Delegates or Trustees? A Theory of Political Accountability

Justin Fox[1] and Kenneth W. Shotts[2]

July 23, 2007

[1]Assistant Professor, Department of Political Science, Yale University, 77 Prospect Street, New Haven, CT 06520. Phone: (203) 432-5724. Email: justin.fox@yale.edu.

[2]Corresponding Author. Associate Professor, Stanford GSB, 518 Memorial Way, Stanford, CA 94305. Phone: (650) 725-4510. Email: kshotts@stanford.edu.

Abstract

When do elections induce politicians to act as delegates, and when do they induce them to act as trustees? To answer this question, we develop a model of political accountability in which politicians vary in both their policy preferences and their competence. We show that elections are more likely to induce politicians to behave as trustees when uncertainty about incumbent policy preferences is low. Otherwise, voters are often unable to credibly commit to vote retrospectively, and incumbents are electorally rewarded for the positions they take as opposed to the outcomes that they generate. Our theoretical results help us understand a number of empirical puzzles, such as why voters sometimes re-elect politicians who are known not to share their preferences. The model also yields predictions about several other factors that determine whether elected officials will act as delegates or trustees.

Suppose the public believes that a particular country is helping a terrorist organization acquire weapons of mass destruction. And suppose that the president's intelligence strongly suggests otherwise. If the president fails to attack the regime, she runs the risk of appearing out of step with public opinion. Yet, if the president attacks, she runs the risk of appearing inept if it is later revealed that the accused country was innocent. What will she do – attack or not attack?

The answer to this question depends on how the public reacts when an executive's policy choice indicates that she shares the public's policy preferences, yet, at the same time, suggests that she has poor judgement. If voters reward the executive with re-election, they encourage her to act as their *delegate*, i.e., she has an electoral incentive to ignore her expert judgement about which policies promote the general welfare, and instead simply to pursue whatever policies happen to be popular at the moment. Alternatively, if voters punish the executive for having chosen an initially-popular policy that turned out to be inappropriate, they encourage her to act as their *trustee*, i.e., she has an electoral incentive to pursue policies that she believes promote the general welfare, based on her expertise. Thus, whether the public evaluates the executive based on the policies she chooses or the outcomes that her policies generate determines whether elections encourage her to behave as a delegate or a trustee.

In this research note, we determine when voters will judge politicians on the basis of policies and when they will judge politicians on the basis of outcomes. In doing so, we are able to determine how voters' actions affect the form of representation provided by elected executives. Our analysis focuses on two factors – the competence and preferences of politicians – that political theorists who study delegate and trustee representation have long recognized as playing a central role in elections. For example, in his critique of delegate representation, J.S. Mill (1962 [1861], 241) focused on the need for wise leaders, but noted that preferences also matter since voters "cannot be expected to postpone their particular opinions, unless in order that they be served by a person of superior

knowledge to their own."[1]

Formal theorists have also recognized the importance of preferences and competence, but existing models of accountability analyze these factors in isolation from each other.[2] In preference-based models (e.g., Morris 2001, Maskin and Tirole 2004, Fox 2007, Canes-Wrone and Shotts 2007), executives have incentives to act as delegates, and may shy away from pursuing policies that are in the public interest. In competence-based models (e.g., Canes-Wrone, Herron, and Shotts 2001; Prat 2005), when the appropriateness of an executive's policy choice is revealed before the next election, citizens vote on the basis of whether the policy generated good outcomes. That is, they employ outcome-based *retrospective voting*, taking "past performance as a prima facie indicator of the government's judgement and competence (or lack thereof)" (Fiorina 1981, 12).

Here, we develop a new model of executive policymaking, in which politicians vary in both their policy preferences and their competence. We use the model to analyze a question that cannot be addressed by previous models: how do voters' perceptions of an incumbent's preferences interact with their perceptions of her competence to determine the form of representation she provides?

Our analysis produces several novel results. Specifically, we show that citizens are most likely to make electoral decisions on the basis of outcomes, as opposed to policy positions, when uncertainty about the executive's policy preferences is minimal. The first implication of this result

---

[1]See Stokes 2001, for an excellent discussion of the thoughts of Edmund Burke, James Madison, Bernard Manin, and John Stuart Mill on whether citizens should evaluate candidates based on competence or preferences.

[2]One line of research analyzes policy competition and valence (Ansolabehere and Snyder 2000, Groseclose 2001, Aragones and Palfrey 2004). But politicians in valence models lack private information about optimal policies and precommit to platforms, so the models cannot be used to analyze retrospective voting and elected officials' roles as trustees or delegates.

is that when voters are very confident that the executive shares their preferences, elections will induce her to behave as a trustee. This is broadly consistent with previous analyses, in which preference congruence between a policymaker and her constituents promotes trust (Bianco 1994) and communication (Crawford and Sobel 1982; Mansbridge 1999). However, the second implication of the result is more surprising: elections can also promote trustee representation when the public is quite certain that the incumbent does *not* share its preferences. Accordingly, an increase in the probability that politicians share the public's preferences may fail to increase, and actually may decrease, the probability that the incumbent's policy choices promote the public interest. We also analyze how the form of representation – delegate or trustee – is affected by factors such as the probability that a politician is competent, her level of ambition, and the probability that future policy decisions will be on matters of common value.

The Model

We analyze a two-period model, with one policy decision in each period. The incumbent politician must stand for re-election after the first period, and thus can be held accountable for her actions. Voters are modelled as a single representative citizen. The citizen wants the incumbent to select a policy in each period that promotes his interests, but ultimately he seeks to elect a politician who will choose good policies in the future.[3]

When deciding whether to re-elect the incumbent, the citizen thus uses all information available at the time of the election, including the incumbent's policy choice as well as information about the effects of the policy choice. We assume that the citizen perfectly learns before the next election whether the incumbent's policy choice promoted his interests. This assumption stacks the deck in favor of the citizen being able to use the electoral process to give the incumbent an incentive to act

---

[3]We use male pronouns for citizens and female pronouns for politicians.

in the public interest.

Issues and Policies. On some issues, all actors share the same interests – these are issues of *common value*. On other issues, actors have different policy aims – these are issues of *private values*. In a given period, the probability that the government takes up a common value issue is β, and all actors know whether the policy issue is one of private values or common value. We focus on the case where the first period involves a private values issue, since the common value case for the first period is similar to previous work (Canes-Wrone, Herron, and Shotts 2001).

Once Nature determines the type of issue to be considered in a given period, the office holder selects one of two policies, $p \in \{A, B\}$. Which policy benefits the citizen depends on the underlying state of the world, which is determined by Nature. The state of world ω is either A or B, with $\Pr(\omega = A) = \gamma \in \left(\frac{1}{2}, 1\right)$. Policy p benefits the citizen when $p = \omega$, i.e., the citizen wants policy to match the state of the world. When the executive does this, we say that her policy choice is *appropriate*.

Politicians. Politicians differ along three characteristics: policy preferences, competence, and desire for holding office. These characteristics determine a politician's type, which we denote by (t_p, t_c, t_o). On each of these characteristics, a politician can be one of two types, as follows.

All politicians share the public's policy preferences on matters of common value, but some politicians have preferences that diverge from the public's on matters of private values. In particular, these divergent politicians have preferences that are independent of the state of the world, with a bias in favor of policy B.[4] Such politicians are *non-congruent*, denoted $t_p = n$, whereas politicians who share the public's preferences on issues of private values are *congruent*, denoted $t_p = c$. All else equal, the citizen prefers congruent politicians.

Politicians also differ in their competence. Specifically, some can ascertain perfectly the state

---

[4]Our results can be extended to the case where this bias is in favor of policy A.

of the world, and others are less able to do so. We refer to the former as *skilled*, denoted $t_c = s$, and the latter as *unskilled*, denoted $t_c = u$. All else equal, the citizen prefers skilled politicians.

Asymmetries in skill affect policy choice as follows: At the start of each period, the office holder observes a signal $s \in \{A, B\}$ about the state of the world, the accuracy of which depends on her competence. For a skilled politician, $s = \omega$, i.e., the state of the world is revealed to her with certainty. In contrast, an unskilled politician receives a signal that is informative but noisy, where $\Pr(s = \omega) = q$. We assume that $q > \gamma$, so that for either $s = A$ or $s = B$, Bayes's Rule implies that $\Pr(\omega = s | s) > \frac{1}{2}$, i.e., the executive believes that her signal is probably correct. Thus, a politician aiming to choose the appropriate policy $(p = \omega)$ maximizes her chances of doing so by following her signal.

Finally, politicians vary in the weight they attach to re-election. Those who care about re-election are *ambitious*, whereas those who do not care about re-election are *unambitious*. These are polar types; an ambitious politician is a single-minded seeker of re-election, whereas an unambitious politician cares only about first period policy.

We now summarize the incumbent's preferences. If she is ambitious, she gets 1 unit of utility if re-elected and 0 otherwise. If she is unambitious and congruent, she gets 1 unit of utility if the first period policy is $p = \omega$ and 0 otherwise, whereas if she is unambitious and non-congruent, she gets 1 unit of utility if $p = B$ and 0 otherwise. To reduce mathematical clutter, we do not include second period policy in the politician's utility function, and instead simply assume that the second period politician chooses the policy she most prefers.[5]

---

[5]To be precise, the incumbent has lexicographic preferences. An ambitious politician's first priority is to maximize her probability of winning re-election, and her second priority is second period policy. An unambitious politician's first priority is first period policy, and her second priority is second period policy.

A politician's congruence, skill, and ambition are naturally known better by the politician herself than by the public, so we treat each politician's type as private information. Let $\phi \in (0,1), \theta \in (0,1)$, and $\alpha \in (0,1)$ be the prior probabilities that she is congruent, skilled, and ambitious, respectively. Both the incumbent and the challenger are drawn from the same pool of types, and a politician's three characteristics are drawn independently.

The sequence of the model is as follows: (1) Nature draws the incumbent's type and the challenger's type. (2) Nature draws the first period state of the world ω, and the incumbent's signal s. (3) The incumbent chooses policy p. (4) The citizen observes p, learns ω, and decides whether to re-elect the incumbent. (5) Nature determines the type of issue to be considered in the second period; additionally, Nature draws the second period state of the world, and the second period office holder's signal. (6) The second period office holder chooses policy.

Results

As mentioned in the introduction, the literature on political representation has explored in depth two distinct relationships a policymaker can have with her constituents: she can act as their *trustee*, selecting policy based on her best judgement about how to promote the public interest, or she can act as a *delegate*, selecting policy based on the public's (potentially-mistaken) beliefs about the policy that best serves their interests.

Philosophers have debated the normative merits of each form of representation for centuries. In our model, trustee representation is clearly preferable, in terms of promoting good first-period policy decisions, since the citizen always wants the elected official to match $p = \omega$ and the best way for her to do this is by setting $p = s$. However, the mode of representation she adopts depends on electoral incentives. We show that even though the citizen would like a trustee, in equilibrium, elections often give the incumbent an incentive to act as a delegate. We refer to this phenomenon

6

as the *delegate trap*.

To determine when the delegate trap arises, we proceed in three steps. First, we describe the citizen's selection problem as he decides whether to re-elect the incumbent based on her policy choice and the appropriateness of her policy choice. Second, we demonstrate that our model has two types of equilibria, one in which the incumbent has an electoral incentive to act as a trustee and one in which she has an electoral incentive to act a delegate. Third, we characterize how equilibrium electoral incentives depend on the probabilities that politicians are congruent, skilled, and ambitious, as well as the probability that an issue of common value arises in the second period.

The Citizen's Selection Problem

The citizen's expected second-period payoffs are determined by the election winner's type, particularly her competence and preferences. Since this is a two period game, the election winner simply selects her preferred policy in the second period. With probability β, the second period issue is one of common value, and the winner attempts to match policy to the state. With probability $(1 - \beta)$ the second period issue is one of private values; in such situations, congruent politicians attempt to match policy to the state, and non-congruent politicians simply select policy B. Let $y(t_p, t_c)$ denote the citizen's expected second-period payoff from electing a politician whose policy preferences and competence are characterized by t_p and t_c, respectively.

The ideal type of politician, from the voter's perspective, is one who is both congruent and skilled, because a (c, s) politician always selects the appropriate policy, thereby yielding $y(c, s) = 1$. The worst type of politician is non-congruent and unskilled. When the second period issue is one of common value, a (n, u) politician selects the appropriate policy with probability q, and when the issue involves private values, she selects the appropriate policy only when the state of the world is B, which occurs with probability $(1 - \gamma)$, so that $y(n, u) = \beta q + (1 - \beta)(1 - \gamma)$. Similar

7

derivations yield $y(c,n) = q$ and $y(n,s) = \beta + (1-\beta)(1-\gamma)$. The citizen's preference between a congruent and unskilled (c,u) versus non-congruent and skilled (n,s) politician depends on the model's parameters.

Recall that the citizen does not directly observe the incumbent's or challenger's type. However, at the time that he votes, he has observed the incumbent's policy choice and the first-period state of the world. Hence, the citizen updates his prior about the incumbent's type using his knowledge of the incumbent's strategy, her policy choice, and the realized state of the world. Using this updated belief, the citizen calculates his expected payoff from re-electing the incumbent. If his utility from the incumbent is strictly greater than that from electing the challenger, the citizen keeps the incumbent; otherwise, he replaces her.[6] Before proceeding to our main results, we state a lemma that clarifies the citizen's selection problem.

Lemma 1 *Fix the incumbent's strategy and fix the state of the world ω. Then either **(1)** the voter strictly prefers to re-elect the incumbent if $p = A$ and to remove her if $p = B$, **(2)** the voter strictly prefers to re-elect the incumbent if $p = B$ and to remove her if $p = A$, or **(3)** the voter is indifferent both when $p = A$ and when $p = B$.*

Details of all proofs are in the appendix. The intuition for this lemma is the following. The incumbent's type is independent of the state of the world. Thus, given ω, if the voter updates

---

[6]To be precise, our solution concept is Perfect Bayesian Equilibrium. In evaluating a politician, the citizen takes the expectation of his second-period utility $y(\cdot, \cdot)$ with respect to his belief about the politician's type. Write $\pi(t_p, t_c | p, \omega)$ for the probability the citizen assigns to the incumbent's type being (t_p, t_c) when the incumbent's policy choice is p and the state of the world is ω. Having observed (p, ω), the citizen's expected payoff from re-electing the incumbent is $U_I(p, \omega) \equiv \sum_{t_p, t_c} y(t_p, t_c) \pi(t_p, t_c | p, \omega)$. The citizen re-elects the incumbent if $U_I(p, \omega)$ is greater than that from electing the challenger, $U_{CH} \equiv y(c,s)\phi\theta + y(c,u)\phi(1-\theta) + y(n,s)(1-\phi)\theta + y(n,u)(1-\phi)(1-\theta)$.

positively on the incumbent when $p = A$, he must update negatively when $p = B$, and vice versa. And since the incumbent and challenger are drawn from the same pool, this means that for any given state of the world, at most one policy, A or B, can guarantee re-election.

Delegate and Trustee Equilibria

We focus on pure strategy equilibria, which always exist in our model. From Lemma 1, we know there are four possible types of citizen behavior: re-elect if and only if $p = A$, re-elect if and only if $p = \omega$, re-elect if and only if $p = B$, and re-elect if and only if $p \neq \omega$. We focus on the first two types, at least one of which occurs in equilibrium for any parameter values.[7] Clearly, which type of equilibrium occurs will have a major effect on electoral incentives, which in turn determine the policymaking behavior of electorally ambitious incumbents.

In the first type of equilibrium, the citizen selects on the basis of policy, re-electing the incumbent whenever she chooses the policy that the citizen believes is best, $p = A$. Thus, the incumbent maximizes her chances of re-election by following public opinion, always choosing policy A, even though doing so when her signal is B requires her to go against her expert judgement about how best to promote the citizen's interests. Because the incumbent is electorally rewarded for simply following the public's prior belief about which policy is best, we call this a *delegate equilibrium.*

---

[7]There cannot exist an equilibrium in which the citizen re-elects if and only if $p = B$, since only non-ambitious congruent incumbents would ever choose $p = A$, and the voter would prefer to re-elect when $p = \omega = A$. An equilibrium in which the citizen re-elects if and only if $p \neq \omega$ can exist if politicians are sufficiently likely to be ambitious. Such an equilibrium is similar to the equilibrium in which the incumbent wins re-election if and only if $p = \omega$, with one perverse, and empirically implausible, difference – the incumbent strives to show that she is high quality by picking the *wrong* policy.

In the second type of equilibrium, the citizen votes retrospectively, on the basis of performance. Since the incumbent is re-elected when her policy choice matches the state of the world, she maximizes her re-election prospects by following her private signal, choosing policy $p = s$. Because the incumbent is electorally rewarded for using her expert judgement when choosing policy, we call this a *trustee equilibrium.* The following lemmas, which follow from our preceding discussion, summarize behavior in delegate and trustee equilibria.

Lemma 2 *(Delegate Equilibria) In an equilibrium where the citizen re-elects if and only if the incumbent selects policy A:*

1. *All ambitious politicians select policy A regardless of their signals.*

2. *All unambitious politicians select their preferred policy; congruent politicians follow their signals and non-congruent politicians select policy B.*

Lemma 3 *(Trustee Equilibria) In an equilibrium where the citizen votes retrospectively, re-electing if and only if the incumbent's policy choice was appropriate $(p = \omega)$:*

1. *All ambitious politicians follow their signals.*

2. *All unambitious politicians select their preferred policy; congruent politicians follow their signals and non-congruent politicians select policy B.*

Representation and The Delegate Trap

We now characterize equilibria for our model and show that the citizen's inability to credibly commit to an outcome-based retrospective voting rule can give politicians an incentive to behave as delegates even though the public would prefer them to act as trustees; this is the delegate trap.

The main result of this section demonstrates that the delegate trap can be avoided – i.e., trustee equilibria exist – only if the fraction of skilled politicians in the candidate pool is sufficiently large.

To formally state this result, we use the following notation: let

$$g(\alpha, \beta, \phi) \equiv \max \left\{ \frac{(1-\phi)[y(c,u) - y(n,u)]}{\phi[y(c,s) - y(c,u)] + (1-\phi)[y(n,s) - y(n,u)]} \cdot \frac{\phi(1-\alpha)}{\phi + \alpha(1-\phi)}, 1 \right\},$$

and note that g is non-increasing in the fraction of ambitious politicians in the candidate pool (α), so for all $\alpha \in (0,1)$, $0 < g(\alpha, \beta, \phi) \leq g(0, \beta, \phi)$.

Proposition 1 *(Existence) There exists a delegate equilibrium if and only if $\theta \leq g(0, \beta, \phi)$, and there exists a trustee equilibrium if and only if $\theta \geq g(\alpha, \beta, \phi)$.*

[[Figure 1 about here]]

As shown in Figure 1, the minimum proportion of skilled politicians necessary to sustain a trustee equilibrium is $g(\alpha, \beta, \phi)$, and the maximum proportion that can sustain a delegate equilibrium is $g(0, \beta, \phi)$.[8] Thus, only when politicians are sufficiently likely to be skilled will the public judge the incumbent on the basis of the consequences of her policy choice, as opposed to the policy itself. Otherwise, outcome-based retrospective voting is not time-consistent, and the delegate trap arises: politicians have an incentive to blindly follow public opinion despite the fact that the citizen would prefer that they use their expertise.

---

[8]For intermediate proportions of skilled politicians, i.e., $\theta \in [g(\alpha, \beta, \phi), g(0, \beta, \phi)]$, it is possible to have either a trustee equilibrium or a delegate equilibrium. In such circumstances, if the citizen expects the incumbent to use her private information to promote his interests, then ambitious incumbents will indeed do so, anticipating electoral rewards whenever $p = \omega$. If, however, the citizen expects the incumbent to ignore her private information and posture in an attempt to prove that she shares his preferences, then ambitious incumbents will indeed do so, setting $p = A$.

11

To see why outcome-based retrospective voting, and hence a trustee equilibrium, cannot be sustained when politicians are unlikely to be skilled, suppose that politicians expect to be rewarded when they choose appropriate policies and punished when they choose inappropriate policies. Given these electoral incentives, all incumbents will follow their signals, with the exception of unambitious non-congruent ones, who select policy B regardless of their private information.

Note that retrospective voting requires the citizen to remove an incumbent who inappropriately chooses policy $p = A$ when the state of the world is $\omega = B$. However, given the policymaking behavior that arises when incumbents expect to be judged on the basis of outcomes, the citizen will not be willing to vote retrospectively when skill is a sufficiently rare trait among politicians ($\theta < g(\alpha, \beta, \phi)$).

The reason the citizen might stand by an incumbent who chooses policy A inappropriately is that although he updates *unfavorably* about her competence because her policy choice was inappropriate, he updates *favorably* about her congruence because her policy was in step with public opinion. Hence, from the citizen's perspective, when $p = A$ and $\omega = B$, removing the incumbent is is a risky gamble, with both a potential upside – the challenger may turn out to be both congruent and skilled – and a potential downside – the challenger may turn out to be both non-congruent and unskilled. And when skill is rare, i.e., θ is low, the citizen's concern about downgrading on congruence outweighs the potential benefit of upgrading on competence.

Thus, when skill is a rare trait, the citizen cannot credibly commit to outcome-based retrospective voting. As a result, electoral decisions are based solely on the incumbent's policy choices, regardless of whether these policy choices produced good results. In sum, when politicians are unlikely to be skilled, policymaking degenerates into feckless posturing and position taking, since executives are rewarded for pandering to public opinion even when they believe that voters misperceive their true interests.

Comparative Statics

We now examine how $g(\alpha, \beta, \phi)$, the number of skilled politicians in the candidate pool necessary to sustain a trustee equilibrium, depends on other parameters of the model.

Proposition 2 *(Comparative Statics) The minimum fraction of skilled politician in the candidate pool necessary to sustain a trustee equilibrium is **(1)** decreasing in the fraction of ambitious politicians in the candidate pool (α), **(2)** decreasing in the likelihood that future issues are of common value (β), and **(3)** single peaked in the fraction of congruent politicians in the candidate pool (ϕ), with $\lim_{\phi \to 0} g(\alpha, \beta, \phi) = \lim_{\phi \to 1} g(\alpha, \beta, \phi) = 0$.*

Before we explain why this proposition holds, note that in any trustee equilibrium, three conditions must be satisfied. First, all politicians follow their respective signals, with the exception of unambitious non-congruent ones, who rigidly select policy B. Second, the citizen votes retrospectively when the state of the world is A. Third, the citizen votes retrospectively when the state of the world is B. Only the third condition is sensitive to the model's parametrization. By Lemma 1, this condition holds if the citizen's expected payoff from re-electing the incumbent upon observing policy $p = A$ chosen inappropriately, i.e., when the state is $\omega = B$, is less than his expected payoff from the challenger.[9]

When all incumbents except the non-congruent and unambitious follow their signals, as is the case in a trustee equilibrium, a citizen who observes $p = A$ and $\omega = B$ draws a favorable inference about the incumbent's congruence and an unfavorable inference about her skill. Thus, in deciding whether to replace the incumbent when $p = A$ and $\omega = B$, the citizen must weigh the cost of downgrading on congruence against the benefit of upgrading on skill. Accordingly, the key to understanding the model's comparative statics is to understand how variation in the model's

---

[9]Formally, in a trustee equilibrium $U_I(A, B) = \frac{\phi}{\phi + \alpha(1-\phi)} y(c, u) + \frac{\alpha(1-\phi)}{\phi + \alpha(1-\phi)} y(n, u) < U_{CH}$.

parameters influences these costs and benefits.

The Proportion of Ambitious Politicians (α)**.** In a trustee equilibrium, increasing the likelihood that the incumbent is ambitious decreases the signaling effect of selecting policy A when the state of the world is B.[10] If $\alpha = 0$, meaning that no politicians are ambitious, only politicians who are both unskilled and congruent select $p = A$ when $\omega = B$; hence, upon observing $p = A$ and $\omega = B$, the citizen knows for sure that the incumbent is congruent. However, as α increases, the probability that an unskilled and non-congruent politician selects A when the state is B in a trustee equilibrium increases, and, in turn, the weight the citizen attaches to the incumbent being congruent decreases. This reduces the citizen's concern over downgrading on congruence if she elects the challenger, so she is more willing to elect the challenger in the hopes of upgrading on skill after observing $p = A$ and $\omega = B$. Thus, the more likely politicians are to be ambitious, the lower is the skill cutoff $g(\alpha, \beta, \phi)$ necessary to sustain a trustee equilibrium.

The Probability of a Common Value Issue in the Second Period (β)**.** Increasing the likelihood that future issues are of common value increases the importance of electing a skilled politician. Hence, the greater is β, the greater is the benefit of upgrading on skill, so as β increases, $g(\alpha, \beta, \phi)$ decreases, i.e., a trustee equilibrium can be sustained for a smaller probability θ that politicians are skilled.

The Proportion of Congruent Politicians (ϕ)**.** Proposition 2 shows that $g(\alpha, \beta, \phi)$ is smallest when ϕ is very close to either 0 or 1. In other words, a trustee equilibrium is easiest to sustain when politicians are either extremely likely to be congruent or extremely unlikely to be congruent.

The former result is not surprising, given previous work. As Mansbridge (1999) notes in her analysis of race, gender, and representation, descriptive representation by politicians who are highly likely to share their constituents' preferences facilitates trust between a lawmaker and her con-

---

[10]This can be seen directly from inspection $U_I(A, B)$, which was formally defined in footnote 9.

14

stituents, and this trust enhances the lawmaker's ability to pursue policies she believes to be in her constituents' best interest. Bianco (1994) goes further; in his model, a high degree of preference congruence is not only a sufficient condition, but also a necessary condition, for an equilibrium in which the incumbent is not electorally punished when she goes against voters' prior beliefs about what policies best promote their interests.[11] Both of these works have a flavor similar to Crawford and Sobel (1982), which shows that communication between two parties becomes more difficult as their preferences diverge.

However, from Figure 1, we see that although high congruence is a *sufficient* condition for a trustee equilibrium in our model, it is not a necessary condition. Specifically, if congruence is extremely likely then the only equilibrium is a trustee equilibrium, and if congruence is extremely unlikely there exist both a delegate equilibrium and a trustee equilibrium, whereas for moderate probabilities of congruence the only equilibrium is a delegate equilibrium.[12]

---

[11]Bianco's model differs from ours in that he assumes there is no variation in incumbent competence, voters never learn whether a policy succeeded, and voters are not forward-looking. However, in insightful discussions outside the scope of his model (e.g., p. 79) he explores the implications of the fact that voters may draw inferences about an incumbent's type based on her actions.

[12]To be precise, fix $\theta < \max_{\phi \in (0,1)} g(\alpha, \beta, \phi)$. Then there exist ϕ_1 and ϕ_2, where $\phi_1 < \phi_2$, and the equilibrium depends on congruence as follows: if $\phi < \phi_1$ then there exist both a delegate equilibrium and a trustee equilibrium, if $\phi \in (\phi_1, \phi_2)$ there is only a delegate equilibrium, and if $\phi > \phi_2$ there is only a trustee equilibrium. It is interesting to note that voter welfare need not be increasing in the probability that politicians are congruent because the probability that $p = \omega$ in the first period suffers a discontinuous drop at ϕ_1 when the increase in congruence switches the equilibrium from a trustee to a delegate equilibrium.

Why is it possible to have a trustee equilibrium in the absence of congruence? When the incumbent is extremely unlikely to be congruent, it is difficult for her to signal that she is in fact congruent. Even if she picks the policy the public favors ex-ante, they remain unconvinced that she shares their preferences. However, the incumbent *can* signal that she is *competent*.[13] In other words, when the incumbent is believed to be non-congruent, she has little hope of convincing the public otherwise. This forces the incumbent to focus on attempting to generate good outcomes, to convince voters that she is competent and thus worthy of re-election.

It is interesting to note that there are examples of politicians with preferences that are clearly non-congruent with their electorates, who win re-election precisely because they have a proven track record of competence, the canonical example being Rudy Giuliani as the Republican mayor of New York City.

Conclusion

This research note develops a new model of electoral accountability, in which voters are concerned about both the competence and preferences of elected officials. The model helps clarify the relationship between retrospective voting and trustee and delegate forms of representation. Additionally, we use the model to identify conditions under which elections provide incentives for one form of representation over another. Voters in our model always prefer trustee representation, but when a politician's actions reveal information about both her policy preferences and her competence, voters are often unable to credibly commit to outcome-based retrospective voting. In such circumstances, voters judge politicians on the basis of their policy choices, as opposed to the outcomes that result

---

[13]Note that this trustee equilibrium is predicated on the public's ability to learn whether the incumbent's policy choice was appropriate; otherwise there would be no way for the incumbent to signal her competence, and the only possible equilibrium would be a delegate equilibrium.

from their choices. Consequently, politicians wind up catering to public opinion, even when doing so is not in the best interest of their constituents, a pathology we call the delegate trap.

We show that elections are most likely to provide incentives for elected officials to behave as trustees when uncertainty about politicians' policy preferences is minimal. Hence, if the voters believe executives are either *very likely* or *very unlikely* to share their preferences, they will be willing to judge incumbents on the basis of the outcomes of their policy choices, reelecting even those who choose policies that are initially unpopular, so long as their choices ultimately turn out to be appropriate. The possibility of trustee representation in the absence of preference congruence is often overlooked in contemporary scholarship, but was suggested by J.S. Mill (1962 [1861], 238) almost a hundred and fifty years ago:

> Individuals, and peoples, who are acutely sensible of the value of superior wisdom, are likely to recognize it, where it exists, by other signs than thinking exactly as they do, and even in spite of considerable differences of opinion: and when they have recognized it they will be far too desirous to secure it, at any admissible cost, to be prone to impose their own opinion as a law upon persons whom they look up to as wiser than themselves.

Although conditions exist where elected officials are rewarded for using their expert judgement, such incentives are far from inevitable in our model. In circumstances where incumbents may be rewarded for acting as trustees, there is often another equilibrium, in which incumbents are rewarded for disregarding their expert judgement and catering to public opinion; hence, there is no guarantee that the delegate trap will be avoided. Moreover, when there is substantial uncertainty about politicians' policy preferences, unless politicians are sufficiently likely to be skilled, they will always be electorally rewarded for acting as delegates.

We conclude by discussing two additional implications of our analysis.

On Institutional Reforms to Encourage Retrospective Voting. To induce officials to act as trustees, citizens must vote retrospectively, judging incumbents by the outcomes they generate as opposed to the policies they pursue. However, often in American politics, it appears that the public does not judge officials on the consequences of their policy initiatives. Almost three decades ago, Morris Fiorina (1980, 46) argued that

> We hold our politicians individually accountable for the proposals they advocate, but less so for the adoption of those proposals, and not at all for overseeing the implementation of those proposals and the evaluation of their results. In contemporary America officials do not govern, they merely posture.

Many would assert that the above quote aptly characterizes contemporary politics as well.

Fiorina attributed the lack of retrospective voting to the fact that decentralized political parties and America's system of separated powers obscure responsibility for policy outcomes. While being able to assign responsibility for outcomes is surely a necessary condition for citizens to vote retrospectively, our work demonstrates that such knowledge is not sufficient. Accordingly, attempts to resurrect retrospective voting via increased governmental accountability, whether through the civic mindedness of the electorate or via institutional reforms that decrease the incidence of divided government, may simply be of no avail.

Recall that in our model, a unitary executive has sole control over the instruments of public policy. In addition, the public always learns whether the incumbent's policy choice was appropriate. Together, these factors enable the public to assign responsibility for outcomes, yet outcome-based retrospective voting is often time-inconsistent. There are three root causes of this tension in our model: First, the executive possesses better information than the public about the appropriateness of alternative policy courses. Second, the executive's information is not necessarily perfect. Third, the public is uncertain of the executive's underlying policy preferences.

18

Little can be done to alleviate the first two causes; both are inevitable aspects of policymaking in numerous policy domains, particularly in matters of foreign policy. Thus, institutional change can only realistically address the third issue, preference uncertainty. Accordingly, in thinking about how political institutions structure electoral incentives, one must pay careful attention to their effects on the public's information about the policy preferences of lawmakers. Institutional features that reduce uncertainty about politicians' preferences encourage them to behave as trustees, whereas institutions that increase uncertainty will encourage them to behave as delegates.

For example, there has been much debate about the normative consequences of the recent trend towards increased ideological homogeneity within the Democratic and Republican congressional delegations (McCarty, Poole, and Rosenthal 2006). While there are surely many costs associated with this trend, one potential benefit our model suggests is that as uncertainty about each party's underlying preferences declines, the public becomes more able to credibly commit to judging the parties on the basis of the success, or failure, of their policies.

On the Relationship Between Electoral Ambition and Congruence. Politicians are often criticized both for being too concerned about re-election (e.g., Fearon 1999) *and* for being insufficiently concerned about re-election (e.g., Dickerson 2005). Our model provides an explanation for this paradox, by predicting when ambition will be valued in a politician – namely, when it is clear that she is likely to have policy preferences that differ from her constituents. Otherwise, the perception that she is too ambitious can harm her electoral prospects.

The rationale for this prediction is straightforward. The only reason a non-congruent politician would use her information to promote her constituents' interests is if she wishes to be re-elected. In contrast, a congruent politician will naturally tend to promote her constituents' interests, unless electoral ambition, combined with electoral incentives to act as a delegate, induces her to posture and choose popular policies that actually harm her constituents' interests.

Appendix

This appendix uses notation for payoffs and beliefs introduced in footnote 6.

Proof of Lemma 1. Without loss of generality, let $\omega = A$. Since ϕ and θ are the same for all politicians, and ω is independent of the incumbent's type, $U_{CH} = U_I = \Pr(p = A|\omega = A)U_I(A, A) + \Pr(p = B|\omega = A)U_I(B, A) = \Pr(p = A|\omega = A)U_I(A, A) + (1 - \Pr(p = A|\omega = A))U_I(B, A)$. Since U_{CH} is a convex combination of $U_I(A, A)$ and $U_I(B, A)$ the lemma holds.∎

The next four lemmas prove Proposition 1.

Lemma 4 *For any p, ω, if $\pi(c, s|p, \omega) + \pi(c, u|p, \omega) \geq \phi$, $\pi(c, s|p, \omega) + \pi(n, s|p, \omega) \geq \theta$, and $\pi(c, s|p, \omega) \geq \phi\theta$, with at least one inequality strict, then the citizen's expected utility from the incumbent given p and ω is strictly greater than his expected utility from the challenger.*

Proof The voter's expected utility from re-electing the incumbent given p and ω is

$$
\begin{aligned}
U_I(p, \omega) &= \pi(c, s|p, \omega) y(c, s) + \pi(c, u|p, \omega) y(c, u) + \pi(n, s|p, \omega) y(n, s) + \pi(n, u|p, \omega) y(n, u) \\
&= \pi(c, s|p, \omega)[\beta + (1 - \beta)] + \pi(c, u|p, \omega)[\beta q + (1 - \beta) q] \\
&\quad + \pi(n, s|p, \omega)[\beta + (1 - \beta)(1 - \gamma)] + \pi(n, u|p, \omega)[\beta q + (1 - \beta)(1 - \gamma)].
\end{aligned}
$$

Substituting $\pi(n, u|p, \omega) = 1 - \pi(c, s|p, \omega) - \pi(c, u|p, \omega) - \pi(n, s|p, \omega)$ and simplifying yields $U_I(p, \omega) = \beta q + (1 - \beta)(1 - \gamma) + (1 - \beta)(q + \gamma - 1)[\pi(c, s|p, \omega) + \pi(c, u|p, \omega)] + \beta(1 - q)[\pi(c, s|p, \omega) + \pi(n, s|p, \omega)] + (1 - q)(1 - \beta)\pi(c, s|p, \omega)$. Since $(1 - \beta)(q + \gamma - 1) > 0$, $\beta(1 - q) > 0$, and $(1 - q)(1 - \beta) > 0$, the result holds. ∎

Lemma 5 *In either a delegate or trustee equilibrium the voter strictly prefers to re-elect the incumbent when $p = A$ and $\omega = A$.*

Proof. By Lemma 2 it suffices to show **(1)** $\phi < \pi(c, s|A, A) + \pi(c, u|A, A)$, **(2)** $\theta < \pi(c, s|A, A) + \pi(n, s|A, A)$, and **(3)** $\phi\theta < \pi(c, s|A, A)$. For a delegate equilibrium, **(1)** $\phi < \frac{\phi[\alpha + (1 - \alpha)(\theta + (1 - \theta)q)]}{\phi[\alpha + (1 - \alpha)(\theta + (1 - \theta)q)] + (1 - \phi)\alpha}$,

20

as $\alpha < \alpha + (1-\alpha)(\theta + (1-\theta)q)$, **(2)** $\theta < \frac{\theta[\alpha+(1-\alpha)\phi]}{\theta[\alpha+(1-\alpha)\phi]+(1-\theta)[\alpha+(1-\alpha)\phi q]}$, as $\alpha + (1-\alpha)\phi q <$

$\alpha + (1-\alpha)\phi$, and **(3)** $\phi\theta < \frac{\phi\theta}{\phi\theta+\phi(1-\theta)[\alpha+(1-\alpha)q]+(1-\phi)\alpha}$, as the denominator is less than one. For

a trustee equilibrium, **(1)** $\phi < \frac{\phi[\theta+(1-\theta)q]}{\phi[\theta+(1-\theta)q]+(1-\phi)\alpha[\theta+(1-\theta)q]}$, as $\alpha[\theta + (1-\theta)q] < \theta + (1-\theta)q$, **(2)**

$\theta < \frac{\theta[\phi+(1-\phi)\alpha]}{\theta[\phi+(1-\phi)\alpha]+(1-\theta)q[\phi+(1-\phi)\alpha]}$, as $q[\phi + (1-\phi)\alpha] < \phi + (1-\phi)\alpha$, and **(3)** $\phi\theta <$

$\frac{\phi\theta}{\phi\theta+\phi(1-\theta)q+(1-\phi)\theta\alpha+(1-\phi)(1-\theta)\alpha q}$, as the denominator is less than one.∎

Lemma 6 *There exists a delegate equilibrium if and only if $\theta \leq g(0,\beta,\phi)$.*

Proof. Lemma 2 (see the main text) establishes optimality of the incumbent's strategy and Lemmas 1 and 5 prove optimality of voter behavior when $\omega = A$. What remains is to confirm that it is optimal for the voter to re-elect the incumbent iff $p = A$ when $\omega = B$. By Lemma 1, it is sufficient to show that $U_I(B,B) \leq U_{CH}$. First, we rearrange terms:

$$
\begin{aligned}
U_I(B,B) &= \pi(c,s|B,B)y(c,s) + \pi(c,u|B,B)y(c,u) + \pi(n,s|B,B)y(n,s) + \pi(n,u|B,B)y(n,u) \\
&= \frac{1}{\phi\theta + q\phi(1-\theta) + (1-\phi)\theta + (1-\phi)(1-\theta)} \cdot \\
&\quad [\phi\theta y(c,s) + q\phi(1-\theta)y(c,u) + (1-\phi)\theta y(n,s) + (1-\phi)(1-\theta)y(n,u)] \\
&= \frac{1}{1 - \phi(1-q)(1-\theta)} \cdot [U_{CH} - (1-q)\phi(1-\theta)y(c,u)].
\end{aligned}
$$

Using this fact, algebra establishes that $U_I(B,B) \leq U_{CH}$ if and only if $\theta \leq$

$\frac{(1-\phi)[y(c,u)-y(n,u)]}{\phi(y(c,s)-y(c,u))+(1-\phi)(y(n,s)-y(n,u))}$. Accordingly, a delegate equilibrium exists if and only if $\theta \leq$ $g(0,\beta,\phi)$.∎

Lemma 7 *There exists a trustee equilibrium if and only if $\theta \geq g(\alpha,\beta,\phi)$.*

Proof. By reasoning as in the proof of Lemma 6, we only need $U(A,B) \leq U_{CH}$. When $\omega = B$, there are two ways the incumbent plays $p = A$ in a trustee equilibrium: she is congruent, unskilled,

and $s = A$ or she is non-congruent, unskilled, ambitious, and $s = A$. Thus

$$
\begin{aligned}
U_I\left(A, B\right) &= \pi\left(c, u|A, B\right) y\left(c, u\right) + \pi\left(n, u|A, B\right) y\left(n, u\right) \\
&= \frac{1}{\phi + \left(1 - \phi\right)\alpha}\left[\phi y\left(c, u\right) + \left(1 - \phi\right)\alpha y\left(n, u\right)\right].
\end{aligned}
$$

A trustee equilibrium requires $U_I\left(A, B\right) \leq U_{CH}$. Algebra establishes that this is so if and only if $\theta \geq \frac{\phi(1-\alpha)}{\alpha+\phi(1-\alpha)} \cdot \frac{(1-\phi)[y(c,u)-y(n,u)]}{\phi(y(c,s)-y(c,u))+(1-\phi)(y(n,s)-y(n,u))}$. Accordingly, a trustee equilibrium exists if and only if $\theta \geq g(\alpha, \beta, \phi)$.∎

Proof of Proposition 2. Substituting for $y(\cdot, \cdot)$ we have $g\left(\alpha, \beta, \phi\right) = \frac{(1-\phi)(1-\beta)(q+\gamma-1)}{\phi(1-q)+(1-\phi)\beta(1-q)} \cdot \frac{\phi(1-\alpha)}{\phi+\alpha(1-\phi)}$.
Part 1. The sign of $\frac{\partial g(\alpha,\beta,\phi)}{\partial \alpha}$ depends only on the sign of the derivative of $\frac{\phi(1-\alpha)}{\phi+\alpha(1-\phi)}$ with respect to α, which is negative. **Part 2.** The sign of $\frac{\partial g(\alpha,\beta,\phi)}{\partial \beta}$ depends only on the sign of the derivative of $\frac{(1-\phi)(1-\beta)(q+\gamma-1)}{\phi(1-q)+(1-\phi)\beta(1-q)}$ with respect to β, which is negative. **Part 3.** By inspection, it is obvious that $\lim_{\phi \to 0} g\left(\alpha, \beta, \phi\right) = \lim_{\phi \to 1} g\left(\alpha, \beta, \phi\right) = 0$. To show that g is single peaked we begin by differentiating: $\frac{\partial g}{\partial \phi} = \frac{-(1-\beta)(q+\gamma-1)(1-q)}{[\phi(1-q)+(1-\phi)\beta(1-q)]^2} \cdot \frac{\phi(1-\alpha)}{\phi+\alpha(1-\phi)} + \frac{(1-\phi)(1-\beta)(q+\gamma-1)}{\phi(1-q)+(1-\phi)\beta(1-q)} \cdot \frac{(1-\alpha)\alpha}{(\phi+\alpha(1-\phi))^2}$. Algebra establishes that $\frac{\partial g}{\partial \phi} \geq 0 \Leftrightarrow (1-\phi)^2 \alpha\beta - \phi^2 \geq 0$. The left hand side is positive when $\phi = 0$, negative when $\phi = 1$, and decreasing in ϕ, so g is single peaked.∎

References

Ansolabehere, Stephen, and James M. Snyder, Jr. 2000. "Valence Politics and Equilibrium in Spatial Election Models." *Public Choice* 103(June):327-336.

Aragones, Enriqueta, and Thomas R. Palfrey. 2004. "The Effect of Candidate Quality on Electoral Equilibrium: An Experimental Study." *American Political Science Review* 98(February):77-90.

Bianco, William T. 1994. *Trust: Representatives and Constituents.* Ann Arbor: University of Michigan Press.

Canes-Wrone, Brandice, Michael C. Herron, and Kenneth W. Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45(July):532-550.

Canes-Wrone, Brandice, and Kenneth W. Shotts. 2007. "When Do Elections Encourage Ideological Rigidity?" *American Political Science Review* 101(May):273-288.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(November):1431-1451.

Dickerson, John. 2005. "Cheney's Dreadful Lack of Ambition." Slate.com (Monday, November 21).

Fearon, James C. 1999. "Electoral Accountability and the Control of Politicians." In Adam Przeworski, Bernard Manin, and Susan C. Stokes, eds. *Democracy, Accountability, and Representation.* New York: Cambridge University Press.

Fiorina, Morris P. 1980. "The Decline of Collective Responsibility in American Politics." *Daedalus* 109(Summer):25-46.

Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections.* New Haven: Yale University Press.

Fox, Justin. 2007. "Government Transparency and Policymaking." *Public Choice* 131(April):23-44.

Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science* 45(October):862-886.

Mansbridge, Jane. 1999. "Should Blacks Represent Blacks and Women Represent Women? A Contingent 'Yes.'" *Journal of Politics* 61(August):628-657.

Maskin, Eric, and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94(September):1034-1054.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Political Ideology and Unequal Riches.* Cambridge: MIT Press.

Mill, John Stuart. 1962 [1861]. *Considerations on Representative Government.* South Bend: Gateway.

Morris, Stephen. 2001. "Political Correctness." *Journal of Political Economy* 109(April):231-265.

Prat, Andrea. 2005. "The Wrong Kind of Transparency." *American Economic Review* 95(June):862-877.

Stokes, Susan C. 2001. *Mandates and Democracy.* New York: Cambridge University Press.

Figure 1: Delegate and Trustee Equilibria
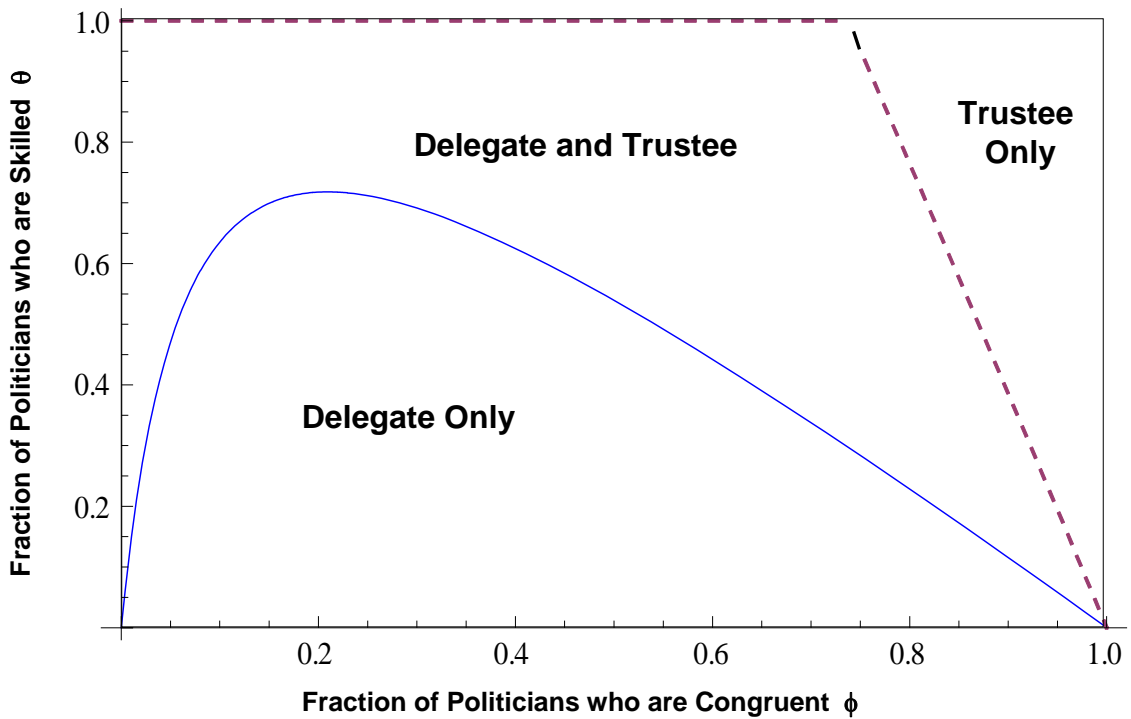
Key: solid line is $g(\phi, \alpha, \beta)$, dashed line is $g(\phi, 0, \beta)$
Parameter values for the figure: $\alpha = 0.7$, $\beta = 0.1$, $\gamma = 0.8$, q = 0.85