

On the optimal allocation of responsibilities among national and subnational governments

Ron Shani[^], Yaniv Reingewertz^{^*}

Abstract

This paper introduces a model for optimizing local public goods provision across national, regional, and local government tiers. We analyze how spillovers, preference heterogeneity, and economies of scale affect centralization versus decentralization decisions. Three key insights emerge: (a) Adding regional governments to the model creates subnational solutions for inter-local spillovers, (b) Preference heterogeneity incentivizes decentralization, (c) Economies of scale favor centralization for efficiency gains. The optimal responsibility distribution depends on these three competing forces' interplay. Our model integrates previous studies' insights into a unified theory encompassing the primary forces influencing centralization-decentralization decisions across government tiers.

Keywords: Fiscal Federalism; subnational governments, decentralization, local public goods
JEL Codes: H77, H11, H41, R13

[^]Division of Public Administration and Policy, University of Haifa, Haifa, Israel. ^{*}corresponding author. yanivrein@poli.haifa.ac.il , +972-48240599. We would like to thank participants of the Regional Inequalities workshop in Bochum, the Israeli Economic Association Annual meeting and seminar participants at Ariel University, as well as Thushyanthan Baskaran, Oliver Lorz, and Robert Dur for their valuable comments and suggestions.

1 Introduction

Sub national governments in general, and regional governments in particular, are an important yet neglected part of the fiscal federalism literature. An OECD survey finds that out of 135 countries, representing 93% of the world's population, 72% have at least two sub-national government (SNG) levels, resulting in three levels of government (OECD, 2022).¹ Alas, the theoretical literature on the optimal allocation of powers between those sub-national governments is very limited, and most of the literature is only providing models with two levels of government (see, e.g. Geys & Konrad, 2010; Hämäläinen & Moisio, 2015; Oates, 1972).

This line of literature predominantly follows Oates' Decentralization Theorem, which argues that "it will always be more efficient for local governments to provide the Pareto-efficient level of output ... than for the central government" (Oates, 1972, p. 35). suggesting that most powers should be decentralized to the municipal level. However, intermediate SNG levels may balance local responsiveness with economies of scale. A somewhat more contemporary literature also stresses that centralization is preferred if spillovers between local governments are pervasive (Besley & Coate, 2003; Lockwood, 2002).

While the classic fiscal federalism literature assumes that benevolent officials make policy decisions, 'second generation' theories consider officials pursuing political interests (Geys & Konrad, 2010; Martinez-Vazquez et al., 2017; Oates, 2005, 2008; Reingewertz, 2014; Slavinskaitė et al., 2019; Weingast, 2009). Newer studies examine various political decision-making procedures and their impact on budgets and project allocation between government levels, particularly in the context of spillovers (Besley & Coate, 2003; Cheikbossian, 2016; Dur & Roelfsema, 2005; Feidler & Staal, 2012; Frère & Védrine, 2024; Gregor & Stastna, 2012; Koethenbuerger, 2008; Lockwood, 2002; Lorz & Willmann, 2005). Despite this rich literature on national-subnational power dynamics, research examining the allocation of responsibilities among different tiers of subnational governments remains scarce.

We aim at providing a theory that illustrates how to allocate the powers between several SNGs levels. The first contribution of our model to these strands of literature is to add the regional level to a fairly standard model of fiscal federalism. We demonstrate that the most efficient provision of an impure public good may reside at the regional level.² Moreover, the

¹ The OECD sample represents 94% of the world GDP. See also Council of European Municipalities and Regions (CEMR) for a description of subnational government structures in Europe (www.ccre.org), and Chatry and Vincent (2019) for subnational governments in Asia.

² In this study we refer to impure public goods and services as those that impose marginal costs, which are absorbed by the local government.

optimal tier depends on careful consideration of preference heterogeneity, spillover effects, and economies of scale. By integrating regional governments into the model, the role of central governments in providing local public goods is significantly diminished, thereby extending the traditional decentralization model.

Another contribution of our model is in allowing the preferences of individuals to vary not only between local governments but also within them, where previous studies assume homogeneity within local governments. This assumption enriches existing models, relaxes the assumption of homogeneity within local governments and allows for a more realistic representation of actual preferences. When spillovers are relatively small and local preference heterogeneity is fairly large, we demonstrate that the aggregate surplus of providing local public goods is greater under local provision. Additionally, we model a political-economy decision making process at the regional level for public products where elected representatives negotiate the number of product units the region provides. We distinguish between pro-public spending representatives who decide to provide more product units than the optimal regional average, and averse public spending representatives who decide to provide less. We show that if they decide to provide more product units they need to subsidize the cost. Yet when they decide to provide less product units, the perceived value of such public product should be low.

We introduce a comprehensive "forces model" that systematically integrates the multidimensional factors influencing centralization and decentralization decisions. Unlike existing literature that typically examines isolated aspects of government service provision, our framework allows for economic, political, structural, and preference-based determinants within a unified analytical approach. Our model offers a sophisticated diagnostic tool that bridges theoretical insights with practical policy design, enabling policymakers to evaluate the optimal governmental tier for providing specific public goods beyond simplistic binary choices between centralization and decentralization.

The rest of the paper is organized as follows. Section 2 reviews the theoretical literature on fiscal federalism. Section 3 describes the model and extensions. Section 4 discusses the findings and their policy implications and concludes.

2 Literature review

The classic Fiscal Federalism theory suggests that the lowest tier of government should provide local public goods because of information asymmetries between the local and central levels

(Oates, 1972). Roughly speaking, the central government should stabilize the economy and control the distribution-equalization of resources, whereas subnational governments (SNG) should provide local or regional public goods adjusted to the preferences of their residents (Musgrave, 1959; Oates, 1972; Olson, 1969). Oates' classic Decentralization Theorem states that under equal conditions when maximizing the aggregate surplus of residents, local governments will provide local public goods more efficiently than the central government (Oates, 1972, p. 35). When the public good exhibits economies of scale, or when it spills over jurisdiction borders, the optimal allocation favors central provision. The decentralization theorem assumes that central governments provide equal goods to all residents (equal level of services, same number of product units), regardless of their local preferences.³ Oates asserts that when the central government possesses the knowledge of individual tastes, it can provide Pareto-efficient levels of consumption in all jurisdictions and adds that in terms of the economic definition of federalism, it is equivalent to decentralization (Oates, 1972, p. 37).

A rich literature evolved from Oates' work, extending his model in various ways. First, many studies analyze the case of spillovers between different jurisdictions (Cheikbossian, 2016; Dur & Staal, 2008; Feidler & Staal, 2012; Gregor & Stastna, 2012; Koethenbuerger, 2008; Lockwood, 2002; Lorz & Willmann, 2005; Seabright, 1996). Second, many studies add the political arena to the model in order to describe a more realistic decision making (Besley & Coate, 2003; Dur & Roelfsema, 2005; Feidler & Staal, 2012; Lockwood, 2002; Lorz & Willmann, 2005; Ponzetto et al., 2016). Third, some studies state that economies of scale lead to centralization (Gregor & Stastna, 2012; Ponzetto et al., 2016); Forth, some add heterogeneity of preferences in each local government (Gregor & Stastna, 2012; Lockwood, 2002; Ponzetto et al., 2016). Finally, only few studies enrich the model by discussing more than one sub-national government level (Geys & Konrad, 2010; Halásková & Halásková, 2015; Seabright, 1996).⁴

While the classic decentralization theorem posits that decentralization of power to the lower level is preferred, it is silent about the regional level. This stands in sharp contrast to actual division of responsibilities across SNG levels (OECD, 2022). **Error! Reference source not found.** is based on this OECD classification of allocation of responsibilities between three SNG levels. This allocation considers mainly functional responsibilities which are geographic

³ Ponzetto et al. (2016) further state that they “find that centralization can be welfare-maximizing only if it is accompanied by a uniformity constraint that requires the central government to provide identical public goods to all regions.”

⁴ Empirical studies confirm that decentralization, which may increase efficiency, positively affects the provision of local public goods (Barankay & Lockwood, 2007; Martinez-Vazquez et al., 2017)

in nature. As can be seen in the table, in practice the regional level is in charge of various public goods and services.

Table 1: Breakdown of responsibilities across SNG levels

National level	Regional level	Local level
Country	County	Municipality
National interests	Supra-local interest	Local interests
Foreign affairs policy	Assistance to small municipalities	Education (all levels)
National defense and armed forces	Regional infrastructure: secondary roads, public transportation	Urban/local planning, zoning regulation
Stabilize the economy	Higher education, universities	Local utilities
Control equalization and distribution of resources	Spatial planning	Local roads and public transportation
National infrastructure	Secondary hospitals	Social affairs
Highways	Supra-municipality social welfare	Primary and preventive healthcare
Master planning	Regional public order and safety	Recreation and culture
National utilities	Local governments supervision	Public order, security and safety
National public transportation	Waste treatment	Local economic development
National health	Environment	Environment, sustainability
		Administrative and permit services
		Waste collection

Notes: (1) Federal countries exhibit also a state level. The table splits states' responsibilities between the central and regional levels. (2) We adapted the table from OECD (2022, p. 43 Table 2.1) while adding the National level with the observations of Musgrave (1959).

Possibly the most important contribution within this literature is Besley and Coate (2003), which add the political point-of-view. They assume that each district requests different local public goods (e.g., parks in the first and roads in the second) and that there are spillovers between districts. A key contribution of Besley and Coate (2003) is the analysis of a political equilibrium at the central level, where minimum winning coalitions dictate the geographical distribution of local public goods. They conclude that, to some extent, their analysis corroborates Oates' insights that decentralization is the more efficient arrangement when spillovers are minor, or regions are heterogeneous, while centralization proves more advantageous when there is a high level of spillovers.

In a complementary paper, Lockwood (2002) studies allocating projects between the central and the local level in a political economy setting, assuming homogenous regional residents. The study concludes that small spillovers and significant heterogeneity between regions still mean that decentralization is preferred, but when spillovers are significant, centralization is more efficient. Dur and Roelfsema (2005) show that the centralization of decision-making results in overspending in some regions and too low in others. They also note that when the central government finances local public goods, the common pool problem may arise.

Focusing on local public goods with inter-region spillovers, Lorz and Willmann (2005) show that centralization is sub-optimally low depending on the elected representatives' approach, as representatives prefer decentralization to reduce their region's cost share through sharing the cost with other regions. With small spillovers, decision-making should be decentralized, and when spillovers are significant, a decision at the center is better.

Several other studies discuss the role of spillovers. Koethenbuerger (2008) concludes that a higher degree of spillovers may reduce the difference in the utility of centralization and decentralization. Gregor and Stastna (2012) study when a spill-in is a complement good and conclude that centralization dominates in most settings, yet preference homogeneity is not a sufficient condition for centralization dominance. Cheikbossian (2016) differentiates between local public goods which are substitutes or complements. He concludes that with a common financing rule, centralization improves welfare if spillovers are significant and the local public goods are complementary. Most studies agree that centralization allows for the internalization of spillovers, but this varies by the level of spillovers.

Several studies add issues of size to the analysis. Feidler and Staal (2012) add that the size of regions influences the decision-making process. Residents of larger local government (LG) favor decentralization, as they can finance the local public good without intergovernmental grants, at least when spillovers are small enough. Smaller districts favor centralization, as other districts' residents participate in the cost of the local public good. They state that decentralization leads to a higher surplus when spillovers are small, and when spillovers are significant, centralization leads to a higher surplus. Frère and Védrine (2024) study voluntary intermunicipal cooperation as an alternative to centralization ("local centralization"). They emphasize the "Zoo effect", where small LGs cannot provide local public goods requiring significant resources (Oates, 1988). Shani et al. (2023) present somewhat similar ideas by discussing a mezzanine tier that agglomerates small rural villages to "locally centralize" and benefit from economies of scale. They find that such structures benefit from economies of scale, yet there are limits to those benefits.

Summing up the literature, we observe that several forces influence the allocation of powers between different tiers of subnational governments. The actual policy decision needs to balance those forces. Table 2 lists the influencing forces and their magnitude towards (de)centralization while referencing the studies that discuss each force. We note that second-generation Fiscal Federalism studies, which add the political economy and decision-making processes, describe the behavior of residents and their elected representatives and the

representatives' negotiations in the House. We allocated the political economy behavior into one force in Table 2.

Table 2: Forces influencing centralization - decentralization allocation of powers

Force	Pro-Decentralization	Pro-Centralization	Reference
Information asymmetry	High	Low	Oates (1972)
Spillovers	Small	Significant	Besley and Coate (2003) Dur and Staal (2008) Feidler and Staal (2012) (Koethenbuerger, 2008) Lockwood (2002) Lorz and Willmann (2005)
Residents' preferences variations between regions	Heterogenous	Homogenous	Gregor and Stastna (2012) Lockwood (2002) Oates (1972)
Economies of scale, size and scope	Insignificant	Significant	Oates (1972)
Zoo effect: good scale	Small	Large	Frère and Védrine (2024) Oates (1988)
Size of local government	Large	Small	Feidler and Staal (2012)
Complementarity of spillovers	Substitute	Complementary	Cheikbossian (2016) Gregor and Stastna (2012)
Representation	Averse public spending	Pro public spending	Besley and Coate (2003) Lorz and Willmann (2005)

Note: Gregor and Stastna, (2012) add that decentralization may dominate even in homogeneity of preferences.

3 The model

Our model introduces a novel approach to analyzing public good provision across three government tiers: central, regional, and local. Unlike existing two-tier frameworks in the fiscal federalism literature, we explicitly model the complexities of multi-level governance by examining how different levels of government can optimize public good delivery. The core innovation lies in our treatment of residents' preferences. We move beyond the traditional assumption of homogeneous preferences by modeling how individual residents value public goods differently. Each resident has a preferred quantity of public goods, and their utility is determined by the minimum of their preferred quantity and the quantity actually provided, net of the cost of provision.

Central to our analysis is the utility function that captures the way residents experience public good provision. When a government provides fewer units than a resident desires, the resident values only the units received. Conversely, when more units are provided than

preferred, the resident only values units up to their preference level. This approach reflects real-world scenarios in services like education and healthcare, where overconsumption does not necessarily increase utility.

The model's key contribution is demonstrating how preference heterogeneity, spillovers, and economies of scale interact across different government tiers. By introducing a regional level, we provide a more sophisticated mechanism for addressing local variations in public good provision. The analysis reveals how the optimal tier for providing a public good depends on a complex interplay of factors, including preference dispersion, cost structures, and inter-jurisdictional dynamics.

Assume a country is divided into regions that are subdivided into local governments. C, R, and L denote the central, regional, and local tiers. For simplicity, we assume that all local governments have the same number of residents and that all regions have the same number of local governments. These governments need to decide on the level of local public goods provision.

Residents of each local government have different preferences for local public goods, where the mean preference and the standard deviation characterize their preferences in each local government. We focus on a local public good, assuming that each government can provide this good, yet with different characteristics, such as the quality of the good or the number of product units each resident prefers to consume. For simplicity, we refer to the number of product units each resident prefers or each government provides as a proxy for the bundle of service quality-quantity.

When a sub-national government (SNG) provides Q^x product units to each resident in its jurisdiction, some residents receive more product units than their preferences $Q^x > Q_i$ and some residents receive less $Q^x < Q_i$. Residents assign value to the local public good according to their preferences. If they receive more product units than their preferences, they attribute value only to the number of units that match their preferences, yet they pay for the number of units the government provides (usually through taxation). If, however, they receive fewer units than their preferences, they assign value only to the number of units they receive. Therefore, the utility each resident assumes is $U_i^x = V^x \min(Q_i, Q^x) - Q^x C^x$, where V^x denotes the value that residents attribute to a public product unit, and C^x the cost of a product unit.

The aggregate utility of all residents in a government jurisdiction is

$$U^x = \sum_{i=1}^{N^x} U_i^x = \sum_{\substack{i=1 \\ x \in (C, R, L)}}^{N^x} [V^x \min(Q_i, Q^x) - Q^x C^x] = V^x \sum_{i=1}^{N^x} \min(Q_i, Q^x) - N^x Q^x C^x, \quad (1)$$

where Q^x denotes the per-resident number of product units the government provides, and N^x the number of residents in the government's jurisdiction, $x \in (C, R, L)$. There are n regions and m local governments in the country.

The level of government which should provide the local public good is the one which maximizes U^x , or $\max(U^C, U^R, U^L)$. In the following analysis we maximize the per-resident utility across central, regional and local provision of services by calculating the per-resident utility difference between central and regional and compare it to the difference between regional and local: $\Delta U^{C-R} = \frac{U^C}{N^C} - \frac{U^R}{N^C}$ and $\Delta U^{L-R} = \frac{U^L}{N^L} - \frac{U^R}{N^L}$. We first calculate the utility for each government level (U^C, U^R, U^L), and then calculate the differences as Equations (11), (12) and (15) present.

According to the Decentralization Theorem (Oates, 1972, p. 60), each government maximizes its residents' aggregate surplus (or minimizes the welfare loss). The optimal provision of the local public good is the arithmetic mean of the quantity across all residents, which is also Pareto-efficient. We follow Oates (1972), by looking at the average number of product units each government provides (Equation (2)). However, while Oates assumes welfare losses occur when fewer product units are provided than some residents prefer, we also account for welfare losses arising when more product units are provided than other residents prefer, forcing them to pay for unwanted services through additional taxation.⁵

$$\begin{aligned} Q^C &= \frac{1}{N^C} \sum_{i=1}^{N^C} Q_i = \frac{1}{n} \sum_{j=1}^n Q_j^R = \frac{1}{m} \sum_{k=1}^m Q_k^L \\ Q_j^R &= \frac{1}{N_j^R} \sum_{i=1}^{N_j^R} Q_i : \text{for residents of region } j \\ Q_k^L &= \frac{1}{N_k^L} \sum_{i=1}^{N_k^L} Q_i : \text{for residents of local government } k \end{aligned} \tag{2}$$

In any distribution, there are residents whose preferences are below the average and others who are above. Let $\gamma \in (0,1)$ denote the ratio of residents whose preferences are below the average preferences of the jurisdiction population. Therefore, there are γN^x residents that experience $Q_i < Q^x$, and $(1 - \gamma)N^x$ residents which face $Q_i > Q^x$.⁶ The left argument of Equation (1) can now be

⁵ See Appendix B

⁶ γ is a measure of the skewness of the distribution. In a symmetrical distribution $\gamma = \Phi(0) = 1/2$. In exponential distribution $\gamma = \Phi(1) = 0.63$.

$$\sum_{i=1}^{N^x} \min(Q_i, Q^x) = \sum_{\substack{i=1 \\ \text{s.t. } Q_i < Q^x}}^{\gamma N^x} Q_i + \sum_{\substack{i=1 \\ \text{s.t. } Q_i > Q^x}}^{(1-\gamma)N^x} Q^x = \sum_{\substack{i=1 \\ \text{s.t. } Q_i < Q^x}}^{\gamma N^x} Q_i + (1-\gamma)N^x Q^x \quad (3)$$

We can express the left sum argument of the right-hand side of Equation (3) as the average of the residents whose $Q_i < Q^x$ times the number of those residents. We can further express the average of the residents below the total average using the definition of the Mean Absolute Deviation around a central point (MAD) using Equation A4.⁷

$$\sum_{i=1}^{N^x} \min(Q_i, Q^x) = \gamma N^x \left(Q^x - \frac{1}{2\gamma} \text{MAD} \right) + (1-\gamma)N^x Q^x \quad (4)$$

Substituting Equation (4) into (1) yields

$$U^x = N^x Q^x (V^x - C^x) - \frac{1}{2} V^x N^x \text{MAD} \quad (5)$$

Proposition 1: the aggregate surplus of providing local public goods under utility maximizing conditions is proportional to the utility from the good minus the loss of utility resulting from the relative dispersion of residents' preferences.

Proof: using Jensen's inequality, one can show that the MAD is less than or equal to the standard deviation: $\text{MAD}: E(|x - \mu|) \leq \sigma$. Embedding it into Equation (5) yields $U^x \geq N^x Q^x (V^x - C^x) - \frac{1}{2} V^x N^x \sigma$. The greater the standard deviation, the smaller the aggregate residents' utility. Note that the aggregate surplus is independent of the skewness of the distribution.

The rationale behind Proposition 1 is that residents have different tastes and preferences for the number of products the government provides. Therefore, the aggregate surplus subtracts the loss of utility due to mismatch between residents' preferences and the actual service provision. The more the residents' preferences are close to homogeneity, the higher the aggregate utility.

The Decentralization Theorem (Oates, 1972) assumes homogeneous preferences within each jurisdiction. We extend this by showing that utility losses resulting from heterogeneous resident preferences within jurisdictions significantly impact the aggregate surplus from decentralization. This is somewhat similar to Proposition 3.3.2 of Frère and Védrine (2024), yet it takes the probability distribution function of residents preferences into account, while

⁷ MAD is defined as $\frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$. For the normal distribution MAD is $\frac{E|x|}{\sqrt{E(x^2)}} = \sigma \sqrt{\frac{2}{\pi}}$. See Appendix A.

also differentiating between when the local government provides less product units than the resident prefers, and when it provides more.

Corollary 1a: *The greater the homogeneity of residents within a jurisdiction, the higher the aggregate benefit they derive from local public goods.*

Proof: with Proposition 1, when the residents in a defined jurisdiction have similar preferences, the standard deviation becomes small, resulting in higher aggregate utility. \square

Corollary 1b: *when the cost of providing a product unit is the same across government tiers, the highest utility is when the lower tier provides the local public good.*⁸

Proof: assume that the perceived value and the cost of a public product unit are equal at all government levels: $V^C = V^R = V^L \equiv V$, and $C^C = C^R = C^L \equiv C$. Assuming even a small Tiebout sorting, the between LGs variance is greater than the within, implying that the standard deviation becomes smaller as the level of government tier is closer to the resident: $\sigma^C > \sigma^R > \sigma^L$, yielding from Equation (5) that $U^C < U^R < U^L$.⁹

3.1 Comparing local, regional and central provision of services

Equation (5) describes the aggregate utility for the whole country if the central government provides the local public good for a specific region or local government ($x=c$ in Equation (5)). Let us now aggregate all the regions and all the local governments to the country level. According to Equation (5), the specific regional utility of region j is:

$$U_j^R = N_j^R Q_j^R (V_j^R - C_j^R) - \frac{1}{2} V_j^R N_j^R \text{MAD}_j^R, j = 1 \dots n \quad (6)$$

Let us assume that all regions are similar except for the residents' preferences, i.e., all have the same number of residents, the same good value and cost (N^R, V^R, C^R), but different distributions ($Q_j^R, v_j^R, \text{MAD}_j^R$). Aggregating all n regions to the country level when the regions provide the public good and using Equations (2) and (6) yields

$$U^R = \sum_{j=1}^n U_j^R = N^C Q^C (V^R - C^R) - \frac{1}{2} N^C V^R \sum_{j=1}^n \frac{\text{MAD}_j^R}{n} \quad (7)$$

Similarly, the expression for the aggregate surplus when local governments provide the local public good is

$$U^L = \sum_{k=1}^m U_k^L = N^C Q^C (V^L - C^L) - \frac{1}{2} N^C V^L \sum_{k=1}^m \frac{\text{MAD}_k^L}{m} \quad (8)$$

⁸ Another proof of the Decentralization Theorem

⁹ With the Tiebout sorting assumption, the residents of both LGs are at a steady state of location choices, and do not relocate as a result of local provision of services.

The left argument of Equations (5), (7), and (8) describes the surplus from the local public good. The right arguments describe the average variability within each government tier. If the local public good has some local merit, we expect the variability (or dispersion) to increase as we move from the lower to the higher tier. When the valuation and the cost are similar at all tiers of government, the tier that better provides the local public good is the one where the average variability is minimal.

3.2 Spillovers

Residents of a local government may consume local public goods that an adjacent local government provides if they fit their preferences better than similar goods their local government provides.

Most papers that discuss spillovers use a model of two LGs, each providing a different local public good, where residents have limited resources. Therefore, residents decide about the mix of goods they will consume from the two LGs. We model spillovers based on a different behavior: both LGs provide similar impure public products, yet with different number of product units. Figure 1 illustrates an example of such two LGs. In this sense, our analysis refers to a complementary public good that governments at the same level provide, such as education, sanitation, cultural etc. Some residents are satisfied with the level of services their local government provides, yet others request better services (or more product units), and may consume the product from another local government. We chose this setting as it complements previous studies.

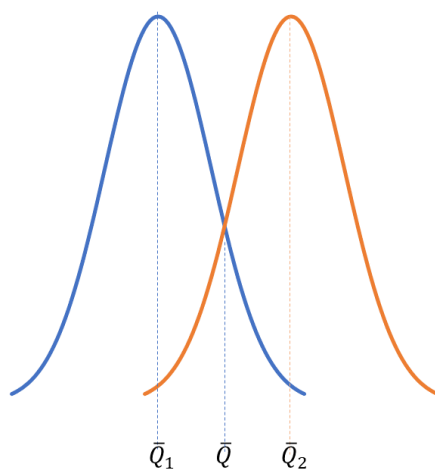


Figure 1: Product competition and spillovers

We assume that there are two local governments, LG₁ and LG₂, that belong to the same region R. Each of the LGs provides the same local public good but with different quantity $\bar{Q}_1 < \bar{Q}_2$. LG₁ and LG₂ are close enough so that residents of LG₁ can consume the local public good LG₂ provides, yet they pay for the good LG₁ provides, most commonly in taxes. There are two options: either each LG provides its residents with the locally average good, or the region provides the good to residents of both LGs with the average number of product units $\bar{Q} = \frac{1}{2}(\bar{Q}_1 + \bar{Q}_2)$.

When each LG provides the impure local public good, there are residents of LG₁ who wish to consume the local public good LG₂ provides, as they prefer a more significant number of product units. All residents of LG₂ consume the local good LG₂ provides. Equation (9) illustrates the aggregate surplus of the residents of each LG, where the two left-hand arguments are derived from Equation (5). The right-hand argument of U_1^L adds the utility of those residents from LG₁ who consumes the public goods LG₂ provides, where the right-hand argument of U_2^L subtracts the additional cost to LG₂.^{10 11}

$$\begin{aligned} U_1^L &= N^L \bar{Q}_1 (V^L - C^L) - \frac{1}{2} V^L N^L MAD_1 + V^L (\bar{Q}_2 - \bar{Q}_1) N^L (1 - \Phi(\alpha)) \\ U_2^L &= N^L \bar{Q}_2 (V^L - C^L) - \frac{1}{2} V^L N^L MAD_2 - C^L \bar{Q}_2 N^L (1 - \Phi(\alpha)) \end{aligned} \quad (9)$$

Equation (10) illustrates the aggregate surplus of both LGs, where $\Phi(\alpha)$ denotes the cumulative distribution function (CDF) of the residents' preferences for the local public good, and α is the number of standard deviations where residents whose preferences lie above $\alpha\sigma$ consume goods from LG₂. Therefore, there are $1 - \Phi(\alpha)$ residents of LG₁ who spillover to consume the goods that LG₂ provides. **Figure 2** shows two examples of $1 - \Phi(\alpha)$, one for a normal distribution and the second for an exponential distribution.

$$U^L = 2N^L \bar{Q} (V^L - C^L) - \frac{1}{2} V^L N^L (MAD_1 + MAD_2) + [\bar{Q}_2 (V^L - C^L) - V^L \bar{Q}_1] N^L (1 - \Phi(\alpha)) \quad (10)$$

¹⁰ We assume that both LGs have the same number of residents, that the value and the cost of the public good is the same in both LGs, and that there is no congestion and LG₂ can provide the spillover to LG₁.

¹¹ An upper-level government may compensate LG₂ for the spillovers of this public good through intergovernmental grants. In such a case the right-hand argument of U_2^L will be zero, increasing the aggregate surplus. In addition, ΔU^{L-R} increases, increasing the likelihood of decentralization policy.

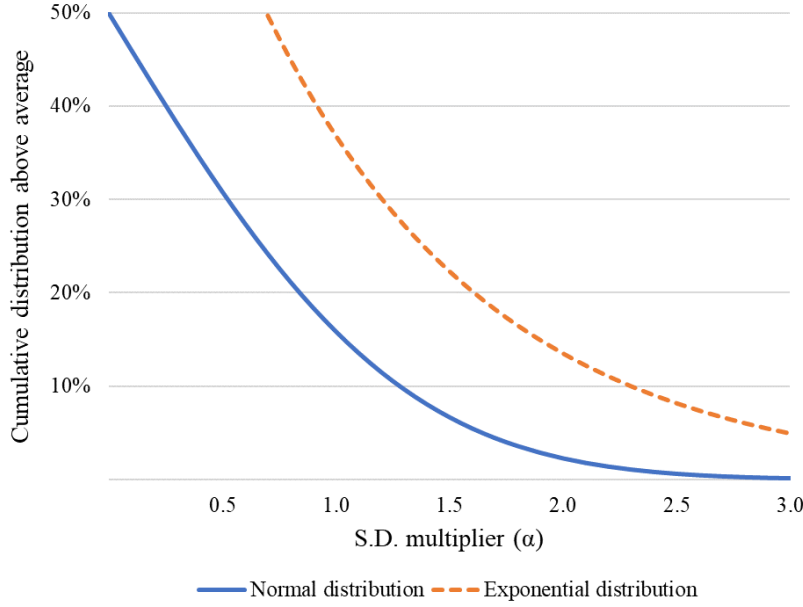


Figure 2: Share of residents of LG₁ requesting more than $\alpha\sigma$ product units and consume spillover from LG₂

When the regional government directly provides the local public good, it internalizes spillovers, yet the number of product units it provides is the average across all constituents. For simplicity, we assume that the number of LG₁ residents that their preferences exceed the average (\bar{Q}) is equal to the number of LG₂ residents whose preferences are lower than the average (see Figure 1), therefore, Equation (13) shows the aggregate surplus when the regional government provides the local public good.

To determine which level should provide the local public good, we subtract the aggregate surplus of residents under local provision from their aggregate surplus under regional government's provision.

$$\begin{aligned} \Delta U^{L-R} &= \frac{U^L}{N^L} - \frac{U^R}{N^L} = \\ &= V \left(\frac{1}{2} (\bar{Q}_2 - \bar{Q}_1) - \frac{1}{2} (MAD_1 + MAD_2) \right) + (\bar{Q}_2 (V - C) - V \bar{Q}_1) (1 - \Phi(\alpha)) \end{aligned} \quad (11)$$

Proposition 2: Centralization to one upper level is preferred when residents' preferences are sufficiently similar and spillovers are sufficiently large.

Proof: when $\bar{Q}_2 - \bar{Q}_1 < \sigma$, and noting that $MAD \leq \sigma$, the left-hand argument of Equation (11) is negative. We can write the right-hand argument as $\bar{Q}_2 \left(V(1 - \frac{\bar{Q}_1}{\bar{Q}_2}) - C \right) (1 - \Phi(\alpha))$. When preferences are sufficiently similar, we can assume that $\frac{C}{V} > (1 - \frac{\bar{Q}_1}{\bar{Q}_2})$, therefore,

the right-hand argument is also negative. When spillovers are high (small α), $1 - \Phi(\alpha) \gg 0$, resulting in $\Delta U^{L-R} < 0$. \square

Corollary 2: *when residents' preferences are significantly different, it is more beneficial for the local governments to provide the local public good, regardless of spillovers.*

Proof: when $\bar{Q}_2 \gg \bar{Q}_1$, the left-hand argument of Equation (11) is positive. When preferences are sufficiently different, we can assume that $\frac{c}{v} < (1 - \frac{\bar{Q}_1}{\bar{Q}_2})$, therefore, the right-hand argument is also positive. When spillovers are high (small α), $1 - \Phi(\alpha) \gg 0$, resulting in $\Delta U^{L-R} > 0$. \square

The rationale behind Proposition 2 and Corollary 2 is that when spillovers are significant, yet residents' preferences are sufficiently similar, it's better to internalize them at an upper level. But when residents' preferences diverge, it's better that local governments provide the public service, as they can customize the public good locally.

3.3 Economies of scale

So far, we assumed that the cost of providing the local public good is the same across the different government tiers. However, there may be differences in the cost based on economies of scale (see e.g., Andrews & Boyne, 2009). When the regional or central government provides the goods, it produces greater quantities compared to the number of product units a local government provides, resulting in lower average unit costs.¹²

Assuming that the per-unit cost is a declining function of quantity, we obtain that $C^C < C^R < C^L$.

The optimal government tier to provide the public good is the one that maximizes the aggregate surplus, namely $\max(U^C, U^R, U^L)$. To find out which is that tier, we subtract Equation (7) from Equation (5). Assuming that the good value is similar, and adding the economies of scale shows that:¹³

$$\Delta U^{C-R} = \frac{U^C}{N^C} - \frac{U^R}{N^C} = Q^C(C^R - C^C) - \frac{1}{2}V \left(MAD^C - \sum_{j=1}^n \frac{MAD_j^R}{n} \right) \quad (12)$$

¹² Caveat: lower-tier governments which face a smaller number of product units, may form a joint venture to purchase the good at greater quantities, or may outsource, which may result in a lower cost, exhibiting returns to scale. Yet, there can be differences in the management overhead, where lower-tier government, especially smaller local governments exhibit higher per resident costs.

¹³ Equation (12) presents the difference between U^C and U^R . The difference between U^R and U^L is analogous.

Proposition 3: *An upper-tier government is more effective in providing a local public good when the cost savings from economies of scale exceed the variability in residents' preferences.*

Proof: Given that $C^R > C^C$, the left-hand expression in Equation (12) remains positive. The right-hand term reflects the difference between national and regional average preference variability. Consequently, ΔU^{C-R} is positive when this difference is relatively small, indicating a preference for centralization. Conversely, if the cost saving due to economies of scale are insufficient to outweigh the difference in average preference variability between the national and regional levels, decentralization to a lower-tier government is preferred. \square

The underlying rationale of Proposition 3 is that greater variability in preferences suggests a wide range of needs across different regions, signaling heterogeneity in demand. Since centralization thrives on preferences homogeneity, highly diverse requirements may make decentralized governance more effective.

Equation (12) illustrates the contrast between national and regional governments. The same analytical framework applies to the distinction between regional and local governments, as generalized in Proposition 3.

3.4 Bi-modal distribution

In the previous analysis, we discuss general distributions characterized by mean, standard deviation, and MAD. Here we focus on a specific type of distribution – the bi-modal distribution. This specific case is interesting since there may be jurisdictions where the distribution of the residents' preferences is bi-modal. We may experience a bi-modal distribution in two options. The first is significant differences in residents' preferences within one local government. The second is when two local governments are part of the same regional government, and the regional government provides the goods and faces a bi-modal distribution of preferences (Jones et al., 2022).

We start with the first option. We may look at a bi-modal distribution as a joint of two distant-enough uni-modal distributions, each characterized by the mean (μ) and standard deviation (σ), and the mixing ratio (p). A bi-modal distribution requires that $|\mu_1 - \mu_2| \gg 2\sigma$. Without loss of generality, let us assume that the mixing ratio is half, the two resident groups have the same number of residents and exhibit the same standard deviation, yet they exhibit different preferences for the number of units of the local public good (\bar{Q}_1, \bar{Q}_2). The government

faces an average preference $\bar{Q} = p\bar{Q}_1 + (1 - p)\bar{Q}_2 = \frac{1}{2}(\bar{Q}_1 + \bar{Q}_2)$. We also assume that $\bar{Q}_1 < \bar{Q} < \bar{Q}_2$.

For clarity, let us denote by R the government which faces the bi-modal distribution and provides the local public good, and by L_1 and L_2 the residents' groups requesting \bar{Q}_1 and \bar{Q}_2 . Each group receives the local public good as a uni-modal per their preferences.¹⁴

When the R government provides the local public good, all of its residents receive the average number of product units \bar{Q} . Half of the population will receive more than their preference, and half will receive less. **Figure 3** shows an example of a bi-modal distribution.

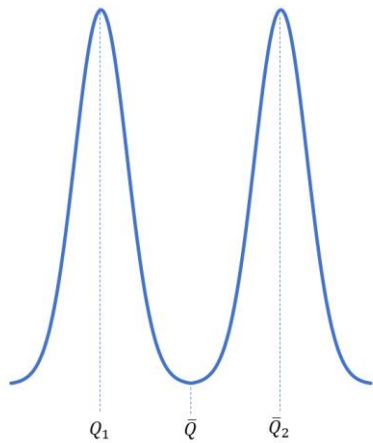


Figure 3: Normal bi-modal distribution

Note: the normal distribution is presented as an example. Our model is agnostic to the distribution function.

The aggregate surplus of the first group is $\frac{1}{2}N^R\bar{Q}_1V^R - \frac{1}{2}N^R\bar{Q}C^R$, and of the second group $\frac{1}{2}N^R\bar{Q}(V^R - C^R)$. The total aggregate surplus of the R government residents is

$$U^R = \frac{1}{2}N^RV^R(\bar{Q}_1 + \bar{Q}) - N^R\bar{Q}C^R \quad (13)$$

Now, we move to the second option and calculate the total aggregate surplus when each group is in a different local government (i.e., there is a geographic distinction). In such an arrangement, each group's residents receive the number of product units per the group's average. The surplus of each group will be according to Equation (5)

$$U^L = 2N^L\bar{Q}(V^L - C^L) - \frac{1}{2}V^LN^L(MAD_1 + MAD_2) \quad (14)$$

¹⁴ We use the second option of bi-modal occurrence, where two LGs (L_1 and L_2) are a part of a region. Therefore, we use L to symbolize LGs which exhibit uni-modal distribution, and R the government which face a bi-modal distribution.

Subtracting Equation (13) from Equation (14), assuming that the good value and cost are similar, shows that

$$\Delta U^{L-R} = \frac{U^L}{N^L} - \frac{U^R}{N^L} = V \left[\bar{Q} - \bar{Q}_1 - \frac{1}{2}(\text{MAD}_1 + \text{MAD}_2) \right] \quad (15)$$

Proposition 4: when residents' preferences exhibit a bi-modal distribution, it is more beneficial to split them into two groups and provide each group with its average preference.

Proof: in a bi-modal distribution $|\bar{Q}_1 - \bar{Q}_2| \gg 2\sigma$. Using the MAD characteristic $E(|x - \mu|) \leq \sigma$ yields that $\Delta U^{L-R} > 0$. Proposition 4 is an extension of Corollary 1a. \square

3.5 Political economy analysis

The classic approach to fiscal federalism assumes a benevolent planner, where officials make decisions based on the common good and aim to maximize aggregate welfare.¹⁵ The classic approach shows that the number of public good product units a government provides reflects the average preferences of its residents. Interestingly, when the distribution of preferences is symmetrical, this aligns with the median voter theorem (see Appendix B). The second generation of Fiscal Federalism relaxes this assumption, and incorporates political decision-making processes (Besley & Coate, 2003; Cheikbossian, 2016; Feidler & Staal, 2012; Lockwood, 2002; Oates, 2008). Previous studies show that the number of product units local governments provide may vary based on political negotiations, yet do not quantify the influence on residents. In this section we study the implication of a political decision-making process, where a higher-level government requests the provision of a specific number of product units from their local governments whose elected representatives form a coalition.

This part of our model builds on the minimum winning coalition (MWC) model (Besley & Coate, 2003; Riker, 1962). In MWC models, political actors form coalitions that are just large enough to secure a majority, maximizing per-member benefits while excluding unnecessary partners. The equilibrium emerges through strategic interactions where potential coalitions form and dissolve until no member can gain by defecting. MWCs systematically skew resource distribution, channeling benefits primarily to coalition members' constituencies while relatively neglecting opposition voters. This favoritism helps maintain electoral support from core constituents. We build on this model using public goods provision as a revealed

¹⁵ Oates (2008) indicates that the median voter model provides outcomes which do not deviate much from the benevolent planner model, therefore, the first assumption of the Decentralization Theorem does not require benevolence of public officials.

preference measure, analyzing scenarios where coalition voters prefer either high or low levels of public goods provision (compared to our baseline results depicted in Equation (5)). While we do not model the specific bargaining process of coalition formation, focusing on these contrasting outcomes allows us to examine how MWCs shape public goods allocation. This approach captures the fundamental tension between serving the preferences of coalition voters versus serving the preferences of the median voter. It provides theoretical traction on how political coalitions influence distributional outcomes through public goods provision.

In previous parts of the model, we assumed that regional ~~local~~ governments decide on the number of public product units based on the average preferences of their residents, which aligns with the median voter in a symmetrical distribution. In contrast, in this part of the model the decision-making process occurs in the house of representatives of a regional government, in two stages. In the first stage, residents of each local government elect their representative, reflecting the median voter preferences of that constituency, and in the second stage, representatives negotiate, form a coalition, and decide upon the number of product units to provide based on a MWC decisions. In this scenario, the regional government will provide Q^R product units based on the perception of the elected representatives, pro or averse public spending, and coalition negotiation results as prior studies describe (Besley & Coate, 2003; Cheikbossian, 2016; Feidler & Staal, 2012; Lockwood, 2002).

In any case, there are two options: (a) the region provides $Q^R > \bar{Q}^R$, or (b) the region provides $Q^R < \bar{Q}^R$.

The aggregate surplus of residents of region j is¹⁶

$$U_j^R = V \sum_{i=1}^{N_j^R} \min(Q_i, Q^R) - N_j^R C Q^R, \quad (16)$$

$i \in \{\text{residents of region } j\}$

Let us split Equation (16) into the above two options: $Q^R < \bar{Q}^R$, and $Q^R > \bar{Q}^R$. Let us denote the number of product units the region provides as α multiples of the standard deviation, $Q^R = \bar{Q}^R + \alpha\sigma$. **Figure 4** shows the case that $\alpha > 0$, and **Figure 5** when $\alpha < 0$. Let us further split the distribution into three areas as **Figure 4** shows for $\alpha > 0$, and into two areas as **Figure 5** shows for $\alpha < 0$. We then compute the aggregate surplus for each option.

When $Q^R > \bar{Q}^R$, the aggregate surplus described in Equation (16) becomes¹⁷

¹⁶ For convenience and clarity, we omit the subscript j of the region in the following equations.

¹⁷ For simplicity, when $\alpha > 0$ we approximate the average of area II as the mean between Q^R and \bar{Q}^R .

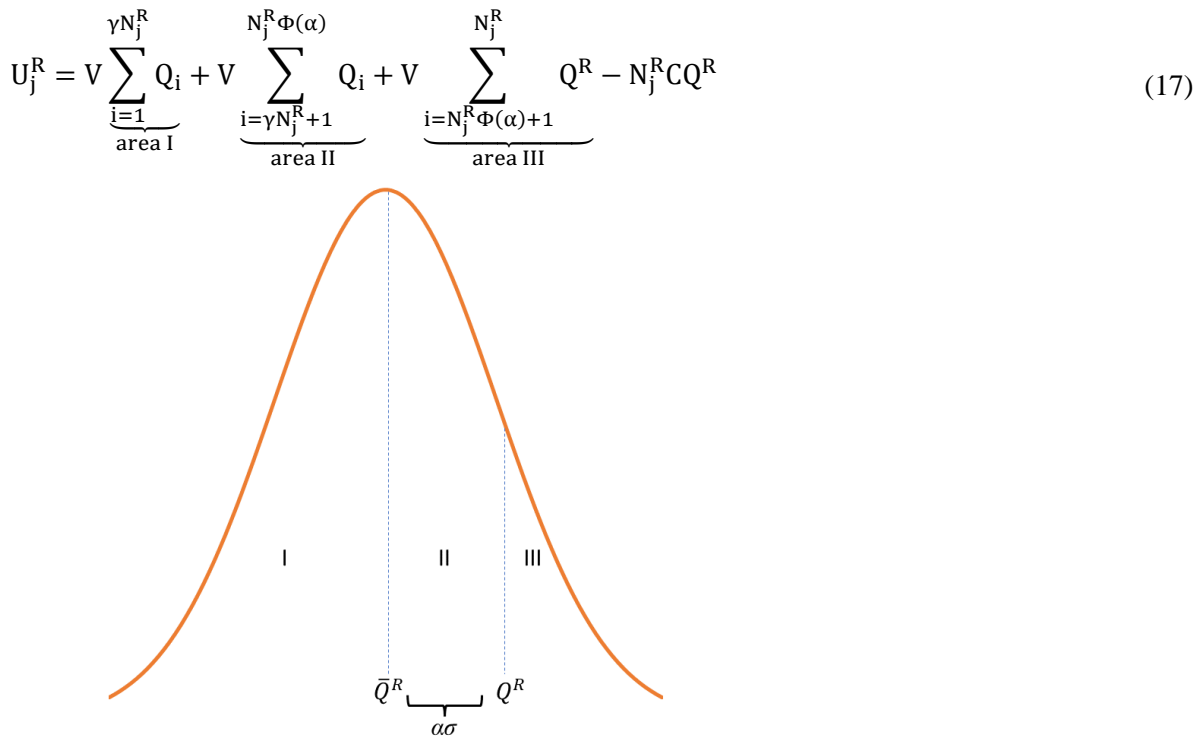


Figure 4: The region provides more product units than the average preferences of its residents

Note: The normal distribution shape is for clarity only

Rearranging Equation (17) gives

$$U^R = \frac{1}{2} V N^R \{ (\bar{Q}^R - Q^R) (\gamma + \Phi(\alpha)) + 2Q^R - M A D^R \} - N^R C Q^R \quad (18)$$

The difference between the classic approach, as shown in Equation (5), and the political economy approach shown in Equation (18) is

$$\Delta U^R = N^R \left\{ (\bar{Q}^R - Q^R) \left[\frac{1}{2} V (\gamma + \Phi(\alpha) - 2) + C \right] \right\} \quad (19)$$

The coalition elected representatives will benefit their constituencies if $\Delta U^R > 0$, which yields

$$\frac{V}{C} > \frac{2}{2 - \gamma - \Phi(\alpha)} > 1 \quad (20)$$

The right-hand side of **Figure 6** illustrates examples of Equation (20) for both the normal and exponential distributions. It can be observed that, with a normal distribution, if the elected representatives aim to benefit their constituencies by providing higher quality of the public product (e.g., $\alpha > 1$), they need to subsidize the cost to achieve $\frac{V}{C} > 2.5$. They may accomplish this through dedicated subsidies that have price effect.¹⁸ Such policy may reduce

¹⁸ It is assumed that the House cannot discriminate residents and local governments based on different cost, yet they may subsidize LGs differently.

the magnitude of the spillovers of that specific public product, as the region provides more product units.

When the elected representatives are averse to public spending, they may decide that the region will provide less product units than the preferences average, namely $Q^R < \bar{Q}^R$. In area I of **Figure 5**, residents' preferences are lower than Q^R , while in area II residents receive less product units than they prefer. To compute the surplus of residents in area I let us define the ratio between the truncated mean preferences of residents in area I to the number of products the region provides as

$$\int_{-\infty}^{Q^R} Q f(Q) dQ = \delta Q^R, \delta \in (0,1), f(Q) \text{ is the PDF of } Q \quad (21)$$

The aggregate surplus in this case is

$$U^R = V \underbrace{\sum_{i=1}^{N^R \Phi(\alpha)} Q_i}_{\text{area I}} + V \underbrace{\sum_{i=N^R \Phi(\alpha)+1}^{N^R} Q^R}_{\text{area II}} - N^R C Q^R \quad (22)$$

and rearranging Equation (22) gives

$$U^R = V N^R Q^R (1 - \Phi(\alpha)(1 - \delta)) - N^R C Q^R \quad (23)$$

Combining Equations (18) and (23) gives¹⁹

$$U^R = \begin{cases} V N^R Q^R (1 - \Phi(\alpha)(1 - \delta)) - N^R C Q^R, & Q^R < \bar{Q}^R \\ \frac{1}{2} V N^R ((\bar{Q}^R - Q^R)(\gamma + \Phi(\alpha)) + 2Q^R - MAD^R) - N^R C Q^R, & Q^R > \bar{Q}^R \end{cases} \quad (24)$$

The difference between the classic approach, as shown in Equation (5), and the political economy approach shown in Equation (23) is

$$\Delta U^R = N^R (V - C)(Q^R - \bar{Q}^R) + V N^R \left(\frac{1}{2} MAD^R - Q^R \Phi(\alpha)(1 - \delta) \right) \quad (25)$$

The coalition elected representatives will benefit their constituencies if $\Delta U^R > 0$, which yields

$$\frac{C}{V} > 1 - \frac{\frac{1}{2} MAD^R - Q^R \Phi(\alpha)(1 - \delta)}{\bar{Q}^R - Q^R} \quad (26)$$

Figure 6 presents a simulation of Equations (20) and (26) for normal distribution, while indicating the frontier where the representatives' decision benefit residents compared to when the region provides the average of its residents' preferences per the classical approach.

¹⁹ When $Q^R = \bar{Q}^R, \gamma = \Phi(\alpha)$, and $\delta = 1 - \frac{1}{2\gamma\bar{Q}^R} MAD^R$, we obtain the same equation on both rows, as Equation (6) presents.

What we deduce from Equations (20) and (26) is: (a) if the elected representatives are pro public spending, they'll decide to provide more product units such that $Q^R > \bar{Q}^R$. Residents will appreciate such approach if the benefits outperform the alternative standard approach, meaning that the ratio of perceived value to cost is higher (area above the curve). As the cost cannot be reduced, the region will have to subsidize at least those LGs which are a part of the coalition through earmarked transfers. (b) if the elected representatives are averse public spending, they may decide to provide less product units, such that $Q^R < \bar{Q}^R$. Similarly, residents will appreciate such policy, provided that the cost will be close to the perceived value (area below the curve). When the perceived product value is so low, we may argue that residents care less about its provision, so the elected representatives may decide to provide fewer product units only to those products that have low perceived value.

As an example of such decision-making when residents highly value a service and Minimum Winning Coalition (MWC) representatives wish to benefit their constituents, consider public education. With MWC, the House may decide to provide more facilities to select schools than what average residents in that jurisdiction prefer. In such cases, the region must subsidize these extra-curricular activities. Conversely, consider street lighting. Assume that while average resident preferences call for a certain light intensity, residents are relatively indifferent to lower intensity levels—in other words, the elasticity of demand for street lighting is low. In this case, the ratio of perceived valuation to cost may be low enough that the House decides to reduce illumination intensity.

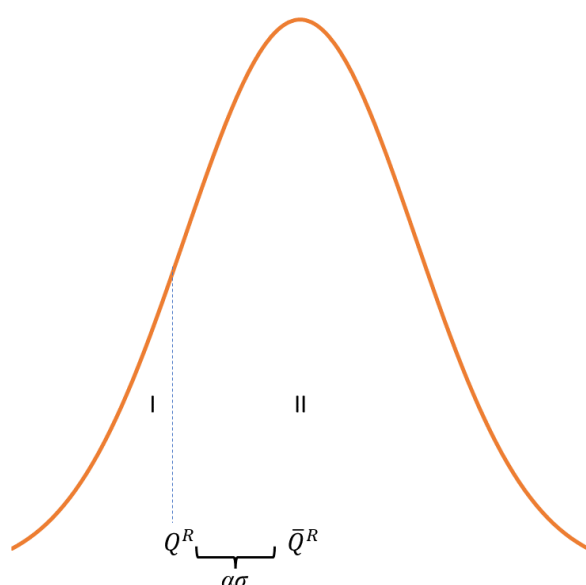


Figure 5: The region provides less product units than the average preferences of its residents

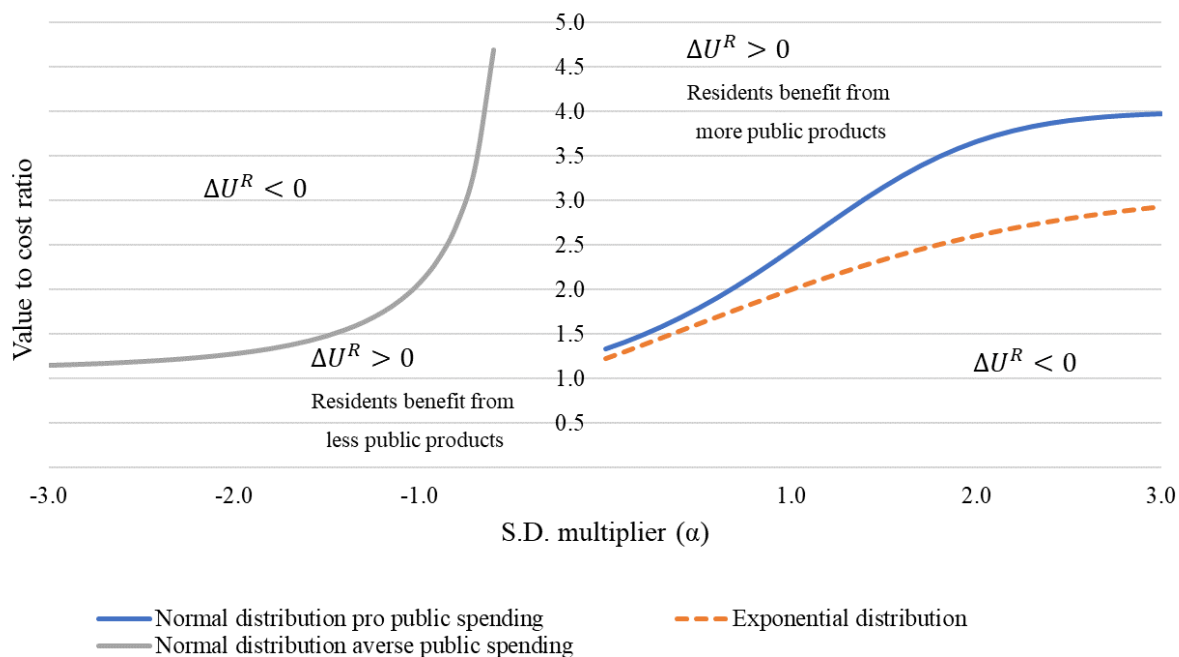


Figure 6: The ratio of perceived value to cost when the region provides more (Equation (20), right-hand side) or less (Equation (26), left-hand side) product units than the residents' average preferences

Note: Figure 6 shows the ratio of perceived value to cost which differentiates between the space where subsidies are economically beneficial and when they are not. The right-hand side illustrates examples of Equation (20) for both the normal and exponential distributions when the region provides more product units than the optimal average. The left-hand side show example of Equation (26) for normal distribution when the region provides less product units than the optimal average.

4 Discussion and conclusion

The literature on fiscal federalism typically assumes a two-tier government structure and examines the allocation of budgets, projects, and powers between those tiers. Our study introduces a third tier - the regional tier - and demonstrate how this intermediate tier effectively addresses local spillovers. Building on prior studies, we relax the assumption of homogenous subnational governments and recognize that residents within SNG units may have diverse preferences for local public goods. Residents whose preferences exceed the level of services their SNG provides may consume those goods from other SNGs. In line with the Decentralization Theorem (Oates, 1972) and the Tiebout model (Tiebout, 1956), we show that heterogeneity of residents' preferences reduces the aggregate surplus compared to a homogenous community. Furthermore, we demonstrate that when the cost of providing the

good is the same across SNGs, the lower tier should provide the local public good, thereby offering another proof of the Decentralization Theorem.

Most studies also assume that the cost of the public good is the same across SNGs. They note that if the product exhibits economies of scale, the central government should provide the good. Our analysis reveals a more nuanced perspective: economies of scale make it more beneficial for an upper-tier government to provide the local public good if the cost savings resulting from scale outweigh the impact of variations in residents' preferences.

Our approach to handling spillovers diverges from that of previous studies. Earlier research typically assumes two distinct public goods, with each SNG providing different goods. Residents of LG_1 consume both the good that their LG provides and the good that LG_2 provides, and vice-versa. In contrast, our model posits that both SNGs provide similar public goods but in different quantities. Each SNG supplies an average number of product units based on residents' preferences. Consequently, residents in LG_1 whose preferences exceed the average can consume the goods that LG_2 provides, provided that they can access them. For example, in education services, some residents may seek high schools offering advanced STEM classes and enroll their children in schools in another SNG that provides such classes. We consider our assumption more realistic and more general because it reflects how elected representatives seek to leverage their region in providing public goods that local residents deem more important. This single-product model is more realistic as it better captures the nuanced variations in service levels across jurisdictions, rather than assuming completely distinct services. It is also more general, as it can accommodate both scenarios where services differ in quality or quantity, thus encompassing a wider range of real-world situations.

We demonstrate that determining which tier is more efficient in providing a specific local public good hinges on the level of heterogeneity between SNGs and the level of homogeneity within each SNG. Moreover, contrary to the findings of most previous studies, we show that it can be advantageous for local governments to provide the public good even in the presence of spillovers. This outcome depends on the nature of the local public good and the characteristics of the residents. Specifically, differences in residents' preferences (i.e., the level of heterogeneity) moderate the relationship between the significance of spillovers and the more beneficial tier in providing the public good.

To provide a summary of our results and how they complement prior studies we have augmented Table 1 with our findings, adding the new forces and the moderation effect. **Table 3** summarizes the additional forces.

Table 3: Additional forces influencing centralization - decentralization allocation of powers

Force	Pro-Decentralization	Pro-Centralization	Comment
Economies of scale	no economies of scale	economies of scale exist	Corollary 1b
Economies of scale vs. preference variability	Cost saving < preference variability	Cost saving > preference variability	Lemma 3
Spillovers High preferences variability	Significant	Insignificant	Corollary 2
Spillovers Small Preferences variability	Insignificant	Significant	Corollary 2
Preferences variability between LGs	Large	Small	Proposition 2 Corollary 2

We can group the forces that influence the decision-making process of the centralization - decentralization allocation of goods, as Tables 2 and 3 show, into five categories: Economic, Residents, Government Structure, Public Good Characteristics, and Politics. **Figure 7** presents the allocation of public goods in the multi-tier subnational government forces model inspired by Porter (1985, p. 5).

Explanations for the categories. (1) Economic: This category includes factors such as economies of scale, cost efficiency, and spillovers. (2) Residents: This category encompasses factors related to citizen preferences, heterogeneity of demand, and information asymmetry. (3) Government Structure: This refers to the organization and levels of government (e.g., local, regional, central) and how they interact in the provision of public goods, and the size of the government. (4) Public Good Characteristics: This category includes aspects specific to the nature of the public goods being provided, such as rivalry, excludability, and the presence of spillovers. (5) Politics: This category covers political factors that influence decision-making, such as voting patterns, political ideologies, and governance structures.

These categories provide a comprehensive framework for analyzing the various forces that affect the allocation of public goods across different tiers of government.

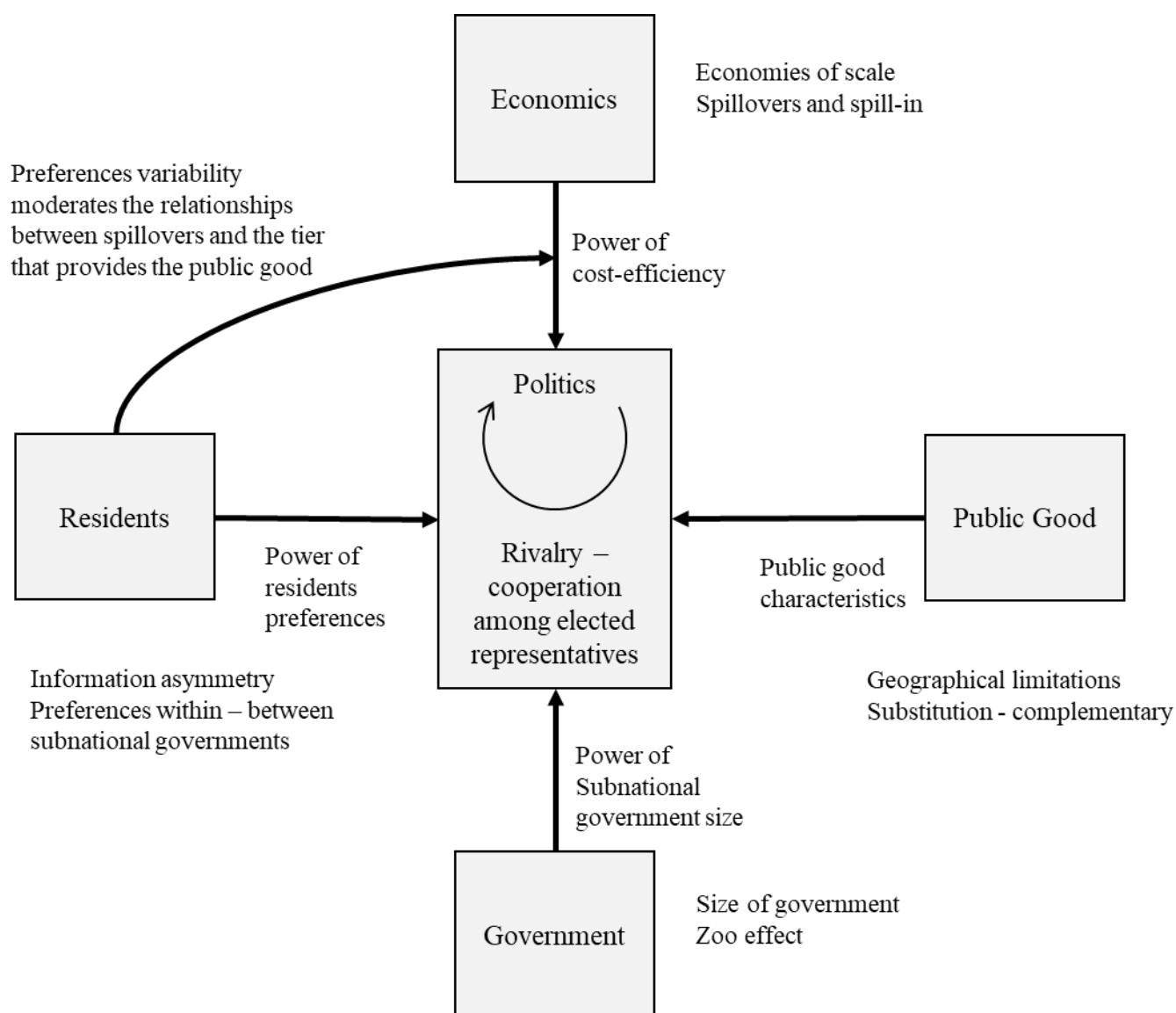


Figure 7: Allocation of public goods in multi-tier subnational governments forces model

The political process at the center of the forces model in **Figure 7** integrates all forces and the moderation effect into the decision-making process. Second-generation Fiscal Federalism studies focusing on the political game assume that the decision-making process occurs through negotiations between elected representatives. These negotiations may result with a decision to provide more or fewer product units than the average. We study the consequences of each option, demonstrating that when representatives decide to provide more product units, the region will have to subsidize the cost. Conversely, they'll decide to provide fewer product units if the perceived value is low.

By combining Tables 2 and 3, we create a decision table that can assist policymakers and administrators in determining which public goods should be decentralized and which will

be better served by an upper tier. Appendix C and Table C1 provide an example of using such a decision table.

Our research introduces a new model of fiscal federalism that incorporates a third, regional tier of subnational government and addresses local spillovers. This model relaxes the assumption of homogeneous SNGs, allowing for diverse resident preferences for local public goods. The model demonstrates that heterogeneity in preferences can reduce aggregate surplus compared to homogeneous communities. It also shows that when costs are uniform across SNGs, lower tiers should provide local public goods, supporting the Decentralization Theorem. The study explores how economies of scale and spillovers affect the optimal tier for providing public goods, introducing a moderation effect between resident preference heterogeneity and spillover significance. It presents a multi-tier SNG forces model, categorizing influences on centralization-decentralization decisions into economic, resident preference, SNG size, public good characteristics, and political powers. The study concludes with a decision table to assist policymakers in determining the appropriate tier for providing specific public goods. By providing both the theoretical underpinnings of decentralization and a practical tool for applying this theory we aim to contribute not only to academic research but also to applied research and policy making. The decision table we've developed offers policymakers a valuable resource for determining the appropriate tier for providing specific public goods, bridging the gap between theory and practice in fiscal federalism.

References

- Andrews, R., & Boyne, G. (2009). Size, Structure and Administrative Overheads: An Empirical Analysis of English Local Authorities. *Urban Studies*, 46(4), 739–759. <https://doi.org/10.1177/0042098009102127>
- Barankay, I., & Lockwood, B. (2007). Decentralization and the productive efficiency of government: Evidence from Swiss cantons. *Journal of Public Economics*, 91(5–6), 1197–1218. <https://doi.org/10.1016/j.jpubeco.2006.11.006>
- Barzel, Y. (1969). Two propositions on the optimum level of producing collective goods. *Public Choice*, 6(1), 31–37. <https://doi.org/10.1007/BF01718576>
- Besley, T., & Coate, S. (2003). Centralized versus decentralized provision of local public goods: A political economy approach. *Journal of Public Economics*, 97, 2611–2637. [https://doi.org/10.1016/S0047-2727\(02\)00141-X](https://doi.org/10.1016/S0047-2727(02)00141-X)
- Chatry, I., & Vincent, R. C. (2019). A global view of sub-national governments in Asia: Structure and finance. In J. Kim & S. Dougherty (Eds.), *Fiscal Decentralisation and Inclusive Growth in Asia* (pp. 27–57). OECD Fiscal Federalism Studies. <https://doi.org/10.1787/239e95db-en>
- Cheikbossian, G. (2016). The political economy of (De)centralization with complementary public goods. *Social Choice and Welfare*, 47(2), 315–348. <https://doi.org/10.1007/s00355-016-0962-3>
- Dur, R., & Roelfsema, H. (2005). Why does centralisation fail to internalise policy externalities? *Public Choice*, 122(3–4), 395–416. <https://doi.org/10.1007/s11127-005-5290-6>
- Dur, R., & Staal, K. (2008). Local public good provision, municipal consolidation, and national transfers. *Regional Science and Urban Economics*, 38(2), 160–173. <https://doi.org/10.1016/j.regsciurbeco.2008.01.005>
- Feidler, J., & Staal, K. (2012). Centralized and decentralized provision of public goods. *Economics of Governance*, 13(1), 73–93. <https://doi.org/10.1007/s10101-011-0104-4>
- Frère, Q., & Védrine, L. (2024). Does decentralisation theorem shape intermunicipal cooperation? *Regional Studies*, 58(9), 1728–1753. <https://doi.org/10.1080/00343404.2023.2282139>
- Geys, B., & Konrad, K. A. (2010). *Federalism and Optimal Allocation Across Levels of Governance* (SP II 2010-09; WZB Markets and Politics). <https://hdl.handle.net/10419/54598>

- Gregor, M., & Stastna, L. (2012). The decentralization tradeoff for complementary spillovers. *Review of Economic Design*, 16(1), 41–69. <https://doi.org/10.1007/s10058-012-0113-y>
- Halásková, M., & Halásková, R. (2015). Fiscal decentralisation and provision of local public services in selected EU countries. *Lex Localis*, 13(3), 595–613. [https://doi.org/10.4335/13.3.595-613\(2015\)](https://doi.org/10.4335/13.3.595-613(2015))
- Hämäläinen, K., & Moisio, A. (2015). *One or Two Tiers of Local Government? – The Cost Effects of a Regional Experiment* (VATT No. 65).
- Jones, M. I., Sirianni, A. D., & Fu, F. (2022). Polarization, abstention, and the median voter theorem. *Humanities and Social Sciences Communications*, 9(1), 1–12. <https://doi.org/10.1057/s41599-022-01056-0>
- Koethenbuerger, M. (2008). Revisiting the “Decentralization Theorem”-On the role of externalities. *Journal of Urban Economics*, 64(1), 116–122. <https://doi.org/10.1016/j.jue.2007.10.001>
- Lockwood, B. (2002). Distributive politics and the costs of centralization. *Review of Economic Studies*, 69(2), 313–337. <https://doi.org/10.1111/1467-937X.00207>
- Lorz, O., & Willmann, G. (2005). On the endogenous allocation of decision powers in federal structures. *Journal of Urban Economics*, 57(2), 242–257. <https://doi.org/10.1016/j.jue.2004.11.002>
- Martinez-Vazquez, J., Lago-Peñas, S., & Sacchi, A. (2017). The Impact of Fiscal Decentralization: a Survey. *Journal of Economic Surveys*, 31(4), 1095–1129. <https://doi.org/10.1111/joes.12182>
- Musgrave, R. A. (1959). *The theory of public finance : a study in public economy*. McGraw-Hill.
- Oates, W. E. (1972). *Fiscal Federalism*. Harcourt Brace Jovanovich.
- Oates, W. E. (1988). On the measurement of congestion in the provision of local public goods. *Journal of Urban Economics*, 24(1), 85–94. [https://doi.org/10.1016/0094-1190\(88\)90048-4](https://doi.org/10.1016/0094-1190(88)90048-4)
- Oates, W. E. (2005). Toward a second-generation theory of fiscal federalism. *International Tax and Public Finance*, 12, 349–373. <https://doi.org/10.1007/s10797-005-1619-9>
- Oates, W. E. (2008). On the evolution of fiscal federalism: Theory and institutions. *National Tax Journal*, 61(2), 313–334. <https://doi.org/10.17310/ntj.2008.2.08>
- OECD. (2022). 2022 Synthesis Report World Observatory on Subnational Government Finance and Investment. In *OECD*. <https://doi.org/https://doi.org/10.1787/b80a8cdb-en>
Authors

- Olson, M. (1969). The Principle of “Fiscal Equivalence”: The Division of Responsibilities among Different Levels of Government. *American Economic Review*, 59(2), 479–487. <https://doi.org/10.2307/1823700>
- Ponzetto, G. A. M., Backus, P., Berry, C., Bustos, P., Carvalho, V., Cicala, S., Donder, P. De, Enikolopov, R., Gancia, G., & Nuevo-, A. (2016). Political Centralization and Government Accountability. *The Quarterly Journal of Economics*, 131(1), 381–422. <https://doi.org/10.1093/qje/qjv035>.Advance
- Porter, M. E. (1985). *Competitive Advantage: creating and sustaining superior performance*. The Free Press.
- Reingewertz, Y. (2014). Fiscal Decentralization - A Survey of the Empirical Literature. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2523335>
- Riker, W. H. (1962). The Theory of Political Coalitions. In *The Western Political Quarterly*. Yale University. [https://ia801405.us.archive.org/9/items/RikerTheTheoryOfPoliticalCoalitions/Riker-The Theory of Political Coalitions.pdf](https://ia801405.us.archive.org/9/items/RikerTheTheoryOfPoliticalCoalitions/Riker-TheTheoryofPoliticalCoalitions.pdf)
- Seabright, P. (1996). Accountability and decentralisation in government: An incomplete contracts model. *European Economic Review*, 40, 61–89. [https://doi.org/10.1016/0014-2921\(95\)00055-0](https://doi.org/10.1016/0014-2921(95)00055-0)
- Shani, R., Reingewertz, Y., & Vigoda-gadot, E. (2023). Regional governments in the rural space: the effectiveness of dual-tier regional councils. *Regional Studies*, 57(9), 1816–1831. <https://doi.org/10.1080/00343404.2022.2153825>
- Slavinskaitė, N., Liučvaitienė, A., & Gedvilaitė, D. (2019). Theoretical Analysis of the Fiscal Federalism. *International Journal of Contemporary Economics and Administrative Sciences*, 9(2), 250–267. <https://doi.org/10.5281/zenodo.3595926>
- Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64(5), 416–424. <https://doi.org/10.1086/257839>
- Weingast, B. R. (2009). Second generation fiscal federalism: The implications of fiscal incentives. *Journal of Urban Economics*, 65, 279–293. <https://doi.org/10.1016/j.jue.2008.12.005>

Appendix A: The relations between MAD and mean

The mean absolute deviation around the average indicates the average (absolute) distance from the mean. The mean absolute deviation of a set $\{x_1, x_2, \dots, x_n\}$ is $\frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$.

Let us express the average of all the observations which are either less than the average (\underline{x}^*) or above it (\bar{x}^*) using the mean and MAD for any arbitrary distribution. γ is the ratio of the number of observations less than the average to the total number of observations. For symmetrical distribution $\gamma=1/2$. In general, $\gamma \in (0,1)$, and may be approximated by $\gamma = \frac{\text{mean}}{\text{mean} + \text{median}}$. The averages of the observations, which are either above or below the average, are:

$$\underline{x}^* \equiv \frac{1}{\gamma n} \sum_{\substack{i=1 \\ \text{s.t. } x_i < \bar{x}}}^{\gamma n} x_i ; \bar{x}^* \equiv \frac{1}{(1-\gamma)n} \sum_{\substack{i=\gamma n+1 \\ \text{s.t. } x_i > \bar{x}}}^n x_i \quad (\text{A1})$$

The definition of MAD is

$$MAD = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}| = \bar{x}(2\gamma - 1) - \gamma \underline{x}^* + (1 - \gamma) \bar{x}^* \quad (\text{A2})$$

Similarly, the average

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \gamma \underline{x}^* + (1 - \gamma) \bar{x}^* \quad (\text{A3})$$

Subtracting A2 from A3 yields

$$\underline{x}^* = \bar{x} - \frac{1}{2\gamma} MAD \quad (\text{A4})$$

Adding A2 to A3 yields

$$\bar{x}^* = \bar{x} + \frac{1}{2(1-\gamma)} MAD \quad (\text{A5})$$

Appendix B: The optimal allocation of public products – extending the classical approach

Based on Oates Fiscal Federalism, 1972, page 59, which in turn is based on Barzel (1969) “Two propositions on the optimum level of producing collective goods”.

Assume that the cost of additional unit is the same, and the demand curves are linear, at least over the relevant range as Figure B1 shows. C_0 is the product-unit cost to each resident, which results in the demand of Q_i product units. Let us assume two cases: the government provides $\underline{Q}_g < Q_i$, or $\bar{Q}_g > Q_i$. In the first option, residents receive less product units than desired, and their welfare loss is the shaded upper triangle in Figure B1. In the second option residents receive more product units than they prefer, but pay (in taxes) for all the product units their government provides. Their welfare loss is the added cost of the product units above their desired level, which the lower triangle describes.

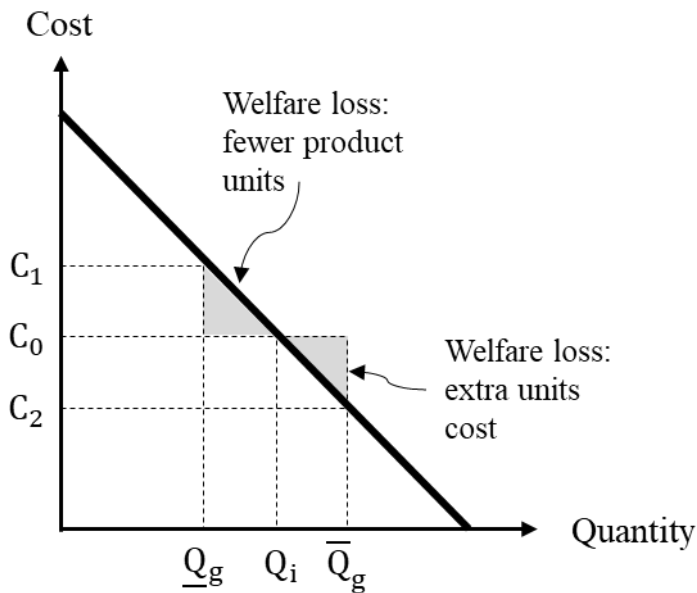


Figure B1: Demand curve of public products

Under the assumption of linear demand curves, the relations between per-product unit cost and the desired quantity is $Q_i = a_i + b_i C_i$, $b_i < 0$. The welfare loss due to fewer product units is the area of the upper triangle:

$$\Delta W_i = \frac{1}{2} (Q_i - \underline{Q}_g) (C_1 - C_0) = \frac{-1}{2b_i} (Q_i - \underline{Q}_g)^2 \quad (B1)$$

Further assuming that all residents have the same slope of the demand curve and the same marginal rate of substitution, we can sum the welfare loss for all residents. Let $\gamma \in (0,1)$

be the ratio of the number of residents whose preferences are below the supplied product units, implying that $n(1 - \gamma)$ is the number of residents whose preferences are above the supplied product units. Therefore, the total welfare loss of the first case is

$$\Delta \underline{W} = \sum_{\substack{i=\gamma n+1 \\ \text{s.t. } Q_i > \underline{Q}_g}}^n \Delta \underline{W}_i = \frac{-1}{2b} \sum_{\substack{i=\gamma n+1 \\ \text{s.t. } Q_i > \underline{Q}_g}}^n (Q_i - \underline{Q}_g)^2 \quad (\text{B2})$$

Minimizing $\Delta \underline{W}$ results in ²⁰

$$\underline{Q}_g^* = \frac{1}{(1 - \gamma)n} \sum_{\substack{i=\gamma n+1 \\ \text{s.t. } Q_i > \underline{Q}_g}}^n Q_i \quad (\text{B3})$$

Using the MAD definition,²¹ we can write Equation (B3) as

$$\underline{Q}_g^* = \bar{Q} + \frac{1}{2(1 - \gamma)} \text{MAD}, \text{ where } \bar{Q} = \frac{1}{n} \sum_{i=1}^n Q_i \quad (\text{B4})$$

Similarly, for the case where residents receive more product units than they prefer, the welfare loss will be the area of the lower triangle.

$$\Delta \bar{W}_i = \frac{1}{2} (\bar{Q}_g - Q_i) (C_0 - C_2) = \frac{-1}{2b_i} (Q_i - \bar{Q}_g)^2 \quad (\text{B5})$$

$$\Delta \bar{W} = \sum_{\substack{i=1 \\ \text{s.t. } Q_i < \bar{Q}_g}}^{\gamma n} \Delta \bar{W}_i = \frac{-1}{2b} \sum_{\substack{i=1 \\ \text{s.t. } Q_i < \bar{Q}_g}}^{\gamma n} (Q_i - \bar{Q}_g)^2 \quad (\text{B6})$$

Minimizing $\Delta \bar{W}$

$$\bar{Q}_g^* = \frac{1}{\gamma n} \sum_{\substack{i=1 \\ \text{s.t. } Q_i < \bar{Q}_g}}^{\gamma n} Q_i = \bar{Q} - \frac{1}{2\gamma} \text{MAD} \quad (\text{B7})$$

Welfare loss for the whole residents is $\Delta W = \Delta \underline{W} + \Delta \bar{W}$. Minimizing ΔW gives the number of product units that minimizes the welfare loss to be the average over all residents' preferences

$$\bar{Q} = \frac{1}{n} \sum_{i=1}^n Q_i = \gamma \bar{Q}_g^* + (1 - \gamma) \underline{Q}_g^* \quad (\text{B8})$$

²⁰ Minimizing ΔW by $\frac{d\Delta W}{dQ_i} = 0, \frac{d^2\Delta W}{dQ_i^2} > 0$

²¹ See Appendix A

Appendix C: Using the decision table

Implementing the theoretical model, we describe in this study, involves estimating the various forces influencing whether to decentralize powers to sub-national governments. Table C1 presents an application of our model as a tool for a decision-making process in a sub-national structure having three tiers. The planner has to estimate the magnitude of each force between central and regional or regional and local tiers. We provide an example of such a decision table, analyzing which tier will better provide education services. Table C1 lists hypothetical estimates of all forces, where column (5) shows the ranking of local vs. regional tiers, and column (6) ranks regional vs. central governments. Each force ranges between 5 (full force to decentralize), and -5 (full force to centralize).

We describe here the hypothetical numbers given in this theoretical example. For economies of scale the local-regional number equals -3, meaning that centralization to the regional level is moderately preferred to local provision due to economies of scale. When we compare the regional to the central level there is a small advantage to centralize the provision of education services (-1). Moving to spillovers, we assume that these are very small, giving a small advantage to the local vs. the regional level (+1). The third item is preference variability, which we assume is large, giving an advantage to the level closer to the residents (+4 for local vs. regional, and +4 for regional vs. central). For each additional item in the table, we plug in numbers which correspond to the intensity of the advantage to centralize/decentralize. Finally, we sum all numbers, getting a value of -0.45 for the local-regional comparison, meaning a small advantage to centralize. The value for the regional-central comparison equals 1.09, i.e., there is an advantage to decentralize to the regional level. thus, for education services which exhibit the magnitudes of forces listed in the table there are efficiency gains in providing these services at the regional level.

The last row of Table C1 shows the average of each column with equal weights. The result of the local-regional column (5) is a negative average force, meaning it will be better to “locally centralize” the good at the regional tier. The result of the regional-central column (6) is a positive average force, meaning it will be better to “locally centralize” at the regional tier and not centralize at the central government level.

Future studies may probe further into the question of weights and develop methods to evaluate each force's magnitude.

Table C1: SNG decision table with education service example

Force group	Force	Decentralization	Centralization	Local - regional	Regional - central
(1)	(2)	(3)	(4)	(5)	(6)
Economics	Economies of scale	insignificant	Significant	-3	-1
Economics	Spillovers	Small	Significant	1	0
Economics	Public good cost vs. preference variability	Cost saving < preference variability	Cost saving > preference variability	3	3
Residence	Information asymmetry	High	Low	4	4
Residents	Residents' preferences between regions	Homogenous	Heterogenous	-3	-1
Residents	Preferences variability between LGs	Large	Small	3	1
Governance	Zoo effect: good scale	Small	Large	-3	3
Governance	Size of local government	Large	Small	-5	2
Public good	Spill-in public goods	Substitute	Complementary	2	2
Public good	Geographic limitations	Bounded	Unlimited	-2	1
Politics	Representation	Averse public spending	Pro public spending	-2	-2
Moderation	Spillovers & High preferences variability	Significant	Insignificant		
Moderation	Spillovers & Small Preferences variability	Insignificant	Significant		
Average				-0.45	1.09