Multitask, Accountability, and Institutional Design

Scott Ashworth* Ethan Bueno de Mesquita†

June 19, 2012
Preliminary Draft. Comments Welcome.

Abstract

We consider a model of political accountability that allows us to examine the implications of unified vs. divided executive authority for the welfare of voters. The government is responsible for different tasks and the voter attempts to learn about the task-specific competences of the incumbent leader(s) in order to make electoral decisions. We identify a variety of trade-offs that shed light on the conditions under which it is optimal to bundle the tasks into a single elected office or unbundle the tasks into separate elected offices. Voter welfare is multi-faceted: voters care both about the strength of the incentives they create for politicians to take good actions and about identifying and retaining high quality politicians, creating the possibility for trade-offs in the institutional comparison. We show that as voter welfare puts greater weight on a particular task or as a politician's task-specific competences become more highly correlated, unbundling becomes more desirable relative to bundling with respect to creating incentives, but less desirable with respect to selecting high quality politicians. For some configurations of parameter values there is an unambiguously optimal institutional arrangement. For other configurations of parameter values, the optimal institutional arrangement depends on the relative weights placed on the two elements of voter welfare.

*Harris School of Public Policy Studies, University of Chicago, email: sashwort@uchicago.edu
†Harris School of Public Policy Studies, University of Chicago, email: bdm@uchicago.edu.

Government, at every level, is responsible for a broad array of tasks. The question of how tasks should be assigned across offices is a fundamental institutional design question for democracies. Important literatures in political science, economics, and law are concerned with the trade-offs associated with unified and divided authority in the executive.[1]

Interestingly, there is considerable empirical variation in this institutional design choice. For instance, in the United States federal government executive authority is essentially unified in a single office. The president has responsibility for all of the tasks assigned to the executive branch. The same, however, is not true of executive authority in the states or at the local level. In state government, executive authority is often divided among multiple elected offices including a governor, attorney general, treasurer, and so on. Similarly, local governments often have elected executives or mayors, sheriffs, assessors, and so on. Of course, there is also considerable variation in terms of the number of executive offices and the division or responsibilities, even at the state and local level (Berry and Gersen, N.d.).

We consider a model of political accountability that allows us to examine the implications of unified vs. divided executive authority for the welfare of voters. In our model, the government is responsible for different tasks and the voter is attempting to learn about the competence of the incumbent leader(s) in order to make electoral decisions. We explore two key features of the environment. First, we allow for the possibility that voters care differentially about the two tasks—i.e., one task may be more important for voter welfare than the other. Second, we allow for the possibility that a politician's competences on the various tasks are correlated. We focus on two institutional arrangements. Under the first, which, following Berry and Gersen (2008), we call "bundling", there is one politician responsible for all of the tasks. Under the second, which we call "unbundling", there is one politician responsible for each task.

We identify a variety of trade-offs that shed light on the conditions under which it is optimal to bundle or unbundle the tasks. An important subtlety is that voter welfare turns out to be multi-faceted in ways that complicate institutional comparisons. In particular, voters care both about the strength of the incentives they create for politicians to take good actions and about how well they (the voters) are doing and identifying and retaining high quality politicians. It can happen that the institution that is optimal for creating incentives is different from the institution that is optimal for retaining high quality incumbents. As such, we treat these two aspects of voter welfare separately.

The key results concern two types of comparative statics. First we ask how making the

[1] See, among many others, Calabresi and Rhodes (1992), Besley and Coate (2003), Marshall (2006), Berry and Gersen (2008), and Gersen (2010).

voter increasingly concerned with only one of the tasks affects which institution is optimal with respect to each aspect of voter welfare. Second we ask how increasing the correlation between the task-specific competences affects which institution is optimal with respect to each aspect of voter welfare. It turns out, there are trade-offs between the two aspects of voter welfare. As either the voter becomes more focused on a particular task or the task-specific competences become more highly correlated, unbundling becomes more desirable relative to bundling with respect to creating incentives, but less desirable with respect to selecting high quality politicians. For some configurations of parameter values there is an unambiguously optimal institutional arrangement. For other configurations of parameter values, the optimal institutional arrangement depends on the relative weights placed on the two elements of voter welfare.

Before turning to the analysis, it is worth commenting briefly on two related literatures. First, our model is clearly related to canonical models of multitask (Holmström and Milgrom, 1991). However, there is an important difference. Standard results about effort distortion in multi-task problems come from cost-complementarity. That is, the agent has a convex cost function and, so, the more she works on task 1, the higher the marginal cost of working on task 2. We explicitly rule out cost-complementarity as a source of distortions in our model by assuming task-specific costs that are additively separable. Thus, all of the multi-task incentives in our model come through the learning process as a result of the correlation between task-specific competences.

There is also an existing literature on multi-task problems in political agency settings. These models differ from ours in a variety of ways. Besley and Coate (2003) are also concerned with whether responsibility for multiple tasks should be located in one or multiple elected offices. They explore this question within a citizen-candidate model, so they are only able to address questions of selection, not of accountability and incentives. Bueno de Mesquita (2007), Hatfield and Padró i Miquel (2006), and Bueno de Mesquita and Landa (2012) focus on multi-task models with pure moral hazard, so there is no learning about politician type. Ashworth's (2005) model is in some ways similar to ours, but all of the multi-task incentives come through cost-complementarity—there is no correlation across task-specific competences. None of these models other than Besley and Coate (2003) consider the institutional design question with which we are concerned.

The paper proceeds as follows. Section 1 describes the formal model. Section 2 lays out our plan for the analysis and anticipates some key results. Sections 3 and 4 characterize equilibrium under bundling and unbundling, respectively. Section 5 characterizes the optimal institution with respect to the aspect of voter welfare that concerns creating incentives.

Section 6 characterizes the optimal institution with respect to the aspect of voter welfare that concerns retaining high quality politicians. In each of these two sections, we show how changing the two key parameters—how much the voter cares about one task versus another and the correlation in task-specific competences—affects the optimal institution. Section 7 offers some discussion and concludes.

1 The Setting

Consider an environment in which a government is undertaking two tasks, labeled 1 and 2, on behalf of a Voter. Each task will be carried out in each of two periods, $t = 1, 2$.

In each period, t, the politician responsible for task j will choose effort, a_j^t. This will result in an outcome for task j in period t given by:

$$s_j^t = a_j^t + \theta_j^t + \epsilon_j^t,$$

where θ_j^t is the task-j-specific competence of the politician responsible for task j in period t, and ϵ_j^t is a random shock that is normally distributed with mean 0 and variance 1. We assume these shocks are independent of each other and of the competence.

The task specific competences of a given politician are correlated. Specifically, any given politician has competences on the two tasks with prior distribution

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

We will consider two different institutions. In each, there is an election at the end of the first period. Prior to the election, the Voter observes the outputs, s_1^1 and s_2^1, but not the actions. No player observes the competences.

Following Berry and Gersen (2008), we refer to the first institution as *bundling*. Under bundling there is one politician in office in each period. She has responsibility for both tasks. At the election stage, she runs against a randomly selected challenger. The winner of the election gets a benefit of holding office $R > 0$.

We refer to the second institution as *unbundling*. Under unbundling, there are two politicians in office in each period. Each has responsibility for one task. At the election stage, each incumbent runs against a separate randomly selected challenger. Each election is for the relevant task-specific office. The winner of the task j election gets a benefit of holding office $R_j \geq 0$, with $R_1 + R_2 = R$.

1.1 Payoffs

Under either institution, in any period, t, the Voter's payoff from an outcome (s_1^t, s_2^t) is:

$$\gamma s_1^t + (1-\gamma)s_2^t,$$

with $\gamma \in [0,1]$. The Voter does not discount the future.

Under bundling the politician in office in period t who takes actions (a_1^t, a_2^t) has a payoff of

$$R - c(a_1^t) - c(a_2^t).$$

The cost function, c, satisfies:

$$c'(a) = a^k,$$

for some $k > 1$.

Under unbundling, the task-j politician in office in period t who takes action a_j^t has a payoff of

$$R_j - c(a_j^t).$$

2 Where We Are Going

The analysis of these models is quite involved. Thus, before turning to the analysis, it makes sense to highlight where we are going so that the reader does not get lost in the details.

We are interested in making an institutional comparison. In particular, we want to know when it is optimal for the Voter to have tasks bundled in a single office and when it is optimal for the Voter to have separate, specialized offices. Our approach to answering this question is as follows.

We begin by characterizing equilibrium under each institution. But notice, unbundling as we describe it above actually constitutes a continuum of institutions, depending on how the rewards of office are divided between the two offices. In order to make the institutional comparison, we find the optimal division for Voter welfare.

We then compare Voter welfare under each institution. In particular, we ask how changes in two parameter values affect which institution is optimal. The two parameters we study are the correlation across the two task-specific competences and the weights the Voter puts on the two tasks.

Importantly, the notion of Voter welfare in this model is multifaceted. In particular, changes in parameter values may have differential effects on Voter welfare in the first period (which depends on politician incentives) and Voter welfare in the second period (which turns out to depend only on how well the Voter does at selecting competent types). For some configurations of parameter values one institution is unambiguously optimal. However, for other configurations of parameter values one institution is optimal with respect to first period Voter welfare while the other is optimal with respect to second period Voter welfare.

In those latter circumstances, we do not offer any unequivocal normative conclusion. The reason is that we think of the second period of our model as, in a stylized way, representing the future. As such, it is unclear how much weight the Voter should put on the first period versus the second. Depending on the Voter's patience and the expected length of the future, any weight between zero and infinity might be appropriate. Thus, we leave seperate the analysis of first and second period welfare, without comparing between them.

The characterization of equilibrium and the following institutional comparisons become fairly technically complicated. So, to give a sense of where we are going, the following figure shows our ultimate result on optimal institutions as a function of the correlations in competences (ρ) and the weight the Voter attaches to task 1 relative to task 2 (γ). It does so for the special case where $c'(a) = a^2$.

3 Equilibrium under Bundling

We now turn to characterizing perfect Bayesian equilibrium in the game with bundling.

3.1 Second Period Effort

In the second period there are no electoral incentives, so the politician in office will choose $a_1 = a_2 = 0$.

3.2 Election

Given that second period effort will be zero, at the time of the election, the Voter's expected second period payoff from politician P is simply:

$$\gamma \mathbb{E}[\theta_1^P | (s_1, s_2)] + (1 - \gamma) \mathbb{E}[\theta_2^P | (s_1, s_2)].$$

For the Challenger, each of these expectations is zero. However, to determine the value of these expectations for the Incumbent, we need to calculate the Voter's posterior beliefs conditional on the first period outcomes.

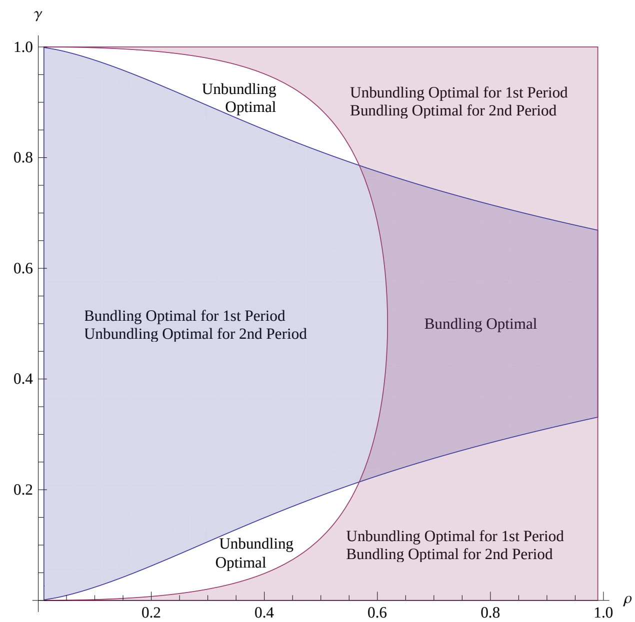

Figure 1: Institutional trade-offs as a function of Voter preference weights (γ) and correlation of competencies (ρ) for $k = 2$.

Because the prior and the signals are normal, the Voter's posterior beliefs will also be normal. Thus, the posterior means are sufficient statistics for optimal behavior by the Voter.

Suppose the Voter believes the Incumbent chose efforts (a_1^b, a_2^b) in the first period. Then $(s_1 - a_1^b, s_2 - a_2^b)$ is an unbiased signal of the true distribution of competence. The Voter combines this signal with his mean zero prior in the standard way to form his posterior (DeGroot, 1970, p. 175). Let (m_1, m_2) be the posterior means. They are given by:

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} + \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1} \right)^{-1} \cdot \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} s_1 - a_1^b \\ s_2 - a_2^b \end{pmatrix} + \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right].$$

Simplifying, this can be rewritten:

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \frac{1-\rho^2}{(2-\rho^2)^2 - \rho^2} \begin{pmatrix} 2-\rho^2 & \rho \\ \rho & 2-\rho^2 \end{pmatrix} \begin{pmatrix} s_1 - a_1^b \\ s_2 - a_2^b \end{pmatrix}. \tag{1}$$

Remark 3.1 It is worth pausing here to see an important intuition captured by this updating. Multiplying the first two factors on the right-hand side of Equation 1 we can see how performance on task i affects the Voter's beliefs about the Incumbent's task i and task j competences. In particular, as the outcome on task i improves (relative to the Voter's expectation) by one unit, the Voter's posterior beliefs about the Incumbent's competence on task i improves by

$$\left(\frac{1-\rho^2}{(2-\rho^2)^2-\rho^2}\right)\left(2-\rho^2\right).$$

As the outcome on task i improves (relative to the Voter's expectation) by one unit, the Voter's posterior beliefs about the Incumbent's competence on task j improves by

$$\left(\frac{1-\rho^2}{(2-\rho^2)^2-\rho^2}\right)\rho.$$

The first thing we learn from this is that increased performance on task i improves the Voter's beliefs about the Incumbent's ability on both tasks. This is intuitive because the two competences are correlated. The second thing we learn is a comparative static about the correlation. The marginal effect of task i outcomes on beliefs about task i competence is decreasing in the amount of correlation. That is

$$\frac{d}{d\rho}\left(\frac{1-\rho^2}{(2-\rho^2)^2-\rho^2}\right)\left(2-\rho^2\right)<0.$$

In contrast, the marginal effect of task i outcomes on beliefs about task j competence is increasing in the amount of correlation. That is

$$\frac{d}{d\rho}\left(\frac{1-\rho^2}{(2-\rho^2)^2-\rho^2}\right)\rho>0.$$

The second of these effects is straightforward. The more correlated are the two competences, the more informative is the task j outcome about the task i competence. The first effect is more subtle. The task j outcome serves essentially as a second signal about the task i competence. The more correlated are the two competences, the more informative is this second signal. Making this second signal more informative is akin to decreasing the Voter's prior uncertainty about task i competence, which leads the Voter to place less weight on the direct signal of task i competence.

Since the politician in office in the second period will choose minimal effort, the Voter makes his reelection decision simply by comparing these posterior beliefs to his prior on the Challenger's competence, weighting appropriately by how much he cares about each task. This implies that the Voter reelects the incumbent if and only if

$$\begin{pmatrix} \gamma & 1-\gamma \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \geq 0. \tag{2}$$

It will be useful to unpack the Voter's expected payoff from reelecting the Incumbent by substituting for (m_1, m_2) from Equation 1. In particular, define

$$\lambda_1 = \left(\frac{1-\rho^2}{(2-\rho^2)^2 - \rho^2} \right) \left(\gamma(2-\rho^2) + (1-\gamma)\rho \right)$$

and

$$\lambda_2 = \left(\frac{1-\rho^2}{(2-\rho^2)^2 - \rho^2} \right) \left(\gamma\rho + (1-\gamma)(2-\rho^2) \right).$$

Then Condition 2 above can be rewritten in terms of observed outcomes, rather than beliefs. In particular, Condition 2 is equivalent to the Voter reelecting the Incumbent if and only if

$$\lambda_1(s_1 - a_1^b) + \lambda_2(s_2 - a_2^b) \geq 0.$$

Remark 3.2 The difference between the λ's and γ provides another important intuition. The Voter's preferences put weight γ on task 1 and $1-\gamma$ on task 2. However, the Voter's equilibrium reelection rule puts weight λ_1 on task 1 and λ_2 on task 2. The fact that $\lambda_1 \neq \gamma$ and $\lambda_2 \neq 1-\gamma$ implies that, in equilibrium, the Voter is not giving the Incumbent precisely the incentives she would like to with respect to how to divide efforts across the two tasks.

The reason for this is that the Voter's reelection decision is forward looking—focused on selecting the politician who provides the highest expected payoff in the future. The λ's do take account of the preference weights, since these matter for the Voter's future payoffs. However, the λ's also incorporate information about the variances of signals and the correlations across the competences. It is this latter set of considerations that drives a wedge between the incentives the

Voter would like to give to maximize first period welfare and the incentives the Voter actually gives.

3.3 First Period Effort

The Incumbent will be reelected if $\lambda_1(s_1 - a_1^b) + \lambda_2(s_2 - a_2^b) \geq 0$. Notice, the left-hand side of this condition is distributed normally with mean

$$\lambda_1(a_1 - a_1^b) + \lambda_2(a_2 - a_2^b)$$

and variance

$$\sigma_b^2 = 2\lambda_1^2 + 2\lambda_2^2 + 2\lambda_1\lambda_2\rho. \tag{3}$$

Hence, the Incumbent believes that if she chooses efforts (a_1, a_2), she is reelected with probability:

$$1 - \Phi\left(\frac{0 - \lambda_1(a_1 - a_1^b) - \lambda_2(a_2 - a_2^b)}{\sigma_b}\right).$$

Thus the incumbent's expected payoff if she chooses a_1 and a_2 is

$$R\left[1 - \Phi\left(\frac{0 - \lambda_1(a_1 - a_1^b) - \lambda_2(a_2 - a_2^b)}{\sigma_b}\right)\right] - c(a_1) - c(a_2).$$

The first-order conditions for maximizing this are:[2]

$$\frac{\lambda_1 R}{\sigma_b}\phi\left(\frac{0 - \lambda_1(a_1 - a_1^b) - \lambda_2(a_2 - a_2^b)}{\sigma_b}\right) = c'(a_1)$$

and

$$\frac{\lambda_2 R}{\sigma_b}\phi\left(\frac{0 - \lambda_1(a_1 - a_1^b) - \lambda_2(a_2 - a_2^b)}{\sigma_b}\right) = c'(a_2).$$

In equilibrium the Voter's beliefs about effort must equal the true effort. Imposing these rational expectations gives the equilibrium conditions:

$$\frac{\lambda_1 R}{\sigma_b}\phi(0) = c'(a_1^*) \quad \text{and} \quad \frac{\lambda_2 R}{\sigma_b}\phi(0) = c'(a_2^*). \tag{4}$$

[2]Second order conditions will be satisfied if the cost function is sufficiently convex. See Ashworth (2005) for more details.

4 Equilibrium under Unbundling

In this section we solve for a perfect Bayesian equilibrium under unbundling. Let the reward for the winner of the task 1 election be $R_1 \equiv \eta R$ and the reward to the winner of the task 2 election be $R_2 \equiv (1-\eta)R$.

4.1 Second Period Effort

As in the case with bundling, in the second period there are no electoral incentives, so the politicians in office will engage in no effort.

4.2 Elections

Given that the second period efforts are zero, at the time of the election, the Voter's expected second-period payoff from having politician P on task 1 and politician P' on task 2 in the second period is:

$$\gamma \mathbb{E}[\theta_1^P | s_1] + (1-\gamma)\mathbb{E}[\theta_2^{P'} | s_2].$$

For each office, the expected competence of the the Challenger is zero. We now turn to calculating these expectations for the Incumbents.

Suppose the Voter believes the task j Incumbent will take action a_j^u in the first period. From the Voter's perspective, $s_j - a_j^u$, is normally distributed with mean θ_j and variance 2. Hence, standard updating with normal priors and normal signals implies that, given an outcome s_j and a belief a_j^u, the Voter's posterior beliefs about the task j incumbent's competence are normally distributed with mean $\lambda_u(s_j - a_j^u)$ and variance $2\lambda_u^2$, where $\lambda_u = \frac{1}{2}$. The Voter will reelect the task j incumbent if and only if the posterior mean is greater than 0. That is, if and only if

$$\lambda_u(s_j - a_j^u) \geq 0.$$

4.3 First Period Effort

The task j Incumbent will be reelected if $\lambda_u(s_j - a_j^u) \geq 0$. Notice, the left-hand side of this condition is distributed normally with mean

$$\lambda_u(a_j - a_j^u))$$

and variance

$$\sigma_u^2 = 2\lambda_u^2. \tag{5}$$

Hence, the task j Incumbent believes that if she chooses efforts a_j she is reelected with probability:

$$1 - \Phi\left(\frac{0 - \lambda_u(a_j - a_j^u)}{\sigma_u}\right).$$

Given this, the task 1 incumbent's expected payoff if she chooses effort a_1 is:

$$\eta R\left[1 - \Phi\left(\frac{0 - \lambda_u(a_1 - a_1^u)}{\sqrt{2\lambda_u^2}}\right)\right] - c(a_1).$$

The first-order condition is

$$\frac{\lambda_u \eta R}{\sqrt{2\lambda_u^2}}\phi\left(\frac{0 - \lambda_u(a_1^* - a_1^u)}{\sqrt{2\lambda_u^2}}\right) = c'(a_1^*).$$

In equilibrium the Voter's beliefs must be correct (i.e., $a_1^* = a_1^u$). Imposing this rational expectations condition and canceling the λ_u gives the equilibrium condition

$$\frac{\eta R}{\sqrt{2}}\phi(0) = c'(a_1^u). \tag{6}$$

An analogous analysis gives the equilibrium condition for task 2 as

$$\frac{(1-\eta)R}{\sqrt{2}}\phi(0) = c'(a_2^u). \tag{7}$$

4.4 Optimal Division of Rewards to Office

We have characterized equilibrium for an arbitrary division of electoral rewards between the two offices. We now ask for the division of rewards that maximizes Voter welfare. Clearly, the choice of η affects first period incentives for effort. The following result shows that η has no effect on Voter's ability to select competent types and, thus, no effect on the Voter's second period welfare.

Lemma 4.1 *The Voter's second period welfare is independent of η.*

(All omitted proofs are in the appendix.)

Lemma 4.1 implies that the optimal η is the one that maximizes the Voter's first period welfare. Write $a_1^u(\eta)$ and $a_2^u(\eta)$ for the equilibrium first period efforts as functions of η. To find the optimal allocation of electoral rewards, consider

$$\max_{\eta} \gamma a_1^u(\eta) + (1-\gamma) a_2^u(\eta).$$

The following result gives the optimum.

Lemma 4.2 *The Voter welfare maximizing η is given by:*

$$\eta^*(\gamma) = \frac{1}{1 + \left(\frac{1-\gamma}{\gamma}\right)^{\frac{k}{k-1}}}.$$

5 Optimal Institution: First Period Welfare

As we highlighted at the outset, it is possible for the welfare consequences of the two institutions to be different for first- and second-period Voter welfare. As such, we divide our analysis of optimal institutions into two parts. In this section we focus on the Voter's first period welfare. In the next section we consider the Voter's second period welfare. In both cases we consider how two parameters affect the optimal institutional choice: (i) the amount of correlation between the task-specific competences (ρ) and (ii) the Voter's weighting of the importance of the two tasks (γ).

Write $W_1^i(\gamma, \rho)$ for the date-1 Voter welfare under institution i when the correlation is ρ and the preference weight is γ. Then we can define

$$\mathcal{B}_1 = \{(\gamma, \rho) \mid W_1^b(\gamma, \rho) > W_1^u(\gamma, \rho)\}$$

and

$$\mathcal{U}_1 = \{(\gamma, \rho) \mid W_1^u(\gamma, \rho) > W_1^b(\gamma, \rho)\};$$

that is, $\mathcal{B}_1$ is the set of parameters for which bundling leads to the highest date-1 welfare and $\mathcal{U}_1$ is the set of parameters for which unbundling does.

We can define cross-sections of these sets in the obvious way. For example,

$$\mathcal{B}_1(\rho) = \{\gamma \mid W_1^b(\gamma, \rho) > W_1^u(\gamma, \rho)\}$$

is the set of preference weights γ for which bundling is optimal at a fixed correlation ρ.

First we show that either institution can be better for date-1 welfare—both $\mathcal{B}_1$ and $\mathcal{U}_1$ are nonempty. In fact, we show the stronger result that $\mathcal{B}_1(\rho)$ includes $\gamma = 1/2$ for all ρ and that $\mathcal{U}_1(\rho)$ includes $\gamma = 1$ for all $\rho \neq 0$.

Proposition 5.1 *There exist parameter values for which bundling is the optimal institution with respect to first-period Voter welfare and there exist parameter values for which unbundling is the optimal institution with respect to first-period Voter welfare. In particular:*

(i) If $\gamma \in \{0, 1\}$, then unbundling is preferred for all $\rho \neq 0$.

(ii) If $\gamma = \frac{1}{2}$, then bundling is preferred to unbundling for all ρ.

Proof.

Define $Q \equiv \frac{1-\rho^2}{(2-\rho^2)^2-\rho^2}$. Notice, with this definition we can write $\lambda_1 = Q\left[\gamma(2-\rho^2) + (1-\gamma)\rho\right]$ and $\lambda_2 = Q\left[\gamma\rho + (1-\gamma)(2-\rho^2)\right]$.

(i) Consider $\gamma = 1$. (An analogous argument holds for $\gamma = 0$.) In this case, Lemma 4.2 tells us that $\eta = 1$. As a result, the equilibrium condition for the unbundled case becomes

$$\frac{R}{\sqrt{2}}\phi(0) = c'(a_1^u).$$

In the bundled case, we have $\lambda_1 = Q(2-\rho^2)$ and $\lambda_2 = Q\rho$. So the equilibrium condition is

$$\frac{(2-\rho^2)R}{\sqrt{2((2-\rho^2)^2+\rho^2+(2-\rho^2)\rho^2)}}\phi(0) = c'(a_1^b).$$

Comparing the equilibrium conditions, we see that the comparison of a_1^u and a_1^b comes down to comparing

$$\frac{2-\rho^2}{\sqrt{2((2-\rho^2)^2+\rho^2+(2-\rho^2)\rho^2)}} \quad \text{to} \quad \frac{1}{\sqrt{2}}.$$

Square the LHS to get

$$\frac{1}{2}\frac{(2-\rho^2)^2}{((2-\rho^2)^2+\rho^2+(2-\rho^2)\rho^2)},$$

which is less than $1/2$ unless $\rho = 0$.

(ii) At $\gamma = 1/2$, Lemma 4.2 implies that $\eta = 1/2$. Then the equilibrium conditions (6) and (7) imply that $a_1^u = a_2^u = a^u$, where a^u satisfies

$$\frac{R}{2\sqrt{2}}\phi(0) = c'(a^u).$$

At $\gamma = 1/2$, we have $\lambda_1 = \lambda_2 \equiv \lambda$. This allows us to simplify the equilibrium conditions in (4): cancel the λ in

$$\frac{\lambda R}{\sqrt{2\lambda^2(2+\rho)}}\phi(0) = c'(a^b)$$

to get

$$\frac{R}{\sqrt{2(2+\rho)}}\phi(0) = c'(a^b)$$

Since c' is strictly increasing, $a^b > a^u$ if and only if

$$\frac{R}{\sqrt{2(2+\rho)}} > \frac{R}{2\sqrt{2}}.$$

This is always true, since $\rho \leq 1$ implies $\sqrt{2(2+\rho)} < 2\sqrt{2}$.

■

We can offer several intuitions about what is driving this result.

Take first the case of extreme preference weights. In this case, the Voter would like effort to be highly focused on the more important task. Under unbundling the Voter can insure this by giving almost all the rewards of office to the important task. However, as we saw in Remark 3.1, the Voter cannot fully achieve this goal under bundling. To see why, suppose the Voter cares only about task 1, but task 1 and task 2 abilities are correlated. To maximize task 1 incentives, the Voter would like to commit to voting only based on the outcome from task 1. However, because the two dimensions of competence are correlated, the Voter cannot make this commitment. The Voter learns about task 1 competence from Incumbent performance on task 2. As such, the Voter, who cares only about task 1, will nonetheless vote based on task 1 and task 2 performance, since both are informative about future task 1 performance. Knowing this, the Incumbent has incentives to spread her effort across the two tasks, which has the effect of decreasing effort on task 1. Thus, when the Voter has extreme preferences, unbundling provides better first period incentives by allowing the Voter to force himself to shut down incentives on the task he doesn't care about.

Now consider the case of a Voter who cares equally about the two tasks. Here, there are two effects of moving from an unbundled to a bundled institutional arrangement, and they cut in opposite directions. The first effect is that the rewards to office for each Incumbent are lower, since R is divided evenly between the two offices. This tends to lower effort under unbundling. The second effect is that the probability that an Incumbent's effort is pivotal in moving her from not being reelected to being reelected is higher under unbundling. This is because, as we've discussed, under bundling, the presence of a second task adds noise (from the Incumbent's perspective) to the Voter's reelection rule with respect to the first task (and vice versa). This tends to lower effort under bundling.

Since rewards from office fall by one half when moving from bundling to unbundling, the question for comparison, when $\gamma = 1/2$, is whether the pivot probability under bundling is more or less than one half of the pivot probability under unbundling. Since Proposition 5.1 says that at $\gamma = 1/2$ incentives are stronger under bundling, we want to see that the pivot probability falls by less than half when moving from unbundling to bundling. Notice that, formally, the pivot probabilities are represented by the densities at 0:

$$\frac{1}{\sigma_b}\phi(0) \quad \text{and} \quad \frac{1}{\sigma_u}\phi(0)$$

Thus the pivot probability will fall by less than half when moving from unbundling to bundling if $2\sigma_u > \sigma_b$. This is true if and only if:

$$4(\sigma_\theta^2 + \sigma_\epsilon^2) > 2(\sigma_\theta^2 + \sigma_\epsilon^2 + \sigma_\theta^2 \rho)$$

which is true iff

$$2(1-\rho)\sigma_\theta^2 + 2\sigma_\epsilon^2 > 0,$$

which is true for $\rho < 1$. (While we've developed these intuitions about the two competing effects and shown the dominance formally, we don't really have any idea why the dominance exists.)

Having seen that either institution can be optimal, we now consider comparative statics.

5.1 Competence Correlation

The first question we ask is how changes in the correlation between task 1 competence and task 2 competence affects the optimal institution. Obviously, correlation has no effect on first period welfare when the two tasks are unbundled, since different politicians have responsibility for the two tasks.

The same is not true when tasks are bundled. As the following result shows, the more correlated are the two dimensions of competence, the lower is the Voter's first period welfare under bundling. The intuition is that, as highlighted in Remark 3.2, correlation drives a wedge between the Voter's preferences over the two dimensions and the incentives across the two dimensions to which the Incumbent is responding. The more correlation there is, the more the Voter takes account of task 2 performance when forming task 1 beliefs and vice-versa. One can see formally by noting that the Incumbent's incentives come from (λ_1, λ_2), which combine γ and ρ. Hence, the Voter cannot fully align how much she cares about the two dimensions and how much she votes based on output on those dimensions. By contrast, under unbundling, the Voter chooses the rewards to office precisely to induce a division of effort across the two tasks that reflects the Voter's preference weights.

Proposition 5.2 *The more highly correlated are the competences, the less likely bundling is to be optimal with respect to first-period Voter welfare. That is, if $\rho > \rho'$, then $\mathcal{B}_1(\rho) \subsetneq \mathcal{B}_1(\rho')$.*

This result is perhaps surprising. One might have thought that, when two tasks are highly correlated, it makes sense to give both tasks to one agent. While we will see that this intuition has some merit with respect to the quality of selection, this result shows that there is a countervailing effect. Giving highly correlated tasks to one agent reduces incentives.

5.2 Preference Weights

Next we ask how the optimal institution, with respect to first-period Voter welfare, changes in the Voter's preference weights. The difference in first-period Voter welfare between the two institutions is given by:

$$W_1^u(\gamma, \rho) - W_1^b(\gamma, \rho) = \gamma(a_1^u(\eta^*(\gamma)) - a_1^b(\gamma)) + (1-\gamma)(a_2^u(\eta^*(\gamma)) - a_2^b(\gamma)),$$

where we have suppressed the dependence of the effort choices on ρ and will continue to do so to avoid clutter.

We want to ask whether bundling becomes more or less attractive relative to unbundling with respect to first-period Voter welfare as the Voter becomes more focused on a particular task. To address this question, we focus on $\gamma \in (\frac{1}{2}, 1)$, so that as γ gets larger the Voter is becoming more focused on task 1. (The case where $\gamma < 1/2$ is analogous). Differentiating with respect to γ and applying the Envelope Theorem, we find:

$$\frac{\partial}{\partial \gamma}\left[W_1^u(\gamma,,\rho) - W_1^b(\gamma,,\rho)\right] = \left(a_1^u\left(\eta^*\left(\gamma\right)\right) - a_1^b\left(\gamma\right)\right) - \left(a_2^u\left(\eta^*\left(\gamma\right)\right) - a_2^b\left(\gamma\right)\right) - \gamma\frac{\partial a_1^b\left(\gamma\right)}{\partial \gamma} - (1-\gamma)\frac{\partial a_2^b\left(\gamma\right)}{\partial \gamma}.$$

In order to get a handle on what is happening here, consider the limiting case of quadratic costs, i.e., $k = 1$. In this case, the optimal η is a corner solution at $\eta^* = 1$ for $\gamma > \frac{1}{2}$. Thus, in this case we have $a_1^u = \frac{R}{\sqrt{2}}\phi(0)$ and $a_2^u = 0$. Substituting these values into the derivative above yields:

$$\frac{\partial}{\partial \gamma}\left[W_1^u(\gamma,\rho) - W_1^b(\gamma,\rho)\right] = \frac{1}{\sqrt{2}} - a_1^b(\gamma) + a_2^b(\gamma) - \gamma\frac{\partial a_1^b(\gamma)}{\partial \gamma} - (1-\gamma)\frac{\partial a_2^b(\gamma)}{\partial \gamma}.$$

There are two competing effects here.

The first three terms of this derivative capture the first effect. Under unbundling, when $\gamma > \frac{1}{2}$, only the task 1 politician has any incentives to exert effort, since all of the electoral rents are allocated to her. The same is not true under bundling. As γ increases, the Voter becomes more focused on that task 1. As such, all else equal, he increasingly benefits from unbundling because his preferences are moving in the direction of alignment with the politician's behavior under unbundling relative to bundling. The following lemma establishes that this effect always goes in the direction of making unbundling more attractive to the Voter relative to bundling.

Lemma 5.1 *Let $k = 1$. Then $\frac{1}{\sqrt{2}} - a_1^b(\gamma) + a_2^b(\gamma) > 0$.*

The final two terms of the derivative capture the second effect. As the Voter becomes increasingly focused on task 1, the Politician's efforts under bundling change, but those under unbundling do not. In particular, as the Voter becomes increasingly focused on task 1, so too does the Politician under bundling. This effect, all else equal, tends to make bundling attractive relative to unbundling, as established in the following result.

Lemma 5.2 *Let $k = 1$. Then $-\gamma\frac{\partial a_1^b(\gamma)}{\partial \gamma} - (1-\gamma)\frac{\partial a_2^b(\gamma)}{\partial \gamma} < 0$.*

Taken together Lemma 5.1 and 5.2 demonstrate that there are competing effects of making the Voter more focused on the first task. The next result shows that, when $k = 1$, the first effect dominates so that as the Voter becomes more focused on one task, unbundling becomes more attractive with respect to first-period Voter welfare.

Proposition 5.3 *Let $k = 1$. For any ρ, $\mathcal{B}_1(\rho)$ is a centrally located interval. That is, for a fixed ρ, there exists a $\xi(\rho)$ such that bundling is optimal if and only if $\gamma \in (1/2 - \xi(\rho), 1/2 + \xi(\rho))$.*

When $k > 1$ we are unable to provide analytical results. The basic problem is that, in these cases, the effort levels and allocations under unbundling are also a function of γ, which adds additional effects. As shown in Figure 2, for a variety of choices of k, computation indicates that the result continues to hold—for every ρ, $B_1(\rho)$ is a centrally located interval. Notice, in this figure, you can also see the result from Proposition 5.2—the more correlated are the competences, the smaller is the interval of γ's on which bundling is optimal.

Remark 5.1 Proposition 5.3 and the computational results reported in Figure 2 have an important substantive implication for optimal institutional design. If one is concerned with first period Voter welfare, then, when one task is significantly more important than the other, create a hierarchy of offices: each is specialized, and the more important one gets significantly greater reward. When the jobs are of roughly equal importance, create a single office with multiple responsibilities. Importantly, it is never optimal to create two offices with roughly equal rewards.

6 Optimal Institution: Second Period Welfare

In this section, we consider second period welfare as a result of selection of good types. We start by characterizing ex ante, expected second-period welfare under each institution.

Lemma 6.1 *Let σ_b and σ_u be as defined in equations 3 and 5, respectively. Ex ante, expected second-period Voter welfare under institution i is:*

$$\sigma_i \phi(0).$$

Write $W_2^i(\gamma, \rho)$ for the date-1 voter welfare under institution i when the correlation is ρ and the preference weight is γ. Then we can define

$$\mathcal{B}_2 = \{(\gamma, \rho) \mid W_2^b(\gamma, \rho) > W_2^u(\gamma, \rho)\}$$

and

$$\mathcal{U}_2 = \{(\gamma, \rho) \mid W_2^u(\gamma, \rho) > W_2^b(\gamma, \rho)\};$$

that is, $\mathcal{B}_2$ is the set of parameters for which bundling leads to the highest ex ante, expected second-period Voter welfare and $\mathcal{U}_2$ is the set of parameters for which unbundling does.

We can define cross-sections of these sets in the obvious way. For example,

$$\mathcal{B}_2(\rho) = \{\gamma \mid W_2^b(\gamma, \rho) > W_2^u(\gamma, \rho)\}$$

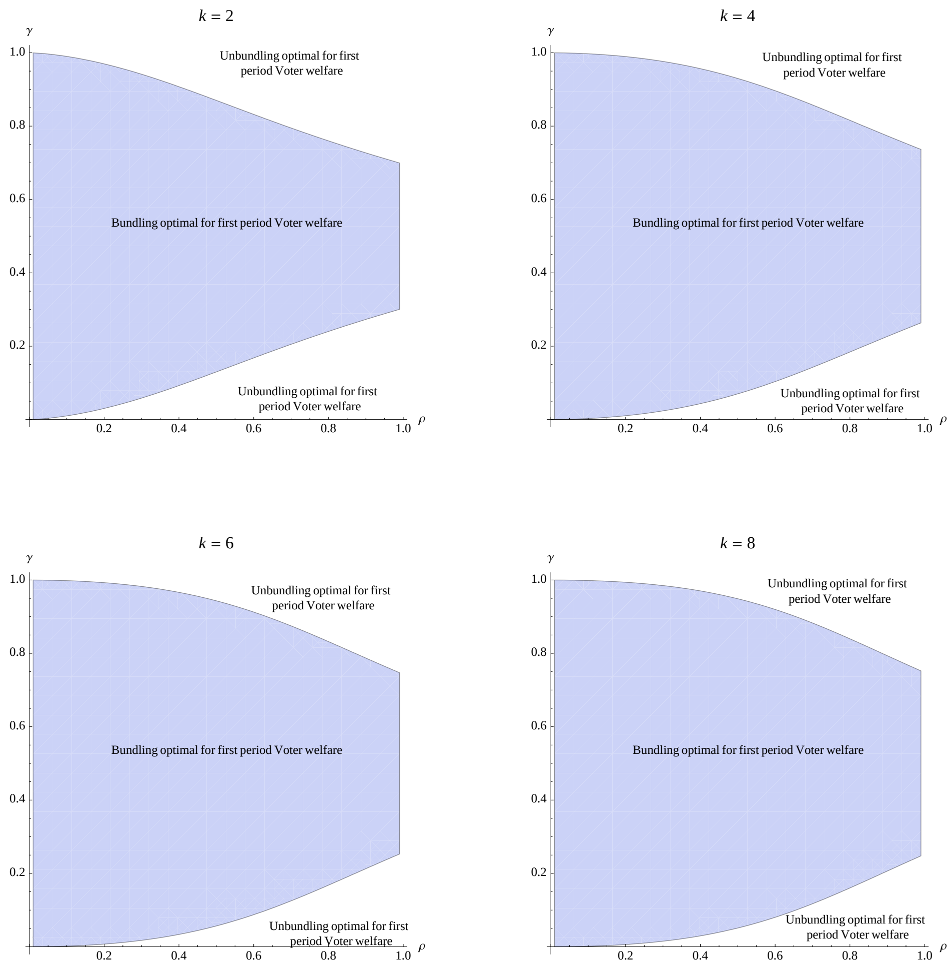

Figure 2: Computational results on optimal institution for first period Voter welfare for four values of k.

is the set of preference weights γ for which bundling is optimal at a fixed correlation ρ.

First we show that either institution can be better for date-2 welfare—both $\mathcal{B}_2$ and $\mathcal{U}_2$ are nonempty.

Proposition 6.1 *There exist parameter values for which bundling is the optimal institution with respect to ex ante, expected second-period welfare and there exist parameter values for which unbundling is the optimal institution with respect to ex ante, expected second-period welfare. In particular:*

(i) If ρ is sufficiently small, then unbundling is preferred for all γ and strictly preferred if $\gamma \notin \{0,1\}$.

(ii) If ρ is sufficiently large, then bundling is preferred for all γ.

Proof. By Lemma 6.1, to establish the results, we need only compare σ_u to σ_b. Note that, since $\lambda_u = 1/2$, $\sigma_u = \frac{1}{\sqrt{2}}$.

(i) Evaluating σ_b at $\rho = 0$ yields:

$$\sigma_b(\rho = 0) = \sqrt{\frac{\gamma^2 + (1-\gamma)^2}{2}}.$$

This is strictly less than $\frac{1}{\sqrt{2}}$ for any $\gamma \notin \{0,1\}$.

(ii) At $\rho = 1$, λ_1 and λ_2 are not defined. However, by l'Hopital's rule, the limit as ρ goes to 1 of each λ is $\frac{1}{3}$. Thus, we have:

$$\lim_{\rho \to 1} \sigma_b(\rho) = \sqrt{\frac{2}{3}},$$

which is greater than $\frac{1}{\sqrt{2}}$.

■

The intuition for this result is as follows. Moving from unbundling to bundling has two effects on the quality of selection. First, under bundling, the Voter has only one vote to select on two dimensions. This decreases the Voter's ability to flexibly select politicians who are high quality on one dimension but not the other. When there is no correlation, this is the only effect of the move and, thus, bundling is unambiguously bad for selection. However, there is a second effect of moving from unbundling to bundling when there is correlation across dimensions. In particular, the correlation means that the Voter gets more total information about the quality of the Incumbent under bundling. The more correlated are the dimensions, the larger this effect. The proposition shows that, if correlation is large enough, this effect dominates and bundling is preferred.

We now turn to comparative statics.

6.1 Competence Correlation

Next consider the effect of correlation on second period Voter welfare from selection. As the following result shows, increased correlation improves the quality of selection. There are two reasons for this. First, the more highly correlated are the two dimensions, the more informative is performance on each dimension about the other dimension. Thus, the Voter has more overall information. Second, the more highly correlated are the two dimensions, the more likely is it that an Incumbent who is strong (resp. weak) on one dimension is also strong (resp. weak) on the other dimension, so the Voter faces weaker cross-dimension trade-offs in selection.

Proposition 6.2 *The more highly correlated are the competences, the more likely bundling is to be optimal with respect to ex ante, expected second-period Voter welfare. That is, if* $\rho > \rho'$*, then* $\mathcal{B}_2(\rho) \subsetneq \mathcal{B}_2(\rho')$*.*

6.2 Preference Weights

Next consider the effect of preference weights on second period Voter welfare from selection. As the following result shows, the more the Voter cares about one task over the other, the better is selection. The reason is as follows. Changing Voter weights doesn't effect the amount of information. But the more the Voter cares about one issue, the more he can focus exclusively on finding a politician who is high competence on that dimension, as formalized in the next result.

Proposition 6.3 *The further from* $1/2$ *are the Voter's preference weights, the more likely bundling is to be optimal with respect to ex ante, expected second-period Voter welfare. That is, if* $|\gamma - \frac{1}{2}| > |\gamma' - \frac{1}{2}|$*, then* $\mathcal{B}_2(\gamma') \subsetneq \mathcal{B}_2(\gamma)$*.*

Figure 3 shows when bundling is optimal for ex ante, expected second-period Voter welfare. Consistent with Proposition 6.2, as ρ increases, bundling becomes more attractive relative to unbundling. Consistent with Proposition 6.3, as γ moves away from one-half, bundling becomes more attractive relative to unbundling.

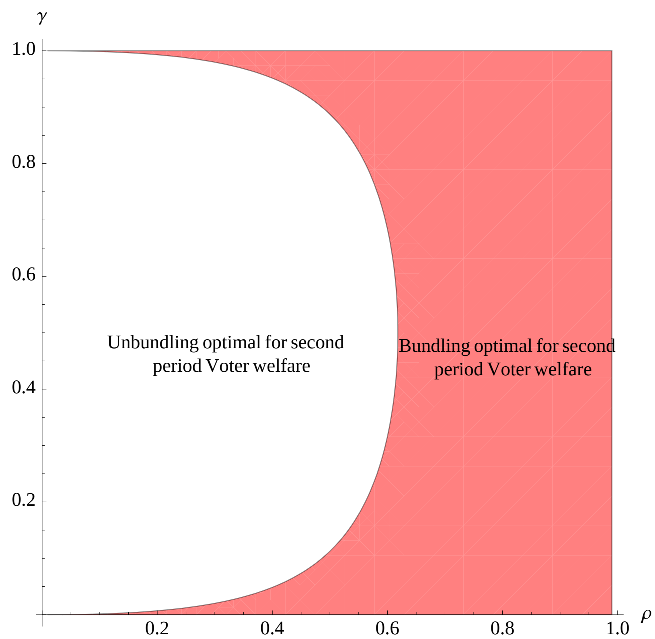

Figure 3: Optimal institution with respect to ex ante, expected second-period Voter welfare.

7 Conclusion

Figure 1 combines Figures 2 and 3. It shows that when the Voter cares about both issues (γ close to one-half) and the task-specific abilities are highly correlated, then bundling is the unambiguously optimal institution. When the Voter cares almost exclusively about one issue (γ close to one or zero) and the task-specific abilities are not too highly correlated, then unbundling is the unambiguously optimal institution. For other configurations of parameter values, an institutional choice poses a trade-off between first- and second-period welfare.

We are going to want some interpretation and implications.

Appendix

Proof of Lemma 4.1. By the martingale property of beliefs, the prior probability that the task j Incumbent is reelected is one-half. Recall that the Challengers' competences are both mean zero, so if an Incumbent is not reelected, the Voter's expected payoff on that task is zero. Thus, the Voter's expected second period welfare:

$$\frac{1}{2}\left(\mathbb{E}[\gamma\mathbb{E}[\theta_1|\lambda_u(s_1 - a_1^u) \geq 0] + (1-\gamma)\mathbb{E}[\theta_2|\lambda_u(s_2 - a_2^u) \geq 0]]\right)$$

which by the law of iterated expectations can be written:

$$\frac{1}{2}\left(\gamma\mathbb{E}[\lambda_u(s_1 - a_1^u)|\lambda_u(s_1 - a_1^u) \geq 0] + (1-\gamma)\mathbb{E}[\lambda_u(s_2 - a_2^u)|\lambda_u(s_2 - a_2^u) \geq 0]\right).$$

Recall that we can write:

$$s_j = a_j + \theta_j + \epsilon_j.$$

In equilibrium, $a_1 = a_1^u$ and $a_2 = a_2^u$. Hence, in equilibrium:

$$\lambda_u(s_j - a_j^u) = \lambda_u(\theta_j + \epsilon_j).$$

Call the right-hand side Z. The prior distribution of Z is normal with mean zero and variance:

$$\sigma^2 = \lambda_u^2 \cdot 2.$$

Now we can rewrite the Voter's expected second period welfare as:

$$\frac{1}{2}\left(\gamma\mathbb{E}[Z|Z \geq 0] + (1-\gamma)\mathbb{E}[Z|Z \geq 0]\right) = \frac{1}{2}\mathbb{E}[Z|Z \geq 0].$$

By a formula in Wooldridge (2002, p. 522), this is equal to:

$$\sigma_u\phi(0).$$

Since λ_u is independent of η, we have that σ_u is independent of η, establishing the result.
■

Proof of Lemma 4.2. The first-order condition is

$$\gamma\frac{da_1^u}{d\eta} + (1-\gamma)\frac{da_2^u}{d\eta} = 0.$$

By the implicit function theorem, the derivatives are

$$\frac{da_1^u}{d\eta} = \frac{\Delta}{c''(a_1^u(\eta))} \quad \text{and} \quad \frac{da_2^u}{d\eta} = \frac{-\Delta}{c''(a_2^u(\eta))},$$

where $\Delta = R\phi(0)/\sqrt{2}$. Rearranging the first-order condition gives:

$$\frac{\gamma}{c''(a_1^u(\eta))} = \frac{(1-\gamma)}{c''(a_2^u(\eta))}.$$

Using the fact that $c'(a) = a^k$, we can solve explicitly for the efforts. In particular, they satisfy:

$$a_2^u(\eta) = \left(\frac{1-\gamma}{\gamma}\right)^{\frac{1}{k-1}} a_1^u(\eta).$$

Substitute $a_1^u(\eta) = (\eta\Delta)^{1/k}$ and $a_2^u(\eta) = ((1-\eta)\Delta)^{1/k}$ to get

$$((1-\eta)\Delta)^{1/k} = \left(\frac{1-\gamma}{\gamma}\right)^{\frac{1}{k-1}} (\eta\Delta)^{1/k},$$

or

$$1-\eta = \left(\frac{1-\gamma}{\gamma}\right)^{\frac{k}{k-1}} \eta,$$

or

$$\eta(\gamma) = \frac{1}{1+\left(\frac{1-\gamma}{\gamma}\right)^{\frac{k}{k-1}}}.$$

■

Proof of Proposition 5.2. Recall $Q = \frac{1-\rho^2}{(2-\rho^2)^2-\rho^2}$.

Since $c'(a) = a^k$ for $k > 1$, date-1 welfare is

$$W(\gamma,\rho) = \gamma\left(\frac{\lambda_1 R\phi(0)}{\sigma}\right)^{1/k} + (1-\gamma)\left(\frac{\lambda_2 R\phi(0)}{\sigma}\right)^{1/k}.$$

Differentiate with respect to ρ to get

$$W_2(\gamma,\rho) = \frac{\gamma}{k}\left(\frac{\lambda_1 R\phi(0)}{\sigma}\right)^{\frac{1-k}{k}} \cdot \frac{d}{d\rho}\frac{\lambda_1}{\sigma} + \frac{1-\gamma}{k}\left(\frac{\lambda_2 R\phi(0)}{\sigma}\right)^{\frac{1-k}{k}} \cdot \frac{d}{d\rho}\frac{\lambda_2}{\sigma}.$$

Factor this to get

$$\frac{1}{k}\left(\frac{R\phi(0)}{\sigma}\right)^{\frac{1-k}{k}}\left[\gamma\lambda_1^{\frac{k-1}{k}}\frac{d}{d\rho}\frac{\lambda_1}{\sigma}+(1-\gamma)\lambda_2^{\frac{1-k}{k}}\frac{d}{d\rho}\frac{\lambda_2}{\sigma}\right].$$

To make progress, we will have to consider two cases: $\gamma \geq 1/2$ and $\gamma \leq 1/2$. But the symmetry of the problem means we only have to do one explicitly. We will see in the calculations below that $\gamma \geq 1/2$ implies

$$\frac{d}{d\rho}\frac{\lambda_1}{\sigma} < 0.$$

In addition, $\gamma \geq 1/2$ implies $\lambda_1 \geq \lambda_2$. Since $\frac{1-k}{k}$ is negative, this implies $\lambda_1^{\frac{1-k}{k}} \leq \lambda_2^{\frac{1-k}{k}}$.

Now there are two sub-cases: either

$$\frac{d}{d\rho}\frac{\lambda_2}{\sigma} \leq 0.$$

or not. In the first case, it is immediate that $W_2(\gamma,\rho) < 0$ for $\gamma \geq 1/2$. In the second case, we have the bound

$$\begin{aligned}&\frac{1}{k}\left(\frac{R\phi(0)}{\sigma}\right)^{\frac{1-k}{k}}\left[\gamma\lambda_1^{\frac{k-1}{k}}\frac{d}{d\rho}\frac{\lambda_1}{\sigma}+(1-\gamma)\lambda_2^{\frac{1-k}{k}}\frac{d}{d\rho}\frac{\lambda_2}{\sigma}\right]\\ \leq\ &\frac{1}{k}\left(\frac{R\phi(0)}{\sigma}\right)^{\frac{1-k}{k}}\lambda_1^{\frac{1-k}{k}}\left[\gamma\frac{d}{d\rho}\frac{\lambda_1}{\sigma}+(1-\gamma)\frac{d}{d\rho}\frac{\lambda_2}{\sigma}\right]\end{aligned}$$

Everything before the square brackets is positive, so showing that the term inside the square brackets is negative will establish that $W_2 < 0$ in the second sub-case. That will complete the argument for the case of $\gamma \geq 1/2$, and by symmetry will establish $W_2(\gamma,\rho) < 0$ for all γ, completing the proof of the entire proposition.

Let's calculate

$$\frac{d}{d\rho}\frac{\lambda_1}{\sqrt{\sigma^2}} = \frac{\frac{d\lambda_1}{d\rho}\sigma-\lambda_1\frac{d\sigma^2/d\rho}{2\sigma}}{\sigma^2} = \frac{\frac{d\lambda_1}{d\rho}\sigma^2-\frac{1}{2}\lambda_1\frac{d\sigma^2}{d\rho}}{\sigma^3}$$

To calculate the components, we take advantage of the simplification that Q factors out of the square root and cancels, so we can ignore the $\frac{dQ}{d\rho}$ components of the derivatives.

Given this, for $i = 1, 2$, define

$$\hat{\lambda}_i = \frac{\lambda_i}{Q}.$$

Similarly, define

$$\hat{\sigma} = \frac{\sigma}{Q}.$$

Notice that $\frac{\hat{\lambda}_i}{\hat{\sigma}} = \frac{\lambda_i}{\sigma}$.

Now, we have

$$\frac{d\hat{\lambda}_1}{d\rho} = -2\gamma\rho + 1 - \gamma$$

Similarly

$$\frac{d\hat{\lambda}_2}{d\rho} = \gamma - 2\rho(1 - \gamma)$$

Looking for symmetry,

$$\hat{\lambda}_2\frac{d\hat{\lambda}_1}{d\rho} = [\gamma\rho+(1-\gamma)(2-\rho^2)][-2\gamma\rho+(1-\gamma)] = -2\gamma^2\rho^2-2\gamma\rho(1-\gamma)(2-\rho^2)+\gamma(1-\gamma)\rho+(1-\gamma)^2(2-\rho^2)$$

$$\hat{\lambda}_1\frac{d\hat{\lambda}_2}{d\rho} = [\gamma(2-\rho^2)+(1-\gamma)\rho][\gamma-2\rho(1-\gamma)] = \gamma^2(2-\rho^2)+\gamma(1-\gamma)\rho-2\rho\gamma(1-\gamma)(2-\rho^2)-2\rho^2(1-\gamma)^2$$

These give us

$$\hat{\lambda}_2\frac{d\hat{\lambda}_1}{d\rho} - \hat{\lambda}_1\frac{d\hat{\lambda}_2}{d\rho} = (\rho^2 + 2)(1 - 2\gamma).$$

Next we have

$$\frac{d\hat{\sigma}^2}{d\rho} = 4\hat{\lambda}_1\frac{d\hat{\lambda}_1}{d\rho} + 4\hat{\lambda}_2\frac{d\hat{\lambda}_2}{d\rho} + 2\frac{d\hat{\lambda}_1}{d\rho}\hat{\lambda}_2\rho + 2\hat{\lambda}_1\frac{d\hat{\lambda}_2}{d\rho}\rho + 2\hat{\lambda}_1\hat{\lambda}_2.$$

Now we can start building up to the full derivative:

$$\frac{d\hat{\lambda}_1}{d\rho}\hat{\sigma}^2 = 2\hat{\lambda}_1^2\frac{d\hat{\lambda}_1}{d\rho} + 2\hat{\lambda}_2^2\frac{d\hat{\lambda}_1}{d\rho} + 2\rho\hat{\lambda}_1\hat{\lambda}_2\frac{d\hat{\lambda}_1}{d\rho}$$

$$\frac{1}{2}\hat{\lambda}_1\frac{d\hat{\sigma}^2}{d\rho} = 2\hat{\lambda}_1^2\frac{d\hat{\lambda}_1}{d\rho} + 2\hat{\lambda}_1\hat{\lambda}_2\frac{d\hat{\lambda}_2}{d\rho} + \frac{d\hat{\lambda}_1}{d\rho}\hat{\lambda}_1\hat{\lambda}_2\rho + \hat{\lambda}_1^2\frac{d\hat{\lambda}_2}{d\rho}\rho + \hat{\lambda}_1^2\hat{\lambda}_2$$

Thus the difference in the numerator of $\frac{d}{d\rho}\frac{\hat{\lambda}_1}{\sqrt{\hat{\sigma}^2}}$ is (after canceling the first terms and grouping)

$$\left(\hat{\lambda}_2\frac{d\hat{\lambda}_1}{d\rho} - \hat{\lambda}_1\frac{d\hat{\lambda}_2}{d\rho}\right)(2\hat{\lambda}_2 + \rho\hat{\lambda}_1) - \hat{\lambda}_1^2\hat{\lambda}_2 = (\rho^2 + 2)(1 - 2\gamma)(2\hat{\lambda}_2 + \rho\hat{\lambda}_1) - \hat{\lambda}_1^2\hat{\lambda}_2$$

Now let's calculate $\frac{d}{d\rho}\frac{\hat{\lambda}_2}{\sqrt{\hat{\sigma}^2}}$. We have

$$\frac{d}{d\rho}\frac{\hat{\lambda}_2}{\sqrt{\hat{\sigma}^2}} = \frac{\frac{d\hat{\lambda}_2}{d\rho}\hat{\sigma}^2 - \frac{1}{2}\hat{\lambda}_2\frac{d\hat{\sigma}^2}{d\rho}}{\hat{\sigma}^3}$$

$$\frac{d\hat{\lambda}_2}{d\rho}\hat{\sigma}^2 = 2\hat{\lambda}_1^2\frac{d\hat{\lambda}_2}{d\rho} + 2\hat{\lambda}_2^2\frac{d\hat{\lambda}_2}{d\rho} + 2\rho\hat{\lambda}_1\hat{\lambda}_2\frac{d\hat{\lambda}_2}{d\rho}$$

$$\frac{1}{2}\hat{\lambda}_2\frac{d\hat{\sigma}^2}{d\rho} = 2\hat{\lambda}_1\hat{\lambda}_2\frac{d\hat{\lambda}_1}{d\rho} + 2\hat{\lambda}_2^2\frac{d\hat{\lambda}_2}{d\rho} + \frac{d\hat{\lambda}_1}{d\rho}\hat{\lambda}_2^2\rho + \hat{\lambda}_1\hat{\lambda}_2\frac{d\hat{\lambda}_2}{d\rho}\rho + \hat{\lambda}_1\hat{\lambda}_2^2$$

The numerator is then

$$\left(\hat{\lambda}_1\frac{d\hat{\lambda}_2}{d\rho} - \hat{\lambda}_2\frac{d\hat{\lambda}_1}{d\rho}\right)(2\hat{\lambda}_1 + \rho\hat{\lambda}_2) - \hat{\lambda}_1\hat{\lambda}_2^2 = (\rho^2+2)(2\gamma-1)(2\hat{\lambda}_1 + \rho\hat{\lambda}_2) - \hat{\lambda}_1\hat{\lambda}_2^2.$$

Putting this mess of calculation all together, we have

$$\begin{aligned}
&\gamma\frac{d}{d\rho}\frac{\hat{\lambda}_1}{\hat{\sigma}} + (1-\gamma)\frac{d}{d\rho}\frac{\hat{\lambda}_2}{\hat{\sigma}} \\
=\ & \frac{\gamma[(\rho^2+2)(1-2\gamma)(2\hat{\lambda}_2+\rho\hat{\lambda}_1) - \hat{\lambda}_1^2\hat{\lambda}_2] + (1-\gamma)[(\rho^2+2)(2\gamma-1)(2\hat{\lambda}_1+\rho\hat{\lambda}_2) - \hat{\lambda}_1\hat{\lambda}_2^2]}{\hat{\sigma}^3} \\
=\ & \frac{(\rho^2+2)(1-2\gamma)[\gamma(2\hat{\lambda}_2+\rho\hat{\lambda}_1) - (1-\gamma)(2\hat{\lambda}_1+\rho\hat{\lambda}_2)] - \gamma\hat{\lambda}_1^2\hat{\lambda}_2 - (1-\gamma)\hat{\lambda}_1\hat{\lambda}_2^2}{\hat{\sigma}^3}.
\end{aligned}$$

We will now show that $\gamma(2\hat{\lambda}_2+\rho\hat{\lambda}_1) - (1-\gamma)(2\hat{\lambda}_1+\rho\hat{\lambda}_2)$ has the opposite sign of $1-2\gamma$, which will complete the proof. The expression $\gamma(2\hat{\lambda}_2+\rho\hat{\lambda}_1) - (1-\gamma)(2\hat{\lambda}_1+\rho\hat{\lambda}_2)$ is positive if and only if

$$2\gamma\hat{\lambda}_2 + \rho\gamma\hat{\lambda}_1 > 2(1-\gamma)\hat{\lambda}_1 + \rho(1-\gamma)\hat{\lambda}_2$$

iff

$$2[\hat{\lambda}_2\gamma - (1-\gamma)\hat{\lambda}_1] > \rho[(1-\gamma)\hat{\lambda}_2 - \gamma\hat{\lambda}_1]$$

iff

$$2[\gamma^2\rho+\gamma(1-\gamma)(2-\rho^2)-\gamma(1-\gamma)(2-\rho^2)-(1-\gamma)^2\rho] > \rho[(1-\gamma)\gamma\rho+(1-\gamma)^2(2-\rho^2)-\gamma^2(2-\rho^2)-\gamma(1-\gamma)\rho]$$

iff

$$2\rho(\gamma^2 - (1-\gamma)^2) > \rho(2-\rho^2)((1-\gamma)^2 - \gamma^2).$$

If $\gamma > 1/2$, then the LHS is positive and the RHS is negative, so the bracketed term is positive. If $\gamma < 1/2$, then the bracketed term is negative. Thus the bracketed term and $(1-2\gamma)$ always have the opposite sign, and the overall derivative is negative.

■

Proof of Lemma 5.1. In the case of quadratic costs, $a_j^b(\gamma) = \frac{\lambda_j}{\sigma_b}$. Hence, we want:

$$\frac{1}{\sqrt{2}} - \frac{\lambda_1}{\sigma_b} + \frac{\lambda_2}{\sigma_b} > 0.$$

Substituting for the definition of σ_b, this is equivalent to:

$$\frac{1}{\sqrt{2}} > \frac{\lambda_1 - \lambda_2}{\sqrt{2}\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_1\lambda_2\rho}},$$

which clearly holds. ■

Proof of Lemma 5.2. Differentiating and allowing Mathematica to simplify, we have:

$$-\gamma\frac{\partial a_1^b(\gamma)}{\partial\gamma} - (1-\gamma)\frac{\partial a_2^b(\gamma)}{\partial\gamma} = \frac{(2\gamma-1)\rho(\rho^2-1)}{2\sqrt{\frac{2\gamma^2(\rho^3-3\rho+2)-2\gamma(\rho^3-3\rho+2)+2}{4-\rho^2}}\,(\gamma^2(\rho^3-3\rho+2)-\gamma(\rho^3-3\rho+2)+1)}.$$

Since $\gamma > \frac{1}{2}$, the numerator is negative. Thus, the result is true if and only if:

$$\gamma^2(\rho^3 - 3\rho + 2) - \gamma(\rho^3 - 3\rho + 2) + 1 > 0.$$

This term is minimized at $\gamma = \frac{1}{2}$. Substituting it is straightforward to verify that this condition holds. ■

Proof of Proposition 5.3. Since the problem is symmetric we can restrict attention to $\gamma > \frac{1}{2}$. Now consider:

$$\Delta(\gamma,\rho) \equiv W_1^u(\gamma,\rho) - W_1^b(\gamma,\rho).$$

From Proposition 5.1, for any ρ, we have $\Delta(\frac{1}{2},\rho) < 0$ and $\Delta(1,\rho) \geq 0$. Thus, it suffices to show that for all ρ, $\Delta(\gamma,\rho)$ is single crossing in γ.

To do so, it suffices to show that Δ is concave in γ. To see this, note that the first crossing must be from below as established just above. Thus, if there was a second crossing it would be from above, implying that the function had become decreasing. But the fact that $\Delta(1, \rho) \geq 0$ would then imply that the function had to become increasing again, contradicting concavity. (Notice, because of quadratic costs, this is equivalent to showing that W_1^b is convex.)

Differentiating and allowing Mathematica to simplify we have:

$$\frac{\partial^2 \Delta}{\partial \gamma^2} = \frac{(1-\rho^2)\left[2\gamma^2(\rho-1)^2(3\rho^2+5\rho-2) - 2\gamma(\rho-1)^2(3\rho^2+5\rho-2) + 3\rho^4 - 9\rho^2 - 2\right]}{4\sqrt{\frac{2\gamma^2(\rho^3-3\rho+2)-2\gamma(\rho^3-3\rho+2)+2}{4-\rho^2}}\left[\gamma^2(\rho^3-3\rho+2) - \gamma(\rho^3-3\rho+2)+1)\right]^2}.$$

The denominator is positive, so the whole second derivative has the same sign as the numerator. Thus, it suffices to show:

$$2\gamma^2(\rho-1)^2(3\rho^2+5\rho-2) - 2\gamma(\rho-1)^2(3\rho^2+5\rho-2) + 3\rho^4 - 9\rho^2 - 2 < 0.$$

Now there are two cases to consider. If $\rho \geq \frac{1}{3}$, the left-hand side attains its maximum at $\gamma = 1$, where the value is $3\rho^4 - 9\rho^2 - 2$. That is decreasing in ρ, so it is maximized at $\rho = 1/3$, where it is clearly negative.

If $\rho \leq \frac{1}{3}$, the left-hand side attains its maximum at $\gamma = \frac{1}{2}$, where the value is $-\frac{(\rho-1)^2(3\rho^2+5\rho-2)}{2} + 3\rho^4 - 9\rho^2 - 2$. Rearranging, this is negative as long as:

$$3\rho^4 + \rho^3 - 9\rho^2 - 9\rho - 2 < 0.$$

We can find an upper bound on the left-hand side as follows. The term $3\rho^4 - 9\rho^2 - 2$ is convex, so it is maximized at either $\rho = 0$ or $\rho = 1/3$. It is straightforward to verify that it is maximized at $\rho = 0$. Next consider the term $\rho^3 - 9\rho$. It is decreasing in ρ, so is maximized at $\rho = 0$. Thus, the whole left hand side is maximized at $\rho = 0$ where it is equal to -2, establishing the claim. ■

Proof of Lemma 6.1.

First consider the case of bundling. By the martingale property of beliefs, the prior probability that the Incumbent is reelected is one-half. Recall that the Challenger's competences are both mean zero, so if the Incumbent is not reelected, the Voter's expected payoff is zero. Thus, the Voter's expected second period welfare:

$$\frac{1}{2}\mathbb{E}[\mathbb{E}[\gamma\theta_1 + (1-\gamma)\theta_2)|\gamma m_1 + (1-\gamma)m_2 \geq 0]]$$

which by the law of iterated expectations can be written:

$$\frac{1}{2}\mathbb{E}[\gamma m_1 + (1-\gamma)m_2)|\gamma m_1 + (1-\gamma)m_2 \geq 0].$$

Recall that we can write:

$$\gamma m_1 + (1-\gamma)m_2 = \lambda_1(a_1 + \theta_1 + \epsilon_1 - a_1^b) + \lambda_2(a_2 + \theta_2 + \epsilon_2 - a_2^b).$$

In equilibrium, $a_1 = a_1^b$ and $a_2 = a_2^b$. Hence, in equilibrium:

$$\gamma m_1 + (1-\gamma)m_2 = \lambda_1(\theta_1 + \epsilon_1) + \lambda_2(\theta_2 + \epsilon_2).$$

Call the right-hand side Z. The prior distribution of Z is normal with mean zero and variance:

$$\sigma^2 = \lambda_1^2 \cdot 2 + \lambda_2^2 \cdot 2 + 2\lambda_1\lambda_2\rho.$$

Now we can rewrite the Voter's expected second period welfare as:

$$\frac{1}{2}\mathbb{E}[Z|Z \geq 0].$$

By a formula in Wooldridge (2002, p. 522), this is equal to:

$$\sigma_b\phi(0).$$

Now consider the case of unbundling. By the martingale property of beliefs, the prior probability that the task j Incumbent is reelected is one-half. Recall that the Challengers' competences are both mean zero, so if an Incumbent is not reelected, the Voter's expected payoff on that task is zero. Thus, the Voter's expected second period welfare:

$$\frac{1}{2}\left(\mathbb{E}[\gamma\mathbb{E}[\theta_1|\lambda_u(s_1 - a_1^u) \geq 0] + (1-\gamma)\mathbb{E}[\theta_2|\lambda_u(s_2 - a_2^u) \geq 0]]\right)$$

which by the law of iterated expectations can be written:

$$\frac{1}{2}\left(\gamma\mathbb{E}[\lambda_u(s_1 - a_1^u)|\lambda_u(s_1 - a_1^u) \geq 0] + (1-\gamma)\mathbb{E}[\lambda_u(s_2 - a_2^u)|\lambda_u(s_2 - a_2^u) \geq 0]\right).$$

Recall that we can write:

$$s_j = a_j + \theta_j + \epsilon_j.$$

In equilibrium, $a_1 = a_1^u$ and $a_2 = a_2^u$. Hence, in equilibrium:

$$\lambda_u(s_j - a_j^u) = \lambda_u(\theta_j + \epsilon_j).$$

Call the right-hand side Y. The prior distribution of Y is normal with mean zero and variance:

$$\sigma^2 = \lambda_u^2 \cdot 2.$$

Now we can rewrite the Voter's expected second period welfare as:

$$\frac{1}{2}\left(\gamma \mathbb{E}[Y|Y \geq 0] + (1-\gamma)\mathbb{E}[Y|Y \geq 0]\right) = \frac{1}{2}\mathbb{E}[Y|Y \geq 0].$$

By a formula in Wooldridge (2002, p. 522), this is equal to:

$$\sigma_u \phi(0).$$

■

Proof of Proposition 6.2. From Lemma 6.1, ex ante, expected second-period voter welfare under bundling is:

$$\sigma_b \phi(0)$$

and under unbundling is

$$\sigma_u \phi(0).$$

Since σ_u is independent of ρ, it suffices to show that σ_b^2 is increasing in ρ. Differentiating σ_b^2 with respect to ρ and allowing Mathematica to simplify, we have:

$$\frac{\partial \sigma_b^2}{\partial \rho} = \frac{2(2\rho + (\rho^2 - 4\rho + 3)(2+\rho)^2 \gamma(1-\gamma))}{(\rho^2 - 4)^2},$$

which is clearly positive for $\rho \in [0, 1]$. ■

Proof of Proposition 6.3. From Lemma 6.1, ex ante, expected second-period voter welfare under budling is:

$$\sigma_b \phi(0)$$

and under unbundling is

$$\sigma_u \phi(0).$$

Since σ_u is independent of γ, it suffices to show that σ_b is decreasing in γ for $\gamma \in (0, \frac{1}{2})$ and increasing in γ for $\gamma \in (\frac{1}{2}, 1)$. It will first be useful to note:

$$\frac{d}{d\gamma}\lambda_1 = Q(2 - \rho^2 - \rho) > 0 \quad \text{and} \quad \frac{d}{d\gamma}\lambda_2 = Q(\rho - 2 + \rho^2) = -\frac{d}{d\gamma}\lambda_1$$

Now we have:

$$\begin{aligned}
\frac{d}{d\gamma}\sigma_b^2 &= 4\lambda_1\frac{d\lambda_1}{d\gamma} + 4\lambda_2\frac{d\lambda_2}{d\gamma} + 2\rho\lambda_1\frac{d\lambda_2}{d\gamma} + 2\rho\lambda_2\frac{d\lambda_1}{d\gamma} \\
&= 4(\lambda_1 - \lambda_2)\frac{d\lambda_1}{d\gamma} - 2\rho(\lambda_1 - \lambda_2)\frac{d\lambda_1}{d\gamma} \\
&= 2Q(\lambda_1 - \lambda_2)\frac{\partial\lambda_1}{\partial\gamma}(2 - \rho).
\end{aligned}$$

Since $\lambda_1 > \lambda_2$ if and only if $\gamma > \frac{1}{2}$, this establishes the result. ■

References

Ashworth, Scott. 2005. "Reputational Dynamics and Political Careers." *Journal of Law, Economics and Organization* 21:441–466.

Berry, Christoperh R. and Jacob E. Gersen. 2008. "The Unbundled Executive." *University of Chicago Law Review* 75:1385.

Berry, Christoperh R. and Jacob E. Gersen. N.d. "Fiscal Consequences of Electoral Institutions." *Journal of Law and Economics*. Forthcoming.

Besley, Timothy and Stephen Coate. 2003. "Elected versus appointed regulators: Theory and evidence." *Journal of the European Economic Association* 1(5):1176–1206.

Bueno de Mesquita, Ethan. 2007. "Politics and the Suboptimal Provision of Counterterror." *International Organization* 61(1):9–36.

Bueno de Mesquita, Ethan and Dimitri Landa. 2012. "Clarity of Responsibility with Sequential Policymaking." University of Chicago typescript.

Calabresi, Steven G. and Kevin H. Rhodes. 1992. "The Structural Constitution: Unitary Executive, Plural Judiciary." *Harvard Law Review* 1153:105.

DeGroot, Morris H. 1970. *Optimal Statistical Decisions*. New York: McGraw-Hill.

Gersen, Jacob E. 2010. "Unbundled Powers." *Virginia Law Review* 96:301–1965.

Hatfield, John William and Gerard Padró i Miquel. 2006. "Multitasking, Limited Liability, and Political Agency." Stanford Typescript.

Holmström, Bengt and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7:24–52.

Marshall, William P. 2006. "Break Up the Presidency? Governors, State Attorneys General, and Lessons from the Divided Executive." *Yale Law Journal* 246:115.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.