Interregional Disparities in Productivity
and the Choice of Fiscal Regime[*]

Kimiko Terai[†]

Abstract

Two districts with divergent productivity levels engage in policy-making on the provision of local public goods that enhance future income and hence create a dynamic linkage across periods. The policy choices of district representatives are derived under alternative fiscal systems, and the problem of system selection is examined. It is shown that a decentralized system is more likely to be selected in a more equal society. On the other hand, when a great deal of benefit spills over from a local public good, or when policy makers are expected to care solely about the immediate effects of their decisions on their districts, a centralized system is more likely to be selected.

Keywords: Interregional and intergenerational spillovers; Decentralization; Centralization; Disparity in productivity; Dynamic political economy model

JEL classification: H23, H41, H73, H77

[†]Associate Professor, Hosei University, and Visiting Researcher, University of California, Irvine. Correspondence to: Kimiko Terai, Department of Economics, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697, USA. E-mail: terai@hosei.ac.jp

1 Introduction

There is a rich literature exploring how best to provide local public goods. The pioneering work of Oates (1972) established a standard informal method for analyzing such problems. Oates' Decentralization Theorem states that in the face of heterogeneous regional preferences, local governments can attain greater efficiency by setting the level of public goods individually than can a central government by providing a uniform level of public goods. This theorem holds "for a public good—the consumption of which is defined over geographical subsets of the total population" (p. 35). Indeed, Oates argues that if the optimal-sized locality to internalize all the benefits and costs associated with a particular public good cannot be determined, loss of welfare will result. It is now well understood that independent local governments, which represent the interests of citizens in their jurisdiction, generally fail to provide the optimal level of local public goods that will internalize spillovers into other regions; this implies that the central government is expected to correct the problem.[1]

Many studies have been based on Oates' arguments. Besley and Coate (2003), in a static framework, deduce that higher interregional spillovers will make a centralized system more efficient in that aggregate surplus is maximized. Without spillovers, on the other hand, a decentralized system will be more efficient. These theoretical results support Oates' Decentralization Theorem. Lockwood (2002) continues this line of inquiry by analyzing alternative legislative decision-making rules and the fiscal regime selected. The present study shares a similar motivation, but analyzes district productivity as the principal indicator of interregional heterogeneity. We explore a causal relationship between productivity disparities and the fiscal regime selected in a dynamic framework, and find results that qualitatively differ from previous studies.

---

[1]See, e.g., Wildasin (1986) and Wellisch (2000), who provide comprehensive surveys on the benefits and problems of decentralization.

The aim of this study is to evaluate the relative merits of decentralization and centralization. In a decentralized system, a government in each district individually sets a tax rate on income to finance public expenditure on the local public good. In a centralized system, the central government, which comprises delegates from the districts, sets a uniform tax rate and provides a uniform level of local public goods across the districts. The centralized system, although costly, enables districts to share funds, and is further expected to implement a cooperative solution that can maximize the welfare of the entire society.

Suppose that there are two districts that differ from each other in their productivity. A local public good supplied in any period enhances the next-period income of citizens in the home district, and also raises the next-period income of citizens in the neighboring district provided that it generates *interregional benefit spillovers*. A local public good in a district also has indirect effects on the income of future citizens in this district. The increase in the next-period income in the home or neighboring district influences the local public good level in that period, which in turn influences income in the following period. This dynamic linkage will expand indefinitely. These effects of a given district's local public good provision upon its future citizens are referred to as *intergenerational benefit spillovers*. We first examine policy choice under alternative fiscal regimes when representatives are shortsighted and care only about the effects of their decisions on their citizens' income in the next period. The policy choices derived from this model analysis are in line with the current literature. We then extend the representatives' horizon and show that a tax rate chosen by representatives with longer sight is higher, since they internalize intergenerational benefit spillovers into the farther future.

The choice between the two fiscal systems is examined on the basis of the expected policy decisions by shortsighted representatives. Our main result is to analytically predict that unanimous consent for a decentralized (centralized) fiscal regime is more often observed in societies with a smaller (larger) productivity dis-

3

parity and a lower (higher) degree of interregional benefit spillovers. When a great deal of benefit spills over from a local public good, financially supporting a low-productivity district is also beneficial to a high-productivity district. In contrast, with a low degree of benefit spillovers, a high-productivity district is unwilling to invest in a low-productivity district. Moreover, the low-productivity district loses less by decentralization if interregional divergence is small. It is well documented both theoretically and empirically that decentralization (centralization) is promoted by higher (lower) income levels,[2] but no previous analysis has shown a theoretical link between interregional disparities in productivity and the choice of fiscal regime.[3] It is important to focus on disparity in studying the political decision-making process in the presence of conflicting interests.

This model further shows a link between representatives' time horizon and the system selected. When policy makers consider future generations, decentralization may be preferred over centralization, even with complete interregional spillovers. Each farsighted policy maker allows for the long-term benefits of increasing income in the other district, even under a decentralized system, which suppresses the usual welfare loss resulting from decentralization. Farsighted policy choice is thus a partial solution to interregional externality problems.[4] In other words, when each policy-

---

[2]For instance, see Panizza (1999) and Arzaghi and Henderson (2005). These authors construct theoretical models and demonstrate empirically the negative correlation between per capita income and centralization.

[3]Among the empirical works, Greene (1985) and Hayashi and Sato (2005) demonstrate the negative effects of income inequality as measured by the Gini coefficient on decentralization. Greene (1985) also considers the range of median family income across regions as a measure of interregional income disparity and obtains its estimated negative coefficient. In contrast, Pommerehne's (1977) empirical analysis reveals an inverse relation between interregional income disparity and *centralization*. He points out, however, that the sample size is too small to regard this result "as safely established" (p. 305).

[4]Wellisch and Richter (1995) show theoretically that decentralization induces internalization of intergenerational pollution externalities, since the future adverse environment causing emigration of residents is capitalized into the current property price, which is the concern of the current local

maker is expected to make a decision when caring only about its immediate effects, an institution to promote policy coordination among districts will be established. These results therefore predict progress towards decentralization in an environment with fiscal equalization policies and foresighted politicians.

A number of researchers have discussed whether decision-making on public expenditure should be centralized or decentralized. Models closely related to ours can be divided into three strands. The most closely related consider which system is superior in solving the problems of interregional externalities, including Lockwood (2002) and Besley and Coate (2003), as mentioned, as well as Ellingsen (1998), Schwager (1999), Alesina, Angeloni, and Etro (2005), and Rubinchik-Pessach (2005).[5] There is also a series of papers that, employing the standard capital tax competition framework, analyze the optimal design of a fiscal system (see, for example, Bucovetsky, Marchand, and Pestieau, 1998; Brueckner, 2004; Wilson and Janeba, 2005). In models belonging to these two strands, however, public goods directly affect citizens' utility, and different preferences for public goods are stated as the reason for heterogeneity among districts, whereas the present model assumes that districts have the same preferences for public goods: districts differ only in productivity, and public goods serve as public inputs.

Finally, the present model contributes to the development of dynamic models incorporating public expenditure. Oates (1993) mentions the possibility of a decentralized system to facilitate economic development that tailors public projects to suit local needs. Bénabou (1996) and Brueckner (2006) examine the relation between systems and economic growth by considering investment in human capital

---

agency. Their model is, however, based on the assumption of no interregional spillovers to leave future generations a place to escape.

[5]In spite of their different motivations, the study of Bolton and Roland (1997) should be mentioned because they also examine fiscal regime design in a two-district model. They describe interregional conflicts over redistribution policies and the resulting breakup of the nation in a static framework.

as the engine driving growth. Nishimura (2006) focuses on the relationship between complementarity in local public good provision and economic growth under alternative systems. The studies of Besley and Coate (1998) and Battaglini and Coate (2005) are similar to the present one in that benefits from public goods directly produce a dynamic linkage, but they do not consider fiscal decentralization. The recent study of Kempf and Rossignol (2005), which extends endogenous growth models with the public input of Barro (1990) and Alesina and Rodrik (1994), also uses a two-district model with heterogeneous income. Although they consider the problem of integration, they do not describe interregional benefit spillovers or the explicit linkage between disparity and system design.

In Section 2, formal definitions of each fiscal system and citizen preferences are set out. In Section 3, policy choices by shortsighted representatives are derived for each fiscal regime, and then the choice of fiscal regime is considered for any pair of productivity disparity and the degree of interregional benefit spillovers. Section 4 extends the horizon of representatives and examines the relationship between their consideration of future generations and system design. Section 5 presents conclusions. Technical details are set out in the Appendix.

2 The Model

2.1 Policy Choice

We define a decentralized fiscal regime as a system in which all decisions on financing and supplying local public goods are made by local governments. Under a centralized fiscal regime, these decisions are made by a central government comprising delegates from all districts that is supposed to impose a uniform tax rate and to provide the same level of local public goods across districts. Below, the superscripts d and c stand for decentralization and centralization.

There are two geographically distinct districts, denoted A and B. Each district

has a continuum of citizens with a population of one. We suppose immobility of citizens across districts. The population of a given district $i \in \{A, B\}$ is internally homogeneous, so that each citizen has an identical level of income. Let x_i denote a constant and district-specific level of productivity, which is heterogeneous across districts. The level of productivity is determined by geographic and historical factors of the district, e.g., whether the district is convenient for transportation and rich in natural resources; whether it is politically stable; and whether it has well-functioning economic and political institutional structures. Suppose without loss of generality that $0 < x_A < x_B$, so that $1 < \frac{x_B}{x_A} < \infty$.

Citizens' real income in a given period depends on their productivity level, as well as the levels of local public goods supplied in both districts during the previous period:

$$y_{is} = x_i \Delta_{is}, \tag{1}$$

where y_{is} denotes the real income of citizens in district i and period s; the factor Δ_{is} is given by

$$\Delta_{is} \equiv g_{i,s-1}^{\alpha} g_{j,s-1}^{\alpha\gamma}, \tag{2}$$

where $g_{i,s-1}$ and $g_{j,s-1}$ denote the levels of local public goods supplied in districts i and j during period $s-1$. In this model, $g_{i,s-1}$ and $g_{j,s-1}$ are therefore public inputs that enhance real income in the next period s. The positive-valued parameter α is a measure of the degree of benefit produced by the home public input, and the parameter $\gamma \in [0, 1]$ indicates the extent of benefit spillovers, i.e., how it affects real income in the neighboring district in comparison with the home district. If $\gamma = 0$, then each local public good produces no benefit spillovers to the neighboring district, whereas if $\gamma = 1$, the benefit of one district's public good spills over completely. In the context of this model, districts are defined as local entities organized by a national government if a centralized system is executed. Local public goods generating interregional benefit spillovers include investments in infrastructure, ed-

ucation, and preservation of the environment. We could equally well interpret each district as a nation and the centralized government as an international institution, provided it can control the provision of global public goods in each nation. In this context, public goods might include peace and security, research, economic development and stabilization, as well as those cited above.[6]

Assumption 1. $0 < \alpha < \frac{1}{2}$.

Assumption 2. $\Delta_{A0} = \Delta_{B0} = \Delta_0$.

Assumption 1 ensures that Δ_{is} converges to a steady-state level, and allows us to compare citizen welfare under alternative fiscal regimes. Assumption 2 provides us with initial conditions. The transitional dynamics of Δ_{is} and its steady-state level under Assumptions 1 and 2 are detailed in Appendix A.1 using policy outcomes derived in Sections 3 and 4.

The real consumption of citizens in district i and period s is given by:

$$c_{is} = y_{is}(1 - \tau_{is}^r), \tag{3}$$

where $r \in \{d, c\}$ is the choice of fiscal regime, and τ_{is}^r denotes the proportional tax rate imposed on income under each regime. Let the utility of every citizen in district i in period s be given by:

$$u_{is} = \log c_{is} - \lambda^r, \tag{4}$$

where λ^r ($\lambda^d = 0 < \lambda^c$) is a measure of each district's disadvantage on losing autonomy. It may capture the effects of sociological factors that likely prompt preferences for political independence, such as ethno-linguistic or religious fragmentation in the society.[7]

---

[6]Alesina, Angeloni, and Etro (2005) study the role of international unions in resolving the problems of externalities arising from such public goods.

[7]Easterly and Levine (1997), Poterba (1997), and Alesina, Baqir, and Easterly (1999) find

8

In every period s, a representative, drawn randomly from the homogeneous citizenry in each district i, sets a tax rate τ_{is}^r either individually or collectively according to fiscal regime r. In choosing τ_{is}^r, the representative counts the utility of all citizens who reside or will reside in his district through period $s+T$, $1 \leq T < \infty$. From (1), (3), and (4), the welfare of his district is thus measured by:

$$V_{is}^T = \sum_{k=0}^{T} \delta^k u_{i,s+k} = \sum_{k=0}^{T} \delta^k \left[\log x_i + \log \Delta_{i,s+k} + \log \left(1 - \tau_{i,s+k}^r\right) - \lambda^r \right], \qquad (5)$$

where $\delta \in (0,1)$ denotes a discount factor common to all citizens.

2.2 Choice of Fiscal Regime

Suppose that, at the beginning of period 0 and prior to the representatives' policy decision, districts collectively choose one of the two systems. We denote the steady-state level of Δ_{is} under regime r by Δ_i^r. At this design stage, the welfare of district i under fiscal regime r is measured by the steady-state level of period utility, which is derived from (1), (3), and (4) as:

$$W_i^r = \log x_i + \log \Delta_i^r + \log \left(1 - \tau^r\right) - \lambda^r. \qquad (6)$$

We demonstrate in the subsequent sections that a tax rate chosen at each policy decision stage is uniform across regions and periods, i.e., $\tau_{is}^r = \tau^r$ holds for each r. Formulation (6) implies that these policy outcomes are expected at the system selection stage.

To determine which system is selected, we rely on the Pareto criterion for binary choice, which is also used in Lockwood's (2002) argument.

evidence showing that ethno-linguistic division is a significant cause of conflicts over public good policies. Panizza (1999) finds that ethno-linguistic division in a country is negatively correlated with the degree of its fiscal centralization. In contrast, Arzaghi and Henderson (2005) show that ethno-linguistic diversity is positively related to fiscal centralization. They mention as a reason that "diverse countries may choose to remain de facto centralized as a strong-arm way of avoiding a complete split of the country into two (or more)" (p. 1181).

Definition 1. System r is Pareto-preferred over system r' if and only if, without any side payment, the allocation of local public goods under r makes at least one district better off, and no district worse off, than the allocation under r'.

According to Lockwood's terminology, a system is *unanimously preferred* if and only if it is Pareto-preferred. The following discussion conforms with this definition, but does not depend on the adoption of a specific legislative rule. We can, however, presume the adoption of the following rule according to this criterion:

Example 1. Let each of two representatives cast a vote in the central legislature for their preferred fiscal regime [that which maximizes (6)]. If a representative has an equal preference for two options, the vote of all such representatives is split among them with equal shares. If a *sole* system receives the highest number of votes, then it is selected; otherwise, the status quo holds. The system selected with the highest number of votes is clearly preferred unanimously over the other system, according to Definition 1.

3   Shortsighted Policy Choice

We first examine policy choice under the alternative fiscal regimes, and subsequently system selection in anticipation of it, when the representative has a short time horizon $(T = 1)$ and cares only about the immediate effects of his policy choice on citizens' income in the next period. We extend this horizon in Section 4.

Decentralization. The period s representative in district i selects the best policy τ_{is}^d against the other policy τ_{js}^d in anticipation of $\tau_{i,s+1}^d$.[8] A pair of simultaneous best-policy choices therefore constitutes a Nash equilibrium in each period game.

---

[8]Due to our assumption of a continuum of citizens, the probability of any representative being reelected is 0, so that his strategy is confined to a choice of τ_{is}^r.

Formally, from (5), the period s representative in district i maximizes

$$V_{is}^1 = \log x_i + \log \Delta_{is} + \log\left(1 - \tau_{is}^d\right) + \delta\left[\log x_i + \log \Delta_{i,s+1} + \log\left(1 - \tau_{i,s+1}^d\right)\right], \quad (7)$$

subject to the balanced-budget constraint

$$g_{is} = \tau_{is}^d y_{is}. \qquad (8)$$

It follows from (7) that the representative's choice of τ_{is}^d is independent of the level of the state variables $\{\Delta_{is}, \Delta_{js}\}$. This logic also applies to the next-period representative's choice of $\tau_{i,s+1}^d$. It will be chosen independently of $\{\Delta_{i,s+1}, \Delta_{j,s+1}\}$ and hence $\{\tau_{is}^d, \tau_{js}^d\}$. By regarding τ_{js}^d and $\tau_{i,s+1}^d$ as constants and by substituting (8) into (7), we can derive the following first-order condition with respect to τ_{is}^d:

$$-\frac{1}{1 - \tau_{is}^d} + \frac{\alpha\delta}{\tau_{is}^d} = 0. \qquad (9)$$

The first term on the left-hand side of (9) is associated with the marginal cost imposed on citizens through tax payment in the current period, and the second term represents the marginal benefit generated in the next period due to local public good provision in the current period.[9] Therefore, for any $i \in \{A, B\}$ and s, we can derive the following tax rule by solving (9):

$$\tau_{is}^d = \tau^d = \frac{\alpha\delta}{1 + \alpha\delta}, \quad T = 1. \qquad (10)$$

Note that given Δ_{is}, the choice of τ_{is}^d indirectly determines the level of g_{is} according to (8). Indeed, we defined choice of τ_{is}^d, and not choice of g_{is}, as the strategy of the game. If g_{is} were a control variable instead, choice of g_{is} would affect choice of $g_{i,s+1}$ in the next period, which would make it difficult to search for an equilibrium. Thus, our definition of the game strategy, as well as the logarithmic function in (4), allows us to derive an equilibrium by utilizing the first-order condition without further specification of the model.

---

[9] It is straightforward to show that the second-order condition holds as in successive maximization problems.

Centralization. Under the centralized fiscal regime, the national tax rate and the allocation of local public goods are uniform across regions, i.e., $\tau_{is}^c = \tau_{js}^c = \tau_s^c$ and $g_{is} = g_{js} = g_s$. This supposition describes the way in which a national tax scheme implements interregional cost-sharing. Policy determination under a centralized fiscal regime occurs after representatives who serve in the central legislature are selected. The present approach to centralized decision-making is normative; a central government must solve the problem of interregional benefit spillovers through the agreement of all concerned.[10] From (5), anticipating τ_{s+1}^c, delegates from the two districts reach an agreement that maximizes, by choice of τ_s^c, the welfare of all citizens:

$$
\begin{aligned}
V_{is}^1 + V_{js}^1 &= \log x_i + \log \Delta_{is} + \log\left(1 - \tau_s^c\right) - \lambda^c \\
&+ \log x_j + \log \Delta_{js} + \log\left(1 - \tau_s^c\right) - \lambda^c \\
&+ \delta \left[\log x_i + \log \Delta_{i,s+1} + \log\left(1 - \tau_{s+1}^c\right) - \lambda^c\right] \\
&+ \delta \left[\log x_j + \log \Delta_{j,s+1} + \log\left(1 - \tau_{s+1}^c\right) - \lambda^c\right],
\end{aligned}
\tag{11}
$$

subject to the balanced-budget constraint

$$
g_s = \frac{\tau_s^c(y_{is} + y_{js})}{2}.
\tag{12}
$$

By considering τ_{s+1}^c as a constant and by substituting (12) into (11), we can derive the first-order condition with respect to τ_s^c as:

$$
-\frac{1}{1 - \tau_s^c} + \frac{\alpha(1 + \gamma)\delta}{\tau_s^c} = 0.
\tag{13}
$$

From (13), we derive the following tax rule for any s:

$$
\tau_s^c = \tau^c = \frac{\alpha(1 + \gamma)\delta}{1 + \alpha(1 + \gamma)\delta}, \quad T = 1.
\tag{14}
$$

---

[10]In addition to their analysis of cooperative decision-making by representatives in a central legislature, Besley and Coate (2003) study non-cooperative representatives who seek to advance only their own interests. However, they justify the cooperative approach on an empirical basis (see p. 2622).

Comparison between results (10) and (14) shows that for $\gamma > 0$ the tax rate under a centralized fiscal regime is higher than that under a decentralized fiscal regime. In the presence of interregional benefit spillovers, policy choice under a centralized system attempts to internalize such spillovers, which leads to a higher tax rate.

Comparison of Two Fiscal Regimes. For any time horizon T, $\log \Delta_i^r$ in (6) is calculated as:

$$\log \Delta_i^d = \frac{\alpha(1+\gamma)}{1 - \alpha(1+\gamma)} \log \tau^d + \frac{\alpha(1+\gamma)}{1 - \alpha(1+\gamma)} \log \left(x_i^{1-\mu} x_j^{\mu} \right),$$

$$\log \Delta_i^c = \log \Delta^c = \frac{\alpha(1+\gamma)}{1 - \alpha(1+\gamma)} \log \tau^c + \frac{\alpha(1+\gamma)}{1 - \alpha(1+\gamma)} \log \overline{x}, \tag{15}$$

where $\overline{x} \equiv \frac{x_i + x_j}{2}$ and $\mu \equiv \frac{\gamma}{(1+\gamma)(1-\alpha(1-\gamma))}$. For a detailed derivation of (15), see Appendix A.1. We have already given τ^d and τ^c in (10) and (14) for $T = 1$; for an arbitrary choice of T, these are derived in Section 4. Note that from Assumption 1,

$$\frac{\partial \mu}{\partial \gamma} = \frac{1 - \alpha - \alpha\gamma^2}{(1+\gamma)^2 (1 - \alpha(1 - \gamma))^2} > 0, \tag{16}$$

and μ takes its lowest value of 0 when $\gamma = 0$ and its highest value of $\frac{1}{2}$ when $\gamma = 1$. Conditions (15) and (16) suggest that under a decentralized fiscal regime, with a higher extent of interregional benefit spillovers, the steady-state level of income in a district is influenced more by the neighboring district's productivity through public good provision by this district.

For any time horizon T, (6) and (15) can be used to express the net welfare of citizens in district i by choosing a centralized fiscal regime rather than a decentralized system:

$$W_i^c - W_i^d = \frac{\alpha(1+\gamma)}{1 - \alpha(1+\gamma)} \log \frac{\tau^c}{\tau^d} + \frac{\alpha(1+\gamma)}{1 - \alpha(1+\gamma)} \log \frac{\overline{x}}{x_i^{1-\mu} x_j^{\mu}} + \log \frac{1 - \tau^c}{1 - \tau^d} - \lambda^c. \tag{17}$$

The first, second, and third terms on the right-hand side of (17) are associated with the advantages of a centralized fiscal regime for public input provision, cost-sharing, and the tax burden, respectively.

13

Employing τ^r in (10) and (14) for each system, the result from calculation (17) for a shortsighted policy choice $(T = 1)$ is as follows:

$$
\begin{aligned}
&W_i^c|_{T=1} - W_i^d|_{T=1} \\
&= \frac{\alpha(1+\gamma)}{1-\alpha(1+\gamma)} \log(1+\gamma) + \frac{1}{1-\alpha(1+\gamma)} \log \frac{1+\alpha\delta}{1+\alpha(1+\gamma)\delta} \\
&+ \frac{\alpha(1+\gamma)}{1-\alpha(1+\gamma)} \log \frac{\overline{x}}{x_i^{1-\mu}x_j^{\mu}} - \lambda^c.
\end{aligned}
\tag{18}
$$

Using (18), decentralization and centralization can be compared as $W_i^c|_{T=1} \gtreqless W_i^d|_{T=1}$, which is equivalent to

$$
\log \frac{\overline{x}}{x_i^{1-\mu}x_j^{\mu}} \gtreqless -\log(1+\gamma) + \frac{1}{\alpha(1+\gamma)} \log \frac{1+\alpha(1+\gamma)\delta}{1+\alpha\delta} + \frac{1-\alpha(1+\gamma)}{\alpha(1+\gamma)} \lambda^c. \tag{19}
$$

The left-hand side of (19) can take different values between the two districts, but the value of the right-hand side of (19) is common. On the left-hand side of (19), note that

$$
\frac{\partial}{\partial\left(\frac{x_i}{x_j}\right)} \log \frac{\overline{x}}{x_i^{1-\mu}x_j^{\mu}} \lesseqgtr 0, \text{ if } \frac{x_i}{x_j} \lesseqgtr \frac{1}{\mu} - 1. \tag{20}
$$

For the derivation of (20), see Appendix A.2. The results in (20) imply that given μ, when the degree of interregional disparity is relatively small, a centralized fiscal regime is less favorable for high-productivity district B as disparity increases: under a national tax scheme, the high-productivity district is effectively forced to transfer part of its tax payment to the low-productivity district, and therefore it is more reluctant to submit to it with greater disparity. Over a certain level of disparity, which is lower with higher γ, district B benefits more from centralization, since it can support a low-productivity district's public good provision that may generate benefit spillovers. On the other hand, for district A, a centralized fiscal regime is more favorable as disparity is more distinct, since $\frac{x_A}{x_B} < 1 \leq \frac{1}{\mu} - 1$ for any μ. Appendix A.3 also shows that the right-hand side of (19) is strictly decreasing in γ. This suggests that a higher level of spillovers makes cooperative decision-making more attractive to both districts.

14

We can examine the case when a centralized fiscal regime is unanimously chosen by utilizing (19). It is straightforward to show that district B's level of the left-hand side of (19) is equal to or lower than that of district A since $\mu \in [0, \frac{1}{2}]$ (equality arises only for $\gamma = 1$ and hence $\mu = \frac{1}{2}$). This means that high-productivity district B, a *transferor*, desires decentralization more strongly than low-productivity district A, a *transferee*. That is, when $\gamma < 1$, if $W_B^c|_{T=1} \geq W_B^d|_{T=1}$, then $W_A^c|_{T=1} > W_A^d|_{T=1}$, so that a centralized system is unanimously preferred over a decentralized system according to Definition 1. When $\gamma = 1$, a centralized system is unanimously preferred over a decentralized system if and only if $W_B^c|_{T=1} > W_B^d|_{T=1}$ (then $W_A^c|_{T=1} > W_A^d|_{T=1}$ also holds).

This relation is now investigated in more detail for any pair of $\left(\frac{x_B}{x_A}, \gamma\right) \in (1, \infty) \times [0, 1]$. According to (17), we define $\rho_i^T \equiv \frac{1-\alpha(1+\gamma)}{\alpha(1+\gamma)} \left(W_i^c - W_i^d\right)$. Then $\rho_i^1 = \frac{1-\alpha(1+\gamma)}{\alpha(1+\gamma)} \left(W_i^c|_{T=1} - W_i^d|_{T=1}\right)$, which corresponds to the left-hand side minus the right-hand side of (19). Suppose that a set of parameters $\{\alpha, \delta, \lambda^c\}$ satisfies $\lim_{\frac{x_B}{x_A} \to 1} \rho_B^1|_{\gamma=1} (= \lim_{\frac{x_B}{x_A} \to 1} \rho_A^1|_{\gamma=1}) = \log 2 - \frac{1}{2\alpha} \log \frac{1+2\alpha\delta}{1+\alpha\delta} - \frac{1-2\alpha}{2\alpha}\lambda^c \geq 0$. Note that this condition is likely to hold with a lower discount factor δ (implying that citizens and representatives are concerned about the current cost rather than the future benefit) and a lower cost of central governance λ^c (the case with $\lim_{\frac{x_B}{x_A} \to 1} \rho_B^1|_{\gamma=1} < 0$ is investigated in Section 4, where decentralization may occur even if the benefit from a local public good spills over completely). As demonstrated below, under this condition we can find a single value of $\gamma \in (0, 1)$ that satisfies $\rho_B^1 = 0$ for any choice of $\frac{x_B}{x_A} \in (1, \infty)$. This value is given by the function $\hat{\gamma}_B^1 = \hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right)$. Then, by employing the implicit function theorem the following relation is derived:

$$\frac{\partial \hat{\gamma}_B^1}{\partial \left(\frac{x_B}{x_A}\right)} = -\frac{\partial \rho_B^1}{\partial \left(\frac{x_B}{x_A}\right)} \Big/ \frac{\partial \rho_B^1}{\partial \gamma}. \tag{21}$$

Note that the left-hand side of (19) for B strictly increases in γ, since, from (16),

$$\frac{\partial}{\partial \mu} \log \frac{\overline{x}}{x_B^{1-\mu} x_A^\mu} \frac{\partial \mu}{\partial \gamma} = \log \frac{x_B}{x_A} \frac{\partial \mu}{\partial \gamma} > 0,$$

while the right-hand side of (19) strictly decreases in γ. These facts lead to $\frac{\partial \rho_B^1}{\partial \gamma} > 0$,

and therefore the sign of (21) depends on the sign of $\frac{\partial \rho_B^1}{\partial \left(\frac{x_B}{x_A} \right)} = \frac{\partial}{\partial \left(\frac{x_B}{x_A} \right)} \log \frac{\overline{x}}{x_B^{1-\mu} x_A^{\mu}}$. Consequently, by relying on (20), we can derive the sign of (21) as:

$$\frac{\partial \hat{\gamma}_B^1}{\partial \left(\frac{x_B}{x_A} \right)} \gtreqless 0, \ \text{if} \ \frac{x_B}{x_A} \lesseqgtr \frac{1}{\mu} - 1. \tag{22}$$

In Figure 1 the thin down-slope curve indicates the magnitude of $\xi(\gamma) \equiv \frac{1}{\mu(\gamma)} - 1$ on the horizontal axis. Below this curve, $\frac{x_B}{x_A} < \frac{1}{\mu} - 1$ and hence $\frac{\partial \hat{\gamma}_B^1}{\partial \left(\frac{x_B}{x_A} \right)} > 0$ from (22); similarly, $\frac{x_B}{x_A} > \frac{1}{\mu} - 1$ and hence $\frac{\partial \hat{\gamma}_B^1}{\partial \left(\frac{x_B}{x_A} \right)} < 0$ above the curve. Thus, we can plot the points $\left(\frac{x_B}{x_A}, \hat{\gamma}_B^1 \left(\frac{x_B}{x_A} \right) \right)$ as a single-peaked curve. The peak is on the down-slope curve at $\frac{x_B}{x_A} = \check{x}$. It follows from these facts and (20) that under the supposition of $\lim_{\frac{x_B}{x_A} \to 1} \rho_B^1|_{\gamma=1} \geq 0$, $0 < \hat{\gamma}_B^1 < 1$ for any $\frac{x_B}{x_A} \in (1, \infty)$.

Figure 1 also shows district B's preference relation for any $\left(\frac{x_B}{x_A}, \gamma \right)$. Points $\left(\frac{x_B}{x_A}, \gamma \right)$ below this single-peaked curve $\hat{\gamma}_B^1 \left(\frac{x_B}{x_A} \right)$ favor decentralization, and those above the curve favor centralization. According to our discussion, if the combination of the disparity and the degree of benefit spillovers is such that $\left(\frac{x_B}{x_A}, \gamma \right)$ is located on or above the curve, then a centralized fiscal regime is unanimously preferred. These results establish the following proposition.

Proposition 1. Suppose $T = 1$ (shortsighted policy choice). Let Assumptions 1 and 2 hold. Suppose that a set of parameters $\{\alpha, \delta, \lambda^c\}$ satisfies $\log 2 - \frac{1}{2\alpha} \log \frac{1+2\alpha\delta}{1+\alpha\delta} - \frac{1-2\alpha}{2\alpha} \lambda^c \geq 0$. Then we can define the function $\hat{\gamma}_B^1 \left(\frac{x_B}{x_A} \right)$, $\frac{x_B}{x_A} \in (1, \infty)$, on which any point $\left(\frac{x_B}{x_A}, \hat{\gamma}_B^1 \left(\frac{x_B}{x_A} \right) \right)$, $\hat{\gamma}_B^1 \in (0, 1)$, means that citizens in district B are indifferent to the choice between two fiscal regimes. There is a threshold value $\check{x} \in (1, \infty)$ such that $\frac{\partial \hat{\gamma}_B^1}{\partial \left(\frac{x_B}{x_A} \right)} \gtreqless 0$ for $\frac{x_B}{x_A} \lesseqgtr \check{x}$. Then a centralized fiscal regime is unanimously preferred over a decentralized fiscal regime if and only if $\left(\frac{x_B}{x_A}, \gamma \right)$ is such that $\gamma \geq \hat{\gamma}_B^1 \left(\frac{x_B}{x_A} \right)$.

According to Proposition 1, the high degree of benefit spillovers is linked to unanimous preference for a centralized fiscal regime. It also suggests that central-ization is more likely under sufficiently small and sufficiently large values of disparity

measure $\frac{x_B}{x_A}$. When the productivity is similar, so that $\frac{x_B}{x_A}$ is small, high-productivity district B is more reluctant to transfer a portion of its income to low-productivity district A as the disparity increases. If the disparity is so serious that it exceeds the threshold, however, financially supporting the neighbor's local public good provision is beneficial to district B, even with the lower degree of spillovers. The merit of cost-sharing inherent to the centralized fiscal regime is dominant.

Corollary 1. For any $\frac{x_B}{x_A} \in (1, \infty)$, $\hat{\gamma}_B^1 \left(\frac{x_B}{x_A} \right)$ is strictly increasing with respect to each of λ^c and δ, while $\check{x}$ is strictly decreasing with respect to each of λ^c and δ.

Proof of Corollary 1. See Appendix A.4.

Corollary 1 describes that the single-peaked curve in Figure 1 shifts upward with increasing conflict by losing autonomy and increasing discount factor, resulting in likely decentralization in spite of the higher degree of benefit spillovers. When the discount factor is so high that policy-makers take greater account of the benefit of local public good provision rather than its cost, the divergence in tax burden under the two systems is greater, and a decentralized system is more favored. Thus, the expected higher tax rate that all districts will agree on to solve the problem of positive externalities may ironically hinder cooperative solution, as well as the cost of conflict.

Remark 1. In our model, interregional cost-sharing and internalization of benefit spillovers can be achieved by delegating power to an institution overseeing interregional coordination, a central government. Without its power to enforce cooperative solutions, each district would consequently behave selfishly and a Nash equilibrium (a tax rate under a decentralized fiscal regime) would arise. Let us interpret each district as a nation and the central government as an international institution. In this case, Proposition 1 demonstrates the circumstances under which an international institution may be established to achieve efficient allocation of global public

goods.

Remark 2. Proposition 1 means that if $\left(\frac{x_B}{x_A}, \gamma\right)$ is such that $\gamma < \hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right)$, a decentralized fiscal regime is unanimously preferred over a centralized fiscal regime, or no fiscal regime is unanimously preferred. Bolton and Roland (1997) suppose that when at least one region favors decentralization, secession occurs. If we complied with their definition, Proposition 1 could give us a necessary and sufficient condition for secession: $\gamma < \hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right)$.

We can examine the unanimous choice of a decentralized fiscal regime in a similar manner. When $\gamma > 0$, if $W_A^d|_{T=1} \geq W_A^c|_{T=1}$, then $W_B^d|_{T=1} > W_B^c|_{T=1}$, so that a decentralized system is unanimously preferred over a centralized system. When $\gamma = 1$, a decentralized system is unanimously preferred over a centralized system if and only if $W_A^d|_{T=1} > W_A^c|_{T=1}$.

We investigate again the relationship between a pair of $\left(\frac{x_B}{x_A}, \gamma\right)$ and district A's preferred system. Although we cannot analytically solve the sign of $\frac{\partial \rho_A^1}{\partial \gamma}$ at each point without further assumptions, it immediately follows from (20) that $\frac{\partial \rho_A^1}{\partial\left(\frac{x_B}{x_A}\right)} > 0$ for any $\left(\frac{x_B}{x_A}, \gamma\right) \in (1, \infty) \times [0, 1]$. This implies that if we have $\frac{x_B}{x_A} \in (1, \infty)$ that satisfies $\rho_A^1 = 0$ for a given γ, it should be single-valued. We represent this value of $\frac{x_B}{x_A}$ by the function $\hat{X}_A^1 = \hat{X}_A^1(\gamma)$. For example, we examine the sign of ρ_A^1 for $\gamma = 0$. Then $\rho_A^1|_{\gamma=0} \gtreqqless 0$ for $\frac{x_B}{x_A} \gtreqqless \hat{X}_A^1(0) = e^{\frac{1-\alpha}{\alpha}\lambda^c + \log 2} - 1$, so that values of the disparity measure higher (lower) than $\hat{X}_A^1(0)$ are associated with centralization (decentralization). Furthermore, from (19), $\lim_{\frac{x_B}{x_A} \to 1} \rho_A^1|_{\gamma=\hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right)} = \lim_{\frac{x_B}{x_A} \to 1} \rho_B^1|_{\gamma=\hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right)} = 0$. These two facts suggest that we can plot a frontier $\hat{X}_A^1(\gamma)$ for an interval $\gamma \in [0, \lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right))$. Figure 2 displays one example of the curve $\hat{X}_A^1(\gamma)$ and district A's preference relations for two fiscal regimes, although we cannot certify that $\hat{X}_A^1$ is monotonically decreasing in γ for any set of $\{\alpha, \delta, \lambda^c\}$.

Proposition 2. Suppose $T = 1$. Let Assumptions 1 and 2 hold. Suppose that a

18

set of parameters $\{\alpha, \delta, \lambda^c\}$ satisfies $\log 2 - \frac{1}{2\alpha} \log \frac{1+2\alpha\delta}{1+\alpha\delta} - \frac{1-2\alpha}{2\alpha}\lambda^c \geq 0$. Then we can define the continuously differentiable function $\hat{X}_A^1(\gamma)$, $\gamma \in [0, \lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1 \left(\frac{x_B}{x_A}\right))$, on which $\hat{X}_A^1(0) = e^{\frac{1-\alpha}{\alpha}\lambda^c + \log 2} - 1$, $\lim_{\gamma \to \lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right)} \hat{X}_A^1(\gamma) = 1$, and any point $\left(\hat{X}_A^1(\gamma), \gamma\right)$ means that citizens in district A are indifferent to the choice between the two fiscal regimes. Then a decentralized fiscal regime is unanimously preferred over a centralized fiscal regime if and only if $\left(\frac{x_B}{x_A}, \gamma\right)$ is such that $\frac{x_B}{x_A} \leq \hat{X}_A^1(\gamma)$.

Continuous differentiability of function $\hat{X}_A^1(\gamma)$ arises from the continuous differentiability of ρ_A^1 in $\frac{x_B}{x_A}$ and γ. According to Proposition 2, a pair of a lower value of the disparity measure and a lower value of the degree of benefit spillovers is likely linked to unanimous preference for a decentralized fiscal regime, because both districts lose less by decentralization. If the disparity measure $\frac{x_B}{x_A}$ is large, divergence between the level of a public good for district A under centralization and decentralization is also large, and it will gain more from centralization even if the tax rate rises. If the degree of benefit spillovers γ is large, it benefits district B to support district A financially.

The following corollary shows the positive relationship between potential friction due to centralized governance and agreement on decentralization.

Corollary 2. $\hat{X}_A^1(0)$ and $\lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right)$ are strictly increasing with respect to λ^c. Furthermore, $\lim_{\lambda^c \to 0} \hat{X}_A^1(0) = 1$ and $\lim_{\lambda^c \to 0} \lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1\left(\frac{x_B}{x_A}\right) = 0$.

Proof of Corollary 2. See Appendix A.5.

4 General Solutions for Policy Choice

A local public good in a district also has indirect effects on the income of future citizens in this district by influencing the dynamic transition of the state variables. Given $\{\Delta_{is}, \Delta_{js}\}$, and hence given tax bases $\{y_{is}, y_{js}\}$, the choice of τ_{is}^r determines

the public good level of district i in period s. With interregional benefit spillovers, this influences the period $s+1$ state variables $\{\Delta_{i,s+1}, \Delta_{j,s+1}\}$, and therefore the period $s+1$ tax bases in both districts $\{y_{i,s+1}, y_{j,s+1}\}$, which in turn affect the period $s+1$ local public good levels. These will subsequently influence the period $s+2$ state variables and the tax bases of both districts, and so on. Let us refer to these spillovers from the district i public good in period s to the district i income in period $s+k$, $2 \leq k < \infty$, as *intergenerational benefit spillovers*.[11] A representative with time horizon T will count these effects through period $s+T$.

Figure 3 describes how the choice of τ_{is}^r contributes to the period utility of each district. Due to the logarithm in the period utility function (4), given $\{\Delta_{is}, \Delta_{js}\}$ and hence $\{y_{is}, y_{js}\}$, the choice of τ_{is}^r, immediately affecting the period s utility of district i by $\log(1 - \tau_{is}^r)$, also influences the period $s+1$ utility of citizens in district i by $\alpha \log \tau_{is}^r$, whereas it influences the period $s+1$ utility of citizens in district j by $\alpha\gamma \log \tau_{is}^r$. In period $s+2$, it contributes to the utility of districts i and j by $(\alpha^2 + (\alpha\gamma)^2) \log \tau_{is}^r$ and $2\alpha(\alpha\gamma) \log \tau_{is}^r$, respectively. Under a decentralized fiscal regime, a district i representative at the policy decision stage in period s considers its effects on $u_{i,s+k}$ up to $k = T$. Under a centralized fiscal regime, its whole effects on $u_{i,s+k}$ and $u_{j,s+k}$ are counted. A stream of these future effects on $u_{i,s+k}$ or $u_{i,s+k} + u_{j,s+k}$ is demonstrated in Appendix A.6. To derive the optimal policy choice under each regime, the balance between the tax burden in the current period and discounted future benefits should be solved.

Decentralization. Given $\{\Delta_{is}, \Delta_{js}\}$, against the other representative's policy choice and in anticipation of future policy decisions in both districts, the period s representative in district i chooses τ_{is}^d to maximize V_{is}^T in (5) subject to (8). The

---

[11]We use the term "intergenerational" to describe the relations permitting generations that did not pay the tax price to benefit from local public goods supplied by preceding generations. A comment by Yukihiro Nishimura on an earlier version of this paper inspired this treatment of intergenerational benefit spillovers.

effect of τ_{is}^d on $u_{i,s+k}$, $1 \leq k \leq T$, is given by:

$$\frac{(\alpha + \alpha\gamma)^k + (\alpha - \alpha\gamma)^k}{2} \log \tau_{is}^d. \tag{23}$$

For derivation of (23), see Appendix A.6. Then, from (23), by taking $\{\tau_{i,s+k}^d, \tau_{j,s+k-1}^d\}_{1 \leq k \leq T}$ as a set of constants, the first-order condition with respect to τ_{is}^d is:

$$-\frac{1}{1 - \tau_{is}^d} + \frac{\theta^d}{\tau_{is}^d} = 0, \tag{24}$$

where

$$\begin{aligned}
\theta^d &= \sum_{k=1}^{T} \delta^k \frac{(\alpha + \alpha\gamma)^k + (\alpha - \alpha\gamma)^k}{2} \\
&= \frac{1}{2}\left(\frac{(\alpha + \alpha\gamma)\delta(1 - (\alpha + \alpha\gamma)^T\delta^T)}{1 - (\alpha + \alpha\gamma)\delta} + \frac{(\alpha - \alpha\gamma)\delta(1 - (\alpha - \alpha\gamma)^T\delta^T)}{1 - (\alpha - \alpha\gamma)\delta} \right). (25)
\end{aligned}$$

Condition (24) provides us with the following tax rule for any $i \in \{A, B\}$ and s:

$$\tau_{is}^d = \tau^d = \frac{\theta^d}{1 + \theta^d}. \tag{26}$$

It follows immediately from (25), (26), and Assumption 1 that a tax rate chosen by a policy-maker with longer sight is higher, since it internalizes intergenerational benefit spillovers into the farther future.

Centralization. Given $\{\Delta_{is}, \Delta_{js}\}$, the period s delegates from both districts cooperate so as to maximize $V_{is}^T + V_{js}^T$ by the choice of $\tau_s^c = \tau_{is}^c = \tau_{js}^c$ subject to (12). The effect of τ_s^c on $u_{i,s+k} + u_{j,s+k}$, $1 \leq k \leq T$, is given by:

$$2(\alpha + \alpha\gamma)^k \log \tau_s^c. \tag{27}$$

Refer to Appendix A.6 for the derivation of (27). Then, by regarding $\{\tau_{s+k}^c\}_{1 \leq k \leq T}$ as a set of constants, we can derive the first-order condition with respect to τ_s^c:

$$-\frac{1}{1 - \tau_s^c} + \frac{\theta^c}{\tau_s^c} = 0, \tag{28}$$

where

$$\theta^c = \sum_{k=1}^{T} \delta^k (\alpha + \alpha\gamma)^k = \frac{(\alpha + \alpha\gamma)\delta(1 - (\alpha + \alpha\gamma)^T\delta^T)}{1 - (\alpha + \alpha\gamma)\delta}. \tag{29}$$

Therefore from (28), for any s the following tax rule is derived:

$$\tau_s^c = \tau^c = \frac{\theta^c}{1 + \theta^c}. \tag{30}$$

Note again that τ^c in (30) is greater with higher T.

Comparison of Two Fiscal Regimes. It is difficult to derive analytical solutions for the choice of fiscal regime for any T from (25), (26), (29), and (30) without further specification. Therefore, we focus on the case with full benefit spillovers ($\gamma = 1$). The purpose of the analysis is to investigate how interregional benefit spillovers, which are completely internalized under a centralized fiscal regime, can be internalized even under a decentralized system by a policy-maker who can consider intergenerational spillover effects.

If $\gamma = 1$, from (17), decentralization and centralization are compared as $W_i^c|_{\gamma=1} \gtrless W_i^d|_{\gamma=1}$, which is equivalent to

$$\log \frac{\overline{x}}{\sqrt{x_i x_j}} \gtrless -\log \frac{\tau^c}{\tau^d} - \frac{1 - 2\alpha}{2\alpha} \log \frac{1 - \tau^c}{1 - \tau^d} + \frac{1 - 2\alpha}{2\alpha} \lambda^c, \tag{31}$$

where

$$\frac{\tau^c}{\tau^d} = 2 \left(\frac{1 + \frac{1 - 2\alpha\delta}{\alpha\delta(1 - (2\alpha\delta)^T)}}{2 + \frac{1 - 2\alpha\delta}{\alpha\delta(1 - (2\alpha\delta)^T)}} \right), \tag{32}$$

and

$$\frac{1 - \tau^c}{1 - \tau^d} = \frac{1 + \frac{1 - 2\alpha\delta}{\alpha\delta(1 - (2\alpha\delta)^T)}}{2 + \frac{1 - 2\alpha\delta}{\alpha\delta(1 - (2\alpha\delta)^T)}}. \tag{33}$$

Equations (32) and (33) are derived from (25), (26), (29), and (30). The left-hand side of (31) suggests that with full benefit spillovers, disagreement between districts with diverse productivity does not arise: they favor the same system. It follows from (20) that with $\gamma = 1$ and hence $\mu = \frac{1}{2}$, $\log \frac{\overline{x}}{\sqrt{x_i x_j}}$ is strictly increasing in $\frac{x_B}{x_A}$, suggesting that a greater disparity induces centralization. A change in T leads to a change in $\frac{\tau^c}{\tau^d}$ and $\frac{1 - \tau^c}{1 - \tau^d}$ in the right-hand side of (31). From (32) and (33), it is straightforward to show that under Assumption 1, as T increases, both $\frac{\tau^c}{\tau^d}$ and $\frac{1 - \tau^c}{1 - \tau^d}$

22

decrease. This suggests that a politician's farsightedness in policy decision-making may compensate for the deficiencies of decentralization to a certain extent.

Lemma 1. Let Assumptions 1 and 2 hold. Then $W_i^c|_{\gamma=1} - W_i^d|_{\gamma=1}$ is strictly decreasing with respect to T.

Proof of Lemma 1. See Appendix A.7.

From (17), we can compute $\rho_i^T|_{\gamma=1} = \rho^T|_{\gamma=1} = \frac{1-2\alpha}{2\alpha}\left(W_i^c|_{\gamma=1} - W_i^d|_{\gamma=1}\right)$, which coincides with the left-hand side minus the right-hand side of (31) and is common for any i. Proposition 1 supposed that the environment represented by $\{\alpha, \delta, \lambda^c\}$ satisfies $\lim_{\frac{x_B}{x_A} \to 1} \rho^1|_{\gamma=1} = \log 2 - \frac{1}{2\alpha}\log\frac{1+2\alpha\delta}{1+\alpha\delta} - \frac{1-2\alpha}{2\alpha}\lambda^c \geq 0$. Then, due to the monotonic increase of $\log\frac{\bar{x}}{\sqrt{x_i x_j}}$ with regard to $\frac{x_B}{x_A}$ for $\gamma = 1$, $\rho^1|_{\gamma=1} > 0$ for any $\frac{x_B}{x_A} \in (1, \infty)$, which implies that with full benefit spillovers, a centralized fiscal regime is unanimously preferred over a decentralized fiscal regime for any degree of disparity measure. Now suppose that $\lim_{\frac{x_B}{x_A} \to 1} \rho^1|_{\gamma=1} < 0$: this condition is likely with a higher δ and a higher λ^c. Then we have a threshold $\tilde{X}^1 \in (1, \infty)$ such that $\rho^1|_{\gamma=1} \gtreqless 0$ for $\frac{x_B}{x_A} \gtreqless \tilde{X}^1$. Furthermore, according to Lemma 1, we can also find a threshold $\tilde{X}^T \in (1, \infty)$ for $T > 1$.

Proposition 3. Let Assumptions 1 and 2 hold. Suppose that a set of parameters $\{\alpha, \delta, \lambda^c\}$ satisfies $\log 2 - \frac{1}{2\alpha}\log\frac{1+2\alpha\delta}{1+\alpha\delta} - \frac{1-2\alpha}{2\alpha}\lambda^c < 0$. For any time horizon $T \in [1, \infty)$, we have $\tilde{X}^T \in (1, \infty)$ such that: with complete interregional spillovers ($\gamma = 1$), a decentralized fiscal regime is unanimously preferred over a centralized fiscal regime if $\frac{x_B}{x_A} \in (1, \tilde{X}^T)$; no system is preferred unanimously if $\frac{x_B}{x_A} = \tilde{X}^T$; and a centralized fiscal regime is unanimously preferred over a decentralized fiscal regime if $\frac{x_B}{x_A} \in (\tilde{X}^T, \infty)$. Then $\tilde{X}^T < \tilde{X}^{T+1}$.

Figure 4 illustrates Proposition 3, which has two messages. One is that centralization is more likely selected under a great disparity. Because the benefit produced

by the neighbor's public good completely spills over, the high-productivity district willingly supports the low-productivity district to realize uniform local public good allocation, in spite of the unequal tax burden. The second is that if policy-makers can make decisions that consider the welfare of successive generations, a decentralized fiscal regime stands a better chance of being adopted. Even if each policy-maker is solely interested in the welfare in his district under a decentralized system, he will consider the secondary effect of his choice on the neighboring district, since its increased income will benefit his district at later intervals. Farsighted policy choice, which internalizes intergenerational benefit spillovers through the far future, thus reduces the usual efficiency loss due to decentralization. In other words, when each policy-maker chooses a policy by considering only its immediate effects, an institution to promote cooperation among districts will be established.

Corollary 3. $\tilde{X}^T$ is strictly increasing with respect to λ^c.

Proof of Corollary 3. See Appendix A.8.

Corollary 3 has the same implication as Corollaries 1 and 2: suppression of conflict induced by centralized governance is a force that drives decentralization.

5   Concluding Remarks

This paper has presented a model by which the relative merits of decentralization and centralization can be evaluated. A centralized fiscal regime can implement cost-sharing between districts with diverse productivity levels, and further allows internalization of interregional benefit spillovers from the provision of local public goods. The centralized fiscal regime is, however, assumed to be inferior, in that citizens suffer on losing autonomy. The key features of this model are its incorporation of interregional dependencies and dynamic linkages. In each period, representatives decide on a proportional income tax rate; this decision determines the level of a local

24

public good and affects future income levels in both districts, which in turn affect future local public good levels. The decisions made by representatives also depend on the extent to which they take into account intergenerational benefit spillovers from a local public good.

A key finding of this paper is that a decentralized (centralized) fiscal regime is more likely to be unanimously preferred when the interregional disparity in productivity is small (large) and the extent of interregional benefit spillovers is low (high). It was also found that even with full interregional spillovers, decentralization is more likely to be promoted when policy-makers have sufficient foresight and consider future generations in making their decisions. Farsighted policy-makers allow for long-term benefits due to the increased income of other districts even under a decentralized system, and this effect suppresses the usual welfare loss caused by decentralization.

This work has supposed that delegates cooperate in a central legislature to maximize the welfare of the entire population. Under this supposition, results have been derived showing that as interregional disparity increases, centralized decision-making is more likely to be preferred. Introduction of strategic behavior by delegates into the model could spawn a deeper examination of the problem.

A   Appendix

A.1   General Solutions for $\log \Delta_{is}$

Decentralization. From (1), (2), and (8), using $\tau_{is}^d = \tau^d$ in (10) (or (26)), the system on $\log \Delta_{is}$ and $\log \Delta_{js}$ is given by:

$$
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \log \Delta_{i,s+1} \\ \log \Delta_{j,s+1} \end{bmatrix} + \begin{bmatrix} -\alpha & -\alpha\gamma \\ -\alpha\gamma & -\alpha \end{bmatrix} \begin{bmatrix} \log \Delta_{is} \\ \log \Delta_{js} \end{bmatrix} = \begin{bmatrix} \log \omega_i \\ \log \omega_j \end{bmatrix}, \quad \text{(A1)}
$$

where

$$
\log \omega_i \;=\; \alpha \log \left(\tau^d x_i \right) + \alpha\gamma \log \left(\tau^d x_j \right),
$$

25

$$\log \omega_j = \alpha \log \left(\tau^d x_j \right) + \alpha \gamma \log \left(\tau^d x_i \right). \tag{A2}$$

From (A1), (A2), and Assumption 1, the steady-state levels of $\log \Delta_{is}$ and $\log \Delta_{js}$ under a decentralized fiscal regime, $\log \Delta_i^d$ and $\log \Delta_j^d$, are given by:

$$\begin{bmatrix} \log \Delta_i^d \\ \log \Delta_j^d \end{bmatrix} = \frac{1}{(1-\alpha)^2 - (\alpha\gamma)^2} \begin{bmatrix} (1-\alpha) \log \omega_i + \alpha\gamma \log \omega_j \\ (1-\alpha) \log \omega_j + \alpha\gamma \log \omega_i \end{bmatrix}. \tag{A3}$$

Result (A3) is arranged as the upper part of (15). Since the eigenvalues associated with (A1) are $\alpha(1+\gamma)$ and $\alpha(1-\gamma)$, general solutions are given by:

$$\log \Delta_{is} = \phi_1 (\alpha(1+\gamma))^s + \phi_2 (\alpha(1-\gamma))^s + \log \Delta_i^d,$$

$$\log \Delta_{js} = \phi_1 (\alpha(1+\gamma))^s - \phi_2 (\alpha(1-\gamma))^s + \log \Delta_j^d,$$

where

$$\phi_1 = \log \Delta_0 - \frac{1}{2} \left(\log \Delta_i^d + \log \Delta_j^d \right),$$

$$\phi_2 = -\frac{1}{2} \left(\log \Delta_i^d - \log \Delta_j^d \right).$$

Centralization. It follows from Assumption 2 that $\Delta_{is} = \Delta_{js} = \Delta_s$, $s \geq 0$, since this scheme provides a uniform level of local public goods across districts. From (1), (2), and (12), using $\tau_s^c = \tau^c$ in (14) (or (30)), the difference equation on $\log \Delta_s$ is given by:

$$\log \Delta_{s+1} = \alpha(1+\gamma) \log \left(\tau^c \overline{x} \right) + \alpha(1+\gamma) \log \Delta_s. \tag{A4}$$

From Assumption 1 and (A4), the steady-state level of $\log \Delta_s$, which is denoted by $\log \Delta^c$, is:

$$\log \Delta^c = \frac{\alpha(1+\gamma)}{1 - \alpha(1+\gamma)} \log \left(\tau^c \overline{x} \right). \tag{A5}$$

Result (A5) can be represented as the lower part of (15). Thus, the general solution is:

$$\log \Delta_s = (\log \Delta_0 - \log \Delta^c) \left(\alpha(1+\gamma) \right)^s + \log \Delta^c.$$

26

A.2 Derivation of (20)

Note that

$$\frac{\overline{x}}{x_i^{1-\mu}x_j^{\mu}} = \frac{1}{2}\left(\frac{x_i}{x_j}\right)^{\mu-1}\left(\frac{x_i}{x_j}+1\right). \tag{A6}$$

From (A6),

$$\frac{\partial}{\partial\left(\frac{x_i}{x_j}\right)}\log\frac{\overline{x}}{x_i^{1-\mu}x_j^{\mu}} = \frac{x_i^{1-\mu}x_j^{\mu}}{x_i+x_j}\left(\frac{x_i}{x_j}\right)^{\mu-2}\mu\left(\frac{x_i}{x_j}+1-\frac{1}{\mu}\right). \tag{A7}$$

Then, from (A7), given $\mu > 0$, $\frac{\partial}{\partial\left(\frac{x_i}{x_j}\right)}\log\frac{\overline{x}}{x_i^{1-\mu}x_j^{\mu}} \lesseqgtr 0$ for $\frac{x_i}{x_j} \lesseqgtr \frac{1}{\mu} - 1$. When $\mu = 0$, $\frac{\partial}{\partial\left(\frac{x_i}{x_j}\right)}\log\frac{\overline{x}}{x_i^{1-\mu}x_j^{\mu}} < 0$ for any $\frac{x_i}{x_j}$. These results are summarized in (20).

A.3 Monotonicity of the RHS of (19)

Define $\psi \equiv -\log\left(1+\gamma\right) + \frac{1}{\alpha(1+\gamma)}\log\frac{1+\alpha(1+\gamma)\delta}{1+\alpha\delta} + \frac{1-\alpha(1+\gamma)}{\alpha(1+\gamma)}\lambda^c$. This represents the RHS of (19). Then we can immediately derive

$$\begin{aligned}\frac{\partial\psi}{\partial\gamma} &= -\frac{1}{1+\gamma}\left(1-\frac{\delta}{1+\alpha(1+\gamma)\delta}\right) \\ &\quad -\frac{1}{\alpha(1+\gamma)^2}\log\frac{1+\alpha(1+\gamma)\delta}{1+\alpha\delta} - \frac{1}{\alpha(1+\gamma)^2}\lambda^c < 0.\end{aligned} \tag{A8}$$

Thus, the RHS of (19) strictly decreases in γ.

A.4 Proof of Corollary 1

According to Proposition 1, given $\frac{x_B}{x_A}$, we have a single value of γ, $\hat{\gamma}_B^1 > 0$, with which equality holds in (19) for B. Then the RHS of (19) exceeds its LHS for a higher value of λ^c. To recover equality, a higher value of γ is needed, since the LHS of (19) for B strictly increases in γ, and the RHS of (19) strictly decreases in γ, as shown in (A8).

We can examine the relation between δ and $\hat{\gamma}_B^1$ in the same manner. Recall that ψ represents the RHS of (19). Around $\hat{\gamma}_B^1$,

$$\frac{\partial\psi}{\partial\delta} = \frac{\gamma}{(1+\gamma)(1+\alpha(1+\gamma)\delta)(1+\alpha\delta)} > 0, \tag{A9}$$

27

and hence the RHS of (19) exceeds its LHS for a higher value of δ. To recover equality, a higher value of γ is needed for the same reasoning.

The last assertion for $\check{x}$ depends on the fact that the peak of $\hat{\gamma}_B^1 \left(\frac{x_B}{x_A} \right)$ is on the down-slope curve.

A.5 Proof of Corollary 2

The assertion for $\hat{X}_A^1(0)$ is immediately derived using $\hat{X}_A^1(0) = e^{\frac{1-\alpha}{\alpha} \lambda^c + \log 2} - 1$ and Assumption 1.

For the monotonicity of $\lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1$ with respect to λ^c, note that from (A6),

$$\rho_B^1 \left(\frac{x_B}{x_A}, \gamma, \lambda^c \right) = -\log 2 + (\mu(\gamma) - 1) \log \frac{x_B}{x_A} + \log \left(\frac{x_B}{x_A} + 1 \right) - \psi(\gamma, \lambda^c), \quad \text{(A10)}$$

where ψ represents the RHS of (19). From the definition of $\hat{\gamma}_B^1$, given λ^c, the following condition holds for any $\frac{x_B}{x_A} \in (1, \infty)$:

$$\rho_B^1 \left(\frac{x_B}{x_A}, \hat{\gamma}_B^1, \lambda^c \right)$$
$$= -\log 2 + \left(\mu \left(\hat{\gamma}_B^1 \right) - 1 \right) \log \frac{x_B}{x_A} + \log \left(\frac{x_B}{x_A} + 1 \right) - \psi \left(\hat{\gamma}_B^1, \lambda^c \right) = 0, \quad \text{(A11)}$$

which also implies

$$\lim_{\frac{x_B}{x_A} \to 1} \rho_B^1 \left(\frac{x_B}{x_A}, \hat{\gamma}_B^1, \lambda^c \right) = -\psi \left(\hat{\gamma}_B^1, \lambda^c \right) = 0, \quad \text{(A12)}$$

since $\mu(\cdot)$ is bounded. It follows from (A12) that $\lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1$ is implicitly given by $\psi(\gamma, \lambda^c) = 0$. Suppose that it holds for a given value of λ^c. Then $\psi(\gamma, \lambda^c) > 0$ with a higher value of λ^c, and to recover equality we need a higher value of γ, since ψ strictly decreases in γ.

Furthermore, note that

$$\lim_{\lambda^c \to 0} \psi(\gamma, \lambda^c) = -\log(1 + \gamma) + \frac{1}{\alpha(1 + \gamma)} \log \frac{1 + \alpha(1 + \gamma)\delta}{1 + \alpha\delta}, \quad \text{(A13)}$$

which equals 0 only for $\gamma = 0$ owing to its monotonicity [refer to (A8)]. The assertion $\lim_{\lambda^c \to 0} \lim_{\frac{x_B}{x_A} \to 1} \hat{\gamma}_B^1 = 0$ is thus proved.

28

A.6 Future Effects of Policy Choice

We calculate the effects of a choice of τ_{is}^r on $u_{i,s+k}$ and $u_{j,s+k}$. The discussion does not rely on an assumption of $\tau_{is}^c = \tau_{js}^c$. Let $\beta \equiv \alpha\gamma$ and let $\kappa_{s+k} = (\alpha + \beta)^k$. Then $\kappa_{s+k} \log \tau_{is}^r$ represents the total contribution of a policy choice τ_{is}^r to the sum of the period $s + k$ utility for both districts, $u_{i,s+k} + u_{j,s+k}$. Let the benefit spilling from τ_{is}^r to $u_{j,s+k}$ be denoted by $\chi_{s+k} \log \tau_{is}^r$. Then χ_{s+k} evolves according to

$$
\begin{aligned}
\chi_{s+k+1} &= \alpha\chi_{s+k} + \beta(\kappa_{s+k} - \chi_{s+k}) \\
&= (\alpha - \beta)\chi_{s+k} + \beta(\alpha + \beta)^k.
\end{aligned} \tag{A14}
$$

From (A14), the general solution for χ_{s+k} is given by:

$$
\chi_{s+k} = \frac{(\alpha + \beta)^k - (\alpha - \beta)^k}{2}. \tag{A15}
$$

From (A15), when choosing τ_{is}^d, a district i representative considers the stream

$$
(\kappa_{s+k} - \chi_{s+k}) \log \tau_{is}^d = \frac{(\alpha + \beta)^k + (\alpha - \beta)^k}{2} \log \tau_{is}^d, \;\; 1 \le k \le T, \tag{A16}
$$

whereas in setting τ_{is}^c he considers the stream

$$
\kappa_{s+k} \log \tau_{is}^c = (\alpha + \beta)^k \log \tau_{is}^c, \;\; 1 \le k \le T. \tag{A17}
$$

A.7 Proof of Lemma 1

We can derive from (17) that

$$
W_i^c|_{\gamma=1} - W_i^d|_{\gamma=1} = \frac{2\alpha}{1 - 2\alpha} \log \frac{\tau^c}{\tau^d} + \frac{2\alpha}{1 - 2\alpha} \log \frac{\overline{x}}{\sqrt{x_i x_j}} + \log \frac{1 - \tau^c}{1 - \tau^d} - \lambda^c. \tag{A18}
$$

Define $\phi \equiv \frac{1 - 2\alpha\delta}{\alpha\delta(1 - (2\alpha\delta)^T)}$. Then, from Assumption 1, ϕ strictly decreases in T. Furthermore, the sign of the partial derivative of $\frac{\tau^c}{\tau^d}$ with respect to ϕ is derived from (32) as

$$
\frac{\partial \left(\frac{\tau^c}{\tau^d} \right)}{\partial \phi} = \frac{\partial 2 \left(\frac{1+\phi}{2+\phi} \right)}{\partial \phi} > 0.
$$

Therefore, $\frac{\tau^c}{\tau^d}$ strictly decreases in T. Also, using (33) we can prove that $\frac{1 - \tau^c}{1 - \tau^d}$ strictly decreases in T in the same manner. It is thus demonstrated that $W_i^c|_{\gamma=1} - W_i^d|_{\gamma=1}$ in (A18) is strictly decreasing with regard to T.

A.8 Proof of Corollary 3

Let equality hold in (31). The RHS exceeds the LHS for a higher value of λ^c. To recover equality a higher value of $\frac{x_B}{x_A}$ is needed, since the LHS of (31) strictly increases in $\frac{x_B}{x_A}$.

References

[1] Alesina, A., I. Angeloni, and F. Etro (2005) International unions, *American Economic Review*, Vol. 95 pp. 602–615.

[2] Alesina, A., R. Baqir, and W. Easterly (1999) Public goods and ethnic divisions, *Quarterly Journal of Economics*, Vol. 114 pp. 1243–1284.

[3] Alesina, A., and D. Rodrik (1994) Distributive politics and economic growth, *Quarterly Journal of Economics*, Vol. 109 pp. 465–490.

[4] Arzaghi, M., and J. V. Henderson (2005) Why countries are fiscally decentralizing, *Journal of Public Economics*, Vol. 89 pp. 1157–1189.

[5] Barro, R. (1990) Government spending in a simple model of endogenous growth, *Journal of Political Economy*, Vol. 98 pp. 103–125.

[6] Battaglini, M., and S. Coate (2007) Inefficiency in legislative policymaking: a dynamic analysis, *American Economic Review*, Vol. 97, pp. 118–149.

[7] Bénabou, R. (1996) Heterogeneity, stratification, and growth: macroeconomic implications of community structure and social finance, *American Economic Review*, Vol. 86 pp. 584–609.

[8] Besley, T., and S. Coate (1998) Sources of inefficiency in a representative democracy: a dynamic analysis, *American Economic Review*, Vol. 88 pp. 139–156.

[9] Besley, T., and S. Coate (2003) Centralized versus decentralized provision of local public goods: a political economy approach, *Journal of Public Economics*, Vol. 87 pp. 2611–2637.

[10] Bolton, P., and G. Roland (1997) The breakup of nations: a political economy analysis, *Quarterly Journal of Economics*, Vol. 112 pp. 1057–1090.

[11] Brueckner, J. (2004) Fiscal decentralization with distortionary taxation: Tiebout vs. tax competition, *International Tax and Public Finance*, Vol. 11 pp. 133–153.

[12] Brueckner, J. (2006) Fiscal federalism and economic growth, *Journal of Public Economics*, Vol. 90 pp. 2107–2120. .

[13] Bucovetsky, S., M. Marchand, and P. Pestieau (1998) Tax competition and revelation of preferences for public expenditure, *Journal of Urban Economics*, Vol. 44 pp. 367–390.

[14] Easterly, W., and R. Levine (1997) Africa growth tragedy: policies and ethnic divisions, *Quarterly Journal of Economics*, Vol. 112 pp. 1203–1250.

[15] Ellingsen, T. (1998) Externalities vs internalities: a model of political integration, *Journal of Public Economics*, Vol. 68 pp. 251–268.

[16] Greene, K. V. (1985) Fiscal decentralization: evidence on the role of income and other determinants, *Public Finance*, Vol. 40 pp. 291–298.

[17] Hayashi, M., and M. Sato (2005) Globalization and decentralization, paper presented at the 61st Congress of the International Institute of Public Finance, ICC Jeju, Jeju, Korea, August 22–25.

[18] Kempf, H., and S. Rossignol (2005) Growth, inequality, and integration: a political economy analysis, *Journal of Public Economic Theory*, Vol. 7 pp. 709–739.

[19] Lockwood, B. (2002) Distributive politics and the costs of centralization, *Review of Economic Studies*, Vol. 69 pp. 313–337.

[20] Nishimura, Y. (2006) Human fallibility, complementarity, and fiscal decentralization *Journal of Public Economic Theory*, Vol. 8 pp. 487–501.

[21] Oates, W. (1972) *Fiscal Federalism*, New York: Harcourt Brace.

[22] Oates, W. (1993) Fiscal decentralization and economic development, *National Tax Journal*, Vol. 46 pp. 237–243.

[23] Panizza, U. (1999) On the determinants of fiscal centralization: theory and evidence, *Journal of Public Economics*, Vol. 74 pp. 97–139.

[24] Pommerehne, W. W. (1977) Quantitative aspects of federalism: a study of six countries, in *The Political Economy of Fiscal Federalism*, W. Oates, Ed., Lexington: D. C. Heath and Company.

[25] Poterba, J. M. (1997) Demographic structure and the political economy of public education, *Journal of Policy Analysis and Management*, Vol. 16 pp. 48–66.

[26] Rubinchik-Pessach, A. (2005) Can decentralization be beneficial? *Journal of Public Economics*, Vol. 89 pp. 1231–1249.

[27] Schwager, R. (1999) Administrative federalism and a central government with regionally based preferences, *International Tax and Public Finance*, Vol. 6 pp. 165–189.

[28] Wellisch, D. (2000) *Theory of Public Finance in A Federal State*, Cambridge: Cambridge University Press.

[29] Wellisch, D., and W. F. Richter (1995) Internalizing intergenerational externalities by regionalization, *Regional Science and Urban Economics*, Vol. 25 pp. 685–704.

[30] Wildasin, D. E. (1986) *Urban Public Finance*, Chur, Switzerland: Harwood Academic Publishers.

[31] Wilson, J. D., and E. Janeba (2005) Decentralization and international tax competition, *Journal of Public Economics*, Vol. 89 pp. 1211–1229.

Figure 1.

Preference Relations of High-Productivity District B
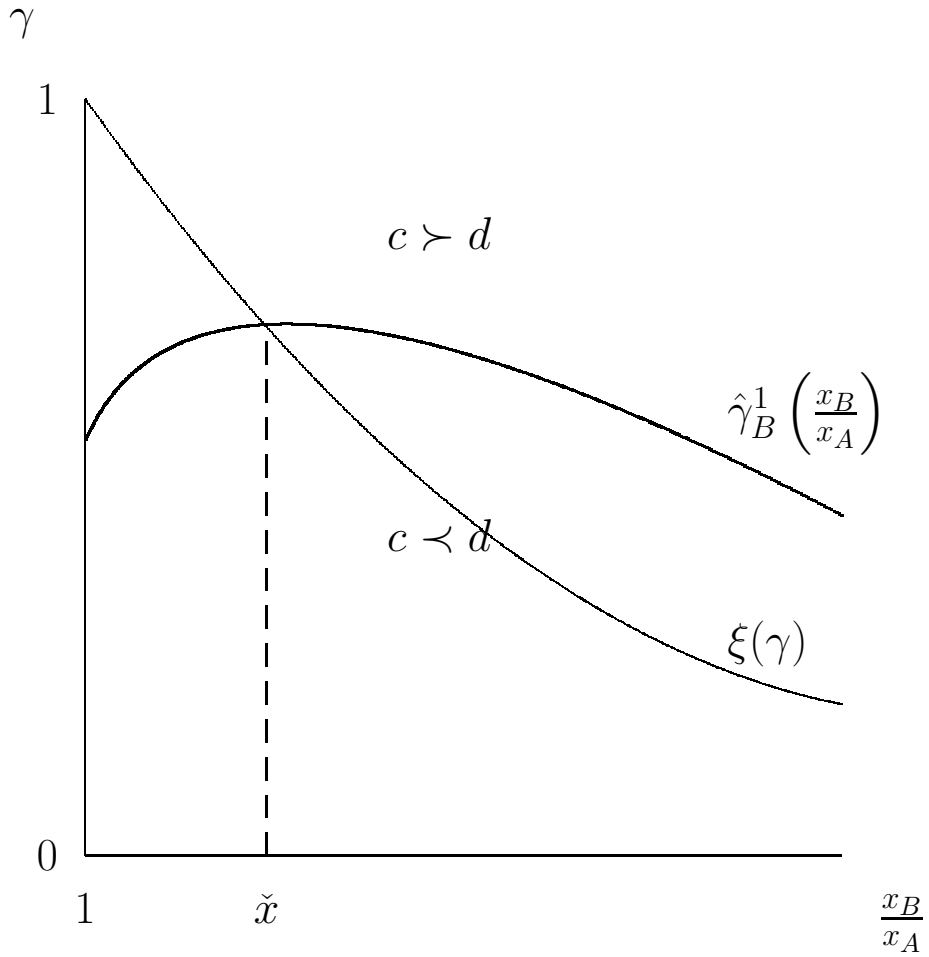
over Two Fiscal Regimes $(T = 1)$

Figure 2.

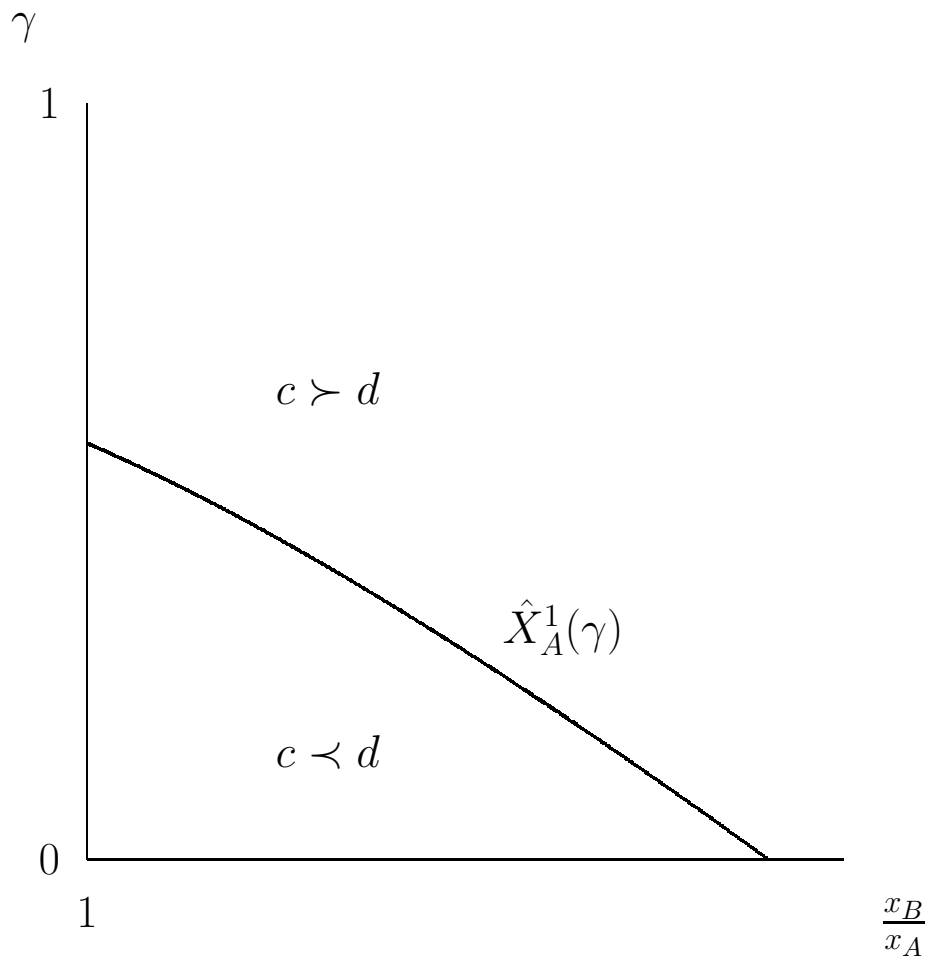## Preference Relations of Low-Productivity District A

over Two Fiscal Regimes $(T = 1)$

Figure 3.

Effects of Policy Choice

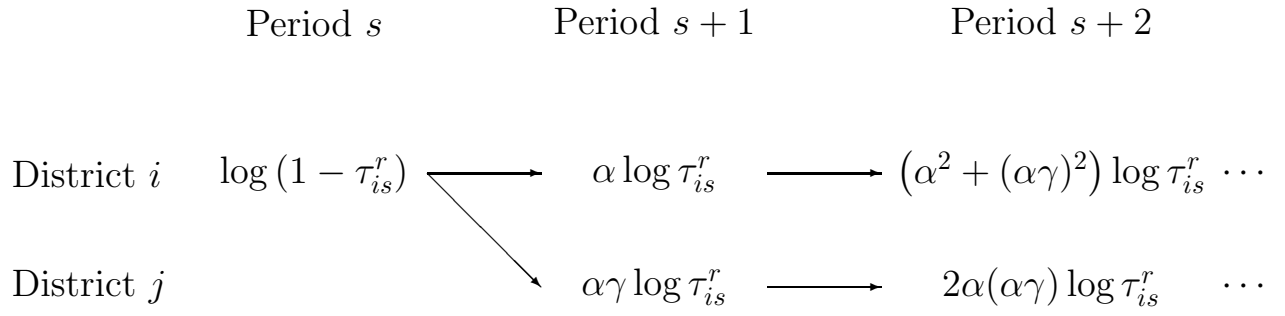|	Period s	Period $s+1$	Period $s+2$
District i	$\log\left(1 - \tau_{is}^r\right)$	$\alpha \log \tau_{is}^r$	$\left(\alpha^2 + (\alpha\gamma)^2\right) \log \tau_{is}^r \;\cdots$
District j		$\alpha\gamma \log \tau_{is}^r$	$2\alpha(\alpha\gamma) \log \tau_{is}^r \quad \cdots$

Figure 4.

Preference Relations of Districts A and B

over Two Fiscal Regimes $(\gamma = 1)$

$c \prec d$ $c \succ d$

1 $\tilde{X}^T$ $\tilde{X}^{T+1}$ $\frac{x_B}{x_A}$