta L}) \end{array} \right.$$

- $\eta \in \left[ \frac{\lambda}{2\lambda-1}, 1 \right]$

$$\left\{ \begin{array}{l} V_1^H = \lambda - (\lambda - 1)b_1^{\theta H} \\ V_2^H = \lambda - \lambda b_2^{\theta H} \\ V_1^L = \lambda - \lambda\alpha g_1^{\theta L} \\ V_2^L = \lambda - (\lambda\alpha - 1)g_2^{\theta L} \end{array} \right. \quad \left\{ \begin{array}{l} V_1^H = \lambda - \frac{\lambda-1}{\gamma\eta} (1 + U_1^{\theta H}) \\ V_2^H = \lambda - \frac{\lambda}{\gamma\eta} (1 + U_2^{\theta H}) \\ V_1^L = \lambda - \frac{\lambda\alpha}{\gamma\eta} (1 + U_1^{\theta L}) \\ V_2^L = \lambda - \frac{\lambda\alpha-1}{\gamma\eta} (1 + U_2^{\theta L}) \end{array} \right.$$

## C Aggregate Welfare

The results follow directly from Lemma 1 and the feasibility constraint, by calculating the weighted sum of the winning principal and the agent's utility (from policy) for a given accepted contract using welfare weight  $\mu \in [0, 1]$ . Ex post welfare is given by  $\mathcal{W}_i^\theta(x; \mu) = \mu V_i^\theta(x) + (1 - \mu)x$ .

- $\eta \in \left[ \frac{1}{2}, \frac{\lambda\alpha}{2\lambda\alpha-1} \right)$

$$\begin{cases} \mathcal{W}_1^H = \mu \left( \lambda - \frac{\lambda-1}{\gamma\eta} \right) - U_1^{\theta H} \left[ \mu \left( 1 + \frac{\lambda-1}{\gamma\eta} \right) - 1 \right] \\ \mathcal{W}_2^H = \mu \left( \lambda - \frac{\lambda-1}{\gamma(1-\eta)} \right) - U_2^{\theta H} \left[ \mu \left( 1 + \frac{\lambda-1}{\gamma(1-\eta)} \right) - 1 \right] \\ \mathcal{W}_1^L = \mu \left( \lambda - \frac{\lambda\alpha-1}{\gamma(1-\eta)} \right) - U_1^{\theta L} \left[ \mu \left( 1 + \frac{\lambda\alpha-1}{\gamma(1-\eta)} \right) - 1 \right] \\ \mathcal{W}_2^L = \mu \left( \lambda - \frac{\lambda\alpha-1}{\gamma\eta} \right) - U_2^{\theta L} \left[ \mu \left( 1 + \frac{\lambda\alpha-1}{\gamma\eta} \right) - 1 \right] \end{cases}$$

- $\eta \in \left[ \frac{\lambda\alpha}{2\lambda\alpha-1}, \frac{\lambda}{2\lambda-1} \right)$

$$\begin{cases} \mathcal{W}_1^H = \mu \left( \lambda - \frac{\lambda-1}{\gamma\eta} \right) - U_1^{\theta H} \left[ \mu \left( 1 + \frac{\lambda-1}{\gamma\eta} \right) - 1 \right] \\ \mathcal{W}_2^H = \mu \left( \lambda - \frac{\lambda-1}{\gamma(1-\eta)} \right) - U_2^{\theta H} \left[ \mu \left( 1 + \frac{\lambda-1}{\gamma(1-\eta)} \right) - 1 \right] \\ \mathcal{W}_1^L = \mu \left( \lambda - \frac{\lambda\alpha}{\gamma\eta} \right) - U_1^{\theta L} \left[ \mu \left( 1 + \frac{\lambda\alpha}{\gamma\eta} \right) - 1 \right] \\ \mathcal{W}_2^L = \mu \left( \lambda - \frac{\lambda\alpha-1}{\gamma\eta} \right) - U_2^{\theta L} \left[ \mu \left( 1 + \frac{\lambda\alpha-1}{\gamma\eta} \right) - 1 \right] \end{cases}$$

- $\eta \in \left[ \frac{\lambda}{2\lambda-1}, 1 \right)$

$$\begin{cases} \mathcal{W}_1^H = \mu \left( \lambda - \frac{\lambda-1}{\gamma\eta} \right) - U_1^{\theta H} \left[ \mu \left( 1 + \frac{\lambda-1}{\gamma\eta} \right) - 1 \right] \\ \mathcal{W}_2^H = \mu \left( \lambda - \frac{\lambda}{\gamma\eta} \right) - U_2^{\theta H} \left[ \mu \left( 1 + \frac{\lambda}{\gamma\eta} \right) - 1 \right] \\ \mathcal{W}_1^L = \mu \left( \lambda - \frac{\lambda\alpha}{\gamma\eta} \right) - U_1^{\theta L} \left[ \mu \left( 1 + \frac{\lambda\alpha}{\gamma\eta} \right) - 1 \right] \\ \mathcal{W}_2^L = \mu \left( \lambda - \frac{\lambda\alpha-1}{\gamma\eta} \right) - U_2^{\theta L} \left[ \mu \left( 1 + \frac{\lambda\alpha-1}{\gamma\eta} \right) - 1 \right] \end{cases}$$

## D The Truth-Telling Constraint

Lemma 4 describes conditions under which TC will be satisfied. The cutoffs in  $\eta$  align with those established in Lemma 1, reflecting where the preference-misaligned principals switch from offering a public good of their preference to one which the

agent prefers.

**Lemma 4 (TC).** *A menu contract  $(U_i^{\theta H}, U_i^{\theta L})$  satisfies TC iff*

$$\begin{aligned} \frac{(\lambda-1)(1-\eta)}{(\lambda\alpha-1)\eta} (1 + U_1^{\theta H}) - 1 \leq U_1^{\theta L} \leq \frac{1-\eta}{\eta} (1 + U_1^{\theta H}) - 1 & \quad \text{if } \eta < \frac{\lambda\alpha}{2\lambda\alpha-1} \\ \frac{(\lambda-1)}{\lambda\alpha} (1 + U_1^{\theta H}) - 1 \leq U_1^{\theta L} \leq \frac{\lambda\alpha-1}{\lambda\alpha} (1 + U_1^{\theta H}) - 1 & \quad \text{if } \eta \geq \frac{\lambda\alpha}{2\lambda\alpha-1} \end{aligned}$$

$$\begin{aligned} \frac{(\lambda-1)\eta}{(\lambda\alpha-1)(1-\eta)} (1 + U_2^{\theta H}) - 1 \leq U_2^{\theta L} \leq \frac{\eta}{(1-\eta)} (1 + U_2^{\theta H}) - 1 & \quad \text{if } \eta < \frac{\lambda}{2\lambda-1} \\ \frac{\lambda}{\lambda\alpha-1} (1 + U_2^{\theta H}) - 1 \leq U_2^{\theta L} \leq \frac{\lambda\alpha}{\lambda\alpha-1} (1 + U_2^{\theta H}) - 1 & \quad \text{if } \eta \geq \frac{\lambda}{2\lambda-1} \end{aligned}$$

*Proof.* TC requires that  $Pi\theta$  always weakly prefers to implement  $\mathcal{C}_i^{\theta\theta}$  rather than  $\mathcal{C}_i^{\theta(-\theta)}$ . We will work through the conditions for  $P1$ , and the case of  $P2$  is analogous.

Recall from Lemma 1 that for all  $\theta$ ,  $\mathcal{C}_1^{\theta H}$  is always such that  $g_1^{\theta H} = 0$ . However,  $\mathcal{C}_1^{\theta L}$  will consist of different public good offers depending on the value of  $\eta$ . Therefore, TC for  $P1H$  will vary depending on the cutoff values of  $\eta$  that  $P1L$  has. We assume that if  $PiH$  implements  $\mathcal{C}_1^{HL}$ , she must carry out the contract terms as listed, so the feasibility constraint will not bind.

1.  $g_1^{\theta L} = 0$  (i.e.  $\eta \in \left[\frac{1}{2}, \frac{\lambda\alpha}{2\lambda\alpha-1}\right)$ )

To attain a level  $\bar{U}$  for the agent,  $P1H$  uses

When  $\mathcal{C}_1^{HH}$  is implemented

When  $\mathcal{C}_1^{HL}$  is implemented

$$\begin{aligned}
b_1^{HH} &= \frac{1}{\gamma\eta} (\bar{U} + 1) & b_1^{HL} &= \frac{1}{\gamma(1-\eta)} (\bar{U} + 1) \\
V_1^H &= \lambda - (\lambda - 1)b_1^{HH} & V_1^L &= \lambda - (\lambda\alpha - 1)b_1^{HH} \\
&= \lambda - \frac{\lambda-1}{\gamma\eta} (\bar{U} + 1) & &= \lambda - \frac{\lambda\alpha-1}{\gamma(1-\eta)} (\bar{U} + 1)
\end{aligned}$$

Therefore, for  $P1H$  not to implement  $\mathcal{C}_1^{HL}$ , we need

$$\begin{aligned}
\frac{\lambda-1}{\gamma\eta} (U_1^{\theta H} + 1) &\leq \frac{\lambda\alpha-1}{\gamma(1-\eta)} (U_1^{\theta L} + 1) \\
U_1^{\theta L} &\geq \frac{(\lambda-1)(1-\eta)}{(\lambda\alpha-1)\eta} (1 + U_1^{\theta H}) - 1
\end{aligned}$$

To attain a level  $\bar{U}$  for the agent,  $P1L$  uses

When  $\mathcal{C}_1^{LL}$  is implemented

$$\begin{aligned}
b_1^{LL} &= \frac{1}{\gamma(1-\eta)} (\bar{U} + 1) \\
V_1^L &= \lambda - (\lambda\alpha - 1)b_1^{LL} \\
&= \lambda - \frac{\lambda\alpha-1}{\gamma(1-\eta)} (\bar{U} + 1)
\end{aligned}$$

When  $\mathcal{C}_1^{LH}$  is implemented

$$\begin{aligned}
b_1^{LH} &= \frac{1}{\gamma\eta} (\bar{U} + 1) \\
V_1^L &= \lambda - (\lambda\alpha - 1)b_1^{LH} \\
&= \lambda - \frac{\lambda\alpha-1}{\gamma\eta} (\bar{U} + 1)
\end{aligned}$$

Therefore, for  $P1L$  not to implement  $\mathcal{C}_1^{LH}$ , we need

$$\begin{aligned}
\frac{\lambda\alpha-1}{\gamma(1-\eta)} (U_1^{\theta L} + 1) &\leq \frac{\lambda\alpha-1}{\gamma\eta} (U_1^{\theta H} + 1) \\
U_1^{\theta L} &\leq \frac{1-\eta}{\eta} (1 + U_1^{\theta H}) - 1
\end{aligned}$$

2.  $b_1^{\theta L} = 0$  (i.e.  $\eta \in \left[ \frac{\alpha}{2\alpha-\lambda}, 1 \right]$ )

With the same argument as above, for  $P1H$  to not implement  $\mathcal{C}_1^{HL}$ ,

$$\begin{aligned}\frac{\lambda-1}{\gamma\eta} (U_1^{\theta H} + 1) &\leq \frac{\lambda\alpha}{\gamma\eta} (U_1^{\theta L} + 1) \\ U_1^{\theta L} &\geq \frac{\lambda-1}{\lambda\alpha} (1 + U_1^{\theta H}) - 1\end{aligned}$$

For  $P1L$  to not implement  $\mathcal{C}_1^{LH}$ ,

$$\begin{aligned}\frac{\lambda\alpha}{\gamma\eta} (U_1^{\theta L} + 1) &\leq \frac{\lambda\alpha-1}{\gamma\eta} (U_1^{\theta H} + 1) \\ U_1^{\theta L} &\geq \frac{\lambda\alpha-1}{\lambda\alpha} (1 + U_1^{\theta H}) - 1\end{aligned}$$

These first condition (so that  $P1H$  follows TC) for each  $\eta$  range gives the lower bounds for  $U_1^{\theta L}$ , and the second condition (so that  $P1L$  follows TC) gives the upper bounds for  $U_1^{\theta L}$ . ■

## E Proof of Proposition 2

To prove Proposition 2, we will first find the set of equilibria in the case of purely valence-driven voter, then show that there exist conditions under which the set of pooling equilibria welfare dominates the set of separating equilibria.

Our strategy for equilibrium characterization is to first identify ranges of one's opponent's offer – for which deviations will be profitable given the agent's beliefs. Having identified these conditions, our characterization algorithm is just an iterated application of the linear programming method to rule out profitable deviations.

**Remark 1 (1P: Profitable Deviation).** *Fix a system of beliefs  $\pi$  and a putative equilibrium offer  $\hat{U}$ .  $P\theta$  has a profitable deviation against  $\hat{U}$  given  $\pi$  iff  $\exists (\tilde{U}^{\theta H}, \tilde{U}^{\theta L})$  s.t. IR-P, TC, and IR-A[ $\pi$ ] hold, and  $\tilde{U}^{\theta\theta} < \hat{U}^{\theta\theta}$ .*

*Proof.* Consider a set of beliefs – which includes the on- and off-path beliefs – and a proposed equilibrium  $\hat{U}$ . By construction  $\hat{U}$  must satisfy IR-P, TC, and IR-A[ $\pi$ ], where  $\pi$  is the on-path belief given  $\hat{U}$ . Any deviation contracts must also satisfy these three conditions. Any off-path contract that satisfies IR-P, TC, and IR-A[ $\tilde{\pi}$ ] will be accepted by the agent<sup>54</sup> (since there is only one principal), and thus constitutes the set of possible deviations - the question is whether any of them is profitable.

A deviation (within the above-mentioned set) is profitable for principal  $i$  if and only if this deviation yields her higher utility. Since TC holds, only the relevant part of the menu will be implemented, so the only part that matters to  $P\theta$  is  $\tilde{U}^{\theta\theta}$  and  $\hat{U}^{\theta\theta}$ . ■

### Algorithm for Equilibrium Characterization in a 1-Principal Model

Fix a system of beliefs, we can characterize the set of separating and pooling equilibria sustainable<sup>55</sup> in a one-principal model using the following algorithm:

1.  $\forall\theta, \mathcal{C}_\theta \equiv \{(U^{\theta H}, U^{\theta L}) : \text{IR-P \& TC hold}\} \setminus \{(U^{\theta H}, U^{\theta L}) : U^{\theta H} < 0 \cap U^{\theta L} < 0\}$
2. Define  $\mathcal{E}_\theta$  and  $\mathcal{D}_\theta$ <sup>56</sup> by restricting  $\mathcal{C}_\theta$  using IR-A[ $\pi$ ] and IR-A[ $\tilde{\pi}$ ] respectively
3. If  $\mathcal{D}_\theta \setminus \mathcal{E}_\theta \neq \emptyset$  for at least one  $\theta$ , then  $\nexists$  equilibrium using this system of beliefs.
4. If  $\mathcal{D}_\theta \setminus \mathcal{E}_\theta = \emptyset$  (i.e.  $\mathcal{D}_\theta \subseteq \mathcal{E}_\theta$ )  $\forall\theta$ , then the set of equilibria is given by:

$$\text{a Separating: } (U^{\theta H}, U^{\theta L}) \in \mathcal{E}_\theta \text{ s.t. } U^{\theta\theta} \leq \min \left\{ U^{(-\theta)\theta}, \arg \min_{U^{\theta\theta}} \mathcal{D}_\theta \right\} \quad \forall\theta$$

<sup>54</sup>Note that IR-A is dependent on the agent's beliefs, therefore it is calculated using the corresponding off-path beliefs when considering deviations.

<sup>55</sup>possibly empty for the given beliefs

<sup>56</sup>The  $\theta$  subscripts in  $\mathcal{C}_\theta$ ,  $\mathcal{E}_\theta$ , and  $\mathcal{D}_\theta$  are used simply to notationally differentiate  $(U^{HH}, U^{HL})$  and  $(U^{LH}, U^{LL})$ . Otherwise, these sets are identical across states.

b Pooling:  $(U^{\theta H}, U^{\theta L}) \in \mathcal{E}_\theta$  s.t.  $U^{\theta H} \leq \min_{U^{\theta H}} \mathcal{D}_\theta \cap U^{\theta L} \leq \min_{U^{\theta L}} \mathcal{D}_\theta \quad \forall \theta$

**Remark 2.** Fix a system of beliefs, the algorithm is necessary and sufficient for the characterization of the sets of separating and pooling equilibria.

*Proof.* We will prove each direction in turn.

## I. Necessity

We show that if the conditions laid in each step is violated, then there exists profitable deviation to the proposed set of equilibria. Steps 1 and 2 just define the relative sets / notation.

In step 3, suppose  $\mathcal{D}_\theta \setminus \mathcal{E}_\theta \neq \emptyset$  for the given set of beliefs. This set of beliefs can be offer-dependent; in that case, we use

$$\mathcal{D}_\theta = \left\{ (U^{\theta H}, U^{\theta L}) : \text{IR-P, TC, \& IR-A}[\tilde{\pi}(U)] \text{ hold} \right\}$$

. Notice here  $\mathcal{D}_\theta \setminus \mathcal{E}_\theta \neq \emptyset$  necessarily implies  $\forall U^\theta \in \mathcal{E}_\theta, \exists \tilde{U}^\theta \in [\mathcal{D}_\theta \setminus \mathcal{E}_\theta]$  such that at least one of the following strict inequalities holds:

$$\tilde{U}^{\theta H} < U^{\theta H} \quad \text{or} \quad \tilde{U}^{\theta L} < U^{\theta L}$$

This means that the set  $[\mathcal{D}_\theta \setminus \mathcal{E}_\theta]$  must lie southwest of  $\mathcal{E}^\theta$ . One can most easily see this graphically. Since  $\mathcal{E}_\theta$  is bounded by IR-P, TC, & IR-A $[\pi]$ , all points northeast of IR-A $[\pi]$  within  $\mathcal{C}_\theta$  belongs to  $\mathcal{E}_\theta$ . Since points in  $\mathcal{D}_\theta$  must also satisfy IR-P and TC, it is in  $\mathcal{C}_\theta$  as well. For  $\mathcal{D}_\theta \setminus \mathcal{E}_\theta \neq \emptyset$ ,  $\tilde{U}^\theta \in [\mathcal{D}_\theta \setminus \mathcal{E}_\theta]$  must be strictly to the southwest of IR-A $[\pi]$ , and therefore to the southwest of  $\mathcal{E}_\theta$  as well.

For each  $U^\theta \in \mathcal{E}_\theta$ , denote as  $\tilde{\theta}$  the type where  $\tilde{U}^{\theta\tilde{\theta}} < U^{\theta\tilde{\theta}}$  for some  $\tilde{U}^\theta \in [\mathcal{D}_\theta \setminus \mathcal{E}_\theta]$ . Then  $P\tilde{\theta}$  will find it profitable to offer  $\tilde{U}^{\theta\tilde{\theta}}$  instead of the proposed equilibrium  $U^{\tilde{\theta}}$ , meaning that  $U^{\tilde{\theta}}$  cannot be an equilibrium offer.

Since we can do so for all  $U^\theta \in \mathcal{E}_\theta$ , there does not exist an equilibrium *given this system of beliefs* if  $\mathcal{D}_\theta \setminus \mathcal{E}_\theta \neq \emptyset$ .

The argument for step 4 is similar to that of step 3, but we will construct deviations for each class of equilibria supposing that the conditions laid out are violated.

First, consider the set of separating equilibria  $U$ , and suppose by way of contradiction that  $\exists \tilde{U}^{\theta\theta} \in \mathcal{D}_\theta$  such that  $U^{\theta\theta} > \tilde{U}^{\theta\theta}$ . By definition of  $\mathcal{D}_\theta$ ,  $\tilde{U}^{\theta\theta}$  satisfies IR-P, TC, and IR-A $[\tilde{\pi}]$ . Since  $V^\theta(U^{\theta\theta})$  is decreasing in  $U^{\theta\theta}$ ,  $P\theta$  will be *strictly* better off offering  $\tilde{U}^{\theta\theta}$  instead of the proposed equilibrium  $U^{\theta\theta}$ . This means that  $U^{\theta\theta}$  cannot be an equilibrium.

Alternatively, suppose that  $U^{\theta\theta} > U^{(-\theta)\theta}$  in a separating equilibrium. Since TC holds,  $P\theta$  will always implement the  $\theta$  part of a menu contract. By definition, the menu contract  $U^{-\theta}$  satisfies IR-P, TC, and IR-A $[\pi]$ , and therefore it will be accepted by the agent. If  $U^{\theta\theta} > U^{(-\theta)\theta}$ , then  $P\theta$  will be strictly better off offering  $U^{-\theta}$ , and thus the originally proposed equilibrium  $U$ , where  $U^\theta \neq U^{-\theta}$  because it is a separating equilibrium, cannot be an equilibrium.

For the set of pooling equilibrium, suppose that  $\exists \hat{\theta}$  such that  $U^{\theta\hat{\theta}} > \tilde{U}^{\theta\hat{\theta}}$  for some  $\tilde{U}^{\theta\hat{\theta}} \in \mathcal{D}_\theta$ . If  $P\hat{\theta}$  offers such  $\tilde{U}^{\theta\hat{\theta}}$ , she can be guaranteed utility higher than under the proposed equilibrium  $U$ , which means that again  $U$  cannot be an equilibrium.



## II. Sufficiency

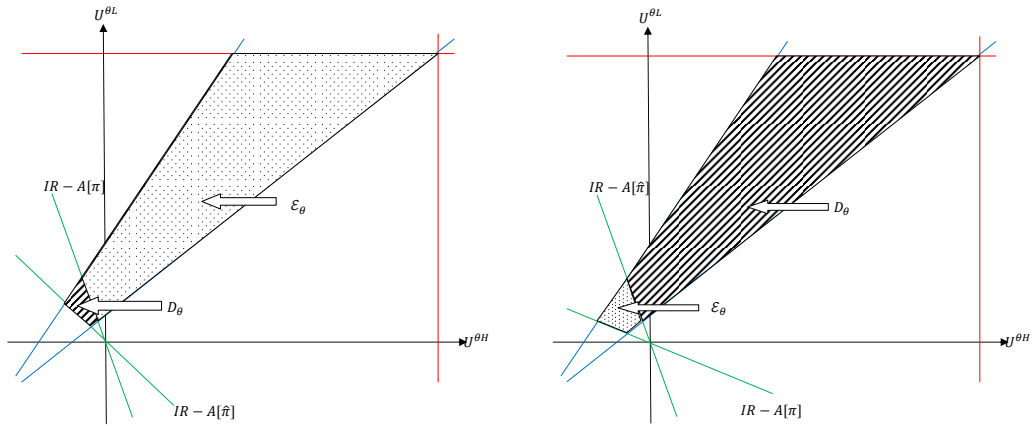
Suppose the conditions laid out in the algorithm is satisfied. We will argue that each contract offer in the proposed classes of equilibria indeed constitutes a Perfect Bayesian Equilibrium. Beliefs are not described because on-path beliefs are given by Bayes' Rule, and off-path beliefs are exogenously given outside of the algorithm.

Fix a proposed equilibrium  $U$  as described in step 4 of the algorithm, and consider any possible deviation by  $P\theta$  that satisfies IR-P, TC, and IR-A $[\bar{\pi}]$  given the system of beliefs. By construction, the set that satisfies the above three constraints is denoted  $\mathcal{D}_\theta$ . Since TC is satisfied,  $U^{\theta\theta}$  is the part of the contract that will be implemented should the contract be accepted, so that is the only payoff-relevant element in the contract for the agent and for all  $P\theta$ .

First we look at the class of separating equilibria. By step 4a,  $U^{\theta\theta} \leq \tilde{U}^{\theta\theta} \vee \tilde{U}^{\theta\theta} \in \mathcal{D}_\theta$ . Since  $V^\theta(U^{\theta\theta})$  is decreasing in  $U^{\theta\theta}$ , this means that  $P\theta$ 's utility is highest under the proposed equilibrium (relative to the entire set of feasible deviations). Moreover, given TC and  $U^{\theta\theta} \leq U^{(-\theta)\theta}$ ,  $P\theta$  also does not prefer to offer  $P(-\theta)$ 's equilibrium contract either. Together with the agent's beliefs, this constitutes a Perfect Bayesian Equilibrium.

For the set of pooling equilibria, both types of the principal offers the same menu contract. Therefore, a pooling equilibrium requires that *neither* type of principal has an incentive to deviate from the proposed equilibrium. For any pooling equilibrium  $U$  given by the algorithm, by step 4b,  $U^{\theta\hat{\theta}} \leq \tilde{U}^{\theta\hat{\theta}} \vee \tilde{U}^{\theta\hat{\theta}} \in \mathcal{D}_{\hat{\theta}}$ , for  $\hat{\theta} = H, L$ . This means that neither type will find it profitable to deviation by any other feasible contracts. Together with the agent's beliefs, this constitutes a Perfect Bayesian Equilibrium. ■

Figure 3 is a demonstration of how steps 1-3 of the algorithm is applied. For simplicity of illustration and notation when outlining the various algorithms, we limit ourselves to describing what are essentially deviation-independent beliefs throughout the paper. The algorithms can be easily extended to any class of beliefs, at the expense of introducing more notation. Moreover, the characterization results are proved for any class of beliefs that the agent might have.



(a) No equilibrium exists for given beliefs. (b) Equilibrium exists for given beliefs.

Figure 3: **Equilibrium Characterization Algorithm Illustration.**

The algorithm introduced is a general method to identify the set of sustainable equilibria (if any) given a system of beliefs. However, the algorithm itself does not provide a complete direct characterization of the equilibrium set, since there is a continuum of possible off-path beliefs to be considered. Our goal is to identify the set of beliefs that sustains the largest set of equilibria.

**Remark 3 (Purely Valence-Driven Voter: Minimal Deviation Set).** *The*

minimal set  $\mathcal{D}_\theta$  for any out-of-equilibrium beliefs is

$$\mathcal{D}_\theta^{min} \equiv \left\{ (U^{\theta H}, U^{\theta L}) : IR-P \ \& \ TC \ hold \right\} \cap \mathbb{R}_+^2$$

Remark 3 says that we cannot shrink the set of potential profitable deviation set to anything smaller than  $\mathcal{D}_\theta^{min}$ . The argument is quite straightforward. First, notice there do not exist any beliefs such that a contract with  $U^{\theta\hat{\theta}} \geq 0 \ \forall \theta, \hat{\theta}$  will not be accepted by the agent. Therefore, any contract within that set can potentially be a deviation, so long as the principal finds it profitable. If there is a type whose putative on-path offer renders herself lower than what she can get by offering some  $\tilde{U} \in \mathcal{D}_\theta^{min}$ , then that putative on-path offer cannot be part of the equilibrium.

An example of a system of beliefs that yields  $\mathcal{D}_\theta^{min}$  is “upon observing  $U^*$ , beliefs are as prescribed in a separating or pooling equilibrium. For any other offers observed, beliefs are such that if any offer  $\tilde{U}_i^{\theta\hat{\theta}} < 0$  for some  $\hat{\theta}$ , then  $prob(\theta = \hat{\theta}) = 1$ .”<sup>57</sup>

From the algorithm, it should be clear that the smaller the set  $\mathcal{D}_\theta$ , the larger the set of sustainable equilibria. This is because the restrictions in steps 3 and 4 are both relaxed if the  $\mathcal{D}_\theta$  considered is smaller. That means the largest possible set of equilibria is supported by the out-of-equilibrium beliefs that result in  $\mathcal{D}_\theta^{min}$ .

Using Remark 3 and the algorithm, we can characterize the largest possible sets of separating and pooling equilibria when there are two informed politicians and voters are purely valence-driven. Figure 4 is a graphical description of these equilibria.

---

<sup>57</sup>Clearly if  $\tilde{U}_i^{\theta\hat{\theta}} < 0 \ \forall \hat{\theta}$ , then this offer will be rejected regardless of beliefs. Also, since we are considering unilateral deviations only, we ignore the possibility of contradicting posterior beliefs, for instance, if  $\tilde{U}_i^{\theta\hat{\theta}} < 0$  and  $\tilde{U}_{-i}^{\theta(-\hat{\theta})} < 0$ .

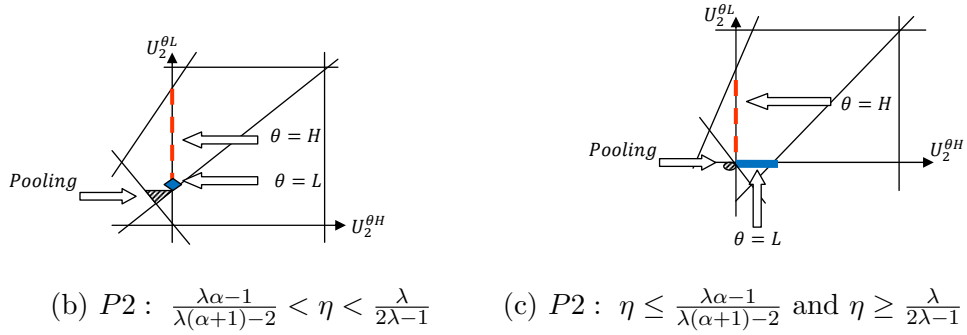
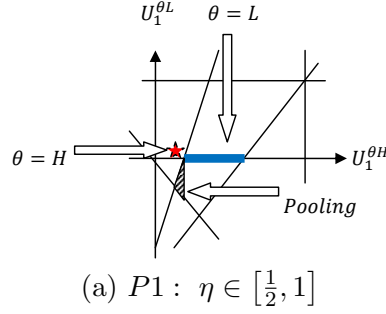


Figure 4: [**2-P Purely Valence-Driven Voter**]. Shaded: Pooling equilibrium ; dotted line: Separating Equilibrium  $\theta = H$  ; thick line: Separating Equilibrium  $\theta = L$ .

We can see that for all  $i, \theta$ ,  $\inf U_i^{S\theta} \geq \sup U_i^{P\theta}$ . Using Appendix C, we can see  $\forall i, \theta$  and  $\mu \in (0, 1]$ ,  $\exists (\gamma, \lambda, \alpha, \eta)$  such that  $\frac{\partial \mathcal{W}_i^\theta(x)}{\partial x} < 0$ . The requirement is that as  $\mu \rightarrow 0$ ,  $\lambda \rightarrow \infty$  (at a faster rate than the former) so that  $\frac{\partial \mathcal{W}_i^\theta(x)}{\partial x} < 0$  still holds. ■

## F Proof of Lemma 2

In the case of purely valence-driven voter,  $(U^*, \pi^*, \tilde{\pi}^*)$  is an equilibrium if and only if

$$Pr(i \text{ wins} | U^{\theta*}) V_i^\theta(U_i^{\theta\theta*}) \geq Pr(i \text{ wins} | \tilde{U}_i^\theta, U_{-i}^{\theta*}) V_i^\theta(\tilde{U}_i^{\theta\theta}) \quad \forall \tilde{U}_i^\theta \in \mathcal{U}_i^\theta, \quad \forall \theta, i \quad (\text{F-1})$$

and on-path beliefs  $\pi^\theta$  are correctly derived from Bayes' Rule. There are no restrictions on off-path beliefs  $\tilde{\pi}^*$  except that (F-1) must hold given  $\pi^*$  and  $\tilde{\pi}^*$ . In addition, since there is no competition in the policy dimension,  $Pr(i \text{ wins} | U^\theta)$  is a constant for all  $U^\theta$ . Therefore, (F-1) becomes

$$V_i^\theta(U_i^{\theta\theta*}) \geq V_i^\theta(\tilde{U}_i^{\theta\theta}) \quad \forall \tilde{U}_i^\theta \in \mathcal{U}_i^\theta, \quad \forall \theta, i \quad (\text{F-2})$$

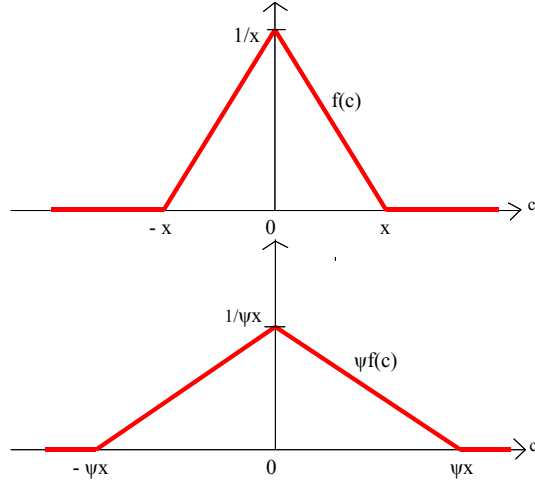


Figure 5: Densities of  $F(\cdot)$  and  $\hat{F}(\cdot; \psi)$  for  $\psi > 1$ .

Fix a continuous distribution  $F(c)$  over a compact support that has density  $f(c)$

and zero expected value. Let  $\hat{F}(c; \psi) \equiv F(c/\psi)$ ;  $\hat{F}(c; \psi)$  is the distribution for  $\psi c$  and a transformation of the original distribution incorporating  $\psi$ , with a corresponding density  $\hat{f}(\cdot; \psi)$ . A graphical example of the densities of  $F(\cdot)$  and  $\hat{F}(\cdot; \psi)$  is given in Figure 5.

Let the subset of the support of  $\psi c$  in which there is policy competition be  $[\underline{z}, \bar{z}]$ . This means

$$\underline{z} \equiv \left[ \min_{U_1^{\theta\hat{\theta}} \in \mathcal{U}_1^{\theta\hat{\theta}}; \theta, \hat{\theta} \in \{H, L\}} U_1^{\theta\hat{\theta}} - \max_{U_2^{\theta\hat{\theta}} \in \mathcal{U}_2^{\theta\hat{\theta}}; \theta, \hat{\theta} \in \{H, L\}} U_2^{\theta\hat{\theta}} \right]$$

and

$$\bar{z} \equiv \left[ \max_{U_1^{\theta\hat{\theta}} \in \mathcal{U}_1^{\theta\hat{\theta}}; \theta, \hat{\theta} \in \{H, L\}} U_1^{\theta\hat{\theta}} - \min_{U_2^{\theta\hat{\theta}} \in \mathcal{U}_2^{\theta\hat{\theta}}; \theta, \hat{\theta} \in \{H, L\}} U_2^{\theta\hat{\theta}} \right]$$

Notice that  $\underline{z}$  and  $\bar{z}$  do not depend on the valence component ( $\psi$  or  $c$ ). For  $\psi c < \underline{z}$ ,  $P1$  wins regardless of the policies offered (as long as IR-P, IR-A, and TC are satisfied); for  $\psi c > \bar{z}$ ,  $P2$  wins regardless of policies offered.

Given  $(\psi, \gamma, \lambda, \alpha, \eta)$ ,  $(U^\diamond, \pi^\diamond, \tilde{\pi}^\diamond)$  is an equilibrium of the model if and only if

$$\begin{aligned} & \mathfrak{F}_i(\underline{z}, \bar{z}; \psi) V_i^\theta \left( U_i^{\theta\theta\diamond} \right) \\ & + \left[ \hat{F}(\bar{z}; \psi) - \hat{F}(\underline{z}; \psi) \right] Pr(i \text{ wins} \mid \psi c \in [\underline{z}, \bar{z}], U^{\theta\diamond}) V_i^\theta \left( U_i^{\theta\theta\diamond} \right) \\ & \geq \mathfrak{F}_i(\underline{z}, \bar{z}; \psi) V_i^\theta \left( \hat{U}_i^{\theta\theta} \right) \\ & + \left[ \hat{F}(\bar{z}; \psi) - \hat{F}(\underline{z}; \psi) \right] Pr(i \text{ wins} \mid \psi c \in [\underline{z}, \bar{z}], \hat{U}_i^\theta, U_{-i}^{\theta\diamond}) V_i^\theta \left( \hat{U}_i^{\theta\theta} \right) \end{aligned}$$

for all  $\hat{U}_i^\theta \in \mathcal{U}_i^\theta$  and for all  $\theta, i$ , where  $\mathfrak{F}_1 = \hat{F}(\underline{z}; \psi)$  and  $\mathfrak{F}_2 = [1 - \hat{F}(\bar{z}; \psi)]$ . The requirements on beliefs are as in the case of purely valence-driven voter. We can

rearrange the above inequality and get

$$\begin{aligned}
V_i^\theta(U_i^{\theta\theta\Diamond}) - V_i^\theta(\hat{U}_i^{\theta\theta}) &\geq \frac{[\hat{F}(\bar{z};\psi) - \hat{F}(\underline{z};\psi)]}{\mathfrak{F}_i(\underline{z}, \bar{z}; \psi)} \\
&\left[ Pr(i \text{ wins} \mid \psi \in [\underline{z}, \bar{z}], \hat{U}_i^\theta, U_{-i}^{\theta\Diamond}) V_i^\theta(\hat{U}_i^{\theta\theta}) \right. \\
&\quad \left. - Pr(i \text{ wins} \mid \psi \in [\underline{z}, \bar{z}], U^{\theta\Diamond}) V_i^\theta(U_i^{\theta\theta\Diamond}) \right] \quad (\text{F-3})
\end{aligned}$$

for all  $\hat{U}_i^\theta \in \mathcal{U}_i^\theta$  and for all  $i$ .

Since  $\underline{z}$  and  $\bar{z}$  are both finite,  $\lim_{\psi \rightarrow \infty} [\hat{F}(\bar{z}; \psi) - \hat{F}(\underline{z}; \psi)] = 0$ , which means that the condition (F-3) will be the same as (F-2) in the limit. Therefore, the set of equilibria for the case of purely valence-driven voter coincides with that in the model for the limiting case of  $\psi$ . ■

## G Proof of Lemma 3

We would like to prove that the set of equilibrium correspondence is upper hemicontinuous in the parameter  $\psi$ . Let  $\psi^m$  be a sequence such that  $\psi^m \rightarrow \psi$ , and suppose  $(U^{*m}, \pi^{*m}, \tilde{\pi}^m)$  is a corresponding equilibrium given  $\psi^m$  for all  $m$ , with each of its elements converging to  $U^*$ ,  $\pi^*$ , and  $\tilde{\pi}$  respectively. We will show that  $(U^*, \pi^*, \tilde{\pi})$  is an equilibrium given  $\psi$ .

Suppose by way of contradiction that  $(U^*, \pi^*, \tilde{\pi})$  is *not* an equilibrium given  $\psi$ . This means one of two things (or both): either at least one principal is not best responding given her opponent's strategies (fixing the agent's beliefs), or the agent's beliefs on equilibrium path are not derived using Bayes' Rule.

A Perfect Bayesian Equilibrium imposes no restrictions on off-path beliefs, what matters is that given these off-path beliefs, the strategies proposed do satisfy equilibrium conditions (no deviations). On-path beliefs in our model that are consistent

with Bayes' Rule can only take on one of three values:  $\pi^* \in \{0, 1, \pi_0\}$  (where  $\pi_0$  is the prior belief). In the limit, they must also take on one of these values (since every element in any sequence – an equilibrium – takes on one of them). Moreover, in order to converge to any one of the three values in the limit, it must be a constant sequence far enough along the sequence. Therefore, the on-path belief in the limit must also be correct and consistent with Bayes' Rule.

Given that the latter condition (regarding on-path beliefs) holds, if  $(U^*, \pi^*, \tilde{\pi})$  is not an equilibrium, it must mean that  $\exists i, \theta$  and  $\tilde{U}_i^\theta$  such that

$$Pr\left(i \text{ wins} | \tilde{U}_i^\theta, U_{-i}^{\theta*}\right) V_i^\theta\left(\tilde{U}_i^{\theta\theta}\right) > Pr\left(i \text{ wins} | U^{\theta*}\right) V_i^\theta\left(U_i^{\theta\theta*}\right) \quad (\text{G-1})$$

where  $Pr(i \text{ wins} | U^\theta) = F(U_1^{\theta\theta} - U_2^{\theta\theta})$  if  $i = 1$ , and  $Pr(i \text{ wins} | U^\theta) = 1 - F(U_1^{\theta\theta} - U_2^{\theta\theta})$  if  $i = 2$ .<sup>58</sup> Since  $F(c)$  is continuous in  $c$  and almost everywhere convergent with respect to  $\psi$ , the principals' payoffs are continuous in  $\psi$  and  $U_i^{\theta\theta} \forall i, \theta$ . Therefore, (G-1) implies that must exist some  $\psi^{\bar{m}}$  in the sequence converging to  $\psi$  such that for the same  $i$  and  $\theta$  used in (G-1),  $\exists \hat{U}_i^\theta$  with which

$$Pr\left(i \text{ wins} | \hat{U}_i^\theta, \left(U_{-i}^\theta\right)^{* \bar{m}}\right) V_i^\theta\left(\bar{U}_i^\theta\right) > Pr\left(i \text{ wins} | \left(U^\theta\right)^{* \bar{m}}\right) V_i^\theta\left(\left(U_i^{\theta\theta}\right)^{* \bar{m}}\right)$$

This contradicts the fact that  $(U^{*\bar{m}}, \pi^{*\bar{m}}, \tilde{\pi}^{\bar{m}})$  is an equilibrium given  $\psi^{\bar{m}}$ . ■

---

<sup>58</sup>This is written for a separating equilibrium; the pooling equilibrium counterpart is analogous, but with expected utility of the agent instead.



## H Proof of Proposition 3

The sets of separating and pooling equilibria satisfying the Intuitive Criterion in the case of purely valence-driven voter are given below; the case of pure policy is available from the author upon request.

- Separating Equilibrium

Same as original set of separating equilibrium. See Figure 4.

- Pooling Equilibrium [ $P1$  wins]

Let the set of Pooling equilibria as given in Figure 4 be  $\mathcal{E}_1^{Pool*}$ .

The set of Pooling equilibria given the Intuitive Criterion is

$$\left\{ \left( \hat{U}_1^{\theta H}, \hat{U}_1^{\theta L} \right) \in \mathcal{E}_1^{Pool*} : \hat{U}_1^{\theta H} = \left\{ \arg \min_{U_1^{\theta H}} \mathcal{E}_1^{Pool*} \text{ s.t. } U_1^{\theta L} = \hat{U}_1^{\theta L} \right\} \right\}$$

- Pooling Equilibrium [ $P2$  wins]

Let the set of Pooling equilibria as given in Figure 4 be  $\mathcal{E}_2^{Pool*}$ .

The set of Pooling equilibria given the Intuitive Criterion is

$$\left\{ \left( \hat{U}_2^{\theta H}, \hat{U}_2^{\theta L} \right) \in \mathcal{E}_2^{Pool*} : \hat{U}_2^{\theta L} = \left\{ \arg \min_{U_2^{\theta L}} \mathcal{E}_2^{Pool*} \text{ s.t. } U_2^{\theta H} = \hat{U}_2^{\theta H} \right\} \right\}$$

For both the separating and pooling equilibria, we will argue the case of  $P1$ , and the case of  $P2$  is symmetric.

### i) Separating Equilibria

For  $\theta = H$ , since TC binds, in order to lower  $U_1^{\theta H}$  so that such deviation may be profitable for  $P1H$ ,  $U_1^{\theta L}$  must be lowered as well. This means that the Intuitive Criterion has no bite – there are beliefs (e.g.  $\tilde{\pi}_1 = 1$ ) such that both types will find it profitable to deviate – and any off-path beliefs can be used. For  $\theta = L$ , if the

deviation is profitable for  $P1L$ , either it is only profitable for  $P1L$  and not  $P1H$  (i.e. by the Intuitive Criterion the only possible off-path belief for that deviation is  $\tilde{\pi}_1 = 0$ , and  $P1L$  will lose with probability 1), or there exists beliefs such that it may be profitable for both types, and the Intuitive Criterion has no bite. In either case, there does not exist any deviation that can break the proposed set of separating equilibria.

## ii) Pooling Equilibria

For P1, the set of pooling equilibria that satisfies the Intuitive Criterion is the “<” shaped line that borders set of pooling equilibria in Figure 4. Formally, the set of equilibria is

$$\left\{ \left( \hat{U}_1^{\theta H}, \hat{U}_1^{\theta L} \right) \in \mathcal{E}_1^{Pool*} : \hat{U}_1^{\theta H} = \left\{ \arg \min_{U_1^{\theta H}} \mathcal{E}_1^{Pool*} \text{ s.t. } U_1^{\theta L} = \hat{U}_1^{\theta L} \right\} \right\}$$

First, notice that any

$$\left\{ \left( \hat{U}_1^{\theta H}, \hat{U}_1^{\theta L} \right) \in \mathcal{E}_1^{Pool*} : \hat{U}_1^{\theta H} < \left\{ \arg \min_{U_1^{\theta H}} \mathcal{E}_1^{Pool*} \text{ s.t. } U_1^{\theta L} = \hat{U}_1^{\theta L} \right\} \right\}$$

is simply not in  $\mathcal{U}_1^\theta$  (graphically, it is the area to the right of the “<” shaped borders in Figure 4). Now consider any

$$\left\{ \left( \hat{U}_1^{\theta H}, \hat{U}_1^{\theta L} \right) \in \mathcal{E}_1^{Pool*} : \hat{U}_1^{\theta H} > \left\{ \arg \min_{U_1^{\theta H}} \mathcal{E}_1^{Pool*} \text{ s.t. } U_1^{\theta L} = \hat{U}_1^{\theta L} \right\} \right\}$$

This is the pooling equilibria that is not part of the “<” shaped borders. It will be

profitable for  $P1H$  to deviate with

$$\left\{ \left( \tilde{U}_1^{\theta H}, \tilde{U}_1^{\theta L} \right) \in \mathcal{E}_1^{Pool*} : \tilde{U}_1^{\theta L} = \hat{U}_1^{\theta L} \text{ and } \tilde{U}_1^{\theta H} = \left\{ \arg \min_{U_1^{\theta H}} \mathcal{E}_1^{Pool*} \text{ s.t. } U_1^{\theta L} = \tilde{U}_1^{\theta L} \right\} \right\}$$

if the agent's belief given this deviation is  $\tilde{\pi}_1 = 1$  (since  $P1H$  still wins with probability 1 and only needs to offer a lower  $U_1^{HH}$ ). We will show that this deviation is never profitable for  $P1L$  regardless of the agent's beliefs:

1. Suppose  $P1L$  deviates with the above, and the agent's belief upon observing this deviation is such that this contract will not be accepted. Then clearly  $P1L$  is strictly worse off. If  $P1L$  does not win in the original putative equilibrium, then she is indifferent.
2. Suppose  $P1L$  deviates with the above, and the agent's belief upon observing this deviation is such that this contract will be accepted. Even then,  $P1L$  is indifferent between deviating or not.

Therefore, for any equilibrium in which

$$\left\{ \left( \hat{U}_1^{\theta H}, \hat{U}_1^{\theta L} \right) \in \mathcal{E}_1^{Pool*} : \hat{U}_1^{\theta H} > \left\{ \arg \min_{U_1^{\theta H}} \mathcal{E}_1^{Pool*} \text{ s.t. } U_1^{\theta L} = \hat{U}_1^{\theta L} \right\} \right\}$$

$P1H$  can break the equilibrium by deviating as above. By the Intuitive Criterion, upon observing this deviation, the agent must assign  $\tilde{\pi}_1 = 1$ , hence allowing  $P1H$  to win with probability 1 and obtain a higher payoff.

Summarizing the above, we can see that the welfare comparison remains qualitatively the same ( $\inf \mathcal{W}^{P\theta} \geq \sup \mathcal{W}^{S\theta}$ ) for the sets of separating and pooling equilibria after applying the Intuitive Criterion. Moreover, both sets of equilibria remain non-empty. ■

## I Description of Data

Table 1: Variables of interest from Polity IV & World Bank WDI.

Source	Variable	Notes
Polity IV	<i>polity2</i> <i>durable</i>	Integer ranges from -10 (autocracy) to 10 (democracy) No. of years since last change of $\geq 3$ in <i>polity2</i> over three years or less
WDI	<i>gdp05</i> <i>urban</i> <i>trade</i>	GDP per capita, PPP (constant 2005 international \$) Urban Population (% of total) Trade (% of GDP)
Generated	<i>change</i> <i>d5</i> <i>d0</i> <i>a5</i> <i>ld5</i>	= 1 if <i>polity2</i> change from last period is 3 points or more = 1 if <i>polity2</i> > 5 ; 0 otherwise = 1 if <i>polity2</i> > 0 ; 0 otherwise = 1 if <i>polity2</i> < -5 ; 0 otherwise One-period lag of the variable <i>d5</i>
	<i>persist</i>	No. of years (as of last year) since last change of $\geq 3$ in <i>polity2</i> over three years or less; it is also a one-period lag of variable <i>durable</i>
	<i>pcd5</i> <i>pcd0</i> <i>pca5</i>	% of countries (for full sample) with <i>d5</i> = 1 % of countries (for full sample) with <i>d0</i> = 1 % of countries (for full sample) with <i>a5</i> = 1
	<i>d5*pcd5</i> <i>d0*pcd0</i> <i>a5*pca5</i> <i>d5*pca5</i> <i>a5*pcd5</i>	Interaction between <i>d5</i> and <i>pcd5</i> Interaction between <i>d0</i> and <i>pcd0</i> Interaction between <i>a5</i> and <i>pca5</i> Interaction between <i>d5</i> and <i>pca5</i> Interaction between <i>a5</i> and <i>pcd5</i>

Table 2: Countries with Complete Data that Experienced Institutional Change between 1980-2007.

Albania	Iran
Algeria	Ivory Coast
Argentina	Jordan
Bangladesh	Kenya
Bhutan	Korea South
Bolivia	Lesotho
Brazil	Madagascar
Bulgaria	Malawi
Cameroon	Mali
Central African Republic	Mauritania
Chad	Mexico
Chile	Mozambique
Comoros	Nepal
Congo Brazzaville	Nicaragua
Congo Kinshasa	Nigeria
Dominican Republic	Pakistan
Ecuador	Panama
Egypt	Paraguay
Fiji	Peru
Gabon	Philippines
Gambia	Senegal
Ghana	Sierra Leone
Guatemala	Sudan
Guinea	Thailand
Guinea-Bissau	Tunisia
Honduras	Turkey
Hungary	Uruguay
Indonesia	Zambia

Table 3: Summary Statistics.

	Original Set (those with data) (1980-2007)	Sample Estimated (56 countries) (1980-2007)
<i>polity2</i>	1.42 (7.34)	0.379 (6.64)
<i>durable</i>	23.37 (29.34)	11.57 (13.66)
<i>gdp05</i>	9290.60 (11222)	3865.48 (3717.37)
<i>urban</i>	52.68 (24.69)	45.38 (20.21)
<i>trade</i>	82.66 (47.72)	65.45 (32.67)
<i>change</i>	0.046 (0.210)	0.080 (0.271)
<i>d5</i>	0.451 (0.500)	0.372 (0.484)
<i>d0</i>	0.540 (0.498)	0.482 (0.500)
<i>a5</i>	0.319 (0.466)	0.318 (0.466)
<i>persist</i>	23.20 (29.20)	11.62 (13.79)
<i>pcd5</i>	0.439 (0.103)	0.439 (0.103)
<i>pcd0</i>	0.524 (0.128)	0.524 (0.128)
<i>pca5</i>	0.334 (0.165)	0.334 (0.165)
no. obs.	varies by var.	1568

Mean is listed for each variable; standard deviation in parentheses.

Table 4: Correlations among Continuous Variables.

	<i>persist</i>	<i>gdp05</i>	<i>urban</i>	<i>trade</i>	<i>pcd5</i>	<i>pcd0</i>	<i>pca5</i>
<i>persist</i>	1						
<i>gdp05</i>	-0.0073	1					
<i>urban</i>	-0.1092	0.7723	1				
<i>trade</i>	0.0869	0.1096	0.0883	1			
<i>pcd5</i>	-0.1107	0.1093	0.1951	0.2345	1		
<i>pcd0</i>	-0.1144	0.1040	0.1934	0.2267	0.9917	1	
<i>pca5</i>	0.1181	-0.0982	-0.1889	-0.2193	-0.9762	-0.9913	1

Table 5: Summary Statistics for Continuous Variables across Values of Binary Variables.

	<i>persist</i>	<i>gdp05</i>	<i>urban</i>	<i>trade</i>
<i>change</i> = 0	11.77 (13.78)	3913.35 (3765.58)	45.65 (20.40)	65.96 (32.89)
<i>change</i> = 1	9.92 (13.81)	3312.92 (3065.97)	42.24 (17.69)	59.51 (29.57)
<i>d0</i> = 0	15.93 (16.97)	3127.88 (3325.73)	40.21 (18.44)	63.00 (31.26)
<i>d0</i> = 1	6.99 (6.63)	4657.73 (3947.97)	50.94 (20.57)	68.08 (33.96)
<i>d5</i> = 0	13.81 (16.29)	3063.33 (3183.45)	40.25 (17.85)	63.33 (30.41)
<i>d5</i> = 1	7.93 (6.46)	5217.05 (4139.94)	54.02 (21.02)	69.02 (35.92)
<i>a5</i> = <i>d5</i> = 0	6.00 (7.87)	3248.91 (3298.55)	43.89 (18.06)	65.70 (30.64)
<i>a5</i> = 0	7.06 (7.20)	4324.12 (3904.83)	49.42 (20.36)	67.52 (33.65)
<i>a5</i> = 1	21.40 (18.62)	2882.96 (3059.97)	36.72 (16.93)	61.02 (30.03)

Mean is listed for each variable; standard deviation in parentheses.



Table 6: Correlations among Binary Variables.

	<i>change</i>		<i>d0</i>		<i>d5</i>		<i>a5</i>		<i>a5 = d5</i>	
	0	1	0	1	0	1	0	1	0	1
<i>change = 0</i> <i>change = 1</i>			52.32%	47.68%	61.95%	38.05%	66.46%	33.54%	28.41%	71.59%
<i>d0 = 0</i> <i>d0 = 1</i>	92.98%	7.02%	45.60%	54.40%	100%	0%	38.55%	61.45%	38.55%	61.45%
<i>d5 = 0</i> <i>d5 = 1</i>	90.85%	9.15%	82.52%	17.48%	22.75%	77.25%	100%	0%	49.29%	50.71%
<i>a5 = 0</i> <i>a5 = 1</i>	89.71%	10.29%	29.28%	70.72%	45.37%	54.63%	100%	0%	45.37%	54.63%
<i>a5 = d5 = 0</i>	84.54%	15.46%	64.54%	35.46%	100%	0%	100%	0%	100%	0%

Percentages sum across each cell to 100 (except  $a5 = d5 = 0$ ).

Figure 6: Democratization between 1980-2007.

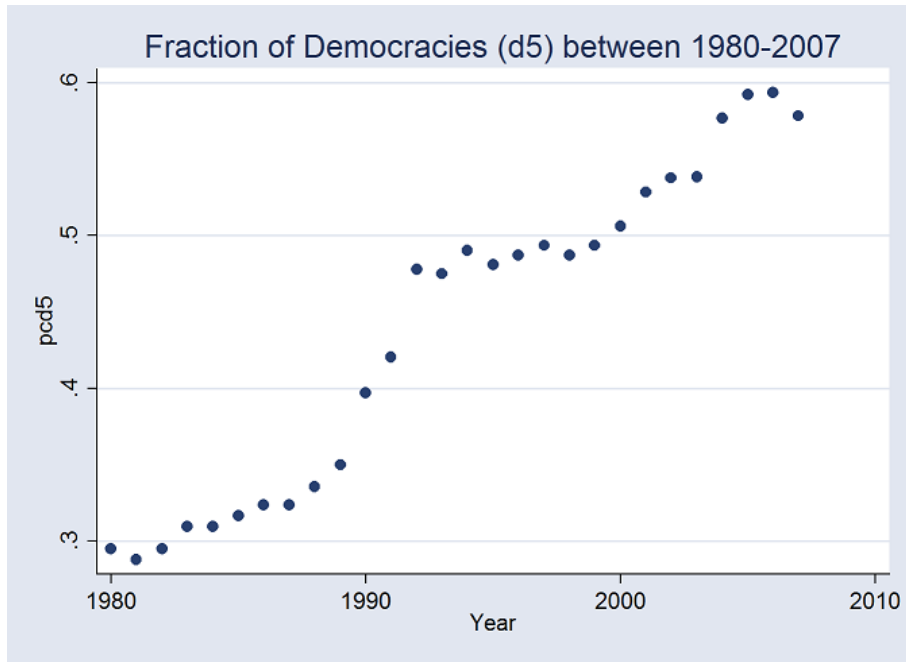


Figure 7: Fraction of Institutional Change observed in Sample between 1980-2007.

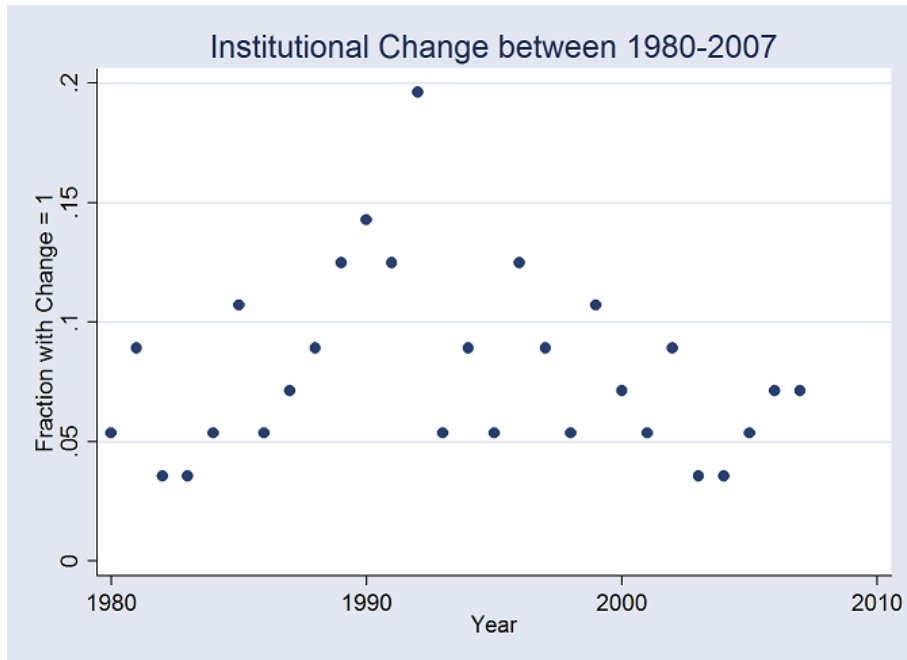


Figure 8: Fraction of Institutional Change among Democracies in Sample.

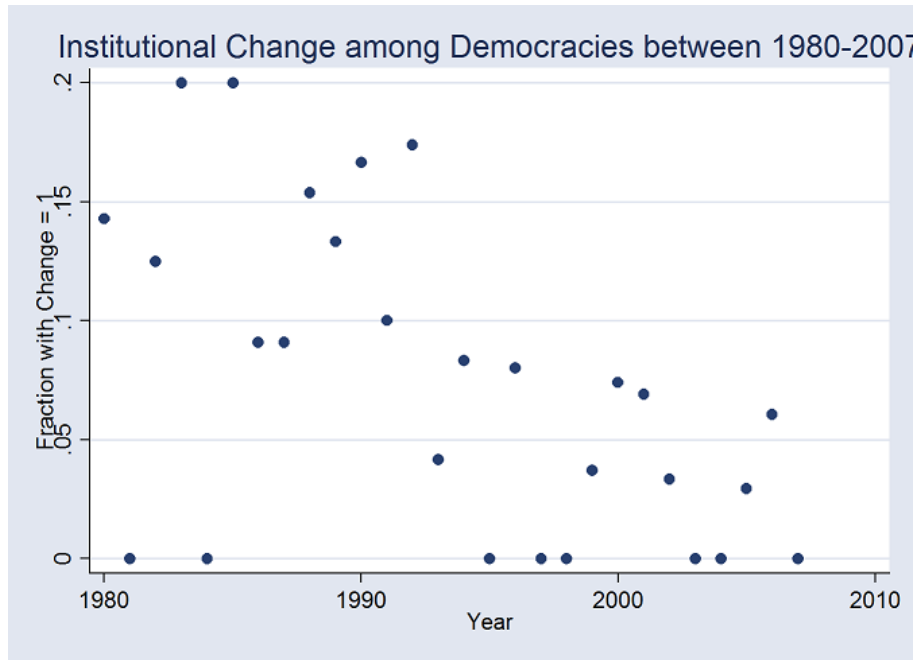
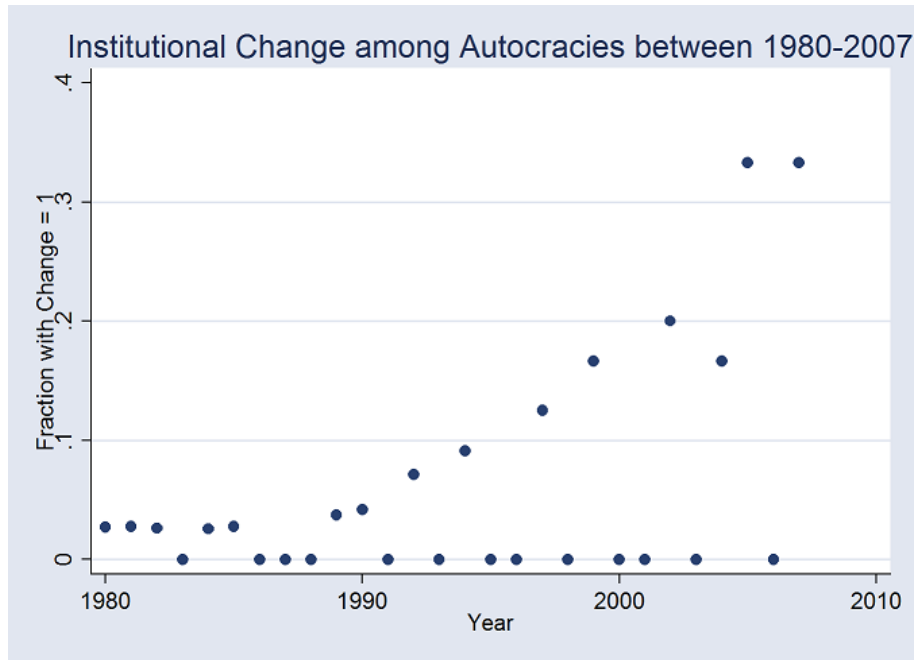


Figure 9: Fraction of Institutional Change among Autocracies in Sample.



## J Tables of Results

Table 7: Fixed Effects Probit with Bias Correction (*d5* only).

	[1]	[2]	[3]	[4]	[5]
<i>d5</i>	−0.207 (0.306)	−0.222 (0.305)	−0.206 (0.304)	−0.277 (0.274)	−0.267 (0.276)
<i>persist</i>	0.233 (0.788)		0.199 (0.785)	0.229 (0.757)	
<i>gdp05</i>	−0.260 (0.449)	−0.228 (0.445)		−0.135 (0.418)	
<i>urban</i>	−0.965* (0.565)	−0.943* (0.558)	−1.16** (0.567)	−0.884* (0.526)	−1.04** (0.526)
<i>trade</i>	−0.208 (0.528)	−0.201 (0.523)	−0.216 (0.525)		
<i>pcd5</i>	0.516 (0.603)	0.456 (0.594)	0.517 (0.600)	0.465 (0.551)	0.411 (0.545)
<i>d5*pcd5</i>	−0.677** (0.294)	−0.603** (0.293)	−0.693** (0.292)	−0.708*** (0.263)	−0.652** (0.264)

Dependent Variable: Institutional Change.

Standard errors in parentheses.

Country Fixed Effects included in all estimations (not shown).

Significance: \* 10% significance \*\* 5% significance \*\*\* 1% significance

Table 8: Probit Estimation without Bias Correction.

	[1]	[2]	[3]	[4]	[5]	[6]
<i>d5</i>	-0.234** (0.109)	-0.206* (0.108)	-0.234** (0.109)	-0.242** (0.109)	-0.215** (0.107)	-0.229** (0.109)
<i>persist</i>	-0.082 (0.051)		-0.083 (0.051)	-0.078 (0.051)	-0.083 (0.051)	-0.081 (0.051)
<i>gdp05</i>	-0.005 (0.084)	-0.021 (0.079)		-0.056 (0.053)	-0.014 (0.066)	-0.011 (0.079)
<i>urban</i>	-0.065 (0.080)	-0.049* (0.079)	-0.068 (0.051)		0.025 (0.075)	-0.051 (0.077)
<i>trade</i>	-0.061 (0.065)	-0.075 (0.065)	-0.061 (0.065)	-0.068 (0.064)		-0.096* (0.052)
<i>pcd5</i>	0.046 (0.051)	0.051 (0.051)	0.047 (0.051)	0.042 (0.050)	0.012 (0.051)	0.045 (0.051)
<i>d5*pcd5</i>	-0.096 (0.109)	-0.090 (0.109)	-0.097 (0.108)	-0.079 (0.106)	-0.206* (0.105)	
<i>constant</i>	-1.34*** (0.059)	-1.35*** (0.059)	-1.34*** (0.059)	-1.34*** (0.059)	-1.31*** (0.059)	-1.34*** (0.059)

Dependent Variable: Institutional Change.

Standard errors in parentheses.

Significance: \* 10% significance \*\* 5% significance \*\*\* 1% significance

Table 9: Fixed Effects Probit with Bias Correction (*d5* and *a5*).

	[1]	[2]	[3]	[4]	[5]
<i>d5</i>	-0.484*	-0.478	-0.483	-0.507**	-0.485*
	(0.290)	(0.293)	(0.294)	(0.252)	(0.267)
<i>a5</i>	-1.12***	-0.831***	-1.13***	-1.11***	-0.830***
	(0.303)	(0.285)	(0.312)	(0.248)	(0.250)
<i>persist</i>	0.419		0.408	0.406	
	(0.763)		(0.768)	(0.719)	
<i>gdp05</i>	0.101	0.133		0.255	
	(0.414)	(0.417)		(0.373)	
<i>urban</i>	-1.40***	-1.37***	-1.50***	-1.24***	-1.29***
	(0.529)	(0.523)	(0.539)	(0.471)	(0.489)
<i>trade</i>	-0.238	-0.211	-0.239		
	(0.472)	(0.476)	(0.480)		
<i>pcd5</i>	0.193	0.096	0.191	0.095	0.002
	(0.537)	(0.533)	(0.546)	(0.466)	(0.488)
<i>pca5</i>	0.107	0.062	0.106	0.102	0.048
	(0.544)	(0.538)	(0.553)	(0.474)	(0.494)
<i>d5*pcd5</i>	-0.546**	-0.407	-0.578**	-0.598**	-0.461*
	(0.278)	(0.279)	(0.282)	(0.241)	(0.254)
<i>a5*pca5</i>	1.20***	1.10***	1.16***	1.19***	1.07***
	(0.318)	(0.303)	(0.328)	(0.256)	(0.265)
<i>d5*pca5</i>	-0.169	-0.159	-0.221	-0.096	-0.126
	(0.276)	(0.276)	(0.278)	(0.245)	(0.255)
<i>a5*pcd5</i>	1.88***	1.84***	1.85***	1.88***	1.83***
	(0.303)	(0.291)	(0.312)	(0.246)	(0.256)

Dependent Variable: Institutional Change.

Standard errors in parentheses.

Country Fixed Effects included in all estimations (not shown).

Significance: \* 10% significance \*\* 5% significance \*\*\* 1% significance



Table 10: Average Marginal Effects on Institutional Change for Country Randomly Drawn from Sample (*d5* only).

	[1] All included	[2] <i>persist</i> excluded	[3] <i>gdp05</i> excluded	[4] <i>trade</i> excluded	[5] <i>persist, gdp05, trade</i> excluded
<i>d5</i>	-0.075	-0.068	-0.077	-0.098	-0.084
<i>persist</i>	0.033		0.028	0.033	
<i>gdp05</i>	-0.037	-0.033		-0.019	
<i>urban</i>	-0.137*	-0.135*	-0.165**	-0.126*	-0.150**
<i>trade</i>	-0.030	-0.029	-0.031		
<i>pcd5</i> for <i>d5</i> = 1	0.017**	-0.016**	-0.019**	-0.025***	-0.026**
<i>pcd5</i> for <i>d5</i> = 0	0.082	0.073	0.082	0.074	0.064

Significance of Point Estimate: \* 10% significance \*\* 5% significance \*\*\* 1% significance  
 Continuous variables: average marginal effect of increasing variable by one standard deviation.

Table 11: Average Marginal Effects on Institutional Change for Country Randomly Drawn from Sample ( $d5$  and  $a5$ ).

	[1] All included	[2] <i>persist</i> excluded	[3] <i>gdp05</i> excluded	[4] <i>trade</i> excluded	[5] <i>persist, gdp05, trade</i> excluded
<i>d5</i>	-0.089*	-0.082	-0.084	-0.100**	-0.081*
<i>a5</i>	-0.319***	-0.239***	-0.328***	-0.328***	-0.254***
<i>persist</i>	0.055		0.053	0.053	
<i>gdp05</i>	0.013	0.018		0.033	
<i>urban</i>	-0.183***	-0.182***	-0.195***	-0.162***	-0.171***
<i>trade</i>	-0.031	-0.028	-0.031		
<i>pcd5</i> for $d5 = 1$	-0.037**	-0.034	-0.040**	-0.052**	-0.049*
<i>pcd5</i> for $a5 = 1$	0.145***	0.127***	0.141***	0.139***	0.120***
<i>pcd5</i> for $d5 = a5 = 0$	0.043	0.022	0.043	0.021	0.0005
<i>pca5</i> for $d5 = 1$	-0.006	-0.011	-0.011	0.0006	-0.008
<i>pca5</i> for $a5 = 1$	0.092***	0.076***	0.088***	0.091***	0.074***
<i>pca5</i> for $d5 = a5 = 0$	0.024	0.014	0.024	0.023	0.011

Significance of Point Estimate: \* 10% significance \*\* 5% significance \*\*\* 1% significance  
Continuous variables: average marginal effect of increasing variable by one standard deviation.

Table 12: Fixed Effects Probit with Bias Correction and Time Trend Controls.

	[1]	[2]	[3]	[4]	[5]	[6]
<i>d5</i>	-0.279 (0.299)	-0.286 (0.303)	-0.209 (0.305)	-0.531* (0.282)	-0.519* (0.286)	-0.487* (0.289)
<i>a5</i>				-1.15*** (0.302)	-1.12*** (0.304)	-1.12*** (0.304)
<i>persist</i>	0.205 (0.785)	0.194 (0.788)	0.233 (0.788)	0.405 (0.756)	0.384 (0.760)	0.418 (0.763)
<i>gdp05</i>	-0.105 (0.449)	-0.053 (0.452)	-0.256 (0.449)	0.228 (0.411)	0.298 (0.416)	0.113 (0.414)
<i>urban</i>	-1.10* (0.564)	-1.17** (0.568)	-0.970* (0.566)	-1.49*** (0.524)	-1.56*** (0.529)	-1.41*** (0.529)
<i>trade</i>	-0.228 (0.522)	-0.222 (0.526)	-0.210 (0.528)	-0.279 (0.460)	-0.274 (0.466)	-0.242 (0.471)
<i>time (linear)</i>			-0.003 (0.598)			0.002 (0.531)
<i>pcd5</i>	-0.425 (0.594)	0.426 (0.594)	0.506 (0.603)	-0.976* (0.528)	0.210 (0.529)	0.166 (0.537)
<i>pca5</i>				-0.638 (0.536)	0.246 (0.536)	0.103 (0.544)
<i>d5*pcd5</i>	-0.680** (0.285)	-0.661** (0.287)	-0.677** (0.294)	-0.585** (0.269)	-0.615** (0.272)	-0.553** (0.276)
<i>a5*pca5</i>				1.60*** (0.314)	1.17*** (0.312)	1.20*** (0.320)
<i>d5*pca5</i>				-0.157 (0.269)	-0.169 (0.271)	-0.170 (0.275)
<i>a5*pcd5</i>				2.30*** (0.300)	1.87*** (0.298)	1.88*** (0.305)
<i>Period Dum</i>	1 yr / period	4 yrs / period		1 yr / period	4 yrs / period	

Dependent Variable: Institutional Change.

Standard errors in parentheses.

Country Fixed Effects included in all estimations (not shown).

Time Fixed Effects included in some estimations (noted in *Period Dum* but not shown).

Significance: \* 10% significance \*\* 5% significance \*\*\* 1% significance

Table 13: Two-Step Fixed Effects Probit with Bias Correction.

	[1]	[2]	[3]	[4]
<b>First Stage – Dependent Variable: <i>d5</i></b>				
<i>ld5</i>	2.52*** (0.254)	2.85*** (0.257)	2.52*** (0.254)	2.85*** (0.257)
<i>gdp05</i>	0.039 (0.356)		0.039 (0.356)	
<i>urban</i>	-0.557 (0.404)		-0.557 (0.404)	
<i>trade</i>	0.161 (0.414)		0.161 (0.414)	
<i>pcd5</i>	0.360 (0.415)		0.360 (0.415)	
<b>Second Stage – Dependent Variable: Institutional Change</b>				
<i>d5</i>	-1.07*** (0.246)	-0.962*** (0.252)	-1.14*** (0.253)	-1.04*** (0.256)
<i>persist</i>	0.252 (0.495)	0.268 (0.493)		
<i>gdp05</i>	-0.455 (0.427)	-0.322 (0.436)		
<i>urban</i>	-1.53*** (0.465)	-1.31*** (0.466)	-1.47*** (0.475)	-1.29*** (0.474)
<i>trade</i>	-0.189 (0.465)	-0.207 (0.466)		
<i>pcd5</i>	0.654 (0.485)	0.611 (0.484)	0.545 (0.495)	0.492 (0.494)
<i>d5*pcd5</i>	-0.547** (0.229)	-0.539** (0.235)	-0.481** (0.238)	-0.479** (0.240)
<i>gen resid</i>	0.662** (0.277)	0.627** (0.294)	0.650** (0.278)	0.607** (0.295)

Standard errors in parentheses.

Country Fixed Effects included in all estimations (not shown).

Significance: \* 10% significance \*\* 5% significance \*\*\* 1% significance

Table 14: Two-Stage Least Squares Estimation.

	[1]	[2]	[3]	[4]	[5]
<b>First Stage – Dependent Variable: <math>d5</math></b>					
<i>ld5</i>	0.811*** (0.024)		0.812*** (0.024)	0.812*** (0.024)	
<i>gdp05</i>	0.004 (0.015)				
<i>urban</i>	-0.022 (0.033)				
<i>trade</i>	0.028** (0.013)		0.029** (0.013)	0.029** (0.013)	
<i>pcd5</i>	0.026*** (0.009)		0.022*** (0.006)	0.022*** (0.006)	
$R^2$	0.847		0.847	0.847	
<b>Second Stage – Dependent Variable: Institutional Change</b>					
<i>d5</i>	-0.157*** (0.034)	-0.022 (0.027)	-0.157*** (0.034)	-0.168*** (0.031)	-0.036 (0.026)
<i>persist</i>	0.027* (0.014)	0.027* (0.015)	0.027* (0.014)		
<i>gdp05</i>	-0.037 (0.027)	-0.052* (0.028)	-0.038 (0.027)		
<i>urban</i>	-0.137*** (0.046)	-0.118** (0.047)	-0.129*** (0.047)	-0.135*** (0.048)	-0.126*** (0.047)
<i>trade</i>	-0.004 (0.018)	-0.017 (0.017)	-0.004 (0.018)		
<i>pcd5</i>	0.082*** (0.014)	0.066*** (0.014)	0.081*** (0.014)	0.074*** (0.014)	0.058*** (0.014)
<i>d5*pcd5</i>	-0.069*** (0.020)	-0.082*** (0.022)	-0.069*** (0.020)	-0.064*** (0.017)	-0.082*** (0.020)
<i>resid</i>	0.384*** (0.114)		0.383*** (0.114)	0.388*** (0.113)	
$R^2$	0.104	0.058	0.104	0.100	0.052

Robust Standard errors in parentheses.

Country Fixed Effects included in all estimations (not shown).

Significance: \* 10% significance \*\* 5% significance \*\*\* 1% significance

The (One-Stage) Least Squares Estimations are given in [2] and [5] for comparison.

## References

- Acemoglu, Daron and James A. Robinson**, “Persistence of Power, Elites, and Institutions,” *American Economic Review*, 2008, 98 (1), 267–293.
- , **Georgy Egorov, and Konstantin Sonin**, “Dynamics and Stability of Constitutions, Coalitions, and Clubs,” Mimeo, Department of Economics, MIT 2008.
- Alston, Lee J., Thráinn Eggertsson, and Douglass C. North**, *Empirical Studies in Institutional Change*, New York: Cambridge University Press, 1996.
- Arrow, Kenneth**, “Uncertainty and the Welfare Economics of Medical Care,” *American Economic Review*, 1963, 53, 941–973.
- Athey, Susan, Andrew Atkeson, and Patrick Kehoe**, “The Optimal Degree of Discretion in Monetary Policy,” *Econometrica*, 2005, 73, 1431–1475.
- Barro, Robert**, “Determinants of Economic Growth: A Cross-Country Empirical Study,” *NBER Working Paper*, 1996, # 5698.
- Boutros-Ghali, Boutros**, “Democracy: A Newly Recognized Imperative,” *Global Governance*, 1995, 1, 3–11.
- Burke, Edmund**, “Speech to the electors of Bristol,” in “The Works of the Right Honourable Edmund Burke,” Vol. I, London: Henry G. Bohn, 1854.
- Caillaud, Bernard and Jean Tirole**, “Parties as Political Intermediaries,” *The Quarterly Journal of Economics*, 2002, 117 (4), 1453–1489.
- Caplin, Andrew and Barry Nalebuff**, “On 64% Majority Rule,” *Econometrica*, 1988, 56 (4), 787–814.
- Cho, In-Koo and David Kreps**, “Signaling Games and Stable Equilibria,” *The Quarterly Journal of Economics*, 1987, 102, 179–222.
- Coate, Stephen and Stephen Morris**, “Policy Persistence,” *American Economic Review*, 1999, 89 (5), 1327–1336.
- Dasgupta, Partha and Eric Maskin**, “Efficient Auctions,” *The Quarterly Journal of Economics*, 2000, 115, 341–388.
- Diamond, Larry**, “Promoting Democracy,” *Foreign Policy*, 1992, 87, 25–46.
- Dollar, David and Aart Kraay**, “Institutions, Trade, and Growth,” *World Bank Policy Research Working Paper*, 2003, # 3004.

- Downs, Anthony**, *An Economic Theory of Democracy*, New York: Harper & Row, 1957.
- Epstein, Larry and Michael Peters**, “A Revelation Principle for Competing Mechanisms,” *Journal of Economic Theory*, 1999, 88, 119–160.
- Feng, Yi**, “Democracy, Political Stability and Economic Growth,” *British Journal of Political Science*, 1997, 27, 391–418.
- Fernández-Val, Ivan**, “Fixed Effects Estimation of Structural Parameters and Marginal Effects in Panel Probit Models,” *Journal of Econometrics*, 2009, 150, 71–85.
- **and Francis Vella**, “Bias Corrections for Two-Step Fixed Effects Panel Data Estimators,” *IZA Discussion Paper*, 2007, no. 2690.
- Fox, Justin and Kenneth J. Shotts**, “Delegates or Trustees? A Theory of Political Accountability,” *Journal of Politics*, 2009, *forthcoming*.
- Fudenberg, Drew and Jean Tirole**, *Game Theory*, Cambridge, MA: MIT Press, 1991.
- Greene, William**, “The Behavior of the Fixed Effects Estimator in Nonlinear Models,” Mimeo, Department of Economics, Stern School of Business, NYU 2002.
- Grossman, Sanford and Oliver Hart**, “An Analysis of the Principal-Agent Problem,” *Econometrica*, 1983, 51, 7–45.
- Hahn, Jinyong and Whitney Newey**, “Jackknife and Analytical Bias Reduction for Nonlinear Panel Models,” *Econometrica*, 2004, 72 (4), 1295–1319.
- Holmstrom, Bengt**, “Moral Hazard and Observability,” *Bell Journal of Economics*, 1979, 10 (1), 74–91.
- Hotelling, Harold**, “Stability and Competition,” *Economic Journal*, 1929, 39, 41–57.
- Ikenberry, G. John**, “Why export democracy? The “hidden grand strategy” of American foreign policy is reemerging into plain view after a long Cold War hibernation,” *The Wilson Quarterly*, 1999, 23 (2), 56.
- Jordan, James**, “Pillage and Property,” *Journal of Economic Theory*, 2006, 131 (1), 26–44.

- Kessler, Anke**, “Representative versus Direct Democracy: The role of informational asymmetries,” *Public Choice*, 2005, *122*, 9–38.
- Klein, Roger and Richard Spady**, “An Efficient Semiparametric Estimator for the Binary Response Model,” *Econometrica*, 1993, *61*, 387–421.
- Lagunoff, Roger**, “Dynamic Stability and Reform of Political Institutions,” *Games and Economic Behavior*, 2009, *67* (2), 569–583.
- Lindbeck, Assar and Jorgen Weibull**, “Balanced-budget Redistribution as the outcome of Political Competition,” *Public Choice*, 1987, *60*, 273–297.
- Marshall, Monty G. and Keith Jagers**, *Polity IV Project: Dataset Users’ Manual* 2005.
- Martimort, David and Humberto Moreira**, “Common Agency with Informed Principals,” *Escola de Pós-Graduação em Economia Working Paper Series*, 2004, No. 551.
- **and Lars Stole**, “The Revelation and Delegation Principles in Common Agency Games,” *Econometrica*, 2002, *70*, 1659–1673.
- Maskin, Eric and Jean Tirole**, “The Principal-Agent Relationship with an Informed Principal: The case of Private Values,” *Econometrica*, 1990, *58*, 379–409.
- **and –**, “The Principal-Agent Relationship with an Informed Principal, II: Common Values,” *Econometrica*, 1992, *60*, 1–42.
- **and –**, “The Politician and the Judge: Accountability in Government,” *American Economic Review*, 2004, *94* (4), 1034–1054.
- Mill, John Stuart**, *Considerations on Representative Government*, New York: Henry Holt & Co., 1882.
- Mirrlees, James**, “The Optimal Structure of Incentives and Authority within an Organization,” *Bell Journal of Economics*, 1976, *7*, 105–131.
- Myerson, Roger**, “Mechanism Design by an Informed Principal,” *Econometrica*, 1983, *51*, 1767–1798.
- Neyman, J. and Elizabeth L. Scott**, “Consistent Estimates Based on Partially Consistent Observations,” *Econometrica*, 1948, *16* (1), 1–32.



- North, Douglass C.**, *Institutions, Institutional Change, and Economic Performance*, New York: Cambridge University Press, 1990.
- Persson, Torsten and Guido Tabellini**, “Democratic Capital: The Nexus of Political and Economic Change,” *American Economic Journal: Macroeconomics*, 2009, 1 (2), 88–126.
- Peters, Michael**, “Common Agency and the Revelation Principle,” *Econometrica*, 2001, 69, 1349–1372.
- Rivers, Douglas and Quang H. Vuong**, “Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models,” *Journal of Econometrics*, 1988, 39 (3), 347–366.
- Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi**, “Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development,” *NBER Working Paper*, 2002, # 9305.
- Rothstein, Paul**, “Order Restricted Preferences and Majority Rule,” *Social Choice and Welfare*, 1990, 7, 331–342.
- Schultz, Christian**, “Information, Polarization and Term Length in Democracy,” *Journal of Public Economics*, 2008, 92, 1078–1091.
- Sen, Amartya**, “Democracy as a Universal Value,” *Journal of Democracy*, 1999, 10 (3), 3–17.
- Vlaicu, Razvan**, “Executive Performance under Direct and Hierarchical Accountability Structures: Theory and Evidence,” *SSRN Working Paper Series*, 2008.