Fiscal Decentralization, Political Heterogeneity and Welfare

Erkmen Giray Aslim† and Bilin Neyapti‡

Abstract

Theoretical and empirical literature on fiscal decentralization (FD) has been thriving, while understanding the welfare implications of FD under political diversity necessitates further investigation. We contribute to this literature by formally modeling the interaction between the central government and local governments, where the latter may have varying degrees of political proximity to the former. The model solution reveals that the optimal tax rate is positively associated with FD, political unison, and spillovers across localities, while the local tax collection effort is negatively associated with all of these parameters. The first novel finding of this study is that both the welfare and the central government's utility peak and income distribution is more equitable at a lower level of FD when spillovers exist than otherwise, which supports the decentralization theorem. The second novel finding is that both the amount of redistributable income and central government utility increases with political unison.

Keywords: Fiscal decentralization; fiscal efficiency; welfare

JEL Codes: E62, H71, H72

†Texas A&M University, Private Enterprise Research Center, College Station, TX 77845 USA. E-mail: easlim@tamu.edu.

‡Bilkent University, Department of Economics, Bilkent, Ankara 06800 TURKEY. E-mail: neyapti@bilkent.edu.tr.

1. Introduction

The literature on the welfare effects of fiscal decentralization (FD) has been expanding since the seminal works of Buchanan (1950) and Tiebout (1956). Oates (1972) stated that FD is welfare-improving if the central government provides a uniform public good and faces the same cost structure as the sub-national or local governments, and if there are no spillovers across localities. Coined as the *decentralization theorem*, the postulate is that local governments (LG)[1] are better positioned to learn and respond to local preferences and, hence, they are likely to be more efficient in public good provision than the central government (CG). The theorem is complemented by the argument that decentralization of fiscal activities may reduce transaction costs since LGs can be more transparent in their actions and more accountable for their decisions than CG.[2] The justification for FD as an institutional mechanism to achieve fiscal efficiency is stronger the more heterogeneous the society is, and thus increases the variation across local preferences.

However, local capacity constraints and reliance of LGs on central transfers for their expenses may prevent the potential welfare gains from FD.[3] Furthermore, CG tends to make transfers strategically either to maximize their re-electability and/or redistribute towards its own constituency.[4] When public spending is prioritized by the political proximity of LGs to the center, transfers may become economically inefficient.[5] Thus, fiscal rules as well as political institutions, such as the nature of the electoral system, legislation and party structures, may play a role on the effectiveness of FD

[1]We will refer to the sub-national governments as *local governments* (LGs), and the sub-national jurisdictions as *localities* throughout the paper.

[2]See Ligthart and van Oudheusden (2015) for empirical evidence on the positive relationship between trust in the government and FD.

[3]The fiscal decentralization indicators of the World Bank demonstrate that central transfers constitute a sizable share of local government resources even in developed countries.

[4]Empirical evidence supports that redistributive choices tend to be systematically related with electoral incentives. Dellmuth and Stoffel (2012), for example, provided evidence in the case of the use of European structural funds in Germany. Also, Solé-Ollé and Sorribas-Navarro (2008), Arulampalam et al. (2009), Brollo and Nannicini (2012), Bracco et al. (2015) all showed a strong effect of political alignment on intergovernmental transfers in Spain, India, Brazil and Italy, respectively. For a broader discussion of the political economy of intergovernmental transfers, see also Inman and Rubinfeld (1996), Khemani (2007) and Sato (2007).

[5]Well-designed and transparent transfer mechanisms are therefore crucial for eliminating discretionary or politically-oriented redistributive policies and to achieve efficient and equitable outcomes. Ma (1997) discussed the types of fiscal transfer rules. In a cross-sectional study, Neyapti (2013) showed that fiscal rules have significant effects on the fiscal disciplining impact of FD. Bulut-Cevik and Neyapti (2014) and Akin, Bulut-Cevik, and Neyapti (2016) also investigated the welfare effects of transfer rules.

in delivering fiscal efficiency.[6] These factors also define the socio-political environment within which, interactive with various other structural and economic factors, the extent of FD may be determined.[7] The theoretical literature on FD generally addresses the question of whether the decentralization theorem survives after relaxing the assumptions of benevolent government and uniform public good provision under centralization.[8] Both the empirical and theoretical literature on the relationship between political factors and the benefits of FD, however, have been inconclusive, calling for further analysis.[9]

We contribute to this literature by presenting an original formal framework to investigate the welfare implications of the interaction between the central and local governments, where the latter are heterogeneous both in their income levels and political orientation. Facing an exogenously given level of FD that is uniform across the localities, we model CG as choosing a general tax rate (t) optimally, whereas LGs choose their tax collection effort.[10] Our framework differs from the earlier studies mainly in two respects. First, LGs' political proximity (denoted by p) to CG is taken into account in the optimization by CG. Second, rather than conducting a comparative welfare analysis of the centralized versus decentralized fiscal regimes, FD is allowed to vary between zero

[6] See, for example, Inman and Rubinfeld (1997), Besley and Coate (2003) and Neyapti (2013).

[7] North and Weingast (1989) conjectured decentralization as an institutional mechanism to constrain the fiscal policies of the opposition in case the incumbent party faces a high probability of losing the elections. O'Neill (2003) argued that the adoption of FD is linked with optimizing the political power via securing higher political support in sub-national elections. In federal states, the allocation of power is specified in the constitution whose amendment may require either regional votes or referendum.

[8] Besley and Coate (2003) argued that, when regional and central governments bargain for delegation, centralization does not necessarily imply uniformity of public good provision; they showed that centralization can welfare dominate FD even when regions are heterogeneous and there are no spillovers across regions. Lockwood (2008) argued that the decentralization theorem fails only when the benevolence assumption is replaced by direct democracy or majority voting; decentralization can be welfare-dominating even when regions are homogeneous and there are positive externalities. Gonzalez et al. (2006) argued that the welfare effects (measured by the extent of political business cycles) of FD, vis-à-vis centralization, depends on the extent of the political rents of the central government in a majority voting model. Janeba and Wilson (2011) argued that optimal fiscal decentralization is affected by tax competition and spillovers.

[9] Enikolopov and Zhuravskaya (2007), for example, showed empirically that the outcomes of FD improve with the strength of national political government, but deteriorate with the extent that local administration is subordinated to the CG. Likewise, Ponce-Rodríguez et al. (2016) argued that democratic decentralization combined with a strong CG contribute to FD's effectiveness even when regional spillovers exist. In accordance with these studies, the positive association of both democracy and governance quality with the desired outcomes of FD has been well-reported in the literature (see, for example, Stepan (1999), De Mello and Barenstein (2001), Arzaghi and Henderson (2005), Kyriacou and Roca-Sagalés (2011) and Altunbaş and Thornton (2012)). Eaton (2001), however, reported no significant relationship between such political factors and the outcomes of FD.

[10] As in Besley and Coate (2003) and Lockwood (2008), the centrally determined tax rate is uniform across the economy. In this setup, the optimal choice of FD by CG is trivial since it would be optimal to centralize all the public good provision when CG is politically oriented.

and one so as to investigate welfare-optimizing level of FD.

Our model's solution implies that t is positively related with FD, while it is negatively related with the local tax collection effort, as in Aslim and Neyapti (2017). The novel findings of the present study are in regard to the interrelationships between FD, welfare, income distribution, political polarization (measured by the squared distances between localities' political proximity to the center) and political unison (denoted by P). We measure P as the sum of p's, which indicates the degree of political or ideological cohesion. Our findings indicate that both optimal t and tax revenue increases in P.[11] Second, we show that LG's tax collection efficiency decreases in FD, and this effect increases in both FD and P. Third, simulations reveal that both tax revenue and CG's utility peak at an intermediate level of FD. The finding that tax revenue declines at the extreme values of FD conforms to the *decentralization-Laffer curve*, as has also been noted in Aslim and Neyapti (2017). Simulation results also demonstrate that CG's utility increases in P.

As an extension, we introduce spillovers into the model, which can then be solved in a leader-follower framework where the leader is a representative LG. It is observed that while optimal tax collection effort declines, the optimal tax rate increases in spillovers across localities, resulting in an ambiguous effect of spillovers on tax revenues. Also, the level of FD for which the maximum central government utility is at its maximum is observed to be lower in the case of spillovers than otherwise, yielding support for the decentralization theorem. A notable observation is that the level of FD that maximizes CG's utility in the model is the observed OECD average of 0.3.[12] The results for polarization, on the other hand, remain ambiguous in the presence of spillovers. Noting that spillovers across localities as well as vertical imbalances are a fact, the positive revenue effect of political cohesion, coupled with its ambiguous effects on LG's utility and income distribution, point at the need for institutional mechanisms, such as transparent fiscal rules, for reducing excess tax increases and possible political biases in fiscal transfers.

In what follows, the model and its solution is presented in Section 2. In Section 3, the comparative statics and simulation results are presented. In Section 4, we present an extension of the benchmark model by introducing spillovers across localities, and provide a comparative analysis of the two models' results. The conclusion and discussion of the

[11] Similarly, the positive impact of socio-economic homogeneity on macroeconomic outcomes have been shown empirically in Knack and Keefer (1997), Knack and Zak (2003), and theoretically in Neyapti (2017).

[12] Based on the World Bank database of "Fiscal Decentralization Indicators" for 1997.

possible extensions of the paper is provided in Section 5.

2. The Model

We aim to explore the relationship of political unison and polarization with welfare given a degree of fiscal decentralization. To this end, we consider a partial equilibrium model comprised of utility maximizing central and local governments (denoted by CG and LGs) facing an exogenous level of FD. In order to focus on the welfare implications of FD in case of political diversity across localities, we make the following simplifying assumptions. The economy is closed and local preferences are symmetric. We consider a static model where the degree of FD, as well as each locality's income (Y_i) and political position vis-à-vis CG (p_i) are given exogenously. We assume that there is no resource mobility or tax competition across the LGs. In the benchmark model, we also assume that there are no spillovers across localities, which we relax later.

The following describes the key variables of the model. Total spending in region i, which is denoted by $\tilde{Y}_i$ (where $i = 1 \ldots n$ is the number of localities),[13] is assumed to be equal to the sum of the private (C_i) and public sector spending. Public sector spending consists of spending by CG and LGs, denoted by G_i^C and G_i^L, respectively, which are considered as the levels of centrally and locally provided public goods. Hence, income and its components in locality i are given by Equations (1) to (4):

$$\tilde{Y}_i = C_i + G_i^L + G_i^C, \tag{1}$$

where

$$C_i = (1 - [a_i\phi + (1-\phi)]t)Y_i; \tag{2}$$

$$G_i^L = \phi a_i t Y_i; \tag{3}$$

$$G_i^C = (1-\phi)t\hat{p}_i \sum_i Y_i \tag{4}$$

for all i. In Equations (2)-(4), ϕ ($\phi \in [0,1]$) stands for the degree of FD. We define a_i ($a_i > 0$) as LG's relative tax collection effort vis-à-vis that of CG in locality i. As we

[13] Total spending ($\tilde{Y}_i$) the initial, exogenously given, level of income (Y_i) by the amount of transfers made by the central government. However, for the whole economy, $\sum_i \tilde{Y}_i = \sum_i Y_i$, since the overall government budget balances.

do not assume ability and cost differentials across LGs and CG, we interpret the case of '$a_i > 1$' merely as LG$_i$'s willingness to collect more tax than CG in its jurisdiction. Accordingly, the case of '$a_i < 1$' indicates that LG's tax collection effort is less than that of CG in locality i.[14]

Equation (2) shows private consumption as the after-tax income, where t ($t \in (0, 1)$) is the flat tax rate set by the central government. The tax revenue in each locality is either collected by the local government ($a_i \phi t Y_i$) or by the center ($(1 - \phi)tY_i$). Hence, the *effective tax rate* for region i (t_i) is given by:

$$t_i = [a_i \phi + (1 - \phi)]t, \tag{5}$$

which is constrained by the unit interval ($t_i \in (0, 1)$) for feasibility.[15]

Equations (3) and (4) imply that both LGs and CG are assumed to follow a balanced budget rule. LG's spending (G_i^L) is thus a fraction of the total local tax revenue, determined by the product of a_i and ϕ. In the extreme case of $\phi = 1$, all the tax revenue is collected and spent by LGs. When $\phi = 0$, on the other hand, CG becomes the agent who collects all the revenues and does all the public spending in locality i.

While the above features of the model follow Aslim and Neyapti (2017), we depart significantly from that model in the way we define G_i^C. In Equation (4), $\hat{p}_i$ (defined as $\hat{p}_i = p_i / \sum_i p_i$, where $\{\hat{p}_i,\, p_i\} \in [0, 1]$), represents the *relative* proximity of LG$_i$ to CG's political position or ideology, taking into account that of the remaining LGs. Because the total of p_i's need not add up to one, we weight each locality's share of the central pool of revenues by $\sum_i p_i$ so that $\sum_i \hat{p}_i = 1$, which ensures that CG's budget balances. The share of total revenue pool that CG spends in locality i is then $\hat{p}_i$; that is, redistribution is assumed to be solely politically-driven. Summing G_i^C in Equation (6) over $i's$, we obtain total central government spending:

$$G^C = \sum_i G_i^C = (1 - \phi)t \sum_i Y_i. \tag{6}$$

Given this identification, when each LG's political position are identical to each other, or is the same as that of CG, CG acts like a benevolent government and spends the same

[14]Differential local capacity constraints and transaction costs may also affect the value of a_i, which we do not take into consideration in the current model for the purpose of simplicity. For an empirical analysis of these issues, see, for example, De Mello (2000), Treisman (2006), and Freinkman and Plekhanov (2009).

[15]Note that t_i is not feasible ($t_i > 1$) when $a_i > 1 + (1/\phi)(1 - t)/t$.

amount in each region. In an extreme case where some of the p_i's are zero, G_i^C in those localities would be zero. In an extreme case of a non-ideological or purely benevolent government or when all p_i's are identical to 1, $\sum_i p_i = n$ and CG's spending in each region is an equal share of the total CG revenue pool: $G_i^C = (1-\phi)\frac{t}{n}\sum_i Y_i$.[16]

The total tax revenue is given by the following expression, where local incomes constitute the only tax base:

$$T = t\sum_i (a_i\phi + (1-\phi))Y_i, \tag{7}$$

which is also equivalent to total government spending, G, where $G = \sum_i G_i^L + G^C$. Hence, the overall government budget constraint holds, as do the central and the local governments'.

We assume that, absent a transfer mechanism, the central government reallocates resources to those localities that are close to its own political ideology. Considering that many countries lack fair and strictly enforced transfer mechanisms, p_i (or $\hat{p}_i$) can be considered to represent a key policy parameter. Net transfers to any region i is then given by the following, which expresses the relation between the *ex-ante* and *ex-post* (i.e., before and after taxes and transfers) local incomes:

$$\left(\tilde{Y}_i - Y_i\right) = (1-\phi)t\left(\frac{p_i}{\sum_i p_i}\sum_i Y_i - Y_i\right). \tag{8}$$

In what follows, the optimization problems of LGs and CG are described. The non-cooperative game, defined by the joint solution of a_i's and t selected optimally by LGs and CG, respectively, yields a Nash equilibrium as demonstrated below.[17]

[16]Note that as different from Aslim and Neyapti (2017), in this paper we consider that CG delivers local, rather than pure, public good; hence $G^C \neq G_i^C$.

[17]The solution to a non-cooperative game between the three agents, local governments, the central government and a social planner (SP), is also explored; this is a problem similar to that of the central government given above, except with p_i=1/2. This scenario is based on the joint solution of the local and central government problems, which, however, fails to yield a solution. That is, there is no common set of parameters that satisfies the set of optimal solutions for the general tax rate, the level of fiscal decentralization and the level of local public good provision (or local tax collection effort). The problem of SP and LG, on the other hand, is separately analyzed in Aslim and Neyapti (2017).

2.1. A Representative Local Government's Problem

We define the optimization problem of LG, following Aslim and Neyapti (2017): a representative LG chooses its tax collection effort in order to maximize its utility, which is composed of public (local and central) and private spending in that locality. It is assumed that LGs do not distinguish between the locally and centrally provided public good, G_i^L and G_i^C, which is fully justified if G_i^C is in the form of open-ended central transfers to local governments. We assume that LG's utility is in a log-linear utility form:

$$U_i^{LG} = \alpha \ln C_i + \beta \ln G_i^L + \beta \ln G_i^C, \tag{9}$$

which is maximized subject to the constraints given in Equations (2) through (4). Hence, the unconstrained problem becomes:

$$\max_{a_i} \; U_i^{LG} = \alpha \ln((1-t_i)Y_i) + \beta \ln(\phi a_i t Y_i) + \beta \ln((1-\phi)t\hat{p}_i \sum_i Y_i). \tag{10}$$

The first order condition for the above problem is:

$$a_i = \left(\frac{\beta}{\alpha+\beta}\right) \frac{1-t+\phi t}{\phi t}. \tag{11}$$

Proof: *Appendix A*

The negative relationship between a_i and t arises due to the substitutability between C_i and G_i in the utility function.

2.2. Central Government's Problem

CG chooses t in order to maximize the aggregate utility, given a level of ϕ.[18] The objective function of the central government differs from the sum of the LGs' objective functions by the relative utility weights on G_i^L's, which are p_i's that indicate the degree of substitutability between G_i^L and G_i^C for CG. Hence, the CG's optimization problem is:

$$\max_{t} \; U_i^{CG} = \sum_i \left(\alpha \ln C_i + p_i \beta \ln G_i^L + \beta \ln G_i^C\right), \tag{12}$$

[18]The inclusion of p_i is similar to Lockwood's (2008) inclusion of special interest groups in the utility function.

which indicates that CG gets higher utility from G_i^L the closer the political position of locality i to itself. This problem is solved after substituting the constraints given in Equations (2) through (4):

$$\max_t \; U_i^{CG} = \sum_i \left(\alpha \ln((1-t_i)Y_i) + p_i \beta \ln(\phi a_i t Y_i) + \beta \ln((1-\phi)t\hat{p}_i \sum_i Y_i) \right). \tag{13}$$

Let's define $P = p_1 + p_2$ as *political unison* ($P \in [0, n]$). Given that LGs are symmetric in their optimizing behavior [see Equation (11)] and, hence, $a_i = a_1 = a_2$, where $n = 2$ in the current specification, the first order condition of the CG problem yields:

$$t = \frac{\beta(P+2)}{(1-\phi(1-a_i))(2\alpha+\beta(P+2))}. \tag{14}$$

Proof: *Appendix A*

Equations (11) and (14) stand for the best responses of LG and CG, respectively, to the other player's action. Given Equations (1) to (4), the Nash equilibrium of the model is defined as the set of $\{t, a_i\}$ that satisfy Equations (11) and (14).

Lemma 1. The joint solution of Equations (11) and (14) for $i \in \{1, 2\}$ yields the following solutions:

$$t^* = \frac{\beta P}{(1-\phi)(2\alpha+\beta(P+2))}; \text{ and } a_i^* = \frac{2(1-\phi)}{\phi P}. \tag{15}$$

Proof: *Appendix A*

Using Equation (5) and the optimal a_i and t expressions in Equation (15), we obtain the effective tax rate in locality i: $t_i = \beta(P+2)/(2(\alpha+\beta)+\beta P)$. The resulting optimal transfers satisfy the efficiency condition: $\sum_i Y_i = \sum_i (C_i + G_i^L + G_i^C)$ for $a_i > 1 + (1/\phi)(1-t)/t$, which ensures the feasibility of t_i based on Equation (5).

For uniqueness of the Nash equilibrium, corresponding to each exogenous quadruple in the set $\{\alpha, \beta, \phi, P\}$ there should exist a unique corresponding pair of endogenous variables $\{t^*, a_i^*\}$. As both reaction functions [Equations (11) and (14)] are negatively sloped, the uniqueness of the optimal solution hinges on the absence of corner solutions. Equation (11) rules out the possibility that $t = 0$, for which a_i is undefined. It is also clear from Equation (15) that $a_i^* = 0$ holds when $\phi = 1$, which is the case of full decentralization; this is also not feasible since it implies that $t_i = T = 0$ [see Equation (7)], which means zero tax collection and thus zero public good provision. Hence, corner solutions are not feasible. Therefore, the joint solution of the problem shown by Lemma 1 exists and is unique. The comparative statics and the welfare and income distribution implications of the model are presented in what follows.

3. Comparative Statics

In this section we investigate how the underlying model parameters $\{\alpha, \beta, \phi$, and P (or p_i) $\}$ affect the optimal choices of the central and local governments. Table 1 presents the signs of the partial derivatives of the optimal solutions given in Equation (15).

Table 1: The Signs of the Partial Derivatives, Benchmark Case

	t^*	a_i^*
ϕ	+	-
p_i	+	-
α	-	0
β	+	0

Table 1 shows that ϕ affects t^* positively, but a_i^* negatively. The first of these arises as CG compensates for its declining revenues when ϕ increases. The second effect is due to LG's effort to compensate for the loss in its utility as disposable income and C_i decrease when t^* increases.

Proposition 1. An increase in ϕ leads to a decrease in a_i^* and an increase in t^*.

Proof: *Appendix A*

Corollary 1. The negative response of a_i^* to ϕ increases in both ϕ and P.

Proof: *Appendix A*

In view of these opposing effects, the net impact of ϕ on the tax revenue, when evaluated based on the optimal values of t and a_i, is ambiguous. The effect of the rest of the model parameters on the sign of this effect is further investigated via simulations in the next section.

Table 1 also indicates that p_i has a negative effect on a_i^*, and a positive effect on t^* (the same results hold for P). These opposing effects can be explained as follows. Ceteris paribus, CG derives higher utility from G_i^L the higher p_i is [see Equation (12)], which compensates for the utility loss arising from a decrease in C_i that arises in response to an increase in t^*. C_i also falls in a_i that is reduced by LG's response to an increase in transfers as p_i increases [see Equation (9)]. Intuitively, this also means that CG accepts a greater degree of crowding out the greater p_i is. The net effect of p_i on the effective tax rate and total tax revenue, however, is ambiguous and will be explored via simulation analysis.

Proposition 2. The greater is P, the higher is t^*.

Proof: *Appendix A*

It is also observed that increasing p_i increases the negative relationship between ϕ and a_i^*.

Corollary 2. The negative response of a_i^* to P increases in P.

Proof: *Appendix A*

Thus, it is optimal for CG to tax more in case of a greater degree of political unison. Given the reduction in a_i^*, however, the net effect of p_i on the effective tax rate t_i, or the tax revenue (T) is ambiguous. Hence, we also resort to a simulation analysis to explore this effect further.

Table 1 also shows that the optimal tax rate is positively related with the utility share of the public good (β), and negatively with that of the private good (α). The explanation for this is straightforward from the utility of the central government (U^{CG}), which increases in t^* the higher the utility share is of the public spending and the lower the share of private spending.[19] The results are the same in nature when t is replaced

[19] While a_i^* does not have a direct relationship with α and β, simulations show that total tax collection

with T; the higher the utility share of the public good, the higher is the tax revenue.

3.1. Simulations

The above comparative statics leave the effects of the model parameters on the effective tax rate (t_i), tax revenue (T) and the rest of the model aggregates ambiguous. In this section, we investigate how $\{\tilde{Y_1}/\tilde{Y_2}, t_i, T, U^{LG}, U^{CG}\}$ respond to the changes in the model parameters: $\{\phi, p_i, \alpha, \beta\}$, where $\tilde{Y_1}/\tilde{Y_2}$ denotes *ex-post* income distribution and U^{LG} denotes the utility of the local government. To obtain simulation data on income distribution, Y_1 is fixed arbitrarily and Y_2 is defined as some multiple (x) of it, where $x \in [0.1, 5]$. Hence, we set locality 1's income such that it can be as small as the one-tenth, or as large as five times the income of locality 2.

The simulations reported in Table 2 correspond to the feasible ranges of the underlying model parameters, given the feasibility constraints: $\{\phi, p_i, \alpha, \beta\} \in [0, 1]$; $\{t^*, t_i\} \in (0, 1)$; and for the remaining endogenous variables: $\{a_i^*, a_i^* t^*, C_i, G, G_i^C, G_i^L, \tilde{Y_i}\} \in \mathbb{R}_+$, for all i. The data set obtained using the Matlab software is composed of 410,310 observations. Based on this data set, several additional observations are made. In the following, we report only the cases of definite relations that are revealed from the boxplot analysis of the simulated data (see *Appendix B*).

Table 2: Calibration of the Parameters

Parameter	Range	Increment
x	[0.1, 5]	0.5
ϕ	[0.1, 1]	0.1
α	[0.1, 1]	0.1
β	[0.1, 1]	0.1
p_i	[0.1, 1]	0.1

Based on the simulation data, we observe that tax revenue (T) shows a positive relationship with t^* since the optimal tax rate is utility maximizing, only the rising portion of the traditional Laffer curve is observed (see Figure B1).

effort increases in α and decreases in β strictly (not shown, the results are available from the author upon request).

Remark 1. T increases in t^*.

It is also observed that T first increases and then decreases in ϕ, supporting the *decentralization-Laffer curve* relationship proposed in Aslim and Neyapti (2017). More precisely, t^* increases in ϕ up to $\phi \lesssim 0.5$ so as to overcompensate for the reduction in the tax collection effort (see Figure B2). For $\phi \gtrsim 0.5$, however, the reduction in a_i^* dominates the increase in t^*. This observation conforms to the recent consensus in the fiscal decentralization literature: the extreme values of ϕ do not contribute to fiscal efficiency and welfare. Cross-country data on tax revenues and ϕ depicts a picture that is consistent with this finding (see Figure C1 in *Appendix C*).[20] By this account, it can be said that the county-year observations on the right side of the curve in Figure C1 may be indicating unoptimally high levels of ϕ.[21]

Figure B3 also demonstrate that U^{LG} (also U^{CG}, not shown)[22] and $(\tilde{Y_1}/\tilde{Y_2})$ depicts a relationship with ϕ that is similar to the relationship of T with ϕ (see Figure B4).

Remark 2. T depicts a *decentralization-Laffer curve* relationship.

Next, we investigate the relationship between the political variables and the endogenous variables of the model. We define *political unison*, denoted by P, as the sum of p_i's; and *political polarization*, denoted by σ_P, where $\sigma_P = (p_1 - p_2)^2$.[23] P can be viewed as the degree to which the society's political choices are in accord with that of CG, while σ_P measures the degree to which LGs are diverted from each other with respect to their ideological or political positions. Simulations reveal that T clearly increases in P, while the relationships of both t^* and T with σ_P are ambiguous.

Remark 3. T increases with P.

[20] Figure C1 is based on 139 country-year observations, where ϕ is measured by expenditure decentralization, expressed as the subnational government shares, that are obtained from the Fiscal Decentralization Indicators dataset of the World Bank, and tax revenues as percentage of GDP are sourced from the World Development Indicators of the World Bank.

[21] The outliers on the left side in Figure C1 are Israel, Madagascar and Zimbabwe during 1970s and 1980s. The unoptimally high FD ratios on the right side, corresponding to low T, include Canada and the US.

[22] We consider that U^{LG} better represents welfare than U^{CG} since the latter is politically weighted.

[23] There is a double-peaked hump-shaped relationship between σ_P and P. Thus, note that for $p_1 = p_2 = 0.5$; P is 1 but σ_P is 0; for $p_1 = p_2 = 0.7$, P is 1.4 but σ_P is still 0. The upper bounds of σ_P and P are equal to 1 and 2, respectively, in the simulations for i=1, 2.

On the other hand, neither a_i or t_i depict a definitive relationship with the political variables. Likewise, income distribution does not depict a clear relationship with the political variables.[24] The effects of P and σ_P on LG's and CG's utilities are also examined, as these relations are not clear at the onset. Simulations reveal show that while U^{CG} clearly increases in P its relation with σ_P is ambiguous; U^{LG} shows no clear relationship with the political variables, however.

3.2. Extension: Introducing Spillovers Across Localities

In this section we investigate how spillovers of local public goods affect our benchmark results reported in the foregoing sections. We consider spillovers as either positive or negative externalities arising from local or central governments' spending in the neighboring locality. Positive externalities may arise in the form of spillovers from increasing incomes and mutually beneficial cultural and trade relations across localities, while negative externalities may arise from tax exporting and spillovers of socio-political instability and environmentally degrading economic activities. An investigation of the spillover effects on welfare, political polarization and FD will also help to reveal whether spillovers lower the benefits from fiscal decentralization, as stated by the decentralization theorem.[25]

Two scenarios are considered for the joint solution of LG and CG problems. In the first, LGs and CG engage in a non-cooperative strategic game, and in the second, they play a leader-follower duopoly game. In the first scenario, two cases are explored: in the case of asymmetric information, where only LG's are fully informed about the extent of the spillovers (s_i) and CG assumes $s_i = 0$, the problem remains the same as in the benchmark model. In the case of full information, however, the optimal solution to the simultaneous-move game does not yield any feasible solution for t.

As a second scenario, LG is assumed to be the leader and makes its decision by taking the reaction function of CG into account. LGs' objective function with spillover effects

[24]Figure C2 shows this empirically, based on 311 country-year observations, where measures of ethnolinguistic and religious polarization are used as a proxy for *political polarization* (based on Montalvo and Reynal-Querol (2005)), and the GINI index, as a measure of income inequality, sourced from the World Development Indicators of the World Bank. Similar observations can be made when the measures of religious polarization and fractionalization are used instead of the ethnolinguistic dimension, which are therefore not reported here.

[25]By contrast, Koethenbuerger (2008) argued that welfare gains of FD may increase under spillovers.

is given by:

$$\max_{a_i} U^{LG}_{i,spillover} = \alpha lnC_i + \beta \left(\left[lnG^L_i + \beta lnG^C_i \right] + s_i \left[lnG^L_j + lnG^C_j \right] \right), \tag{16}$$

where $i, j = 1, \ldots, n$ and $s_i \in [-1, 1]$ stands for the extent to which total public spending in locality j affects locality i. Assuming full information about the spillovers across the localities, we consider that the central government's problem becomes[26]:

$$\max_{t} U^{CG}_{i,spillover} = \sum_{i,j,i \neq j} \left(\alpha lnC_i + p_i \beta (lnG^L_i + s_i lnG^L_j) + \beta lnG^C_i \right). \tag{17}$$

where CG derives utility from LGs' expenditures by the extent of their proximity to its own political values, but benefits fully from its own spending in each region. As in the benchmark model, both of the above problems are subject to the constraints given by Equations (2) to (4), and $n = 2$ is assumed to obtain an explicit solution. The first order condition of LGs' problem is identical to the case of no spillovers. The solution of the CG's problem, given the symmetrical reaction functions of LG' yields the following expression:

$$t = \frac{\beta \sum_i (p_i(1 + s_i) + 1)}{(1 + \phi(a_i - 1)) \left(2\alpha + \beta \sum_i (p_i(1 + s_i) + 1)\right)}. \tag{18}$$

Proof: *Appendix D*

Lemma 2. Substituting Equation (18) into Equation (16), the solution of the LG's problem yields:

$$a^*_{i,spillover} = \frac{1 - \phi}{\phi(1 + 2s_i)}. \tag{19}$$

Substituting $a^*_{i,spillover}$ back in Equation (18) yields:

$$t^*_{spillover} = \frac{\beta(1 + 2s_i) \sum_i (p_i(1 + s_i) + 1)}{2(1 - \phi)(1 + s_i)(2\alpha + \beta \sum_i (p_i(1 + s_i) + 1))}. \tag{20}$$

Proof: *Appendix D*

[26]Not taking into account the spillovers from G^C across localities does not alter the optimal solutions.

The comparative statics of the new equilibrium variables are reported in Appendix D and summarized in Table 3, which matches the signs summarized in Table 1, with the exception that p_i has now have no effect on $a^*_{i,spillover}$. This means that, unlike in the benchmark case of no spillovers, $a^*_{i,spillover}$ is not affected by p_i.

Proposition 3. In case of local spillovers, p_i has no effect on a^*_i but affects t^* positively.

Proof: The proof is trivial for a^*_i as Equation (19) does not contain p_i. See *Appendix D* for t^*.

Two additional observations to those reported in Table 1 pertain to the effect of spillovers; while $t^*_{spillover}$ responds positively to s_i, the sign of the effect is the opposite for $a^*_{i,spillover}$. The interpretation of this observation is similar to the opposing effects of ϕ on a^*_i and t^* in the benchmark case: an increase in spillovers induces CG to increase the optimal tax rate in reaction to the reduced incentives for LGs to exert effort to collect the local taxes. This implies that, ceteris paribus, the effects of spillovers on the effective tax rate, tax revenue, and thus income distribution, are not certain. We find that both of these effects increase in ϕ, indicating that the higher is ϕ, the greater are the effects of s_i on $a^*_{i,spillover}$ and $t^*_{spillover}$. Hence, it is possible to say that spillovers reduce the welfare gains of decentralization, the more so the larger ϕ gets.

Table 3: The Signs of the Partial Derivatives,
Spillovers Case

	$t^*_{spillover}$	$a^*_{i,spillover}$
ϕ	+	-
p_i (or P)	+	0
α	-	0
β	+	0
s_i	+	-

Proposition 4. The higher is s_i, the higher is $t^*_{spillover}$ and the lower is $a^*_{i,spillover}$.

Proof: *Appendix D*

A simulation analysis is performed to compare the rest of the model implications with those of the benchmark case.[27] As implied by the opposite signs for $t^*_{spillover}$ and $a^*_{i,spillover}$ in Table 3, simulations show no direct relationship between the s_i's and T, U^{CG} and U^{LG}; each of these are maximized and $\tilde{Y_1}/\tilde{Y_2}$ is minimized (at 1.45) when $s_i = 0.8$, however.

Next, the relationships of the model's endogenous variables with $\{\phi, P, \sigma_P\}$ are reported in Table D1 to enable a comparison with the benchmark case (see *Appendix D*). A notable change from the benchmark case is that t_i is observed to increase with P in the case of spillovers. In addition, the inverted-U relationship between T and ϕ is converted into a U-shaped relationship. Specifically, this indicates that for levels of ϕ higher than 0.3, the response of $a^*_{i,spillover}$ to ϕ dominates that of $t^*_{spillover}$ in case of spillovers. We also observe that, income distribution improves at a lower level of ϕ in case of spillovers ($\tilde{Y_1}/\tilde{Y_2} = 1.5$ when $\phi = 0.3$) than the benchmark case ($\tilde{Y_1}/\tilde{Y_2} = 1.85$ when $\phi = 0.6$).

Furthermore, a U-shaped relationship between ϕ and both U^{CG} and U^{LG} is observed, where both utilities peak at $\phi \cong 0.3$. This finding implies that CG, as the agent who chooses institutions, would have an incentive to set $\phi \cong 0.3$, which happens to be exactly the OECD average of the expenditure decentralization.[28]

Remark 4. In a model with spillovers, both the welfare and the utility of a politically-oriented CG reach their maximums and income distribution reaches its most equitable point at $\phi = 0.3$.

Taking stock, the current framework reveals that the welfare effects of FD vary greatly with political orientation and spillover effects. We particularly note the following observations. First, both spillovers and FD lower tax collection efficiency and increase the tax rate. Second, while welfare and income distribution do not portray a clear relationship with FD, the central government utility is observed to improve up to an intermediate value of FD that is consistent with the observed level of FD in developed countries. Third,

[27] The size of the dataset obtained from the simulations is 609,376, which result from the addition of two spillover effects: $s_i \in [-1, 1]$, with the increments of 0.1, to the parameters reported in Table 2; in order to economize on the run-time, we increase the increments of α and β to 0.2.

[28] Authors' calculations, based on the World Bank, "Fiscal Decentralization Indicators," `http://www1.worldbank.org/publicsector/decentralization/fiscalindicators.htm`.

while the effect of political polarization on welfare, income distribution and efficiency are ambiguous, it is clear that political unison (P) leads to unambiguous improvements in the level of redistributable resources and central government utility.

4. Conclusions

This study presents a formal model in order to examine the validity of decentralization theorem vis-à-vis its fundamental conditions regarding heterogeneity and spillover effects. Our partial equilibrium model involves a central government that determines a flat tax rate and local governments that determine their tax collection effort. Local governments are assumed to be heterogeneous in both their income levels and political proximity to the central government. We solve the model both with and without spillover effects, where the latter necessitates a different game structure than the former in case of asymmetric information. We then investigate the connections of welfare and income distribution with the degrees of fiscal decentralization (FD), political unison and polarization.

In case of no spillovers, the solution of the strategic non-cooperative game between the local and central governments reveals that optimal tax collection effort and optimal tax rate are negatively related. Both of these variables also depict a negative association with FD. The model reveals that tax revenue has a non-monotonic (inverted-U type) relationship with FD, indicating that FD's effect on the optimal tax rate dominates that on the tax collection effort up to an intermediate level of FD, and the reverse effect dominates thereafter. This observation, termed as the decentralization-Laffer curve, is also supported by the model's simulations (and by the empirical evidence) that depict a non-linear relationship of both tax revenues and central government utility with FD. Simulations also reveal that while polarization has no direct effect on the model outcomes, both central government utility and tax revenue increase in political unison.

The extension of the model reveals that the optimal tax rate increases in spillovers but the optimal tax collection efficiency decreases in spillovers. In addition, utilities of both the central and local governments are observed to peak at a lower level of FD in case of spillovers than in the benchmark case, supporting the decentralization theorem. Income distribution is also observed to become the most equitable at a lower level of FD in case of spillovers than in the benchmark case.

In summary, we demonstrate in this study that when the central and local governments act strategically, increasing FD does not monotonically lead to efficiency and welfare gains. Moreover, spillover effects reduce the welfare-maximizing level of FD. Although welfare, measured by local government utility, is not affected by the political variables investigated here, it responds to FD the same way as the central government utility. Given that political unison is revealed to improve redistributable resources and central government utility, but not income distribution and welfare, this study points at the need for rule-based redistributive mechanisms to accompany FD.[29]

[29] See, for example, Akin, Bulut-Cevik, and Neyapti (2016), and Bulut-Cevik and Neyapti (2014) for the role of transfer rules and equalization target on the efficiency effects of FD. Shah (2006), Budina et al. (2012) and Neyapti (2013) are examples of the studies that emphasize the role of fiscal rules.

References

Akin, Zafer, Zeynep B Bulut-Cevik, and Bilin Neyapti. 2016. “Does Fiscal Decentralization Promote Fiscal Discipline?” *Emerging Markets Finance and Trade* 52 (3):690–705.

Altunbaş, Yener and John Thornton. 2012. “Fiscal Decentralization and Governance.” *Public Finance Review* 40 (1):66–85.

Arulampalam, Wiji, Sugato Dasgupta, Amrita Dhillon, and Bhaskar Dutta. 2009. “Electoral Goals and Center-State Transfers: A Theoretical Model and Empirical Evidence from India.” *Journal of Development Economics* 88 (1):103 – 119.

Arzaghi, Mohammad and J Vernon Henderson. 2005. “Why Countries Are Fiscally Decentralizing.” *Journal of Public Economics* 89 (7):1157–1189.

Aslim, Erkmen Giray and Bilin Neyapti. 2017. “Optimal Fiscal Decentralization: Redistribution and Welfare Implications.” *Economic Modelling* 61:224–234.

Besley, Timothy and Stephen Coate. 2003. “Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach.” *Journal of Public Economics* 87 (12):2611–2637.

Bracco, Emanuele, Ben Lockwood, Francesco Porcelli, and Michela Redoano. 2015. “Intergovernmental Grants as Signals and the Alignment Effect: Theory and Evidence.” *Journal of Public Economics* 123:78–91.

Brollo, Fernanda and Tommaso Nannicini. 2012. “Tying Your Enemy’s Hands in Close Races: The Politics of Federal Transfers in Brazil.” *American Political Science Review* 106 (4):742–761.

Buchanan, James M. 1950. “Federalism and Fiscal Equity.” *American Economic Review* 40 (4):583–599.

Budina, Nina, Andrea Schaechter, Anke Weber, Tidiane Kinda et al. 2012. “Fiscal Rules in Response to the Crisis: Toward the “Next-Generation” Rules: A New Dataset.” IMF Working Paper.

Bulut-Cevik, Zeynep Burcu and Bilin Neyapti. 2014. “Fiscal Efficiency, Redistribution and Welfare.” *Economic Modelling* 41:375–382.

De Mello, Luiz R. 2000. “Fiscal Decentralization and Intergovernmental Fiscal Relations: A Cross-Country Analysis.” *World Development* 28 (2):365–380.

De Mello, Luiz R and Matias Barenstein. 2001. “Fiscal Decentralization and Governance: A Cross-Country Analysis.” IMF Working Paper.

Dellmuth, Lisa Maria and Michael F Stoffel. 2012. “Distributive Politics and Intergovernmental Transfers: The Local Allocation of European Union Structural Funds.” *European Union Politics* 13 (3):413–433.

Eaton, Kent. 2001. “Political Obstacles to Decentralization: Evidence from Argentina and the Philippines.” *Development and Change* 32 (1):101–127.

Enikolopov, Ruben and Ekaterina Zhuravskaya. 2007. “Decentralization and Political Institutions.” *Journal of Public Economics* 91 (11):2261–2290.

Freinkman, Lev and Alexander Plekhanov. 2009. “Fiscal Decentralization in Rentier Regions: Evidence from Russia.” *World Development* 37 (2):503–512.

Gonzalez, Paula, Jean Hindriks, Ben Lockwood, Nicolas Porteiro et al. 2006. “Political Budget Cycles and Fiscal Decentralization.” UCL.

Inman, Robert P and Daniel L Rubinfeld. 1996. “Designing Tax Policy in Federalist Economies: An Overview.” *Journal of Public Economics* 60 (3):307–334.

———. 1997. “Rethinking Federalism.” *Journal of Economic Perspectives* 11 (4):43–64.

Janeba, Eckhard and John Douglas Wilson. 2011. “Optimal Fiscal Federalism in the Presence of Tax Competition.” *Journal of Public Economics* 95 (11):1302–1311.

Khemani, Stuti. 2007. “The Political Economy of Equalization Transfers.” In *Fiscal Equalization*, edited by Jorge Martine-Vazquez and Robert Searle. Springer, 463–484.

Knack, Stephen and Philip Keefer. 1997. “Does Social Capital Have an Economic Payoff? A Cross-Country Investigation.” *Quarterly Journal of Economics* 112 (4):1251–1288.

Knack, Stephen and Paul J Zak. 2003. "Building Trust: Public Policy, Interpersonal Trust, and Economic Development." *Supreme Court Economic Review* 10:91–107.

Koethenbuerger, Marko. 2008. "Revisiting the Decentralization Theorem –On the Role of Externalities." *Journal of Urban Economics* 64 (1):116–122.

Kyriacou, Andreas P and Oriol Roca-Sagalés. 2011. "Fiscal Decentralization and Government Quality in the OECD." *Economics Letters* 111 (3):191–193.

Ligthart, Jenny E and Peter van Oudheusden. 2015. "In Government We Trust: The Role of Fiscal Decentralization." *European Journal of Political Economy* 37:116–128.

Lockwood, Ben. 2008. "Voting, Lobbying, and the Decentralization Theorem." *Economics & Politics* 20 (3):416–431.

Ma, Jun. 1997. "Intergovernmental Fiscal Transfers in Nine Countries: Lessons for Developing Countries." *World Bank Policy Research Working Paper* (1822).

Montalvo, Jose G and Marta Reynal-Querol. 2005. "Ethnic Diversity and Economic Development." *Journal of Development Economics* 76 (2):293–323.

Neyapti, Bilin. 2013. "Fiscal Decentralization, Fiscal Rules and Fiscal Discipline." *Economics Letters* 121 (3):528–532.

———. 2017. "Educate or Adjudicate? Socioeconomic Heterogeneity and Welfare." *Defence and Peace Economics* 28 (5):491–510.

North, Douglass C and Barry R Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49 (04):803–832.

Oates, Wallace E. 1972. *Fiscal Federalism.* Edward Elgar.

O'Neill, Kathleen. 2003. "Decentralization as an Electoral Strategy." *Comparative Political Studies* 36 (9):1068–1091.

Ponce-Rodríguez, Raúl A, Charles Hankla, Jorge Martinez-Vazquez, Eunice Heredia-Ortiz et al. 2016. "Political Institutions and Federalism: A "Strong" Decentralization Theorem." Working Paper 16-03, International Center for Public Policy.

Sato, Motohiro. 2007. "The Political Economy of Interregional Grants." In *Intergovernmental Fiscal Transfers*, edited by Robin Broadway and Anwar Shah. The World Bank, Washington DC, 173–197.

Shah, Anwar. 2006. *A Practitioner's Guide to Intergovernmental Fiscal Transfers*, vol. 4039. World Bank Publications.

Solé-Ollé, Albert and Pilar Sorribas-Navarro. 2008. "The Effects of Partisan Alignment on the Allocation of Intergovernmental Transfers. Differences-in-differences Estimates for Spain." *Journal of Public Economics* 92 (12):2302–2319.

Stepan, Alfred C. 1999. "Federalism and Democracy: Beyond the US Model." *Journal of Democracy* 10 (4):19–34.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (5):416–424.

Treisman, Daniel. 2006. "Fiscal Decentralization, Governance, and Economic Performance: A Reconsideration." *Economics & Politics* 18 (2):219–235.

Appendix

Appendix A: The Benchmark Case

Proof of the LG's Reaction Function. The first order condition of Equation (10) with respect to a_i is

$$\frac{\beta}{a_i} - \frac{\alpha\phi t Y_i}{Y_i - \phi a_i t Y_i - (1-\phi)tY_i} = 0.$$

Rearranging terms yields

$$\alpha\phi a_i t Y_i = \beta(1 - \phi a_i t - (1-\phi)t)Y_i.$$

Simplifying the equation above and solving for a_i yields

$$a_i = \left(\frac{\beta}{\alpha+\beta}\right)\frac{1-t+\phi t}{\phi t}. \qquad \square$$

Proof of the CG's Reaction Function. The following first order condition of Equation (13) with respect to t is obtained for each i, assuming $i \in \{1,2\}$ without loss of generality[30]:

$$\alpha\left(\frac{\phi(1-a_1)-1}{(1-t\phi a_1)-t(1-\phi)} + \frac{\phi(1-a_2)-1}{(1-t\phi a_2)-t(1-\phi)}\right) + \beta\left(\frac{p_1+p_2+2}{t}\right) = 0.$$

Let $P = p_1 + p_2$. Considering the symmetric first order conditions for LGs such that $a_i = a_1 = a_2$ for $i = 1, 2$, we can simplify CG's first order condition as

$$\frac{2\alpha(\phi(1-a_i)-1)t}{1+(\phi(1-a_i)-1)t} = -\beta(P+2)$$

Now let $A = \phi(1-a_i) - 1$. The first order condition above becomes

$$t(2\alpha A + \beta P A + 2\beta A) = -\beta(P+2) \Longrightarrow t = \frac{-\beta(P+2)}{A(2\alpha + \beta(P+2))}.$$

[30]The findings remain robust to the extension of the analysis to the case of "n" localities except that the number "2" in the joint solution would be replaced by "n" and P would be $\{p_1 + p_2 + \cdots + p_n\}$.

Finally, we substitute in the equation for A

$$t = \frac{\beta(P+2)}{(1-\phi(1-a_i))(2\alpha+\beta(P+2))} \text{ for } i = 1, 2. \qquad \square$$

Proof of Lemma 1. In order to obtain the Nash equilibrium, we need to solve LGs' and CG's reaction functions simultaneously.
Substituting t into LGs' reaction function yields

$$a_i = \left(\frac{\beta}{\alpha+\beta}\right)\left(\frac{(1-\phi(1-a_i))(2\alpha+\beta(P+2))}{\phi\beta(P+2)} - \frac{1}{\phi} + 1\right).$$

To simplify the equation above let $A = \phi(1-a_i) - 1$. Then, we have

$$a_i = \frac{2A\alpha + A\beta(P+2) - \beta(P+2) + \phi\beta(P+2)}{\phi(\alpha+\beta)(P+2)}.$$

Cross-multiplying the terms and dividing both sides by $(P+2)$ yield

$$a_i\phi(\alpha+\beta) = \frac{2A\alpha}{P+2} + A\beta + \beta(\phi-1).$$

Replacing the equation for A and rearranging the terms yield

$$a_i\phi(\alpha+\beta) - \left(\frac{2\alpha}{P+2} + \beta\right)(1-\phi+\phi a_i) = \beta(\phi-1).$$

Further simplification yields

$$a_i\left(\frac{\alpha(P+2)-2\alpha}{P+2}\right) = \left(\frac{\phi-1}{\phi}\right)\left(\frac{2\alpha}{P+2}\right).$$

Then multiplying both sides by $(P+2)$ and isolating a_i yields the optimal a_i (denoted as a_i^*)

$$a_i^* = \frac{2(1-\phi)}{\phi P}.$$

Now we substitute a_i into CG's reaction function

$$t = \frac{\beta(P+2)}{\left(1-\phi\left(1-\left(\frac{\beta}{\alpha+\beta}\right)\left(\frac{1}{\phi t} - \frac{1}{\phi} + 1\right)\right)\right)(2\alpha+\beta(P+2))}.$$

To simplify the equation above let $R = \frac{\beta}{\alpha+\beta}$ and $C = 2\alpha + \beta(P+2)$. Then,

$$t = \frac{\beta(P+2)}{\left(1 + \frac{R}{t} - R + (R-1)\phi\right) C}.$$

Cross-multiplying and rearranging the terms yield

$$t(R-1)(\phi-1) + R = \frac{\beta(P+2)}{C}.$$

By replacing the equation for R and further simplifying the equation above, we obtain

$$t(\phi-1) = \frac{\beta}{\alpha}\left(1 - \frac{(\alpha+\beta)(P+2)}{C}\right).$$

Isolating t and replacing the equation for C, we conclude that

$$t^* = \frac{\beta P}{(1-\phi)(2\alpha + \beta(P+2))}. \qquad \square$$

Proof of Proposition 1. For $\phi \in (0,1)$, taking the partial derivatives of t^* and a^* with respect to ϕ yield

$$\frac{\partial t^*}{\partial \phi} = \frac{\beta P}{(\phi-1)^2(2\alpha + \beta(P+2))} > 0 \text{ and } \frac{\partial a_i^*}{\partial \phi} = -\frac{2}{\phi^2 P} < 0. \qquad \square$$

Proof of Corollary 1. For $\phi \in (0,1)$, the second partial derivative of a_i^* with respect to ϕ is

$$\frac{\partial^2 a_i^*}{\partial \phi^2} = \frac{4}{\phi^3 P} > 0.$$

Additionally, the cross partial derivative of $\partial a_i^*/\partial \phi$ with respect to P is

$$\frac{\partial^2 a_i^*}{\partial \phi \partial P} = \frac{2}{\phi^2 P^2} > 0. \qquad \square$$

Proof of Proposition 2. For $\phi \in (0,1)$, taking the partial derivative of t^* with respect to P yield

$$\frac{\partial t^*}{\partial P} = \frac{-2\beta(\alpha+\beta)}{(\phi-1)(2\alpha + \beta(P+2))^2} > 0. \qquad \square$$

Proof of Corollary 2. For $\phi \in (0, 1)$, the second partial derivative of a_i^* with respect to P is

$$\frac{\partial^2 a_i^*}{\partial P^2} = \frac{4(1-\phi)}{\phi P^3} > 0. \qquad \square$$

Appendix B: Simulations

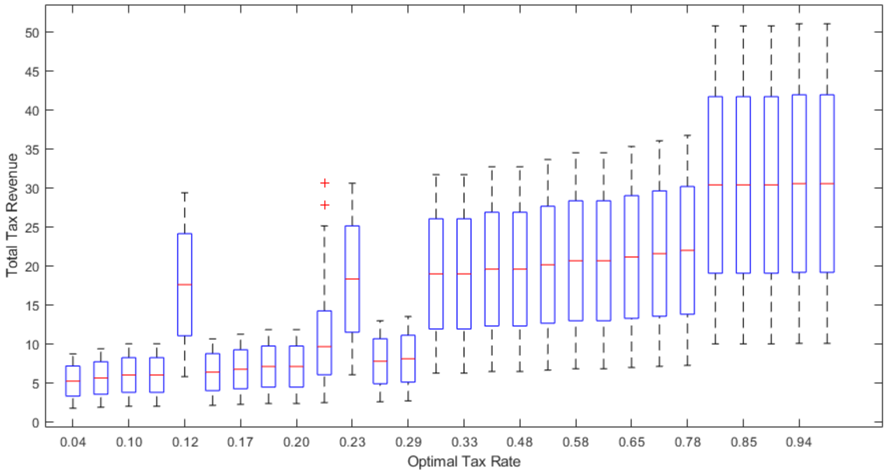

Figure B1: Optimal Tax Rate and Total Tax Revenue

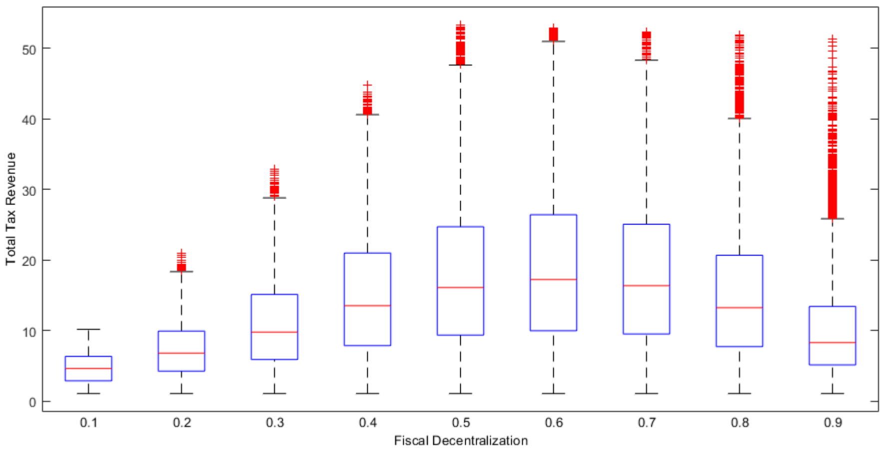

Figure B2: Fiscal Decentralization and Total Tax Revenue

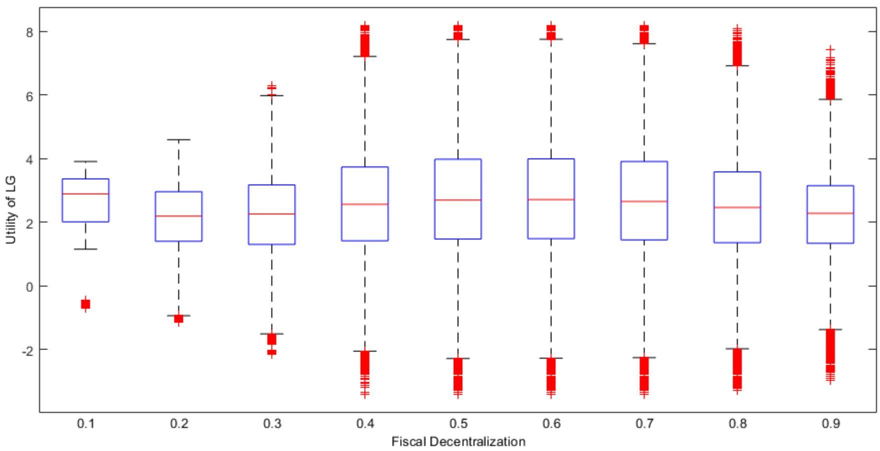

Figure B3: Fiscal Decentralization and Utility of LG

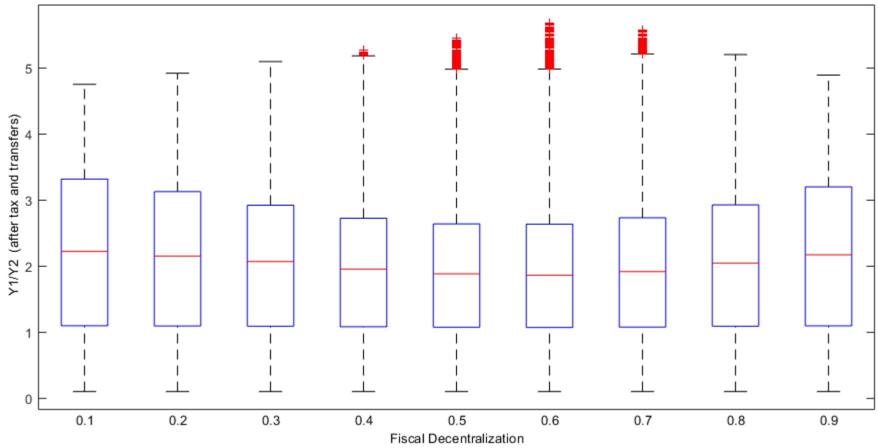

Figure B4: Fiscal Decentralization and Income Distribution

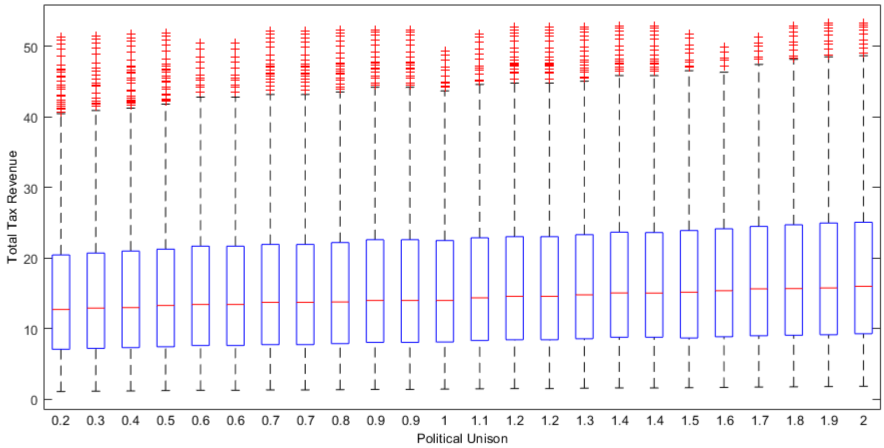

Figure B5: Political Unison and Total Tax Revenue

Appendix C: Empirical Evidence

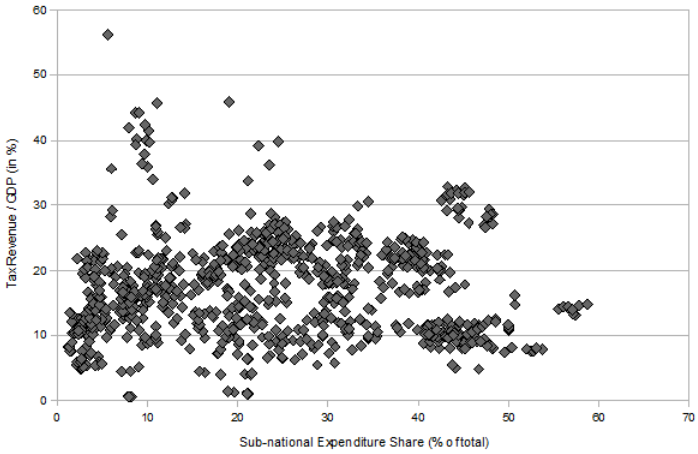

Figure C1: Fiscal Decentralization and Tax Revenue

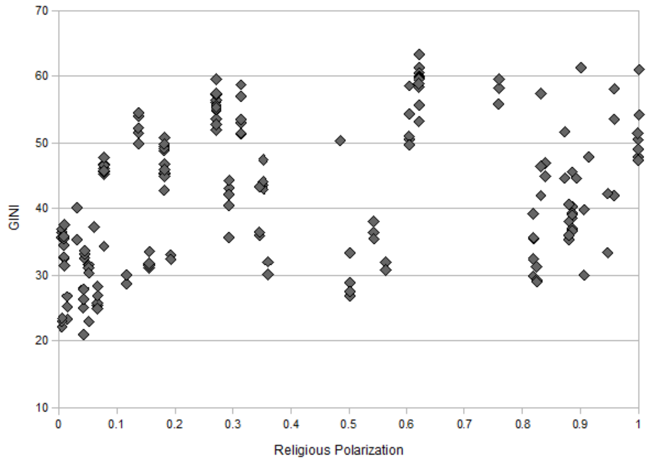

Figure C2: Polarization and Income Distribution

Appendix D: The Spillovers Case

Proof of the CG's Reaction Function. CG maximizes the following utility function:

$$U^{CG}_{i,spillover} = \sum_{i,j,i\neq j} \left(\alpha ln((1-t_i)Y_i) + p_i\beta(ln(\phi a_i t Y_i) + s_i ln(\phi a_j t Y_j)) + \beta ln((1-\phi)t\hat{p}_i \sum_i Y_i) \right).$$

Assuming $i \in \{1,2\}$ without loss of generality, we obtain the following first order condition with respect to t:

$$\alpha \left(\frac{\phi(1-a_1)-1}{(1-t\phi a_1)-t(1-\phi)} + \frac{\phi(1-a_2)-1}{(1-t\phi a_2)-t(1-\phi)} \right) + \frac{\beta}{t}(p_1(1+s_1)+p_2(1+s_2)+2) = 0.$$

Considering symmetry in LGs' tax collection efforts ($a_i = a_1 = a_2$), the reaction function of CG becomes

$$t = \frac{\beta\sum_i(p_i(1+s_i)+1)}{(1+\phi(a_i-1))\left(2\alpha+\beta\sum_i(p_i(1+s_i)+1)\right)} \text{ for } i \in \{1,2\}. \qquad \square$$

Proof of Lemma 2. In order to obtain the Nash equilibrium in a leader-follower framework, we solve LGs' utility maximization problem by taking into account the CG's reaction function denoted above. Hence, the first order condition from LGs' problem is as follows:

$$\underbrace{\frac{-\alpha\left(\phi t + (a_i\phi + (1-\phi))\frac{\partial t}{\partial a_i}\right)}{1 - a_i\phi t - (1-\phi)t}}_{LHS} + \underbrace{\frac{2\beta}{t}\frac{\partial t}{\partial a_i}(1+s_i) + \frac{\beta}{a_i}}_{RHS} = 0$$

Let $D = \beta\sum_i(p_i(1+s_i)+1)$. Then, the partial derivative of t with respect to a_i is

$$\frac{\partial t}{\partial a_i} = \frac{-\phi D}{(2\alpha+D)(a_i\phi+(1-\phi))^2}$$

Substituting $\frac{\partial t}{\partial a_i}$ and t into LHS and RHS yield

$$\underbrace{\frac{\alpha\left(\frac{\phi D}{(2\alpha+D)(a_i\phi+(1-\phi))} - \frac{\phi D}{(2\alpha+D)(a_i\phi+(1-\phi))}\right)}{1-\frac{D}{2\alpha+D}}}_{LHS=0} - \underbrace{\frac{2\beta\phi(1+s_i)}{a_i\phi+(1-\phi)} + \frac{\beta}{a_i}}_{RHS} = 0.$$

Rearranging the terms yield

$$a^*_{i,spillover} = \frac{1-\phi}{\phi(1+2s_i)} \text{ for } i \in \{1,2\}.$$

We now substitute $a^*_{i,spillover}$ into CG's reaction function

$$t = \frac{\beta \sum_i (p_i(1+s_i)+1)}{(1-\phi)\left(1+\frac{1}{1+2s_i}\right)(2\alpha + \beta \sum_i (p_i(1+s_i)+1))}$$

Simplifying the equation above yields

$$t^*_{spillover} = \frac{\beta(1+2s_i)\sum_i (p_i(1+s_i)+1)}{2(1-\phi)(1+s_i)(2\alpha+\beta\sum_i(p_i(1+s_i)+1))} \text{ for } i \in \{1,2\}. \qquad \square$$

Proof of Proposition 3. For $\phi \in (0,1)$, taking the partial derivative of $t^*_{spillover}$ with respect to p_i yield

$$\frac{\partial t^*_{spillover}}{\partial p_i} = \frac{-\alpha\beta(2s_i+1)}{(\phi-1)(2\alpha+\beta\sum_i(p_i(1+s_i)+1))^2} > 0. \qquad \square$$

Proof of Proposition 4. For $\phi \in (0,1)$, taking the partial derivatives of $t^*_{spillover}$ and $a^*_{i,spillover}$ with respect to s_i yield

$$\begin{aligned}
\frac{\partial a^*_{i,spillover}}{\partial s_i} &= \frac{2(\phi-1)}{\phi(1+2s_i)^2} < 0 \text{ and} \\
\frac{\partial t^*_{spillover}}{\partial s_i} &= \underbrace{\frac{(2s_i+1)\beta\sum_i(p_i(1+s_i)+1)}{2(\phi-1)(s_i+1)^2(2\alpha+\beta\sum_i(p_i(1+s_i)+1))}}_{F} - \\
&\quad \underbrace{\frac{\beta p_j(2s_i+1)+2\beta\sum_i(p_i(1+s_i)+1)}{2(\phi-1)(s_i+1)(2\alpha+\beta\sum_i(p_i(1+s_i)+1))}}_{G} + \\
&\quad \underbrace{\frac{p_j(2s_i+1)\beta^2\sum_i(p_i(1+s_i)+1)}{2(\phi-1)(s_i+1)(2\alpha+\beta\sum_i(p_i(1+s_i)+1))^2}}_{H} > 0, \text{ where } i \neq j.
\end{aligned}$$

We want to show that F + H < G. Again let $D = \beta\sum_i(p_i(1+s_i)+1)$, and also let $E = \beta p_j(2s_i+1)(s_i+1)$. Thus,

$$\text{F} + \text{H} < \text{G} \Leftrightarrow DE + 2(s_i+1/2)(2\alpha+D)D < (2\alpha+D)E + 2(s_i+1)(2\alpha+D)D.$$

Since $\alpha \in [0, 1]$, F + H < G holds true. $\square$

Table D1: Comparing the Benchmark Model with the Case of Spillovers

	a_i^*		t^*		t_i		T		U^{CG}		U^{LG}		$\tilde{Y}_1/\tilde{Y}_2$	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
P	-	0	+	+	?	+	+	+	+	+	?	?	?	?
ϕ	-	-	+	+	?	max: 0.3-0.4	max: 0.6	max: 0.3	max: 0.5	max: 0.3	max:0.6	max: 0.3	min:0.6	min: 0.3

Notes: (1) denotes the benchmark case. (2) denotes the spillover case.