Career Concerns and Policy Experimentation under Decentralization[*]

Chen Cheng[†]   Christopher Li[‡]

March 13, 2015

Preliminary, Comments are Welcome

Abstract

We examine policy experimentation in an electoral environment with imperfect information. The voters face both uncertainty about the competency of policymakers as well as a lack of transparency of policymakers' actions. The outcome of experimental policies serves as a noisy signal of politicians' competency, and voters reelect those they believe to be competent. This creates career concerns like incentives (a la Holmstrom, 1999) for the policymakers, which induces policymakers to increase the variance of signals voters receive. Interestingly, the incentive to increase variance leads to more experimentation under decentralization than under centralization. Another result of paper is that centralization does have an advantage in terms of selecting competent policymakers. The implications of transparency of policy choice, non-career-concern payoff of politicians are also discussed.

1 Introduction

Experimentation is essential for discovering better policies. Its importance is vividly captured by an anecdote Margaret Thatcher revealed in the documentary *Commanding Heights*:

> I remember the foreign minister and the finance minister from another country saying to me: "You are the first prime minister who has ever tried to roll back the frontiers of socialism. We want to know what's going to happen. Because if you succeed others will follow.

Political institutions, through shaping the incentives of policymakers, affect policy experimentation. One strand of literature in political economy examines how decentralized policymaking (e.g. federalism) affects experimentation. A consistent message found in this literature is that decentralization discourages policy experimentation.[1] The broad argument is similar to the one on public goods provision: decentralization creates free-riding incentives and thus reduces experimentation. Nonetheless, there is evidence that suggests experimentation by local governments is prevalent, and perhaps more so than by national government. Regarding the experiences of the United States, Oates (1999) notes that "since the dawn of the nation, programs successfully developed at the state level have often provided models for subsequent federal programs." One recent example is the health care reform in Massachusetts, which later became a blue print for the Affordable Care Act. There are instances where the policymaking power is delegated to local governments with the hope of more experimentation. Decentralization of welfare programs in the US in the late 90's is one such example.

To accommodate the stylized facts, this paper explores the connection between decentralization and policy experimentation from a new prospective. We consider an environment where both the competency of policymakers and the outcome of policies (experiments) affect voter welfare, but the voters have difficulty disentangling the two effects because of imperfect information. Moreover, the voters can only discipline self-interested policymakers through reelection. Against this backdrop, we examine the implication of decentralization on voter's information, and consequently the policymaker's incentive to experiment. In doing so, we identify forces that induces higher experimentation under decentralization, thus providing a counterpoint to existing theoretical works.

---

[1]Early discussions such as Oates (1972) argue that decentralization promotes experimentation, but they assume policy uniformity under centralization. See the section below for a more detailed discussion.

The key feature of the model is that policymakers have *career concerns* like incentives. Because the voters care about policymakers' competence, the policymakers have incentives to manipulate voters' belief via policy choices. More specifically, voters only reelect politicians perceived to be of high ability, and this creates a discontinuity in payoffs. Thus the policymaker has an incentive to decrease the informativeness of the signals voters observe. Interestingly, this incentive manifests itself differently under centralization and decentralization. Under centralization, the central policymaker undertakes the same policy (experiment) in all regions. When a central policymaker chooses the same policies in all regions, the observed outcomes convey less information about the policymaker's ability than if those policies differ. This is because outcomes of the same experiment across regions are correlated. On the other hand, local policymakers have incentives to adopt different policies under decentralization. Here, adopting the same policies conveys more information regarding the local policymakers' abilities because the voter in one region can extract information from outcomes in other regions.

In addition to policy experimentation, this paper also speaks to the relationship between electoral accountability and decentralization in Section 4. While decentralization encourages more experimentation in our model, it performs worse in the selection of competent policymakers. Moreover, this accountability gap between centralization and decentralization is increasing in the heterogeneity among districts. Our finding on accountability is supported by empirical studies. Treisman (2000) for example, using an average of various corruption indices produced by Transparency International, finds that federal countries are more likely to be corrupt. We thus point out a new tradeoff in terms of institutional design: decentralization encourages more experimentation, but the selection of politicians suffers.

We also discuss the implications of several assumptions in the model. First we see what will happen if we allow policy choices to be observable. Although it seems apparent that voters will benefit from more transparent policies, our analysis shows that it actually depends on politicians' payoff. When politicians' payoff is only dependent on his expected competency, our results in previous sections reverses: politicians now want to increase (rather than decrease) the precision of voter's information, which results in a better selection of politicians. While our results in previous sections remain if politicians' payoff is linearly dependent on outputs of whole economy.

We then discuss how our results depend on politicians' career concern incentive. We show when politician simply get a fixed benefit if in office and nothing if replaced, whether they still have incentive to increase the variance of voter's information depends on if they have *in-*

cumbency advantage. When they have incumbency advantage, they actually want to increase the precision of voter's information; and vice versa if they don't.

The rest of the paper is organized as following. In the next section, we discuss the relevant literature. Section 2 sets up the model. Section 3 characterizes the equilibrium. In particular, we analyze voters' electoral rule and policymaker's policy choice under centralization and decentralization, for different specifications of policymakers' payoff. Section 4 discusses electoral accountability and its comparative statics with regards with district heterogeneity. Section 5 explores various extensions of the model. Section 6 concludes.

1.1   Related Literature

While there is a sentiment in the policy circle that decentralization/federalism promotes experimentation,[2] formal analysis challenges this view. The pioneering work of Rose-Ackerman (1980) argues that when policy experimentation is costly, decentralization creates free-riding and reduces experimentation relative to centralization. Free riding incentive also plays a central role in most other papers in this topic (e.g.Strumpf (2002), Cai and Treisman (2009), and Callander and Harstad (2013)). Note that while policymakers in Rose-Ackerman (1980) and Cai and Treisman (2009) have less incentive to experiment under decentralization even though they face reelection pressure.[3] Strumpf (2002) and Callander and Harstad (2013) observe that when the districts are sufficiently heterogeneous, decentralization may induce more experimentation. However, their results rely crucially on the assumption that policy must be uniform across regions under centralization. We do not impose this assumption. Moreover, they consider benevolent policymaker whereas the misaligned interest between the politician and voter is at the core of our argument. Therefore, the forces at work are inherently different between our model and theirs. The paper that is closest in spirit to ours is Kotsogiannis and Schwager (2006). They examine policymakers' incentive to experiment when there is a prospect of running for higher office. It is shown that when the benefit of experiment is small (i.e. probability of success is low), experimentation is greater under decentralization than under centralization. While their motivation is similar to ours, the model and mechanism is different. In particular, they study a signaling game where the competency of policymaker is private information and policy choice is observable.

As mentioned before, our modeling approach contains many elements from works on career

---

[2]See discussion in Oates (1999).

[3]They consider an electoral environment without any informational friction.

concerns, e.g. Holmström (1999). It also share similarities with electoral accountability models that focuses on candidate selection, e.g. Ashworth and de Mesquita (2012). We show centralization ensures higher electoral accountability when compared to decentralization. This is in contrary to the insights of previous papers on this topic. Lockwood (2005) and Persson and Tabellini (2002) argue that centralization induces weaker selection of policymakers, though the environments they consider are quite different from ours. In Persson and Tabellini (2002), the policy outcome in a particular region is only informative of the policymaker's ability in that region. In our model, the voter may use outcomes in other regions since they can be informative. Lockwood (2005) studies asymmetric information between voter and politician on competence, while we focus on symmetric uncertainty.

To summarize, this paper contributes to the literature by being the first to study the interaction between career concern incentive and policy experimentation under different political regimes. More importantly, we obtain implications of decentralization on experimentation and accountability that are different from the existing literature.

2 The Model

Consider a two-period economy that is comprised of two districts $d \in \{x, y\}$. Each district is populated with a representative voter (her). Voters choose policymakers, who are characterized by their competency/ability θ. Policymakers implement policy (experiment) $p \in \{a, b\}$ for each district d.

Voter in each district consumes the output in her district per period. Let c_t^d denote that output of district d at time t. The output at time 1 depends on the competency of the policymaker and the outcome of the policy. In particular, for district d:

$$c_1^d = \theta_1^d + \epsilon_p^d$$

where θ_t^d denote the policymaker's competency at date t, and ϵ_p^d is the outcome of policy p^d. The output at date 2 depends only on the competency of the policymaker:

$$c_2^d = \theta_2^d$$

We consider two forms of government:

- *Centralization*: There is one central policymaker. The policymaker chooses (possibly different) policies for the two districts.

- *Decentralization*: There is a policymaker for each district. The policymaker in each district independently selects policy for his own district.

Centralization captures the governance structure of unitary states such as France and the United Kingdom. Decentralization, on the other hand, corresponds the governance structure of federalist states, such as United States and Canada.

Elections and Information

There is an incumbent policymaker at period 1, whose ability is symmetrically unknown.[4] All players share the prior for the incumbent's competency, which we assume to be a normal distribution $N(0, v_\theta)$. Thus, under centralization, the competency of the incumbent, $\theta_1^x = \theta_1^y = \theta_1$, is drawn from $N(0, v_\theta)$. Under decentralization, the competence district policymakers are two independent draws from $N(0, v_\theta)$. When no confusion arises, we will simply use θ to refer to the competency of the incumbents.

The policymaker(s) implements a policy for each district. The outcomes of policies are uncertain at the time of implementation. This captures the idea that the policies are experiments. We assume that competency (θ^x, θ^y) are independent of policy outcomes $(\epsilon_p^x, \epsilon_p^y)$. Moreover, if policies in the districts are different, then $\epsilon_p^x, \epsilon_p^y$ are two independently draws from the normal distribution $N(0, 1)$. If the policy in the two districts are the same, ϵ_p^x is correlated with ϵ_p^y. More specifically, the two shocks are distributed bivariate normal:

$$\begin{bmatrix} \epsilon_p^x \\ \epsilon_p^y \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$$

where $\rho \in [0, 1)$ is the covariance/correlation coefficient. One can interpret ρ as a measure of homogeneity of the districts: the more homogeneous the districts, the higher the correlation between the outcomes of the same policy.

There is an election at the end of the first period, where voters decide whether to reelect incumbents. In the case of decentralization, each district holds an independent election. In the case of centralization, it will be obvious from below that voters from both districts would adopt the same action. At the time of the election, the first period outputs in the two

---

[4]This is a standard assumption in career concerns models as well as many electoral accountability models e.g. Ashworth and de Mesquita (2012).

districts $s \equiv \{c_1^x, c_1^y\}$ are the only information observable to voters. In particular, we assume for now that in addition to the uncertainties described above, the voters do not observe which policy is implemented. This captures the lack of transparency typically faced by voters. The relevance of this assumption is examined in Section 5.1. Note that under either forms of government, the competency of the policymaker at date 2 is unchanged if the voter reelect the policymaker. If replacement happens, the competency of the new policymaker is a draw from the prior distribution $N(0, v_\theta)$.

Payoffs

The voters make reelection decision to maximize their expected stream of consumption (i.e. $c_1^d + c_2^d$).[5] The voter's strategy, as a function of observable information s, is denoted $\lambda(s) \in \{0, 1\}$, where the value 0 represents replacement and 1 means reelection.

The incumbent's payoff depends on reelection. If he is not reelected, he obtains a payoff of 0, while if he is reelected, he obtains a payoff that is a function of realized outputs, denoted as $w(s)$. The policymaker is expected utility maximizer with Bernoulli utility function:

$$u(s) = \begin{cases} 0 & \lambda(s) = 0 \\ w(s) & \lambda(s) = 1 \end{cases}$$

We consider explicit forms of $w(s)$ in the analysis of our model:

1. *Rent-Extraction:* For some $\kappa > 0$, let $w(s) = \kappa(c_1^x + c_1^y)$ under centralization; $w(s) = \kappa c_1^d$ for district d government under decentralization.

2. *Reputation:* $w(s) = \mathbb{E}_{voter}(\theta | s)$, where the expectation is induced by voter's posterior belief over θ.

In the first formulation of $w(s)$, we imagine the scenario where the politician can extract some rent from the economy while in office. Thus, in the case of centralization, the rent is a function of the entire national output; while under decentralization, the district government extracts rent from its own district.

In the second formulation, we assume the policymaker's payoff is the expectation of the voter's posterior belief of their competency. This captures the "reputational concern" story found in standard career concern models (e.g. Holmström, 1999). In the context of our

---

[5]We assume, without the loss of generality, there is no time discounting.

model, this can be interpreted as measure of public opinion of the incumbent. Thus, the second formulation captures the scenario where the incumbent cares about his reputation and obtains ego/legacy rent.

It is worth noting here that, unlike in standard experimentation, we assume the policy is costless to the politician. This is without the loss of generality as long as policy have equal costs, which follows from our assumption that the policies are ex-ante identical. Note that the zero-cost assumption precludes free-riding incentive and allows us to isolate the incentives of experimentation due to reelection. In particular, the policy choice induces a distribution over s. This affects the policymaker's reelection prospect and consequently his payoff since voter's decision is based on observed information s.

Timing

The sequence of events in our model is as following:

- 0. Nature draws incumbent's competency, which is unobserved to any actor.

- 1. The incumbent(s) chooses policies, and then policy shocks $\{\epsilon_p^x, \epsilon_p^y\}$ are realized. Note that under decentralization, the district incumbents choose policy simultaneously.

- 2. The voters observes output from both districts, i.e. $s = \{c_1^x, c_1^y\}$, and decide whether to reelect incumbent(s).

- 3. Reelection happens and last period output realizes. The incumbent consumes $w(s)$ if gets reelected, 0 otherwise.

We focus on pure strategy equilibria of this game.

3 Policy Experimentation

In this section, we analyze the policy choice of the incumbent(s) under centralization and decentralization. We will show that driven by career concern and the possibility of losing reelection, politicians always want to increase the variance of signals that voters utilize. However, this incentive manifests them differently under centralization than under decentralization. The policymaker chooses the same policy in both districts under centralization; while under decentralization, the policymakers adopt different policies. Since the outcome

of the policies are unknown a priori, there is more experimentation when diverse policies are implemented in the districts.

We will begin the analysis by commenting on the strategic incentives of various players in this game.

The Voter: Voters have two dimensions of decisions to make: what signals s to use to optimally deduce politician's competence, and based on those signals, what electoral rule $\lambda(s)$ to use to replace/re-elect the incumbent(s). We will discuss the choice of electoral rule here, and leave the discussion of optimal choice of signal s later.

Since either policy induces the same distribution of payoff, the voter's strategy $\lambda(s)$ is driven only by considerations of date 2 payoff. Moreover, since the expected competency of the replacement at date 2 is zero, the voter would reelect the incumbent if the expected competency conditional on observing s is greater than 0. Formally, let θ stand for the competency of the relevant incumbent, then given a particular posterior, voters' incentive compatible electoral rule $\lambda^*(s)$ is:

$$\lambda^*(s) = \begin{cases} 1 & \text{if } \mathbb{E}_{voter}\left(\theta|s\right) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\mathbb{E}_{voter}$ denotes expectation based on voter's posterior. Note that since policy choice is unobservable to the voter, the posterior of θ and subsequently $\mathbb{E}_{voter}$ are based on her belief about the policy choice. In equilibrium, her belief is correct. Nonetheless, this is important to keep in mind when considering potential deviations by the incumbent.

The Incumbent(s): Note that ex-ante, the policies are identical. Thus, the incentive from choosing particular policy is not due to intrinsic differences in policies. Rather, the incentives are determined by how policy choice affects voters' beliefs about policymakers' ability. In other words, the incumbent faces career concerns like incentives. Formally, given the representative voter adopting a strategy $\lambda(s)$ and in the case of decentralization, the policy choice of the incumbent in the other district, the incumbent chooses policy p in the following optimization problem:

$$\max_{\rho} \int \lambda(s)w(s)dF\left(s|p\right)$$

where $F\left(s|p\right)$ is the distribution of output s if policy p is chosen. Note that s and p are both vectors, implying that under either decentralization or centralization, voters can utilize information from both districts.

3.1 Rent-Extraction

In this subsection, we explore policymakers' choice when they have rent-extraction payoff. We will show that experimentation is greater under decentralization in the sense that incumbent adopts uniform policies under centralization but diverse policies under decentralization.

Centralization

There is one incumbent who chooses policy for each district. This implies that $\theta^x = \theta^y = \theta$. There are two implications to this. First, it is without the loss of generality to consider only one representative voter. This follows from the fact that the voter in each district have common second period consumption, and they observe the same information prior to election. Second and more importantly, output from both districts, c_1^x and c_1^y, are informative of politician's competence θ regardless of the policy choice of the incumbent.

In particular, given the symmetry in the distribution of $\{\epsilon_p^x, \epsilon_p^y\}$, results on Normal Bayesian updating (see Theorem 1 in DeGroot (1970) page 175) indicates that the sample mean,

$$s_c \equiv \frac{c_1^x + c_1^y}{2},$$

is a sufficient statistics for θ.

Lemma 3.1. *Under centralization (i.e. $\theta^x = \theta^y = \theta$), the sample mean of district output $s_c = \frac{c_1^x + c_1^y}{2}$ is a sufficient statistic for $\theta|s$. In other words, $\theta|s = \theta|s_c$.*

Proof. $\theta|s$ is distributed normal with mean and variance given by formula in Theorem B.7 in Greene. One can verify that this is the same distribution as $\theta|s_c$ which is a normal with mean $\frac{\text{cov}(\theta, s_c)}{\text{var}(s_c)} \cdot s_c$ and variance $\frac{\text{cov}^2(\theta, s_c)}{\text{var}(s_c)}$.

□

Given the above analysis, we will show in the following proposition, that politician under centralization has incentive to choose the same policy in both districts instead of diverse policies.

Proposition 3.1. *Under centralization, the incumbent chooses the same policy for both districts in equilibrium.*

Proof. We will first consider voter's equilibrium strategy. Suppose the voter believe the the covariance between ϵ_p^x and ϵ_p^y to be ρ, Lemma 3.1 allows us to express voter's posterior mean of θ as a product of s_c and a function of ρ:

$$\mathbb{E}[\theta|s] = \delta(\rho) \cdot s_c$$

where $\delta(\rho) \equiv \frac{v_\theta}{\frac{1+\rho}{2}+v_\theta}$. Recall that the voter reelect the incumbent if and only if $\mathbb{E}[\theta|s] \geq 0$. Since $\delta(\rho) > 0$, it follows that in equilibrium, the voter's strategy $\lambda^*(s) = 1$ if and only if $s_c \geq 0$. , the voters have a dominant strategy where they reelect the incumbent if and only if $s_c \geq 0$. Denote that strategy as λ^*.

We now consider the incumbent's policy choice in equilibrium. Note that the distribution of s_c depends on the incumbent's policy choice, whether the policy is uniform or different between districts. Observing that $w(s) = 2\kappa s_c$, we can write the incumbent's expected utility as:

$$\begin{aligned}
\mathbb{E}[u(s)] &= \int_{s_c} \lambda^*(s_c)w(s_c)dF(s_c|p) \\
&= \int_{s_c>0} 2\kappa s_c dF(s_c|p) \\
&= 2\kappa \cdot \frac{\sigma(s_c|p)}{\sqrt{2\pi}}
\end{aligned}$$

where $p = (p_x, p_y)$ denotes the policy for each district, and $\sigma(s_c|p)$ is the standard deviation of s_c given policy p. Note the last equality comes from the formula $\mathbb{E}[x|x > 0] = \sigma\sqrt{\frac{2}{\pi}}$ for $x \sim N(0, \sigma^2)$, and the observation that s_c is a normal with mean 0.

In equilibrium, the incumbent chooses p to maximizes $\mathbb{E}[u(s)] = 2\kappa \cdot \frac{\sigma(s_c|p)}{\sqrt{2\pi}}$. Now, if $p_x = p_y$ (uniform policies), then there is correlation between c_1^x and c_1^y through correlation between ϵ_p^x and ϵ_p^y. In this case, $\sigma^2(s_c|p) = v_\theta + \frac{\rho+1}{2}$. If $p_x \neq p_y$ (diverse policy), then there is no correlation between c_1^x and c_1^y. In this case, $\sigma^2(s_c|p) = v_\theta + \frac{1}{2}$. Clearly, the variance is higher under uniform policy than diverse policy. It follows that the incumbent have strict incentive to adopt uniform policy.

$\square$

Decentralization

Now each district has an independent government who implements the policy. Denote θ^d as the competency of the incumbent in district d. We will sometimes abuse the notation and let θ^d represent the incumbent in district d.

Recall that for the voter, θ^x and θ^y are independent draws from $N(0, v_\theta)$. Clearly, for voter in district d, the own district output c_1^d is always informative of θ^d. Whether the output in

the other district is informative depends on policymakers' policy choices. In particular, if different policies were chosen, then c_1^x and c_1^y are statistically independent; whereas if the same policies were chosen, c_1^x and c_1^y are correlated through the correlation between ϵ_ρ^x and ϵ_ρ^y. Thus, the output in the other district is informative of the competency of own district's government if policies are uniform.

If voter in district d believes that policies in the districts are different, then his posterior belief satisfies $\mathbb{E}[\theta^d|s] \geq 0$ if and only if $c_1^d \geq 0$. It follows that the optimal electoral rule is $\lambda(s) = 1$ if and only if $c_1^d \geq 0$.

The posterior belief is more complicated when voters believe that policies are uniform across districts.

Lemma 3.2. *Without the loss of generality, we consider the posterior belief of district x voter only. Define $s_\alpha = c_1^x - \left(\frac{\alpha}{v_\theta+1}\right) c_1^y$. For any $\alpha \in [0,1)$, the posterior of incumbent's ability θ conditional on s is normal with:*

$$\mathbb{E}[\theta^x|s] = \left(\frac{v_\theta}{v_\theta + \frac{1+v_\theta-\alpha^2}{1+v_\theta}}\right) s_\alpha$$

$$\sigma^2[\theta^x|s] = v_\theta - \frac{v_\theta^2}{v_\theta + \frac{1+v_\theta-\alpha^2}{1+v_\theta}}.$$

In particular, s_α is a sufficient statistic for $\theta^x|s$.

Proof. Applying Theorem 1 in DeGroot (1970) page 175, we obtain that the posterior for (θ^x, θ^y) is a bivariate normal with mean vector given by:

$$\left(\begin{bmatrix} \frac{1}{v_\theta} & 0 \\ 0 & \frac{1}{v_\theta} \end{bmatrix} + \begin{bmatrix} \frac{1}{1-\alpha^2} & \frac{-\alpha}{1-\alpha^2} \\ \frac{-\alpha}{1-\alpha^2} & \frac{1}{1-\alpha^2} \end{bmatrix}\right)^{-1} \left(\begin{bmatrix} \frac{1}{1-\alpha^2} & \frac{-\alpha}{1-\alpha^2} \\ \frac{-\alpha}{1-\alpha^2} & \frac{1}{1-\alpha^2} \end{bmatrix} \cdot \begin{bmatrix} c_1^x \\ c_1^y \end{bmatrix}\right)$$

after some algebra, one obtains that:

$$\mathbb{E}[\theta^x|s] = \frac{v_\theta}{v_\theta + \frac{1+v_\theta-\alpha^2}{1+v_\theta}} \underbrace{\left(c_1^x - \frac{\alpha}{1+v_\theta} c_1^y\right)}_{s_\alpha}$$

and;

$$V[\theta^x|s] = v_\theta - \frac{v_\theta^2}{v_\theta + \frac{1+v_\theta-\alpha^2}{1+v_\theta}}$$

12

One can verify that those two moments is the same as the moments of $\theta|s_\alpha$, thus proving s_α is a sufficient statistic.

Note that doing this way is equivalent to do the following: first get the posterior of ϵ^x using c^y as signal, and then calculate the posterior of θ^x based on c^x taking into account the new "prior" of ϵ^x. Note also, that s_α is indeed a sufficient statistic for s i.e. $\theta^x|s_\alpha = \theta^x|s$.

<div align="right">□</div>

There is empirical evidence that supports voters in reality do in fact take the outcome in other districts into account (see Bordignon, Cerniglia and Revelli, 2004).

Corollary 3.1. *For a given $\alpha \in [0, 1)$, the inequality $\mathbb{E}(\theta|s) \geq 0$ holds if and only if $s_\alpha \geq 0$. Thus, the optimal electoral rule for the voter is such that $\lambda(s) = 1$ if and only if $s_\alpha \geq 0$.*

Proof. It follows from the fact that $\mathbb{E}[\theta^x|s_\alpha] = \frac{v_\theta}{v_\theta + \frac{1 + v_\theta - \alpha^2}{1 + v_\theta}} \cdot s_\alpha$.

<div align="right">□</div>

We are now ready to prove that under decentralization, the district governments will adopt different policies in equilibrium.

Proposition 3.2. *Under decentralization, the incumbents choose different policies for the districts in equilibrium.*

Proof. First we will show that adopting the same policies in both district cannot be an equilibrium. Suppose the voters believe the incumbents adopt the same policy in both districts. Then according to the lemma above, (for voter in district x), $\alpha = \rho > 0$ and $s_\alpha = c_1^x - \left(\frac{\rho}{v_\theta + 1}\right) \cdot c_1^y$ is a sufficient statistic for $\theta|s$, and $\lambda(s) = 1$ if and only if $s_\alpha \geq 0$. We will now examine the incentives for the incumbent in district x (without the loss of generality). Using the law of iterated expectation, we can rewrite the expected utility as:

$$\begin{aligned}
\mathbb{E}\left[u(s)\right] &= \int_s \lambda(s)w(s)dF(s|p) \\
&= \kappa \int_{s_\alpha > 0} \mathbb{E}\left(c_1^x|s_\alpha\right) dF(s_\alpha|p) .
\end{aligned}$$

where $p = (p^x, p^y)$. While s_α is determined through voter's belief, the incumbent can manipulate both $\mathbb{E}\left(c_1^x|s_\alpha\right)$ and $F(s_\alpha|p)$ through his policy choice. More specifically, one can write

<div align="center">13</div>

the formula using standard results:

$$\mathbb{E}\left(c_1^x | s_\alpha\right) = \frac{\text{cov}\left(c_1^x, s_\alpha\right)}{\sigma^2(s_\alpha)} s_\alpha$$

$$F(s_\alpha | p) \sim N\left(0, \sigma^2(s_\alpha)\right)$$

where the variance $\sigma^2(s_\alpha)$ is the variance of s_α that's induced by the strategy. We can now write the expected utility as:

$$\mathbb{E}\left[u(s)\right] = \kappa \cdot \frac{\text{cov}\left(c_1^x, s_\alpha\right)}{\sigma^2(s_\alpha)} \cdot \frac{\sigma(s_\alpha)}{\sqrt{2\pi}}$$

$$= \kappa \cdot \frac{\text{cov}\left(c_1^x, s_\alpha\right)}{\sigma(s_\alpha) \cdot \sqrt{2\pi}}$$

Define $A \equiv 1 + v_\theta$ and $C \equiv \frac{\rho^2}{1+v_\theta}$, one can show using standard properties of normal that $\text{cov}\left(c_1^x, s_\alpha\right)$ is equal to $A - C$, if $\rho^x = \rho^y$ and equal to A if $\rho^x \neq \rho^y$. In addition, $\sigma(s_\alpha)$ is equal to $\sqrt{A - C}$ if $\rho^x = \rho^y$ and is equal to $\sqrt{A + C}$ if $\rho_x \neq \rho_y$. Thus, we can simplify the expression to.

$$\frac{\text{cov}\left(c_1^x, s_\alpha\right)}{\sigma(s_\alpha)} = \begin{cases} \sqrt{A - C} & \text{if } \rho_x = \rho_y \\ \frac{A}{\sqrt{A+C}} & \text{if } \rho_x \neq \rho_y \end{cases}$$

We divide the bottom expression by the top one, and obtains $\frac{A}{\sqrt{(A+C)(A-C)}} = \frac{A}{\sqrt{A^2-C^2}} > 1$. Thus, the expected utility for incumbent of district s is higher if he were to choose a policy different than the one in district y. Therefore, uniform policy cannot be sustained in an equilibrium.

We will now show choosing diverse policy is an equilibrium. Note that in this case, the voter will only use her own district's output in updating the incumbent's competence. And since the policy induces the same distribution on payoffs, the incumbent have no incentive to deviate from equilibrium.

$\square$

3.2 Reputation

In this section, we explore a alternative specification of politician's reelection payoff $w(s)$. Politicians care about how voters perceive their ability. More specifically, let

$$w(s) = \mathbb{E}_{voter}[\theta|s]$$

where $\mathbb{E}_{voter}[\theta|s]$ is expectation of voter's posterior belief. This is the standard preference for the agents in career concerns literature (see Holmström, 1999 and Dewatripont, Jewitt and Tirole, 1999). It is important to note that voter's posterior and therefore $\mathbb{E}_{voter}$ depends on voter's belief of policy choice. While in general actual policy choice may differ from voter belief, in equilibrium this belief is correct. We show in the following analysis that our results in previous section continue to hold.

Proposition 3.3. *Under centralization, the incumbent adopt uniform policy in both districts in equilibrium. Under decentralization, the incumbents adopt diverse policies in equilibrium. In other words, experimentation is greater under decentralization than under centralization.*

Proof. Much of the machinery of this proof is borrowed from section 3. We first discussion the case of centralization. As shown above, $s_c \equiv \frac{c_1^x + c_1^y}{2}$ is a sufficient statistic for $\theta|s$, i.e. $\mathbb{E}_{voter}[\theta|s] = \mathbb{E}_{voter}[\theta|s_c]$. We would like to show uniform policy is an equilibrium. Using Lemma 3.1, if the voter believes the district policies to be the same, then the voter's posterior is :

$$\theta|s_c \sim N\left(\frac{v_\theta}{\frac{1+\rho}{2} + v_\theta} \cdot s_c, \frac{\left(\frac{1+\rho}{2}\right)v_\theta}{\frac{1+\rho}{2} + v_\theta}\right)$$

Clearly, the voter would reelect the incumbent if and only if $s_c \geq 0$. Thus, the incumbent's expected utility is:

$$\mathbb{E}[u(s)] = \int_{s_c \geq 0} \frac{v_\theta}{\frac{1+\rho}{2} + v_\theta} \cdot s_c dF(s_c)$$

$$= \frac{v_\theta}{\frac{1+\rho}{2} + v_\theta} \int_{s_c \geq 0} s_c dF(s_c)$$

where the distribution of s_c depends on the incumbent's policy choice. It easy to shown that $\int_{s_c \geq 0} s_c dF(s_c)$ is maximized when the incumbent chooses uniform policy. This in turn justifies the voter's (equilibrium) belief.

To show that diverse policy is not an equilibrium under centralization. Notice that if the voters were to believe that policies are diverse, then $\mathbb{E}_{voter}[\theta|s_c] = \frac{v_\theta}{1+2v_\theta} \cdot s_c$. It is straightforward to see that the incumbent maximizes his expected utility by choosing uniform policies. Thus the voters' belief cannot be sustained in equilibrium.

Now, we would like to show that under decentralization, diverse policy is the feature of equilibrium. If the voter believes that the policy is diverse, then she would treat only her own district's output as informative signal. And since the distribution of the own district's output is unaffected by the policy choice, the (district) incumbent has no incentive to deviate. Thus making diverse policy an equilibrium. On the other hand, suppose the voter believes the incumbents chose the same policy. Then Lemma 3.2 tells us that for voter in district x:

$$\mathbb{E}_{voter}[\theta|s] = \frac{v_\theta}{v_\theta + \frac{1+v_\theta-\rho^2}{1+v_\theta}} \cdot s_\alpha$$

where $s_\alpha = c_1^x - \left(\frac{\rho}{v_\theta+1}\right) \cdot c_1^y$. Clearly, the voter would reelect the incumbent if and only if $s_\alpha \geq 0$. Thus, the incumbent's expected utility is:

$$\mathbb{E}[u(s)] = \int_{s_\alpha \geq 0} \frac{v_\theta}{v_\theta + \frac{1+v_\theta-\rho^2}{1+v_\theta}} \cdot s_\alpha dF(s_\alpha)$$

$$= \left(\frac{v_\theta}{v_\theta + \frac{1+v_\theta-\rho^2}{1+v_\theta}}\right) \int_{s_\alpha \geq 0} s_\alpha dF(s_\alpha)$$

Again, let $A \equiv 1 + v_\theta$ and $C \equiv \frac{\rho^2}{1+v_\theta}$. If the incumbents adopt diverse policies, then

$$s_\alpha \sim N(0, A + C)$$

If the incumbents adopt uniform policy, then

$$s_\alpha \sim N(0, A - C)$$

Notice that the variance of s_α is higher under diverse policies, which implies that $\int_{s_\alpha \geq 0} s_\alpha dF(s_\alpha)$ is greater under diverse policies than under uniform policies. Therefore, the voter's belief about uniform policy cannot be sustained in equilibrium.

$\square$

We have shown that the specific forms of payoffs are not essential to our results. The

informational structure as well as the discontinuity of payoffs are the essential driver of our insights. Below, we provide intuitions for our results.

3.3   Discussion

We have shown that decentralization induces greater level of experimentation than centralization. To understand the intuition behind our result, first note two aspects of the incumbent's payoff. First, the incumbent's payoff is, roughly speaking, increasing in the realized output, which serves as signal of the competency. Second, the payoff is discontinuous because of the possibility of being out of office. Finally, the the threshold in terms of expected ability is above the prior. These three features induce incumbent(s) to increase the dispersion of the signals (or equivalently, decreasing the informativeness of the signal). Below, we formalize this intuition.

Proposition 3.4. *Let s be a normally distributed random variable (e.g. signal) with mean 0, and the incumbent's bernoulli utility be:*

$$u(s) = \begin{cases} 0 & s < c \\ w\left(s\right) & s \geq c \end{cases}$$

where $c \geq 0$ and $w(s)$ is an increasing function. Then, the incumbent's expected utility is increasing in the variance of s.

Proof. We only need to show politician's expected payoff:

$$\int_{s>c} w(s)f(s)ds$$

is increasing in $\sigma^2 \equiv \text{var}(s)$, where $f(s) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}$. This can be proved by showing that

$$\frac{\partial \int_{s>c} w(s)f(s)ds}{\partial \sigma} = \frac{1}{\sqrt{2\pi}\sigma^2}\int_{s>c} w(s)(\frac{s^2}{\sigma^2} - 1)e^{-\frac{s^2}{2\sigma^2}}ds > 0.$$

Assume without the loss of generality that $w(0) \geq 0$, clearly, when $c \geq \sigma$ the above inequality

17

holds, since each quantity in the intergrand is positive. Now, suppose $c < \sigma$, then

$$
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{s>c} w(s)(\frac{s^2}{\sigma^2} - 1)e^{-\frac{s^2}{2\sigma^2}} dF(s)
$$

$$
= \frac{1}{\sqrt{2\pi\sigma^2}} [\int_{s>\sigma} w(s)(\frac{s^2}{\sigma^2} - 1)e^{-\frac{s^2}{2\sigma^2}} dF(s) + \int_{c<s<\sigma} w(s)(\frac{s^2}{\sigma^2} - 1)e^{-\frac{s^2}{2\sigma^2}} dF(s)]
$$

$$
\geq \frac{1}{\sqrt{2\pi\sigma^2}} [\int_{s>\sigma} w(\sigma)(\frac{s^2}{\sigma^2} - 1)e^{-\frac{s^2}{2\sigma^2}} dF(s) + \int_{c<s<\sigma} w(\sigma)(\frac{s^2}{\sigma^2} - 1)e^{-\frac{s^2}{2\sigma^2}} dF(s)]
$$

$$
= \frac{w(\sigma)}{\sqrt{2\pi\sigma^2}} [\int_{s>c} (\frac{s^2}{\sigma^2} - 1)e^{-\frac{s^2}{2\sigma^2}} dF(s)]
$$

$$
> 0
$$

The equation after the first inequality utilizes both the condition that $s > 0$, so that $\frac{s^2}{\sigma^2} - 1 < 0$ for all $0 < s < \sigma$, and the condition that $w(s)$ is an increasing function of s. This results to all increasing functions $w(s)$ since expected utility preference is invariant to positive affine transformation of the bernoulli utility function. Finally, note that as long as the replacement threshold s^* is larger than $-\sigma$, politicians have incentive to increase the variance of signals.

$\square$

We have shown that the incumbent has an incentive to increase the variance of signal in the abstract setting described above. In our model, this incentive manifests itself differently under centralization and decentralization. Under centralization, the output in both districts are informative signals of incumbent's competency. In this scenario, choosing uniform policies induces more correlation between outputs than choosing diverse policies. Since the informativeness of the signal is decreasing in the correlation between district outputs, the incumbent has an incentive to adopt uniform policies. Under decentralization, the outputs are independent when the policies are different, while they are partially correlated under uniform policies. In the former case, each voter treats only his own district's output as an informative signal; in the latter case, each voter takes into account the output in the other district in addition to the own district's output. Thus adopting diverse policy decrease the informativeness of signals, and therefore the incumbent has an incentive to adopt diverse policies in equilibrium.

4 Electoral Accountability

In this section, we compare electoral accountability between centralization and decentralization. Since the voters have a direct interest in the selection of incumbent, we define electoral accountability to be the expected quality of the policymaker at date 2, i.e. $\mathbb{E}[\theta_2]$. This is equivalently the ex-ante voter welfare. We show below that electoral accountability is higher under centralization than under decentralization.

Proposition 4.1. *Electoral Accountability is higher under centralization than under decentralization.*

Proof. Since policy choice have no effect on the first period payoff (the policies induced the same distribution of output), it is without the loss of generality to consider the voter's second period utility, which is the competency of the second period government. Under centralization, the incumbent chooses uniform policy in equilibrium and the voter reelect the incumbent if and only if $\frac{c_1^x + c_1^y}{2} \geq 0$. Since the expected competency of the replacement is zero. The expected second period utility for the voter is:

$$U_c = \Pr\left\{\frac{c_1^x + c_1^y}{2} \geq 0\right\} \cdot \mathbb{E}\left[\theta_1 | \frac{c_1^x + c_1^y}{2} \geq 0\right]$$

$$= \frac{v_\theta}{\sqrt{\frac{1+\rho}{2} + v_\theta}} \cdot \sqrt{\frac{1}{2\pi}}$$

where the last equality comes from using the law of iterated expectations: i.e. $\Pr\left\{\frac{c_1^x + c_1^y}{2} \geq 0\right\} \cdot \mathbb{E}\left[\theta_1 | \frac{c_1^x + c_1^y}{2} \geq 0\right] = \int_{\frac{c_1^x + c_1^y}{2} \geq 0} \mathbb{E}\left[\theta_1 | \frac{c_1^x + c_1^y}{2}\right] dF$. Now, under decentralization, the incumbents choose diverse policies in equilibrium, and therefore the voter's expected second period utility is (for district x voter):

$$U_{dc} = \Pr\{c_1^x \geq 0\} \mathbb{E}[\theta_1 | c_1^x \geq 0]$$

$$= \frac{v_\theta}{\sqrt{1 + v_\theta}} \cdot \sqrt{\frac{1}{2\pi}}$$

Since $\rho < 1$, we have $U_c > U_{dc}$.

$\square$

It's worthwhile to note that, this finding is different from most paper on decentralization

and electoral accountability. Seabright (1996) and Persson and Tabellini (2002) model the idea that centralization makes the incumbent less dependent on pleasing the voters in any particular jurisdiction, and thus less accountable overall. However, they do not discuss the possibility that more jurisdictions might provide more signals (better information) concerning central official's ability. Others argue the tax and yardstick competition between jurisdictions under decentralization might improve accountability (Belleflamme and Hindriks, 2005; Besley and Case, 1995; Besley and Coate, 2003; Bordignon, Cerniglia and Revelli, 2004). While the optimality of yardstick competition is exogenously given in their models, it is endogenously abandoned by voters under decentralization in our paper. The important message here is, when politicians can choose policies strategically, it may no longer optimal for voters to use yardstick competition as a one to discipline government officials.

Next we see how above results affected by heterogeneity across districts.

Corollary 4.1. *The difference in electoral accountability is decreasing in the homogeneity across districts ρ.*

Proof. From proof above, it is easy to see the derivative of the expression $U_c - U_{dc}$ is negative in ρ.

$\square$

Thus we have shown that electoral control in terms of the selection of policymakers is greater under centralization than under decentralization. Furthermore, the advantage of centralization is decreasing in the homogeneity of districts, as measured by ρ. This result is quite different than existing views as noted in the discussion of literature.

We now provide some discussion of the result and its intuition. The mechanism depends on the fact that the selection of candidates depends on the amount of information accessible to voters at date 1. Under centralization, output from both districts are (in equilibrium) informative of the incumbent's competency. While under decentralization, own districts output is the sole informative signal about the (district) incumbent. Intuitively speaking, a centralization provides better information regarding incumbent competency. To formally show this, let $\tau(\theta|s) = \frac{1}{\text{var}(\theta|s)}$ denote the precision the posterior distribution of competency in equilibrium.[6] We shall show that $\tau(\theta|s)$ is higher under centralization.

---

[6]For centralization, θ is the competency of the national incumbent. For decentralization, θ is the competency of the incumbent for a given district.

Under centralization, define $s_c \equiv \frac{c_1^x + c_1^y}{2}$. Therefore, in equilibrium,

$$\tau(\theta | s_c) = \frac{1}{v_\theta} + \frac{2}{1 + \rho},$$

where ρ is the covariance between the shocks. Under decentralization, define $s \equiv c_1^d$. Therefore, in equilibrium,

$$\tau(\theta | c_1^d) = \frac{1}{v_\theta} + 1.$$

Since $\rho \in (0, 1)$, we immediately have

$$\tau(\theta | s_c) > \tau(\theta | c_1^d),$$

which means voters possess better information about the incumbent under centralization than under decentralization.

Since the voters are less mistakenly to reelect an incumbent of low competency i.e. $\theta < 0$ given more information, it follows that electoral control is greater under centralization than decentralization. One way to interpret this result is that greater scope of power results in better selection of incumbents. In particular, the voters observe more/better signals regarding the incumbent when he is the decision maker on a bigger set of policies.

The intuition behind Corollary 4.1 is also based in understanding the interaction between information and district homogeneity. Using the notation from above. One observes that:

$$\partial \tau(\theta | s_c) / \partial \rho < 0.$$

Thus, under centralization the more homogeneous the districts, the less informed the voters will be. One way to see this is to contrast two extreme cases. If the outcomes in both districts are highly correlated ($\rho \approx 1$), then it's as if the voters are observing only one signal about θ. If the outcomes have low correlation ($\rho \approx 0$), then the voters are observing almost two independent signals. Clearly, the latter case is more informative than the former. Thus, the advantages (in terms of electoral accountability) of centralization is decreasing in the homogeneity of districts.

5 Extensions

In this section, we explore the implications of our key assumptions in the paper: non-observability of policy choices and career-concern like payoffs. We shall show immediately below, that observability of policy choices doesn't have a bite when politicians care about rent extracted, but affects our results if politicians have career concerns.

5.1 Transparency of Actions

So far we have assumed that the policy choice of the incumbent is unobserved. We interpret this as a lack of transparency of governmental actions. In this section, we explore how important this assumption is to our conclusion. We analyze the game under observable actions. The timeline of the model implies a stackelberg game where the incumbent is the first mover. We show that our main results are unaffected given rent-extraction payoffs for the incumbent. However, given reputation payoffs, the level of experimentation is higher under centralization than under decentralization. More specifically, the diverse policy is the equilibrium under both centralization while uniform policy is the equilibirium under decentralization.

Proposition 5.1. *When policy choice is observable to voters: given rent-extraction payoffs for politicians, there is greater policy experimentation under decentralization.*

Proof. Under centralization, $s_c = \frac{c_1^x + c_1^y}{2}$ is a sufficient statistic used by the voter. In particular, $\lambda(s) = 1$ if and only if $s_c \geq 0$. And the expected utility for the incumbent is;

$$\mathbb{E}\left[u(s)\right] = \int_{s_c \geq 0} 2s_c dF(s_c)$$

Now s_c is always distributed as a normal distribution with mean 0. However, the variance is greater under uniform policy than under diverse policy. Therefore for the incumbent, choosing uniform policy is a strictly dominant strategy.

Under decentralization, the voter uses own district's output as sufficient statistic if she observes diverse policy. In this case, the incumbent's expected utility is: $\sqrt{\frac{1+v_\theta}{2\pi}}$. If the voter observes uniform policy, then the voter (say of district x) uses $s_\alpha = c_1^x - \left(\frac{\rho}{v_\theta + 1}\right) \cdot c_1^y$ as the

sufficient statistic, and the incumbent's expected utility is:

$$\mathbb{E}\left[u(s)\right] = \int_{s_\alpha > 0} \mathbb{E}\left(c_1^x | s_\alpha\right) dF(s_\alpha)$$

$$= \sqrt{\frac{1 + v_\theta - \frac{\rho^2}{1 + v_\theta}}{2\pi}}$$

which is less than the expected utility under diverse policies. Therefore, the incumbent will adopt diverse policies.

$\square$

Proposition 5.2. *When policy choice is observable to voters: given reputational payoffs for politicians, there is more experimentation under centralization than under decentralization.*

Proof. First, we will examine the case of centralization. Note given career concern payoffs, the incumbent's ex-ante utility is the same as that of the voter. If the incumbent chose identical policies, then he gets payoff $\frac{v_\theta}{\sqrt{\frac{1+\rho}{2} + v_\theta}} \cdot \sqrt{\frac{1}{2\pi}}$ (see c]Electoral Accountability section). Where as if the incumbent chose different policies, then his payoff becomes $\int_{\frac{c_1^x + c_1^y}{2} > 0} \frac{v_\theta}{\frac{1}{2} + v_\theta} \left(\frac{c_1^x + c_1^y}{2}\right) dF = \frac{v_\theta}{\sqrt{\frac{1}{2} + v_\theta}} \cdot \sqrt{\frac{1}{2\pi}}$. So the incumbent chooses diverse policy under centralization in equilibrium.

Now, under decentralization, the incumbent obtains payoff $\frac{v_\theta}{\sqrt{1 + v_\theta - \frac{\rho^2}{1 + v_\theta}}} \cdot \sqrt{\frac{1}{2\pi}}$ if choosing uniform policy (see lemma 3.2), and payoff $\frac{v_\theta}{\sqrt{1 + v_\theta}} \cdot \sqrt{\frac{1}{2\pi}}$ if choosing diverse policies (see electoral accountability). Therefore, the incumbent chooses uniform policies in equilibrium.

$\square$

Below, we provide some intuition as to why the observability of actions have different implication in the rent-extraction case v.s. the reputation case.

Rent-extraction: When politicians care about the aggregate effect of competence and policy, as in the case of rent extraction, observability of their policy choice does not affect the benefit of increasing the variance of signals. It can be seen from the fact that politician's expected payoff is the same under both situations. Hence our results from previous sections does not change when we allow policy choice to be observable.

Reputation: Above observation is not true when politicians only care about $\mathbb{E}[\theta|s]$. Politicians' benefit from increasing variance now changes with the observability of actions. This

is because increasing variance has two countervailing effects: $\mathbb{E}[\theta|s] = \frac{v_\theta}{v_{\theta+\sigma_s^2}} \cdot s$ is decreasing in σ_s, while $\mathbb{E}[s|s \geq 0] = \sigma_s\sqrt{\frac{2}{\pi}}$ is increasing in σ_s.

When policy choice is not observable, *given voters' belief*, $\mathbb{E}[\theta|s]$ remains unchanged when politicians increase σ_s, only $\mathbb{E}[s|s \geq 0]$ increases. Hence there is positive gain for choosing a larger variance. However when policy choice is observable, increasing variance affects $\mathbb{E}[s|s \geq 0]$ as well. Moreover, the cost of a larger variance is bigger than the benefit. So politician's incentive reverses when the observability assumption changes.

5.2 Heterogeneous Outside Options

Previous sections assume every politician receive the same outside option once they get replaced. However, in reality it's highly plausible that politician with higher ability receive higher outside option as well. It's interesting to see how politician's incentive to choose policy will be affected in this situation.

More specifically, let politician's benefit from staying in office be a constant $w > 0$, whereas the their outside option be proportional to his ability.

$$\mathbb{E}u(s) = \mathbb{E}(\theta \mid s < 0) \cdot \Pr(s < 0) + w \cdot \Pr(s > 0)$$

By the symmetric nature of the structure, $\Pr(s > 0) = \Pr(s < 0) = 0.5$, hence $\mathbb{E}u(s)$ increases in $\mathbb{E}(\theta \mid s < 0)$. Again given

$$\begin{pmatrix} \theta \\ s \end{pmatrix} = \begin{pmatrix} \theta \\ \theta + \epsilon_p \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} v_\theta & v_\theta \\ v_\theta & v_\theta + v_p \end{pmatrix} \right)$$

we have $\mathbb{E}(\theta \mid s < 0) = -\frac{v_\theta}{\sqrt{v_\theta + v_p}}$, hence politician would like to choose a larger v_p, i.e. he prefers indention policy than diverse policies.

5.3 Constant In-office Payoff

Unlike most specifications in the literature, policymakers' payoff is not fixed after re-election in this paper. What if politician's after election payoff is a constant? We explore this possibility in this section.

Let politician's Bernoulli utility function be:

$$u(s) = \begin{cases} 0 & \lambda(s) = 0 \\ w & \lambda(s) = 1 \end{cases},$$

where $\lambda(s)$ is the electoral rule as before, and $w > 0$ is a constant representing politician's in-office benefits. Voters' strategy won't change, i.e. they still have the incentive to reelect any incumbent with perceived competence $\mathbb{E}[\theta_{in}|s]$ greater than that from a new draw of politician, i.e. $\mathbb{E}[\theta_{new}]$. All incumbents then try to maximize his expected utility given voters' reelection policy:

$$\mathbb{E}[u(s)] = \int_{s>\mathbb{E}[\theta_{new}]} w \, dF(s)$$

We analyze three situations of $\mathbb{E}[\theta_{new}]$:

- (i) $\mathbb{E}[\theta_{new}] = 0$: this is the case where the competence of new politician is a random draw from the same distribution as the incumbent's. We have $\mathbb{E}[u(s)] = \frac{1}{2}w$. Hence any policy choice by politicians can be an equilibrium choice.

- (ii) $\mathbb{E}[\theta_{new}] = -k < 0$: imagine incumbents have incumbency advantage, hence new politician's type can be seen as a random draw from a distribution with lower mean than the incumbent's. We now have $\mathbb{E}[u(s)] = w\Pr(s > -k) = w(1 - \Phi(k))$, where $\Phi(k) = \frac{1}{\sigma\sqrt{2\pi}} \int_k^\infty e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$ and $s \sim N(\mu, \sigma)$. It can be seen from incumbent's expected utility that he has incentive to decrease the variance σ of the signal s, which is exactly the opposite of what we see when his payoff is non-constant after re-election. Therefore the policy choice is the opposite of our previous results.

- (iii) $\mathbb{E}[\theta_{new}] = k > 0$: imagine that new politicians have advantage over incumbent. It's direct from above case that $\mathbb{E}[u(s)] = w\Pr(s > k) = w \cdot \Phi(k)) = w \cdot \frac{1}{\sigma\sqrt{2\pi}} \int_k^\infty e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$. Now incumbent has incentive to increase the variance of signal s, which coincides with incumbent's incentive when he has non-constant reelection payoff as in our previous analysis. Hence our results from previous sections hold in this case.

Above analysis has several implications. First, incumbency advantage can be a good thing to voters. When newly elected politician is thought to be less competent than the incumbent, the incumbent has an incentive to increase the informativeness of signals voters see. Secondly,

when politicians don't have career concern incentive, the disciplinary power of reelection is limited in terms of selecting good politicians (there is multiplicity of equilibria).

5.4   Incentives for Experimentation and Heterogeneity

Now, consider an extension where the incumbent(s) incur an additional cost if diverse policies are implemented. This can be interpreted as the incumbent sharing some administrative cost when adopting the same policy, hence captures the free-riding incentive in previous literature.

Formally, under decentralization, the Bernoulli utility function for the district incumbent is:

$$u(s) - \tau \mathbb{1} \{\rho_x \neq \rho_y\}$$

where $\mathbb{1}$ is the indicator function. In the case of centralization, the incumbent faces the cost of 2τ in the case of diverse policies.

Proposition 5.3. *Uniform policy remains the equilibrium outcome under centralization. For decentralization, there exists τ^* such that for $\tau > \tau^*$, uniform policy is the equilibrium outcome.*

Proof. Since under centralization, politician already has incentive to adopt the same policy, adding cost for diverse policies τ can only reinforce previous incentive. Hence it is straightforward to see that uniform policies remains the equilibrium outcome for centralization. This is true under either definition of $w(s)$.

For decentralization, diverse policy is no longer an equilibrium for all positive values of τ, because there is a strict incentive to deviate to gain τ if they choose the same policy (true for either specification of $w(s)$). Now, consider the case where $w(s)$ is rent extraction. Recall in proving that uniform policies cannot be sustained, we show the incumbent can deviate to adopt diverse policy and obtains $\frac{A}{\sqrt{A+C}}$ as opposed to $\sqrt{A-C}$ where $A \equiv 1 + v_\theta$ and $C \equiv \frac{\rho^2}{1+v_\theta}$. Now, if we set $\tau^* = \frac{A}{\sqrt{A+C}} - \sqrt{A-C}$, then we can prevent deviation and sustain uniform policy as an equilibrium for $\tau \geq \tau^*$. The same argument can be made for the case of reputation payoff.

□

Corollary 5.1. *τ^* is increasing in homogeneity across districts, ρ.*

Proof. τ^* is equal to the difference between the payoff for diverse policy and uniform policy (decentralization, given voter believing uniform policy). In the case of rent-extraction, the difference is an constant multiple of $\left(\frac{A}{\sqrt{A+C}} - \sqrt{A-C}\right)$. One can verify the derivative of this expression with respect to C is positive. In the case of reputation, the difference is a constant multiple of $\left(\frac{\sqrt{A+C}-\sqrt{A-C}}{A-C}\right)$, which clearly increasing in C. Finally, observing that C is an increasing function of ρ, we have proven our result.

$\square$

In other words, the incentive to experiment under decentralization is increasing in the homogeneity of the districts. Note that this finding is in contrary to previous papers. For example, it is shown in Strumpf (2002) and Callander and Harstad (2013) that the more heterogenous the districts are, the more incentives local governments have to try different policies. This is so in the former paper because when districts are heterogeneous, district A have incentive to try say policy A, then given the spillover effect, district B need only to learn policy B. The mechanism for the latter paper is that when districts are sufficiently different hence suitable policies are different, each place simply has no one to free ride. We have different prediction because officials initially will choose diverse policies under decentralization if there is no additional cost of doing so (i.e. $\tau = 0$). When $\tau > 0$, the benefit of choosing diverse policies (decrease voter's information) increases with the homogeneity measure ρ. Because the larger the ρ, more information can voters extract if politicians choose the same policy, which is contrary to politician' wish. Hence in our model, the incentive to choose different policies decreases with the heterogeneity.

6   Conclusion

The allocation of power between national and local government is a much discussed topic in both public and academic discourse on democratic governance. These discussions remain relevant today considering recent events such as the referendum of Scottish independence in the UK and Catalonia independence in Spain, as well as debate on the future of the European Union.

Our paper contributes to the debate of decentralization on two dimensions: policy experimentation and electoral accountability. We show that forces arising from an informationally complex electoral environment can have interesting effects in those two regards. More specif-

ically, when voters face uncertainty regarding policymaker competency and a lack of transparency of actions, a career-concern like incentive for policymakers arises from the prospect of reelection. This incentive induces higher experimentation under decentralization. This result provides an alternative view than current literature. Interestingly, centralization provides more information on the quality of candidates and allows the voters to better select competent policymakers. In other words, centralization is good for electoral control. This advantage however, is decreasing in the homogeneity of districts. These two effects provides a new tradeoff in terms of institutional design: decentralization provides more policy diversity, while centralization is better at selecting more competent politicians.

References

Ashworth, Scott, and Ethan Bueno de Mesquita. 2012. "Multitask, Accountability, and Institutional Design." 1.1, 4

Belleflamme, Paul, and Jean Hindriks. 2005. "Yardstick competition and political agency problems." *Social Choice and Welfare*, 24(1): 155–169. 4

Besley, Timothy, and Anne Case. 1995. "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term..." *Quarterly Journal of Economics*, 110(3). 4

Besley, Timothy, and Stephen Coate. 2003. "Centralized versus decentralized provision of local public goods: a political economy approach." *Journal of public economics*, 87(12): 2611–2637. 4

Bordignon, Massimo, Floriana Cerniglia, and Federico Revelli. 2004. "Yardstick competition in intergovernmental relationships: theory and empirical predictions." *Economics letters*, 83(3): 325–333. 3.1, 4

Cai, Hongbin, and Daniel Treisman. 2009. "Political decentralization and policy experimentation." *Quarterly Journal of Political Science*, 4(1): 35–58. 1.1

Callander, Steven, and Bård Harstad. 2013. "Experimentation in Federal Systems." National Bureau of Economic Research. 1.1, 5.4

DeGroot, Morris H. 1970. "Optimal statistical decisions." 3.1, 3.1

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999. "The economics of career concerns, part I: Comparing information structures." *The Review of Economic Studies*, 66(1): 183–198. 3.2

Holmström, Bengt. 1999. "Managerial incentive problems: A dynamic perspective." *The Review of Economic Studies*, 66(1): 169–182. 1.1, 2, 3.2

Kotsogiannis, Christos, and Robert Schwager. 2006. "On the incentives to experiment in federations." *Journal of Urban Economics*, 60(3): 484–497. 1.1

Lockwood, Ben. 2005. "Fiscal decentralization: A political economy perspective." 1.1

Oates, Wallace E. 1972. "Fiscal federalism." 1

Oates, Wallace E. 1999. "An essay on fiscal federalism." *Journal of economic literature*, 1120–1149. 1, 2

Persson, Torsten, and Guido Enrico Tabellini. 2002. *Political economics: explaining economic policy.* MIT press. 1.1, 4

Rose-Ackerman, Susan. 1980. "Risk taking and reelection: Does federalism promote innovation?" *The Journal of Legal Studies*, 593–616. 1.1

Seabright, Paul. 1996. "Accountability and decentralisation in government: An incomplete contracts model." *European economic review*, 40(1): 61–89. 4

Strumpf, Koleman S. 2002. "Does government decentralization increase policy innovation?" *Journal of Public Economic Theory*, 4(2): 207–241. 1.1, 5.4

Treisman, Daniel. 2000. "The causes of corruption: a cross-national study." *Journal of public economics*, 76(3): 399–457. 1