A Tale of Two Decentralizations: Volatility and Economic Regimes[*]

Shiyu Bo,[†] Liuchun Deng,[‡] Yufeng Sun,[§] Boqun Wang[¶]

May 23, 2018

Preliminary, please do not cite

Abstract: In this paper, we develop a formal model of inter-governmental communication to study the impact of decentralization on output volatility and economic performance in general under an authoritarian regime. Decentralization shifts the decision power of the policy making from the central government to the local government. Equipped with a framework of rational inattention, we find that two types of decentralization have distinct effects on output volatility: When promotion is mainly based on political loyalty, decentralization leads to higher output volatility; when promotion is determined by economic performance, decentralization yields lower output volatility. Various theoretical extensions highlight the main channels through which decentralization impacts volatility. A case study on two decentralization episodes in China provides empirical support for our model.

Keywords: Regime change, output volatility, economic transition, decentralization

JEL Classification Numbers: F20, P20, P40.

[*] We thank Yu-Hsiang Lei, Jiwei Qian, Zhiyong Yao and participants at Georgetown Center for Economic Research 2017 Conference, Econometric Society Asian Meeting, and Econometric Society China Meeting, Henan University Symposium on Institutional and Development Economics, CES North America Conference, and International Workshop on Economic Analysis of Institutions at Xiamen University for helpful comments. All errors are ours.

[†] Institute for Economic and Social Research, Jinan University, Guangzhou 510632, China. **E-mail:** boshiyu@outlook.com. Tel: +086 8522-5002.

[‡] Halle Institute for Economic Research (IWH) and Friedrich Schiller University Jena. **E-mail:** Liuchun.Deng@iwh-halle.de.

[§] Department of Economics, The Chinese University of HongKong. **E-mail:** pkuskyfree@gmail.com.

[¶] China Financial Policy Research Center, Renmin University of China. **E-mail:** boqunwang@jhu.edu.

1 Introduction

Decentralization, being political or economic, has been a catchword nowadays in the general discussion about reforms in the developing world. As one of most prominent features in their policy reform packages, many emerging market countries, transition economies in Asia and Eastern Europe in particular, embraced decentralization reforms since the last two decades in the twentieth century. A large literature studies the economic impact of decentralization with a focus on the level terms of economic outcomes.[1] Perhaps surprisingly, little work has been devoted to the understanding about the impact of decentralization on the variance terms,[2] most notably, the output volatility, especially given the latter being identified as a key outcome variable in the process of economic development (Koren and Tenreyro, 2007). It is mainly due to the lack of theoretical underpinning that connects decentralization with volatility. This paper attempts to fill this void.

The starting point of our analysis can be summarized by Figure 1, which plots China's economic growth together with its economic volatility over the past fifty years.[3] The output volatility, measured by the standard deviation of real gross domestic product (GDP), skyrocketed in the 1950s, when China implemented its first wave of decentralization. In sharp contrast, following the second wave of decentralization in the reform era since 1978, the output volatility declined sharply afterwards. The contrasting dynamics following two decentralization episodes seems to suggest the relationship between decentralization and output volatility is not monotonic.

Motivated by the experience from China, we build a political-economic model of decentralization to investigate under what condition decentralization tends to worsen output volatility, and economic performance in general, in an authoritarian regime under which the (economic) policy making is always subject to severe pressure from the higher authorities.

In our baseline model, there is a central government and a local government. They decide economic policy under uncertainty. There are two economic regimes. Under the centralized regime, the central government is the decision maker, while under the decentralized regime, the local government decides the policy. To reduce the economic uncertainty, the decision maker can collect information from two sources: It can collect information directly from the economy and indirectly from the other government through the inter-governmental communication. Given the limited resource, the policy maker has to allocate its resource optimally between these two sources.[4] On the other hand, the non-policy-making government can decide on how much resource it wants to allocate to direct information acquisition, thus influencing the policy making through the quality of the information that it sends to the policy maker. We assume that the central government wants to maximize the expected output, but besides the economic motive, the local government also has the incentive to follow the policy prescription from the central government, which we label as the "loyalty concern".

We demonstrate that decentralization always deteriorates economic performance, raising volatility in particular, when the local government is strongly motivated by the loyalty concern. Under

[1] For earlier empirical work on the relationship between (fiscal) decentralization and economic growth, see Davoodi and Zou (1998) among many others. For policy-oriented discussion, see Wetzel (2001).

[2] A few exceptions include earlier work by Akai et al. (2009) and more recent work by Wang and Yang (2016)

[3] We follow Wang and Yang (2016) to calculate the output volatility. We first Construct province-year output shocks by removing the trend in the GDP using Hodrick-Prescott (HP) filter and remove the influence of conventional economic factors that determine volatility such as financial development, openness, inventory management and monetary policy.

[4] Formally, we model the decision problem in the fashion of rational inattention as in Sims (2003).

the assumption that the local government enjoys information advantage,[5] we show that "good" decentralization emerges with lower output volatility, provided that the economic motive is sufficiently strong.

In our baseline model, the increase of volatility under decentralization comes from the distortion in both policy making (adherence to the policy prescription from the central government) and information acquisition (increased effort to communicating with the central government). Our first extension shuts down the second channel by studying a model without endogenous information acquisition. We find the result that decentralization raises volatility no longer always holds, thus suggesting the importance of the margin of information acquisition. Moreover, in this simplified setting, a seemingly paradoxical relationship arises: the output volatility decreases with the noisiness of the inter-governmental communication friction, as higher exogenous communication friction incentivizes the local government to focus on its own and more precise information in the policy marking. Our second and third extensions further weaken the assumptions in the baseline model. Our main results survive.

The last part of our paper revisits the motivating fact of this paper. We provide further evidence on the contrasting dynamics of volatility following two waves of decentralization, supplemented with historical discussions. In addition, we substantiate the key assumption of our model, the presence of loyalty concern, with data on promotion outcomes of provincial leaders in China.

Relation to the Literature

Our work is related to three strands of literature.

First, our work joins the large literature on the economic consequences of decentralization.[6] Our point of departure is the outcome variable of interest. In contrast to the voluminous work on the impact on the first moment of economic performance, growth rate for example, we focus on the second moment. While the focus on output volatility inevitably complicates the decision problem of the government even in its simplest possible form, we abstract from the elaboration of various forms of decentralization. The literature on decentralization can be broadly classified into fiscal (economic) decentralization (Qian and Roland, 1998; Jin et al., 2005) and political decentralization (Mookherjee, 2015).[7] Our work is to some extent at the crossroad between these two categories and therefore, we always use the general term "decentralization" in our discussion.

Second, this paper is related to the literature on the determinants of output volatility. (Ramey and Ramey, 1995) document that development economies are usually more volatile and high volatility is detrimental to economic growth. A variety of factors, including exchange rate regimes (Bleaney and Fielding, 2002), geography (Malik and Temple, 2009), trade openness (Buch et al., 2009), and country size (Di Giovanni and Levchenko, 2012) are identified as sources of cross country differences in volatility. In an influential paper, Koren and Tenreyro (2013) both theoretically and quantitatively demonstrate that the relationship between development and volatility is a by-product of technological diversification. Wang and Yang (2016) provide the first empirical evidence concerning the relationship between decentralization and volatility. Since they focus mainly on the second wave of decentralization in China, they find an unambiguously negative impact of decentralization

[5] For recent empirical evidence, see Huang et al. (2017).

[6] For a recent review, see Bardhan (2016).

[7] In a recent work, Bellofatto and Besfamille (2018) go one step further to study the optimal degree of fiscal decentralization. Regional state capacity plays a key role in the efficiency comparison between partial and full decentralizations.

on output volatility. Following their approach, we enrich their findings by examining both waves of decentralization, thus suggesting a more nuanced view of decentralization under an authoritarian regime.

Third, our theoretical framework adds to the literature on information transmission. On the one hand, our model investigates strategic (inter-governmental) communication, a common theme that has been extensively studied since the seminal work of Crawford and Sobel (1982).[8] However, since in our model the conflict of interest lies on the first moment of the economic outcome while the action set is restricted to the second moment, we obtain a sharp characterization of the equilibrium in which the usual cheap-talk outcome does not arise. On the other hand, from a technical point of view, our work borrows heavily from the literature of rational inattention (Sims, 2003).[9] To our knowledge, this paper for the first time marries the rational inattention framework with strategic communication in a political-economy setting. We feel our model might have broader applications beyond the topics of decentralization and volatility.

The rest of this paper is organized as follows. Section 2 describes the baseline model. Section 3 presents the main results of the baseline model. Section 4 discusses theoretical extensions. A simplified model is presented to highlight the key mechanisms through which decentralization impacts volatility. We also weaken the key assumption of the baseline model in an alternative setting and main results carry over. Section 5 provides a case study of China to provide some supporting evidence of the model. Section 6 concludes.

2 The Motivating Evidence

3 The Model

We model policy making under uncertainty. There are two players, a central government and a local government. They want to implement an economic policy that hinges on the true state of the economy subject to uncertainty. There are two channels through which the governments can reduce the uncertainty. Each government can directly acquire information of the true state of the economy. It can also acquire information from the other government through the inter-governmental communication which is nevertheless subject to communication friction.[10] Each government has a fixed amount of resource, which can be allocated between the two activities: direct information acquisition and indirect information acquisition by reducing the friction in inter-governmental communication.

We consider two economic regimes. Under a centralized regime, the communication is bottom-up. The local government directly acquires information and then sends a noisy signal to the central government. Facing the trade-off between two information channels, the central government decides how to allocate its attention resource and chooses the economic policy accordingly. Under a decentralized regime, the communication is top-down. The central government directly acquires information and sends a noisy signal to the local government. The local government allocates its

[8] For work that we argue share similar spirit, see Board et al. (2007).

[9] For a recent pedagogical overview of the subject, see Gabaix (2017).

[10] In the baseline setting, the communication friction is endogenously determined. We will present a version of the model with exogenously given communication friction in the discussion section, demonstrating the robustness of the model and highlighting the relative importance of alternative mechanisms.

attention resource and then implements its desired economic policy. Figure 2 illustrates the timeline of the model under two regimes.

We now proceed to formally specify the information structure, economic regimes, and the decision problem of the governments under each regime.

3.1 Information Structure

Denote the true state of the economy by θ. Both the local and central governments hold the same prior of θ, which follows a normal distribution with mean zero and variance σ^2, denoted by $\mathcal{N}(0, \sigma^2)$. Due to information imperfections, governments cannot observe θ perfectly. Instead, they observe θ with a white noise.

$$\begin{aligned} \theta_c &= \theta + z_c, \quad z_c \sim \mathcal{N}(0, \sigma_c^2), \\ \theta_\ell &= \theta + z_\ell, \quad z_\ell \sim \mathcal{N}(0, \sigma_\ell^2), \end{aligned}$$

where θ_c and θ_ℓ are the noisy signals for the central and local governments. [11] The governments can choose to reduce σ_c^2 and σ_ℓ^2 by directly acquire information of the state of the economy, so both σ_c^2 and σ_ℓ^2 will be endogenously determined.

Alternatively, a government can acquire information from the other government through inter-governmental communication subject to friction. This will be specified under two different economic regimes.

3.2 Two Economic Regimes

Denote the exogenous communication friction by ϵ. We assume $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. The signal receiver under each regime can reduce this friction by allocating attention resource to the communication channel.

Under the centralized regime, the local government sends a signal s_ℓ. The central government receives a signal s'_ℓ with $s'_\ell = s_\ell + \epsilon$. Upon receiving the signal, the central government can acquire additional information from the local government s''_ℓ with $s''_\ell = s_\ell + \epsilon_c$ and $\epsilon_c \sim \mathcal{N}(0, \sigma_{\epsilon c}^2)$, where $\sigma_{\epsilon c}^2$ will be endogenously determined as an outcome of the trade-off between two information channels which will be formally specified later. Based on its private information θ_c and two signals received,[12] the central government makes the policy choice a_c.

Similarly, under the decentralized regime, the central government sends a signal s_c. The local government receives a signal s'_c with $s'_c = s_c + \epsilon$. The local government also decides how much resource to be spent on the second signal $s''_c = s_c + \epsilon_\ell$ with $\epsilon_\ell \sim \mathcal{N}(0, \sigma_{\epsilon\ell}^2)$. Given the resulting information set, the local government then picks its preferred policy a_ℓ based on θ_ℓ, s'_c, and s''_c.

From now on, we assume that $s_\ell = \theta_\ell$ and $s_c = \theta_c$. In the discussion section, we will allow the signal sender to strategically introduce noise into the inter-governmental communication. It turns out in our setting, the signal sender always has incentive to truthfully reveal its information.

[11] Throughout the paper, we will use subscript "c" for variables associated with the central government and subscript "ℓ" for variables associated with the local government.

[12] The specification with two signals facilitates the introduction of entropy reduction constraint in what follows. One can think of the first signal as the signal receiver's prior of the signal sent by the other government. Alternatively, we can introduce only one inter-governmental signal with its variance bounded above by σ_ϵ^2, but the decision problem for the government is effectively the same, that is, to decide the preciseness of inter-governmental communication.

We assume all the white noises z_c, z_ℓ, ϵ, ϵ_ℓ, and ϵ_c are independent.

3.3 Decision Problem for the Government

The decision problem for the government that decides the economic policy, that is, the signal receiver, has two layers. It has to first decide the resource allocation over two channels of information acquisition and then choose the optimal policy based the information gathered.

We first formalize the resource allocation problem.

Following the framework of rational inattention,[13] we introduce a constraint on the information flow. To formalize the notion of information, we define the differential entropy as in the standard information theory as a measure of the uncertainty of a continuous random variable.[14]

Definition 1. The differential entropy $H(X)$ of a continuous random variable X with a probability density function $f(x)$ is defined as

$$H(X) = E[-\log_2 f(x)] = -\int f(x) \log_2 f(x) dx.$$

If X follows a multivariate normal distribution with a covariance matrix Σ, it can be shown that the entropy of X is given by

$$H(X) = \frac{n}{2} \log_2(2\pi e) + \frac{1}{2} \log_2 |\Sigma|,$$

where n is the dimensionality of the random variable and $|\Sigma|$ is the determinant of Σ.

Definition 2. The conditional differential entropy $H(X|Y)$ of two continuous random variables X and Y with a joint probability density function $f(x, y)$ is defined as

$$H(X|Y) = -\int f(x, y) \log_2 f(x|y) dx dy.$$

In general, we have[15]

$$H(X|Y) = H(X, Y) - H(Y).$$

Hence, if one is interested in X, the informativeness of an observation Y can be captured by the difference between $H(X) - H(X|Y)$. In other words, the difference between $H(X)$ and $H(X|Y)$ is the reduction of uncertainty with respect to X when Y is observed. We assume that each economic agent has limited attention resource, i.e., $H(X) - H(X|Y) < \kappa$. We now specialize this resource constraint to our specific setting.

Under the centralized regime, the information flow constraint for the signal sender, the local government, is given by

$$H(\theta) - H(\theta|\theta_\ell) \leq \kappa_\ell \;\; \Leftrightarrow \;\; \frac{1}{2} \log_2 \left(\frac{\text{Var}(\theta)}{\text{Var}(\theta|\theta_\ell)} \right) = \frac{1}{2} \log_2 \left(\frac{\sigma^2 + \sigma_\ell^2}{\sigma_\ell^2} \right) \leq \kappa_\ell, \tag{1}$$

where $\kappa_\ell > 0$ is the capacity of information acquisition of the local government.

[13] See, for example, Sims (2003) and Mackowiak and Wiederholt (2009), among many others.
[14] See, for example, chapter 8 in Cover and Thomas (2012) for a standard treatment.
[15] See Equation 8.33 in Cover and Thomas (2012).

For the central government, the reduction of entropy comes from two sources: improved information about both θ and ϵ. The constraint on the entropy reduction is then given by

$$H(\theta,\epsilon|s'_\ell) - H(\theta,\epsilon|\theta_c, s'_\ell, s''_\ell) \le \kappa_c,$$

where $\kappa_c > 0$ is the capacity of information acquisition of the central government.

Notice that even though θ and ϵ are independent, we cannot write the constraint in an additively separable form for θ and ϵ as they might not be independent conditional on the acquired information $(\theta_c, s'_\ell, s''_\ell)$.[16] The following lemma provides a closed form solution to this information flow constraint.[17]

Lemma 1. The information flow constraint of the central government under the centralized regime is given by

$$\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_{\epsilon c}^2}\right)\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2}\right) \le \frac{2^{2\kappa_c}(\sigma^2 + \sigma_\ell^2 + \sigma_\epsilon^2)}{\sigma^2\sigma_\ell^2\sigma_\epsilon^2} + \frac{1}{\sigma_\ell^4} \equiv K_c(\sigma_\ell^2). \tag{2}$$

However complicated Equation 2 appears, the choice variables, σ_c^2 and $\sigma_{\epsilon c}^2$ for the central government are multiplicatively separable in the information flow constraint, which is very important for a sharp characterization of the attention allocation problem. Moreover, we have $K_\ell \ge \left(1/\sigma_\epsilon^2 + 1/\sigma_\ell^2\right)\left(1/\sigma^2 + 1/\sigma_\ell^2\right)$ with the equality if and only if $\kappa_c = 0$. That said, when the central government has zero information capacity, then it is impossible to acquire any information $(\sigma_c^2 = \sigma_{\epsilon c}^2 = \infty)$.

In what follows, we sometimes simply write $K_c(\sigma_\ell^2)$ as K_c, but it should be noticed that K_c is a function of σ_ℓ^2, which is particularly important when we study the decision problem of the local government whose choice variable is σ_ℓ^2.

Symmetrically, under the decentralized regime, the information flow constraint for the signal sender, the central government, is given by

$$H(\theta) - H(\theta|\theta_c) \le \kappa_c \quad \Leftrightarrow \quad \frac{1}{2}\log_2\left(\frac{\mathrm{Var}(\theta)}{\mathrm{Var}(\theta|\theta_c)}\right) = \frac{1}{2}\log_2\left(\frac{\sigma^2 + \sigma_c^2}{\sigma_c^2}\right) \le \kappa_c. \tag{3}$$

For the local government, the constraint on the entropy reduction is given by

$$H(\theta,\epsilon|s'_c) - H(\theta,\epsilon|\theta_\ell, s'_c, s''_c) \le \kappa_\ell.$$

Following the proof of Lemma 1, we can rewrite the information flow constraint for the local government in a multiplicatively separable form for the two choice variables σ_ℓ^2 and $\sigma_{\epsilon\ell}^2$.

Lemma 2. The information flow constraint of the local government under the decentralized regime is given by

$$\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_{\epsilon \ell}^2}\right)\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2}\right) \le \frac{2^{2\kappa_\ell}(\sigma^2 + \sigma_c^2 + \sigma_\epsilon^2)}{\sigma^2\sigma_c^2\sigma_\epsilon^2} + \frac{1}{\sigma_c^4} \equiv K_\ell(\sigma_c^2). \tag{4}$$

[16] This stands in sharp contrast with the setting in Mackoviak and Wiederholt (2009).

[17] All the proofs are relegated to the appendix.

In what follows, we sometimes simply write $K_\ell(\sigma_c^2)$ as K_ℓ, but again it should noticed that K_ℓ depends on the choice variable of the central government under this decentralized regime.[18]

We now impose the key assumption of this baseline setting.

Assumption 1. $\kappa_\ell > \kappa_c$.

In words, we assume that the local government has higher information capacity than the central government. This assumption is useful to deliver the main results of the paper in the sharpest form. We will discuss various ways to relax this assumption.

3.4 Output Level and Volatility

We define the output level Y in a quadratic form

$$Y \equiv Y^* - (a_i - \theta)^2, \quad i = c, \ell,$$

where Y^* is the ideal output level if the policy choice a_c or a_ℓ perfectly matches the true state of the economy θ. In this paper, we are particularly interested in the ex ante expected output level $E(Y)$ and its variance $Var(Y)$.

3.5 Government Objectives

The second layer of the government's decision problem is to pick the desired economic policy. Under the centralized regime, the central government attempts to maximize solely the expected economic output, so its decision problem, consisting of two layers, is given by

$$\max_{\sigma_c^2, \sigma_{\epsilon c}^2} E\left[\max_{a_c} \; E(Y|\theta_c, s'_\ell, s''_\ell)\right] = \max_{\sigma_c^2, \sigma_{\epsilon c}^2} E\left[\max_{a_c} \; Y^* - E((a_c - \theta)^2|\theta_c, s'_\ell, s''_\ell)\right],$$

subject to Constraint 2.

The local government, who sends the signal to the central government under this regime, cares about both the economic output and whether its policy suggestion is actually implemented by the government. More precisely, its decision problem is given by

$$\max_{\sigma_\ell^2} (1-\gamma)\left(Y^* - E(a_c - \theta)^2\right) - \gamma E(\theta_\ell - a_c)^2,$$

subject to Constraint 1, where $0 \leq \gamma \leq 1$. The first term in the payoff function is the utility that the local government directly derives from the economic output. The second term captures the fact that the local government also cares about how closely the central government follows its policy suggestion. This is a reduced-form way to incorporate additional promotion incentive beyond the merit-based rules. Therefore, the local government faces a trade-off between economic welfare and career concern. The parameter γ measures the relative importance of career concern. When $\gamma = 0$,

[18] According to Propositions 1 and 2, given the same attention budget, the central government can get a more precise signal θ_c under the centralized regime (setting $\sigma_{\epsilon c}^2 = \infty$) than the decentralized regime, while the local government can get a more precise signal θ_ℓ under the decentralized regime (setting $\sigma_{\epsilon\ell}^2 = \infty$) than the centralized regime. In some sense, being a signal receiver softens the information flow constraint. Our main results are not driven by this *de facto* difference in capacity across regimes. This can be seen when we discuss a variant of the model with σ_ℓ^2 and σ_c^2 being exogenous.

the objective of the local government is perfectly aligned with that of the central government. When $\gamma = 1$, the local government attaches no importance to economic output and only attempts to induce the central government to adopt its policy recommendation.

Under the decentralized regime, the central government, who now becomes the sender of the signal, has the same payoff function but with different choice variables. Its decision problem is given by

$$\max_{\sigma_c^2} \; E(Y) = Y^* - E(a_\ell - \theta)^2,$$

subject to Constraint 3. In words, the central government chooses the signal and its precision in order to induce the local government to maximize the expected economic output.

The local government now has a two-layer decision problem, which is given by

$$\max_{\sigma_\ell^2, \sigma_{\epsilon\ell}^2} E\left\{ \max_{a_\ell} \, (1-\gamma)\left(Y^* - E[(a_\ell - \theta)^2 | \theta_\ell, s_c', s_c'']\right) - \gamma E[(a_\ell - \theta_c)^2 | \theta_\ell, s_c', s_c''] \right\},$$

subject to Constraint 4. Despite having a similar form, the second term of the payoff function entails a different interpretation: We implicitly assume that the local government has a loyalty concern here, which is quite common in an authoritarian regime like China.

3.6 Equilibrium Definition

We close this section by formally defining the (perfect Bayesian Nash) equilibrium of the sequential game under the two regimes.

The equilibrium under the centralized regime is defined as a quadruplet $(\sigma_\ell^{*2}, \sigma_c^{*2}(\cdot), {\sigma_{\epsilon c}^*}^2(\cdot), a_c^*(\cdot))$ such that for any sextuplet $(\sigma_\ell^2, \sigma_c^2, \sigma_{\epsilon c}^2; \theta_c, s_\ell', s_\ell'') \in \mathbb{R}_+^3 \times \mathbb{R}^3$,

$$a_c^*(\sigma_\ell^2, \sigma_c^2, \sigma_{\epsilon c}^2; \theta_c, s_\ell', s_\ell'') \in \arg\max_{a_c} E(Y^* - (a_c - \theta)^2 | \theta_c, s_\ell', s_\ell'');$$

for any $\sigma_\ell^2 \in \mathbb{R}_+$,

$$(\sigma_c^{*2}(\sigma_\ell^2), {\sigma_{\epsilon c}^*}^2(\sigma_\ell^2)) \in \arg\max_{\sigma_c^2, \sigma_{\epsilon c}^2} E(Y^* - (a_c^*(\sigma_\ell^2, \sigma_c^2, \sigma_{\epsilon c}^2; \theta_c, s_\ell', s_\ell'') - \theta)^2)$$

subject to Constraint 2; and

$$\begin{aligned} \sigma_\ell^{*2} \in \arg\max_{\sigma_\ell^2} \quad & \Big\{(1-\gamma)(Y^* - E(a_c^*(\sigma_\ell^2, \sigma_c^{*2}(\sigma_\ell^2), {\sigma_{\epsilon c}^*}^2(\sigma_\ell^2); \theta_c, s_\ell', s_\ell'') - \theta)^2) \\ & -\gamma E(\theta_\ell - a_c^*(\sigma_\ell^2, \sigma_c^{*2}(\sigma_\ell^2), {\sigma_{\epsilon c}^*}^2(\sigma_\ell^2); \theta_c, s_\ell', s_\ell''))^2 \Big\} \end{aligned}$$

subject to Constraint 1.

The equilibrium under the decentralized regime is defined as a quadruplet $(\sigma_c^{*2}, \sigma_\ell^{*2}(\cdot), {\sigma_{\ell c}^*}^2(\cdot), a_\ell^*(\cdot))$ such that for any sextuplet $(\sigma_c^2, \sigma_\ell^2, \sigma_{\epsilon\ell}^2; \theta_\ell, s_c', s_c'') \in \mathbb{R}_+^3 \times \mathbb{R}^3$,

$$a_\ell^*(\sigma_c^2, \sigma_\ell^2, \sigma_{\epsilon\ell}^2; \theta_\ell, s_c', s_c'') \in \arg\max_{a_\ell} (1-\gamma) E(Y^* - (a_\ell - \theta)^2 | \theta_\ell, s_c', s_c'') - \gamma E[(a_\ell - \theta_c)^2 | \theta_\ell, s_c', s_c''];$$

for any $\sigma_c^2 \in \mathbb{R}_+$,

$$(\sigma_\ell^{*2}(\sigma_c^2), {\sigma_{\epsilon\ell}^*}^2(\sigma_c^2)) \in \arg\max_{\sigma_\ell^2, \sigma_{\epsilon\ell}^2} \quad \{(1-\gamma)E(Y^* - (a_\ell^*(\sigma_c^2, \sigma_\ell^2, \sigma_{\epsilon\ell}^2; \theta_\ell, s_c', s_c'') - \theta)^2) \\ -\gamma E(a_\ell^*(\sigma_c^2, \sigma_\ell^2, \sigma_{\epsilon\ell}^2; \theta_\ell, s_c', s_c'') - \theta_c)^2)$$

subject to Constraint 4; and

$$\sigma_c^{*2} \in \arg\max_{\sigma_c^2}(Y^* - E(a_\ell^*(\sigma_c^2, \sigma_\ell^{*2}(\sigma_c^2), {\sigma_{\epsilon\ell}^*}^2(\sigma_c^2); \theta_\ell, s_c', s_c'') - \theta)^2)$$

subject to Constraint 3.

In both regimes, we require the belief updating follows the Bayes' rule.[19]

4 Basic Results

Under each regime, the government that receives the signal solves its decision problem backwards. The resource allocation of the two channels of information acquisition hinges on the determination of optimal economic policy. For each regime, we first fully characterize the optimal economic policy for any given resulting information set. We then solve backwards the optimal allocation of the attention resource. The last step is to characterize the optimal decision of the signal sender. After we solve the equilibrium under each regime, we turn to the comparison of economic output and volatility between two regimes.

4.1 Optimal Economic Policy

Lemma 3. Under the centralized regime, the optimal policy of the central government is chosen as a linear combination of signals

$$a_c = E(\theta|\theta_c, s_\ell', s_\ell'') = \frac{\frac{\theta_c}{\sigma_c^2} + \frac{1}{\sigma_\ell^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon c}^2)^{-1}}\left(\frac{s_\ell'/\sigma_\epsilon^2}{1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon c}^2} + \frac{s_\ell''/\sigma_{\epsilon c}^2}{1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon c}^2}\right)}{\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon c}^2)^{-1}}}.$$

Remark 1. Given the expression of $f(\theta, \theta_c, s_\ell', s_\ell'')$ in the proof, it is easy to see

$$E(a_c - \theta)^2 = Var(\theta|\theta_c, s_\ell', s_\ell'') = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon c}^2)^{-1}}\right)^{-1}.$$

Lemma 4. Under the decentralized regime, the optimal policy of the local government is chosen as a linear combination of signals

$$a_\ell = (1-\gamma)E(\theta|\theta_\ell, s_c', s_c'') + \gamma E(\theta_c|\theta_\ell, s_c', s_c'') = k_1\theta_\ell + k_2 s_c' + k_3 s_c'',$$

with $s_c' = \theta_c + \epsilon$, $s_c'' = \theta_c + \epsilon_\ell$, and

$$k_1 \equiv \frac{\frac{1-\gamma}{\sigma_\ell^2}}{\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}} + \frac{\frac{\gamma}{\sigma_c^2 + \sigma_\ell^2 + \sigma_c^2\sigma_\ell^2/\sigma^2}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}},$$

[19]For simplicity, we omit the prior of θ whenever we state the information set.

$$k_2 \equiv \frac{\frac{1-\gamma}{\sigma_c^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2)^{-1}}\frac{1/\sigma_\epsilon^2}{1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2}}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2)^{-1}}} + \frac{\frac{\gamma}{\sigma_\epsilon^2}}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}},$$

$$k_3 \equiv \frac{\frac{1-\gamma}{\sigma_c^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2)^{-1}}\frac{1/\sigma_{\epsilon\ell}^2}{1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2}}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2)^{-1}}} + \frac{\frac{\gamma}{\sigma_{\epsilon\ell}^2}}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}}.$$

Remark 2. Similar to the centralized regime, we have

$$E(a_\ell-\theta)^2\Big|_{\gamma=0} = Var(\theta|\theta_\ell, s_c', s_c'') = \left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2)^{-1}}\right)^{-1}$$

$$E(a_\ell-\theta_c)^2\Big|_{\gamma=1} = Var(\theta_c|\theta_\ell, s_c', s_c'') = \left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}\right)^{-1}$$

Since the optimal action is always a linear combination of the signal obtained by the signal receiver, we obtain a tight relationship between the expected output and output volatility.

Lemma 5. Let $a_i = m_{i1}\theta_\ell + m_{i2}\theta_c + m_{i3}\epsilon + m_{i4}\epsilon_c + m_{i5}\epsilon_\ell$ $(i = c, \ell)$. The expected output is given by

$$E(Y) \equiv Y^* - E(a_i-\theta)^2 = Y^* - \left(m_{i1}^2\sigma_\ell^2 + m_{i2}^2\sigma_c^2 + m_{i3}^2\sigma_\epsilon^2 + m_{i4}^2\sigma_{\epsilon c}^2 + m_{i5}^2\sigma_{\epsilon\ell}^2 + (1-m_{i1}-m_{i2})^2\sigma^2\right). \tag{5}$$

Moreover, the output volatility strictly decreases with the expected output, which is given by

$$Var(Y) = 2\left(m_{i1}^2\sigma_\ell^2 + m_{i2}^2\sigma_c^2 + m_{i3}^2\sigma_\epsilon^2 + m_{i4}^2\sigma_{\epsilon c}^2 + m_{i5}^2\sigma_{\epsilon\ell}^2 + (1-m_{i1}-m_{i2})^2\sigma^2\right)^2 = 2(Y^*-E(Y))^2. \tag{6}$$

The lemma provides the closed-form solution of the expected output level and its variance. More importantly, it shows that the output level and volatility move in the opposite direction, and as a result, the comparative statics concerning the output level can easily re-interpreted in terms of volatility. From now on, we will mainly work with $E(Y)$, or more directly, $E(a_i-\theta)^2$, given its simpler expression, and interpret our results in both level and volatility terms.

4.2 The Equilibrium under the Centralized Regime

According to Lemma 3 and its remark, the resource allocation problem for the central government can be simplified to

$$\max_{\sigma_c^2,\sigma_{\epsilon c}^2} \frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2)^{-1}},$$

subject to Constraint 2. By inspection, it is observed that the constraint has to be binding. Given the binding constraint, we can further simplify the constrained optimization problem to

$$\max_{\sigma_c^2,\sigma_{\epsilon c}^2} \left(1-\frac{1}{\sigma_\ell^4 K_c}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2}\right)$$

subject to Constraint 2. Since $K_c > 1/\sigma_\ell^4$, the maximum is attained when σ_c^2 attains its minimum under the constraint, or equivalently, $\sigma_{\epsilon c}^2 = \infty$. Hence, under the centralized regime, the central government always allocates all of its attention resource to acquiring economic information directly.

Lemma 6. Under the centralized regime, for any given σ_ℓ^2, the central government always devotes itself to the direct information acquisition with $\sigma_{\epsilon c}^2 = \infty$.

Lemma 7. Under the centralized regime, for any γ, the local government always spends all of its attention resource on information acquisition (Constraint 1 is binding), which leads to

$$\sigma_\ell^2 = \sigma^2/(2^{2\kappa_\ell} - 1).$$

The intuition behind this sharp characterization is twofold. On the one hand, the economic motive (corresponding to $(1-\gamma)EY$) incentivizes the local government to increase the precision of its signal. On the other hand, since the central government attaches more importance to the signal sent by the local government if the quality of the signal is higher, the political motive (corresponding to $-\gamma E(a_c - \theta_\ell)^2$) gives additional incentive for the local government to maximize its information acquisition.

Collecting the results from Lemmas 3, 6, and 7, we obtain the following equilibrium characterization for the centralized regime.

Proposition 1. Under the centralized regime, we have

$$E(a_c - \theta)^2 = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2 + \sigma_\epsilon^2}\right)^{-1}$$

with $\sigma_\ell^2 = \sigma^2/(2^{2\kappa_\ell} - 1)$, $\sigma_{\epsilon c}^2 = \infty$, and

$$\sigma_c^2 = \left[K_c\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_\ell^2}\right)^{-1} - \frac{1}{\sigma^2} - \frac{1}{\sigma_\ell^2}\right]^{-1} = \frac{\sigma^2(\sigma_\epsilon^2 + \sigma_\ell^2)}{(2^{2\kappa_c} - 1)(\sigma^2 + \sigma_\epsilon^2 + \sigma_\ell^2)}.$$

4.3 The Equilibrium under the Decentralized Regime

We first consider the resource allocation problem for the local government for two extreme cases: (i) $\gamma = 0$; (ii) $\gamma = 1$.

4.3.1 No Loyalty Concern ($\gamma = 0$)

In the absence of loyalty concern ($\gamma = 0$), we know from the remark of Lemma 4,

$$E(a_\ell - \theta)^2 = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}\right)^{-1} = Var(\theta|\theta_\ell, s_c', s_c'').$$

The resource allocation of the local government can now be written as

$$\max_{\sigma_\ell^2, \sigma_{\epsilon\ell}^2} \frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}$$

subject to Constraint 4. By inspection, we notice the constraint must be binding. Given the binding constraint, we can further simplify the constrained optimization problem to

$$\max_{\sigma_\ell^2, \sigma_{\epsilon\ell}^2} \left(1 - \frac{1}{\sigma_c^4 K_\ell}\right)\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2}\right)$$

subject to Constraint 4. Since $K_\ell > 1/\sigma_c^4$, the maximum is attained when σ_ℓ^2 attains its minimum under the constraint, or equivalently, $\sigma_{\epsilon\ell}^2 = \infty$. Hence, under the decentralized regime, the local government always allocates all of its attention resource to acquiring economic information directly.

Lemma 8. Under the decentralized regime with $\gamma = 0$, for any given σ_c^2, the local government always devotes itself to the direct information acquisition with $\sigma_{\epsilon\ell}^2 = \infty$.

4.3.2 Pure Loyalty Concern ($\gamma = 1$)

If the local government has only the loyalty concern ($\gamma = 1$), we know from the remark of Lemma 4 that

$$E(a_\ell - \theta_c)^2 = \left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}\right)^{-1} = Var(\theta_c|\theta_\ell, s_c', s_c'').$$

The resource allocation of the local government can now be written as

$$\max_{\sigma_\ell^2, \sigma_{\epsilon\ell}^2} \frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}$$

subject to Constraint 4. Similarly, the constraint must be binding. Given the binding constraint, we can further simplify the constrained optimization problem to

$$\max_{\sigma_\ell^2, \sigma_{\epsilon\ell}^2} \left(1 - \frac{1}{\sigma_c^4 K_\ell}\right)\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_{\epsilon\ell}^2}\right)$$

subject to Constraint 4. Since $K_\ell > 1/\sigma_c^4$, the maximum is attained when $\sigma_{\epsilon\ell}^2$ attains its minimum under the constraint, or equivalently, $\sigma_\ell^2 = \infty$.

Lemma 9. Under the decentralized regime with $\gamma = 1$, for any given σ_c^2, the local government always devotes to the inter-governmental communication with $\sigma_\ell^2 = \infty$.

We now turn to the case with a general $\gamma \in [0, 1]$ under the decentralized regime.

4.3.3 The General Case ($\gamma \in [0, 1]$)

Lemma 10. Under the decentralized regime, there exist two cutoffs $\underline{\gamma}$ and $\bar{\gamma}$ such that $0 < \underline{\gamma} < \bar{\gamma} < 1$. If $\gamma \leq \underline{\gamma}$, the local government specializes in direct information acquisition ($\sigma_{\epsilon\ell}^2 = \infty$). If $\gamma \geq \bar{\gamma}$, the local government specializes in intergovernmental communication ($\sigma_\ell^2 = \infty$). If $\gamma \in (\underline{\gamma}, \bar{\gamma})$, the local government allocates its budget to both activities ($\sigma_\ell^2 < \infty$ and $\sigma_{\epsilon\ell}^2 < \infty$) with

$$\frac{1}{\sigma_\ell^2} + \frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} = \frac{K_\ell^{1/2}(1-\gamma)}{\gamma}, \tag{7}$$

which implies that $\partial\sigma_\ell^2/\partial\gamma > 0$. Moreover, $\underline{\gamma}$ and $\bar{\gamma}$ are the unique roots to the following two equations, respectively.

$$\begin{aligned} \underline{\gamma}^2 - (1-\underline{\gamma})^2 K_\ell^{-1}(1/\sigma_\epsilon^2 + 1/\sigma_c^2)^2 &= 0 \\ \bar{\gamma}^2 - (1-\bar{\gamma})^2 K_\ell (1/\sigma^2 + 1/\sigma_c^2)^{-2} &= 0 \end{aligned}$$

The optimal resource allocation of the local government is very much in line with the intuition. The local government focuses exclusively on direct information acquisition provided that the economic motive is sufficiently strong ($\gamma \leq \underline{\gamma}$, while it focuses exclusively on intergovernmental communication if its loyalty concern is sufficiently strong ($\gamma \geq \bar{\gamma}$. If its loyalty concern is in the intermediate range, then the attention resource will be allocated to both dimensions with the effort on intergovernmental communication strictly increasing with the intensity of the loyalty concern.

We illustrate the relationship between σ_ℓ^2 $(1/\sigma_\ell^2)$) with γ in Figure 3.

We now turn to the optimal behavior of the central government under the decentralized regime.

Lemma 11. Under the decentralized regime, for any γ, the central government always spends all of its attention resource on information acquisition (Constraint 3 is binding), which leads to

$$\sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1).$$

Given the sharp characterization of the optimal action of the central government, we have the following characterization results for the two extreme cases under the decentralized regime.

Proposition 2. Under the decentralized regime with $\gamma = 0$, we have

$$E(a_\ell - \theta)^2\Big|_{\gamma=0} = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + \sigma_\epsilon^2}\right)^{-1}$$

with $\sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1)$, $\sigma_{\epsilon\ell}^2 = \infty$, and

$$\sigma_\ell^2 = \left[K_\ell\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2}\right)^{-1} - \frac{1}{\sigma^2} - \frac{1}{\sigma_c^2}\right]^{-1} = \frac{\sigma^2(\sigma_\epsilon^2 + \sigma_c^2)}{(2^{2\kappa_\ell} - 1)(\sigma^2 + \sigma_\epsilon^2 + \sigma_c^2)}.$$

Proposition 3. Under the decentralized regime with $\gamma = 1$, we have

$$E(a_\ell - \theta)^2\Big|_{\gamma=1} = \frac{\sigma_c^2(1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^2 + 1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2 + \sigma^2/(\sigma_c^2 + \sigma^2)^2}{\left[1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2 + 1/(\sigma_c^2 + \sigma^2)\right]^2}$$

with $\sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1)$, $\sigma_\ell^2 = \infty$, and

$$\sigma_{\epsilon\ell}^2 = \left[K_\ell\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2}\right)^{-1} - \frac{1}{\sigma_\epsilon^2} - \frac{1}{\sigma_c^2}\right]^{-1} = \frac{\sigma_\epsilon^2(\sigma^2 + \sigma_c^2)}{(2^{2\kappa_\ell} - 1)(\sigma^2 + \sigma_\epsilon^2 + \sigma_c^2)}.$$

Moreover, we establish the following monotonicity result that is crucial for the comparison between two regimes.

Proposition 4. Under the decentralized regime, $E(a_\ell - \theta)^2$ strictly increases with γ.

4.4 Comparison between Two Regimes

The following result suggests that, compared with the centralized regime, the decentralized regime with $\gamma = 0$ leads to higher expected output and lower volatility, provided that the local government has higher capacity of information acquisition.[20]

[20] As made clear in the proof, with $\gamma = 1$, decentralization improves economic performance if and only if $\kappa_\ell > \kappa_c$.

Lemma 12. $E(a_c - \theta)^2 > E(a_\ell - \theta)^2\Big|_{\gamma=0}$.

On the other hand, decentralization with $\gamma = 1$ always deteriorate economic performance.

Lemma 13. $E(a_c - \theta)^2 < E(a_\ell - \theta)^2\Big|_{\gamma=1}$.

The following result directly follows from Proposition 4 and Lemmas 5, 12, and 13.

Theorem 1. There exists a unique $\tilde{\gamma}$ in $(0, 1)$ such that $E(a_c - \theta)^2 = E(a_\ell - \theta)\Big|_{\gamma=\tilde{\gamma}}$. If $\gamma > \tilde{\gamma}$, decentralization leads to higher economic volatility (and lower economic output); if $\gamma < \tilde{\gamma}$, decentralization leads to lower economic volatility (and higher economic output). Moreover, economic volatility strictly increases with γ under the decentralized regime.

Figure 4 illustrates the comparison between two economic regimes in relation to the degree of loyalty concern γ. Notice there are two kinks on the curve of $E(a_\ell - \theta)^2$ at which $\gamma = \underline{\gamma}$ or $\bar{\gamma}$. The curve is much steeper in the middle range because the rise of γ induces the local government to increase σ_ℓ^2 when $\gamma \in (\underline{\gamma}, \bar{\gamma})$.

A natural question arises: what are the relationships between $\tilde{\gamma}$ and $\bar{\gamma}$ and between $\tilde{\gamma}$ and $\underline{\gamma}$? The following result suggests that decentralization always leads to worse economic performance if the local government focuses exclusively on the inter-governmental communication.

Corollary 1. $\bar{\gamma} > \tilde{\gamma}$.

On the other hand, there is no clear relationship between $\tilde{\gamma}$ and $\underline{\gamma}$. We conduct two numerical simulation. In the first simulation, we let $\sigma^2 = \sigma_\epsilon^2 = 100$, $\kappa_\ell = 2\kappa_c = 2$. We find that $\underline{\gamma} \approx 0.27$ and $\tilde{\gamma} \approx 0.47$. In the second simulation, we reduce σ_ϵ^2 to be 10. We then obtain $\underline{\gamma} \approx 0.33$ and $\tilde{\gamma} \approx 0.26$.

5 Extension and Discussion

5.1 Exogenous Communication Frictions

We now turn to a model without rational inattention. The basic setup is similar. The main departure is that σ_ℓ^2 and σ_c^2 are exogenously given. In particular, we assume that each government receives a private signal about θ from the nature. The private signal of the central government θ_c follows $\mathcal{N}(\theta, \sigma_c^2)$, while the private signal of the local government θ_ℓ follows $\sim \mathcal{N}(\theta, \sigma_\ell^2)$ with σ_ℓ^2 and σ_c^2 exogenously given. We assume $\sigma_c^2 > \sigma_\ell^2$, that is, the local government receives a more precise signal than the central government. The quality of the signal (σ_ℓ^2 or σ_c^2) is the same under both regimes.

Under the centralized regime, the local government sends its signal $s_\ell = \theta_\ell$. The central government receives a signal s'_ℓ with $s'_\ell = s_\ell + \epsilon$. Upon receiving the signal, the central government makes the policy choice a_c based on the private information θ_c and the signal received s'_ℓ. Similarly, under the decentralized regime, the central government sends a signal $s_c = \theta_c$. The local government receives a signal s'_c with $s'_c = s_c + \epsilon$. The local government then picks its preferred policy a_ℓ based on θ_ℓ and s'_c. We assume $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ with σ_ϵ^2 being exogenous. The model otherwise follows the baseline setting. Figure 5 illustrates the timeline of the model without rational inattention.

Abstracting from the strategic behavior in this simplified framework, the decision problem of the signal receiver is only to determine the desired economic policy.

Under the centralized regime, the decision problem of the central government is given by

$$\max_{a_c} \; E(Y|\theta_c, s'_\ell) = Y^* - E[(a_c-\theta)^2|\theta_c, s'_\ell],$$

Under the decentralized regime, the decision problem of the local government is given by

$$\max_{a_\ell} \; (1-\gamma)\left(Y^* - E[(a_\ell-\theta)^2|\theta_\ell, s'_c]\right) - \gamma E[(a_\ell - s_c)^2|\theta_\ell, s'_c].$$

Lemma 14. Under the centralized regime, the optimal policy for the central government is given by

$$a_c = E(\theta|\theta_c, s'_\ell) = \frac{\frac{\theta_c}{\sigma_c^2} + \frac{s'_\ell}{(\sigma_\ell^2+\sigma_\epsilon^2)}}{\frac{1}{(\sigma_\ell^2+\sigma_\epsilon^2)} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma^2}}, \tag{8}$$

with $s'_\ell = \theta_\ell + \epsilon$.

Under the decentralized regime, the optimal policy for the local government is given by

$$a_\ell = (1-\gamma)E(\theta|\theta_\ell, s'_c) + \gamma E(\theta_c|\theta_\ell, s'_c) = k'_1\theta_\ell + k'_2 s'_c, \tag{9}$$

with $s'_c = \theta_c + \epsilon$ and

$$\begin{aligned} k'_1 &\equiv \frac{\frac{1-\gamma}{\sigma_\ell^2}}{\frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}} + \frac{\frac{\gamma}{\sigma_\ell^2+\sigma_c^2+\sigma_c^2\sigma_\ell^2/\sigma^2}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}} \\ k'_2 &\equiv \frac{\frac{1-\gamma}{\sigma_c^2+\sigma_\epsilon^2}}{\frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}} + \frac{\frac{\gamma}{\sigma_\epsilon^2}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}}. \end{aligned}$$

According to the lemma, the government policy is always a linear combination of two private signals, θ_c and θ_ℓ, and the communication friction ε. According to Lemma 5, to analyze economic volatility and output, we can again just focus on $E(a_c-\theta)^2$ and $E(a_\ell-\theta)^2$.

Lemma 15. Under the centralized regime,

$$E(a_c-\theta)^2 = Var(\theta|\theta_c, s'_\ell) = \left(\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)^{-1}.$$

Under the decentralized regime,

$$\begin{aligned} E(a_\ell-\theta)^2\Big|_{\gamma=0} &= Var(\theta|\theta_\ell, s'_c) = \left(\frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)^{-1} \\ E(a_\ell-\theta)^2\Big|_{\gamma=1} &= \frac{\frac{\sigma_c^2}{\sigma_\epsilon^2} + \frac{(1/\sigma^2+1/\sigma_\ell^2)^{-1}}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}} \end{aligned}$$

Lemma 16. In the absence of the loyalty concern ($\gamma = 0$), decentralization leads to lower economic volatility (and higher economic output).

The assumptions $\sigma_\ell^2 < \sigma_c^2$ and the presence of communication friction play a crucial role in this result. Notice in the absence of communication friction, both governments would share the same information set $\{\theta_c, \theta_\ell\}$ through communication, thus leading to the same economic outcome under different regimes.[21]

Lemma 17. Under the decentralized regime with $\gamma = 1$, there exists a unique $\bar{\sigma}_\epsilon^2 > 0$ such that $E(a_\ell - \theta)^2 = E(a_c - \theta)^2$ for $\sigma_\epsilon^2 = \bar{\sigma}_\epsilon^2$ and $E(a_\ell - \theta)^2 > E(a_c - \theta)^2$ if and only if $\sigma_\epsilon^2 < \bar{\sigma}_\epsilon^2$.

When the objective function of the local government is purely loyalty-based, it makes every effort to match its policy with whatever is sent by the central government. If the communication friction vanishes, the local government will ignore its own, more precise signal and simply follow the policy prescription from the central government. In this case, the local government makes its economic decision as if it received only one signal from the nature, θ_c, thus leading to worse economic outcome. To the other extreme, if the communication is prohibitively noisy ($\sigma_\epsilon^2 \to \infty$), the local government cannot rely on the signal it receives from the central government to predict s_c from the central government. Instead, it has to rely on its own signal, which is also correlated with s_c (via the mutual component θ). In this case, it is possible that decentralization leads to better economic performance even in the presence of pure loyalty concern.

This result stands in sharp contrast with Lemma 13 in which decentralization with $\gamma = 1$ always leads to worse economic performance. In the benchmark with rational inattention, loyalty-driven distortion comes from two margins: (1) the local government assigns suboptimal (from the standpoint of the social welfare) weights to the signals it obtains; (2) the local government endogenously decreases the precision of its own signal about the state of the economy as the loyalty concern rises. Our results in this non-strategic environment demonstrate the importance of the second margin in order to generate contrasting dynamics following different types of decentralization.

It can be easily seen that the expected output strictly decreases with σ_ϵ^2 under the centralized regime and the decentralized regime with $\gamma = 0$. When governments attempt to maximize expected output, additional communication friction always worsens economic outcome. However, the intuition gets reversed when we turn to a loyalty-driven local government. In fact, we can prove the following seemingly paradoxical result: *higher communication friction could be welfare improving under decentralization.*

Proposition 5. Under the decentralized regime with $\gamma = 1$, $\partial E(a_\ell - \theta)^2 / \partial(\sigma_\epsilon^2) < 0$.

We now prove a counterpart of Proposition 4 in this alternative setup. In the absence of strategic considerations, the proof is much simpler.

Corollary 2. Under the decentralized regime in a non-strategic environment, $E(a_\ell - \theta)^2$ strictly increases with γ.

Now we are ready to provide the main theorem in this non-strategic environment, a complete characterization of the relative economic performance under two regimes.

Theorem 2. In a non-strategic environment, there exists a unique $\bar{\sigma}_\epsilon^2 > 0$ such that

1. if $\sigma_\epsilon^2 > \bar{\sigma}_\epsilon^2$, decentralization always leads to lower economic volatility
2. if $\sigma_\epsilon^2 < \bar{\sigma}_\epsilon^2$, there exists a unique $\hat{\gamma} \in (0, 1)$ such that (i) if $\gamma < \hat{\gamma}$, decentralization lowers output volatility; (ii) if $\gamma > \hat{\gamma}$, decentralization increases output volatility.

[21] Technically, the policy in our model is a weighted average of the two private signals(-cum-noise). In the presence of communication friction, each government tends to attach higher weight to its own signal. Assuming the local government has a more precise signal, the differential equilibrium outcome arises ex post.

5.2 Strategic Communication

In the baseline setting, we assume that the signal sender always reveals its information truthfully: $s_\ell = \theta_\ell$ under the centralized regime and $s_c = \theta_\ell$ under the decentralized regime.

We now relax this assumption by allowing the signal sender to manipulate its signal. In particular, under the decentralized regime, we assume the local government sends a signal of the form

$$s_\ell = \theta_\ell + \delta_\ell$$

with $\delta_\ell \sim \mathcal{N}(0, \sigma^2_{\delta\ell})$. Under the centralized regime, the central government sends a signal of the form

$$s_c = \theta_c + \delta_c$$

with $\delta_c \sim \mathcal{N}(0, \sigma^2_{\delta c})$. We assume that the signal sender can choose any $\sigma^2_{\delta\ell}$ (or $\sigma^2_{\delta c}$) without incurring any cost.

In an equilibrium, the signal receiver correctly expects the variance of the white noise added by the signal sender and acts accordingly. Therefore, the decision problem of the signal receiver is the same as that in the baseline setting by simply replacing σ^2_ℓ with $\sigma^2_\ell + \sigma^2_{\delta\ell}$ under the centralized regime and replacing σ^2_c with $\sigma^2_c + \sigma^2_{\delta c}$ under the decentralized regime. The following result then immediately follows from Lemmas 7 and 11.

Proposition 6. In the case of strategic communication, there is no incentive for the signal sender to introduce additional noise into its signal: $\sigma^2_{\delta\ell} = \sigma^2_{\delta c} = 0$.

5.3 Comparative Advantage of Information Acquisition

In our baseline setting, we assume that the local government enjoys absolute advantage in information acquisition over the central government ($\kappa_\ell > \kappa_c$). This assumption may appear too strong under certain economic situations. We now assume, instead, the local government only enjoys comparative advantage in the direct acquisition of economic information. In particular, we assume that $\kappa_\ell = \kappa_c = \kappa$, and we replace Constraint 4 with

$$\left(\frac{1}{\sigma^2_\epsilon} + \frac{1}{\sigma^2_c} + \frac{1}{\sigma^2_{\epsilon\ell}}\right)\left(\frac{1}{\sigma^2} + \frac{1}{\sigma^2_c} + \frac{\lambda}{\sigma^2_\ell}\right) \leq \frac{2^{2\kappa}(\sigma^2 + \sigma^2_c + \sigma^2_\epsilon)}{\sigma^2\sigma^2_c\sigma^2_\epsilon} + \frac{1}{\sigma^4_c} \equiv K_\ell(\sigma^2_c), \tag{10}$$

and replace Constraint 1 with

$$\frac{1}{\sigma^2} + \frac{\lambda}{\sigma^2_\ell} \leq \frac{2^{2\kappa}}{\sigma^2}, \tag{11}$$

where $0 < \lambda < 1$. Under this assumption, the lowest attainable $\sigma^2_{\epsilon\ell}$ under the decentralized regime is the same as the lowest attainable $\sigma^2_{\epsilon c}$ under the centralized regime, while the local government has information advantage over the central government if both governments devote their attention to direct information acquisition.[22]

[22] We introduce λ in a rather reduced-form way, which conveys intuition in a transparent way, but the drawback is that we deviate from the standard entropy reduction framework. Strictly speaking, what we deal with is no longer an entropy. In fact, it is difficult to introduce the notion of comparative advantage in information acquisition into this framework without deviating from the formal definition of entropy, because as we point out earlier, θ and ϵ are not independent conditional on θ_ℓ, s'_c and s''_c

We now turn to the equilibrium characterization of this alternative setting.

First, notice that the optimal policy does not depend on λ, so Lemmas 3 and 4 carry over.

Under the centralized regime, the decision problem of the central government is unchanged. The only departure from the baseline setting is the introduction of the scaling parameter λ into the constraint of the signal sender (Constraint 11), so the following result immediately follows from Proposition 1.

Proposition 7. Let $0 < \lambda < 1$. Under the centralized regime, we have

$$E(a_c - \theta)^2 = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2 + \sigma_\epsilon^2}\right)^{-1}$$

with $\sigma_\ell^2 = \lambda\sigma^2/(2^{2\kappa} - 1)$ and

$$\sigma_c^2 = \left[K_c\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_\ell^2}\right)^{-1} - \frac{1}{\sigma^2} - \frac{1}{\sigma_\ell^2}\right]^{-1} = \frac{\sigma^2(\sigma_\epsilon^2 + \sigma_\ell^2)}{(2^{2\kappa} - 1)(\sigma^2 + \sigma_\epsilon^2 + \sigma_\ell^2)}.$$

Under the decentralized regime, since now the local government, which is the signal receiver, enjoys information comparative advantage, the trade-off of the resource allocation for the local government hinges on λ. If λ is sufficiently close to zero, then the local government always devotes itself to direct information acquisition. To rule out this less interesting case, unless explicitly stated, we always impose the following regularity condition on λ throughout this subsection.

$$\lambda(1/\lambda - 1)^2 < \frac{2^{4\kappa}\left(\frac{1}{2^{2\kappa}-1} + \frac{2^{2\kappa}\sigma^2}{(2^{2\kappa}-1)^2\sigma_\epsilon^2}\right)^2}{2^{2\kappa}\left(\frac{1}{2^{2\kappa}-1} + \frac{2^{2\kappa}\sigma^2}{(2^{2\kappa}-1)^2\sigma_\epsilon^2}\right) + 1} \tag{12}$$

for given $\kappa, \sigma^2, \sigma_\epsilon^2 > 0$. Since λ has to be in $(0, 1)$, this condition is equivalent to imposing a lower bound on λ. The condition may appear complicated, so we provide the following technical result that sharpens the lower bound for λ.

Lemma 18. If $\lambda \geq 1/2$, then the regularity condition 12 holds for any κ, σ^2, and σ_ϵ^2.

We prove the following counterpart of Lemma 10. It can be seen that the the characterization of the optimal policy for a general γ is qualitatively unchanged in this alternative setting.

Lemma 19. Let $0 < \lambda < 1$. Under the decentralized regime, there exist two cutoffs $\underline{\gamma}'$ and $\bar{\gamma}'$ such that $0 < \underline{\gamma}' < \bar{\gamma}' < 1$. If $\gamma \leq \underline{\gamma}'$, the local government specializes in direct information acquisition ($\sigma_{\epsilon\ell}^2 = \infty$). If $\gamma \geq \bar{\gamma}'$, the local government specializes in intergovernmental communication ($\sigma_\ell^2 = \infty$). If $\gamma \in (\underline{\gamma}', \bar{\gamma}')$, the local government allocates its budget to both activities ($\sigma_\ell^2 < \infty$ and $\sigma_{\epsilon\ell}^2 < \infty$) with $\partial\sigma_\ell^2/\partial\gamma > 0$.

Remark 3. As is made clear in the proof of Lemma 19, if the central government devotes itself to the direct information acquisition, then Condition 12 is necessary and sufficient for the local government to specialize in inter-governmental communication when $\gamma = 1$.

It turns out it is very challenging to obtain the counterpart of Lemma 11. Instead, we provide the following weaker extension.

Lemma 20. Let $\lambda \in (0,1)$. Under the decentralized regime, the economic volatility strictly increases with σ_c^2 for σ_c^2 and γ such that $\gamma < \underline{\gamma}'$ or $\gamma > \bar{\gamma}'$.

Notice that $\underline{\gamma}'$ and $\bar{\gamma}'$ are functions of σ_c^2. Although according to the simulation, the monotonicity property also applies to the case that $\gamma \in [\underline{\gamma}', \bar{\gamma}']$, we are not able to formally establish this result. Since we know $0 < \underline{\gamma}' < \bar{\gamma}' < 1$, we then have a partial characterization of the optimal strategy of the central government.

Corollary 3. Let $\lambda \in (0,1)$. Under the decentralized regime, the central government devotes itself to information acquisition with $\sigma_c^2 = \sigma^2/(2^{2\kappa} - 1)$ if $\gamma = 0$ or $\gamma = 1$.

Lemma 21. Let $0 < \lambda < 1$. We have

$$\begin{aligned} E(a_c - \theta)^2 &> E(a_\ell - \theta)^2\Big|_{\gamma=0}, \\ E(a_c - \theta)^2 &< E(a_\ell - \theta)^2\Big|_{\gamma=1}. \end{aligned}$$

We are now ready to state the main result in the setting with comparative advantage of information acquisition.

Theorem 3. Let $\lambda < 1$ satisfy Condition 12. If γ is sufficiently close to one, decentralization leads to higher economic volatility (and lower economic output); if γ is sufficiently close to zero, decentralization leads to lower economic volatility (and higher economic output).

According to our simulation results, the comparison between two regimes is very similar to what has been illustrated by Figure 4: economic volatility increases monotonically with γ and there are two kinks on the curve of $E(a_\ell - \theta)^2$, representing the structural changes of the local government's optimal strategy when $\gamma = \underline{\gamma}'$ and $\gamma = \bar{\gamma}'$.

To close this subsection, we consider the case the local government has very strong comparative advantage such that Condition 12 does not hold. In this case, $\bar{\gamma}'$ disappears. Since there is no equilibrium characterization for the middle range of γ, we perform a battery of numerical experiments. As shown in Figure 6, other things equal, when σ_ϵ^2 is relatively large, then decentralization always leads to improvement of economic performance; when σ_ϵ^2 is relatively small, then the economic outcome of decentralization hinges on γ. The numerical results are qualitatively similar to what we have shown for setting with exogenous communication friction.[23]

5.4 Other Extensions

6 A Case Study

In this section, we present a case study of China and discuss how the historical reform experience and empirical evidence support the basic assumption and main implications of our theory.[24]

[23] Since $\bar{\gamma}'$ disappears, there is only one kink on the curve of $E(a_\ell - \theta)^2$.
[24] This section is to be completed.

6.1 Background and Stylized Facts

China started to build a central-planning economy in the Soviet style since its establishment and finished it in 1956. But then decentralization has been implemented at various time and in different ways because people realized that centralization has its intrinsic flaws. The first decentralization practice is to transfer SOEs to government at lower levels, one of the primary measure for reforming SOEs in China from 1958 until 1978 (Wu, 2005). Chairman Mao Zedong proposed "It is not good to concentrate everything in the hands of the central authorities and impose rigid controls."[25] The primary feature of this decentralization is that only the power to execute local economy was decentralized, the promotion rule still remained similar. The central government did not regard performance of local economy as an important criterion of promotion, but still highlight the importance of political accordance. Provincial leaders were rewarded for following instructions from the central government. The Great Famine from 1959 to 1961 is an example. Provincial governments followed a wrong decision by the central government to express loyalty and finally incurred the tragedy. Related to the theme of this paper, surging output volatility has been another consequence of the first decentralization practice. Wu (2005) noticed that there is a relationship between decentralization and volatility. He made the comment "Administrative orders inherently required a high degree of consistency in command and control by the government, administrative decentralization inevitably led to malfunction of the centrally planned system and to economic disorder."

The second decentralization practice began from 1978. It has been regarded as the most important factor in the recent growth of China (Xu, 2011). Different from the first one, this decentralization regime was featured by giving the economic incentive to local leaders. The central government emphasized the importance of economic performance in promotion criterion (Li and Zhou, 2005). This wave of decentralization leads to rapid economic growth, which has been a central focus of a voluminous literature. However, little attention has been paid to the economic volatility following these two episodes of decentralization.

As we have shown in Figure 1, the economic volatility was rocket high in the 1960s while the economic growth reached its trough. "In November 1957, the state council $\cdots$ out of a total of 9,300 enterprises and public institutions directly under the central departments, 8,100 were transferred to local governments in 1958." "Economic disorder resulting from the comprehensive delegation of administrative power together with the Great Leap Forward campaign forced the government to recentralized the control over SOEs". [26]. In sharp contrast, following the second wave of decentralization, the output volatility of the economy steadily decreased. Figure 7 depicts the output volatility by regions. We can find the trends are similar across different regions.

6.2 Supportive Evidence

We first provide evidence on the assumption that the central government put different weights on promotion incentives in 1950s and 1980s. We construct a dataset on promotion outcomes of provincial leaders in China and try to use two sets of independent variables to explain the pattern

[25] "Letter on Farm Mechanization (Guanyu nongye jixiehua wenti de yi feng xin)" (March 12, 1966), People's Daily (Renmin ribao), December 26, 1977.

[26] The share of enterprises directly under the central departments in national total of gross value of industrial output (GVIO) decreased from 39.7 percent in 1957 to 13.8 percent in 1958. Zhou Taihe et al., Economic System Reform in Contemporary China (Dangdai Zhongguo de jingji tizhi gaige), Beijing: China Social Sciences Press, 1984, p. 70

of promotion. The first set are personal backgrounds of these officials, including whether he was promoted from local grassroots, whether he was a local official, whether he had experience from the central government and whether he had experience from the army. These variables measure how these officials are reliable for the central government. The role of them in promotion reveals the weight given by the central government to political loyalty in the process of promotion. The other set includes the GDP per capita growth rate, measuring the economic weight in contrast. Table 1 shows that the first set of variables are more important in the 1950s and the second is more important in the 1980s. The empirical results confirm that the central government put more weights on political consideration in the first wave of decentralization, which is believed to be responsible for the high volatility in our paper.

We then check the cross-sectional variation in promotion incentive will affect the regional output volatility. We use the percentage of state-owned manufacturing output in 1950 as a proxy for the political incentive of a province. Intuitively, for those highly nationalized region, the industrial production will follow instructions from the central government more. As a result, the volatility of those provinces in 1950s would be higher. However, after 1978, when the nation became more market-oriented, the gap in volatility will be declining. Figure 8 confirms this.

7 Conclusion

In this paper, we theoretically demonstrate that the impact of decentralization on economic volatility in an authoritarian regime hinges on the degree of loyalty concern of the local government. Decentralization could be welfare-reducing and lead to higher volatility if the local government is a loyal follower of the central government. The empirical analysis suggests that the contrasting dynamics following the two waves of decentralization in China can be rationalized by our theory. However, it then raises another question of why the central government would choose to decentralize despite knowing its detrimental effects to the economy. This question is out of the scope of this paper, but we think it will be a fruitful direction to investigate the deep, non-economic roots of decentralization.

A Proofs

A.1 Proof of Lemma 1

Proof. Since $H(\theta,\epsilon|s'_\ell) = H(\theta,\epsilon,s'_\ell) - H(s'_\ell)$ and $H(\theta,\epsilon|\theta_c,s'_\ell,s''_\ell) = H(\theta,\epsilon,\theta_c,s'_\ell,s''_\ell) - H(\theta_c,s'_\ell,s''_\ell)$, we have[27]

$$\begin{aligned} H(\theta,\epsilon|s'_\ell) - H(\theta,\epsilon|\theta_c,s'_\ell,s''_\ell) &= H(\theta,\epsilon,s'_\ell) + H(\theta_c,s'_\ell,s''_\ell) - H(s'_\ell) - H(\theta,\epsilon,\theta_c,s'_\ell,s''_\ell) \\ &= \frac{1}{2}\left(\log_2|\Sigma_{\theta,\epsilon,s'_\ell}| + \log_2|\Sigma_{\theta_c,s'_\ell,s''_\ell}| - \log_2|\Sigma_{s'_\ell}| - \log_2|\Sigma_{\theta,\epsilon,\theta_c,s'_\ell,s''_\ell}|\right) \le \kappa_c. \end{aligned}$$

[27] Alternatively, for two multivariate normal distributions X and Y, we have $|\Sigma_{X|Y}||\Sigma_Y| = |\Sigma_X|$.

Under the assumption of the model, we have $|\Sigma_{s'_\ell}| = \sigma^2 + \sigma_\ell^2 + \sigma_\epsilon^2$ and

$$|\Sigma_{\theta,\epsilon,s'_\ell}| = \begin{vmatrix} \sigma^2 & 0 & \sigma^2 \\ 0 & \sigma_\epsilon^2 & \sigma_\epsilon^2 \\ \sigma^2 & \sigma_\epsilon^2 & \sigma^2+\sigma_\ell^2+\sigma_\epsilon^2 \end{vmatrix} = \sigma^2\sigma_\epsilon^2\sigma_\ell^2,$$

$$|\Sigma_{\theta_c,s'_\ell,s''_\ell}| = \begin{vmatrix} \sigma^2+\sigma_c^2 & \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2+\sigma_\ell^2+\sigma_\epsilon^2 & \sigma^2+\sigma_\ell^2 \\ \sigma^2 & \sigma^2+\sigma_\ell^2 & \sigma^2+\sigma_\ell^2+\sigma_{\epsilon c}^2 \end{vmatrix} = (\sigma^2+\sigma_c^2)(\sigma_\ell^2\sigma_{\epsilon c}^2+\sigma_\epsilon^2\sigma_{\epsilon c}^2+\sigma_\epsilon^2\sigma_\ell^2)+\sigma_c^2\sigma^2(\sigma_{\epsilon c}^2+\sigma_\epsilon^2),$$

$$|\Sigma_{\theta,\epsilon,\theta_c,s'_\ell,s''_\ell}| = \begin{vmatrix} \sigma^2 & 0 & \sigma^2 & \sigma^2 & \sigma^2 \\ 0 & \sigma_\epsilon^2 & 0 & \sigma_\epsilon^2 & 0 \\ \sigma^2 & 0 & \sigma^2+\sigma_c^2 & \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma_\epsilon^2 & \sigma^2 & \sigma^2+\sigma_\ell^2+\sigma_\epsilon^2 & \sigma^2+\sigma_\ell^2 \\ \sigma^2 & 0 & \sigma^2 & \sigma^2+\sigma_\ell^2 & \sigma^2+\sigma_\ell^2+\sigma_{\epsilon c}^2 \end{vmatrix} = \sigma^2\sigma_\epsilon^2\sigma_c^2\sigma_\ell^2\sigma_{\epsilon c}^2.$$

Plugging the three determinants into the information flow constraint, we obtain

$$\left(\frac{1}{\sigma_c^2\sigma_\epsilon^2}+\frac{1}{\sigma^2\sigma_\epsilon^2}+\frac{1}{\sigma_c^2\sigma_\ell^2}+\frac{1}{\sigma^2\sigma_\ell^2}+\frac{1}{\sigma_c^2\sigma_{\epsilon c}^2}+\frac{1}{\sigma^2\sigma_{\epsilon c}^2}+\frac{1}{\sigma_\epsilon^2\sigma_\ell^2}+\frac{1}{\sigma_\ell^2\sigma_{\epsilon c}^2}\right) \le 2^{2\kappa_c}\frac{\sigma^2+\sigma_\ell^2+\sigma_\epsilon^2}{\sigma^2\sigma_\epsilon^2\sigma_\ell^2}.$$

Simplifying the expression above, we obtain the desired conclusion. □

A.2 Proof of Lemma 2

Proof. Since $H(\theta,\epsilon|s'_c) = H(\theta,\epsilon,s'_c) - H(s'_c)$ and $H(\theta,\epsilon|\theta_\ell,s'_c,s''_c) = H(\theta,\epsilon,\theta_\ell,s'_c,s''_c) - H(\theta_\ell,s'_c,s''_c)$,

$$\begin{aligned} H(\theta,\epsilon|s'_c) - H(\theta,\epsilon|\theta_\ell,s'_c,s''_c) &= H(\theta,\epsilon,s'_c) + H(\theta_\ell,s'_c,s''_c) - H(s'_c) - H(\theta,\epsilon,\theta_\ell,s'_c,s''_c) \\ &= \frac{1}{2}\left(\log_2|\Sigma_{\theta,\epsilon,s'_c}| + \log_2|\Sigma_{\theta_\ell,s'_c,s''_c}| - \log_2|\Sigma_{s'_c}| - \log_2|\Sigma_{\theta,\epsilon,\theta_\ell,s'_c,s''_c}|\right) \le \kappa_\ell. \end{aligned}$$

Under the assumption of the model, we have $|\Sigma_{s'_c}| = \sigma^2 + \sigma_c^2 + \sigma_\epsilon^2$ and

$$|\Sigma_{\theta,\epsilon,s'_c}| = \begin{vmatrix} \sigma^2 & 0 & \sigma^2 \\ 0 & \sigma_\epsilon^2 & \sigma_\epsilon^2 \\ \sigma^2 & \sigma_\epsilon^2 & \sigma^2+\sigma_c^2+\sigma_\epsilon^2 \end{vmatrix} = \sigma^2\sigma_\epsilon^2\sigma_c^2,$$

$$|\Sigma_{\theta_\ell,s'_c,s''_c}| = \begin{vmatrix} \sigma^2+\sigma_\ell^2 & \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2+\sigma_c^2+\sigma_\epsilon^2 & \sigma^2+\sigma_c^2 \\ \sigma^2 & \sigma^2+\sigma_c^2 & \sigma^2+\sigma_c^2+\sigma_{\epsilon\ell}^2 \end{vmatrix} = (\sigma^2+\sigma_\ell^2)(\sigma_c^2\sigma_{\epsilon\ell}^2+\sigma_\epsilon^2\sigma_{\epsilon\ell}^2+\sigma_\epsilon^2\sigma_c^2)+\sigma_\ell^2\sigma^2(\sigma_{\epsilon\ell}^2+\sigma_\epsilon^2),$$

$$|\Sigma_{\theta,\epsilon,\theta_\ell,s'_c,s''_c}| = \begin{vmatrix} \sigma^2 & 0 & \sigma^2 & \sigma^2 & \sigma^2 \\ 0 & \sigma_\epsilon^2 & 0 & \sigma_\epsilon^2 & 0 \\ \sigma^2 & 0 & \sigma^2+\sigma_\ell^2 & \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma_\epsilon^2 & \sigma^2 & \sigma^2+\sigma_c^2+\sigma_\epsilon^2 & \sigma^2+\sigma_c^2 \\ \sigma^2 & 0 & \sigma^2 & \sigma^2+\sigma_c^2 & \sigma^2+\sigma_c^2+\sigma_{\epsilon\ell}^2 \end{vmatrix} = \sigma^2\sigma_\epsilon^2\sigma_c^2\sigma_\ell^2\sigma_{\epsilon\ell}^2.$$

Plugging the three determinants into the information flow constraint, we obtain

$$\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}\right)+\frac{1}{\sigma_c^2}\left(\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_\epsilon^2}\right)+\frac{1}{\sigma_c^4}\leq 2^{2\kappa_\ell}\frac{\sigma^2+\sigma_c^2+\sigma_\epsilon^2}{\sigma^2\sigma_\epsilon^2\sigma_c^2}+\frac{1}{\sigma_c^4}.$$

Simplifying the expression above, we obtain the desired conclusion. □

A.3 Proof of Lemma 3

Proof. Under the centralized regime, given the quadratic form of the objective function, the optimal policy for the central government is given by

$$a_c = E(\theta|\theta_c, s'_\ell, s''_\ell).$$

To find the expression of the conditional expectation above, we first notice the probability density function of the joint distribution of $(\theta, \theta_c, \theta_\ell, s'_\ell, s''_\ell)$ can be written as

$$\begin{aligned} f(\theta,\theta_c,\theta_\ell,s'_\ell,s''_\ell) &= f(\theta)f(\theta_c,\theta_\ell,s'_\ell,s''_\ell|\theta)=f(\theta)f(\theta_c|\theta)f(\theta_\ell,s'_\ell,s''_\ell|\theta)\\ &= f(\theta)f(\theta_c|\theta)f(\theta_\ell|\theta)f(s'_\ell,s''_\ell|\theta_\ell)=f(\theta)f(\theta_c|\theta)f(\theta_\ell|\theta)f(s'_\ell|\theta_\ell)f(s''_\ell|\theta_\ell), \end{aligned}$$

where the second to last equation stems from the fact that conditional on θ_ℓ, ϵ and ϵ_c are independent of θ. More explicitly, we have[28]

$$\begin{aligned} f(\theta,\theta_c,\theta_\ell,s'_\ell,s''_\ell) &\sim \exp\left\{-\frac{1}{2}\left[\frac{\theta^2}{\sigma^2}+\frac{(\theta_c-\theta)^2}{\sigma_c^2}+\frac{(\theta_\ell-\theta)^2}{\sigma_\ell^2}+\frac{(s'_\ell-\theta_\ell)^2}{\sigma_\epsilon^2}+\frac{(s''_\ell-\theta_\ell)^2}{\sigma_{\epsilon c}^2}\right]\right\}\\ &= \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon c}^2}\right)\theta_\ell^2-2\left(\frac{\theta}{\sigma_\ell^2}+\frac{s'_\ell}{\sigma_\epsilon^2}+\frac{s''_\ell}{\sigma_{\epsilon c}^2}\right)\theta_\ell\right.\right.\\ &\qquad\left.\left.+\frac{\theta^2}{\sigma^2}+\frac{(\theta_c-\theta)^2}{\sigma_c^2}+\frac{\theta^2}{\sigma_\ell^2}+\frac{s_\ell'^2}{\sigma_\epsilon^2}+\frac{s_\ell''^2}{\sigma_{\epsilon c}^2}\right]\right\}\\ &= \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon c}^2}\right)\left(\theta_\ell-\frac{\frac{\theta}{\sigma_\ell^2}+\frac{s'_\ell}{\sigma_\epsilon^2}+\frac{s''_\ell}{\sigma_{\epsilon c}^2}}{\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon c}^2}}\right)^2\right.\right.\\ &\qquad\left.\left.+\frac{\theta^2}{\sigma^2}+\frac{(\theta_c-\theta)^2}{\sigma_c^2}+\frac{\theta^2}{\sigma_\ell^2}+\frac{s_\ell'^2}{\sigma_\epsilon^2}+\frac{s_\ell''^2}{\sigma_{\epsilon c}^2}-\frac{\left(\frac{\theta}{\sigma_\ell^2}+\frac{s'_\ell}{\sigma_\epsilon^2}+\frac{s''_\ell}{\sigma_{\epsilon c}^2}\right)^2}{\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon c}^2}}\right]\right\} \end{aligned}$$

[28] Alternatively, we know $f(\theta,\theta_c,\theta_\ell,s'_\ell,s''_\ell)\sim\exp\left(-\frac{1}{2}(\theta,\theta_c,\theta_\ell,s'_\ell,s''_\ell)\Sigma^{-1}_{\theta,\theta_c,\theta_\ell,s'_\ell,s''_\ell}(\theta,\theta_c,\theta_\ell,s'_\ell,s''_\ell)^T\right)$ with

$$\begin{aligned} \Sigma^{-1}_{\theta,\theta_c,\theta_\ell,s'_\ell,s''_\ell} &= \begin{pmatrix} \sigma^2 & \sigma^2 & \sigma^2 & \sigma^2 & \sigma^2\\ \sigma^2 & \sigma^2+\sigma_c^2 & \sigma^2 & \sigma^2 & \sigma^2\\ \sigma^2 & \sigma^2 & \sigma^2+\sigma_\ell^2 & \sigma^2+\sigma_\ell^2 & \sigma^2+\sigma_\ell^2\\ \sigma^2 & \sigma^2 & \sigma^2+\sigma_\ell^2 & \sigma^2+\sigma_\ell^2+\sigma_\epsilon^2 & \sigma^2+\sigma_\ell^2\\ \sigma^2 & \sigma^2 & \sigma^2+\sigma_\ell^2 & \sigma^2+\sigma_\ell^2 & \sigma^2+\sigma_\ell^2+\sigma_{\epsilon c}^2 \end{pmatrix}^{-1}\\ &= \begin{pmatrix} \frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2} & -\frac{1}{\sigma_c^2} & -\frac{1}{\sigma_\ell^2} & 0 & 0\\ -\frac{1}{\sigma_c^2} & \frac{1}{\sigma_c^2} & 0 & 0 & 0\\ -\frac{1}{\sigma_\ell^2} & 0 & \frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon c}^2} & -\frac{1}{\sigma_\epsilon^2} & -\frac{1}{\sigma_{\epsilon c}^2}\\ 0 & 0 & -\frac{1}{\sigma_\epsilon^2} & \frac{1}{\sigma_\epsilon^2} & 0\\ 0 & 0 & -\frac{1}{\sigma_{\epsilon c}^2} & 0 & \frac{1}{\sigma_{\epsilon c}^2} \end{pmatrix}. \end{aligned}$$

Integrating out θ_ℓ, we obtain

$$f(\theta,\theta_c,s'_\ell,s''_\ell)\sim\exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\left(\frac{1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2}{1/\sigma_\ell^2+1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2}\right)\right)\theta^2 -2\left(\frac{\theta_c}{\sigma_c^2}+\frac{(s'_\ell/\sigma_\epsilon^2+s''_\ell/\sigma_{\epsilon c}^2)/\sigma_\ell^2}{1/\sigma_\ell^2+1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2}\right)\theta\right]\right\}.$$

This leads to

$$f(\theta|\theta_c,s'_\ell,s''_\ell)\sim f(\theta,\theta_c,s'_\ell,s''_\ell)\sim\exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2)^{-1}}\right)\cdot\left(\theta-\frac{\frac{\theta_c}{\sigma_c^2}+\frac{1}{\sigma_\ell^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2)^{-1}}\left(\frac{s'_\ell/\sigma_\epsilon^2}{1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2}+\frac{s''_\ell/\sigma_{\epsilon c}^2}{1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2}\right)}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2+(1/\sigma_\epsilon^2+1/\sigma_{\epsilon c}^2)^{-1}}}\right)^2\right\},$$

which yields the closed-form solution to $E(\theta|\theta_c,s'_\ell,s''_\ell)$. □

A.4 Proof of Lemma 4

Proof. Under the centralized regime, given the quadratic form of the objective function, the optimal policy for the local government is given by

$$a_\ell=(1-\gamma)E(\theta|\theta_\ell,s'_c,s''_c)+\gamma E(\theta_c|\theta_\ell,s'_c,s''_c).$$

$E(\theta|\theta_\ell,s'_c,s''_c)$ can be obtained similarly as $E(\theta|\theta_c,s'_\ell,s''_\ell)$. For $E(\theta_c|\theta_\ell,s'_c,s''_c)$, we have

$$\begin{aligned}
f(\theta,\theta_c,\theta_\ell,s'_c,s''_c) &= f(\theta)f(\theta_c|\theta)f(\theta_\ell|\theta)f(s'_c|\theta_c)f(s''_c|\theta_c)\\
&\sim \exp\left\{-\frac{1}{2}\left[\frac{\theta^2}{\sigma^2}+\frac{(\theta_c-\theta)^2}{\sigma_c^2}+\frac{(\theta_\ell-\theta)^2}{\sigma_\ell^2}+\frac{(s'_c-\theta_c)^2}{\sigma_\epsilon^2}+\frac{(s''_c-\theta_c)^2}{\sigma_{\epsilon\ell}^2}\right]\right\}\\
&= \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)\theta^2-2\left(\frac{\theta_c}{\sigma_c^2}+\frac{\theta_\ell}{\sigma_\ell^2}\right)\theta+\frac{\theta_c^2}{\sigma_c^2}+\frac{\theta_\ell^2}{\sigma_\ell^2}+\frac{(s'_c-\theta_c)^2}{\sigma_\epsilon^2}+\frac{(s''_c-\theta_c)^2}{\sigma_{\epsilon\ell}^2}\right]\right\}\\
&= \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)\left(\theta-\frac{\frac{\theta_c}{\sigma_c^2}+\frac{\theta_\ell}{\sigma_\ell^2}}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}}\right)^2\right.\right.\\
&\quad\left.\left.+\frac{\theta_c^2}{\sigma_c^2}+\frac{\theta_\ell^2}{\sigma_\ell^2}+\frac{(s'_c-\theta_c)^2}{\sigma_\epsilon^2}+\frac{(s''_c-\theta_c)^2}{\sigma_{\epsilon\ell}^2}-\frac{\left(\frac{\theta_c}{\sigma_c^2}+\frac{\theta_\ell}{\sigma_\ell^2}\right)^2}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}}\right]\right\}.
\end{aligned}$$

Integrating out θ, we obtain

$$f(\theta_c,\theta_\ell,s'_c,s''_c)\sim\exp\left\{-\frac{1}{2}\left[\frac{\theta_c^2}{\sigma_c^2}+\frac{\theta_\ell^2}{\sigma_\ell^2}+\frac{(s'_c-\theta_c)^2}{\sigma_\epsilon^2}+\frac{(s''_c-\theta_c)^2}{\sigma_{\epsilon\ell}^2}-\frac{(\theta_c/\sigma_c^2+\theta_\ell/\sigma_\ell^2)^2}{1/\sigma^2+1/\sigma_c^2+1/\sigma_\ell^2}\right]\right\}.$$

This leads to

$$f(\theta_c|\theta_\ell,s'_c,s''_c)\sim f(\theta_c,\theta_\ell,s'_c,s''_c)\sim\exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}\right)\right.$$

$$\cdot \left(\theta_c - \frac{\frac{s_c'}{\sigma_\epsilon^2} + \frac{s_c''}{\sigma_{\epsilon\ell}^2} + \frac{\theta_\ell}{\sigma_c^2+\sigma_\ell^2+\sigma_c^2\sigma_\ell^2/\sigma^2}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}} \right)^2 \Bigg\},$$

which implies the closed form solution to $E(\theta_c|\theta_\ell, s_c', s_c'')$. Plugging the expressions of $E(\theta|\theta_\ell, s_c', s_c'')$ and $E(\theta_c|\theta_\ell, s_c', s_c'')$ into the equation for a_ℓ, we obtain the desired conclusion. □

A.5 Proof of Lemma 5

Proof. Writing $(a_i - \theta)$ as a linear combination of six independent normal random variables with mean zero,

$$a_i - \theta = m_{i1}(\theta_\ell - \theta) + m_{i2}(\theta_c - \theta) + m_{i3}\epsilon + m_{i4}\epsilon_c + m_{i5}\epsilon_\ell + (m_{i1} + m_{i2} - 1)\theta.$$

Then according to the central moments of a normal distribution, we have

$$E(a_i - \theta)^2 = m_{i1}^2\sigma_\ell^2 + m_{i2}^2\sigma_c^2 + m_{i3}^2\sigma_\epsilon^2 + m_{i4}^2\sigma_{\epsilon c}^2 + m_{i5}^2\sigma_{\epsilon\ell}^2 + (1 - m_{i1} - m_{i2})^2\sigma^2,$$

$$E(a_i-\theta)^4 = 3\left(m_{i1}^2\sigma_\ell^2 + m_{i2}^2\sigma_c^2 + m_{i3}^2\sigma_\epsilon^2 + m_{i4}^2\sigma_{\epsilon c}^2 + m_{i5}^2\sigma_{\epsilon\ell}^2 + (1 - m_{i1} - m_{i2})^2\sigma^2\right)^2 = 3\left(E(a_i - \theta)^2\right)^2.$$

Therefore, by definition,

$$E(Y) \equiv Y^* - E(a_i-\theta)^2 = Y^* - \left(m_{i1}^2\sigma_\ell^2 + m_{i2}^2\sigma_c^2 + m_{i3}^2\sigma_\epsilon^2 + m_{i4}^2\sigma_{\epsilon c}^2 + m_{i5}^2\sigma_{\epsilon\ell}^2 + (1 - m_{i1} - m_{i2})^2\sigma^2)\right),$$

$$Var(Y) = E(a_i - \theta)^4 - \left(E(a_i - \theta)^2\right)^2 = 2\left(E(a_i - \theta)^2\right)^2 = 2(Y^* - E(Y))^2.$$

□

A.6 Proof of Lemma 7

Proof. According to Lemma 6, under the centralized regime, $\sigma_{\epsilon c}^2 = \infty$, so we can simplify a_c as

$$a_c = \frac{\frac{\theta_c}{\sigma_c^2} + \frac{s_\ell'}{\sigma_\ell^2+\sigma_\epsilon^2}}{\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}},$$

Using the backward induction, the local government solves its decision problem

$$\max_{\sigma_\ell^2} (1-\gamma)\left(Y^* - E(a_c - \theta)^2\right) - \gamma E(\theta_\ell - a_c)^2,$$

or equivalently,

$$\min_{\sigma_\ell^2} (1-\gamma)E(a_c - \theta)^2 + \gamma E(\theta_\ell - a_c)^2 \equiv F(\sigma_\ell^2),$$

subject to Constraint 1. Plugging in the expression of a_c, we can write the objective function explicitly as

$$\begin{aligned} F(\sigma_\ell^2) &= (1-\gamma)E(a_c - \theta)^2 + \gamma E(\theta_\ell - a_c)^2 \\ &= E(a_c - \theta)^2 + 2\gamma E(\theta_\ell - \theta)(\theta - a_c) + \gamma E(\theta_\ell - \theta)^2 \end{aligned}$$

$$= \frac{1}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}} - \frac{\frac{2\gamma\sigma_\ell^2}{\sigma_\ell^2+\sigma_\epsilon^2}}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}} + \gamma\sigma_\ell^2$$

Since we know $\sigma_{\epsilon c}^2 = \infty$, Constraint 2 gives us

$$\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2}\right) = K_c(\sigma_\ell^2) = 2^{2\kappa_c}\frac{\sigma^2+\sigma_\ell^2+\sigma_\epsilon^2}{\sigma^2\sigma_\ell^2\sigma_\epsilon^2} + \frac{1}{\sigma_\ell^4}.$$

The difficulty of this optimization problem arises from the equation above: Despite the fact that the central government always devotes to direct information acquisition, the resulting σ_c^2 is still a function of σ_ℓ^2 due to the nature of our information flow constraint.

Using the binding constraint, then the objective function can be rewritten as

$$\begin{aligned} F(\sigma_\ell^2) &= \frac{1}{\frac{K_c}{1/\sigma_\epsilon^2+1/\sigma_\ell^2}-\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}} - \frac{\frac{2\gamma\sigma_\ell^2}{\sigma_\ell^2+\sigma_\epsilon^2}}{\frac{K_c}{1/\sigma_\epsilon^2+1/\sigma_\ell^2}-\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}} + \gamma\sigma_\ell^2 \\ &= \frac{(1-2\gamma)\sigma_\ell^4+\sigma_\epsilon^2\sigma_\ell^2}{K_c\sigma_\epsilon^2\sigma_\ell^4-\sigma_\epsilon^2} + \gamma\sigma_\ell^2 \\ &= \frac{(1-2\gamma)\sigma_\ell^4+\sigma_\epsilon^2\sigma_\ell^2}{2^{2\kappa_c}\sigma_\epsilon^2\left((1/\sigma^2+1/\sigma_\epsilon^2)\sigma_\ell^2+\sigma_\ell^4/(\sigma^2\sigma_\epsilon^2)\right)} + \gamma\sigma_\ell^2. \end{aligned}$$

$$\begin{aligned} F'(\sigma_\ell^2) &= \frac{[(2-4\gamma)\sigma_\ell^2+\sigma_\epsilon^2]\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)\sigma_\ell^2+\frac{\sigma_\ell^4}{\sigma^2\sigma_\epsilon^2}\right] - [(1-2\gamma)\sigma_\ell^4+\sigma_\epsilon^2\sigma_\ell^2]\left[\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}+\frac{2\sigma_\ell^2}{\sigma^2\sigma_\epsilon^2}\right]}{2^{2\kappa_c}\sigma_\epsilon^2\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)\sigma_\ell^2+\frac{\sigma_\ell^4}{\sigma^2\sigma_\epsilon^2}\right]^2} + \gamma \\ &= \frac{\left[\frac{1-2\gamma}{\sigma_\epsilon^2}-\frac{2\gamma}{\sigma^2}\right]\sigma_\ell^4}{2^{2\kappa_c}\sigma_\epsilon^2\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)\sigma_\ell^2+\frac{\sigma_\ell^4}{\sigma^2\sigma_\epsilon^2}\right]^2} + \gamma \\ &\geq \frac{2^{2\kappa_c}\gamma\sigma_\epsilon^2\left[2\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)\frac{\sigma_\ell^6}{\sigma^2\sigma_\epsilon^2}+\frac{\sigma_\ell^8}{\sigma^4\sigma_\epsilon^4}\right] + \left[\frac{1}{\sigma_\epsilon^2}+\gamma\sigma_\epsilon^2\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)^2-\frac{2\gamma}{\sigma_\epsilon^2}-\frac{2\gamma}{\sigma^2}\right]\sigma_\ell^4}{2^{2\kappa_c}\sigma_\epsilon^2\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)\sigma_\ell^2+\frac{\sigma_\ell^4}{\sigma^2\sigma_\epsilon^2}\right]^2} \\ &= \frac{2^{2\kappa_c}\gamma\sigma_\epsilon^2\left[2\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)\frac{\sigma_\ell^6}{\sigma^2\sigma_\epsilon^2}+\frac{\sigma_\ell^8}{\sigma^4\sigma_\epsilon^4}\right] + \left[\frac{1-\gamma}{\sigma_\epsilon^2}+\frac{\gamma\sigma_\epsilon^2}{\sigma^4}\right]\sigma_\ell^4}{2^{2\kappa_c}\sigma_\epsilon^2\left[\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\epsilon^2}\right)\sigma_\ell^2+\frac{\sigma_\ell^4}{\sigma^2\sigma_\epsilon^2}\right]^2} > 0 \end{aligned}$$

where the first inequality follows from $\kappa_c > 0$ and the last inequality follows from $\gamma \in [0,1]$. Therefore, the objective function is minimized if σ_ℓ^2 attains its minimum, $\sigma^2/(2^{2\kappa_\ell}-1)$. □

A.7 Proof of Lemma 10

Proof. According to Lemma 4, we can rewrite the constrained optimization problem as

$$\min_{\sigma_\ell^2,\sigma_{\epsilon\ell}^2} E\left\{(1-\gamma)E[(a_\ell-\theta)^2|\theta_\ell,s_c',s_c''] + \gamma E[(a_\ell-\theta_c)^2|\theta_\ell,s_c',s_c'']\right\} \equiv F(\sigma_\ell^2,\sigma_{\epsilon\ell}^2),$$

with $a_\ell = (1-\gamma)E(\theta|\theta_\ell, s'_c, s''_c) + \gamma E(\theta_c|\theta_\ell, s'_c, s''_c) = k_1\theta_\ell + k_2 s'_c + k_3 s''_c$, subject to Constraint 4. Given the expression of a_ℓ, we have

$$\begin{aligned}
F(\sigma_\ell^2, \sigma_{\epsilon\ell}^2) &= E\left\{(1-\gamma)E[((1-\gamma)(E(\theta|\theta_\ell, s'_c, s''_c) - \theta) + \gamma(E(\theta_c|\theta_\ell, s'_c, s''_c) - \theta))^2|\theta_\ell, s'_c, s''_c]\right. \\
&\quad \left. + \gamma E[((1-\gamma)(E(\theta|\theta_\ell, s'_c, s''_c) - \theta_c) + \gamma(E(\theta_c|\theta_\ell, s'_c, s''_c) - \theta_c))^2|\theta_\ell, s'_c, s''_c]\right\} \\
&= E\left\{(1-\gamma)[(1-\gamma)^2 Var(\theta|\theta_\ell, s'_c, s''_c) + 2\gamma(1-\gamma)Var(\theta|\theta_\ell, s'_c, s''_c)\right. \\
&\quad +\gamma^2 E((E(\theta_c|\theta_\ell, s'_c, s''_c) - \theta_c + \theta_c - \theta)^2|\theta_\ell, s'_c, s''_c)] \\
&\quad +\gamma[\gamma^2 Var(\theta_c|\theta_\ell, s'_c, s''_c) + 2\gamma(1-\gamma)Var(\theta_c|\theta_\ell, s'_c, s''_c) \\
&\quad \left. +(1-\gamma)^2 E((E(\theta|\theta_\ell, s'_c, s''_c) - \theta + \theta - \theta_c)^2|\theta_\ell, s'_c, s''_c)]\right\} \\
&= (1-\gamma)^2(1+\gamma)Var(\theta|\theta_\ell, s'_c, s''_c) + \gamma^2(2-\gamma)Var(\theta_c|\theta_\ell, s'_c, s''_c) \\
&\quad +(1-\gamma)\gamma^2[Var(\theta_c|\theta_\ell, s'_c, s''_c) + 2E((E(\theta_c|\theta_\ell, s'_c, s''_c) - \theta_c)(\theta_c - \theta)) + \sigma_c^2] \\
&\quad +(1-\gamma)^2\gamma[Var(\theta|\theta_\ell, s'_c, s''_c) + 2E((E(\theta|\theta_\ell, s'_c, s''_c) - \theta)(\theta - \theta_c)) + \sigma_c^2] \\
&= (1-\gamma)^2(1+2\gamma)Var(\theta|\theta_\ell, s_c, s''_c) + \gamma^2(3-2\gamma)Var(\theta_c|\theta_\ell, s'_c, s''_c) + (1-\gamma)\gamma\sigma_c^2 \qquad (13) \\
&\quad -2\sigma_c^2\left\{(1-\gamma)\gamma^2 \frac{\frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}} + (1-\gamma)^2\gamma \frac{\frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}}{\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}}\right\}
\end{aligned}$$

To see that Constraint 4 has to be binding, we rewrite the objective function as

$$\begin{aligned}
F(\sigma_\ell^2, \sigma_{\epsilon\ell}^2) &= \frac{(1-\gamma)^2}{\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}} + \frac{\gamma^2}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}} + (1-\gamma)\gamma\sigma_c^2 \\
&\quad + \frac{\frac{2\gamma(1-\gamma)^2(1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}}{\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}} + \frac{\frac{2(1-\gamma)\gamma^2(1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}},
\end{aligned}$$

which strictly increases with σ_ℓ^2 or $\sigma_{\epsilon\ell}^2$.

Since Constraint 4 is binding, from Equation 13, the objective function can be further simplified to be

$$\begin{aligned}
F(\sigma_\ell^2, \sigma_{\epsilon\ell}^2) &= \frac{1}{K_\ell - 1/\sigma_c^4}\left[(1-\gamma)^2(1+2\gamma)\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2}\right) + \gamma^2(3-2\gamma)\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2}\right)\right] \\
&\quad +(1-\gamma)\gamma\sigma_c^2 - 2(1-\gamma)\gamma^2\frac{1/\sigma^2 + 1/\sigma_\ell^2}{K_\ell - 1/\sigma_c^4} - 2(1-\gamma)^2\gamma\frac{1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2}{K_\ell - 1/\sigma_c^4} \\
&= \frac{1}{K_\ell - 1/\sigma_c^4}\left[(1-\gamma)^2 K_\ell\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2}\right)^{-1} + \gamma^2\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2}\right)\right] \\
&\quad +(1-\gamma)\gamma\sigma_c^2 - \frac{2(1-\gamma)\gamma}{\sigma_c^2(K_\ell - 1/\sigma_c^4)}
\end{aligned}$$

Since $K_\ell > 1/\sigma_c^4$ and K_ℓ is constant with respect to σ_ℓ^2 and $\sigma_{\epsilon\ell}^2$, we can simplify the optimization problem as

$$\min_x \gamma^2 x + (1-\gamma)^2 K_\ell/x \equiv H(x)$$

with $x = 1/\sigma^2 + 1/\sigma_\ell^2 + 1/\sigma_c^2$ taking value from $[1/\sigma^2 + 1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)]$. Let the minimizer be x^*.

It is easy to see that when $\gamma = 0$, $x^* = K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)$ with $\sigma_{\epsilon\ell}^2 = \infty$, echoing Lemma 8, and when $\gamma = 1$, $x^* = 1/\sigma^2 + 1/\sigma_c^2$ with $\sigma_\ell^2 = \infty$, echoing Lemma 9. Moreover, we have

$$H'(x) = \gamma^2 - (1-\gamma)^2 K_\ell / x^2 \equiv G(\gamma; x).$$

Since $\frac{dG}{d\gamma}(\gamma; x) = 2\gamma + 2(1-\gamma)K_\ell/x^2 > 0$ for any $\gamma \in [0,1]$ and it is easy to see that $G(0;x) < 0$ and $G(1;x) > 1$, there exists a unique $\gamma \in (0,1)$ such that $G(\gamma;x) = 0$ for any given $x > 0$. Define $\underline{\gamma}$ such that $G(\underline{\gamma};x) = 0$ for $x = K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)$ and $\bar{\gamma}$ such that $G(\bar{\gamma};x) = 0$ for $x = 1/\sigma^2 + 1/\sigma_c^2$. Since $1/\sigma^2 + 1/\sigma_c^2 < K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)$, by construction, we have

$$\left(\frac{\bar{\gamma}}{1-\bar{\gamma}}\right)^2 = K_\ell \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2}\right)^{-2} > K_\ell \left(\frac{K_\ell}{1/\sigma_\epsilon^2 + 1/\sigma_c^2}\right)^{-2} = \left(\frac{\underline{\gamma}}{1-\underline{\gamma}}\right)^2,$$

which implies $\underline{\gamma} < \bar{\gamma}$.

Since $\frac{dG}{d\gamma}(\gamma; x) > 0$, for any $\gamma \le \underline{\gamma}$ and $x < K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)$, we must have

$$H'(x) = G(\gamma;x) \le G(\underline{\gamma};x) = \underline{\gamma}^2 - (1-\underline{\gamma})^2 K_\ell/x^2 < \underline{\gamma}^2 - (1-\underline{\gamma})^2 K_\ell/(K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2))^2 = 0.$$

Therefore, for $\gamma \le \underline{\gamma}$, $x^* = K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)$ with $\sigma_{\epsilon\ell}^2 = \infty$: The local government specializes in direct information acquisition provided that γ is sufficiently small.

Similarly, for any $\gamma \ge \bar{\gamma}$ and $x > 1/\sigma^2 + 1/\sigma_c^2$,

$$H'(x) = G(\gamma;x) \ge G(\bar{\gamma};x) = \bar{\gamma}^2 - (1-\bar{\gamma})^2 K_\ell/x^2 > \bar{\gamma}^2 - (1-\bar{\gamma})^2 K_\ell/(1/\sigma^2 + 1/\sigma_c^2)^2 = 0.$$

Therefore, for $\gamma \ge \bar{\gamma}$, $x^* = 1/\sigma^2 + 1/\sigma_c^2$ with $\sigma_\ell^2 = \infty$: The local government devotes all its attention budget to inter-governmental communication provided that γ is sufficiently large.

For $\gamma \in (\underline{\gamma}, \bar{\gamma})$, we have

$$\begin{aligned} H'(1/\sigma^2 + 1/\sigma_c^2) &= G(\gamma; 1/\sigma^2 + 1/\sigma_c^2) < G(\bar{\gamma}; 1/\sigma^2 + 1/\sigma_c^2) = 0 \\ H'(K_\ell/(1/\sigma^2 + 1/\sigma_c^2)) &= G(\gamma; K_\ell/(1/\sigma^2 + 1/\sigma_c^2)) > G(\underline{\gamma}; K_\ell/(1/\sigma^2 + 1/\sigma_c^2)) = 0. \end{aligned}$$

Further, $H''(x) > 0$ for any $x \in [1/\sigma^2 + 1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)]$. Therefore, there exists a unique $x^* \in (1/\sigma^2 + 1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2))$ with $\sigma_\ell^2 < \infty$ and $\sigma_{\epsilon\ell}^2 < \infty$: When γ is in the intermediate range, the government allocates its attention budget to both dimensions.

Furthermore, if $\gamma \in (\underline{\gamma}, \bar{\gamma})$, we have

$$H'(x^*) = 0 \Leftrightarrow x^* = K_\ell^{1/2}(1-\gamma)/\gamma.$$

Clearly, x^* strictly decreases with γ for $\gamma \in (\underline{\gamma}, \bar{\gamma})$, or equivalently, σ_ℓ^2 strictly increases with γ. Thus, we have obtained the desired conclusion. □

A.8 Proof of Lemma 11

Proof. Following a similar derivation of Equation 13 as in the proof of Lemma 10, we have

$$\begin{aligned} E(a_\ell - \theta)^2 &= E\left\{E[((1-\gamma)(E(\theta|\theta_\ell, s_c', s_c'') - \theta) + \gamma(E(\theta_c|\theta_\ell, s_c', s_c'') - \theta))^2 | \theta_\ell, s_c', s_c'']\right\} \\ &= (1-\gamma)^2 Var(\theta|\theta_\ell, s_c', s_c'') + 2\gamma(1-\gamma)Var(\theta|\theta_\ell, s_c', s_c'') \end{aligned}$$

$$
\begin{aligned}
& +\gamma^2 E\{E[(E(\theta_c|\theta_\ell, s_c', s_c'') - \theta_c + \theta_c - \theta)^2|\theta_\ell, s_c', s_c'']\} \\
= \; & (1-\gamma^2)Var(\theta|\theta_\ell, s_c', s_c'') + \gamma^2 Var(\theta_c|\theta_\ell, s_c', s_c'') + \gamma^2\sigma_c^2 - 2\gamma^2\sigma_c^2 \frac{\frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}} \\
= \; & \frac{1}{K_\ell - 1/\sigma_c^4}\left[\frac{(1-\gamma^2)K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2}} + \gamma^2\left(\frac{1}{\sigma_c^2} - \frac{1}{\sigma^2} - \frac{1}{\sigma_\ell^2}\right)\right] + \gamma^2\sigma_c^2 \\
= \; & \frac{1}{K_\ell - 1/\sigma_c^4}\left[\frac{(1-\gamma^2)K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2}} + K_\ell\gamma^2\sigma_c^2 - \gamma^2\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2}\right)\right]. \qquad (14)
\end{aligned}
$$

where the second to last equality follows from the fact that Constraint 4 is binding. Then the optimization problem of the central government can be rewritten as

$$
\min_{\sigma_c^2} E(a_\ell - \theta)^2 = \frac{1}{K_\ell - 1/\sigma_c^4}\left[\frac{(1-\gamma^2)K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2}} + K_\ell\gamma^2\sigma_c^2 - \gamma^2\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2}\right)\right] \equiv F(\sigma_c^2)
$$

subject to Constraint 3, where it should be emphasized that both σ_ℓ^2 and K_ℓ are functions of σ_c^2.

According to Lemma 10, $\underline{\gamma}$ and $\bar{\gamma}$ are, by construction, continuous functions of σ_c^2. For any given σ_c^2, we can divide the $[0,1]$ interval for γ into three regions: $[0,\underline{\gamma})$, $(\underline{\gamma},\bar{\gamma})$, and $(\bar{\gamma},1]$. Since $\underline{\gamma}$ and $\bar{\gamma}$ are continuous in σ_c^2, for a given γ that is in any of three regions, a small change of σ_c^2 will not change the region that the given γ belongs to.

We arbitrarily pick a σ_c^2 subject to Constraint 3 and consider four possible cases: (1) $\gamma < \underline{\gamma}$; (2) $\underline{\gamma} < \gamma < \bar{\gamma}$; (3) $\gamma > \bar{\gamma}$; (4) $\gamma = \bar{\gamma}$ or $\gamma = \underline{\gamma}$.

Case (1): $\gamma < \underline{\gamma}$.

According to Lemma 10, we have $\sigma_{\epsilon\ell}^2 = \infty$, which implies that Constraint 4 can be rewritten as

$$
\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2} = K_\ell\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2}\right)^{-1}.
$$

The objective function of the central government can then be written as

$$
\begin{aligned}
F(\sigma_c^2) &= \frac{1}{K_\ell - 1/\sigma_c^4}\left[(1-\gamma^2)\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}\right) + K_\ell\gamma^2\sigma_c^2 - \gamma^2\left(K_\ell\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}\right)^{-1} - \frac{1}{\sigma_c^2}\right)\right] \\
&= \frac{1}{K_\ell - 1/\sigma_c^4}\left[(1-\gamma^2)\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2} + \frac{\frac{\gamma^2 K_\ell \sigma_c^2}{\sigma_\epsilon^2}}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}}\right] \\
&= \frac{(1-\gamma^2)\sigma^2\sigma_c^2 + \sigma^2\sigma_\epsilon^2 + \frac{\gamma^2(2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)\sigma_c^4 + \sigma^2\sigma_\epsilon^2\sigma_c^2)}{\sigma_c^2+\sigma_\epsilon^2}}{2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)} \\
&= \frac{[(1-\gamma^2)\sigma^2\sigma_c^2 + \sigma^2\sigma_\epsilon^2](\sigma_c^2+\sigma_\epsilon^2) + \gamma^2(2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)\sigma_c^4 + \sigma^2\sigma_\epsilon^2\sigma_c^2)}{2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2)}
\end{aligned}
$$

where the second to last inequality follows from the definition of K_ℓ ($K_\ell = 2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)/(\sigma_c^2\sigma^2\sigma_\epsilon^2) + 1/\sigma_c^4$). Since $\gamma < \underline{\gamma}$ continues to hold for a small change of σ_c^2, the objective function is differentiable and its first derivative is given by

$$F'(\sigma_c^2) = \frac{G_1(\sigma_c^2)}{2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)^2(\sigma_c^2+\sigma_\epsilon^2)^2}$$

with the numerator $G_1(\sigma_c^2)$ given by

$$\begin{aligned}
G_1(\sigma_c^2) &= \left\{2(1-\gamma^2)\sigma^2\sigma_c^2+\sigma^2\sigma_\epsilon^2+(1-\gamma^2)\sigma^2\sigma_\epsilon^2+\gamma^2[2^{2\kappa_\ell}(3\sigma_c^4+2(\sigma^2+\sigma_\epsilon^2)\sigma_c^2)+\sigma^2\sigma_\epsilon^2]\right\} \\
&\quad \cdot(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2) \\
&\quad -\left\{[(1-\gamma^2)\sigma^2\sigma_c^2+\sigma^2\sigma_\epsilon^2](\sigma_c^2+\sigma_\epsilon^2)+\gamma^2(2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)\sigma_c^4+\sigma^2\sigma_\epsilon^2\sigma_c^2)\right\}(2\sigma_c^2+\sigma^2+2\sigma_\epsilon^2) \\
&= 2(1-\gamma^2+\gamma^2 2^{2\kappa_\ell})\sigma^2\sigma_c^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2)+2\sigma^2\sigma_\epsilon^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2) \\
&\quad +3\gamma^2 2^{2\kappa_\ell}\sigma_c^4(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2)+2\gamma^2 2^{2\kappa_\ell}\sigma_\epsilon^2\sigma_c^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2) \\
&\quad -[(1-\gamma^2)\sigma^2\sigma_c^2+\sigma^2\sigma_\epsilon^2](\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2)-[(1-\gamma^2)\sigma^2\sigma_c^2+\sigma^2\sigma_\epsilon^2](\sigma_c^2+\sigma_\epsilon^2)^2 \\
&\quad -\gamma^2 2^{2\kappa_\ell}\sigma_c^4(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(2\sigma_c^2+\sigma^2+2\sigma_\epsilon^2)-\gamma^2\sigma^2\sigma_\epsilon^2\sigma_c^2(2\sigma_c^2+\sigma^2+2\sigma_\epsilon^2) \\
&= (2\gamma^2 2^{2\kappa_\ell}+1-\gamma^2)\sigma^2\sigma_c^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2)+\sigma^2\sigma_\epsilon^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2) \\
&\quad +\gamma^2 2^{2\kappa_\ell}\sigma_c^4(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2-\sigma^2+\sigma_\epsilon^2)+2\gamma^2 2^{2\kappa_\ell}\sigma_\epsilon^2\sigma_c^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2) \\
&\quad -\sigma^2[(1-\gamma^2)\sigma_c^2+\sigma_\epsilon^2](\sigma_c^2+\sigma_\epsilon^2)^2-\gamma^2\sigma_\epsilon^2\sigma^2\sigma_c^2(2\sigma_c^2+\sigma^2+2\sigma_\epsilon^2) \\
&= 2\gamma^2 2^{2\kappa_\ell}\sigma^2\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2)^2+(2\gamma^2 2^{2\kappa_\ell}+1-\gamma^2)\sigma^4\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2)+(1-\gamma^2)\sigma^4\sigma_\epsilon^2\sigma_c^2+\sigma^4\sigma_\epsilon^4 \\
&\quad +\gamma^2 2^{2\kappa_\ell}\sigma_c^4(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)(\sigma_c^2+\sigma_\epsilon^2)+2\gamma^2 2^{2\kappa_\ell}\sigma_\epsilon^2\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2)^2+2\gamma^2(2^{2\kappa_\ell}-1)\sigma_\epsilon^2\sigma^2\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2) \\
&\quad -\gamma^2 2^{2\kappa_\ell}\sigma_c^4\sigma^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2) \\
&> 2\gamma^2 2^{2\kappa_\ell}\sigma^2\sigma_c^4(\sigma_c^2+\sigma_\epsilon^2)+(2\gamma^2 2^{2\kappa_\ell}+1-\gamma^2)\sigma^4\sigma_c^4+(1-\gamma^2)\sigma^4\sigma_\epsilon^2\sigma_c^2+\sigma^4\sigma_\epsilon^4 \\
&\quad +2\gamma^2 2^{2\kappa_\ell}\sigma_\epsilon^2\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2)^2+2\gamma^2(2^{2\kappa_\ell}-1)\sigma_\epsilon^2\sigma^2\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2)-\gamma^2 2^{2\kappa_\ell}\sigma_c^4\sigma^2(\sigma_c^2+\sigma^2+\sigma_\epsilon^2) \\
&= \gamma^2 2^{2\kappa_\ell}\sigma^2\sigma_c^4(\sigma_c^2+\sigma_\epsilon^2)+(\gamma^2 2^{2\kappa_\ell}+1-\gamma^2)\sigma^4\sigma_c^4+(1-\gamma^2)\sigma^4\sigma_\epsilon^2\sigma_c^2+\sigma^4\sigma_\epsilon^4 \\
&\quad +2\gamma^2 2^{2\kappa_\ell}\sigma_\epsilon^2\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2)^2+2\gamma^2(2^{2\kappa_\ell}-1)\sigma_\epsilon^2\sigma^2\sigma_c^2(\sigma_c^2+\sigma_\epsilon^2)>0,
\end{aligned}$$

where the last inequality follows from $\kappa_\ell > 0$ and $\gamma \in [0,1]$. Since $G_1(\sigma_c^2) > 0$, $F'(\sigma_c^2) > 0$.

Case (2): $\underline{\gamma} < \gamma < \bar{\gamma}$.

According to Lemma 10, we obtain the first order condition for the local government,

$$\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2} = \frac{1-\gamma}{\gamma}K_\ell^{1/2}.$$

Plugging in the expression of $1/\sigma_\ell^2$, the objective function F can be simplified as

$$\begin{aligned}
F(\sigma_c^2) &= \frac{\gamma(1+\gamma)K_\ell^{1/2}+K_\ell\gamma^2\sigma_c^2-\gamma(1-\gamma)K_\ell^{1/2}+\gamma^2/\sigma_c^2}{K_\ell-1/\sigma_c^4} \\
&= \gamma^2\frac{\sigma_c^2K_\ell+2K_\ell^{1/2}+1/\sigma_c^2}{K_\ell-1/\sigma_c^4} = \gamma^2\left(\sigma_c^2+\frac{2K_\ell^{1/2}+2/\sigma_c^2}{K_\ell-1/\sigma_c^4}\right) \\
&= \gamma^2\left(\sigma_c^2+\frac{2}{K_\ell^{1/2}-1/\sigma_c^2}\right)
\end{aligned}$$

Again, since γ is still in the middle range for a small change of σ_c^2, the objective function is differentiable and its first derivative is given by

$$
\begin{aligned}
F'(\sigma_c^2) &= \gamma^2\left(1-\frac{K_\ell^{-1/2}dK_\ell/d(\sigma_c^2)+2/\sigma_c^4}{(K_\ell^{1/2}-1/\sigma_c^2)^2}\right)\\
&= \gamma^2\left(\frac{K_\ell-2K_\ell^{1/2}/\sigma_c^2-K_\ell^{-1/2}dK_\ell/d(\sigma_c^2)-1/\sigma_c^4}{(K_\ell^{1/2}-1/\sigma_c^2)^2}\right)\\
&= \gamma^2\left(\frac{2^{2\kappa_\ell}\frac{\sigma^2+\sigma_c^2+\sigma_\epsilon^2}{\sigma^2\sigma_c^2\sigma_\epsilon^2}-\frac{2K_\ell^{1/2}}{\sigma_c^2}+\frac{2^{2\kappa_\ell}(\sigma^2+\sigma_\epsilon^2)\sigma_c^2+2\sigma^2\sigma_\epsilon^2}{K_\ell^{1/2}\sigma_c^6\sigma^2\sigma_\epsilon^2}}{(K_\ell^{1/2}-1/\sigma_c^2)^2}\right)\\
&= \frac{\gamma^2}{K_\ell^{1/2}(K_\ell^{1/2}-1/\sigma_c^2)^2}\left(\frac{K_\ell^{1/2}2^{2\kappa_\ell}(\sigma^2+\sigma_c^2+\sigma_\epsilon^2)-2\sigma^2\sigma_\epsilon^2K_\ell+2^{2\kappa_\ell}(\sigma^2+\sigma_\epsilon^2)/\sigma_c^2+2\sigma^2\sigma_\epsilon^2/\sigma_c^4}{\sigma^2\sigma_c^2\sigma_\epsilon}\right)\\
&= \frac{2^{2\kappa_\ell}\gamma^2}{K_\ell^{1/2}(K_\ell^{1/2}-1/\sigma_c^2)^2}\left(\frac{K_\ell^{1/2}(\sigma^2+\sigma_c^2+\sigma_\epsilon^2)-(\sigma^2/\sigma_c^2+\sigma_\epsilon^2/\sigma_c^2+2)}{\sigma^2\sigma_c^2\sigma_\epsilon^2}\right)\\
&> \frac{2^{2\kappa_\ell}\gamma^2}{K_\ell^{1/2}(K_\ell^{1/2}-1/\sigma_c^2)^2}\left(\frac{(1/\sigma_\epsilon^2+1/\sigma_c^2)^{1/2}(1/\sigma^2+1/\sigma_c^2)^{1/2}(\sigma^2+\sigma_c^2+\sigma_\epsilon^2)-(\sigma^2/\sigma_c^2+\sigma_\epsilon^2/\sigma_c^2+2)}{\sigma^2\sigma_c^2\sigma_\epsilon^2}\right),
\end{aligned}
$$

where the last inequality follows from $\kappa_\ell > 0$. There are two possibilities. If $1/\sigma_\epsilon^2 \geq 1/\sigma^2$, then

$$F'(\sigma_c^2) \geq \frac{2^{2\kappa_\ell}\gamma^2}{K_\ell^{1/2}(K_\ell^{1/2}-1/\sigma_c^2)^2}\left(\frac{(1/\sigma^2+1/\sigma_c^2)(\sigma^2+\sigma_c^2+\sigma_\epsilon^2)-(\sigma^2/\sigma_c^2+\sigma_\epsilon^2/\sigma_c^2+2)}{\sigma^2\sigma_c^2\sigma_\epsilon^2}\right) > 0.$$

If $1/\sigma_\epsilon^2 < 1/\sigma^2$, then

$$F'(\sigma_c^2) > \frac{2^{2\kappa_\ell}\gamma^2}{K_\ell^{1/2}(K_\ell^{1/2}-1/\sigma_c^2)^2}\left(\frac{(1/\sigma_\epsilon^2+1/\sigma_c^2)(\sigma^2+\sigma_c^2+\sigma_\epsilon^2)-(\sigma^2/\sigma_c^2+\sigma_\epsilon^2/\sigma_c^2+2)}{\sigma^2\sigma_c^2\sigma_\epsilon^2}\right) > 0.$$

Hence, we must have $F'(\sigma_c^2) > 0$.

Case (3): $\gamma > \bar{\gamma}$.

In this case, according to Lemma 10, $\sigma_\ell^2 = \infty$. Then we have

$$
\begin{aligned}
F(\sigma_c^2) &= \frac{1}{K_\ell-1/\sigma_c^4}\left[\frac{(1-\gamma^2)K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}}+K_\ell\gamma^2\sigma_c^2-\frac{\gamma^2}{\sigma^2}\right]\\
&= \frac{K_\ell+K_\ell\gamma^2\sigma_c^2/\sigma^2-\gamma^2(1/\sigma^2+1/\sigma_c^2)/\sigma^2}{(K_\ell-1/\sigma_c^4)(1/\sigma^2+1/\sigma_c^2)}
\end{aligned}
$$

Again, since $\gamma > \bar{\gamma}$ continues to hold for a small change of σ_c^2, the objective function is differentiable and its first derivative is given by

$$F'(\sigma_c^2) = \frac{G_2(\sigma_c^2)}{\left(K_\ell-\frac{1}{\sigma_c^4}\right)^2\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)^2}$$

with the numerator $G_2(\sigma_c^2)$ given by

$$G_2(\sigma_c^2) = \left[\frac{dK_\ell}{d(\sigma_c^2)}\left(1+\frac{\gamma^2\sigma_c^2}{\sigma^2}\right)+\frac{K_\ell\gamma^2}{\sigma^2}+\frac{\gamma^2}{\sigma_c^4\sigma^2}\right]\left(K_\ell-\frac{1}{\sigma_c^4}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)$$

$$-\left[K_\ell\left(1+\frac{\gamma^2\sigma_c^2}{\sigma^2}\right)-\frac{\gamma^2}{\sigma^2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\right]\left[\frac{dK_\ell}{d(\sigma_c^2)}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)+\frac{2}{\sigma_c^6\sigma^2}-\frac{K_\ell}{\sigma_c^4}+\frac{3}{\sigma_c^8}\right]$$

$$= \left(\gamma^2\left(\frac{1}{\sigma^4}+\frac{2}{\sigma^2\sigma_c^2}\right)+\frac{1}{\sigma_c^4}\right)K_\ell^2+\left[\frac{\gamma^2}{\sigma_c^4\sigma^2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)-\frac{\gamma^2}{\sigma^2\sigma_c^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\right.$$

$$\left.-\left(\frac{2}{\sigma_c^6\sigma^2}+\frac{3}{\sigma_c^8}\right)\left(1+\frac{\gamma^2\sigma_c^2}{\sigma^2}\right)-\frac{\gamma^2}{\sigma_c^4\sigma^2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\right]K_\ell+\left[-\frac{1}{\sigma_c^4}\left(1+\frac{\gamma^2\sigma_c^2}{\sigma^2}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\right.$$

$$\left.+\frac{\gamma^2}{\sigma^2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)^2\right]\frac{dK_\ell}{d(\sigma_c^2)}-\frac{\gamma^2}{\sigma_c^8\sigma^2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)+\frac{\gamma^2}{\sigma^2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{2}{\sigma_c^6\sigma^2}+\frac{3}{\sigma_c^8}\right)$$

$$= \frac{K_\ell^2}{\sigma_c^4}-\left(\frac{2}{\sigma_c^6\sigma^2}+\frac{3}{\sigma_c^8}\right)K_\ell-\frac{1}{\sigma_c^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\frac{dK_\ell}{d(\sigma_c^2)}+\gamma^2\left\{\left(\frac{1}{\sigma^4}+\frac{2}{\sigma^2\sigma_c^2}\right)K_\ell^2-\left(\frac{3}{\sigma_c^4\sigma^4}+\frac{4}{\sigma_c^6\sigma^2}\right)K_\ell\right.$$

$$\left.+\frac{1}{\sigma^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\frac{dK_\ell}{d(\sigma_c^2)}+\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{2}{\sigma_c^6\sigma^4}+\frac{2}{\sigma_c^8\sigma^2}\right)\right\}$$

$$= \frac{1}{\sigma_c^4}\left\{\left[\frac{2^{2\kappa_\ell}}{\sigma_c^2\sigma_\epsilon^2\sigma^2}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)-\frac{2}{\sigma_c^2}\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma^2}\right)\right]\left[\frac{2^{2\kappa_\ell}}{\sigma_c^2\sigma_\epsilon^2\sigma^2}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)+\frac{1}{\sigma_c^4}\right]\right.$$

$$\left.+\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left[\frac{2^{2\kappa_\ell}}{\sigma_c^4}\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma^2}\right)+\frac{2}{\sigma_c^6}\right]\right\}$$

$$+\gamma^2\left\{\left(\frac{2^{2\kappa_\ell}}{\sigma_c^2\sigma_\epsilon^2\sigma^2}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)+\frac{1}{\sigma_c^4}\right)\left[\left(\frac{1}{\sigma^4}+\frac{2}{\sigma^2\sigma_c^2}\right)\frac{2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)}{\sigma_c^2\sigma_\epsilon^2\sigma^2}-\frac{2}{\sigma_c^4\sigma^4}-\frac{2}{\sigma_c^6\sigma^2}\right]\right.$$

$$\left.-\frac{1}{\sigma^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{2^{2\kappa_\ell}}{\sigma_c^4}\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma^2}\right)+\frac{2}{\sigma_c^6}\right)+\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{2}{\sigma_c^6\sigma^4}+\frac{2}{\sigma_c^8\sigma^2}\right)\right\}$$

$$= \frac{1}{\sigma_c^4}\left\{\frac{2^{4\kappa_\ell}}{\sigma_c^4\sigma_\epsilon^4\sigma^4}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)^2-\frac{1}{\sigma_c^2}\left(\frac{1}{\sigma_c^2}+\frac{2}{\sigma^2}\right)\cdot\frac{2^{2\kappa_\ell}}{\sigma_c^2\sigma_\epsilon^2\sigma^2}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)\right.$$

$$\left.+\frac{2^{2\kappa_\ell}}{\sigma_c^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma^2}\right)\right\}$$

$$+\gamma^2\left\{\left(\frac{2^{2\kappa_\ell}}{\sigma_c^2\sigma_\epsilon^2\sigma^2}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)+\frac{1}{\sigma_c^4}\right)\left(\frac{1}{\sigma^4}+\frac{2}{\sigma^2\sigma_c^2}\right)\frac{2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)}{\sigma_c^2\sigma_\epsilon^2\sigma^2}\right.$$

$$\left.-\frac{2^{2\kappa_\ell}}{\sigma_c^4\sigma^4}\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma^2}\right)\left(\frac{3}{\sigma_\epsilon^2}+\frac{2}{\sigma_c^2}+\frac{1}{\sigma^2}+\frac{2\sigma^2}{\sigma_c^2\sigma_\epsilon^2}\right)\right\}$$

$$= \frac{2^{2\kappa_\ell}}{\sigma_c^4}\left\{\frac{2^{2\kappa_\ell}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)^2}{\sigma_c^4\sigma_\epsilon^4\sigma^4}-\frac{2\sigma_c^2+2\sigma^2+\sigma_\epsilon^2}{\sigma_c^4\sigma^4\sigma_\epsilon^2}\right\}$$

$$+\gamma^2 2^{2\kappa_\ell}\left\{\left(\frac{2^{2\kappa_\ell}}{\sigma_c^2\sigma_\epsilon^2\sigma^2}(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)+\frac{1}{\sigma_c^4}\right)\left(\frac{1}{\sigma^4}+\frac{2}{\sigma^2\sigma_c^2}\right)\frac{\sigma_c^2+\sigma^2+\sigma_\epsilon^2}{\sigma_c^2\sigma_\epsilon^2\sigma^2}\right.$$

$$\left.-\frac{1}{\sigma_c^4\sigma^4}\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma^2}\right)\left(\frac{3}{\sigma_\epsilon^2}+\frac{2}{\sigma_c^2}+\frac{1}{\sigma^2}+\frac{2\sigma^2}{\sigma_c^2\sigma_\epsilon^2}\right)\right\}$$

$$> \frac{2^{2\kappa_\ell}}{\sigma_c^4}\left\{\frac{(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)^2}{\sigma_c^4\sigma_\epsilon^4\sigma^4}-\frac{2\sigma_c^2+2\sigma^2+\sigma_\epsilon^2}{\sigma_c^4\sigma^4\sigma_\epsilon^2}\right\}+\gamma^2 2^{2\kappa_\ell}\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma^2}\right)$$

$$\cdot\left\{\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)\left(\frac{1}{\sigma^4}+\frac{2}{\sigma^2\sigma_c^2}\right)\frac{\sigma_c^2+\sigma^2+\sigma_\epsilon^2}{\sigma_c^2\sigma_\epsilon^2\sigma^2}-\frac{1}{\sigma_c^4\sigma^4}\left(\frac{3}{\sigma_\epsilon^2}+\frac{2}{\sigma_c^2}+\frac{1}{\sigma^2}+\frac{2\sigma^2}{\sigma_c^2\sigma_\epsilon^2}\right)\right\}$$

$$> \frac{2^{2\kappa_\ell}}{\sigma_c^4}\left\{\frac{(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)^2}{\sigma_c^4\sigma_\epsilon^4\sigma^4}-\frac{2\sigma_c^2+2\sigma^2+\sigma_\epsilon^2}{\sigma_c^4\sigma^4\sigma_\epsilon^2}\right\}+\gamma^2 2^{2\kappa_\ell}\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma^2}\right)$$

$$\cdot\left\{\frac{1}{\sigma_c^2}\left(\frac{1}{\sigma^4}+\frac{2}{\sigma^2\sigma_c^2}\right)\frac{\sigma_c^2+\sigma^2+\sigma_\epsilon^2}{\sigma_c^2\sigma_\epsilon^2\sigma^2}-\frac{1}{\sigma_c^4\sigma^4}\left(\frac{3}{\sigma_\epsilon^2}+\frac{2}{\sigma_c^2}+\frac{1}{\sigma^2}+\frac{2\sigma^2}{\sigma_c^2\sigma_\epsilon^2}\right)\right\}$$

$$= \frac{2^{2\kappa_\ell}}{\sigma_c^4}\frac{(\sigma_c^2+\sigma^2)^2}{\sigma_c^4\sigma^4\sigma_\epsilon^4}+\frac{\gamma^2 2^{2\kappa_\ell}}{\sigma_c^4\sigma^4}\frac{\sigma_c^2}{\sigma_\epsilon^2\sigma^2}>0$$

where the first inequality follows from $\kappa_\ell > 0$. Therefore, we have $F'(\sigma_c^2) > 0$.

Case (4): $\gamma = \underline{\gamma}$ or $\gamma = \bar{\gamma}$.

Suppose $\gamma = \underline{\gamma}$. A small change of σ_c^2 will make $\gamma < \underline{\gamma}$ or $\gamma \in (\underline{\gamma}, \bar{\gamma})$. Since whether γ ends up in Case (1) or (2) depends on the direction of the change of σ_c^2, the left or right derivatives of F at σ_c^2 may not be equal to each other. If an infinitesimal negative change of σ_c^2 leads to Case (1), we know that $F'_-(\sigma_c^2)$ is equal to $F'(\sigma_c^2)$ for Case (1) and therefore $F'_-(\sigma_c^2) > 0$. If an infinitesimal negative of σ_c^2 leads to Case (2), we know that $F'_-(\sigma_c^2)$ is equal to $F'(\sigma_c^2)$ for Case (2) and therefore $F'_-(\sigma_c^2) > 0$. The same argument applies to $F'_+(\sigma_c^2)$ and we have $F'_+(\sigma_c^2) > 0$. Similarly, we can also show that $F'_-(\sigma_c^2) > 0$ and $F'_+(\sigma_c^2) > 0$ for $\gamma = \bar{\gamma}$.

In sum, for an arbitrarily picked σ_c^2, we have shown that the objective function is strictly increasing for any $\gamma \in [0, 1]$. Then the optimal strategy for the central government is to minimize σ_c^2. Therefore, Constraint 3 must be binding and $\sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1)$. We have obtained the desired conclusion. □

A.9 Proof of Proposition 2

Proof. It directly follows from Lemma 8 that $\sigma_{\epsilon\ell}^2 = \infty$ and from Lemma 11 that $\sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1)$. The expression of σ_ℓ^2 can then be derived from the binding Constraint 4. We know the expression of $E(a_\ell - \theta)^2\Big|_{\gamma=0}$ from Remark 2. □

A.10 Proof of Proposition 3

Proof. It directly follows from Lemma 9 that $\sigma_\ell^2 = \infty$ and from Lemma 11 that $\sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1)$. The expression of $\sigma_{\epsilon\ell}^2$ can then be derived from the binding Constraint 4. We can derive the expression of $E(a_\ell - \theta)^2\Big|_{\gamma=0}$ by invoking the formula in Lemma 5 and specializing it with the expression of a_ℓ for $\gamma = 1$ in Lemma 4:[29]

$$E(a_\ell - \theta)^2\Big|_{\gamma=1} = \frac{\sigma_c^2(1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^2 + 1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2 + \sigma^2/(\sigma_c^2+\sigma^2)^2}{\left[1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2 + 1/(\sigma_c^2+\sigma^2)\right]^2} = \frac{\sigma_c^2(1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2) + \sigma^2/(\sigma_c^2+\sigma^2)}{1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2 + 1/(\sigma_c^2+\sigma^2)}.$$

□

A.11 Proof of Proposition 4

Proof. In the proof of Lemma 11, we have obtained Equation 14:

$$E(a_\ell - \theta)^2 = \frac{1}{K_\ell - 1/\sigma_c^4}\left[\frac{(1-\gamma^2)K_\ell}{\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2}} + K_\ell\gamma^2\sigma_c^2 - \gamma^2\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2}\right)\right] \equiv F(\sigma_\ell^2(\gamma), \gamma),$$

with $\sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1)$ being constant with respect to γ.

It is easy to see that $\partial F/\partial(\sigma_\ell^2) > 0$, and

$$\frac{\partial F}{\partial \gamma} = \frac{2\gamma}{K_\ell - 1/\sigma_\ell^4}\left(-\frac{K_\ell}{\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2}} + K_\ell\sigma_c^2 - \frac{1}{\sigma^2} - \frac{1}{\sigma_\ell^2}\right)$$

[29] The expression can alternatively be obtained from Equation 14.

$$
\begin{aligned}
&= \frac{2\gamma}{K_\ell - 1/\sigma_\ell^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}\right)\left(\frac{K_\ell \sigma_c^2}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2}} - 1\right) \\
&= \frac{2\gamma\sigma_c^2}{K_\ell - 1/\sigma_\ell^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}\right)
\end{aligned}
$$

where the last equality follows again from the fact that Constraint 4 is binding. Then $\partial F/\partial\gamma \geq 0$ with the equality if and only if $\gamma = 0$. Since σ_ℓ^2 is a function of γ and we know from Lemma 10 that σ_ℓ^2 weakly increases with γ, we conclude that $E(a_\ell - \theta)^2$ strictly increases with γ. □

A.12 Proof of Lemma 12

Proof. According to Propositions 1 and 2, $E(a_c-\theta)^2 > E(a_\ell-\theta)^2\Big|_{\gamma=0}$ if and only if

$$
\frac{(2^{2\kappa_\ell}-1)(\sigma^2+\sigma_\epsilon^2+\sigma_c^2)}{\sigma^2(\sigma_\epsilon^2+\sigma_c^2)} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2} > \frac{(2^{2\kappa_c}-1)(\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2)}{\sigma^2(\sigma_\epsilon^2+\sigma_\ell^2)} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2},
$$

where $\sigma_c^2 = \sigma^2/(2^{2\kappa_c}-1)$ and $\sigma_\ell^2 = \sigma^2/(2^{2\kappa_\ell}-1)$. Simplifying the expression above, we obtain

$$
\frac{\sigma^2+\sigma_\epsilon^2+\sigma_c^2}{\sigma_\ell^2(\sigma_\epsilon^2+\sigma_c^2)} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2} > \frac{\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2}{\sigma_c^2(\sigma_\epsilon^2+\sigma_\ell^2)} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}.
$$

The inequality holds if and only if $\sigma_c^2 > \sigma_\ell^2$, which follows from $\kappa_\ell > \kappa_c$. □

A.13 Proof of Lemma 13

Proof. According to Proposition 3, we know

$$
\begin{aligned}
E(a_\ell-\theta)^2\Big|_{\gamma=1} &= \frac{\sigma_c^2(1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2)^2 + 1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2+\sigma^2/(\sigma_c^2+\sigma^2)^2}{\left[1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2+1/(\sigma_c^2+\sigma^2)\right]^2} \\
&= \frac{\sigma_c^2(1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2)+\sigma^2/(\sigma_c^2+\sigma^2)}{1/\sigma_\epsilon^2+1/\sigma_{\epsilon\ell}^2+1/(\sigma_c^2+\sigma^2)} \\
&= \sigma_c^2 + \frac{\sigma^2-\sigma_c^2}{\sigma_c^2+\sigma^2}\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_{\epsilon\ell}^2}+\frac{1}{\sigma_c^2+\sigma^2}\right)^{-1}
\end{aligned}
$$

If $\sigma_c^2 \leq \sigma^2$, we have $E(a_\ell-\theta)^2\Big|_{\gamma=1} > \sigma_c^2$. If $\sigma_c^2 > \sigma^2$, then $E(a_\ell-\theta)^2\Big|_{\gamma=1}$ strictly decreases with σ_ϵ^2 and $\sigma_{\epsilon\ell}^2$. We know $\lim_{\sigma_\epsilon^2\to\infty,\sigma_{\epsilon\ell}^2\to\infty} E(a_\ell-\theta)^2\Big|_{\gamma=1} = \sigma^2$, so when $\sigma_c^2 > \sigma^2$, $E(a_\ell-\theta)^2\Big|_{\gamma=1} > \sigma^2$. Therefore, we must have $E(a_\ell-\theta)^2\Big|_{\gamma=1} > \min\{\sigma^2, \sigma_c^2\}$ where $\sigma_c^2 = \sigma^2/(2^{2\kappa_c}-1)$.

According to Proposition 1, with $\sigma_\ell^2 = \sigma^2/(2^{2\kappa_\ell}-1)$, we have

$$
E(a_c-\theta)^2 = \left(\frac{1}{\sigma^2} + \frac{\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2}{\sigma_c^2(\sigma_\epsilon^2+\sigma_\ell^2)} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}\right)^{-1} = \left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\sigma^2+\sigma_c^2}{\sigma_c^2(\sigma_\epsilon^2+\sigma_\ell^2)}\right)^{-1}.
$$

It is easy to see that $E(a_c-\theta)^2 < 1/\sigma^2$ and $E(a_c-\theta) < 1/\sigma_c^2$. Therefore, $E(a_c-\theta)^2 < \min\{\sigma^2,\sigma_c^2\} < E(a_\ell-\theta)^2\Big|_{\gamma=1}$. We have obtained the desired conclusion. □

A.14 Proof of Corollary 1

Proof. Using Equation 14 in the proof of Lemma 11, we obtain

$$\begin{aligned} E(a_\ell-\theta)^2\Big|_{\gamma=\bar\gamma} &= \frac{1}{K_\ell-1/\sigma_c^4}\left[\frac{(1-\bar\gamma^2)K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{1}{\sigma_c^2}}+K_\ell\bar\gamma^2\sigma_c^2-\bar\gamma^2\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}\right)\right] \\ &= \frac{1}{K_\ell-1/\sigma_c^4}\left[\frac{(1-\bar\gamma^2)K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}}+K_\ell\bar\gamma^2\sigma_c^2-\bar\gamma^2\left(\frac{1}{\sigma^2}\right)\right] \\ &= \frac{1}{K_\ell-1/\sigma_c^4}\frac{K_\ell+\frac{\bar\gamma^2\sigma_c^2}{\sigma^2}\left(K_\ell-\frac{1}{\sigma_c^2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\right)}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}} \\ &> \frac{K_\ell}{K_\ell-1/\sigma_c^4}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)^{-1} > \left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)^{-1} \end{aligned}$$

where $\sigma_c=\sigma^2/(2^{2\kappa_c}-1)$ (Lemma 11), the second inequality follows from the fact that $\sigma_\ell^2=\infty$ when $\gamma=\bar\gamma$ (Lemma 10), and the first inequality follows from the definition of K_ℓ. From Proposition 1, we know

$$E(a_c-\theta)^2=\left(\frac{1}{\sigma^2}+\frac{(2^{2\kappa_c}-1)(\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2)}{\sigma^2(\sigma_\epsilon^2+\sigma_\ell^2)}+\frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}\right)^{-1}$$

with $\sigma_\ell^2=\sigma^2/(2^{2\kappa_\ell}-1)$. Hence, we have

$$E(a_c-\theta)^2=\left(\frac{1}{\sigma^2}+\frac{\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2}{\sigma_c^2(\sigma_\epsilon^2+\sigma_\ell^2)}+\frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}\right)^{-1}<\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)^{-1}<E(a_\ell-\theta)^2\Big|_{\gamma=\bar\gamma}.$$

By definition, $E(a_c-\theta)^2=E(a_\ell-\theta)^2\Big|_{\gamma=\tilde\gamma}$. Then $\bar\gamma>\tilde\gamma$ directly follows from Proposition 4. □

A.15 Proof of Lemma 14

Proof. The proof directly follows from Lemmas 3 and 4 by letting $\sigma_{\epsilon\ell}^2=\sigma_{\epsilon c}^2=\infty$.

Alternatively, to derive the expressions of $E(\theta|\theta_c,s'_\ell)$ and $E(\theta|\theta_\ell,s'_c)$, we can invoke the following Bayesian updating rule with a normal prior.

Lemma 22. Let $\mu\sim\mathcal{N}(\mu_0,\sigma_0^2)$ and $x_i|\mu\sim\mathcal{N}(\mu,\sigma_i^2)$ with $i=1,2,...,n.$. Conditional on μ, x_1,

x_2,..., and x_n are independent. If μ_0, σ_0^2, and σ_i^2 are known, then[30]

$$\mu|x_1, x_2, ..., x_n \sim \mathcal{N}\left(\frac{\sum_{i=1}^n (x_i/\sigma_i^2) + \mu_0/\sigma_0^2}{\sum_{i=1}^n (1/\sigma_i^2) + 1/\sigma_0^2}, \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} + \frac{1}{\sigma_0^2}\right)^{-1}\right).$$

Since $\theta \sim \mathcal{N}(0, \sigma^2)$, $\theta_c|\theta \sim \mathcal{N}(\theta, \sigma_c^2)$, $s'_\ell|\theta \sim \mathcal{N}(\theta, \sigma_\ell^2 + \sigma_\epsilon^2)$, and θ_c and s'_ℓ are conditionally independent, applying Lemma 22, we obtain

$$\theta|\theta_c, s'_\ell \sim \mathcal{N}\left(\frac{\theta_c/\sigma_c^2 + s'_\ell/(\sigma_\ell^2 + \sigma_\epsilon^2)}{1/\sigma_c^2 + 1/(\sigma_\ell^2 + \sigma_\epsilon^2) + 1/\sigma^2}, \left(\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2 + \sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)^{-1}\right).$$

Similarly, given that $\theta_\ell|\theta \sim \mathcal{N}(\theta, \sigma_\ell^2)$, $s'_c|\theta \sim \mathcal{N}(\theta, \sigma_c^2 + \sigma_\epsilon^2)$, and θ_ℓ and s'_c are conditionally independent, we have

$$\theta|\theta_\ell, s'_c \sim \mathcal{N}\left(\frac{\theta_\ell/\sigma_\ell^2 + s'_c/(\sigma_c^2 + \sigma_\epsilon^2)}{1/\sigma_\ell^2 + 1/(\sigma_c^2 + \sigma_\epsilon^2) + 1/\sigma^2}, \left(\frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + \sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)^{-1}\right).$$

□

A.16 Proof of Lemma 15

Proof. The expressions are obtained from taking the expressions of a_c and a_ℓ from Lemma 14 and applying the results in Lemma 5. Under the decentralized regime, when $\gamma = 1$, we have

$$E(a_\ell - \theta)^2\Big|_{\gamma=1} = \frac{\frac{\sigma_\ell^2}{(\sigma_c^2+\sigma_\ell^2+\sigma_c^2\sigma_\ell^2/\sigma^2)^2} + \frac{\sigma_c^2+\sigma_\epsilon^2}{\sigma_\epsilon^4} + \frac{\sigma_\ell^4/\sigma^2}{(\sigma_c^2+\sigma_\ell^2+\sigma_c^2\sigma_\ell^2/\sigma^2)^2}}{\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}\right)^2}$$

$$= \frac{\frac{\sigma_c^2}{\sigma_\epsilon^4} + \frac{1}{\sigma_\epsilon^2} + \frac{(1/\sigma^2+1/\sigma_\ell)^{-1}}{(\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1})^2}}{\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}\right)^2} = \frac{\frac{\sigma_c^2}{\sigma_\epsilon^2} + \frac{(1/\sigma^2+1/\sigma_\ell^2)^{-1}}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell)^{-1}}}.$$

□

[30] To see this result, writing the probability density function explicitly, we have

$$f(x_1, x_2, ..., x_n, \mu) = f(x_1, x_2, ..., x_n|\mu)f(\mu) \propto \exp\left\{-\sum_{i=1}^n (x_i - \mu)^2/2\sigma_i^2 - (\mu - \mu_0)^2/2\sigma_0^2\right\}.$$

Given $f(\mu|x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_n, \mu)/\int f(x_1, x_2, ..., x_n, \mu)d\mu$, ignoring the normalizing constant, then

$$f(\mu|x_1, x_2, ..., x_n) \propto f(x_1, x_2, ..., x_n, \mu) \propto \exp\left\{-\frac{\left(\mu - \frac{\sum_{i=1}^n (x_i/\sigma_i^2)+\mu_0/\sigma_0^2}{\sum_{i=1}^n (1/\sigma_i^2)+1/\sigma_0^2}\right)^2}{2\left(\sum_{i=1}^n (1/\sigma_i^2) + 1/\sigma_0^2\right)^{-1}}\right\},$$

which implies the posterior distribution in the lemma. Moreover, if $\sigma_i^2 = \sigma^2$ for $i = 1, 2, ..., n$, then we obtain the familiar posterior distribution, $\mu|x_1, x_2, ..., x_n \sim \mathcal{N}(\frac{n\bar{x}/\sigma^2+\mu_0/\sigma_0^2}{n/\sigma^2+1/\sigma_0^2}, (\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2})^{-1})$. with $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$.

A.17 Proof of Lemma 16

Proof. Since we assume $\sigma_\epsilon^2 > 0$ and $\sigma_c^2 > \sigma_\ell^2$, we have

$$\frac{1}{\sigma_c^2\sigma_\ell^2} > \frac{1}{(\sigma_c^2+\sigma_\epsilon^2)(\sigma_\ell^2+\sigma_\epsilon^2)} \Leftrightarrow \frac{1}{\sigma_\ell^2} - \frac{1}{\sigma_c^2} = \frac{\sigma_c^2-\sigma_\ell^2}{\sigma_c^2\sigma_\ell^2} > \frac{\sigma_c^2-\sigma_\ell^2}{(\sigma_c^2+\sigma_\epsilon^2)(\sigma_\ell^2+\sigma_\epsilon^2)} = \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} - \frac{1}{\sigma_c^2+\sigma_\epsilon^2}.$$

Therefore, we have

$$\left(\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)^{-1} > \left(\frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)^{-1}.$$

The desired conclusion directly follows from Lemma 15. □

A.18 Proof of Lemma 17

Proof. According to Lemma 15, we have

$$\begin{aligned} E(a_\ell-\theta)^2\Big|_{\gamma=1} - E(a_c-\theta)^2 &= \frac{\frac{\sigma_c^2}{\sigma_\epsilon^2} + \frac{(1/\sigma^2+1/\sigma_\ell^2)^{-1}}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}} - \left(\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)^{-1} \\ &= \frac{F(\sigma_\epsilon^2)}{\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}\right)\left(\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right)}, \end{aligned}$$

with the numerator given by

$$\begin{aligned} F(\sigma_\epsilon^2) &= \left(\frac{\sigma_c^2}{\sigma_\epsilon^2} + \frac{(1/\sigma^2+1/\sigma_\ell^2)^{-1}}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}\right)\left(\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right) - \left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}\right) \\ &= \frac{\sigma_c^2}{\sigma_\epsilon^2}\left(\frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} + \frac{1}{\sigma^2}\right) + \frac{1/\sigma_c^2 - 1/\sigma_\ell^2 + 1/(\sigma_\ell^2+\sigma_\epsilon^2)}{\sigma_c^2(1/\sigma^2+1/\sigma_\ell^2)+1}. \end{aligned}$$

We have $F'(\sigma_\epsilon^2) < 0$, $\lim_{\sigma_\epsilon^2\to 0} F(\sigma_\epsilon^2) > 0$ and

$$\lim_{\sigma_\epsilon^2\to\infty} F(\sigma_\epsilon^2) = \frac{1/\sigma_c^2 - 1/\sigma_\ell^2}{\sigma_c^2(1/\sigma^2+1/\sigma_\ell^2)+1} < 0,$$

where the inequality follows from $\sigma_c^2 > \sigma_\ell^2$. Applying the intermediate value theorem, there must exist a unique $\bar{\sigma}_\epsilon^2 > 0$ such that $F(\bar{\sigma}_\epsilon^2) = 0$, or equivalently, $E(a_\ell-\theta)^2\Big|_{\gamma=1} = E(a_c-\theta)^2$. Given the monotonicity of F, we know that $F(\sigma_\epsilon^2) > 0$, or equivalently, $E(a_\ell-\theta)^2\Big|_{\gamma=1} > E(a_c-\theta)^2$ if and only if $\sigma_\epsilon^2 < \bar{\sigma}_\epsilon^2$. □

A.19 Proof of Proposition 5

Proof. According to Lemma 15, we have

$$E(a_\ell-\theta)^2\Big|_{\gamma=1} = \frac{\frac{\sigma_c^2}{\sigma_\epsilon^2} + \frac{(1/\sigma^2+1/\sigma_\ell)^{-1}}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell)^{-1}}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell)^{-1}}} = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2}\right)^{-1} + \frac{\sigma_c^2 - (1/\sigma^2+1/\sigma_\ell^2)^{-1}}{1 + \frac{\sigma_\epsilon^2}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}}.$$

Since $\sigma_c^2 > \sigma_\ell^2$, $\sigma_c^2 > (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}$, which suggests that $E(a_\ell - \theta)\Big|_{\gamma=1}$ strictly decreases with σ_ϵ^2. $\square$

A.20 Proof of Corollary 2

Proof. According to Lemma 14, $a_\ell = (1-\gamma)E(\theta|\theta_\ell, s_c') + \gamma E(\theta_c|\theta_\ell, s_c')$. Then we have

$$\begin{aligned} E(a_\ell - \theta)^2 &= E\left([E(\theta|\theta_\ell, s_c') - \theta] + \gamma[E(\theta_c|\theta_\ell, s_c') - E(\theta|\theta_\ell, s_c')]\right)^2 \\ &= E[E(\theta|\theta_\ell, s_c') - \theta]^2 + \gamma^2 E[E(\theta_c|\theta_\ell, s_c') - E(\theta|\theta_\ell, s_c')]^2 \\ &\quad +2\gamma E\left([E(\theta|\theta_\ell, s_c') - \theta]\cdot[E(\theta_c|\theta_\ell, s_c') - E(\theta|\theta_\ell, s_c')]\right) \\ &= Var(\theta|\theta_\ell, s_c') + \gamma^2 E((\theta_c - \theta)|\theta_\ell, s_c')^2 \end{aligned}$$

Since $E((\theta_c - \theta)|\theta_\ell, s_c')^2 > 0$, we must have $\partial E(a_\ell - \theta)^2/\partial(\gamma^2) > 0$. $\square$

A.21 Proof of Lemma 18

Proof. Let $x \equiv \frac{\sigma^2}{(2^{2\kappa}-1)\sigma_\epsilon^2} > 0$ and $k \equiv 2^{2\kappa} > 1$, we can write the right hand side of Condition 12 as

$$F(k,x) \equiv \frac{k^2\left(\frac{1}{k-1} + \frac{kx}{k-1}\right)^2}{k\left(\frac{1}{k-1} + \frac{kx}{k-1}\right) + 1} = \frac{k^2(kx+1)^2}{(k-1)(k(kx+1)+k-1)}.$$

It is easy to see that $F(k,x)$ strictly increases with x for $k > 1$. So we must have

$$F(k,x) > F(k,0) = \frac{k^2}{(k-1)(2k-1)}.$$

It is to show that $F(k,0)$ attains its minimum on $(1,\infty)$ when $k \to \infty$, which implies $F(k,0) > 1/2$. Therefore, for the regularity condition 12 to hold, it suffices to have

$$\lambda(1/\lambda - 1)^2 \leq 1/2.$$

Solving the inequality with the constraint that $\lambda \in (0,1)$, we then obtain $\lambda \geq 1/2$. $\square$

A.22 Proof of Lemma 19

Proof. Following the proof of Lemma 10, we can write the decision problem of the local government as

$$\min_{\sigma_\ell^2, \sigma_{\epsilon\ell}^2} F(\sigma_\ell^2, \sigma_{\epsilon\ell}^2),$$

subject to Constraint 10, with $F(\sigma_\ell^2, \sigma_{\epsilon\ell}^2)$ given by

$$\begin{aligned} F(\sigma_\ell^2, \sigma_{\epsilon\ell}^2) &= \frac{(1-\gamma)^2(1+2\gamma)}{\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2} + \frac{1}{\sigma_c^2 + (1/\sigma_\epsilon^2 + 1/\sigma_{\epsilon\ell}^2)^{-1}}} + \frac{\gamma^2(3-2\gamma)}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}} + (1-\gamma)\gamma\sigma_c^2 \\ &\quad - \frac{2(1-\gamma)\gamma^2\sigma_c^2}{\left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2}\right)\left(\sigma_c^2 + \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2}\right)^{-1}\right) + 1} - \frac{2(1-\gamma)^2\gamma\sigma_c^2}{\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2}\right)\left(\sigma_c^2 + \left(\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2}\right)^{-1}\right) + 1} \end{aligned}$$

We know Constraint 10 must be binding, so we can rewrite the objective function as

$$
\begin{aligned}
F(\sigma_\ell^2, \sigma_{\epsilon\ell}^2) &= \frac{(1-\gamma)^2(1+2\gamma)}{\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}+\frac{\frac{K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}}-\frac{1}{\sigma_c^2}}{\sigma_c^2\left(\frac{K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}}-\frac{1}{\sigma_c^2}\right)+1}} + \frac{\gamma^2(3-2\gamma)}{\frac{K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}}-\frac{1}{\sigma_c^2}+\frac{1}{\sigma_c^2+(1/\sigma^2+1/\sigma_\ell^2)^{-1}}} + (1-\gamma)\gamma\sigma_c^2 \\
&\quad - \frac{2(1-\gamma)\gamma^2\sigma_c^2}{\left(\frac{K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}}-\frac{1}{\sigma_c^2}\right)\left(\sigma_c^2+\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}\right)^{-1}\right)+1} \\
&\quad - \frac{2(1-\gamma)^2\gamma\sigma_c^2\left(\frac{K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}}-\frac{1}{\sigma_c^2}\right)}{\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}\right)\left(\sigma_c^2\left(\frac{K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}}-\frac{1}{\sigma_c^2}\right)+1\right)+\left(\frac{K_\ell}{\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}}-\frac{1}{\sigma_c^2}\right)} \\
&= \frac{(1-\gamma)^2(1+2\gamma)K_\ell\sigma_c^4}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)} + \frac{\gamma^2(3-2\gamma)\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)} \\
&\quad +(1-\gamma)\gamma\sigma_c^2 - \frac{2(1-\gamma)\gamma^2\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_\ell^2}\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)} \\
&\quad - \frac{2(1-\gamma)^2\gamma\left(\sigma_c^4K_\ell-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)\sigma_c^2\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)} \\
&= \frac{(1-\gamma)^2\left(K_\ell\sigma_c^4+2\gamma\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)\sigma_c^2\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)} + (1-\gamma)\gamma\sigma_c^2 \\
&\quad + \frac{\gamma^2\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}+\frac{2(1-\gamma)}{\sigma_c^2}\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)}.
\end{aligned}
$$

Dropping terms that are constant with respect to σ_ℓ^2 and $\sigma_{\epsilon\ell}^2$ and letting $x = \lambda/\sigma_\ell^2 + 1/\sigma^2 + 1/\sigma_c^2$, we can rewrite the decision problem as

$$
\min_x \frac{(1-\gamma)^2\left(K_\ell\sigma_c^4+2\gamma\sigma_c^2x\right)+\gamma^2\sigma_c^4x\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda+\frac{2(1-\gamma)}{\sigma_c^2}\right)}{K_\ell\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda\right)-x} \equiv H(x)
$$

with x taking value from $[1/\sigma^2+1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2+1/\sigma_c^2)]$. Denote the minimizer be x^*.

$$
\begin{aligned}
H'(x) &= \frac{\left[2(1-\gamma)^2\gamma\sigma_c^2+\gamma^2\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+2x/\lambda+\frac{2(1-\gamma)}{\sigma_c^2}\right)\right]\left[K_\ell\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda\right)-x\right]}{\left[K_\ell\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda\right)-x\right]^2} \\
&\quad - \frac{\left[(1-\gamma)^2\left(K_\ell\sigma_c^4+2\gamma\sigma_c^2x\right)+\gamma^2\sigma_c^4x\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda+\frac{2(1-\gamma)}{\sigma_c^2}\right)\right]\left(K_\ell\sigma_c^4/\lambda-1\right)}{\left[K_\ell\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda\right)-x\right]^2}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\frac{\gamma^2\sigma_c^4}{\lambda}\left(\frac{K_\ell\sigma_c^4}{\lambda}-1\right)x^2+\gamma^2K_\ell\sigma_c^8\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+2x/\lambda+\frac{2(1-\gamma)}{\sigma_c^2}\right)\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}\right)}{\left[K_\ell\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda\right)-x\right]^2} \\
&\quad+\frac{2(1-\gamma)^2\gamma K_\ell\sigma_c^6\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}\right)-(1-\gamma)^2K_\ell\sigma_c^4(K_\ell\sigma_c^4/\lambda-1)}{\left[K_\ell\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda\right)-x\right]^2} \\
&= \frac{\gamma^2\sigma_c^4}{K_\ell\sigma_c^4-\lambda}-\frac{A}{\left[K_\ell\sigma_c^4\left(\frac{1-1/\lambda}{\sigma^2}+\frac{1-1/\lambda}{\sigma_c^2}+x/\lambda\right)-x\right]^2},
\end{aligned}
$$

with A, being constant with respect to x, is given by

$$
A \equiv B^2\frac{\gamma^2K_\ell\sigma_c^8\lambda}{K_\ell\sigma_c^4-\lambda}+2(1-\gamma)\gamma K_\ell\sigma_c^6B+(1-\gamma)^2K_\ell\sigma_c^4\left(\frac{K_\ell\sigma_c^4}{\lambda}-1\right),
$$

where $B \equiv (1/\lambda-1)(1/\sigma^2+1/\sigma_c^2) > 0$. Since $B>0$ and $K_\ell\sigma_c^4 > 1 > \lambda$, $A>0$. We then have

$$
H''(x) = \frac{2A(K_\ell\sigma_c^4/\lambda-1)}{[(K_\ell\sigma_c^4/\lambda-1)x-K_\ell\sigma_c^4B]^3} > 0,
$$

which implies that x^* must be unique.

Following the proof of Lemma 10, we define

$$
G(\gamma;x) \equiv \frac{\gamma^2\sigma_c^4}{K_\ell\sigma_c^4-\lambda}-\frac{B^2\frac{\gamma^2K_\ell\sigma_c^8\lambda}{K_\ell\sigma_c^4-\lambda}+2(1-\gamma)\gamma K_\ell\sigma_c^6B+(1-\gamma)^2K_\ell\sigma_c^4\left(\frac{K_\ell\sigma_c^4}{\lambda}-1\right)}{[(K_\ell\sigma_c^4/\lambda-1)x-K_\ell\sigma_c^4B]^2} = H'(x)
$$

with x taking value from $[1/\sigma^2+1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2+1/\sigma_c^2)]$.

Claim 1. For any x in $[1/\sigma^2+1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2+1/\sigma_c^2)]$, $G(1;x)>0$.

Given the definition of G, it is equivalent to show

$$
\frac{\sigma_c^4}{K_\ell\sigma_c^4-\lambda} > \frac{B^2\frac{K_\ell\sigma_c^8\lambda}{K_\ell\sigma_c^4-\lambda}}{[(K_\ell\sigma_c^4/\lambda-1)x-K_\ell\sigma_c^4B]^2} \Leftrightarrow \left[(K_\ell\sigma_c^4/\lambda-1)x-K_\ell\sigma_c^4B\right]^2 > B^2K_\ell\sigma_c^4\lambda,
$$

for any x in $[1/\sigma^2+1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2+1/\sigma_c^2)]$. Since the left hand side increases with x, it is equivalent to have the inequality with x being replaced with $1/\sigma^2+1/\sigma_c^2$. Then we have

$$
\frac{(K_\ell\sigma_c^4-1)^2}{K_\ell\sigma_c^4} > \lambda(1/\lambda-1)^2.
$$

Since $K_\ell\sigma_c^4>1$, $(K_\ell\sigma_c^4-1)^2/(K_\ell\sigma_c^4)$ increases with $K_\ell\sigma_c^4$ which itself increases with σ_c^2. Then it suffices for the inequality above to hold if it holds for the smallest $\sigma_c^2=\sigma^2/(2^{2\kappa}-1)$, that is,

$$
\lambda(1/\lambda-1)^2 < \frac{2^{4\kappa}\left(\frac{1}{2^{2\kappa}-1}+\frac{2^{2\kappa}\sigma^2}{(2^{2\kappa}-1)^2\sigma_\epsilon^2}\right)^2}{2^{2\kappa}\left(\frac{1}{2^{2\kappa}-1}+\frac{2^{2\kappa}\sigma^2}{(2^{2\kappa}-1)^2\sigma_\epsilon^2}\right)+1},
$$

which coincides with regularity condition 12 for λ.

We now have shown that $G(1;x)>0$. This implies that $x^*=1/\sigma^2+1/\sigma_c^2$ or equivalently

$\sigma_\ell^2 = \infty$ when $\gamma = 1$, which echoes the result in the baseline setting.

Since $G(1;x) > 0$, we claim that there must exist a unique $\gamma \in (0,1)$ such that $G(\gamma;x) = 0$ for any x in $[1/\sigma^2 + 1/\sigma_c^2, K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)]$. To see this, $G(\gamma;x) = 0$ is equivalent to

$$
\begin{aligned}
L(\gamma) &\equiv 2K_\ell\sigma_c^6 B\left(\frac{1-\gamma}{\gamma}\right) + K_\ell\sigma_c^4\left(\frac{K_\ell\sigma_c^4}{\lambda} - 1\right)\left(\frac{1-\gamma}{\gamma}\right)^2 \\
&= \frac{\sigma_c^4\left[(K_\ell\sigma_c^4/\lambda - 1)x - K_\ell\sigma_c^4 B\right]^2}{K_\ell\sigma_c^4 - \lambda} - B^2\frac{K_\ell\sigma_c^8\lambda}{K_\ell\sigma_c^4 - \lambda} > 0,
\end{aligned}
$$

where the inequality follows from $G(1;x) > 0$. Given $\gamma \in [0,1]$, it is easy to see that $L(\gamma)$ strictly decreases with γ and $L(1) = 0$ and $\lim_{\gamma\to 0} L(\gamma) \to \infty$. Therefore, there must exist a unique $\gamma \in (0,1)$ such that $G(\gamma;x) = 0$. We now define $\underline{\gamma}'$ such that $G(\underline{\gamma}';x) = 0$ for $x = K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)$ and $\bar{\gamma}'$ such that $G(\bar{\gamma}';x) = 0$ for $x = 1/\sigma^2 + 1/\sigma_c^2$. Since $1/\sigma^2 + 1/\sigma_c^2 < K_\ell/(1/\sigma_\epsilon^2 + 1/\sigma_c^2)$, by construction, we have $L(\underline{\gamma}') > L(\bar{\gamma}')$. Further, we know L is strictly decreasing, so $\underline{\gamma}' < \bar{\gamma}'$.

The rest of the proof directly follows from the proof of Lemma 10. When $\gamma \in (\underline{\gamma}', \bar{\gamma}')$, we have $H'(x^*) = 0$, or equivalently

$$
\begin{aligned}
\frac{\gamma^2\sigma_c^4}{K_\ell\sigma_c^4 - \lambda} &= \frac{B^2\frac{\gamma^2 K_\ell\sigma_c^8\lambda}{K_\ell\sigma_c^4 - \lambda} + 2(1-\gamma)\gamma K_\ell\sigma_c^6 B + (1-\gamma)^2 K_\ell\sigma_c^4\left(\frac{K_\ell\sigma_c^4}{\lambda} - 1\right)}{\left[(K_\ell\sigma_c^4/\lambda - 1)x^* - K_\ell\sigma_c^4 B\right]^2} \\
\left[(K_\ell\sigma_c^4 - \lambda)x^* - \lambda K_\ell\sigma_c^4 B\right]^2 &= \lambda\left(B^2 K_\ell\lambda^2\sigma_c^4 + \frac{2(1-\gamma)}{\gamma}BK_\ell\lambda\sigma_c^2(K_\ell\sigma_c^4 - \lambda) + \left(\frac{1-\gamma}{\gamma}\right)^2 K_\ell(K_\ell\sigma_c^4 - \lambda)^2\right) \\
\left[(K_\ell\sigma_c^4 - \lambda)x^* - \lambda K_\ell\sigma_c^4 B\right]^2 &= \lambda K_\ell\left(B\lambda\sigma_c^2 + \left(\frac{1-\gamma}{\gamma}\right)(K_\ell\sigma_c^4 - \lambda)\right)^2 \\
(K_\ell\sigma_c^4 - \lambda)x^* - \lambda K_\ell\sigma_c^4 B &= \lambda^{1/2}K_\ell^{1/2}\left(B\lambda\sigma_c^2 + \left(\frac{1-\gamma}{\gamma}\right)(K_\ell\sigma_c^4 - \lambda)\right) \\
x^* &= \frac{1-\gamma}{\gamma}(\lambda K_\ell)^{1/2} + \frac{\lambda BK_\ell^{1/2}\sigma_c^2(K_\ell^{1/2}\sigma_c^2 + \lambda^{1/2})}{K_\ell\sigma_c^4 - \lambda} \\
x^* &= \frac{1-\gamma}{\gamma}(\lambda K_\ell)^{1/2} + \frac{\lambda BK_\ell^{1/2}\sigma_c^2}{K_\ell^{1/2}\sigma_c^2 - \lambda^{1/2}}, \qquad (15)
\end{aligned}
$$

where the last equation nests Equation 7 as a special case. □

A.23 Proof of Lemma 20

Proof. Following the proof of Lemma 11, we first obtain

$$
\begin{aligned}
E(a_\ell - \theta)^2 &= (1-\gamma^2)Var(\theta|\theta_\ell, s_c', s_c'') + \gamma^2 Var(\theta_c|\theta_\ell, s_c', s_c'') + \gamma^2\sigma_c^2 - 2\gamma^2\sigma_c^2\frac{\frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma_{\epsilon\ell}^2} + \frac{1}{\sigma_c^2 + (1/\sigma^2 + 1/\sigma_\ell^2)^{-1}}} \\
&= \frac{(1-\gamma^2)K_\ell\sigma_c^4}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2}\right) - \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{\lambda}{\sigma_\ell^2}\right)} + \frac{\gamma^2\sigma_c^4\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{\lambda}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2}\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2}\right) - \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{\lambda}{\sigma_\ell^2}\right)} \\
&\quad + \gamma^2\sigma_c^2 - \frac{2\gamma^2\sigma_c^4\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{\lambda}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\ell^2}\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{1}{\sigma_\ell^2}\right) - \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2} + \frac{\lambda}{\sigma_\ell^2}\right)}
\end{aligned}
$$

$$= \frac{(1-\gamma^2)K_\ell\sigma_c^4 + \gamma^2\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma_c^2}-\frac{1}{\sigma^2}-\frac{1}{\sigma_\ell^2}\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)} + \gamma^2\sigma_c^2$$

where the second to last equality follows from the fact that Constraint 10 is binding. Then the optimization problem of the central government can be rewritten as

$$\min_{\sigma_c^2} E(a_\ell-\theta)^2 = \frac{(1-\gamma^2)K_\ell\sigma_c^4 + \gamma^2\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)\left(\frac{1}{\sigma_c^2}-\frac{1}{\sigma^2}-\frac{1}{\sigma_\ell^2}\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\ell^2}\right)-\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{\lambda}{\sigma_\ell^2}\right)} + \gamma^2\sigma_c^2 \equiv F(\sigma_c^2)$$

subject to Constraint 3, where it should be emphasized that both σ_ℓ^2 and K_ℓ are functions of σ_c^2.

Consider two cases: (1) $\gamma > \bar{\gamma}'$; (2) $\gamma < \underline{\gamma}'$.

Case (1): $\gamma > \bar{\gamma}'$.

In this case, according to Lemma 10, $\sigma_\ell^2 = \infty$. Then we have

$$F(\sigma_c^2) = \frac{K_\ell + K_\ell\gamma^2\sigma_c^2/\sigma^2 - \gamma^2(1/\sigma^2+1/\sigma_c^2)/\sigma^2}{(K_\ell - 1/\sigma_c^4)(1/\sigma^2+1/\sigma_c^2)},$$

which does not depend on λ and the objective function coincides with that in the baseline setting. Then we know $F'(\sigma_c^2) > 0$.

Case (2): $\gamma < \underline{\gamma}'$.

According to Lemma 10, we have $\sigma_{\epsilon\ell}^2 = \infty$, which implies that Constraint 4 can be rewritten as

$$\frac{1}{\sigma^2} + \frac{\lambda}{\sigma_\ell^2} + \frac{1}{\sigma_c^2} = K_\ell\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}\right)^{-1}.$$

The objective function of the central government can then be written as

$$\begin{aligned}
F(\sigma_c^2) &= \frac{(1-\gamma^2)K_\ell\sigma_c^4 + \frac{\gamma^2\sigma_c^4K_\ell}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}}\left(\frac{1}{\sigma_c^2}-\frac{1}{\sigma^2}-\frac{1}{\lambda}\left(\frac{K_\ell}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}}-\frac{1}{\sigma^2}-\frac{1}{\sigma_c^2}\right)\right)}{K_\ell\sigma_c^4\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}+\frac{1}{\lambda}\left(\frac{K_\ell}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}}-\frac{1}{\sigma^2}-\frac{1}{\sigma_c^2}\right)\right)-\frac{K_\ell}{\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}}} + \gamma^2\sigma_c^2 \\
&= \frac{(1-\gamma^2)K_\ell\sigma_c^4\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)^2 + \gamma^2\sigma_c^4K_\ell\left(\left(\frac{1+\lambda}{\lambda\sigma_c^2}+\frac{1-\lambda}{\lambda\sigma^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)-\frac{K_\ell}{\lambda}\right)}{K_\ell\sigma_c^4\left(\frac{\lambda-1}{\lambda}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)^2+\frac{K_\ell}{\lambda}\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)\right)-K_\ell\left(\frac{1}{\sigma_\epsilon^2}+\frac{1}{\sigma_c^2}\right)} + \gamma^2\sigma_c^2 \\
&= \frac{(1-\gamma^2)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)^2 + \gamma^2\left(\left(\frac{1+\lambda}{\lambda\sigma_c^2}+\frac{1-\lambda}{\lambda\sigma^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)-\frac{K_\ell}{\lambda}\right)}{\left(\frac{K_\ell}{\lambda}-\frac{1}{\sigma_c^4}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)-\frac{1-\lambda}{\lambda}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)^2} + \gamma^2\sigma_c^2 \\
&\equiv (1-\gamma^2)F_1(\sigma_c^2) + \gamma^2F_2(\sigma_c^2)
\end{aligned}$$

with

$$F_1(\sigma_c^2) \equiv \frac{\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)^2}{\left(\frac{K_\ell}{\lambda}-\frac{1}{\sigma_c^4}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)-\frac{1-\lambda}{\lambda}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)^2}$$

$$= \left(\frac{\frac{K_\ell}{\lambda} - \frac{1}{\sigma_c^4}}{\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\epsilon^2}} - \frac{1-\lambda}{\lambda}\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_c^2}\right)\right)^{-1}$$

$$= \frac{\frac{1}{\sigma_c^2} + \frac{1}{\sigma_\epsilon^2}}{\frac{2^{2\kappa}-1+\lambda}{\lambda}\left(\frac{\sigma_c^2+\sigma^2+\sigma_\epsilon^2}{\sigma_c^2\sigma^2\sigma_\epsilon^2}\right)} = \frac{\lambda\sigma^2(\sigma_c^2+\sigma_\epsilon^2)}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma^2+\sigma_\epsilon^2)},$$

$$F_2(\sigma_c^2) \equiv \frac{\left(\frac{1+\lambda}{\lambda\sigma_c^2} + \frac{1-\lambda}{\lambda\sigma^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right) - \frac{K_\ell}{\lambda}}{\left(\frac{K_\ell}{\lambda} - \frac{1}{\sigma_c^4}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right) - \frac{1-\lambda}{\lambda}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)^2} + \sigma_c^2$$

$$= \frac{\sigma_\epsilon^2[(1+\lambda)\sigma^2 + (1-\lambda)\sigma_c^2]\left(\sigma_c^2+\sigma_\epsilon^2\right) - K_\ell\sigma^2\sigma_\epsilon^4\sigma_c^4}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)(\sigma_c^2+\sigma_\epsilon^2)} + \sigma_c^2$$

It is easy to see that $F_1'(\sigma_c^2) > 0$ and

$$\begin{aligned} F_2'(\sigma_c^2) &= 1 + \frac{\sigma_\epsilon^2[(1+\lambda)\sigma^2 + (1-\lambda)\sigma_\epsilon^2 + 2(1-\lambda)\sigma_c^2] - 2^{2\kappa}\sigma_\epsilon^2(\sigma^2+2\sigma_c^2+\sigma_\epsilon^2)}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)(\sigma_c^2+\sigma_\epsilon^2)} \\ &\quad - \frac{\{\sigma_\epsilon^2[(1+\lambda)\sigma^2+(1-\lambda)\sigma_c^2]\left(\sigma_c^2+\sigma_\epsilon^2\right) - K_\ell\sigma^2\sigma_\epsilon^4\sigma_c^4\}(2\sigma_c^2+2\sigma_\epsilon^2+\sigma^2)}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)^2(\sigma_c^2+\sigma_\epsilon^2)^2} \\ &= \frac{2\lambda\sigma^2\sigma_\epsilon^2 + (2^{2\kappa}-1+\lambda)\sigma_c^2(\sigma_c^2+\sigma^2)}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)(\sigma_c^2+\sigma_\epsilon^2)} \\ &\quad - \frac{\sigma_\epsilon^2\{[(1+\lambda)\sigma^2+(1-\lambda)\sigma_c^2]\left(\sigma_c^2+\sigma_\epsilon^2\right) - K_\ell\sigma^2\sigma_\epsilon^2\sigma_c^4\}(2\sigma_c^2+2\sigma_\epsilon^2+\sigma^2)}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)^2(\sigma_c^2+\sigma_\epsilon^2)^2} \\ &= \frac{\sigma_c^2(\sigma_c^2+\sigma^2)}{(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)(\sigma_c^2+\sigma_\epsilon^2)} + \frac{2\lambda\sigma^4\sigma_\epsilon^2}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)^2(\sigma_c^2+\sigma_\epsilon^2)} \\ &\quad + \frac{\sigma_\epsilon^2\{K_\ell\sigma^2\sigma_\epsilon^2\sigma_c^4 - [\sigma^2+(1-\lambda)\sigma_c^2]\left(\sigma_c^2+\sigma_\epsilon^2\right)\}(2\sigma_c^2+2\sigma_\epsilon^2+\sigma^2)}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)^2(\sigma_c^2+\sigma_\epsilon^2)^2} \\ &> \frac{\sigma_c^2(\sigma_c^2+\sigma^2)}{(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)(\sigma_c^2+\sigma_\epsilon^2)} + \frac{2\lambda\sigma^4\sigma_\epsilon^2}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)^2(\sigma_c^2+\sigma_\epsilon^2)} \\ &\quad + \frac{\sigma_\epsilon^4\sigma^2\sigma_c^4\left\{K_\ell - \left(\frac{1}{\sigma^2}+\frac{1}{\sigma_c^2}\right)\left(\frac{1}{\sigma_c^2}+\frac{1}{\sigma_\epsilon^2}\right)\right\}(2\sigma_c^2+2\sigma_\epsilon^2+\sigma^2)}{(2^{2\kappa}-1+\lambda)(\sigma_c^2+\sigma_\epsilon^2+\sigma^2)^2(\sigma_c^2+\sigma_\epsilon^2)^2} > 0, \end{aligned}$$

where the first inequality follows from $\lambda > 0$ and the second inequality follows from $K_\ell > (1/\sigma^2 + 1/\sigma_c^2)(1/\sigma_c^2 + 1/\sigma_\epsilon^2)$.

Since $\gamma < \underline{\gamma}$ continues to hold for a small change of σ_c^2, the objective function is differentiable and its first derivative is given by

$$F'(\sigma_c^2) = (1-\gamma^2)F_1'(\sigma_c^2) + \gamma^2 F_2'(\sigma_c^2) > 0.$$

□

A.24 Proof of Lemma 21

Proof. We first consider $\gamma = 0$. We have

$$E(a_\ell - \theta)^2\Big|_{\gamma=0} = \left(\frac{1}{\sigma^2} + \frac{\lambda(2^{2\kappa}-1)(\sigma^2+\sigma_\epsilon^2+\sigma_c^2)}{\sigma^2(\sigma_\epsilon^2+\sigma_c^2)} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2}\right)^{-1},$$

with $\sigma_c^2 = \sigma^2/(2^{2\kappa}-1)$.

$$E(a_c-\theta)^2 = \left(\frac{1}{\sigma^2} + \frac{(2^{2\kappa}-1)(\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2)}{\sigma^2(\sigma_\epsilon^2+\sigma_\ell^2)} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2}\right)^{-1},$$

with $\sigma_\ell^2 = \lambda\sigma^2/(2^{2\kappa}-1)$. Then to show $E(a_c-\theta)^2 > E(a_\ell-\theta)^2\Big|_{\gamma=0}$, it is equivalent to show

$$\frac{1}{\sigma^2} + \frac{(2^{2\kappa}-1)(\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2)}{\sigma^2(\sigma_\epsilon^2+\sigma_\ell^2)} + \frac{1}{\sigma_\ell^2+\sigma_\epsilon^2} < \frac{1}{\sigma^2} + \frac{\lambda(2^{2\kappa}-1)(\sigma^2+\sigma_\epsilon^2+\sigma_c^2)}{\sigma^2(\sigma_\epsilon^2+\sigma_c^2)} + \frac{1}{\sigma_c^2+\sigma_\epsilon^2}$$
$$\Leftrightarrow \quad \frac{\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2+\sigma_c^2}{\sigma_c^2(\sigma_\epsilon^2+\sigma_\ell^2)} < \frac{\sigma^2+\sigma_\epsilon^2+\sigma_\ell^2+\sigma_c^2}{\sigma_\ell^2(\sigma_\epsilon^2+\sigma_c^2)},$$

where the second inequality directly follows from $\lambda < 1$ ($\sigma_\ell^2 < \sigma_c^2$).

Now consider $\gamma = 1$. The expression of $E(a_\ell-\theta)^2\Big|_{\gamma=1}$ coincides with that in the baseline setting. Following the proof of Lemma 13, we can show that $E(a_\ell-\theta)^2\Big|_{\gamma=1} > \min\{\sigma^2,\sigma_c^2\} > E(a_c-\theta)^2$. $\square$

References

Akai, Nobuo, Masayo Hosoi, and Yukihiro Nishimura, "FISCAL DECENTRALIZATION AND ECONOMIC VOLATILITY: EVIDENCE FROM STATE-LEVEL CROSS-SECTION DATA OF THE USA," *The Japanese Economic Review*, 2009, *60* (2), 223–235.

Bardhan, Pranab, "State and development: The need for a reappraisal of the current literature," *Journal of Economic Literature*, 2016, *54* (3), 862–92.

Bellofatto, Antonio Andrés and Martin Besfamille, "Regional state capacity and the optimal degree of fiscal decentralization," *Journal of Public Economics*, 2018.

Bleaney, Michael and David Fielding, "Exchange rate regimes, inflation and output volatility in developing countries," *Journal of development economics*, 2002, *68* (1), 233–245.

Board, Oliver J, Andreas Blume, and Kohei Kawamura, "Noisy talk," *Theoretical Economics*, 2007, *2* (4), 395–440.

Buch, Claudia M, Jörg Döpke, and Harald Strotmann, "Does Export Openness Increase Firm-level Output Volatility?," *The World Economy*, 2009, *32* (4), 531–551.

Cover, Thomas M and Joy A Thomas, *Elements of information theory*, John Wiley & Sons, 2012.

Crawford, Vincent P and Joel Sobel, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1431–1451.

Davoodi, Hamid and Heng fu Zou, "Fiscal decentralization and economic growth: A cross-country study," *Journal of Urban economics*, 1998, *43* (2), 244–257.

Gabaix, Xavier, "Behavioral inattention," Technical Report, National Bureau of Economic Research 2017.

Giovanni, Julian Di and Andrei A Levchenko, "Country size, international trade, and aggregate fluctuations in granular economies," *Journal of Political Economy*, 2012, *120* (6), 1083–1132.

Huang, Zhangkai, Lixing Li, Guangrong Ma, and Lixin Colin Xu, "Hayek, local information, and commanding heights: Decentralizing state-owned enterprises in china," *American Economic Review*, 2017, *107* (8), 2455–78.

Jin, Hehui, Yingyi Qian, and Barry R Weingast, "Regional decentralization and fiscal incentives: Federalism, Chinese style," *Journal of public economics*, 2005, *89* (9-10), 1719–1742.

Koren, Miklós and Silvana Tenreyro, "Volatility and development," *The Quarterly Journal of Economics*, 2007, *122* (1), 243–287.

_ **and** _ , "Technological diversification," *American Economic Review*, 2013, *103* (1), 378–414.

Li, Hongbin and Li-An Zhou, "Political turnover and economic performance: the incentive role of personnel control in China," *Journal of public economics*, 2005, *89* (9), 1743–1762.

Mackowiak, Bartosz and Mirko Wiederholt, "Optimal sticky prices under rational inattention," *American Economic Review*, 2009, *99* (3), 769–803.

Malik, Adeel and Jonathan RW Temple, "The geography of output volatility," *Journal of Development Economics*, 2009, *90* (2), 163–178.

Mookherjee, Dilip, "Political decentralization," *Annual Review of Economics*, 2015, *7* (1), 231–249.

Qian, Yingyi and Gerard Roland, "Federalism and the soft budget constraint," *American economic review*, 1998, pp. 1143–1162.

Ramey, Garey and Valerie Ramey, "Cross-Country Evidence on the Link between Volatility and Growth," *American Economic Review*, 1995, *85* (5), 1138–51.

Sims, Christopher A, "Implications of rational inattention," *Journal of monetary Economics*, 2003, *50* (3), 665–690.

Wang, Boqun and Tao Yang, "Volatility and Economic Systems: Evidence from a Large Transitional Economy," *Working Paper*, 2016.

Wetzel, Deborah, "Decentralization in the transition economies: Challenges and the road ahead," 2001.

Wu, Jinglian, *Understanding and Interpreting Chinese Economic Reform*, Texere, 2005.

Xu, Chenggang, "The fundamental institutions of China's reforms and development," *Journal of Economic Literature*, 2011, *49* (4), 1076–1151.

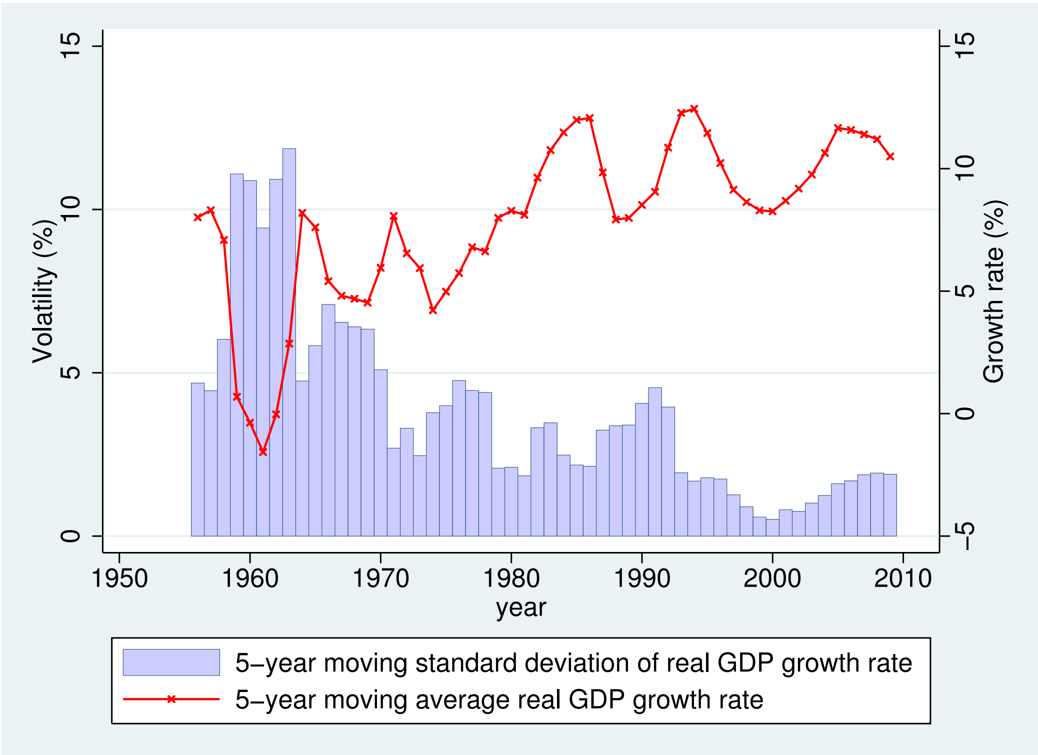
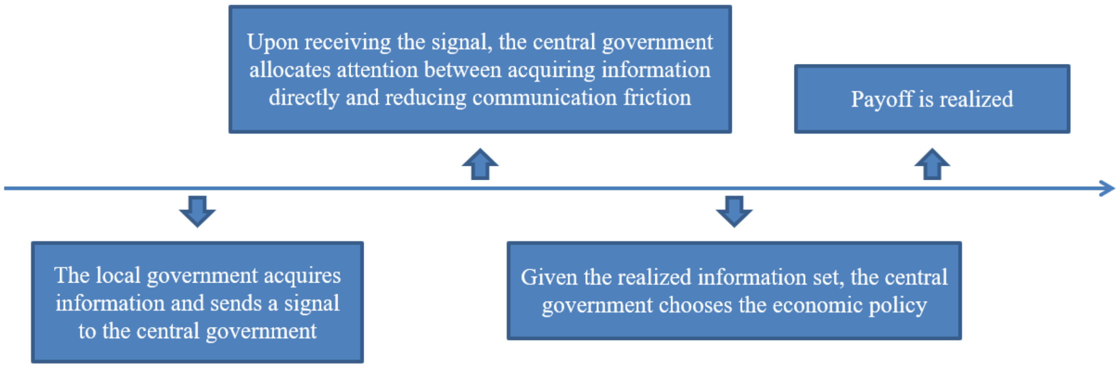
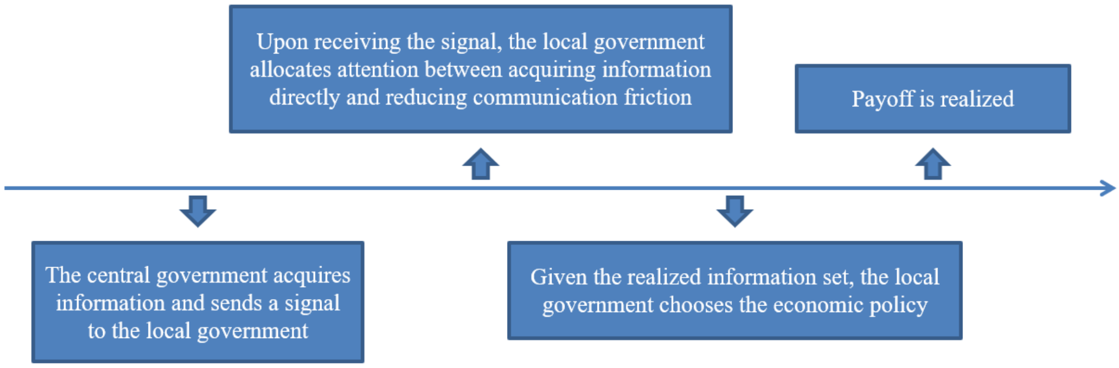


Figure 1: China's economic growth and volatility



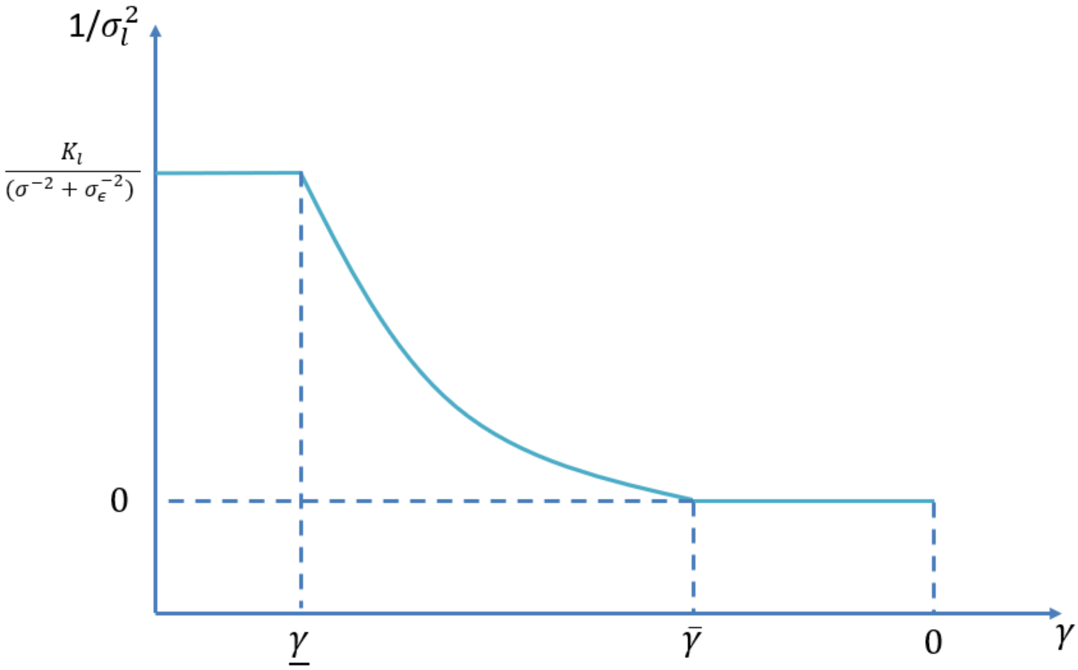
(1) The Centralized Regime



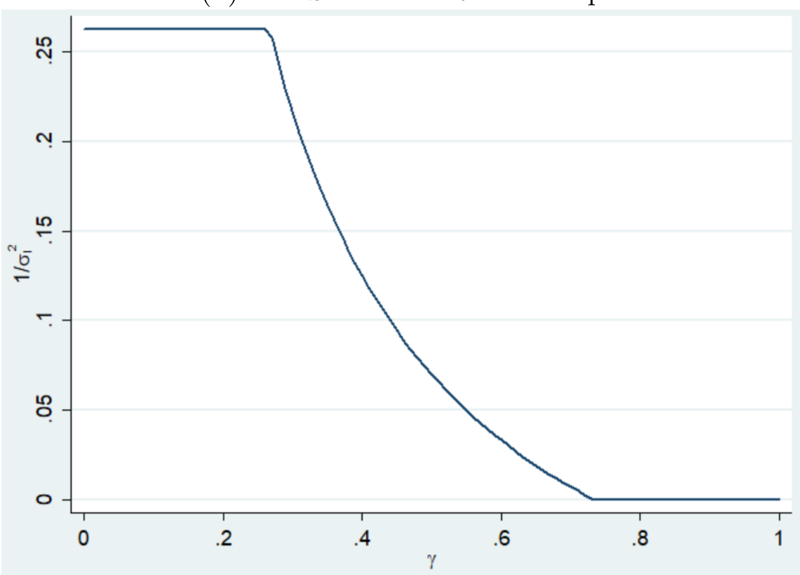
(2) The Decentralized Regime

Figure 2: Timeline: The Centralized and Decentralized Regimes

(a) The Analytical Form



(b) The Simulated Relationship



(Note: $\sigma^2 = \sigma_\epsilon^2 = 100; \kappa_c = 1; \kappa_\ell = 2; \sigma_c^2 = \sigma^2/(2^{2\kappa_c} - 1)$.)

Figure 3: The Relationship between $1/\sigma_\ell^2$ ang γ under the Decentralized Regime

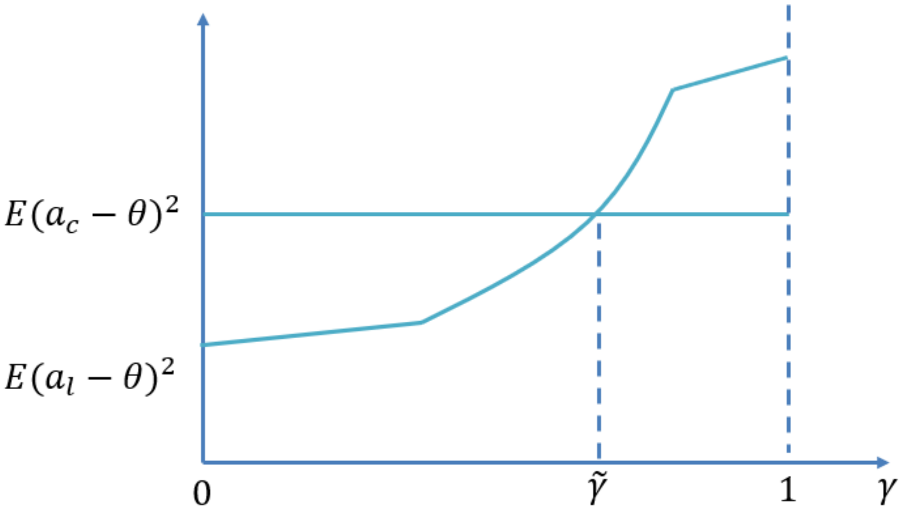


Figure 4: The Comparison between Two Regimes

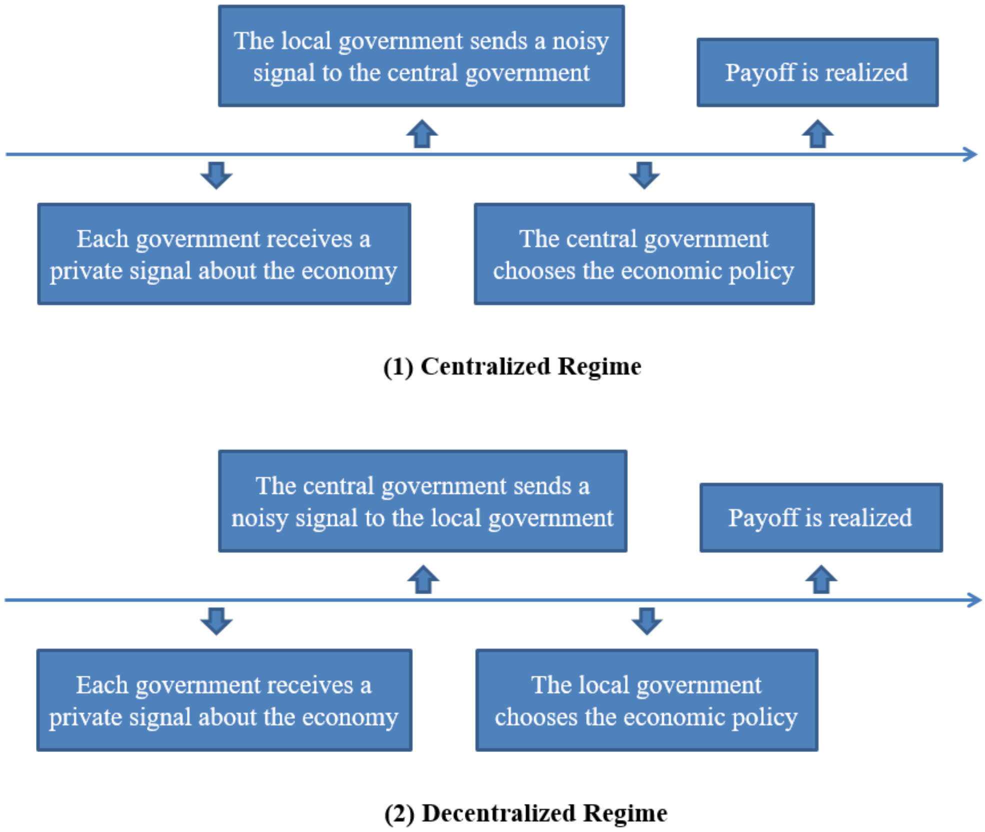
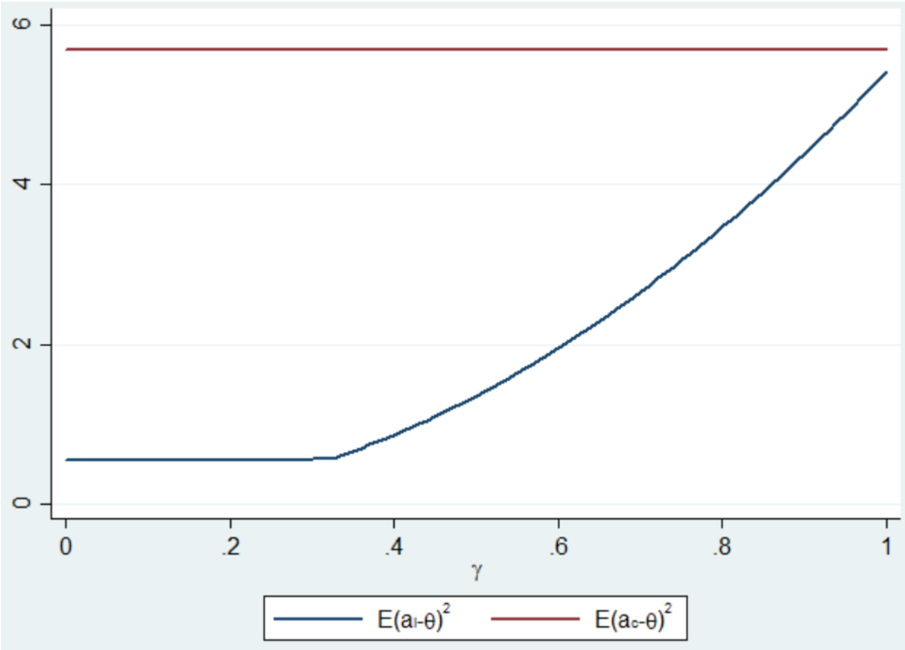
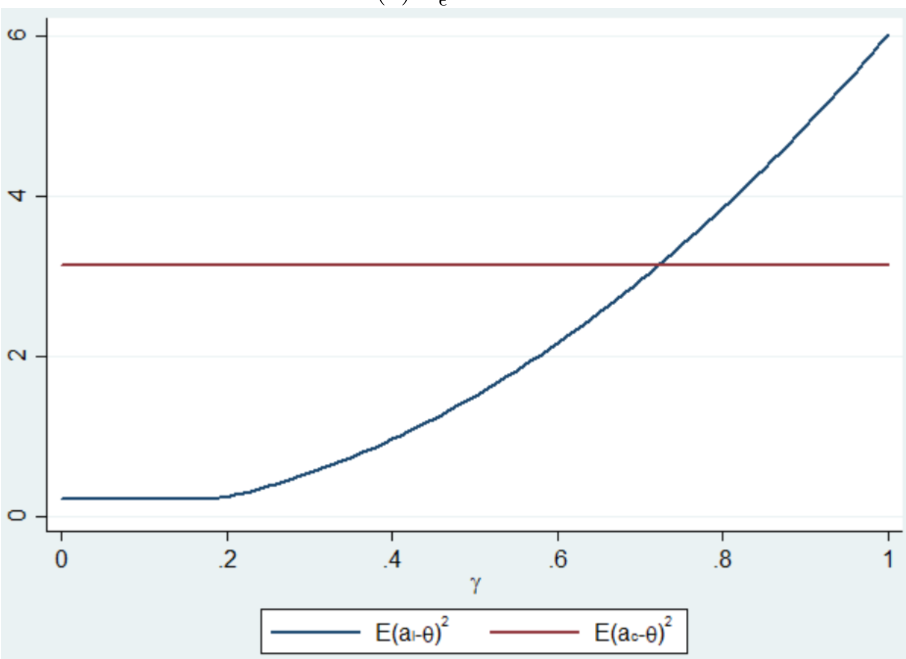


Figure 5: Timeline: Centralized and Decentralized Regimes

(1) $\sigma_\epsilon^2 = 1000$



(2) $\sigma_\epsilon^2 = 100$



(Note: $\sigma^2 = 100; \kappa = 2.$)

Figure 6: Comparison between Two Regimes When Condition 12 Is Violated

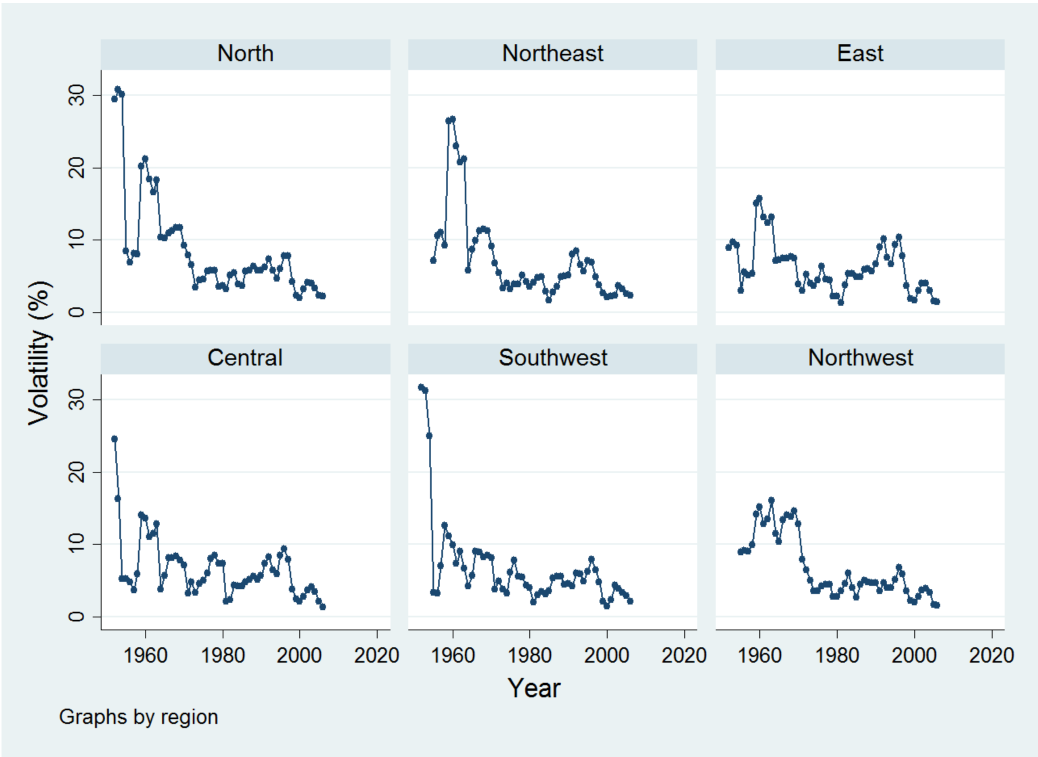


Figure 7: Volatility by regions

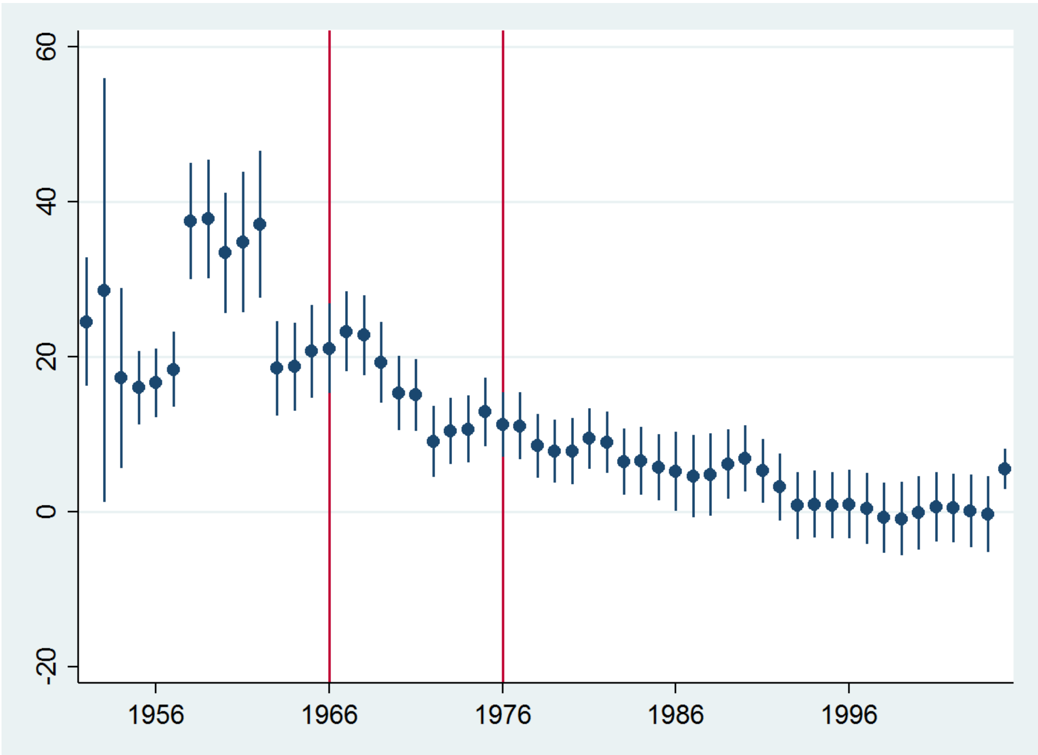


Figure 8: Dynamic effects of percentage of state-owned outputs (1950) on volatilty

Table 1: Evidence on changing promotion incentive

VARIABLES	(1) promotion	(2) promotion	(3) promotion	(4) promotion
local	-0.064		-0.071	
	(0.125)		(0.097)	
localofficial	0.086		0.003	
	(0.145)		(0.151)	
wailai	-0.054		0.071	
	(0.140)		(0.142)	
army	0.234**		-0.136	
	(0.096)		(0.118)	
grppc_growth		26.295		1.271**
		(50.447)		(0.537)
Observations	78	60	185	188
R-squared	0.087	0.006	0.018	0.064
Period	1950s	1950s	1990s	1990s

† Notes: *** denotes significance at 0.01, ** at 0.05, and * at 0.1. Outcomes variable is whether a provincial leader got promoted. Main independent variable is a set of variables measuring political background and GDP growth rate. Standard errors are clustered on province level.