# Decentralization in Government: Career Concerns and Yardstick Competition

**Hideki Konishi***

Faculty of Political Science and Economics, Waseda University
1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, Japan
Email: h.konishi@waseda.jp, Phone & Fax: +81-3-5286-8599

**September 10, 2010**

**Abstract**

This paper applies the career-concerns model of policy-making to a political economy of fiscal federalism and examines the effect of decentralization on policy-makers' effort incentives and its relative efficiency over centralization. In the single-task case, decentralization induces larger effort than centralization owing to the effect of focus and yardstick competition, though it is not necessarily more efficient and in some cases partial integration is the most efficient. In the two-task case, three modes of federalism, i.e., full centralization, full decentralization, and partial (de)centralization are considered and it is shown that there occurs a trade-off on effort incentives via the effects of focus and noise. The number of regions in a nation turns out to play a key role for partial (de)centralization to induce larger effort on each task than full centralization and full decentralization.

*Key words*: centralization, decentralization; yardstick competition; fiscal federalism; career concerns

---

*The author is grateful to Amihai Glazer, Toshihiro Ihori, Yimiyoshi Kamada, Shigeki Kunieda, Motohiro Sato, Therese McGuire, Yukihiro Nishimura, and Kimiko Terai for their comments and suggestions.

# 1 Introduction

This paper analyzes the effect of decentralization on effort incentives of policy-makers in central and regional governments and examines the relative efficiency of decentralization over centralization, constructing a model of career concerns and yardstick competition.

Yardstick competition among local policy-makers has recently attracted a lot of attention in both theoretical and empirical studies on the effects of federalism on policy formation and economic development. Such a surge of research interest is partly related to the rapid economic development in China. Many authors including Maskin, Qian, and Xu (2000), Blanchard and Shleifer (2001), and Li and Zhou (2005) pointed out that one of the important sources for the Chinese economic success is its personnel control embedded in the centralized political hierarchy, which is called "M-form" structure (Maskin, Qian, and Xu (2000)), "Market-preserving federalism" (Weingast (1995)), or "Federalism, Chinese style" (Montinola, Qian, and Weigast (1995)) in the literature. In this system, local policy-makers, who are given primary authority over the economy within their jurisdictions, compete for promotion in the centralized political hierarchy via their relative economic performance. Taking account of these features, we could say that the Chinese-style federalism is a framework of yardstick competition among career-concerned local policy-makers.

A similar personnel control system has also been made use of in the bureaucratic society in post-war Japan. As Inoki (2001) reported, staff loans and transfers between central and local governments has functioned as a mechanism that screens competent bureaucrats out, based on relative performance evaluation. Those who had climbed higher in the bureaucratic hierarchy used to be, in general, guaranteed with higher-paid jobs in private or semi-private corporations upon retirement from public services; a system called *Amakudari* in Japanese.[1]

A huge number of studies have been piled up in the literature on the political economy of fiscal federalism, dating back to Tiebout (1956) and Oates (1972). This paper belongs to the more recent strand of the literature, the incomplete contract approach, pioneered by Seabright (1996). He analyzed how decentralization affects effort incentives of reelection-seeking incumbents

[1]The word, *Amakudari*, literally means going down from the heaven to the earth, with a feudalistic sense of usage that the heaven means the government while the earth does the private sector.

in undertaking tasks in each region. Although voters cannot observe incumbents' effort, they instead observe the outcomes with noise, based on which they decide whom to vote for in the forthcoming election held in each region. His main point is that under centralization a politician has incentives to target regions in which she exerts effort because she will get reelected if she wins only the majority of all regional elections. Maskin, Qian, and Xu (2000), on the other hand, took a complete contract approach to the relative efficiency of decentralization with different organizational forms, called M-form and U-form, where the former corresponds to the case of the Chinese-style federalism.

In this paper we attempt to extend the incomplete contract approach toward the government of career-concerned policy-makers. Other than the realistic interest in studying the economic functions of the Chinese-style federalism and/or the Japanese bureaucratic system, the reasons for our attention to policy-makers' career concerns are, first, Seabright's (1996) model cannot apply to lame-duck politicians who have no chance of reelection, and second, not politicians but bureaucrats may be responsible for policy outcomes, especially if we consider effort exerted at the stage of policy administration and implementation.

In the career-concerns model of this paper, policy-makers' hidden talent and unobservable effort, accompanied by region-common and region-specific stochastic disturbances, determine the outcomes of tasks in each region, and policy-makers are rewarded in the future according to the expectation of their hidden talent, which is calculated conditional on realizations of the outcomes and expected effort levels. More specifically, a policy-maker is rewarded the more as she produces the larger outcome relative to the average. Accordingly, policy-makers have incentives to disguise themselves as a more talented one by increasing effort for larger outcomes. The relative performance evaluation also leads them to engage in yardstick competition under decentralization.[2]

The inference of a policy-maker's hidden talent is an application of the classical signal extraction problem. The allocation of tasks between central and regional policy-makers alters the structure of signal and noise, on which the inference of their talent is based, and as a result,

[2]Although the model formulation differs from ours, Besley and Case (1995) is the first paper that modeled and empirically analyzed the effect of yardstick competition on policy-makers' effort in the context of fiscal federalism.

produces effort incentives differentiated between them.

The results and organization of this paper are as follows. It consists of two parts. In the first part we deal with the case of a single task being undertaken in each region and analyze how the integration of regions affects effort incentives. We will show that decentralization always induces larger effort on each task than centralization because effort incentives decrease as a policy-maker undertakes the larger number of tasks. This is the effect of *focus*, termed by Dewatripont et al. (1999). Encouragement of effort does not mean, however, that decentralization is more efficient. Depending on the strength of policy-makers' career concerns, the equilibrium effort levels may be socially excessive. We will show that in some cases partial integration, where regions are grouped into several provinces and a policy-maker appointed in each province has authorities on all tasks in the regions belonging to it, is the most efficient. In the second part, we consider the case of two tasks being undertaken in each region, and analyzes effort incentives under three modes of federalism, full centralization, full decentralization, and partial (de)centralization. Under full centralization the authorities on the two tasks are allocated to the central policy-maker, under full decentralization they are to regional policy-makers, and under partial (de)centralization, they are split between the two types of policy-makers. Applying the technique developed in the first part of this paper, we will show that full decentralization induces larger effort on every task than full centralization. Under partial (de)centralization, there is a trade-off on effort incentives caused by the effects of focus and noise. We will show a sufficient condition for which partial (de)centralization induces larger effort on every task than full centralization and a necessary condition for which it does so than full decentralization. As well as the variances of policy-makers' uncertain talent and disturbances associated with policy outcomes, the number of regions turns out to play a key role. The implications for efficiency are also discussed. All the proofs of propositions are collected in the appendix.

# 2 The case of a single task in each region

## 2.1 The basic model

Consider a nation consisting of $n$ ($\geq 2$) symmetric regions indexed by $i = 1, 2, \ldots, n$. The utility of a representative citizen in region $i$, $y_i$, which is simply referred to as policy outcome or just outcome in what follows, is determined stochastically, depending on the talent of a policy-maker

who has policy-making right in the region, $\theta_i$, and her effort exerted there, $a_i$, as follows:

$$y_i = y(\theta_i, a_i) + \eta_i + \varepsilon,$$

where $\eta_i$ is a disturbance term specific to region $i$, and $\varepsilon$ is one common across regions. To simplify the analysis, we will assume that $y(\cdot)$ is additively separable:

$$y(\theta_i, a) = \theta_i + f(a_i) \tag{1}$$

with $f' > 0$ and $f'' < 0$. To have internal solutions, we assume $\lim_{a\to 0} f'(a) = +\infty$ and $\lim_{a\to +\infty} f'(a) = 0$ as well. The random variables have independent and identical normal distributions. Specifically, a politician's talent distributes with mean $\theta$ and variance $\sigma_\theta^2$, region-specific disturbance has mean zero and variance $\sigma_\eta^2$, and region-common disturbance has also mean zero and variance $\sigma_\varepsilon^2$.

Policy-making is driven by career concerns of policy-makers.[3] In the case of decentralization, there is an independent policy-maker appointed in each region, who we will call a regional policy-maker, while in the case of centralization a single policy-maker, who we will call a central policy-maker, has policy-making rights in all regions.

The policy-making game goes with the following timing of events in the case of decentralization: (i) regional policy-makers choose effort levels independently and $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ is simultaneously determined; (ii) The combination of policy outcomes across regions, $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$, is realized along with the stochastic terms and observed by risk-neutral outsiders, who we will call firms; (iii) Firms competitively bid for hiring policy-makers. They end up offering a wage, $w_i$, to the policy-maker in region $i$, in accordance with the expectation of her talent conditional on policy outcomes realized across regions, $\boldsymbol{y}$, and the anticipated effort levels, $\boldsymbol{a}^e = (a_1^e, a_2^e, \ldots, a_n^e)$.

The policy-making game is the same in the case of centralization except that a single, central policy-maker determines the effort level in each region. This means that in the case of centralization a policy-maker's talent perfectly correlates across regions.[4]

---

[3]The career concerns model have been made use of to describe bureaucratic policy-making in Dewatripont *et al.* (1999a, b) and Alesina and Tabellini (2007a, b).

[4]Since policy-makers' effort is not rewarded today but after they are retired from policy-making, it might be suitable to regard policy-makers in this model as bureaucrats, rather than politicians. Nonetheless, we might be allowed to consider them as office-seeking politicians or lame-duck

Assume that when employing a policy-maker, firms pay a wage proportional to the conditional expectation of her talent, i.e., $w_i = \alpha \mathrm{E}(\theta_i | \boldsymbol{y}, \boldsymbol{a}^e)$, in which we interpret $\alpha > 0$ as the "market value" of a policy-maker's talent, following Alesina and Tabellini (2007a, b). If a policy-maker discounts future rewards with a rate $\delta$ $(0 < \delta < 1)$, then her expected reward at the stage of policy making is reduced to $\delta\alpha \mathrm{E}[\mathrm{E}(\theta_i | \boldsymbol{y}, \boldsymbol{a}^e)]$, where the first expectation is taken over $\boldsymbol{y}$. In what follows we will define $\beta := \delta\alpha > 0$ and interpret it as a parameter representing the strength of a policy-maker's career concerns.

Firms are assumed to know the strength of career concerns as well as the distribution functions of the random variables. They also observe realized policy outcomes, but cannot observe realizations of each random variables or policy-makers' effort levels. They estimate policy-makers' talent holding expectations about the effort levels they exerted, based on the realizations of policy outcomes. Note that there is no asymmetric information between policy-makers and firms with respect to their talent; at the stage of decision making on effort levels, policy-makers are not aware of their own talent either. In such a situation policy-makers try to disguise a more talented one by exerting more effort to make a better policy outcome realized, though the effort levels that firms anticipate coincides with what policy-makers have actually chosen, i.e., $\boldsymbol{a}^e = \boldsymbol{a}$, in the perfect Bayesian Nash equilibrium of the game, on which we will focus our analysis throughout the paper.

## 2.2 The signal extraction problem

Firms competitively bid an after-retirement wage for policy-maker $i$ by inferring her talent based on the information of realized policy outcomes across regions, holding expectations of exerted effort levels. This is a signal extraction problem with two different disturbance terms, and we will solve it by making use of the theorem developed in Degroot (1970, p.167).

To analyze the signal extraction problem in both cases of centralization and decentralization in a unified way as well as to deduce implications about political integration of regions, let us suppose that each $m$ regions are grouped into a province, where a single policy-maker is appointed and has jurisdiction over policies in the regions belonging to it. Here we assume the integer problem away

---

incumbents who like to be thought of as a talented politician for reelection or for good status after their term of office.

for simplicity, and suppose that there are $p$ $(= n/m)$ policy-makers appointed for policy-making in provinces. Centralization is the case of $m = n$ and decentralization is the case of $n = 1$. We will refer to the case of $1 < m < n$ as partial integration.

Let $\theta_k$ be the talent of policy-maker $k$, who is appointed in province $k = 1, 2, \ldots, p$, and $P_k$ be the set of regions belonging to it. If we assume that the expected effort levels are symmetric across regions, i.e., $a_i^e = a^e$ for all $i$, then the following proposition reveals how we can calculate the conditional expectation of a policy-maker's talent as a solution to the signal extraction problem.

**Proposition 1** *Suppose that policy-makers' effort level are symmetric, i.e.,* $a_i^e = a^e$ *for all* $i$*. Then, the conditional expectation of policy-maker* $k$*'s talent is given by*

$$\mathrm{E}(\theta_k \,|\, \boldsymbol{y}, a^e) = \frac{m\sigma_\theta^2}{\sigma_\eta^2 + m\sigma_\theta^2}\left(\overline{y}_k - \frac{n\sigma_\varepsilon^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + m\sigma_\theta^2}\overline{y}\right) + \Theta(m, a^e), \tag{2}$$

*where*

$$\overline{y}_k = \frac{1}{m}\sum_{i \in P_k} y_i, \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,$$

*and*

$$\Theta(m, a^e) = \frac{\theta(\sigma_\eta^2 + n\sigma_\varepsilon^2) - m\sigma_\theta^2 f(a^e)}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + m\sigma_\theta^2}.$$

**Proof**: See the appendix.

According to the formula in (2), the conditional expectation of policy-maker $k$'s talent is affected by realizations of policy outcomes in two ways. One is through changes in what we will call her *relative performance*, representing the term in the parentheses, and the other is through the *response coefficient*, representing the coefficient before the first parenthesis.

Policy-maker $k$'s relative performance is based on the comparison between the provincial average outcome, $\overline{y}_k$, and the national average outcome, $\overline{y}$, except for the case of centralization in which both average outcome coincide. Specifically, a larger national average outcome decreases the relative performance of every policy-maker by its proportion of $n\sigma_\varepsilon^2/(\sigma_\eta^2 + n\sigma_\varepsilon^2 + m\sigma_\theta^2)$. This proportion follows from calculating the conditional expectation of region-common disturbance based on a realization of the national average outcome, as we can see from the proof of proposition 1 provided in the appendix. We may regard it as a measure of how informative the national average outcome, $\overline{y}$, is for inference of region-common disturbance. In fact, the proportion is equal

to the ratio in variance between region-common disturbance and the national average outcome when effort levels are taken as given.[5] Accordingly, a rise in $m$ reduces the impact of $\overline{y}$ on the evaluation of relative performance, for regional policy outcomes correlate more as a result of the larger number of regions having a common policy-maker and thus region-common disturbance explains the smaller part of fluctuation in the national average policy outcome.

The response coefficient in (2), on the other hand, follows from calculation of the expectation of policy-maker $k$'s conditional on a provincial average outcome and region-common disturbance, as shown in the proof provided in the appendix. In fact, the response coefficient is equal to the ratio in variances between the region-specific disturbance and the provincial average outcome when the region-common disturbance and effort levels are taken as given.[6] Given the relative performance, therefore, policy-maker $k$' talent is judged to be the higher as $m$ increases, for her talent uncertainty has come to explain the larger part of fluctuation in the provincial average outcome.

In this model, except the case of centralization, policy-makers face yardstick competition through the relative performance evaluation, which is endogenously built into the determination of their after-retirement wages. This is in sharp contrast to the traditional literature on fiscal federalism, in which citizens' residential choice and factor movement across regions produce competitive environment of local governments. The thrust of yardstick competition on policy-makers' effort incentives depends on how large impacts the national average outcome has on the evaluation of relative performances, and the composition of its variance plays a crucial roles. The response coefficient also affects effort incentives of policy-makers with its impact depending on the composition of the variance of the provincial average outcome.

## 2.3 The efficiency of decentralization

### 2.3.1 Efficient and equilibrium effort levels

Now let us analyze policy-makers' effort choice. We have $m$ symmetric regions grouped into each province, within which a single policy-maker decides how much policy-making effort to exert in

[5]Since $\overline{y} = (m/n)\sum_{k=1}^{p}\theta_k + (1/n)\sum_{i=1}^{n}(f(a_i)+\varepsilon+\eta_i)$, we have $\mathrm{Var}(\overline{y}\,|\,\boldsymbol{a}) = (m\sigma_\theta^2+\sigma_\eta^2)/n+\sigma_\varepsilon^2$, and thus the proportion is equal to $\sigma_\varepsilon^2/\mathrm{Var}(\overline{y})$.

[6]Since $\overline{y}_k = \theta_k + (1/m)\sum_{i\in P_k}(f(a_i)+\eta_i+\varepsilon)$, we have $\mathrm{Var}(\overline{y}_k\,|\boldsymbol{a},\varepsilon) = \sigma_\theta^2+\sigma_\eta^2/m$, and hence the response coefficient is equal to $\sigma_\theta^2/\mathrm{Var}(\overline{y}_k\,|\,\boldsymbol{a},\varepsilon)$.

each region. We will also suppose that the marginal cost of effort is unity so that a policy-maker in province $k$ incurs cost $\sum_{i \in P_k} a_i$.

First of all, as a benchmark, let us consider the efficient effort level, $a^O$, which is defined as the solution that maximizes the sum of expected utilities across regions minus the cost of effort, $\mathrm{E}(\sum_{i=1}^n y_i) - \sum_{i=1}^n a_i$, subject to (1). From the first order condition, it is characterized by

$$f'(a^O) = 1. \tag{3}$$

Consider next policy-maker $k$, who chooses $a_i$, $i \in P_k$, to maximize her expected payoff,

$$\delta\alpha \mathrm{E}(w_k \,|\, a^e) - \sum_{i \in P_k} a_i = \beta \mathrm{E}[\mathrm{E}(\theta_k \,|\, \boldsymbol{y}, a^e) \,|\, a^e] - \sum_{i \in P_k} a_i, \tag{4}$$

subject to (1) and (2), taking other policy-makers' effort levels and firms' expectation as given.[7] Note that the first expectation operator on the right hand side of (4), as well as the one on the left hand side, is applied over policy outcomes.

Let us confine our attention to the symmetric perfect Bayesian Nash equilibria, in which provincial policy-makers exert the same level of effort in every region, and denote by $a_m^*$ the equilibrium effort level. From the first order condition for the maximization problem formulated in (4), we find that $a_m^*$ is uniquely determined by the following condition:

$$\frac{m\sigma_\theta^2}{\sigma_\eta^2 + m\sigma_\theta^2}\left(\frac{1}{m} - \frac{\sigma_\varepsilon^2}{n\sigma_\varepsilon^2 + \sigma_\eta^2 + m\sigma_\theta^2}\right) f'(a_m^*) = \frac{1}{\beta}, \tag{5}$$

which balances the discounted marginal revenue with the marginal cost of effort.[8] The bracketed term on the left hand side shows how an improvement in a region's policy outcome affects the policy-maker's relative performance evaluation. A rise in $y_i$ by one unit increases the relative performance evaluation of the policy-maker in charge of region $i$ through increasing the provincial average of policy outcome by $1/m$, but decreases it through increasing the national average by $1/n$, which hurts her evaluation by $\sigma_\varepsilon^2/(n\sigma_\varepsilon^2 + \sigma_\eta^2 + m\sigma_\theta^2)$.

From (5), we can show some comparative static properties of $a_m^*$.

**Proposition 2** *$a_m^*$ is decreasing in $m$, $\sigma_\eta^2$, and $\sigma_\varepsilon^2$, but increasing in $\sigma_\theta^2$ and $n$.*

---

[7]The participation constraint is assumed to be satisfied with the wages paid for their current position in government.

[8]Since the second term in the braces is less than $1/m$, the second order condition is met.

It is worth paying special attention to how an increase in $m$ affects $a_m^*$. (5) shows that it has two channels with counteracting effects on a policy-maker's effort. One is via the change in the relative performance evaluation, through which a larger $m$ discourages effort. As easily verified, the bracketed term on the left hand side of (5) is decreasing in $m$. This means that an increase in effort improves relative performance evaluation less as $m$ increases. The other is via the change in the response coefficient, through which a larger $m$ encourage effort. Overall, however, the former outweighs the latter, and thus an increase in $m$ discourages effort. If we interpret $m$ as the number of tasks assigned to each policy-maker, then our result can be seen as a statement that an increase in the number of tasks per policy-maker discourages effort. This is what Dewatripont et al.(1999) termed the effect of *focus*, and our result is an extension of theirs into the case with relative performance evaluation.

The effects of other factors on effort are easy to interpret. A rise in $\sigma_\theta^2$ encourages effort since it increases the response coefficient and decreases the effect of a higher national average hurting a policy-maker's relative performance evaluation. On the other hand, an increase in $\sigma_\varepsilon^2$ discourages effort because it increases the latter effect, without affecting the response coefficient. Though a rise in $\sigma_\eta^2$ decreases the response coefficient but also decreases the effect of a higher national average hurting a policy-maker's evaluation, it overall discourages effort since the former effect outweighs the latter. Finally, an increase in the entire number of regions encourages effort because it also reduces the effect of a higher national average hurting a policy-maker's evaluation.

Now we will examine the implications of proposition 2 from the choice of efficient mode of federalism. Let $a^C$ and $a^D$ denote the symmetric equilibrium effort levels chosen in each region under centralization and decentralization respectively. The conditions they satisfy are obtained from (5) by substituting $m = n$ and $m = 1$:

$$\frac{\sigma_\theta^2 f'(a^C)}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + n\sigma_\theta^2} = \frac{1}{\beta} \tag{6}$$

and

$$\frac{\sigma_\theta^2 f'(a^D)}{\sigma_\eta^2 + \sigma_\theta^2}\left\{1 - \frac{1}{n + (\sigma_\eta^2 + \sigma_\theta^2)/\sigma_\varepsilon^2}\right\} = \frac{1}{\beta} \tag{7}$$

Making use of proposition 2 then straightforwardly yields the following proposition.[9]

[9]Proposition 3 is part of the results obtained in the companion paper, Konishi (2010), which deals with the case of a single task generating spillover effects across regions.

**Proposition 3** *Decentralization induces more effort in each region than centralization, i.e.,* $a^D > a^C$.

Needless to say, we have $a^C = a^D$ whenever $n = 1$. Moreover, from (6) and (7), we observe that $a^C$ is decreasing but $a^D$ is increasing in $n$. The former follows from the effect of lost focus, and the latter does from the increasing thrust of yardstick competition.[10]

We will next consider the welfare implications. Compare (5) with (3), and we find that $a^*_m < a^O$ if and only if

$$\beta < \left(m + \frac{\sigma_\eta^2}{\sigma_\theta^2}\right)\left(1 + \frac{m\sigma_\varepsilon^2}{\sigma_\eta^2 + (n-m)\sigma_\varepsilon^2 + m\sigma_\theta^2}\right) := \phi_m, \tag{8}$$

where $\phi_m$ turns out to be increasing in $m$, i.e., $\phi_1 < \phi_2 < \cdots < \phi_n$. Hence, combining these observations with proposition 3, we have the following proposition.

**Proposition 4** *Decentralization is the most efficient if* $\beta \leq \phi_1$*, while centralization is the most efficient if if* $\beta \geq \phi_n$*. In the case of* $\phi_1 < \beta \leq \phi_n$*, partial integration is the most efficient with each province consisting of* $m$ *or* $m+1$ *regions such that* $\phi_m \leq \beta < \phi_{m+1}$.

Proposition 4 shed light on the relative efficiency of centralization versus decentralization, or the optimal size of jurisdictions, from the point of task allocation and effort incentives, in contrast to the traditional arguments, which rest on externalities, economies of scale, and heterogeneity in preferences across regions (See Oates(1972) and Alesina and Spolaore (1997) among others). Our point is that although the effects of focus and yardstick competition provide the larger effort incentives as the government is the more decentralized, they might be large enough to produce excessively large effort levels from the point of efficiency. The market value of policy-makers' talent and their discount rate play a crucial role in determining if decentralization improves efficiency. Moreover, proposition 4 addresses the desirability of partial integration; there are some cases in which partial integration is the most efficient structure of federalism.

[10]More precisely, as we can see from (2), an increase in $n$ affects $a^D$ through two channels. One is through a higher information value of the national average, which discourages effort, and the other is through a lower effect of a unilateral effort increase on the average outcome, which encourages effort. In total, the latter effect outweighs the former, so that an increase in $n$ reduces $a^D$.

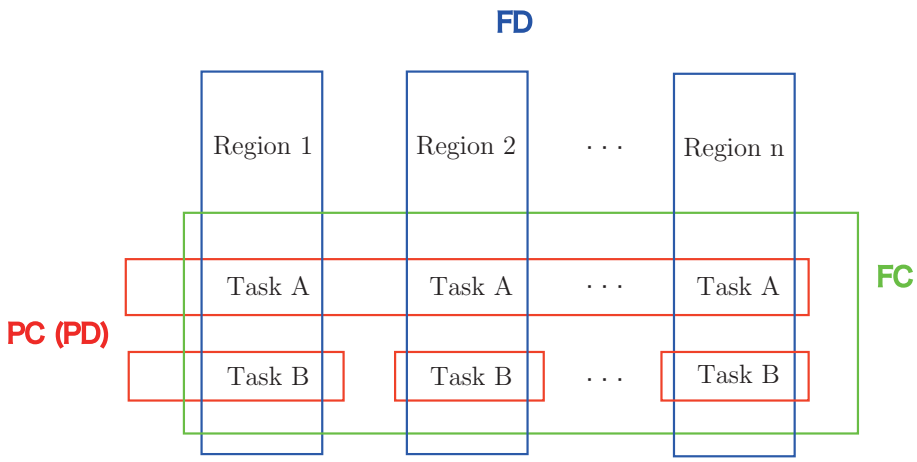


Figure 1: Three modes of federalism

## 3 The case of multiple tasks in each region

We will next examine the case of multiple tasks being performed in each region. To simplify the analysis, we will assume that there are two tasks in each region, whose authorities are potentially allocated between the regional and central policy-makers. An outcome of task $t = A, B$ in region $i$, $y_i^t$, is determined by $y_i^t = \theta_i^t + a_i^t + \eta_i^t + \varepsilon^t$, where $\theta_i^t$ is the talent of a policy-maker who undertakes task $t$ in region $i$, $a_i^t$ is the effort level exerted by the policy-maker, $\eta_i^t$ is region-specific disturbance, and $\varepsilon_i^t$ is region-common disturbance. The random variables follow the same independent and identical normal distributions that we have made use of in the previous section. We further assume that firms can observe only the sum of the policy outcomes in each region, $y_i := y_i^A + y_i^B$, on which evaluation of policy-makers' talent is based. That is, the observable for outsiders is $(y_1, y_2, \ldots, y_n)$, where

$$y_i = \sum_{t=A,B} \theta_i^t + \sum_{t=A,B} f(a_i^t) + \sum_{t=A,B} \eta_i^t + \sum_{t=A,B} \varepsilon^t, \qquad \text{for } i = 1, 2, \ldots, n. \tag{9}$$

Let $\theta_i$ be the talent of a regional policy-maker in region $i$, and $\theta_c$ be that of a central policy-maker. As shown in figure 1, we will distinguish three modes of federalism according to the allocation of authorities over the tasks.

The first is full centralization (FC), in which case two tasks are undertaken solely by the central policy-maker, and hence $\theta_i^A = \theta_i^B = \theta_c$. The second is full decentralization (FD), in which case authorities over two tasks are all allocated to regional policy-makers, and hence $\theta_i^A = \theta_i^B = \theta_i$.

The third is what we call either partial decentralization (PC) or partial centralization (PD), in which case authorities over two tasks are separated between central and regional policy-makers. Let us suppose that the authority over task $A$ is given to the central policy-maker, while that over task $B$ is to regional policy-makers, i.e., $\theta_i^A = \theta_c$ and $\theta_i^B = \theta_i$. We will solve the signal extraction problem in each case and characterize the equilibrium effort levels, confining our attention to symmetric equilibria.

### 3.1 Full centralization versus full decentralization

In the case of full centralization, the policy outcome in region $i$ is determined by (9) with $\sum_{t=A,B} \theta_i^t = 2\theta_c$. Let us denote the symmetric equilibrium effort level on each task by $a^{FD}$. Then, applying proposition 1 and rearranging terms, we obtain the equilibrium condition,

$$\frac{\sigma_\theta^2 f'(a^{FC})}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + 2n\sigma_\theta^2} = \frac{1}{\beta}. \tag{10}$$

For its derivation we can make use of (2), replacing $\theta_k$ with $2\theta_c$, $\sigma_\theta^2$ with $4\sigma_\theta^2$, $\sigma_\eta^2$ with $2\sigma_\eta^2$, and $\sigma_\varepsilon^2$ with $2\sigma_\varepsilon^2$.

Comparing with the corresponding equilibrium condition in the single-task case, (6), we find an only difference that the coefficient of $\sigma_\theta^2$ in the denominator on the left hand side is changed from $n$ to $2n$. This is because a policy-maker's talent correlates perfectly between tasks, whereas other disturbances are independent.[11]

Next consider the case of full decentralization. Let us denote the symmetric equilibrium effort on each task by $a^{FD}$. Applying proposition 1 to (9) with $\theta_i^A + \theta_i^B = 2\theta_i$ and arranging terms, we then obtain

$$\frac{\sigma_\theta^2 f'(a^{FD})}{\sigma_\eta^2 + 2\sigma_\theta^2} \left\{ 1 - \frac{1}{n + (\sigma_\eta^2 + 2\sigma_\theta^2)/\sigma_\varepsilon^2} \right\} = \frac{1}{\beta}. \tag{11}$$

To derive this condition, we can makes use of (2), replacing $\theta_k$ with $2\theta_i$, $\sigma_\eta^2$ with $2\sigma_\eta^2$, $\sigma_\theta^2$ with $4\sigma_\theta^2$, and $\sigma_\varepsilon^2$ with $2\sigma_\varepsilon^2$. Similar to (10), because each regional policy-maker undertakes two tasks, the coefficient of $\sigma_\theta^2$ in each denominator on the left hand side is doubled, as compared to the corresponding equilibrium condition in the single-task case, (7).[12]

[11] Notice that (10) coincides with (6) when we replace $\sigma_\varepsilon^2$ with $\sigma_\varepsilon^2 + \sigma_\theta^2$ in the latter. Then, $a^{FC} < a^C$ follows from proposition 2.

[12] Notice that (11) coincides with (7) when we substitute $\sigma_\eta^2 + \sigma_\theta^2$ for $\sigma_\eta^2$ in the latter. Then, applying proposition 2, $a^{FD} < a^D$ follows.

Now compare full centralization and full decentralization in terms of the equilibrium effort level, and we obtain the following proposition.

**Proposition 5** *Full decentralization induces more effort on each task in each region than full centralization, i.e.,* $a^{FD} > a^{FC}$.

The reason is basically the same with proposition 3, because the essential difference between full decentralization and full centralization is the number of tasks allocated to a policy-maker as in the analysis of the single-task case.

### 3.2 Partial (de)centralization

In the case of partial (de)centralization, two types of policy-maker, central and regional, face different incentive structures, which we can attribute to the effective changes in region-common and region-specific disturbance.

From the point of estimating the central policy-maker's talent, partial (de)centralization changes the fluctuation in policy outcome in two ways, relative to the single-task case. One is that the variances of region-specific disturbance is doubled as well as region-common disturbances due to the introduction of a new task allocated to the authority of regional policy-makers. The other is that regional policy-makers' talent uncertainty plays the same role as region-specific disturbance in the estimation of the central policy-maker's talent.

Let us denote by $a^{PC}$ the symmetric equilibrium effort level chosen on task $A$ by the central policy-maker under partial (de)centralization. To obtain the equilibrium condition, we can make use of (6), replacing $\sigma_\eta^2$ with $2\sigma_\eta^2 + \sigma_\theta^2$ and $\sigma_\varepsilon^2$ with $2\sigma_\varepsilon^2$. Then, rearranging terms yields

$$\frac{\sigma_\theta^2 f'(a^{PC})}{2\sigma_\eta^2 + 2n\sigma_\varepsilon^2 + (n+1)\sigma_\theta^2} = \frac{1}{\beta}. \tag{12}$$

Comparing this with the counterpart in the single-task case, (6), the denominator in (12) turns out to be larger by the amount of $\sigma_\eta^2 + n\sigma_\varepsilon^2 + \sigma_\theta^2$. This is owing to the additional disturbances introduced along with new tasks allocated to regional policy-makers.

Now compare the equilibrium effort levels between full centralization and partial (de)centralization. From (10) and (12), we obtain the following proposition.

**Proposition 6** *With respect to task A, (i) partial (de)centralization induces more effort than full centralization, i.e., $a^{PC} > a^{FC}$, if and only if $\sigma_\theta^2 > \sigma_\varepsilon^2$ and $n > \hat{n} := (\sigma_\eta^2 + \sigma_\theta^2)/(\sigma_\theta^2 - \sigma_\varepsilon^2)$; (ii) partial (de)centralization induces less effort than full decentralization, i.e, $a^{PC} < a^{FD}$.*

With respect to the comparison between $a^{PC}$ and $a^{FC}$, if we look at the denominators on the left hand sides of (10) and (12), we notice that there is a trade-off between an improvement in focus and an increase in noise. More specifically, when full centralization is converted to partial (de)centralization, effort is encouraged through the effect of focus, as the denominator on the left hand side decreases by the amount of $n\sigma_\theta^2$. On the other hand, additional region-specific disturbance produced by the allocation of new tasks to regional policy-makers discourages effort, as the denominator increases by the amount of $\sigma_\eta^2 + n\sigma_\varepsilon^2 + \sigma_\theta^2$. Thus, $\sigma_\theta^2 > \sigma_\varepsilon^2$ is necessary for partial (de)centralization to induce more effort than full centralization, and provided that $\sigma_\theta^2 > \sigma_\eta^2$, the more likely we have $a^{PD} > a^{FD}$, the larger is the number of regions, since the effect of focus increases. The reason for $a^{PC} < a^{FD}$ is that as $n$ increases, the yardstick competition strengthens the effort incentive under full decentralization, while the noise effect decreases it under partial decentralization.[13]

Turn to the effort incentives of regional policy-makers on task $B$. In inferring their talent under partial (de)centralization, not only region-specific and region-common disturbances, but also the central policy-maker's talent is added as a new source of region-common uncertainty. Accordingly, we can use proposition 1 to partial (de)centralization by replacing $\sigma_\eta^2$ with $2\sigma_\eta^2$ and $\sigma_\varepsilon^2$ with $2\sigma_\varepsilon^2 + \sigma_\theta^2$. If we denote by $a^{PD}$ the equilibrium effort level that each regional policy-maker chooses in the symmetric equilibrium, invoking (7), we obtain the following equilibrium condition,

$$\frac{\sigma_\theta^2 f'(a^{PD})}{2\sigma_\eta^2 + \sigma_\theta^2}\left\{1 - \frac{1}{n + (\sigma_\eta^2 + \sigma_\theta^2/2)/(\sigma_\varepsilon^2 + \sigma_\theta^2/2)}\right\} = \frac{1}{\beta}. \tag{13}$$

Comparing this with the equilibrium condition for effort made under decentralization in the

[13]Specifically, let $a^{FD}(n)$ and $a^{PC}(n)$ be the equilibrium effort levels under full decentralization and partial (de)centralization respectively, and $a^C(n)$ be the one under centralization in the single-task case with $n$ regions. Then, comparing the first order conditions, we first notice that $a^{FD}(1) = a^C(2)$ because of the two tasks being symmetric. Similarly, $a^{PC}(2) < a^C(2)$ holds owing to an increase in noise. Finally, if we notice that as $n$ increases, $a^{FD}(n)$ increases owing to the force of yardstick competition while $a^P C(n)$ decreases owing to the loss of focus, we have $a^{PC}(n) < a^{FD}(n)$ for all $n \geq 2$.

single-task case, (7), we notice that effort is discouraged due to increases in region-specific and region-common disturbances.

The next comparison is between $a^{FD}$ and $a^{PD}$. As in the comparison between $a^{FC}$ and $a^{PC}$, there potentially occurs a trade-off in the effects on effort incentives between focus and noise. From (11) and (13), we can show the following proposition.

**Proposition 7** *With respect to task B, (i) there exists $\tilde{n}$ such that partial (de) centralization induces more effort than full centralization, i.e., $a^{PD} > a^{FC}$, if and only if $n > \tilde{n}$; (ii) there exists $\overline{n}$ such that partial (de)centralization induce more effort than full decentralization, i.e., $a^{PD} > a^{FD}$, if and only if $\sigma_\theta^2 > \sigma_\eta^2$ and $n > \overline{n}$.*

With respect to the comparison between $a^{PD}$ and $a^{FC}$, notice that a regional policy-maker exerts more effort under partial (de)centralization as $n$ increases, because the expansion of noise reduces the thrust of yardstick competition.[14] On the other hand, a central policy-maker exerts less effort under full centralization as $n$ increases, because she undertakes the larger number of tasks. With respect to the comparison between $a^{PD}$ and $a^{FD}$, as we can see from the comparison between the terms in the braces in (11) and (13), the expected marginal effect of effort on the relative performance turns out lower under partial (de)centralization than under full decentralization. This is because the fluctuation in the national average outcome has come to be explained in a larger part by region-common disturbance including the uncertainty of the central policy-maker's talent. Accordingly, $\sigma_\theta^2 > \sigma_\eta^2$ is necessary for partial (de)centralization to induce more effort than full decentralization; a condition which guarantees a larger response coefficient in (13) than in (11). Further, if $n$ is sufficiently large, then the term in the braces in (13) discussed above has only a negligible difference from that in (11). This is why $a^{FD} < a^{PD}$ holds when the number of regions is sufficiently large[15].

[14] An increase in the number of regions raises the information value of the national average in estimating each regional policy-maker's talent, but at the same time it makes the expected value of the average outcome less responsive to each policy-maker's effort. As we can see from (2), the latter effect outweighs the former, so that a larger number of regions leads regional policy-makers to exert less effort through yardstick competition.

[15] We can consider another mode of federalism, the U-form decentralization, in which a different policy-maker undertakes the same task in every region, and show that the equilibrium effort under

Finally, we will discuss the welfare implications. As is similar to the analysis in proposition 4, the strength of policy-makers' career concerns, represented by parameter $\beta$, is crucial in determining whether an equilibrium effort level falls short of the efficient level or not. Suppose that $\beta$ is so low that every equilibrium effort level is insufficient from the point of efficiency. Then, from proposition 5, full centralization is always less efficient than full decentralization. From propositions 6(i) and 7(i), it is also less efficient than partial (de)centralization if $\sigma_\theta^2 > \sigma_\varepsilon^2$ and $n > \max\{\hat{n}, \tilde{n}\}$. From propositions 6(ii) and 7(ii), full decentralization is the most efficient if $\sigma_\theta^2 \leq \sigma_\eta^2$ or $n \leq \overline{n}$; otherwise, as compared to full decentralization, partial (de)centralization produces a trade-off in that it encourages effort on the task allocated to regional policy-makers, while discouraging it on the task allocated to the central policy-maker.

# 4 Concluding remarks

This paper analyzed the effect of decentralization in government on effort incentives for policy-making in a model of yardstick competition among career-concerned policy-makers. We found some trade-offs in effort incentives occurring due to the changes in focus and noise through decentralization, and revealed the conditions that determine the relative efficiency among several mode of federalism. Nonetheless, our analysis is just a starting point and important extensions, such as the case of asymmetric multiple tasks, introduction of externalities of effort across regions and collusion of regional policy-makers, are left for further research.

---

U-form decentralization always falls short of one under full centralization since the noise effect discourages effort more than the focus effect.

## Appendix

### Proof of proposition 1

To derive (2), we will use the following theorem, which is borrowed from Theorem 1 in Degroot (1970, p.167).

**Theorem** *Suppose that two independent stochastic variables, $x$ and $z$, distribute normally such that $x \sim \mathcal{N}(z, \sigma_x^2)$ and $z \sim \mathcal{N}(z_0, \sigma_z^2)$, where $z_0$, $\sigma_x^2$ and $\sigma_z^2$ are known parameters. Then, given $n$ observations of $x$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, the conditional expectation of $z$ is given by*

$$E(z \mid \boldsymbol{x}) = \frac{\sigma_x^2}{\sigma_x^2 + n\sigma_z^2} z_0 + \frac{n\sigma_z^2}{\sigma_x^2 + n\sigma_z^2} \left( \frac{1}{n} \sum_{i=1}^{n} x_i \right). \tag{A.1}$$

Now we will prove proposition 1. Consider region $i$ included in province $k$, i.e., $i \in P_k$. The policy outcome there is determined by

$$y_i = \theta_k + a^e + \eta_i + \varepsilon, \tag{A.2}$$

where we assume a symmetric effort level, $a^e$. If $\varepsilon$ is given, then we can suppose that the random variable $y_i - \varepsilon$ follows $\mathcal{N}(\theta_k + a^e, \sigma_\eta^2)$ and $\theta_k + a^e$ does $\mathcal{N}(\theta + a^e, \sigma_\theta^2)$. Since we have $m$ observations of $y_i - \varepsilon$ realized within the province, applying the above theorem demonstrates

$$E(\theta_k + a^e \mid \boldsymbol{y}, a^e, \varepsilon) = \frac{\sigma_\eta^2}{\sigma_\eta^2 + m\sigma_\theta^2}(\theta + a^e) + \frac{m\sigma_\theta^2}{\sigma_\eta^2 + m\sigma_\theta^2}(\overline{y}_k - \varepsilon),$$

and taking its expectation over $\varepsilon$ and arranging terms, we have

$$E(\theta_k \mid \boldsymbol{y}, a^e) = \frac{\theta\sigma_\eta^2}{\sigma_\eta^2 + m\sigma_\theta^2} + \frac{m\sigma_\theta^2}{\sigma_\eta^2 + m\sigma_\theta^2} \left\{ \overline{y}_k - a^e - E(\varepsilon \mid \boldsymbol{y}, a^e) \right\} \tag{A.3}$$

for $k = 1, 2, \ldots, p$.

To obtain the conditional expectation of $\varepsilon$, suppose that a combination of policy-makers' talent, $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$, is given. Then, from (A.2), the random variable $y_i - \theta_k$ follows $\mathcal{N}(\varepsilon + a^e, \sigma_\eta^2)$ and $\varepsilon + a^e$ does $\mathcal{N}(a^e, \sigma_\varepsilon^2)$. Because we have $n$ observations of $y_i - \theta_k$ realized across regions, applying the above theorem yields

$$E(\varepsilon + a^e \mid \boldsymbol{y}, a^e, \boldsymbol{\theta}) = \frac{\sigma_\eta^2 a^e}{\sigma_\eta^2 + n\sigma_\varepsilon^2} + \frac{n\sigma_\varepsilon^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2} \left( \overline{y} - \frac{1}{p} \sum_{k=1}^{p} \theta_k \right).$$

Hence, taking expectation over $\boldsymbol{\theta}$ and arranging terms, we have

$$E(\varepsilon \mid \boldsymbol{y}, a^e) = \frac{n\sigma_\varepsilon^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2}(\overline{y} - a^e) - \frac{n\sigma_\varepsilon^2}{p(\sigma_\eta^2 + n\sigma_\varepsilon^2)} \sum_{k=1}^{p} E(\theta_k \mid \boldsymbol{y}, a^e). \tag{A.4}$$

Finally, we obtain

$$E(\varepsilon \,|\, \boldsymbol{y}, a^e) = \frac{n\sigma_\varepsilon^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + m\sigma_\theta^2}(\overline{y} - \theta - a^e)$$

and

$$E(\theta_k \,|\, \boldsymbol{y}, a^e) = \frac{m\sigma_\theta^2}{\sigma_\eta^2 + m\sigma_\theta^2}\left(\overline{y}_k - \frac{n\sigma_\varepsilon^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + m\sigma_\theta^2}\overline{y}\right) + \Theta(m, a^e)$$

by combining (A.3) and (A.4) and rearranging terms. ||

**Proof of proposition 2**: From (5), since the coefficient of $f'(a_m^*)$ turns out to be increasing in $\sigma_\theta^2$ and $n$, while decreasing in $\sigma_\varepsilon^2$, it follows that $a_m^*$ is increasing in $\sigma_\theta^2$ and $n$, while decreasing in $\sigma_\varepsilon^2$. Similarly, since we can rewrite (5) into

$$\frac{\sigma_\theta^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + m\sigma_\theta^2}\left\{1 + \frac{(n-m)\sigma_\varepsilon^2}{\sigma_\eta^2 + m\sigma_\theta^2}\right\} f'(a_m^*) = \frac{1}{\beta},$$

we can verify that $a_m^*$ is decreasing in $m$ and $\sigma_\eta^2$. ||

**Proof of proposition 3**: Since $a_m^*$ is decreasing in $m$ from proposition 2, and by definition $a^D = a_1^*$ and $a^O = a_n^*$, $a^O < a^D$ follows straightforwardly. ||

**Proof of proposition 4**: From the definition of $\phi_m$ given in (8), it is straightforward that $a^C < a^D \leq a^O$ if $\beta \leq \phi_1$, and $a^O < a^C < a^D$ if $\beta \geq \phi_n$. If $\phi_1 < \beta < \phi_n$, there exists $m \in [1, n]$ such that $\phi_m \leq \beta < \phi_{m+1}$, which implies that $a_{m+1}^* < a^O \leq a_m^*$. ||

**Proof of proposition 5**: Rearranging terms on the right hand sides of (10) and (11) respectively yields

$$\frac{\sigma_\theta^2}{(\sigma_\eta^2 + n\sigma_\varepsilon^2)/n + 2\sigma_\theta^2} f'(a^{FC}) = \frac{1}{\beta}$$

and

$$\frac{\sigma_\theta^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + 2\sigma_\theta^2}\left\{1 + \frac{(n-1)\sigma_\varepsilon^2}{\sigma_\eta^2 + 2\sigma_\theta^2}\right\} f'(a^{FD}) = \frac{1}{\beta}. \tag{A.5}$$

Then, comparing the coefficients of $f'(a^{FC})$ and $f'(a^{DC})$, we have $a^{FD} > a^{FC}$. ||

**Proof of proposition 6**: (i) Rewrite (11) into (A.5) and compare it with (12). Then, because the denominators in the first terms on the left hand sides satisfy $2\sigma_\eta^2 + 2n\sigma_\varepsilon^2 + (n+1)\sigma_\theta^2 > \sigma_\eta^2 + n\sigma_\varepsilon^2 + 2\sigma_\theta^2$,

$a^{FD} > a^{PC}$ follows. (ii) Compare (10) with (12), and it then follows that $a^{PC} > a^{FC}$ if and only if $2\sigma_\eta^2 + 2n\sigma_\varepsilon^2 + (n+1)\sigma_\theta^2 > \sigma_\eta^2 + n\sigma_\varepsilon^2 + 2n\sigma_\theta^2$, the condition of which is reduced to $\sigma_\theta^2 > \sigma_\varepsilon^2$ and $n > (\sigma_\theta^2 + \sigma_\eta^2)/(\sigma_\theta^2 - \sigma_\varepsilon^2)$. ||

**Proof of proposition 7**

(i) Comparing the left hand sides of (10) and (13) and rearranging terms, it follows that $a^{PD} > a^{FC}$ if and only if

$$1 - \frac{(3n-1)\sigma_\theta^2}{\sigma_\eta^2 + n\sigma_\varepsilon^2 + 2n\sigma_\theta^2} < \frac{(\sigma_\varepsilon^2 + \sigma_\theta^2)(n-1)}{\sigma_\eta^2 + \sigma_\theta^2/2}.$$

Since the left hand side monotonically decreases and tends to $(\sigma_\varepsilon^2 - \sigma_\theta^2)/(\sigma_\varepsilon^2 + 2\sigma_\theta^2) < 1$, taking a positive value at $n = 1$, whereas the right hand side monotonically increases without bound, taking zero at $n = 1$. This guarantees a unique existence of $\tilde{n}$ such that $a^{PD} > a^{FC}$ if and only if $n > \tilde{n}$. (ii) Comparing the left hand sides of (11) and (13) and rearranging terms, it is straightforward to see that $a^{FD} < a^{PD}$ if and only if

$$\frac{2\sigma_\eta^2 + \sigma_\theta^2}{\sigma_\eta^2 + 2\sigma_\theta^2} < \frac{\sigma_\eta^2 + n\sigma_\varepsilon^2 + 2\sigma_\theta^2}{\sigma_\eta^2 + (n-1)\sigma_\varepsilon^2 + 2\sigma_\theta^2} \frac{2\sigma_\eta^2 + 2(n-1)\sigma_\varepsilon^2 + n\sigma_\theta^2}{2\sigma_\eta^2 + 2n\sigma_\varepsilon^2 + (n+1)\sigma_\theta^2}. \tag{A.6}$$

If we denote the right hand side of (A.6) by $q$, we can rewrite it into

$$q = \left(1 + \frac{1}{n+b-1}\right)\left(1 - \frac{1}{n+c}\right) = 1 - \frac{b-c}{(n+b-1)(n+c)},$$

where $b = (\sigma_\eta^2 + 2\sigma_\theta^2)/\sigma_\varepsilon^2$ and $c = (\sigma_\eta^2 + \sigma_\theta^2/2)/(\sigma_\varepsilon^2 + \sigma_\theta^2/2)$. Then, because $b > c > 0$, it follows that $q < 1$ and $q$ monotonically converges to unity as $n$ increases. On the other hand, the left hand side of (A.6) is less than unity if and only if $\sigma_\theta^2 < \sigma_\eta^2$. ||

## References


Alesina, A. and Spolaore, E. (1997), On the number and size of nations, *Quarterly Journal of Economics*, pp. 1027–1056.

Alesina, A. and Tabellini, G. (2007a), Bureaucrats or politician? part I: a single policy task, *American Economic Review* 97, pp.169–179.

Alesina, A. and Tabellini, G. (2007b), Bureaucrats or politician? part II: multiple policy tasks, *Journal of Public Economics* 92, pp.426–447.

Besley, T. and Case, A. (1995a), Incumbent behavior: vote-seeking, tax-setting, and yardstick competition, *American Economic Review* 85, pp.25–45.

Bewley, T. (1981), A critique of Tiebout's theory of local public expenditures, *Econometrica* 49, pp.713–739.

Blanchard, O. and Shleifer, A. (2001), Federalism with and without political centralization: China versus Russia, *IMF Staff Papers* 48, pp.171–179.

Degroot, M.H. (1970), *Optimal Statistical Decisions*, McGraw-Hill.

Dewatripont, M., Jewitt, I., and Tirole, J. (1999a), The economics of career concerns, part I: comparing information structures, *Review of Economic Studies* 66, pp.183–198.

Dewatripont, M., Jewitt, I., and Tirole, J. (1999b), The economics of career concerns, part II: application to missions and accountability of government agencies, *Review of Economic Studies* 66, pp.199–217.

Hayek, F.A. (1945), The use of knowledge in society, *American Economic Review* 4, pp.519–530.

Inoki, T. (2001), Staff loans and transfers among central and local governments in Japan, in Muramatsu, M., Igbal, F., and Kume, I. (eds.), *Local Government Development in Post-War Japan*, Oxford University Press.

Konishi, H. (2010), Yardstick competition and the optimal number of jurisdictions in a political economy of fiscal federalism, presented at IIPF annual meeting, 2010.

Li, H. and Zhou, L.-A. (2005), Political turnover and economic performance: the incentive role of personnel control in China, *Journal of Public Economics* 89, pp.1743–1762.

Maskin, E., Qian, Y., and Xu, C. (2000), Incentives, information, and organizational form, *Review of Economic Studies* 67, pp.359–378.

Montinola, G., Qian, Y., and Weingast, B.R. (1995), Federalism, Chinese style: the political basis for economic success in China, *World Politics* 48, pp.50-81.

Oates, W.E.(1972), *Fiscal Federalism*, Harcourt Brace Jovanovich, New York.

Sand-Zantman, W. (2004), Economic integration and political accountability, *European Economic Review* 48, pp.1001–1025.

Seabright, P. (1996), Accountability and decentralization in government: an incomplete contracts approach, *European Economic Review* 40, pp.61–89.

Tiebout, C.(1956), A pure theory of local public expenditure, *Journal of Political Economy*, 64, pp.416–424.

Weingast, B.R. (1995), The economic role of political institutions: market-preserving federalism and economic development, *Journal of Law, Economics, and Organization* 11, pp.1–31.