Competing Informed Principals and Representative Democracy*

Helena Aten†

This Version: February 11, 2009

Abstract

This paper proposes a model in which representative democracy can be preferable to direct democracy. Voters are uninformed about the value of a policy-relevant state. Two informed politicians compete for votes by committing to platforms that may or may not reveal information about the underlying state.

We 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, our 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. The result may be viewed as a welfare explanation for why voters may defer policy choices to an elected representative rather than directly select policy based on the information revealed by the political competition itself.

JEL Classification: C72, D61, D82, D86

Keywords: Informed Principal, Representative Democracy, Contract Theory, Political Competition.

---

*I am greatly indebted to my advisors Luca Anderlini and Roger Lagunoff for their guidance, patience, and encouragement. All errors are my own.

†Department of Economics, Georgetown University, Washington DC 20057 USA. hrp3@georgetown.edu

1 Introduction

The choice of direct versus representative democracy has been the source of much interest among philosophers and social scientists. Direct democracy is commonly known as a system in which "members of a society vote directly on policy options"(Grossman and Helpman, 2001, p.42). Even where officials may be elected to legislate in a direct democracy, voters have much more power in directly influencing policies through referendum or representative recall. Dating back to the Athenian city-state, current examples of direct democracy can be found in Switzerland and a number of states in the US that utilize referenda. Representative democracy, on the other hand, is an institution in which "the citizens elect a subset of their number and endow them with the authority to make policy decisions on their behalf" (Grossman and Helpman, 2001, p.53). Once elected, the representatives are given discretion to implement the policy of their choice during their term, and voters will have little or no means of affecting policies ex post. Representative democracy was first adopted in the Roman Republic around 500 BC[1]; current examples include all legislatures and parliaments in which representatives are elected to serve the people.

While representative democracy is ubiquitous among political institutions today, intuitively direct democracy seems a more straightforward way to aggregate voter preferences. Why would a society instead adopt an indirect system in which policy-making authority is delegated to someone else whose preference likely deviates from the aggregated social preference?

One common explanation for why representative democracy is used is that it economizes on transaction costs. Buchanan and Tullock (1962) maintains that "direct democracy, under almost any decision-making rule, becomes too costly in other than very small political units when more than a few isolated issues must be considered." While this is likely true, a model of exogenous transaction costs directly implies the adoption and persistence of representative democracy. Also, as these transaction costs continue to decrease as a result of technological advances, should we shift away from representative to direct democracy as Matsusaka (2003) and Beedham (1996) suggest, or are there other reasons for the continued adoption of representative democracy?

[1]Though the voting franchise was very limited by today's standards, the new Republic embraced the general idea of government of the people; representatives were chosen among those eligible to serve in the Senate.

This paper proposes an alternative welfare-based justification for representative democracy. We start with a simple illustrative model of two public goods – guns and butter. Two politicians seeking office must choose how much of each public good to produce as part of their policy platform. There are two states of the world – good times and bad – and voters' preferences over the public good can be dependent or independent of the state. Politicians know the state of the world, voters do not; however, voters may update their beliefs about the state after observing the policy platforms offered by the politicians.

How is this model linked to the question of direct versus representative democracy? In the process of competing in an election, the information possessed by the politicians – or more generally, political competitors[2] – may or may not be revealed to voters. The two possible scenarios parallel the cases of direct and representative democracy, and provide basis for welfare comparison between these two archetypes.

In an equilibrium in which the electoral process induces both politicians to reveal what they know, voters have all the information they need to vote on issues directly. The winning politician's role is only to ensure the implementation of the promised policy. Politicians thus either become obsolete or are reduced to mere vehicles for information transmission. In this sense, a fully revealing equilibrium boils down to the case of direct democracy.

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 voting franchise *commits* to entrust the elected representative with full policy-making authority prior to knowing the state of the world, eliminating voters' ability to influence policy ex post. In this sense, a non-revealing equilibrium corresponds to the case of representative democracy.

This paper addresses two main questions:

(i) Are voters always better off if the state is revealed to them through the electoral process than if it is not?

(ii) Are the right kinds and amounts of public goods offered in equilibrium?

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

We find that if voters' policy preferences are not too sensitive to the policy-relevant state, then policy platforms offered will differ from the voters' ideal point. More importantly, we show that if the median voter is sufficiently ideologically driven and has relative state-independent policy preferences, then aggregate welfare is higher when the state is *not* revealed to voters in the process of political competition. If delegating the decision making process to politicians yields higher ex ante aggregate welfare, then one may interpret representative democracy as a welfare-enhancing commitment mechanism adopted to prevent voters from revising chosen policies after the state is realized. The role of a politician extends beyond that of an expert or information provider – she must withhold the relevant information, which is to be revealed only after the election.

The model in this paper is built on the standard Hotelling-Downsian (Hotelling, 1929; Downs, 1957) multi-dimensional political competition model, and borrows from the probabilistic voting model of Lindbeck and Weibull (1987) to represent the intensity of voters' ideological predisposition. The tradition in voting theory since Hotelling-Downs has been to solve the problem of information aggregation among voters, when the size of the electorate is large, and voters may be strategic[3].

Even where politicians may possess private information, the interaction between politicians (or the expert) and "the government" (whose interest aligns with that of the voters or the public) is modeled as a standard principal-agent problem[4]. In these models, the government is the principal while the politician is the agent. Others argue that representative democracy is adopted as a result of informational asymmetries between politicians and voters. In the case of Kessler (2005), informational asymmetry arises as a result of costly information acquisition. The choice of direct versus representative democracy essentially boils down to one of costs (whether it is worthwhile to hire the expert for information in a direct democracy), but the role of politicians is merely that of an information provider. Schultz (2008) has a model similar to that of the current 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. This paper argues that representative democracy is not only a *result* of the information asymmetry between politicians and voters, it may also be beneficial to maintain the information asymmetry through representative democracy.

[3]For example, see Feddersen and Pesendorfer (1996, 1997).

[4]For instance, see Athey et al. (2005), Grossman and Hart (1983), and Mirrlees (1976).

A novelty of this paper is that politicians take on the role of an informed principal rather than that of an agent or expert. This is more in line with the spatial political competition models, and highlights the strategic decisions made by competing politicians when they possess payoff-relevant information[5]. This paper extends the informed principal framework in the existing literature by having two informed principals compete to contract with a single agent. In addition to applying this to a model of electoral competition, the extension that this paper provides may also be useful for understanding other applications with multiple competing principals[6].

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). Myerson (1983) establishes the well-known Inscrutability Principle, in which 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. They classify the problem into the cases of private and common values[7]. Equilibrium characterization differs between the two cases because in common values, there is a fundamental conflict between different types of the informed principal in choosing how much information to reveal. Such conflict is absent in private values.

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 and incentive compatibility 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. The main result of this paper follows a similar line of

[5]While it is true that voters may also have private information, this paper abstracts from it to highlight the results driven by the principals being the ones informed. In fact, Maskin and Tirole (1990, 1992) allow for having two-sided private information, and show that the qualitative results are not different from that obtained when only the principal has private information.

[6]For example, there is usually more than one contractor competing for a contract to carry out construction work; you can choose to sign a franchising agreement with either McDonald's or Burger King; or when purchasing a car, you have multiple brands to choose from, each with uncertain quality and purchasing plans.

[7]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.

logic. We show that even with two competing principals, the set of pooling equilibria will welfare dominate the set of separating equilibria for parameter configurations "close to" the case of Maskin and Tirole (1990).

The rest of the paper is organized as follows: Section 2 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

To map the structure of political competition between two informed politicians into the informed principal framework, we start with the two informed politicians as the principals. A female pronoun is used throughout to denote a principal, while a male pronoun denotes an agent; in addition, a policy platform promised by the politicians and a contract are used interchangeably. We assume preferences of voters are such that a median voter exists[8] – this median voter is the agent. The agent has state-dependent preferences for the public goods, and an ideological predisposition for each political candidate. The timing of the game is as follows:

1. Nature draws the state of the world θ. It is observed only by the politicians (indexed $i = 1, 2$). The median voter's ideology is drawn from a known distribution $F(\cdot)$.

2. The two politicians simultaneously offer policy platforms.

3. Upon observing the policies promised, the median voter updates his beliefs about the state and votes for at most one politician.

4. If a politician is elected, the relevant parties will carry out the details of the contract[9].

A 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, which can also be interpreted as "pork barrel spending" or

[8]See Rothstein (1990) for a general result on the existence of a median voter in models with multidimensional policy space.

[9]We assume that the contract is binding on both sides.

earmarks that benefit specific subsets of the electorate (from which the politician directly or indirectly gains). Finally, t is the tax that the median voter pays to finance both the public goods and the pork. The preferences of all involved in the game are described in detail below.

Principals' Preferences (Politicians)

$$\begin{array}{l} V_1 = \lambda b + p \\ V_2 = \lambda g + p \end{array} \quad ; \quad \lambda < 1$$

The principals' preferences are state-independent and commonly known. V_i is the utility of principal i if her contract with terms (b, g, p, t) is accepted, otherwise her payoff is normalized to zero. We assume $\lambda < 1$[10], denoting the principal's relative preference of direct transfer over the public good. Note that each principal prefers a different public good (principal 1 likes public good b, while principal 2 likes g), but that preference does not vary by state.

Agent's Preference (Median Voter)

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

- $\gamma\eta > 1$;
- $0 \leq t \leq 1$;
- $\eta \in \left[\frac{1}{2}, 1\right]$; and
- $c_1 \equiv 0$ and $c_2 \sim \psi F(c)$ over $\mathbb{R}$,[11] where $\psi \in \mathbb{R}_+$, and $F(c)$ is a continuously differentiable distribution with density $f(c)$.

U_i^θ is the utility of the agent *if he accepts type θ of Principal i's* (henceforth $Pi\theta$) *contract.* The agent's preference for the public goods can take on any level of state dependence, parameterized by η. Regardless of the value of η, the agent (at least weakly)

[10] $\lambda < 1$ is needed since we allow p to be positive or negative (See discussion on page 8). If $\lambda \geq 1$ and $p \in \mathbb{R}$, the principal 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.

[11] From the point of view the principals.

prefers b in the high state, and g in the low state. He dislikes taxes, but in relative terms he always prefers to have one unit in taxes in exchange for one unit of his preferred public good ($\gamma\eta > 1$). The principal-specific utility that the agent derives from choosing principal i is given by c_i, or what we call the agent's "ideology." The agent and the principals all have a normalized reservation utility of zero.

The two variables that make the informed principal framework vary in two dimensions are η and c. Along the dimension of public good specificity, η represents the intensity of the agent's relative public good preference across the two states. If $\eta = \frac{1}{2}$, the agent likes both public goods equally regardless of the state. If $\eta = 1$, the agent *only* wants b (respectively g) when $\theta = H$ (respectively $\theta = L$). Along the dimension of political competition, c_i is the ideological component in the agent's preference, as used in a probabilistic voting model. The purpose of this ideological component is to smooth out discontinuous jumps in the probability of winning when one principal offers a contract just infinitestimally better than that of her opponent. We normalize c_1 to zero, and $c_2 \equiv \psi c$, where c is a random variable representing the voter's relative ideological predisposition. The parameter ψ is a positive number that measures the voter's *intensity* of ideological preference relative to policy. If $\psi = 0$, then the median voter cares only about policy; a larger ψ indicates a higher level of importance placed on ideology compared to policy.

As mentioned above, our goal is to use a simple but illustrative model to provide a possible welfare explanation for representative democracy. 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 it seems 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

$$\begin{array}{rcc} b+g+p \leq t & \text{if} & \theta = H \\ \alpha(b+g)+p \leq t & & \theta = L \end{array} \quad ; \quad \text{where} \quad \begin{array}{c} \gamma\eta > \alpha > \lambda+1;^{12} \\ b \geq 0 \ ; \ g \geq 0 \end{array}$$

[12] The inequality $\alpha > \lambda+1$ makes $P2H$ "more able to compete" than $P2L$, otherwise $\alpha > 1$ is sufficient. See Section 4.1 for details.

The feasibility constraint describes the technology with which the principals convert taxes into public goods and pork. It is identical across principals, 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 agent'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 various contract terms (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; here we set it between zero and one. The "net pork" p that each politician receives can be positive or negative for two reasons: (1) implicit or indirect transfers are often used in politics to elicit votes from a subset 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. (2) 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 $p < 0$.

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[13], 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 after a contract is accepted. Of course, in a separating equilibrium, the agent will know the true state with certainty on equilibrium path. In terms of notation, $b_i^{\theta\widehat{\theta}}$ denotes the level of b that $Pi\theta$ promises to implement in state $\widehat{\theta}$. Other policy variables follow this convention as well. A menu contract for $Pi\theta$ is therefore $\left(b_i^{\theta\widehat{\theta}}, g_i^{\theta\widehat{\theta}}, p_i^{\theta\widehat{\theta}}, t_i^{\theta\widehat{\theta}}\right)_{\widehat{\theta}\in\{H,L\}}$.

3 Characterization

Our goal is to characterize the set of pure strategy, weakly undominated Perfect Bayesian Equilibria. As in most signaling games, Sequential Equilibrium has no bite in this model.

[13]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.

A menu contract for each $Pi\theta$ is an eight-dimensional object. Fortunately, given the linear structure of the model, Lemma 1 below demonstrates that we can restrict the contract space to only two dimensions without loss of generality. The difficulty is in ruling out all other possible contract terms for *any* beliefs that the agent may have. The key to the result is the use of menu contracts, so we can argue that the policy variables chosen *within each part*[14] of the menu contract must be as described in the lemma, regardless of any on- or off-equilibrium-path beliefs that the agent may assign.

Lemma 1 (**Reduced Strategy Space**). *For any contract offered by any politician, there exists another contract, with terms listed below, that gives the politician a weakly higher expected payoff, regardless of what beliefs the agent has and what her opponent is offering:* $\forall\theta$,

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

$$g_1^{\theta\widehat{\theta}} = b_2^{\theta\widehat{\theta}} = 0 \;\; \forall\widehat{\theta} \; ; \qquad t_i^{\theta\widehat{\theta}} = 1 \;\; \forall\widehat{\theta}, \; \forall i$$

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

$$b_i^{\theta L} = 0 \;\; \forall i \; ; \qquad g_1^{\theta H} = b_2^{\theta H} = 0 \; ; \qquad t_i^{\theta\widehat{\theta}} = 1 \;\; \forall\widehat{\theta}, \; \forall i$$

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

$$b_i^{\theta L} = g_i^{\theta H} = 0 \;\; \forall i \; ; \qquad t_i^{\theta\widehat{\theta}} = 1 \;\; \forall\widehat{\theta}, \; \forall i$$

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

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

Second, the preference-aligned politicians ($P1H$ and $P2L$) will always offer the public good of their choice, exactly because there is no conflict between their and the median

[14] There are two "parts" in each menu contract: one referring to what will be implemented in $\theta = H$, and the other referring to what will be implemented in $\theta = L$.

voter's preferences for the public goods. The preference-misaligned politicians ($P1L$ and $P2H$), on the other hand, offer either the good of their choice or that of the median voter, depending on the relative intensity of the voter's preference for the two public goods, parameterized by η. When η is small, 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 median voter has a marked preference for one of the public goods given the state (η large), then the preference-misaligned politician will offer the public good that the voter prefers.

The case of η 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 median voter's utility, which in turn increases her probability of being elected; and (2) to increase her own utility conditional on contract acceptance. Ideally the politicians would like to offer a public good that attains both of these goals; this is possible for all politicians when η is relatively small. However, when η 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 offering the public good that the voter prefers, and we return to the policy convergence result that obtains in most models of political competition.

Lemma 1 simplifies the structure of the game tremendously, in that we have two out of four of the elements in $\left(b_i^{\theta\widehat{\theta}}, g_i^{\theta\widehat{\theta}}, p_i^{\theta\widehat{\theta}}, t_i^{\theta\widehat{\theta}}\right)$ pinned down, and the final two are bound by the feasibility constraint. This reduces our problem from a 32-dimensional strategy profile to just eight. It follows that Lemma 1 and the feasibility constraint can be combined to describe the relationship between any contract that the principal may offer and her utility level should the contract be accepted (details in Appendix B). Provided that the winning principal will truthfully reveal the state ex post, we can also compute the agent's utility for a given accepted contract, and hence the relationship between the payoffs of the winning principal and the agent. Hence we can without loss of generality think of the game as the principals competing in the $\left(U_i^{\theta\widehat{\theta}}\right)_{\widehat{\theta}\in\{H,L\}}$ dimension.

Since the goal of the paper is to provide a welfare explanation for the adoption of representative democracy, we need to define some notion of welfare. Let $\mu \in [0,1]$ be the welfare weight society places on the winning politician (principal). For a given welfare weight μ, we denote the total welfare of an equilibrium in which the agent gets some payoff or utility level x as $\mathcal{W}(x;\mu)$[15]. $\mathcal{W}(x;\mu)$ again follows from Lemma 1 and the feasibility constraint; details are in Appendix C. $\mathcal{W}^S$ and $\mathcal{W}^P$ (with dependence on x and μ suppressed for brevity) refer to welfare ranges for the sets of separating and pooling equilibria, respectively.

To summarize the results in Appendix C, welfare is increasing in the payoff of the voter when μ is sufficiently small, or for intermediate ranges of μ and η sufficiently large. Intuitively, when society places non-negligible weight on the winning politician (μ is not too small), welfare aligns with the voter's payoff if and only if his preference for the public goods is sufficiently state-specific. In this case, since the voter has intense preference for one of the two public goods, total welfare is higher when the "right" type of good (the type that the voter prefers given the state) is offered.

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

First, in the interest of brevity, two components of the equilibrium description are omitted throughout the paper:

(1) The agent's strategy – the agent's action in the game is trivial; his beliefs, on the other hand, are extremely important. Though omitted, it is understood that the agent's choice of accepting or rejecting offers is sequentially rational given his beliefs: he accepts the contract that gives him higher expected utility given his updated beliefs, provided that this contract also gives him above reservation utility.

(2) On-path beliefs – using Bayes' Rule, the agent'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 $\left(U_i^{\theta H}, U_i^{\theta L}\right)$. Throughout the paper, when the relevant subscripts or superscripts are not included, we imply a vector of payoffs which includes all elements of the omitted sub/superscripts[16].

Finally, we consider contracts that have the following standard properties:

[15] We are using the notion of *ex ante* welfare here – that is, in addition to μ, we average welfare over different types using prior beliefs.

[16] For example, U denotes a vector of $\left(U_i^{\theta\widehat{\theta}}\right)_{\widehat{\theta}\in\{H,L\},\ \theta\in\{H,L\},\ i=\{1,2\}}$, whereas U_i denotes vector $\left(U_i^{\theta\widehat{\theta}}\right)_{\widehat{\theta}\in\{H,L\},\ \theta\in\{H,L\}}$ for Pi.

1. $V_i^{\widehat{\theta}}\left(U_i^{\theta\widehat{\theta}}\right) \geq 0$ for at least one $\widehat{\theta}$ **(IR-P)**

IR-P is the individual rationality of the principal. $V_i^{\widehat{\theta}}(\cdot)$ is a function (see Appendix B for details) that gives $Pi\widehat{\theta}$'s payoff for a level of utility promised to the agent, conditional on the contract being accepted. The individual rationality constraint says that no principal can offer a menu contract that gives her negative payoffs for *both* parts of the contract, should it be accepted and implemented. Any such contract will be dominated by some other contract in which the principal breaks even for at least one part of the menu.

2. $\pi U_i^{\theta H} + (1-\pi)U_i^{\theta L} \geq 0$ **(IR-A[π])**

π is the agent's posterior belief of $pr(\theta = H)$. Individual rationality must also hold for the agent, but since the agent does not know the state of the world, he calculates his expected utility given his beliefs (hence IR-A is a function of π). While we defined the utility of the agent to include the ideological component c_i, for the purpose of IR-A, c_i is excluded from the calculation of $U_i^{\theta\hat{\theta}}$. This is to avoid having to consider random reservation levels for IR-A. In our application, IR-A can be viewed as a normalized constitutional guarantee for the voters' basic rights – these rights are usually defined with respect to tangible components of policies, and should not depend on each voter's realized ideology.

3. $V_i^{\theta}\left(U_i^{\theta\theta}\right) \geq V_i^{\theta}\left(U_i^{\theta(-\theta)}\right)$ **(TC)**

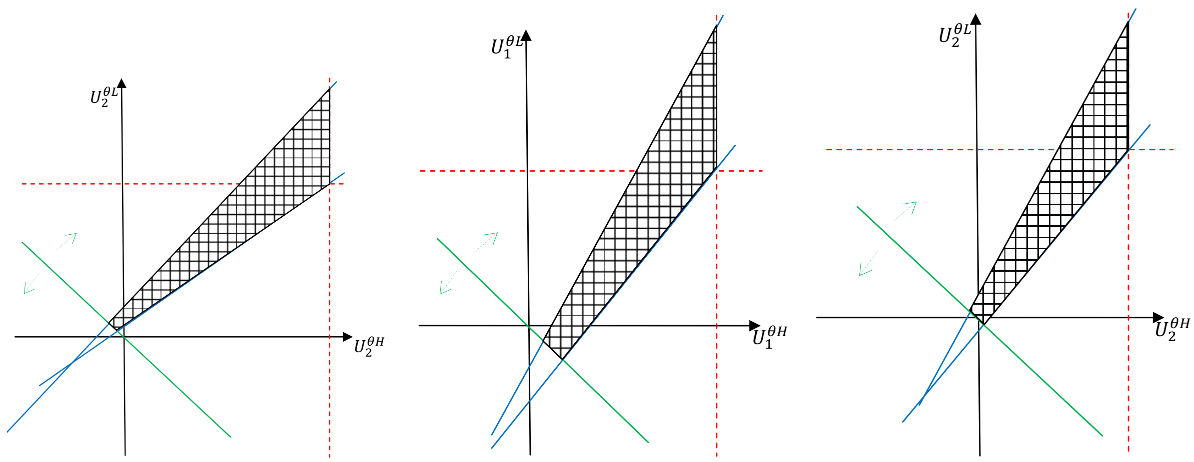
TC is the truth-telling constraint, meaning that each type of each principal will implement the part of the menu contract that corresponds to her type, so that the true state is always revealed ex post. Details of TC are given in Appendix D. TC is without loss of generality, using similar arguments as those used for the Revelation Principle.

On the subject of the Revelation Principle, it must be noted that more complex contract structures such as escalation clauses[17] 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 of commitment and detail that are unnatural in this context. Moreover, the qualitative features of the results are robust to the following extension of the current model that is similar to escalation clauses: before

[17]See Epstein and Peters (1999), Maskin and Dasgupta (2000), Peters (2001), and Martimort and Stole (2002) for examples of more complex contracts and the problem of the Revelation Principle with competing principals.

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[18].

Henceforth we will denote the set of menu contracts that satisfies IR-P, IR-A[π], and TC by $\mathcal{U}_i^\theta$ (with its dependence on the agent's beliefs π suppressed), with the corresponding vector notations $\mathcal{U}_i$, $\mathcal{U}^\theta$, or $\mathcal{U}$ as before. We can graphically represent the contract space in which the two principals compete. Recall that the contract terms can be summarized by the agent's utility conditional on the contract being accepted (U_i^θ). Since we consider menu contracts, $Pi\theta$'s contract is given by the pair $\left(U_i^{\theta H}, U_i^{\theta L}\right)$. A generic menu contract will therefore span a two-dimensional space. Qualitative representations of each principal's contract space are given in Figure 1.



(a) $P2: \ \eta < \frac{1}{2-\lambda}$ (b) $P1: \ \eta \in \left[\frac{1}{2}, 1\right]$ (c) $P2: \ \eta \geq \frac{1}{2-\lambda}$

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

IR-P puts maxima on the agent's utility for a given contract, with the bounds given by *either* of the dotted lines. Recall from our earlier discussion that IR-P rules out menu contracts with which the principal gets negative utility for both parts of the contract. Graphically, IR-P rules out the northeast quadrant of the dotted lines. IR-A[π] is given by the negatively-sloped line. The agent'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

[18]The results also do not depend on which principal being the first or last to announce her platform.

believes $\pi = 1$). A feasible contract must lie northeast of IR-A[π]. 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 $\left|U_i^{\theta H} - U_i^{\theta L}\right|$ is too large, $Pi\theta$ will have an incentive to implement the part of the contract that yields her the higher utility regardless of the true state.

4 Welfare-Dominant Pooling Equilibria

With the basic machinery in place, we now proceed to examine the equilibrium and welfare implications of our model. The goal is to identify conditions for which pooling equilibria result in higher welfare relative to the entire set of separating equilibria. Provided that these conditions are met, elected representation may be justified on welfare grounds.

Proposition 1 (**Welfare-Dominating Pooling Equilibria**). $\forall\, \mu \in (0, 1]$, $\exists\, (F(\cdot), \gamma, \lambda, \alpha, \eta)$ *such that*

$$\inf \mathcal{W}^P \geq \sup \mathcal{W}^S$$

Proposition 1 says provided that society places *some* weight on the winning politician's payoff (μ 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 representative democracy: if a society is better off in a system with ex ante commitment to a policy menu than one in which policy is chosen after the relevant information is known, then representative democracy – an institution in which authority is delegated to elected officials – can be viewed as a commitment mechanism adopted on the basis of welfare considerations.

Maskin and Tirole (1990) (henceforth MT90) show the existence of Pareto superior pooling equilibria (with respect to all types of the principal) in the case of private values. Intuitively, when the relevant information does not pose a conflict between the two types of principal – when information does not enter separately in the agent's utility function – the different types of principal can "trade" slacks in IR-A and TC constraints in a pooling equilibrium, since these constraints need to hold only in expectation, instead of state-by-state in a separating equilibrium. Loosening these constraints simultaneously allows all types of the principal to obtain a higher payoff relative to the set of separating equilibria.

The key difference between this paper and MT90 is that 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 equilibrium. What we will proceed to show is that the insight of MT90 still applies for an appropriate range of preference parameters and the voter's ideological predisposition.

To understand how we obtain Proposition 1, let's 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?[19] In our model, the level of policy competition is given by the intensity of the voter's ideological predisposition, parameterized by ψ. Intuitively, the intensity of policy competition determines how the surplus is split between the winning politician and the voters. 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 proceed by first considering the case of pure ideology, then prove upper hemicontinuity of the equilibrium correspondence for the more general case.

4.1 Purely Ideologically Driven Voter

First, consider the equivalent of the one-informed-principal model: an agent whose preference is overwhelmingly dominated by his ideological predisposition. Suppose ψ is sufficiently large, so that the agent will accept principal 1's contract if $c < 0$, and principal 2's if $c > 0$[20]. This is a case in which there is no competition along the policy dimension for either principal.

The case of pure ideology in this paper is an application of MT90. 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 should the draw of c turn out in her favor[21]. The proposition below follows from Proposition 1

[19]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 median voter's posterior belief is $pr\left(\hat{\theta}\right) > 0$.

[20]A sufficient condition is $|\psi c| > \max\left\{U_i^{\theta\hat{\theta}} : \text{IR-P, IR-A, and TC hold}\right\} \quad \forall c \in supp\ F(c)$.

[21]IR-P does not bind and TC is implicitly assumed here; see Section 3.

of MT90; one can readily check that all the assumptions needed for MT90 are satisfied in this model.

Proposition 0 (**Maskin & Tirole (1990)**). *For all i, θ, and $\hat{\theta}$, let $\bar{U}_i^{\theta\hat{\theta}}$ be such that $V_i^{\hat{\theta}}\left(\bar{U}_i^{\theta\hat{\theta}}\right) \equiv 0$. If $\left[-\max_{\hat{\theta}} \bar{U}_2^{\theta\hat{\theta}}\ ,\ \max_{\hat{\theta}} \bar{U}_1^{\theta\hat{\theta}}\right] \notin$ supp $F(\cdot)$, then $\exists(\gamma, \lambda, \alpha, \eta)$ s.t.*

$$\inf \mathcal{W}^P \geq \sup \mathcal{W}^S$$

In words, $\bar{U}_i^{\theta\hat{\theta}}$ is the agent's promised payoff that corresponds to $Pi\hat{\theta}$ getting her reservation utility. If any draw of ideology necessarily lies outside the range $\left[-\max_{\hat{\theta}} \bar{U}_2^{\theta\hat{\theta}}\ ,\ \max_{\hat{\theta}} \bar{U}_1^{\theta\hat{\theta}}\right]$, neither principal can affect her probability of winning regardless of the contract she offers. The condition laid out is sufficient for there to be no policy competition between the principals.

A proof specific to our model, as well as details of the equilibrium characterization, are given in an earlier version of this paper, but skipped for brevity and to focus on our key results. Consistent with MT90, Proposition 0 states that for an appropriate range of parameter values, the set of pooling equilibria welfare-dominates the set of separating equilibria.

4.2 A Voter Both Ideologically- and Policy-Driven

Next, consider the intermediate case in which the median voter cares about a mixture of ideology and policy. That is, while politicians' policy platforms alone do not determine who wins the election, they have a positive probability of affecting each voter's choice over the two politicians.

Since politicians only know the distribution of the ideological component, if we assume that $F(\cdot)$ is a continuously differentiable function with density $f(\cdot)$, then offering a contract that yields the pivotal voter infinitestimally 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 0. 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 2 (UHC of PBE Correspondence). *The set of Perfect Bayesian Equilibria is upper hemicontinuous in ψ.*

Recall the upper hemicontinuity of the equilibrium correspondence implies that at points of ψ 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 $\psi^n \to \check{\psi}$. Since the welfare comparison in the case of pure ideology is $\inf \mathcal{W}^P \geq \sup \mathcal{W}^S$, even if lower hemicontinuity fails and the set of equilibria "very close to" the case of pure ideology is smaller, the welfare comparison between the sets of separating and pooling equilibria will still hold. Therefore, provided that the median voter is sufficiently ideologically 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, unfortunately, is 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 (available from the author upon request) 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 median 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 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). In a pooling equilibrium, however, since each politician must offer the same policy menu across the two states, the same principal must win in both states. It also means that there is always one preference-misaligned politician that is winning in one state. Purely from the voter's perspective the pooling equilibria seem worse than the separating, although the multiplicity of equilibria makes a direct comparison inconclusive.

The calculation and comparison of aggregate welfare are further complicated by the fact that different principals win in each class of equilibria. A lower equilibrium payoff for the agent does not necessarily imply a higher equilibrium payoff for the winning principal, since the identity of the winner may be different in each class. This problem does not arise in the case of pure ideology, because once we fix a realized ideological predisposition, 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 .

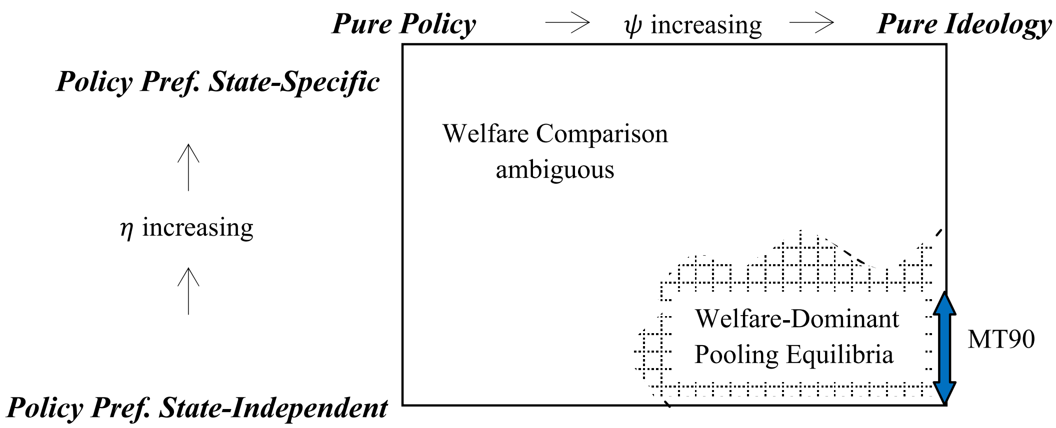


Figure 2: Summary of the Welfare Implications of the Model.

Figure 2 lays out the welfare results of the model. The model spans two dimensions parameterized by ψ and η: ψ denotes the intensity of policy competition, while η describes the level of state-specificity for the voter's public good preference. The thick two-sided arrow illustrates where Proposition 0 (MT90) applies – in the case of pure policy, given an appropriate set of parameter configurations. Upper hemicontinuity of the equilibrium correspondence implies the existence of welfare-dominant pooling equilibria for an area sufficiently close to where Proposition 0 applies. Finally, 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.

4.3 Equilibrium Refinement

A major concern of sequential equilibria in signaling games, especially those that are obtained using "punishment by beliefs", is the huge multiplicity of equilibria, including some that seem unreasonable and are only sustainable using unappealing beliefs. This spurs the need to discuss refinement of the Perfect Bayesian Equilibrium, most notably the Intuitive Criterion of Cho and Kreps (1987). Specific to our model, we would like to know whether the sets of separating and pooling equilibria satisfying the Intuitive Criterion are non-empty, and whether applying the Intuitive Criterion would destroy the results established earlier.

The Intuitive Criterion posits the following question about the equilibrium. Upon observing an out-of-equilibrium action by a principal, the agent asks, "Is there a type such that *regardless* of beliefs that the agent may have, and that the principals know the agent will best respond given those beliefs, this type will never find it profitable to take this out-of-equilibrium action?" 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-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 the misaligned type with probability zero.

Since the set of equilibria satisfying the Intuitive Criterion is a subset of the set of Perfect Bayesian Equilibria, the fact that we have $\inf \mathcal{W}^P \geq \sup \mathcal{W}^S$ in the case of pure ideology implies that Proposition 1 is robust to the Intuitive Criterion. In fact, the Intuitive Criterion can only *strengthen* our results by possibly widening the gap between

$\inf \mathcal{W}^P$ and $\sup \mathcal{W}^S$. The remaining concern is the non-emptiness of each class of equilibria. The full description of the sets of separating and pooling equilibria given the Intuitive Criterion and the proofs are relegated to the appendix.

Proposition 2 (**Robustness with respect to the Intuitive Criterion**). *Using the Intuitive Criterion as a refinement of the equilibrium obtained in section 4, the qualitative welfare features, rankings and comparisons preserve:*

- *In the case of pure ideology,* $\forall \mu \in (0,1],\ \exists(\gamma, \lambda, \alpha, \eta)$ *s.t.* $\inf \mathcal{W}^P \geq \sup \mathcal{W}^S$
- *In the case of pure policy, the welfare ranges across the sets of separating and pooling equilibria generally overlap.*

5 Conclusion

The goal of this paper is to examine the relative merits of direct versus representative democracy. The idea that the observed adoption of institutions may be explained by welfare comparisons goes back to at least Arrow (1963). We model competition between two politicians, who are informed about a payoff-relevant state. These politicians seek to win the vote of the median voter, who has state-dependent policy preferences and state-independent ideological preferences.

This model of information asymmetry allows us to compare welfare across two classes of equilibria: separating and pooling. We argue that these two classes of equilibria parallel the systems of direct and representative democracy respectively. In a separating equilibrium, having the information fully revealed to them, voters vote directly on policy by proxy of the politician they choose. In a pooling equilibrium, voters must defer policy-making authority to the elected official, whose roles as the provider *and keeper* of information remain 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 median voter's preference for the public goods. Politicians offer divergent policy platforms if the median 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 median voter's public good specificity (η), while the latter depends on the intensity of policy competition between the two politicians (ψ). If the median voter is not very ideologically driven and 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.

The main result of this paper identifies the existence of welfare dominant pooling equilibria, and interprets it as a possible welfare justification for the adoption of representative democracy. Welfare dominant pooling equilibria exist if the median voter is sufficiently ideologically driven and has relative state-independent policy preferences. In this case, ex ante aggregate welfare is higher if voters eliminate their ability to influence policy ex post by delegating the final policy choice to the politician in office. The welfare justification is robust to having any amount of friction (cost) in the voting procedure, and equilibrium refinement using the Intuitive Criterion.

A few simplifying modeling assumptions were made in this paper. First, as mentioned we have linear utility functions and feasibility constraints; second, we assume that the principals have perfect information about the state; third, we assume full commitment for both the politicians and the voters. The first two assumptions seem relatively innocuous; qualitative features of our results would likely generalize, though possibly at the expense of a less precise equilibrium characterization and welfare comparison. Whether or not the third assumption can be relaxed would depend on the structure of non-commitment and renegotiations.

References

Anderlini, Luca and Leonardo Felli (2008), "Agency problems." In *New Palgrave Dictionary of Economics*, 2nd edition, Palgrave Macmillan, New York.

Arrow, Kenneth (1963), "Uncertainty and the welfare economics of medical care." *American Economic Review*, 53, 941–973.

Athey, Susan, Andrew Atkeson, and Patrick Kehoe (2005), "The optimal degree of discretion in monetary policy." *Econometrica*, 73, 1431–1475.

Beedham, Brian (1996), "It means government by the people, and we are the people." *Economist*, 341 (December 21), 3–4.

Buchanan, James and Gordon Tullock (1962), *The Calculus of consent: Logical Foundations of constitutional democracy*. University of Michigan Press, Ann Arbor.

Cho, In-Koo and David Kreps (1987), "Signaling games and stable equilibria." *The Quarterly Journal of Economics*, 102, 179–222.

Downs, Anthony (1957), *An Economic Theory of Democracy*. Harper & Row, New York.

Epstein, Larry and Michael Peters (1999), "A revelation principle for competing mechanisms." *Journal of Economic Theory*, 88, 119–160.

Feddersen, Timothy and Wolfgang Pesendorfer (1996), "The swing voter's curse." *American Economic Review*, 86, 408–424.

Feddersen, Timothy and Wolfgang Pesendorfer (1997), "Voting behavior and information aggregation in elections with private information." *Econometrica*, 65, 1029–1058.

Fudenberg, Drew and Jean Tirole (1991), *Game Theory*. MIT Press, Cambridge, MA.

Grossman, Gene and Elhanan Helpman (2001), *Special Interest Politics*. MIT Press, Cambridge, MA.

Grossman, Sanford and Oliver Hart (1983), "An analysis of the principal-agent problem." *Econometrica*, 51, 7–45.

Hotelling, Harold (1929), "Stability and competition." *Economic Journal*, 39, 41–57.

Kessler, Anke (2005), "Representative versus direct democracy: The role of informational asymmetries." *Public Choice*, 122, 9–38.

Lindbeck, Assar and Jorgen Weibull (1987), "Balanced-budget redistribution as the outcome of political competition." *Public Choice*, 60, 273–297.

Martimort, David and Lars Stole (2002), "The revelation and delegation principles in common agency games." *Econometrica*, 70, 1659–1673.

Maskin, Eric and Partha Dasgupta (2000), "Efficient auctions." *The Quarterly Journal of Economics*, 115, 341–388.

Maskin, Eric and Jean Tirole (1990), "The principal-agent relationship with an informed principal: The case of private values." *Econometrica*, 58, 379–409.

Maskin, Eric and Jean Tirole (1992), "The principal-agent relationship with an informed principal, II: Common values." *Econometrica*, 60, 1–42.

Matsusaka, John (2003), "Direct democracy." In *Enclyclopedia of Public Choice* (Charles K. Rowley and Friedrich Schneider, eds.), 1st edition, Springer, New York.

Mirrlees, James (1976), "The optimal structure of incentives and authority within an organization." *Bell Journal of Economics*, 7, 105–131.

Myerson, Roger (1983), "Mechanism design by an informed principal." *Econometrica*, 51, 1767–1798.

Peters, Michael (2001), "Common agency and the revelation principle." *Econometrica*, 69, 1349–1372.

Rothstein, Paul (1990), "Order restricted preferences and majority rule." *Social Choice and Welfare*, 7, 331–342.

Schultz, Christian (2008), "Information, polarization and term length in democracy." *Journal of Public Economics*, 92, 1078–1091.

Appendix

A. Proof of Lemma 1

We will first prove this lemma *assuming that the truth-telling constraint is satisfied*, then prove that for any contract in which truth-telling does not hold, there must be an alternative contract where truth-telling holds and the principal can get weakly higher payoffs, regardless of her opponent's action and the agent's beliefs. The truth-telling constraint guarantees that the principal will implement the part of the menu contract that corresponds to the true state of the world. Since the relevant notation has not been introduced, suffice to say that for all i and θ, $Pi\theta$ will get a weakly higher payoff implementing the type θ part of her menu than the other (the type $(-\theta)$) part of the menu.

Since we are considering menu contracts, provided that the truth-telling constraint holds, so that the correct part of the accepted menu contract will be implemented, the agent can calculate what his utility is from *each part of the menu contract* (i.e. given a particular state). $U_i^{\theta\hat{\theta}}$ is the agent's utility level from accepting $Pi\theta$'s contract and the $\hat{\theta}$ part of this contract is implemented. Given the agent's posterior belief π, his expected utility from $Pi\theta$'s menu contract is $\left[\pi U_i^{\theta H} + (1-\pi)U_i^{\theta L}\right]$.

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).

While the agent updates his beliefs after observing the pair of menu contracts offered, once we fix these beliefs, only his expected utility is relevant, the specific contract details are not directly so. As such, the agent selects the contract that yields him higher expected utility given his beliefs, provided that it also gives him above reservation utility.

Similarly, each type of each principal seeks to maximize her expected utility, which is given by

$$[Pr(Pi\theta \text{ wins } |U)] \left[V_i^{\theta}\left(U_i^{\theta\theta}\right)\right]$$

Fixing the agent's beliefs, offering a higher $U_i^{\theta\theta}$ weakly increases $Pi\theta$'s probability of winning; however, the corresponding $V_i^{\theta}\left(U_i^{\theta\theta}\right)$ will be lower. Each principal's objective is to offer contract terms that for *each element* of her menu contract maximizes her utility given the agent's utility.

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 $\hat{V}_i^{\theta}\left(U_i^{\theta\hat{\theta}}\right)$ respectively, should this contract be accepted and the $\hat{\theta}$ part of the menu is implemented. Now consider an alternative contract in which all other contract terms are the same as the above, except now $\tilde{t}_{\hat{i}}^{\theta\hat{\theta}} = 1$. Denote $\delta = \tilde{t}_{\hat{i}}^{\theta\hat{\theta}} - t_i^{\theta\hat{\theta}}$. Depending on the state, use δ 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 $\hat{\theta}$, θ.

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, $V_1^H = \lambda b_1^{\theta H} + p_1^{\theta H}$, we get $V_1^H = 1 - (1-\lambda) b_1^{\theta H} - g_1^{\theta H}$. Also, the agent's utility is given by $U_1^{\theta H} = \gamma \left[\eta b_1^{\theta H} + (1-\eta) g_1^{\theta H}\right] - 1$.

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[22]. 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 $b_1^{\theta H} = \frac{1}{\gamma\eta}\left(\bar{U}+1\right)$. This gives

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

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

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

$\frac{1}{\gamma(1-\eta)} > \frac{1-\lambda}{\gamma\eta}$ 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.

[22] This is the standard constrained maximization problem:

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

Finally, we need to show that we can eliminate the qualifier "assuming that the truth-telling constraint holds" without changing the results. We appeal to the Revelation Principle, which argues among other things that the truth-telling (or incentive compatibility) constraint is, indeed, without loss of generality.

For any menu contract in which truth-telling does not hold, one part of the menu is always implemented regardless of the state. Construct an alternative contract as follows: take the implemented part of the original menu, and replicate it for *both* parts (states) of the alternative menu contract. This alternative contract is payoff-equivalent to the original contract and satisfies truth-telling[23]. ■

B. Strategy and Outcome Equivalence

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

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

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

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

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

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

[23] 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. 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 "pork" (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.

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

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

C. Aggregate Welfare

The results follow directly from Lemma 1 and the feasibility constraint, by calculating the weighted sum of the principal and the agent's utility for a given accepted contract using welfare weight $\mu \in [0, 1]$.

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

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

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

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

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

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

D. The Truth-Telling Constraint

Lemma 3 describes conditions under which TC will be satisfied. The cutoffs in η 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 3 (TC). *A menu contract* $\left(U_i^{\theta H}, U_i^{\theta L}\right)$ *satisfies TC iff*

$$\begin{array}{rcl} \frac{(1-\lambda)(1-\eta)}{(\alpha-\lambda)\eta}\left(1+U_1^{\theta H}\right)-1 & \leq U_1^{\theta L} \leq & \frac{1-\eta}{\eta}\left(1+U_1^{\theta H}\right)-1 \\ \frac{(1-\lambda)}{\alpha}\left(1+U_1^{\theta H}\right)-1 & \leq U_1^{\theta L} \leq & \frac{\alpha-\lambda}{\alpha}\left(1+U_1^{\theta H}\right)-1 \end{array} \quad if \quad \begin{array}{ll} \eta & < \frac{\alpha}{2\alpha-\lambda} \\ \eta & \geq \frac{\alpha}{2\alpha-\lambda} \end{array}$$

$$\begin{array}{rcl} \frac{(1-\lambda)\eta}{(\alpha-\lambda)(1-\eta)}\left(1+U_2^{\theta H}\right)-1 & \leq U_2^{\theta L} \leq & \frac{\eta}{(1-\eta)}\left(1+U_2^{\theta H}\right)-1 \\ \frac{1}{\alpha-\lambda}\left(1+U_2^{\theta H}\right)-1 & \leq U_2^{\theta L} \leq & \frac{\alpha}{\alpha-\lambda}\left(1+U_2^{\theta H}\right)-1 \end{array} \quad if \quad \begin{array}{ll} \eta & < \frac{1}{2-\lambda} \\ \eta & \geq \frac{1}{2-\lambda} \end{array}$$

Proof. TC requires that type θ of the principal always weakly prefers to implement the θ part of the menu rather than the $(-\theta)$ part. This means any contract $\left(U_i^{\hat{\theta}\theta}\right)_{\theta\in\{H,L\}}$ must be such that

$$V_i^{\hat{\theta}}\left(U_i^{\hat{\theta}\hat{\theta}}\right) \geq V_i^{\hat{\theta}}\left(U_i^{\hat{\theta}(-\hat{\theta})}\right) \qquad \forall\hat{\theta}, \forall i$$

We will work through the case for $P1H$, and all other cases are analogous.

Recall from Lemma 1 that the $\hat{\theta} = H$ part of $P1$'s menu contract (regardless of type) will always be such that $g_1^{\theta\hat{\theta}} = 0$. The $\hat{\theta} = L$ part of $P1$'s menu contract, however, will consist of different public good offers depending on the value of η. Therefore, TC for $P1H$ will vary depending on the cutoff values of η that **P1L** has.

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

   For $P1H$ to not implement the $\theta = L$ part of the contract,

$$\begin{array}{rl} 1 - \frac{1-\lambda}{\gamma\eta}\left(U_1^{\theta H}+1\right) & \geq 1 - \frac{\alpha-\lambda}{\gamma(1-\eta)}\left(U_1^{\theta L}+1\right) \\ U_1^{\theta L} & \geq \frac{(1-\lambda)(1-\eta)}{(\alpha-\lambda)\eta}\left(1+U_1^{\theta H}\right)-1 \end{array}$$

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

   For $P1H$ to not implement the $\theta = L$ part of the contract,

$$\begin{array}{rl} 1 - \frac{1-\lambda}{\gamma\eta}\left(U_1^{\theta H}+1\right) & \geq 1 - \frac{\alpha}{\gamma\eta}\left(U_1^{\theta L}+1\right) \\ U_1^{\theta L} & \geq \frac{1-\lambda}{\alpha}\left(1+U_1^{\theta H}\right)-1 \end{array}$$

These two conditions give lower bounds for $U_1^{\theta L}$ for each η range. The TC constraints for $P1L$ (so that P1L will implement the $\theta = L$ instead of $\theta = H$ part of the menu) give upper bounds for $U_1^{\theta L}$. ■

E. Proof of Lemma 2

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

Suppose by way of contradiction that $(U^*, \pi^*, \widetilde{\pi})$ is *not* an equilibrium given ψ. 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 π_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) must hold, the fact that $(U^*, \pi^*, \widetilde{\pi})$ is not an equilibrium must be because at least one principal has an incentive to deviate and obtain a strictly higher payoff, given all other components of the putative equilibrium. Since the principals' payoffs are continuous in ψ, there must exist some $\psi^{\bar{m}}$ in the sequence converging to ψ in which the same deviation would also be profitable for that principal. This contradicts the fact that $(U^{*\bar{m}}, \pi^{*\bar{m}}, \widetilde{\pi}^{\bar{m}})$ is an equilibrium given $\psi^{\bar{m}}$.

Thus concludes our proof. ∎

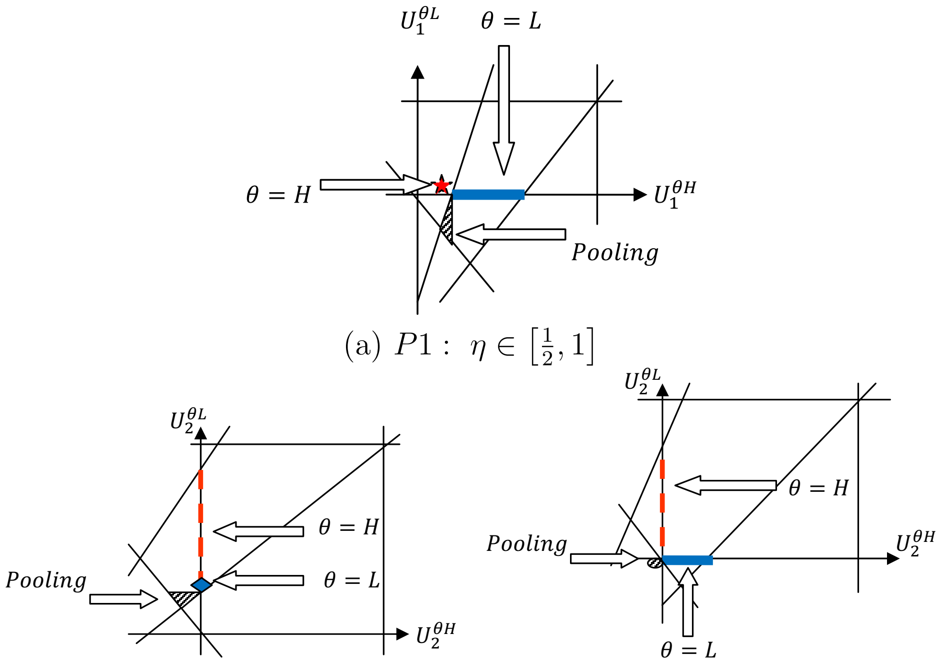
F. Proof of Proposition 2

The sets of separating and pooling equilibria in the case of pure ideology are given below; for the case of pure policy, refer to an earlier draft of this paper.

- Separating Equilibrium
  As in Figure 3.
- Pooling Equilibrium [$P1$ wins]
  Let the set of Pooling equilibria as given in Figure 3 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 3 be $\mathcal{E}_2^{Pool*}$.

The set of Pooling equilibria given the Intuitive Criterion is

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



(a) $P1: \ \eta \in \left[\frac{1}{2}, 1\right]$

(b) $P2: \ \eta < \frac{1}{2-k}$

(c) $P2: \ \eta \geq \frac{1}{2-k}$

Figure 3: **[2-P Pure Ideology].** Shaded: Pooling equilibrium ; dotted line: Separating Equilibrium $\theta = H$; thick line: Separating Equilibrium $\theta = L$.

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 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 3. 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 3). 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 ($\inf \mathcal{W}^P \geq \sup \mathcal{W}^S$) for the sets of separating and pooling equilibria even if the Intuitive Criterion is used. ■