Unified vs. Divided Political Authority*

Scott Ashworth[†] Ethan Bueno de Mesquita[‡]

July 15, 2015

Abstract

Is unified or divided authority optimal for voter welfare? We study this question in a political agency model where multi-task distortions arise from correlation in task-specific competences. The model highlights trade-offs both within and across facets of voter welfare. Regarding incentives, unified authority yields higher total effort, but an allocation of that effort across tasks less fully aligned with voter preferences. Regarding the selection of good types, unified authority yields more voter information, but constrains voters to use that information less flexibly. Our comparative static analysis highlights a fundamental trade-off between determinants of the optimal institution—factors that make divided authority more attractive for incentives (e.g., voters focused on one task, highly correlated competences) make unified authority more attractive for selecting good types. For some parameter values there is nonetheless an unambiguously optimal institution. For other parameter values, the overall optimal institution depends on the heterogeneity of politician competences.

Keywords: Political Agency, Elections, Career Concerns, Multitask, Unified vs. Divided Authority

*We have received valuable feedback from Chris Berry, Steve Coate, Alex Frankel, Matt Gentzkow, Jacob Gersen, Gilat Levy, Pablo Montagnes, John Patty, Carlo Prato, Jesse Shapiro, Ken Shotts, Lars Stole, Stephane Wolton, and seminar audiences at the Priorat Workshop on Theoretical Political Science, the University of Chicago, and Washington University.

[†]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.

The institutions that govern the accountability relationship between policy-makers and voters have far reaching effects on the quality of governance. For instance, term limited governors in the U.S. spend and tax more and borrow at higher interest rates than their reelection-eligible peers (Besley and Case, 1995, 2003; Alt, Bueno de Mesquita and Rose, 2011). Similarly, term-limited mayors in Brazil implement education reforms less effectively and engage in more corruption than their reelection-eligible peers (Finan and Ferraz, 2011; de Janvry, Finan and Sadoulet, 2011). Beyond term limits, regulators who are directly elected are more pro-consumer than are regulators who are appointed by an elected official (Besley and Coate, 2003). U.S. states with more centralized pension administration have less generous public sector pensions than do states with more decentralized systems (Glaeser and Ponzetto, 2014). And recent evidence suggests that centralization of regulatory authority under the Clean Air Act affected the distribution of benefits from pollution reduction (Boffa, Piolatto and Ponzetto, 2013).

At least since the time of the American founding, the division of executive authority across policy domains has been a key issue in debates about the design of institutions to promote electoral accountability. Reflecting their view that unified authority facilitated accountability and good governance, the Founders chose unified executive authority for the federal government. As Alexander Hamilton wrote in Federalist 70:

> [T]he executive power is more easily confined when it is one: that it is far more safe there should be a single object for the jealousy and watchfulness of the people; and, in a word, that all multiplication of the executive is rather dangerous than friendly to liberty.

To emphasize that the Founders were making an explicit institutional design choice, Calabresi and Terrell (2009) note that neither the majority of state governments, nor the federal government under the Articles of Confederation, had a unitary executive.

This heterogeneity in the institutional design of executive accountability persists to the present day. While the executive branch of any government has multiple responsibilities, the number of elected executive-branch officials amongst whom these responsibilities are divided varies tremendously across polities. To get a sense of the magnitudes, consider the following facts about U.S. state and local governments, reported by Berry and Gersen (2009). The county at the seventy-fifth percentile for number of elected officials per capita has over five times as many elected officials (1.6 per 1,000 people) as the county at the twenty-fifth percentile (0.3 per 1,000). In U.S. state governments, there were more than 10,000 elected executive officials (not including state legislatures or elected boards) as of 1992, with enormous variation in executive unbundling from state to state. For instance, New Jersey elected only its governor, Delaware elected 80 officials, and Pennsylvanian elected 1,200. Elected executive offices range from governors (in every state) to attorneys general, treasurers, and secretaries of state (in most states) to utility regulators, education commissioners, and comptrollers (in a handful of states).

In light of both the long-standing debate and the considerable empirical variation in institutional structure, the positive and normative implications of the division of executive authority are important topics for political economy analysis.

Two aspects of the desirability of unified vs. divided executive authority have been understood at least since Hamilton. These are that dividing authority leads to problems of free riding and coordination. And, indeed Berry (2009) and Berry and Gersen (2009) show that divided authority creates a fiscal common pool that results in larger government. More recently, Besley and Coate (2003) have identified an additional welfare consideration. They show that, relative to appointment by an elected official, separate elections for regulators tend to select more ideologically pro-consumer regulators.

Because Besley and Coate (2003) work within a citizen-candidate model, they are only able to address questions of selection, not how concerns over reelection affect politicians' actions. And because traditional arguments about the effect of unified vs. divided authority on free riding and coordination are concerned only with politicians' actions, they do not address issues of selection.

Of course, the literature on political agency suggests that these two sets of concerns are deeply intertwined in equilibrium (Fearon, 1999; Persson and Tabellini, 2000; Maskin and Tirole, 2004; Besley, 2005). To further the debate, we provide an analysis of unified vs. divided executive authority in the context of a canonical model of electoral accountability, one that incorporates both incentives and selection (although abstracting away from issues of coordination that are already well understood).[1]

In our model, the executive is responsible for two different tasks, while the voters are attempting to learn about the competence of the incumbent leader(s) in order to make electoral decisions. We explore three 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. Third, we allow for variation

[1]We build on the literature applying Holmström's (1999) career concerns model to electoral incentives. For other models applying career concerns to elections see, for instance, Lohmann (1998), Persson and Tabellini (2000), Ashworth (2005), Ashworth and Bueno de Mesquita (2006), Gehlbach (2007), and Alesina and Tabellini (2007, 2008).

in how heterogenous the population of politicians is in terms of competence.

We focus on two institutional arrangements. Under the first, which, following Berry and Gersen (2008), we call *bundling*, there is one politician responsible for both tasks. Under the second (*unbundling*), there are separate politicians responsible for each task.

We identify a variety of trade-offs that help illuminate when bundling or unbundling is optimal for voter welfare. An important subtlety is that voter welfare is multi-faceted in ways that complicate institutional comparisons. Voters care both about the incentives they create for politicians and about identifying and retaining high quality politicians. It is useful to think about each of these facets of voter welfare separately.

For selection, the key trade-off is information versus flexibility. The more information the voters have, the better they can identify high quality politicians. But, all else equal, voters are better off if they can flexibly keep a politician who is good at one task while at the same time replacing a politician who is bad at the other task. Under bundling, the voter decides whether to reelect a single incumbent about whom he observes two informative signals—performance on each task. Under unbundling, the voter makes two decoupled decisions, each on an incumbent about whom he has one informative signal. Thus, bundling creates more information for the voter, while unbundling allows greater flexibility.

The way the interplay of information and decoupling affects first-period incentives is more subtle, but the net effects are these. First, the voters would like the politician to have strong incentives to exert a high level of total effort. Second, the voters would like the politician's allocation of effort across the two tasks to be aligned with the voters' own weighting of the relative importance of the tasks. Bundling creates more total effort, but reduces alignment, so again there is a trade-off.

In addition to these trade-offs within each facet of voter welfare, it can happen that there are trade-offs across the facets. That is, sometimes the institution that is optimal for incentives is different from the institution that is optimal for identifying and retaining high quality incumbents.

Our key welfare results concern how two comparative statics affect the trade-offs both within and across facets of voter welfare. First, we ask how making the voter increasingly concerned with only one of the tasks affects which institution is optimal with respect to each facet of voter welfare. Second we ask how increasing the correlation between the task-specific competences affects which institution is optimal with respect to each facet of voter welfare. These comparative statics reveal a fundamental trade-off: changes that make bundling more attractive with respect to one facet of voter welfare make unbundling more attractive with respect to the other. In particular, 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 incentives, but less desirable with respect to selecting high quality politicians.

In light of these results, we turn to an analysis of overall voter welfare. For some configurations of parameter values there is an unambiguously optimal institutional arrangement. For other configurations of parameter values, one institution is optimal for one facet of voter welfare and the other institution is optimal for the other facet of voter welfare. We show that, as the pool of potential politicians becomes either very heterogeneous or very homogenous, overall voter welfare is entirely determined by one of the two facets of voter welfare. When the pool is very homogenous, the optimal institution for overall welfare is entirely determined by which institution is better for incentives. When the pool is very heterogeneous, the optimal institution for overall welfare is entirely determined by which institution yields better selection. As we show, in the case of extreme heterogeneity, flexibility dominates information, and so unbundling is unambiguously preferred.

The paper proceeds as follows. Section 1 briefly reviews related theoretical literatures. Section 2 describes the formal model. Sections 3 and 4 characterize equilibrium under unbundling and bundling, respectively. Section 5 characterizes the optimal institution with respect to the facet of voter welfare that concerns incentives. Section 6 characterizes the optimal institution with respect to the facet 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 explores an analysis of overall voter welfare. Section 8 offers some discussion and concludes.

1 Related Literature

Existing models of multitask problems in political agency settings differ from our model in a variety of ways. Besley and Coate (2003) are also concerned with whether responsibility for different 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. Swank and Visser (2006), Hatfield and Padró i Miquel (2006), Bueno de Mesquita (2007), and Bueno de Mesquita and Landa (2012) focus on multitask models with pure moral hazard, so there is no learning about politician type. Ashworth (2005) studies a model in which multitask distortions 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.[2]

Our finding that there are trade-offs between institutions that are best for first-period welfare and those that are best for second-period welfare is reminiscent of results in Besley (2005, ch. 3) and Maskin and Tirole (2004).[3] In these models, politicians are either congruent types, who share the voters' policy preferences, or non-congruent types, who have preferences opposed to the voters'. Institutions that create strong incentives lead to pooling, which is good for first-period voter welfare, but makes it impossible for voters to select congruent types for the future. Institutions that create weak incentives allow voters to select congruent types for the future, but at the cost of lower first-period welfare. In our model, the first-period actions have no direct impact on how much the voter learns. Instead, the trade-off across periods results from the different ways that effort and selection are affected by information and flexibility.

Our model is also, of course, related to discussions of multitask principal-agent problems outside of elections. Unlike Holmström and Milgrom's (1991) and Dewatripont, Jewitt and Tirole's (1999) canonical models, multitask incentives in our model do not arise because of cost complementarity. Indeed, we assume additively separable costs. Our incentive alignment problem is similar to that studied in Feltham and Xie (1994), Baker (2002), and Gibbons (2010). In these papers there are multiple tasks, but explicit contracts are restricted to condition only on lower-dimensional signals. Misalignment arises when the low-dimensional signals weight the tasks differently than does the principal's welfare. In our model the voter has access to a signal for each action, but the career concerns nature of the model endogenously creates incentives based on a lower-dimensional weighting of these signals. Hence, a similar misalignment occurs endogenously in our model. Our comparative statics then show (among other things) how this endogenous misalignment is affected by changes in the correlation across task-specific competences and the voters' weighting of the two tasks. In addition, the electoral setting of our model introduces a second dimension of welfare (namely, selection) that isn't present in those models.

---

[2] Ahn and Oliveros (2012, 2013) study voting over multiple issues in a pure information aggregation environment with no agency relationship. They show that bundling and unbundling are equivalent in terms of asymptotic efficiency in that environment.

[3] See Prat (2005) for a related contribution outside the electoral context.

2 The Setting

A government undertakes 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 takes effort, a_j^t. The outcome for task j in period t is

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

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 competences.

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

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma_\theta^2 \cdot \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).$$

Here, σ_θ^2 measures heterogeneity of the pool of politicians is with respect to competence, while ρ measures the individual-level correlation between the task-specific competences.

We consider two different institutions. Under 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 or shocks. No player observes the competences.

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 $R > 0$.

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 $R_1 + R_2 = R$.

Under either institution, in any 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) gets payoff

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

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

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

Costs are quadratic:

$$c(a) = \frac{1}{2}a^2.$$

Throughout, we focus on pure-strategy perfect Bayesian equilibria. Such an equilibrium exists so long as the first-order conditions we derive below characterize optimal effort choices by the incumbent(s); Lemma A.2 in the appendix shows that a sufficient condition for this is that R not be too large.

3 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 $R_1 = \eta R$ and the reward to the winner of the task 2 election be $R_2 = (1-\eta)R$, where $0 \leq \eta \leq 1$.

3.1 Second Period Effort

In the second period there are no electoral incentives, so the politicians in office will engage in no effort.

3.2 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 $\sigma_\theta^2 + 1$. Hence, standard results on updating normal priors with normal signals imply 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

$(\sigma_\theta^2+1)\lambda_u^2$, where $\lambda_u = \frac{\sigma_\theta^2}{\sigma_\theta^2+1}$. 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

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

3.3 First Period Effort

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

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

and variance

$$\sigma_u^2 = (\sigma_\theta^2 + 1)\left(\frac{\sigma_\theta^2}{\sigma_\theta^2+1}\right)^2 = \frac{\sigma_\theta^4}{\sigma_\theta^2+1}. \tag{1}$$

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:

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

The first-order condition for maximizing this payoff is:

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

In equilibrium the Voter's conjecture (a_1^u) must be correct. Imposing this rational expectations condition, substituting for c', and noting that

$$\frac{\lambda_u}{\sigma_u} = \frac{1}{\sqrt{\sigma_\theta^2+1}}$$

gives the equilibrium condition:

$$\frac{\eta R}{\sqrt{\sigma_\theta^2+1}}\phi(0) = a_1^u. \tag{2}$$

An analogous analysis gives the equilibrium condition for task 2 as

$$\frac{(1-\eta)R}{\sqrt{\sigma_\theta^2+1}}\phi(0) = a_2^u. \tag{3}$$

3.4 Equilibrium Voter Welfare

Unbundling creates a whole family of institutions, one for each value of η. For the purpose of comparing unbundling to bundling, we focus on the unbundled institution whose division of rewards, η, maximizes Voter welfare. This allows us to compare bundling to the best possible version of unbundling.

To this end, we next calculate the Voter's ex-ante expected first-period and second-period welfare as a function of the division of rewards.

Lemma 3.1

(i) The Voter's ex-ante expected first-period welfare under unbundling when the division of rewards is η is

$$V_1^u(\eta,\rho,\gamma,\sigma_\theta^2) = \gamma\frac{\eta R}{\sqrt{\sigma_\theta^2+1}}\phi(0) + (1-\gamma)\frac{(1-\eta)R}{\sqrt{\sigma_\theta^2+1}}\phi(0)$$

(ii) The Voter's ex-ante expected second-period welfare under unbundling when the division of rewards is η is

$$V_2^u(\eta,\rho,\gamma,\sigma_\theta^2) = \frac{\sigma_\theta^2}{\sqrt{\sigma_\theta^2+1}}\phi(0).$$

Proof.

(i) Using Equations 2 and 3, the Voter's expected first-period welfare is

$$\gamma a_1^u + (1-\gamma)a_2^u = \gamma\frac{\eta R}{\sqrt{\sigma_\theta^2+1}}\phi(0) + (1-\gamma)\frac{(1-\eta)R}{\sqrt{\sigma_\theta^2+1}}\phi(0).$$

(ii) The Voter's expected second-period welfare conditional on posterior means (m_1, m_2) is

$$\gamma\max\{m_1,0\} + (1-\gamma)\max\{m_2,0\}.$$

Along the equilibrium path,

$$m_1 = \lambda_u(\theta_1 + \epsilon_1) \quad \text{and} \quad m_2 = \lambda_u(\theta_2 + \epsilon_2).$$

Call these random variables Z_1 and Z_2, respectively. Each Z has a prior distribution that is normal with mean 0 and variance

$$\sigma_u^2 = \frac{\sigma_\theta^4}{\sigma_\theta^2 + 1}.$$

By a formula in Wooldridge (2002, p. 522),

$$\mathbb{E}[\max\{Z_i, 0\}] = \sigma_u \phi(0)$$

for each i. Thus

$$\gamma \max\{m_1, 0\} + (1-\gamma)\max\{m_2, 0\} = \sigma_u \phi(0).$$

■

Voter welfare, $V_1^u(\eta, \rho, \gamma, \sigma_\theta^2) + V_2^u(\eta, \rho, \gamma, \sigma_\theta^2)$, is affine in $\eta \in [0,1]$, so the optimal division of rewards is bang-bang, as characterized in the next result.

Lemma 3.2 *Under unbundling, the Voter welfare maximizing η is given by:*

$$\eta^* = \begin{cases} 1 & \text{if } \gamma > \frac{1}{2} \\ 0 & \text{if } \gamma < \frac{1}{2}. \end{cases}$$

If $\gamma = 1/2$, then any η is optimal.

Proof. Follows from the argument in the text. ■

4 Equilibrium under Bundling

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

4.1 Second Period Effort

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

4.2 Election

Given that second period effort will be zero, the Voter's (election-date) expected second period payoff from politician P is

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

For the Challenger, each of these expectations is zero. 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.

The prior and the signals are normal, so the Voter's posterior beliefs are also normal. 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 to form his posterior (DeGroot, 1970, p. 175). Let (m_1, m_2) be the posterior means. They are given by:

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} + \begin{pmatrix} \sigma_\theta^2 & \rho\sigma_\theta^2 \\ \rho\sigma_\theta^2 & \sigma_\theta^2 \end{pmatrix}^{-1} \right)^{-1} \cdot \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} s_1 - a_1^b \\ s_2 - a_2^b \end{pmatrix} + \begin{pmatrix} \sigma_\theta^2 & \rho\sigma_\theta^2 \\ \rho\sigma_\theta^2 & \sigma_\theta^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right].$$

Simplifying, this can be rewritten:

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \frac{1}{\sigma_\theta^4(1-\rho^2) + 2\sigma_\theta^2 + 1} \begin{pmatrix} \sigma_\theta^4(1-\rho^2) + \sigma_\theta^2 & \rho\sigma_\theta^2 \\ \rho\sigma_\theta^2 & \sigma_\theta^4(1-\rho^2) + \sigma_\theta^2 \end{pmatrix} \begin{pmatrix} s_1 - a_1^b \\ s_2 - a_2^b \end{pmatrix}. \tag{4}$$

Remark 4.1 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 4 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

$$\frac{\sigma_\theta^4(1-\rho^2) + \sigma_\theta^2}{\sigma_\theta^4(1-\rho^2) + 2\sigma_\theta^2 + 1}.$$

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

$$\frac{\rho\sigma_\theta^2}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}.$$

The first thing we learn from this is that because the two competences are correlated, increased performance on task i improves the Voter's beliefs about the Incumbent's competence on both tasks. 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 correlation:

$$\frac{d}{d\rho}\left(\frac{\sigma_\theta^4(1-\rho^2)+\sigma_\theta^2}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}\right)<0.$$

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

$$\frac{d}{d\rho}\left(\frac{\rho\sigma_\theta^2}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}\right)>0.$$

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

$$\begin{pmatrix} \gamma & 1-\gamma \end{pmatrix}\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}\geq 0. \tag{5}$$

It will be useful to unpack the Voter's expected payoff from reelecting the Incumbent by substituting for (m_1, m_2) from Equation 4. In particular, define

$$\lambda_1 = \left(\frac{1}{\sigma_\theta^4(1-\rho^2) + 2\sigma_\theta^2 + 1}\right)\left(\gamma(\sigma_\theta^4(1-\rho^2) + \sigma_\theta^2) + (1-\gamma)\rho\sigma_\theta^2\right) \tag{6}$$

and

$$\lambda_2 = \left(\frac{1}{\sigma_\theta^4(1-\rho^2) + 2\sigma_\theta^2 + 1}\right)\left(\gamma\rho\sigma_\theta^2 + (1-\gamma)(\sigma_\theta^4(1-\rho^2) + \sigma_\theta^2)\right). \tag{7}$$

Then Condition 5 above can be rewritten in terms of observed outcomes, rather than beliefs. In particular, Condition 5 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 4.2 The Voter's equilibrium reelection rule puts weight λ_1 on task 1 and λ_2 on task 2. From the perspective of first-period Voter welfare, however, it would be optimal for the Incumbent to put all her effort into whichever task the Voter puts greater weight on (i.e., task 1 if $\gamma > 1/2$ and task 2 if $\gamma < 1/2$). The fact that both λ_1 and λ_2 can be positive, and that the marginal cost of effort on each task is zero at zero effort on that task, imply that, in equilibrium, the Voter's preferences and the Incumbent's incentives are not perfectly aligned.

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 that maximize first period welfare and the incentives the Voter actually gives in equilibrium.

4.3 Reelection Probabilities

To calculate the probabilities of reelection that enter the Incumbent's optimization problem, we must consider the prior distribution of the posterior means, conditional on both the actual actions a_1 and a_2 and the expected actions a_1^b and a_2^b. Standard results imply that this distribution is normal (because the updating is linear) with mean

$$\begin{pmatrix} \overline{m}_1 \\ \overline{m}_2 \end{pmatrix} = \frac{1}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1} \begin{pmatrix} \sigma_\theta^4(1-\rho^2)+\sigma_\theta^2 & \rho\sigma_\theta^2 \\ \rho\sigma_\theta^2 & \sigma_\theta^4(1-\rho^2)+\sigma_\theta^2 \end{pmatrix} \begin{pmatrix} a_1 - a_1^b \\ a_2 - a_2^b \end{pmatrix}$$

and variance Σ equal to the prior variance minus the posterior variance, or

$$\Sigma = \begin{pmatrix} \sigma_\theta^2 & \rho\sigma_\theta^2 \\ \rho\sigma_\theta^2 & \sigma_\theta^2 \end{pmatrix} - \frac{1}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1} \begin{pmatrix} \sigma_\theta^4(1-\rho^2)+\sigma_\theta^2 & \rho\sigma_\theta^2 \\ \rho\sigma_\theta^2 & \sigma_\theta^4(1-\rho^2)+\sigma_\theta^2-\rho^2\sigma_\theta^4 \end{pmatrix}.$$

Write

$$\sigma_m^2 = \sigma_\theta^2 - \frac{\sigma_\theta^4(1-\rho^2)+\sigma_\theta^2}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}$$

for the diagonal elements of Σ (the prior variance of a single dimension's posterior mean belief about competence) and write

$$\mathrm{cov}_m = \rho\sigma_\theta^2 \left(1 - \frac{1}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}\right)$$

for the off-diagonal elements (the prior correlation of the posterior mean beliefs about competence).

At the time of the election, the Incumbent will be reelected if $\gamma m_1 + (1-\gamma) m_2 \geq 0$. The left-hand side of this condition is distributed normally, with mean

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

and variance

$$\sigma_b^2 \equiv (\gamma^2 + (1-\gamma)^2)\sigma_m^2 + 2\gamma(1-\gamma)\,\mathrm{cov}_m\,. \tag{8}$$

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

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

Remark 4.3 The prior variance of the Voter's expected payoff from the Incumbent under bundling, σ_b^2, will play an important role in the comparative statics to follow. It is a weighted average of two terms, σ_m^2 and cov_m. Two important facts follow from this observation.

First, the Cauchy-Schwarz inequality implies that $\text{cov}_m \leq \sigma_m^2$, strictly so if $\rho < 1$. Thus σ_b^2 is increasing in the distance of γ from $\frac{1}{2}$, unless $\rho = 1$. Intuitively, this *variance reduction* comes from averaging two imperfectly correlated random variables. In particular, the Voter's second-period welfare, which determines her electoral behavior, is the γ-weighted average of the two dimensions of politician competence. It is also important to note, then, that the magnitude of the variance reduction is decreasing in the correlation, ρ.

Second, each of σ_m^2 and cov_m are increasing in the prior correlation ρ. Since σ_b^2 is a weighted average of these two quantities, it is also increasing in ρ. This *information effect* reflects the fact that, the more correlated are the competences, the more the Voter learns from the signals and, thus, the more spread out is the distribution of the mean of her posterior beliefs.

4.4 First Period Effort

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 payoff are:

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

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 and substituting for c' gives the equilibrium conditions:

$$\frac{\lambda_1 R}{\sigma_b}\phi(0) = a_1^b \quad \text{and} \quad \frac{\lambda_2 R}{\sigma_b}\phi(0) = a_2^b. \tag{9}$$

The marginal benefit of increased effort reflected on the left-hand side of these equilibrium conditions is the product of three terms that we will refer back to later:

(i) The *probability impact*: the change in probability of reelection resulting from a small increase in the Voter's posterior beliefs about the γ-weighted average of the Incumbent's competences, given by $\frac{\phi(0)}{\sigma_b}$;

(ii) The *belief impacts*: λ_1 (resp. λ_2) is the amount a small increase in effort on task 1 (resp. task 2) raises the Voter's expected posterior beliefs about the γ-weighted average of the Incumbent's competences.

(iii) The *benefit of reelection*: R.

It is worth pausing to comment on the meaning of the probability impact both because it will be important in the sequel and because it is subtle.

> **Remark 4.4** The probability impact is decreasing in σ_b. Recall that σ_b is the standard deviation of the prior distribution of the γ-weighted average of posterior mean beliefs about competences. The incumbent is reelected if that γ-weighted average ends up being greater than zero. Since the distribution of the γ-weighted average has mean zero, when σ_b is larger it is less likely that the γ-weighted average will end up close to zero. As such, it is less likely that a small change in beliefs will change the election result, and so the probability impact is decreasing in σ_b. Notice, this also implies from our earlier Remark 4.3, that the probability impact is decreasing in ρ and in $|\gamma - 1/2|$.

4.5 Equilibrium Voter Welfare

The equilibrium that we have characterized gives the Voter expected first-period welfare:

$$\gamma \frac{\lambda_1 R}{\sigma_b} \phi(0) + (1-\gamma) \frac{\lambda_2 R}{\sigma_b} \phi(0).$$

The next result calculates ex-ante expected second-period welfare.

Lemma 4.1 *The Voter's ex-ante expected second-period welfare is $\sigma_b \phi(0)$.*

Proof. The result follows from the same argument that proves Lemma 3.1, with the modification that expected second period welfare conditional on posterior means (m_1, m_2) is

$$\max\{\gamma m_1 + (1-\gamma) m_2, 0\}.$$

■

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 three parts. In this section we focus on the Voter's first period welfare, which depends on incentives. In the next section we consider the Voter's second period welfare, which depends on selection. In both cases we consider how two parameters affect the optimal institutional choice: the correlation between the task-specific competences (ρ) and the relative weight the Voter puts on each task (γ). Finally, we consider the optimal institution for overall Voter welfare.

We will occasionally use the following notation. Let $W_t^i(\rho, \gamma, \sigma_\theta^2)$ be the Voter's welfare in period t under institution i at parameter values $(\rho, \gamma, \sigma_\theta^2)$.

Under either institution, expected first period welfare is :

$$\gamma a_1 + (1-\gamma) a_2 = \frac{1}{2}(a_1 + a_2) + \frac{1}{2}(2\gamma - 1)(a_1 - a_2). \tag{10}$$

In this decomposition, first-period welfare is an average of two factors: (i) the *total effort* $(a_1 + a_2)$ and (ii) the *alignment* between the Incumbent's division of effort across the tasks and the Voter's weighting of the tasks' relative importance $((2\gamma - 1)(a_1 - a_2))$. Note, when $\gamma > 1/2$, the Voter considers task 1 more important and the alignment term is increasing in $a_1 - a_2$. When $\gamma < 1/2$, the Voter considers task 2 more important and the alignment term is decreasing in $a_1 - a_2$.

In the case of unbundling, first period welfare is then very simple. Recall, when $\gamma > 1/2$, we have that $a_1^u = \frac{\phi(0)R}{\sigma_u}\lambda_u$ and $a_2^u = 0$. If $\gamma < 1/2$, these two efforts are reversed. Moreover, recall that $\lambda_u = \frac{\sigma_\theta^2}{\sigma_\theta^2 + 1}$.

In the case of bundling we have that

$$a_1^b + a_2^b = \frac{\phi(0)R}{\sigma_b}(\lambda_1 + \lambda_2)$$

and

$$a_1^b - a_2^b = \frac{\phi(0)R}{\sigma_b}(2\gamma - 1)(\lambda_1 - \lambda_2).$$

Notice, in these calculations, we again have a decomposition into three effects, as in the discussion of the first-order conditions. Total effort is the product of the probability impact, the benefit of reelection, and the sum of the belief impacts of actions 1 and 2. Alignment is the product of the probability impact, the benefits of reelection, and the difference in the belief impacts of actions 1 and 2.

The key to understanding the institutional tradeoffs for period 1 welfare is to observe that the two institutions are ranked in opposite ways by the two components of first-period welfare. In particular, bundling induces more total effort, but unbundling better aligns effort with Voter preferences.

Total effort under bundling is larger than under unbundling for two reasons. First, the probability impact is larger because of the variance reduction highlighted in Remark 4.3. Second, the sum of the belief impacts is strictly higher under bundling when there is positive correlation because the correlation increases the total amount of information, reducing the weight the Voter's posteriors put on her priors.

It is straightforward from Equation 10 that for any $\gamma \neq 1/2$, alignment is maximized when all effort is devoted to the more important task. This is exactly what is attained under unbundling because, as we saw in Lemma 3.2, all of the rewards of office are devoted to that task. However, as we saw in Remark 4.1, for any $\rho > 0$, 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. Such a Voter will vote based only on task 1 beliefs. However, this means he will vote based on the output from both tasks, which does not maximize incentives for task 1 effort. Maximizing task 1 incentives would involve the Voter voting only based on the outcome from task 1, but doing so is not sequentially rational given the correlation between the two competences.

These intuitions are formalized in the following result.

Proposition 5.1 *Suppose that prior correlation (ρ) and preference weight on task 1 (γ) are such that both λ_1 and λ_2 are strictly positive. Then, for any $\sigma_\theta^2 > 0$:*

(i) Total effort is higher under bundling than under unbundling—i.e., $a_1^u + a_2^u < a_1^b + a_2^b$.

(ii) Alignment is better under unbundling than under bundling—i.e., $(2\gamma - 1)(a_1^u - a_2^u) > (2\gamma - 1)(a_1^b - a_2^b)$.

Before turning to a comparative static analysis, it is important to see that either institution can be optimal for first-period Voter welfare.

An immediate implication of Proposition 5.1 is that bundling is always better than unbundling when the Voter cares equally about the two tasks (i.e., $\gamma = 1/2$)—in that case, only total effort matters, and both λ_1 and λ_2 are positive for any value of the correlation.

The case where the Voter cares only about one task (i.e., $\gamma = 1$ or $\gamma = 0$) is less immediate. When $\gamma = 1$, first-period welfare is just a_1. Hence, when $\gamma = 1$, unbundling is strictly preferred if and only if $\frac{\lambda_u}{\sigma_u} > \frac{\lambda_1}{\sigma_b}$. The next result establishes that this is indeed the case, for any strictly positive correlation.

Proposition 5.2 *Either institution can be optimal with respect to first-period Voter welfare. In particular, for any $\sigma_\theta^2 > 0$:*

(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.

The ideas underlying Proposition 5.2 can be extended to provide comparative statics of the optimal institution (for first period welfare). As we show in Proposition 5.3 below, we are able to provide clean comparative statics. Nonetheless, we first want to highlight that there are competing effects away from the limits.

Voter Preferences Consider first the comparative static on γ. Recall that bundling induces higher total effort but that unbundling better aligns the Voter's preferences and the Incumbent's allocation of effort between the two tasks. As the Voter's preferences become more extreme, the total effort advantage of bundling shrinks and the alignment advantage of unbundling grows. Hence, the more extreme the Voter, the more attractive is unbundling.

The total effort advantage of bundling depends on γ through the variance reduction described in Remark 4.3. The variance reduction comes from the fact that the Voter's reelection decision is based on a γ-weighted average of her posterior beliefs about the two dimensions of competence. Hence, as the Voter becomes more extreme, the variance reduction is attenuated.

The alignment advantage of unbundling depends on γ through two competing effects. First, recall that under unbundling, for any $\gamma \neq 1/2$ there is only positive effort on the more important dimension, while under bundling there is effort on both dimensions. Since the Voter's utility is linear in the efforts, this tends to make unbundling more attractive, and increasingly so as preferences become more extreme. Second, the extent of mis-alignment under bundling decreases as preferences become more extreme (whereas there is no change under unbundling), since as preferences become more extreme the Incumbent's incentives are increasingly to focus on the more important task. This effect tends to make bundling more attractive as the Voter becomes more extreme. Although there are competing effects, on net, it turns out that, as the Voter becomes more extreme, the alignment advantage of unbundling increases.

Competence Correlation Now consider the comparative static on ρ. Welfare under unbundling is independent of the correlation. Hence, we can get all the intuition by focusing on bundled welfare.

A change in the correlation affects both total effort and alignment under bundling. The overall effect is that bundled welfare is decreasing in ρ.

In the case of total effort, there are two effects that cut in opposite directions. First, as correlation increases, the sum of the belief impacts increases, which tends to increase total effort. Intuitively, as correlation increases, there is more total information. When there is more total information, the posterior beliefs puts less weight on the Voter's priors. And when the weight on the prior is lower the effect of a small change in the signal, and hence the effect of a small change of effort, on posterior beliefs is larger. Second, as correlation increases, the probability impact decreases. This is because, as we've already discussed, the probability impact is determined by one over the prior variance of the γ-weighted average of the posterior means and, when correlation increases so there is more information, this prior variance is larger. The net effect of these two competing effects is to make total effort single-peaked in ρ.

A change in correlation also affects alignment under bundling in two ways (both negative) which are familiar from our discussion of the intuition underlying Proposition 5.2. First, as before, an increase in ρ decreases the probability impact. Second, an increase in ρ decreases alignment for the reasons discussed earlier—an increase in ρ increases the amount of informational spillover across dimensions, thereby increasing incentives to work on the less important task under bundling.

The results of these intuitions are illustrated in Figure 1 and formalized below.

Proposition 5.3 *There is a function $\hat{\gamma}(\rho, \sigma_\theta^2)$ which is strictly decreasing in ρ and which satisfies $\hat{\gamma}(\rho, \sigma_\theta^2) \geq \frac{1}{2}$ for all (ρ, σ_θ^2), such that bundling is optimal for first period welfare if and only if*

$$1 - \hat{\gamma}(\rho, \sigma_\theta^2) \leq \gamma \leq \hat{\gamma}(\rho, \sigma_\theta^2).$$

The optimality is strict if the inequalities are strict.

Proof. Let $\Delta_1(\gamma, \rho)$ be the difference between bundled and unbundled welfare:

$$\Delta_1(\gamma, \rho, \sigma_\theta^2) = W_1^b(\gamma, \rho, \sigma_\theta^2) - W_1^u(\gamma, \rho, \sigma_\theta^2).$$

Unbundling is strictly optimal if $\Delta_1(\gamma, \rho)$ is strictly negative, bundling is strictly optimal if $\Delta_1(\gamma, \rho)$ is strictly positive, and either institution is optimal if $\Delta_1(\gamma, \rho)$ equals zero.

Lemma B.1 shows that W^b is strictly convex in γ. Since W^u is linear in γ on $[1/2, 1]$, this implies that Δ_1 is strictly convex, and hence strictly quasiconvex, in γ on $[1/2, 1]$. That implies that the set of $\gamma \in [1/2, 1]$ such that $\Delta_1(\gamma, \rho, \sigma_\theta^2) \leq 0$ is an interval.

Proposition 5.2 implies that this interval has the form $[1/2, \hat{\gamma}(\rho, \sigma_\theta^2)]$. Symmetry of the payoff around $\gamma = 1/2$ then implies that bundling is optimal for first-period welfare on the

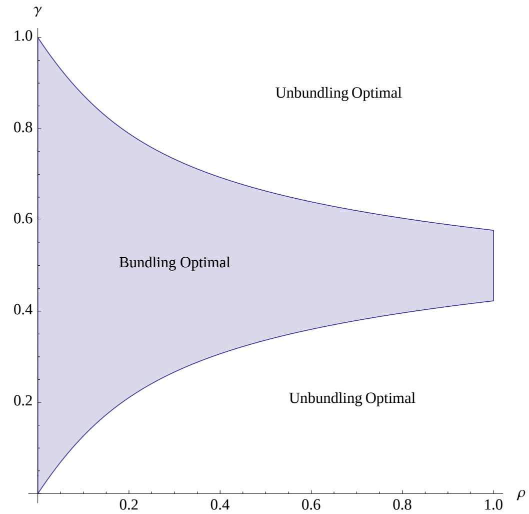

Figure 1: Optimal institution for first period welfare as a function of ρ and γ, for the case of $\sigma_\theta^2 = 1$.

interval $[1 - \hat{\gamma}(\rho, \sigma_\theta^2), \hat{\gamma}(\rho, \sigma_\theta^2)]$.

Finally, Lemma B.2 shows that $W_1^b(\gamma, \rho, \sigma_\theta^2)$ is strictly decreasing in ρ, which implies $\hat{\gamma}(\rho, \sigma_\theta^2)$ is strictly decreasing in ρ. ■

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 first period welfare by creating severe misalignment.

6 Optimal Institution: Second Period Welfare

Lemmas 3.1 and 4.1 show that ex ante, second period Voter welfare under institution i is

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

Critically, notice that ex ante, second period Voter welfare is increasing in the prior variance of the posterior mean of the payoff from reelecting an incumbent, σ_i. The intuition is straightforward—increased dispersion of the posterior means coincides with increased Voter information, and the more information the Voter has, the better job she does at selecting good types.

The welfare comparisons now follow from the two effects highlighted in Remark 4.3.

First, by averaging across the Incumbent's competence on two dimensions, bundling creates variance reduction, which tends to decrease ex ante, second period Voter welfare under bundling relative to unbundling. To see how this works, consider a situation in which an Incumbent is competent on one dimension and incompetent on the other. Bundling forces the Voter to either stick with the Incumbent, who is good at one task but bad at the other, or replace her with the Challenger, who is mean zero at both. In a similar scenario under unbundling—where the Incumbent on one task is competent and the Incumbent on the other task is incompetent—the Voter has the flexibly to stick with the competent Incumbent while replacing the incompetent one.

Second, for any positive correlation across dimensions, there is an information effect associated with bundling. Because both tasks are informative about both dimensions of competence, the Voter has more total information about the quality of the Incumbent under bundling. This increased information allows the Voter to do a better job of selecting good types.

Before turning to a comparative static analysis, it is important to see that either institution can be optimal for ex ante, second period Voter welfare.

When there is no correlation, the only effect of a move from unbundling to bundling is the flexibility effect. Hence, when there is no correlation, unbundling is preferred. When there is perfect correlation, the only effect of a move from unbundling to bundling is the information effect. There are two ways to see why this is the case. First, when correlation is perfect, there is no variance reduction associated with linking the dimensions. Second, and equivalently, when the two dimensions of competence are perfectly correlated, the Voter will never face an incumbent who is competent on one dimension and incompetent on the other, so flexibility is irrelevant. Hence, when correlation is perfect, bundling is preferred.

These intuitions are summarized in the following proposition, which establishes that either institution can be optimal for ex ante, second period Voter welfare.

Proposition 6.1 *Either institution can be optimal with respect to ex ante, expected second-period welfare. In particular, for any* $\sigma_\theta^2 > 0$*:*

(i) If $\rho = 0$*, then unbundling is strictly optimal for all* $\gamma \in (0, 1)$*.*

(ii) If $\rho = 1$*, then bundling is strictly optimal for all* γ*.*

The ideas underlying Proposition 6.1 can be extended to provide an intuition for comparative statics of the optimal institution (for second period welfare). First, consider the effect of γ. As the distance between γ and $1/2$ increases, there is more variance reduction and the information effect is unchanged. Hence, as the Voter cares more about one dimension relative to the other, unbundling becomes relatively more attractive. Second, consider the effect of ρ. As ρ increases, variance reduction is attenuated and the information effect increases. Both of these tend to make bundling relatively more attractive.

These intuitions are formalized in the next proposition and illustrated in Figure 2.

Proposition 6.2 *For any $(\gamma, \sigma_\theta^2)$, there is a unique $\hat{\rho}(\gamma, \sigma_\theta^2) \in [0, 1)$ such that bundling is optimal for second-period welfare if and only if:*

$$\rho \geq \hat{\rho}(\gamma, \sigma_\theta^2).$$

The optimality is strict if the inequality is strict.

Moreover, $\hat{\rho}$ is strictly increasing in γ for $\gamma < 1/2$ and strictly decreasing in γ for $\gamma > 1/2$.

Proof. From Proposition 6.1, at $\rho = 0$ unbundling is optimal and at $\rho = 1$ bundling is optimal. Second period welfare, for institution i is $\sigma_i \phi(0)$. Since, σ_u is independent of ρ, it suffices to show that σ_b is increasing in ρ. Differentiating σ_b^2 with respect to ρ we have:

$$\frac{\partial \sigma_b^2}{\partial \rho} = \frac{2\sigma_\theta^4 \left[\rho\left(1+\sigma_\theta^2\right) + \left(\gamma-\gamma^2\right)(\rho-1)\left((\rho-1)\sigma_\theta^2 - 2\right)\left((1+\rho)\sigma_\theta^2+1\right)^2\right]}{\left(1+2\sigma_\theta^2 - \left(\rho^2-1\right)\sigma_\theta^4\right)^2},$$

which is positive for $\rho \in [0, 1]$, as required.

All that remains is to show that $\hat{\rho}$ is strictly increasing for $\gamma < 1/2$ and strictly decreasing for $\gamma > 1/2$. Since σ_u is independent of γ, this follows from the observation from Remark 4.3 that σ_b is increasing in the distance between γ and $1/2$. ■

These comparative statics are the opposite of those for first-period welfare, suggesting an important tension in designing optimal institutions. We return to this in the next section.

7 Overall Optimal Institution

Thus far, we have studied the comparative static effects of ρ and γ on the optimal institution for first- and second-period welfare separately. Figure 3 summarizes these two welfare analyses by combining Figures 1 and 2. Figure 3 highlights a fundamental trade-off established

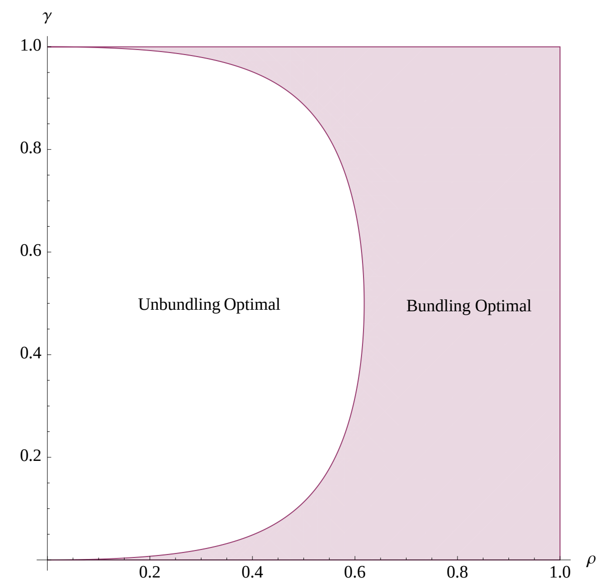

Figure 2: Optimal institution for second period welfare as a function of ρ and γ, for the case of $\sigma_\theta^2 = 1$.

by Propositions 5.3 and 6.2. Factors that increase the attractiveness of one institution for first-period Voter welfare decrease the attractiveness of that institution for second-period Voter welfare. In particular, as γ moves away from 1/2 or ρ goes to 0, unbundling becomes more attractive for first-period welfare and bundling becomes more attractive for second-period welfare.

What we see in Figure 3 is that for some parameter values, there is no trade-off between first- and second-period welfare optimization. If the Voter cares largely about one issue (γ not too far from 1/2) and the task-specific competences are not too highly correlated (ρ not too close to 1), then unbundling is the optimal institution for both periods. If the Voter cares similarly about the two tasks and the task-specific competences are sufficiently correlated, then bundling is the optimal institution for both periods.

However, for other configurations of parameter values, institutional choice poses a trade-off between first- and second-period welfare.

And, indeed, Propositions 5.2 and 6.1, coupled with continuity of the welfares, show this qualitative conclusion is general. For any finite and positive σ_θ^2, there is a region of parameter values such that the optimal institution poses no trade-off between first- and second-period welfare and other regions for which there is such a trade-off.

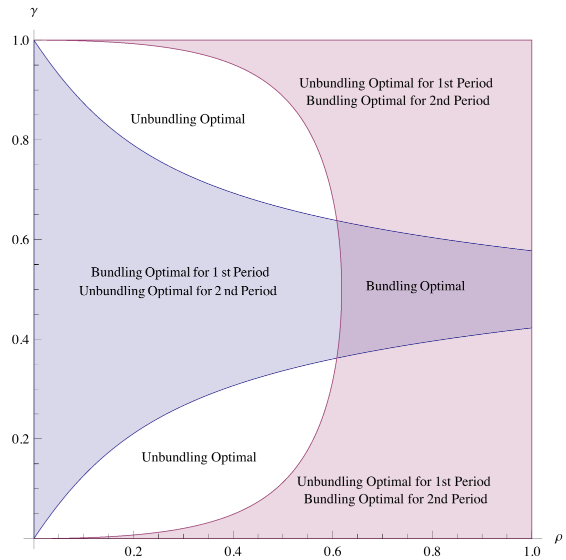

Figure 3: Optimal institution for first- and second-period welfare as a function of ρ and γ, for the case of $\sigma_\theta^2 = 1$.

Of course, the fact that there are trade-offs does not mean that an overall optimal institution does not exist. Here we study this issue by defining overall welfare as the undiscounted sum of first- and second- period welfare. Recall that $W_t^i(\rho, \gamma, \sigma_\theta^2)$ is the Voter's welfare in period t under institution i at parameter values $(\rho, \gamma, \sigma_\theta^2)$. Then, unbundling is the overall optimal institution if and only if:

$$W_1^u(\rho, \gamma, \sigma_\theta^2) + W_2^u(\rho, \gamma, \sigma_\theta^2) \geq W_1^b(\rho, \gamma, \sigma_\theta^2) + W_2^b(\rho, \gamma, \sigma_\theta^2).$$

For some collections of parameter values, this overall welfare analysis depends on how trade-offs between first- and second-period welfare happen to balance out. However, when σ_θ^2 is either very small or very large we can say more.

When σ_θ^2 is very small, there is little heterogeneity in the quality of candidates. As such, electoral selection is relatively unimportant for Voter welfare. Hence, overall welfare is maximized by whichever institution maximizes the Voter's first-period welfare. As the next result shows, which institution maximizes first-period welfare still depends on ρ and γ, but there are no longer any trade-offs between first- and second-period welfare that matter for the overall welfare comparison.

Proposition 7.1 *There is a decreasing function $\hat{\gamma}(\cdot)$ such that:*

- *For any pair (γ, ρ) with $\gamma < \hat{\gamma}(\rho)$, there exists a $\underline{\sigma}(\gamma, \rho)$ such that, for any $\sigma_\theta < \underline{\sigma}(\gamma, \rho)$ bundling is optimal for overall Voter welfare.*
- *For any pair (γ, ρ) with $\gamma > \hat{\gamma}(\rho)$, there exists a $\underline{\sigma}(\gamma, \rho)$ such that, for any $\sigma_\theta < \underline{\sigma}(\gamma, \rho)$ unbundling is optimal for overall Voter welfare.*

Proof.

Recall,

$$\sigma_b^2 = \frac{\sigma_\theta^4}{\sigma_\theta^4(1-\rho^2) + 2\sigma_\theta^2 + 1}\left[\left(\gamma^2 + (1-\gamma)^2\right)\left(\sigma_\theta^2(1-\rho^2) + 1 + \rho^2\right) + 2\gamma(1-\gamma)\left(\sigma_\theta^2\rho(1-\rho^2) + 2\rho\right)\right]$$

Further, we have

$$\sigma_u^2 = \frac{\sigma_\theta^4}{\sigma_\theta^2 + 1}.$$

We can now note several things. First, as $\sigma_\theta^2 \to 0$, both σ_b^2 and σ_u^2 go to zero, so second period welfare goes to the same constant for both. Thus, overall welfare is determined by any differences in first period welfare.

Give this, the result follows from an argument identical to the proof of Proposition 5.3, letting σ_θ^2 equal zero.

■

When σ_θ^2 is very large, there is substantial heterogeneity in candidate quality. This has two implications. First, candidates themselves believe effort is unlikely to be decisive in reelection, so there is little difference across institutions in the power of incentives. Second, electoral selection is very important for the Voter's welfare. Hence, overall Voter welfare is maximized by whichever institution maximizes the Voter's second-period welfare. Further, in this case, the same institution maximizes second-period (and thus overall) Voter welfare for all values of γ and ρ. Recall that the information effect of bundling comes from the fact that, when the Voter gets more informative signals, his posterior beliefs put less weight on his priors, spreading out the prior distribution of posterior means. In the limit, as σ_θ^2 gets infinitely large, the Voter's updating puts no weight on his prior beliefs, for any informative signal. Hence, in the limit, the information effect disappears, leaving only the flexibility effect. This means that, for σ_θ^2 very large, not only does second-period welfare determine the overall optimal institution, but second-period welfare always favors unbundling.

Proposition 7.2 *For any pair (γ, ρ) with $\gamma \notin \{0, 1\}$ and $\rho < 1$, there exists a $\overline{\sigma}(\gamma, \rho)$ such that, for any $\sigma_\theta > \overline{\sigma}(\gamma, \rho)$ unbundling is optimal for overall Voter welfare.*

Proof. The difference between bundled and unbundled overall welfare is the sum of the difference in first period welfares and the difference in second period welfares. The difference in first period welfares is bounded as $\sigma_\theta^2 \to \infty$. We will show that the difference between second period welfare under bundling and second period welfare under unbundling goes to positive infinity as $\sigma_\theta^2 \to \infty$. Hence, unbundling is preferred for σ_θ^2 sufficiently large.

To see that the difference in first-period welfares is bounded, note that any first period effort greater than $\sqrt{2B}$ is dominated by an effort of zero.

Now consider second period payoffs. Recall that

$$\sigma_b^2 = \left(\gamma^2 + (1-\gamma)^2\right) \sigma_m^2 + 2\gamma(1-\gamma)\,\mathrm{cov}_m\,.$$

Hence, we can write the difference in expected second period payoffs as:

$$\phi(0)\left(\sigma_u^2 - \sigma_b^2\right) = \phi(0)\left[\left(\gamma^2 + (1-\gamma)^2\right)\left(\sigma_u^2 - \sigma_m^2\right) + 2\gamma(1-\gamma)\left(\sigma_u^2 - \mathrm{cov}_m\right)\right].$$

Now consider these term-by-term.

$$\sigma_u^2 - \sigma_m^2 = \frac{-\rho^2\sigma_\theta^4}{(1+\sigma_\theta^2)(1+2\sigma_\theta^2+(1-\rho^2)\sigma_\theta^4)},$$

which clearly goes to zero as $\sigma_\theta^2 \to \infty$.

$$\sigma_u^2 - \mathrm{cov}_m = \frac{1+(1-\rho)\sigma_\theta^2+(1-3\rho)\sigma_\theta^4+(2-3\rho+\rho^3)\sigma_\theta^6+(1-\rho)^2(1+\rho)\sigma^8}{1+3\sigma_\theta^2+(3-\rho^2)\sigma_\theta^4+(1-\rho^2)\sigma_\theta^6},0$$

which clearly goes to infinity as $\sigma_\theta^2 \to \infty$. Hence, the difference in second period expected voter welfare goes to infinity as $\sigma_\theta^2 \to \infty$, as required, ■

8 Conclusion

We have analyzed an institutional choice problem within the context of a political agency model in which multi-task issues arise because of correlation in task-specific competences. The model highlights two key trade-offs for Voter welfare when considering a bundled versus unbundled institution.

The first trade-off concerns incentives. Under bundling total effort by the incumbent politician is higher than under unbundling. However, the allocation of that effort across tasks is less fully aligned with Voter preferences than under unbundling.

The second trade-off concerns identifying and retaining high quality politicians. Under bundling, the Voter has more information about the incumbent politician's competences than under unbundling. However, the Voter is constrained to use that information less flexibly under bundling than unbundling.

We then showed how the resolution of each of these trade-offs responds to changes in Voter preference weightings and the underlying information structure. This analysis highlighted a fundamental tension between the two determinants of the optimal institution—factors that push toward making bundling more attractive for incentives (for instance, by making total effort more important than alignment) also push toward making unbundling more attractive for selection (for instance, by making flexibility more important than information). As such, we showed that, while sometimes the same institution is optimal for both facets of Voter welfare, often the Voter faces a real trade-off between choosing an institution that optimizes incentives or one that optimizes selection.

Appendix

A Second-Order Conditions

Lemma A.1 *(i) For any x, we have:*

$$|x\phi(x)| \leq \frac{1}{\sqrt{2\pi e}}.$$

(ii) $\frac{\lambda_u}{\sigma_u} \leq 1$

(iii) $\frac{\lambda_j}{\sigma_b} \leq 1$ *for* $j = 1, 2$

Proof.

(i) The function $x \mapsto x\phi(x)$ is zero at $x = 0$, and it approaches 0 as x tends to either ∞ or $-\infty$, since the Gauss kernel tends to zero faster than any polynomial. The derivative is $\phi(x) - x^2\phi(x)$, so the critical points are 1 and -1. At each of these points, the absolute value of the function is

$$\frac{1}{\sqrt{2\pi}}e^{-(1/2)}.$$

(ii) The main text shows that

$$\frac{\lambda_u}{\sigma_u} = \frac{1}{\sqrt{\sigma_\theta^2 + 1}},$$

which is clearly less than or equal to 1.

(iii) The main text shows that

$$\frac{\lambda_1}{\sigma_b} = \frac{\lambda_1}{\sqrt{\lambda_1^2(\sigma_\theta^2 + 1) + \lambda_2^2(\sigma_\theta^2 + 1) + 2\rho\sigma_\theta^2\lambda_1\lambda_2}}.$$

Square both sides to get

$$\left(\frac{\lambda_1}{\sigma_b}\right)^2 = \frac{\lambda_1^2}{\lambda_1^2(\sigma_\theta^2 + 1) + \lambda_2^2(\sigma_\theta^2 + 1) + 2\rho\sigma_\theta^2\lambda_1\lambda_2},$$

which is clearly less than or equal to 1.

An analogous argument applies to

$$\frac{\lambda_2}{\sigma_b} = \frac{\lambda_2}{\sqrt{\lambda_1^2(\sigma_\theta^2+1) + \lambda_2^2(\sigma_\theta^2+1) + 2\rho\sigma_\theta^2\lambda_1\lambda_2}}.$$

■

Lemma A.2 *(i) Suppose $R < \sqrt{2\pi e}$. Then the function*

$$a \mapsto \eta R\left(1 - \Phi\left(\frac{0 - \lambda_u(a - a_1^u)}{\sigma_u}\right)\right) - \frac{1}{2}a^2 \tag{11}$$

is strictly concave for all a_1^u and all η with $0 \leq \eta \leq 1$.

(ii) Suppose $R < \sqrt{\frac{\pi e}{2}}$. Then the function

$$(a_1, a_2) \mapsto R\left(1 - \Phi\left(\frac{0 - \lambda_1(a_1 - a_1^b) - \lambda_2(a_2 - a_2^b)}{\sigma_b}\right)\right) - \frac{1}{2}a_1^2 - \frac{1}{2}a_2^2 \tag{12}$$

is strictly concave for all a_1^b and a_2^b.

Proof.

(i) Differentiate the function at (11) twice to get

$$\eta R\left(\frac{\lambda_u}{\sigma_u}\right)^2\left[-\left(\frac{0 - \lambda_u(a - a_1^u)}{\sigma_u}\right)\phi\left(\frac{0 - \lambda_u(a - a_1^u)}{\sigma_u}\right)\right] - 1.$$

Thus the function is strictly concave for any a_1^u if

$$\left|\eta R\left(\frac{\lambda_u}{\sigma_u}\right)^2\left[-\left(\frac{0 - \lambda_u(a - a_1^u)}{\sigma_u}\right)\phi\left(\frac{0 - \lambda_u(a - a_1^u)}{\sigma_u}\right)\right]\right| < 1.$$

To see that this holds, use parts (i) and (ii) of Lemma A.1, $\eta \leq 1$, and the hypothesis

$R < \sqrt{2\pi e}$ to write:

$$\begin{aligned}
&\left|\eta R\left(\frac{\lambda_u}{\sigma_u}\right)^2\left[-\left(\frac{0-\lambda_u(a-a_1^u)}{\sigma_u}\right)\phi\left(\frac{0-\lambda_u(a-a_1^u)}{\sigma_u}\right)\right]\right| \\
&\quad= \eta R\left(\frac{\lambda_u}{\sigma_u}\right)^2\left|\left[-\left(\frac{0-\lambda_u(a-a_1^u)}{\sigma_u}\right)\phi\left(\frac{0-\lambda_u(a-a_1^u)}{\sigma_u}\right)\right]\right| \\
&\quad\leq R\frac{1}{\sqrt{2\pi e}} \\
&\quad< 1.
\end{aligned}$$

(ii) The function at (12) is strictly concave if its Hessian is negative definite. This imposes two requirements.

First, the diagonal elements must be negative:

$$R\left(\frac{\lambda_j}{\sigma_b}\right)^2\left[-\left(\frac{0-\lambda_1(a_1-a_1^b)-\lambda_2(a_2-a_2^b)}{\sigma_b}\right)\phi\left(\frac{0-\lambda_1(a_1-a_1^b)-\lambda_2(a_2-a_2^b)}{\sigma_u}\right)\right]-1<0$$

for $j = 1, 2$. The argument from part (i) shows that this holds for $R < \sqrt{\frac{\pi e}{2}} < \sqrt{2\pi e}$.

Second, the determinant of the Hessian must be positive. Writing

$$\alpha = \frac{0-\lambda_1(a_1-a_1^b)-\lambda_2(a_2-a_2^b)}{\sigma_b}$$

to conserve space, this requires

$$\left(R\left(\frac{\lambda_1}{\sigma_b}\right)^2[-\alpha\phi(\alpha)]-1\right)\left(R\left(\frac{\lambda_2}{\sigma_b}\right)^2[-\alpha\phi(\alpha)]-1\right)-\left(R\frac{\lambda_1}{\sigma_b}\frac{\lambda_2}{\sigma_b}[-\alpha\phi(\alpha)]\right)^2>0.$$

Multiply out and cancel terms to reduce this to

$$-R\left(\frac{\lambda_1}{\sigma_b}\right)^2[-\alpha\phi(\alpha)]-R\left(\frac{\lambda_2}{\sigma_b}\right)^2[-\alpha\phi(\alpha)]+1>0.$$

A sufficient condition for this is

$$\left|R\left(\frac{\lambda_1}{\sigma_b}\right)^2[-\alpha\phi(\alpha)]\right|<\frac{1}{2}.$$

A simple modification of the argument for part (i) shows that this condition is satisfied if

$$R < \frac{1}{2}\sqrt{2\pi e} = \sqrt{\frac{\pi e}{2}}.$$

■

B Proofs of Propositions

Proof of Proposition 5.1.

(i) Recall that

$$a_1^u + a_2^u = \phi(0) R\left(\frac{\lambda_u}{\sigma_u}\right).$$

and

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

Since $\frac{\lambda^u}{\sigma_u} = \frac{1}{\sqrt{\sigma_\theta^2 + 1}}$, it suffices to show:

$$\frac{\lambda_1 + \lambda_2}{\sigma_b}\sqrt{\sigma_\theta^2 + 1} > 1.$$

Squaring both sides and using the fact that $\sigma_b = \sqrt{\lambda_1^2(\sigma_\theta^2 + 1) + \lambda_2^2(\sigma_\theta^2 + 1) + 2\rho\sigma_\theta^2\lambda_1\lambda_2}$, this is equivalent to:

$$\lambda_1^2(\sigma_\theta^2 + 1) + \lambda_2^2(\sigma_\theta^2 + 1) + 2\lambda_1\lambda_2(\sigma_\theta^2 + 1) > \lambda_1^2(\sigma_\theta^2 + 1) + \lambda_2^2(\sigma_\theta^2 + 1) + 2\lambda_1\lambda_2\rho\sigma_\theta^2,$$

which follows from $\rho \leq 1$.

(ii) The argument proceeds as in the previous point except we now have

$$\begin{aligned}\left((\lambda_1 - \lambda_2)\sqrt{\sigma_\theta^2 + 1}\right)^2 &= \lambda_1^2(\sigma_\theta^2 + 1) + \lambda_2^2(\sigma_\theta^2 + 1) - 2(\sigma_\theta^2 + 1)\lambda_1\lambda_2 \\ &< \lambda_1^2(\sigma_\theta^2 + 1) + \lambda_2^2(\sigma_\theta^2 + 1) + 2\lambda_1\lambda_2\rho\sigma_\theta^2.\end{aligned}$$

■

Proof of Proposition 5.2.

(i) Recall that $\frac{\lambda_u}{\sigma_u} = \frac{1}{\sqrt{\sigma_\theta^2 + 1}}$. To calculate $\frac{\lambda_1}{\sigma_b}$ for the case of $\gamma = 1$, substitute $\gamma = 1$ into Equation 6 to get

$$\lambda_1 = \frac{\sigma_\theta^4(1-\rho^2)+\sigma_\theta^2}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}$$

and into Equation 8 to get

$$\sigma_b^2 = \frac{\sigma_\theta^4(\sigma_\theta^2(1-\rho^2)+1+\rho^2)}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}.$$

The result follows from

$$\begin{aligned}
\frac{\lambda_1}{\sigma_b} &= \frac{\frac{\sigma_\theta^4(1-\rho^2)+\sigma_\theta^2}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}}{\sqrt{\frac{\sigma_\theta^4(\sigma_\theta^2(1-\rho^2)+1+\rho^2)}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}}} \\
&= \frac{1}{\sqrt{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}} \cdot \frac{\sigma_\theta^2(1-\rho^2)+1}{\sqrt{\sigma_\theta^2(1-\rho^2)+1+\rho^2}} \\
&< \frac{1}{\sqrt{\sigma_\theta^2+1}},
\end{aligned}$$

which holds for all $\rho \neq 0$.

(ii) This follows directly from Proposition 5.1.

■

Proofs of Lemmas for Proposition 5.3

Lemma B.1 *First-period welfare under bundling, $W_1^b(\gamma,\rho,\sigma_\theta^2)$, is a strictly convex function of γ for all ρ and σ_θ^2.*

Proof. Recall that

$$\begin{aligned}
W_1^b(\gamma,\rho,\sigma_\theta^2) &= \gamma\frac{\lambda_1 R}{\sigma_b}\phi(0) + (1-\gamma)\frac{\lambda_2 R}{\sigma_b}\phi(0) \\
&= R\phi(0)\left(\frac{\gamma\lambda_1+(1-\gamma)\lambda_2}{\sigma_b}\right).
\end{aligned}$$

Differentiating the term in parentheses twice and using the definitions of the λs and σ_b, we get

$$\frac{\partial^2 W_1^b}{\partial \gamma^2} \propto (\rho^2-1)\sigma_\theta^2 \left[-1-7\rho^2+(-1-2\rho^2+3\rho^4)\sigma_\theta^2-2\gamma(\rho-1)^2(3\rho-1)(1+(1+\rho)\sigma_\theta^2)+2\gamma^2(\rho-1)^2(3\rho-1)(1+(1+\rho)\sigma_\theta^2)\right]$$

$$\Bigg[\left(1+\sigma_\theta^2-\rho^2(\sigma_\theta^2-1)-2\gamma(\rho-1)^2(1+(1+\rho)\sigma_\theta^2)+2\gamma^2(\rho-1)^2(1+(1+p)\sigma_\theta^2)\right)^2$$

$$\times\sqrt{-\frac{\sigma_\theta^4(1-2\gamma(\rho-1)^2+2\gamma^2(\rho-1)^2+\rho^2+(-1+2(\gamma-1)\gamma(\rho-1))(\rho^2-1)\sigma_\theta^2)}{-1-2\sigma_\theta^2+(\rho^2-1)\sigma_\theta^4}}\Bigg]^{-1}.$$

The denominator is positive, so the whole second derivative has the same sign as the numerator. Since $(\rho^2-1)\sigma_\theta^2<0$, it suffices to show:

$$-1-7\rho^2+(-1-2\rho^2+3\rho^4)\sigma_\theta^2-2\gamma(\rho-1)^2(3\rho-1)(1+(1+\rho)\sigma_\theta^2)+2\gamma^2(\rho-1)^2(3\rho-1)(1+(1+\rho)\sigma_\theta^2)<0. \tag{13}$$

We will show that an upper bound on the left-hand side of Condition 13 is always negative. To get an upper bound, start by maximizing the left-hand side of Condition 13 with respect to γ.

There are two cases.

(i) $\rho>\mathbf{\frac{1}{3}}$: Here the LHS of 13 is maximized at $\gamma=0$ or $\gamma=1$. So the LHS of 13 is less than or equal to

$$-1-7\rho^2+(-1-2\rho^2+3\rho^4)\sigma_\theta^2,$$

which is clearly negative for any ρ.

(ii) $\rho\le\mathbf{\frac{1}{3}}$: Here the LHS of 13 is maximized at $\gamma=1/2$. So the LHS of 13 is less than or equal to

$$-1-7\rho^2+(-1-2\rho^2+3\rho^4)\sigma_\theta^2-\frac{1}{2}(\rho-1)^2(3\rho-1)(1+(1+\rho)\sigma_\theta^2).$$

Collecting terms, this can be rewritten:

$$-\frac{1}{2}(1+\rho)^2(1+3\rho)+\frac{1}{2}(1+\rho)^2(-1-2\rho+3\rho^2)\sigma_\theta^2,$$

which is clearly negative since both terms are negative for any ρ.

■

Lemma B.2 *For any γ and σ_θ^2, first-period voter welfare under bundling is strictly decreasing in ρ.*

Proof. Recall that first-period welfare under bundling is

$$\begin{aligned}
\frac{1}{2}(a_1^b + a_2^b) + \frac{1}{2}(2\gamma - 1)(a_1^b - a_2^b) &= \frac{\phi(0)R}{2}\left(\frac{\lambda_1 + \lambda_2}{\sigma_b} + (2\gamma - 1)\frac{\lambda_1 - \lambda_2}{\sigma_b}\right) \\
&= \frac{\phi(0)R}{2}\left(\frac{\frac{(1-\rho^2)\sigma_\theta^4 + (1+\rho)\sigma_\theta^2}{(1-\rho^2)\sigma_\theta^4 + 2\sigma_\theta^2 + 1} + (2\gamma - 1)^2 \cdot \frac{(1-\rho^2)\sigma_\theta^4 + (1-\rho)\sigma_\theta^2}{(1-\rho^2)\sigma_\theta^4 + 2\sigma_\theta^2 + 1}}{\sqrt{(\gamma^2 + (1-\gamma)^2)\sigma_m^2 + 2\gamma(1-\gamma)\operatorname{cov}_m}}\right) \\
&= \frac{\phi(0)R}{2}\left(\frac{(1-\rho^2)\sigma_\theta^4 + (1+\rho)\sigma_\theta^2 + (2\gamma - 1)^2\left((1-\rho^2)\sigma_\theta^4 + (1-\rho)\sigma_\theta^2\right)}{\left((1-\rho^2)\sigma_\theta^4 + 2\sigma_\theta^2 + 1\right)\sqrt{(\gamma^2 + (1-\gamma)^2)\sigma_m^2 + 2\gamma(1-\gamma)\operatorname{cov}_m}}\right).
\end{aligned}$$

Define $\tilde{\gamma} = 2\gamma(1-\gamma) = 2\gamma - 2\gamma^2$. Then we have

$$(2\gamma - 1)^2 = 4\gamma^2 - 4\gamma + 1 = 1 - 2\tilde{\gamma}$$

and

$$\gamma^2 + (1-\gamma)^2 = 2\gamma^2 - 2\gamma + 1 = 1 - \tilde{\gamma}.$$

With these observations, we can write bundled welfare as

$$\frac{\phi(0)R}{2}\left(\frac{(1-\rho^2)\sigma_\theta^4 + (1+\rho)\sigma_\theta^2 + (1-2\tilde{\gamma})\left((1-\rho^2)\sigma_\theta^4 + (1-\rho)\sigma_\theta^2\right)}{\left((1-\rho^2)\sigma_\theta^4 + 2\sigma_\theta^2 + 1\right)\sqrt{(1-\tilde{\gamma})\,\sigma_m^2 + \tilde{\gamma}\operatorname{cov}_m}}\right).$$

Differentiate with respect to ρ to get

$$\begin{aligned}
&\frac{\phi(0)R}{2\left[(1-\tilde{\gamma})\sigma_m^2 + \tilde{\gamma}\operatorname{cov}_m\right]^{3/2}} \times \\
&\left[\frac{\partial}{\partial\rho}\left(\frac{(1-\rho^2)\sigma_\theta^4 + (1+\rho)\sigma_\theta^2 + (1-2\tilde{\gamma})\left((1-\rho^2)\sigma_\theta^4 + (1-\rho)\sigma_\theta^2\right)}{\left((1-\rho^2)\sigma_\theta^4 + 2\sigma_\theta^2 + 1\right)}\right)\left((1-\tilde{\gamma})\sigma_m^2 + \tilde{\gamma}\operatorname{cov}_m\right)\right. \\
&\left.-\frac{1}{2}\left(\frac{(1-\rho^2)\sigma_\theta^4 + (1+\rho)\sigma_\theta^2 + (1-2\tilde{\gamma})\left((1-\rho^2)\sigma_\theta^4 + (1-\rho)\sigma_\theta^2\right)}{\left((1-\rho^2)\sigma_\theta^4 + 2\sigma_\theta^2 + 1\right)}\right)\frac{\partial}{\partial\rho}\left((1-\tilde{\gamma})\sigma_m^2 + \tilde{\gamma}\operatorname{cov}_m\right)\right].
\end{aligned}$$

The first term is positive, so this derivative has the same sign as the bracketed factor:

$$\frac{\partial}{\partial\rho}\left(\frac{(1-\rho^2)\sigma_\theta^4+(1+\rho)\sigma_\theta^2+(1-2\tilde{\gamma})\left((1-\rho^2)\sigma_\theta^4+(1-\rho)\sigma_\theta^2\right)}{\left((1-\rho^2)\sigma_\theta^4+2\sigma_\theta^2+1\right)}\right)((1-\tilde{\gamma})\sigma_m^2+\tilde{\gamma}\,\mathrm{cov}_m)$$

$$-\frac{1}{2}\left(\frac{(1-\rho^2)\sigma_\theta^4+(1+\rho)\sigma_\theta^2+(1-2\tilde{\gamma})\left((1-\rho^2)\sigma_\theta^4+(1-\rho)\sigma_\theta^2\right)}{\left((1-\rho^2)\sigma_\theta^4+2\sigma_\theta^2+1\right)}\right)\frac{\partial}{\partial\rho}\left((1-\tilde{\gamma})\sigma_m^2+\tilde{\gamma}\,\mathrm{cov}_m\right)$$

Substituting

$$\sigma_m^2=\frac{\sigma_\theta^4\left(\sigma_\theta^2(1-\rho^2)+1+\rho\right)}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}$$

$$\mathrm{cov}_m=\frac{\rho\sigma_\theta^4\left(\sigma_\theta^2(1-\rho^2)+2\right)}{\sigma_\theta^4(1-\rho^2)+2\sigma_\theta^2+1}$$

and simplifying the term in brackets, we have that the derivative of bundled welfare with respect to ρ has the same sign as

$$2\frac{-4\rho+\tilde{\gamma}^2(\rho^2+\rho-2)^2-\tilde{\gamma}(\rho^4-3\rho^2-8\rho+4)}{(\rho^2-4)^2}.$$

So we need

$$-4\rho+\tilde{\gamma}^2(\rho^2+\rho-2)^2-\tilde{\gamma}(\rho^4-3\rho^2-8\rho+4)<0.$$

Observe that

$$(\rho^2+\rho-2)^2=\rho^4+2\rho^3-3\rho^2-4\rho+4.$$

and that

$$\begin{aligned}\rho^4-3\rho^2-8\rho+4&=\rho^4+2\rho^3-3\rho^2-4\rho+4-2\rho^3-4\rho\\&=(\rho^2+\rho-2)^2-2\rho^3-4\rho.\end{aligned}$$

Thus the expression we want to be negative is

$$-4\rho+(\tilde{\gamma}^2-\tilde{\gamma})(\rho^2+\rho-2)^2+\tilde{\gamma}(2\rho^3+4\rho).$$

Since $\tilde{\gamma}<1$, we have $\tilde{\gamma}^2-\tilde{\gamma}<0$, so all that is left is to show that

$$-4\rho+\tilde{\gamma}(2\rho^3+4\rho)<0.$$

Since $\tilde{\gamma} \leq \frac{1}{2}$, the worst case for us is $\tilde{\gamma} = \frac{1}{2}$, where the expression is negative if

$$-2 + \rho^2 < 0,$$

which clearly holds. ■

Proof of Proposition 6.1. By Lemmas 3.1 and 4.1, to establish the results, we need only compare σ_u to σ_b.

Recall that

$$\sigma_u = \sqrt{\frac{\sigma_\theta^4}{\sigma_\theta^2 + 1}}.$$

and

$$\sigma_b = \sqrt{\frac{\sigma_\theta^4}{\sigma_\theta^4(1-\rho^2) + 2\sigma_\theta^2 + 1} \left[(\gamma^2 + (1-\gamma)^2)\left(\sigma_\theta^2(1-\rho^2) + 1 + \rho^2\right) + 2\gamma(1-\gamma)\left(\sigma_\theta^2\rho(1-\rho^2) + 2\rho\right)\right]}$$

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

$$\sqrt{\frac{\sigma_\theta^4\left[\gamma^2 + (1-\gamma)^2\right]}{\sigma_\theta^2 + 1}}.$$

This is strictly less than $\sqrt{\frac{\sigma_\theta^4}{\sigma_\theta^2+1}}$ for any $\gamma \notin \{0, 1\}$.

(ii) Evaluating σ_b at $\rho = 1$ yields:

$$\sqrt{\frac{2\sigma_\theta^4}{2\sigma_\theta^2 + 1}},$$

which is greater than $\sqrt{\frac{\sigma_\theta^4}{\sigma_\theta^2+1}}$.

■

References

Ahn, David S. and Santiago Oliveros. 2012. "Combinatorial Voting." *Econometrica* 80(1):89–141.

Ahn, David S. and Santiago Oliveros. 2013. "The Condorce Jur(ies) Theorem." Unpublished Working Paper.

Alesina, Alberto and Guido Tabellini. 2007. "Bureaucrats or Politicians? Part I: A Single Policy Task." *American Economic Review* 97(1):169–179.

Alesina, Alberto and Guido Tabellini. 2008. "Bureaucrats or Politicians? Part II: Multiple Policy Tasks." *Journal of Public Economics* 92(3):426–447.

Alt, James E., Ethan Bueno de Mesquita and Shanna Rose. 2011. "Disentangling Accountability and Competence in Elections: Evidence from US Term Limits." *Journal of Politics* 73(1):171–186.

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

Ashworth, Scott and Ethan Bueno de Mesquita. 2006. "Delivering the Goods: Legislative Particularism in Different Electoral and Institutional Settings." *Journal of Politics* 68(1):169–179.

Baker, George. 2002. "Distortion and risk in optimal incentive contracts." *Journal of Human Resources* pp. 728–751.

Berry, Christopher R. 2009. *Imperfect Union: Representation and Taxation in Multilevel Governments*. New York, NY: Cambridge University Press.

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

Berry, Christopher R. and Jacob E. Gersen. 2009. "Fiscal Consequences of Electoral Institutions." *Journal of Law and Economics* 52:469–495.

Besley, Timothy. 2005. *Principled Agents: Motivation and Incentives in Politics*. Oxford: Oxford University Press.

Besley, Timothy and Anne Case. 1995. "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits." *Quarterly Journal of Economics* 110(3):769–798.

Besley, Timothy and Anne Case. 2003. "Political Institutions and Policy Choices: Evidence from the United States." *Journal of Economic Literature* 41(1):7–73.

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

Boffa, Federico, Amedeo Piolatto and Giacomo A.M. Ponzetto. 2013. "Centralization and Accountability: Theory and Evidence from the Clean Air Act." CEPR Discussion Paper No. DP9514.

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 Nicholas Terrell. 2009. "The Fatally Flawed Theory of the Unbundled Executive." *Minnesota Law Review* 93:1696–1740.

de Janvry, Alain, Frederico Finan and Elisabeth Sadoulet. 2011. "Local Electoral Incentives and Decentralized Program Performance." *Review of Economics and Statistics* 94(3):672–685.

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

Dewatripont, Mathias, Ian Jewitt and Jean Tirole. 1999. "The Economics of Career Concerns, Part 1: Comparing Information Structures." *Review of Economic Studies* 66:183–198.

Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, ed. Adam Przeworski, Susan Stokes and Bernard Manin. New York: Cambridge University Press.

Feltham, Gerald A and Jim Xie. 1994. "Performance measure congruity and diversity in multi-task principal/agent relations." *Accounting review* pp. 429–453.

Finan, Frederico and Claudio Ferraz. 2011. "Electoral Accountability and Corruption in Local Governments: Evidence from Audit Reports." *American Economic Review* 101:1274–1311.

Gehlbach, Scott. 2007. "Electoral Institutions and the National Provision of Local Public Goods." *Quarterly Journal of Political Science* 2:5–25.

Gibbons, Robert. 2010. "Inside Organizations: Pricing, Politics, and Path Dependence." *Annual Review of Economics* 2:337–365.

Glaeser, Edward L. and Giacomo A.M. Ponzetto. 2014. "Shrouded Costs of Government: the Political Economy of State and Local Public Pensions." *Journal of Public Economics* 116:89–105.

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

Holmström, Bengt. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *The Review of Economic Studies* 66(1):169–182.

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.

Lohmann, Susanne. 1998. "An information rationale for the power of special interests." *American Political Science Review* pp. 809–827.

Maskin, Eric and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94(4):1034–1054.

Persson, Torsten and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge: MIT Press.

Prat, Andrea. 2005. "The Wrong Kind of Transparency." *American Economic Review* 95(3):862–877.

Swank, Otto H. and Bauke Visser. 2006. "Do Elections Lead to Informed Public Decisions?" *Public Choice* 129:435–460.

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