CESifo Working Papers

www.cesifo.org/wp

Public Provision and Local Income Tax Competition

Florian Kuhlmey
Beat Hintermann

CESIFO WORKING PAPER NO. 5789
CATEGORY 1: PUBLIC FINANCE
MARCH 2016

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5789*

Public Provision and Local Income Tax Competition

Abstract

We extend the literature on local income tax competition by allowing for inter-jurisdictional spillovers and imperfect rivalry in consumption of a publicly provided good. Comparing decentralized second-best results of a theoretical model with an efficient benchmark, we identify three inefficiencies: (1) imperfect redistribution; (2) inter-community free-riding; and (3) an inefficient allocation of the population. We quantify the relative size of these inefficiencies in a numerical implementation of the theoretical model, which reveals that free-riding rises unambiguously in the level of spillovers, whereas the welfare losses from (1) and (3) depend nonlinearly on the levels of spillovers and rivalry.

JEL-Codes: H210, H400, H770, Q580.

Keywords: publicly provided goods, tax competition, fiscal federalism, decentralization, free-riding, welfare analysis.

Florian Kuhlmey
University of Basel
Faculty of Business and Economics
Peter-Merian-Weg 6
Switzerland – 4002 Basel
florian.kuhlmey@unibas.ch

Beat Hintermann
University of Basel
Faculty of Business and Economics
Peter-Merian-Weg 6
Switzerland – 4002 Basel
b.hintermann@unibas.ch

February 19, 2016

1 Introduction

According to Oates' decentralization theorem (1972), every publicly provided good or service should be provided by the lowest possible entity which contains all households affected by the public intervention. Since this requires an impracticable plurality of entities, perfect correspondence does usually not apply in practice. In the context of climate change, for example, adaptation measures may be regionally or even locally determined, yet they can affect other jurisdictions (e.g. building a dam also protects regions located downriver). Likewise, the use of public facilities such as airports, sports stadiums or opera houses is not restricted to the local population. Under fiscal federalism, local jurisdictions pay a part or all of the expenses associated with such public provision. To finance these expenses, they tax mobile households or firms and thus compete with each other to offer an attractive package of tax rates and public provision.

There is a large literature about fiscal federalism in the context of mobile households and firms, following the seminal contributions of Tiebout (1956). A subset of this literature focuses on the relative merits of centralized and decentralized provision of public goods; for a review, see Oates (1999) and Boadway & Tremblay (2012). Because many papers use environmental regulation as their example of the public good, this literature has become known under the label of "environmental federalism". Wellisch (1994) shows that decentralized provision of public goods that exhibit interregional spillovers is efficient, provided that migration is costless.[1] This result is driven by an assumption that the production function is concave: The larger the population in a jurisdiction, the lower is the average productivity of labor. Interregional spillovers, along with monetary transfers to other jurisdictions, effectively serve to achieve the desired population level from a region's point of view, by making people in other jurisdictions better off.

Importantly, models of environmental federalism abstract from real-world taxation and instead rely on head taxes to finance the public good. Because people receive exactly the public provision they pay for, these taxes are known as "benefit" taxes. This contrasts with models of "non-benefit" tax competition, where jurisdictions compete with each other about mobile tax bases such as capital or

[1] Hoel & Shapiro (2003, 2004) and Hoel (2004) generalize Wellisch's (1994) model to a broader class of externalities and government instruments. Bucovetsky (2011) derives sufficiency conditions for decentralized actions to be Pareto efficient.

households' income, but which abstract from inter-jurisdictional spillovers.[2] In such a setting, public provision is inefficient even for the typically assumed case of a publicly provided private good, i.e., a good that is strictly local and perfectly rival in consumption. The inefficiency is due to two types of fiscal externalities: (i) Regional governments (or equivalently, median voters) consider the outflow of the tax base in response to an increase in the regional tax as a cost, but they do not consider the benefit of this tax base relocation to other regions; (ii) households choose their preferred residence based on private utility maximization and thus neglect the fiscal impact they impose on other households in the same jurisdiction (with non-benefit taxation, households generally contribute more or less to the public good via their tax payments than they impose in terms of costs).

In this paper, we combine these two strands of the literature and propose a model of income tax competition that allows for inter-jurisdictional spillovers and imperfect rivalry of publicly provided goods.[3] We extend Schmidheiny's (2006*b*) income tax model in two ways: First, we allow for various levels of rivalry of the publicly provided good, ranging from a strictly private to a pure public good. Second, we introduce spillovers to other jurisdictions, again allowing for the entire span ranging from a publicly provided good that is exclusively consumed by the jurisdiction that supplies it, to a publicly provided good that is consumed equally in all jurisdictions.

To our knowledge, the only other residence-based tax competition model that allows for truly public goods is by Oddou (2011, 2016). However, his analysis is focused on existence of equilibria and lacks normative and numerical analysis, which are central to our work. Models which combine spillovers and tax competition have been proposed in the context of capital tax competition, which is the prime example of source-based taxation. Source-based taxation is typically analyzed in the context of an immobile population and thus sidesteps many of the issues related to endogenous sorting encountered in models of residence-based tax competition. Bjorvatn & Schjelderup (2002) find that the presence of inter-jurisdictional spillovers reduce the incentives for capital tax competition, and even completely eliminate them in the special case of perfect spillovers.[4]

[2]For reviews, see Wilson (1999), Ross & Yinger (1999), Wilson & Wildasin (2004) and Bruelhart, Bucovetsky & Schmidheiny (2015).

[3]Regional income taxation is an empirically relevant feature in countries that have a strong federalist structure, e.g. in the USA or Switzerland (see e.g. Feld & Kirchgaessner 2001, for an analysis in the latter).

[4]For the special case of fixed capital supply and positive spillovers that are proportional to the capital stock, Ogawa & Wildasin (2009) show that local policy choices are efficient. However, if capital supply is elastic, the decrease in spillovers

In the theoretical part of the paper, we derive a modified Samuelson condition for the provision of public goods that depends on rivalry in consumption and inter-jurisdictional spillovers. We further derive an expression of the "jurisdictional choice externality" introduced by Calabrese, Epple & Romano (2012) in the context of property tax competition, and show that this externality exists even in a first-best environment, unless the publicly provided good is strictly private in nature (as in their case). In the decentralized equilibrium, communities differ according to the (linear) income tax rate, the level of public consumption, and housing prices, all of which are endogenous. Because the model is too complex to solve even with very simple functional forms, we implement it numerically. We restrict our analysis to equilibria that are characterized by income segregation (i.e., households self-select into households according to their income, such that the richest member of a community is poorer than the poorest member of the next-richer community).

We conduct a normative analysis similar in spirit to that carried out by Calabrese et al. (2012) and quantify the welfare losses associated with decentralization, relative to a first-best model where the social planner makes all relevant policy choices and assigns the population across communities.[5] We find that the welfare loss increases as as the level of spillovers increases, and as the good in question becomes less rival in consumption. To understand the underlying sources that drive our results, we split the move from the first-best solution to the fully decentralized equilibrium into four different steps, at each of which we impose an additional constraint on the welfare maximization problem. We thus decompose the overall welfare loss into four different sources of inefficiency: (i) inability to redistribute income across communities; (ii) inability to redistribute income within communities; (iii) inter-community free-riding; and (iv) jurisdictional choice externality. Depending on the nature of the public good, the relative importance of these inefficiencies varies.

---

from other regions is smaller, making it optimal for regional governments to choose tax rates that are lower than what would be socially optimal (Eichner & Runkel 2012). Dembour & Wauthy (2009) show that if firms can choose to move outside the borders of the model (e.g. to another country, where they receive a reservation profit), local governments may have an incentive to over-invest in public goods that help attract firms, with the degree of over-investment being inversely related to the degree of spillovers.

[5]Throughout this paper, we focus on decentralization from a public finance perspective, where the central planner is benevolent and local policy is set directly by the median voter. Because there are no informational asymmetries in our model, decentralization leads to a welfare loss by construction. We abstract from the benefits associated with decentralization due to informational advantages of the local governments and the mitigation of the principal-agent problem that arises if politicians are self-interested.

Adding spillovers and imperfect rivalry into the income tax competition model has two counteracting effects. First, spillovers and economies of scale introduce an inefficiency due to free-riding on the provision of other jurisdictions, which we call inter-community free-riding. This effect increasing as the good becomes more public in nature. At the same time, spillovers help decrease an inefficiency inherited in tax competition itself: With residence-based tax competition, poorer households "chase the rich" to profit from higher public provision levels in their communities without paying the full price. An increase in the level of inter-community spillovers means that households in the poorer communities are better off, and therefore have a lower incentive to crowd into the rich community. As a consequence, the rich community reduces the entrance barriers (e.g. in the form of higher property prices), which increases efficiency. If spillovers are sufficiently large, tax competition breaks down in the sense that the equilibrium is no longer characterized by income segregation. This result is qualitatively similar to those obtained in the context of capital tax competition and of environmental federalism, in the sense that spillovers serve as a means for redistribution from richer to poorer regions and thus mitigate the consequences of tax competition.

In the next section we present our theoretical model. Section 3 introduces specific functional forms and describes second-best decentralized solutions in a numerical implementation of the model. These are contrasted to the corresponding first-best equilibria in section 4, which also contains a quantification of the different inefficiencies. Section 5 concludes.

2 Theoretical model

We start by describing the general structure of the decentralized model and provide the necessary conditions for an equilibrium characterized by self-selection of mobile households into jurisdictions or communities according to income.[6] We then formulate a first-best version of the model, where a social planner determines the distribution of households, the tax rate and thus the public provision in each region, and also has access to an individualized transfer scheme. This allows us to highlight the inefficiencies of the decentralized income tax equilibrium in the presence of spillovers and different

[6] We will refer to communities for the remainder of the paper because we have a metropolitan model in mind where the choice of the residential location is independent of the choice of the work location (usually assumed to be in the central business district); however, the model also applies to larger jurisdictions such as states or cantons, provided that income is independent of the residence.

degrees of rivalry in consumption.

2.1 Equilibrium conditions in second-best

The model consists of $j = 1, \ldots, J$ communities. Each community decides upon the production of a publicly provided good G_j, which is financed using a linear income tax rate t_j. The community-specific tax rate is determined by majority voting of the inhabitants for the respective community (see below). We choose quantities that set the unit cost of public production to one, such that the level of public production in community j is given by

$$G_j = t_j Y_j, \tag{1}$$

where Y_j refers to aggregate income in community j. With spillovers and imperfect rivalry, consumption of the publicly provided good in community j does not coincide with its production. Consumption is determined by

$$g_j = \frac{G_j + \sigma \sum_{i \neq j} G_i}{\left(N_j + \nu \sum_{i \neq j} N_i\right)^{\rho}}, \tag{2}$$

where N_j is the population residing in community j, $\nu \in [0, 1]$ is a parameter that measures the degree to which households from outside community j have access to consume the publicly provided good in j, $\rho \in [0, 1]$ reflects the the degree of rivalry in consumption, and $\sigma \in [0, 1]$ determines the degree of inter-jurisdictional spillovers.[7] The numerator in (2) contains the amount produced in the home community j and in the other communities $i \neq j$, weighted by the spillover parameter σ, which implies that spillovers are symmetric between jurisdictions.[8] Rivalry in consumption enters in

[7]The definitions of production G_j and consumption g_j imply that they have different units: Whereas G_j is equal to aggregate production in j and therefore measured in units of the publicly provided good, consumption g_j describes the amount that every household in j can consume and is therefore measured in units of the publicly provided good per household.

[8]The model could easily be adopted to the case of asymmetric spillovers. Instead of the single parameters ν and σ we would define both as $J \times J$ matrices, with $\nu_{jj} = \sigma_{jj} = 1$. $0 \leq \sigma_{ij} \leq 1 \;\; \forall\, i \neq j$ is then the degree of spillover from community i to j and $0 \leq \nu_{ij} \leq 1 \;\; \forall\, i \neq j$ describes the extent to which households from community i have access to consume the publicly provided good in j . (2) would then read as $g_j = \sum_{i=1}^{J} \sigma_{ij} G_i / (\sum_{i=1}^{J} \nu_{ij} N_i)^{\rho}$. Since we rely on symmetric spillovers in our numerical application, we decided to keep the general model as simple as possible.

the denominator: If $\rho = 0$, there is no rivalry in consumption of the publicly provided good, whereas rivalry is perfect at $\rho = 1$. The term $N_j + \nu \sum_{i \neq j} N_i$ is the mass of households that has access to the publicly provided good in j. The "neighborhood parameter" ν determines to which extent the inhabitants of the other communities ("neighbors") can cross the border and consume the publicly provided good there. It seems natural that for the limiting cases of the spillover parameter (σ close to 0 or 1), ν should be quite close or equal to the value of σ, but we do not restrict our model at this stage.[9] The combination of ($\rho = 0, \sigma = \nu = 1$) describes the case of a pure public good, where consumption equals the sum of what all communities provide. On the other end of the spectrum, the combination ($\rho = 1, \sigma = \nu = 0$) describes a purely private good, of which a household living in community j consumes $g_j = G_j/N_j$. This is the standard assumption for the publicly provided good in existing local tax competition models (e.g. Schmidheiny 2006*b*, Calabrese et al. 2012). Any parameter combination in between describes an intermediate case.[10]

Besides a numeraire and the publicly provided good, households consume housing, which is supplied elastically by absentee landlords, based on a constant returns to scale (CRS) technology. Housing market clearing implies that

$$H_j^S(p_j) - \int_{\underline{y_j}}^{\overline{y_j}} h^j(y) f(y) \mathrm{d}y = 0 \quad \forall j, \tag{3}$$

where H_j^S is the aggregate housing supply function and the integral describes the aggregate of households' demand for housing $h^j(y)$ in community j. We define this as the first of the necessary conditions to constitute a decentralized equilibrium:

Condition 1. *Housing market clearing: Equation* (3) *holds in every community.*

Because the amount of available land is fixed, the marginal cost of housing is increasing in hous-

[9] For the case of perfect spillovers ($\sigma = 1$), it follows that ν should be 1 too, because everybody is affected equally by the production of a unit of the publicly provided good, regardless of where it is produced. Likewise, if spillovers are not present, the population in other communities is presumably irrelevant. For the case of imperfect spillovers, however, access and spillovers do not necessarily have to be the same. In our numerical application, we assume that $\nu = \sigma$.

[10] Note that not all parameter combinations are equally meaningful. For example, it would make no sense to define the publicly provided good as perfectly rival in consumption ($\rho = 1$), but at the same time allow for inter-jurisdictional spillovers ($\sigma > 0$) or neighbors to consume ($\nu > 0$), because any outsiders could be easily excluded from consuming the good at home or by crossing borders. The combination ($\rho = 0, \sigma = \nu = 0$) would describe a club good for which households from outside the community can be excluded.

ing demand. Thus, the attractiveness of the tax-expenditure package of a community is capitalized into housing prices p_j. No other taxes or publicly provided goods exist, such that from the perspective of households, communities are fully characterized by the triplet (p_j, t_j, g_j).[11]

Households differ with respect to exogenous income y, which is distributed on the domain $(\underline{y}, \overline{y})$ according to the probability density function $f(y)$. The utility function $U(x, h, g_j)$ represents household preferences for a numeraire consumption good x, housing h, and the consumption level of the publicly provided good g_j that is associated with residence in community j. Conditional on their residential choice, households maximize their utility by choosing the preferred level of numeraire consumption and housing, subject to a budget constraint that depends on the housing price and the income tax rate. Solving this problem yields an indirect utility function, which depends on community characteristics and personal income:

$$V(p_j, t_j, g_j; y) = \max_{x,h} U(x, h, g_j) \quad \text{s.t.} \quad y(1 - t_j) = x + p_j h. \tag{4}$$

Households choose the community in which they wish to reside. If this self-sorting leads to an equilibrium where the richest member of any given community is poorer than the poorest member of the next richer community, this is referred to as (complete) segregation by income. In the presence of local income tax competition, segregation by income occurs if and only if preferences comply with the conditions identified by Schmidheiny (2002): First, the marginal rates of substitution between any arguments of the indirect utility function must change monotonically in income, which is known as the "single-crossing condition", because it ensures that the indifference curves of two households with different incomes cross only once. Second, preferences have to satisfy a "proportional shift of relative preferences" for income, meaning that they must support a linear expenditure system, and yet they must be non-homothetic.[12]

The migration behavior of households depends only on their income and on how they trade off private and public consumption, which is determined by their utility function. This trade-off does

[11] This is analogous to Schmidheiny (2006*b*), but differs from models of property tax competition where the tax and the housing price are combined, such that communities differ by two dimensions only: the gross-of-tax housing price, and the level of the publicly provided good. Note that G_j is irrelevant for the moving decision of the households, since what matters for utility is the consumption level g_j and not community j's contribution to the publicly provided good.

[12] A class of utility functions that complies with these conditions is the Stone-Geary utility, which we use in section 3.

not change if we introduce varying degrees of spillovers or congestion. Therefore, the necessary conditions identified by Schmidheiny (2002) carry over to the more general description of the publicly provided good used in this paper.

To characterize the migration equilibrium, and without loss of generality, we order communities in ascending order of average income such that under income segregation, every member in community j is (weakly) richer than every member of community $j-1$. For any two adjacent communities j and $j-1$, there exists a household with income $\widetilde{y}_{j-1,j}$ that is indifferent between them. Formally, this means that $V(p_j, t_j, g_j; \widetilde{y}_{j-1,j}) = V(p_{j-1}, t_{j-1}, g_{j-1}; \widetilde{y}_{j-1,j})$. The indifferent household is thus determined by the characteristics of both communities: $\widetilde{y}_{j-1,j} = \widetilde{y}(p_j, t_j, g_j, p_{j-1}, t_{j-1}, g_{j-1})$. A migration equilibrium under income segregation requires that the indifferent households are the community "borders", i.e. that $\widetilde{y}_{j-1,j} = \underline{y_j}$ and $\widetilde{y}_{j,j+1} = \overline{y_j}$, where the lower and upper bar represents the poorest and the richest household in community j, respectively. A sorting equilibrium has been reached if nobody has an incentive to move and if the only indifferent households are at the communities borders. This implies that

$$\forall\, y \in [\underline{y_j}, \overline{y_j}]: \quad V(p_j, t_j, g_j; y) - V(p_i, t_i, g_i; y) \geq 0 \;\; \forall\, i \neq j \;\forall\, j, \tag{5}$$

where the above holds with equality only for border households between two adjacent communities.[13] This leads to the next equilibrium condition:

Condition 2. *Migration equilibrium: The indifferent households between adjacent communities are at the actual "border" of every community, i.e.* (5) *holds with equality only for* $\underline{y_j}$ *and* $\overline{y_j}$. *All households in between strictly prefer community* j *to any other community.*

The sorting and the housing market equilibrium conditions have the same structure as in standard residence-based tax competition models. The equilibrium condition for majority voting on the tax rate, however, changes with the introduction of spillovers. The reason is that voters consider the amount of the publicly provided good supplied by other communities if $\sigma > 0$. They thus free-ride

[13] The empirically more realistic case of "imperfect sorting" of households requires that households differ with respect to income and additionally some preference parameter, as in Epple & Platt (1998). The model would then require that (5) holds for all values of the second source of heterogeneity, i.e. that there is income segregation conditional on the value of the preference parameter.

on public provision elsewhere.[14]

We formalize this situation as a two-stage game (see Fernandez & Rogerson 1996). In the first stage, households choose their location and buy housing, correctly anticipating the voting outcome in every community. In the second stage, households determine the level of public production along with the necessary tax rate to satisfy the budget constraint by majority vote, and consumption takes place. At this stage, we assume that households perfectly foresee the other communities' contributions to the publicly provided good.[15] When voting, the household distribution and therefore aggregate income and population (defined by $Y_j = \int_{\underline{y_j}}^{\overline{y_j}} y f(y) \mathrm{d}y$ and $N_j = N \int_{\underline{y_j}}^{\overline{y_j}} f(y) \mathrm{d}y$, respectively) is taken as given. The maximization problem for the voter with income y is

$$\max_{t_j, g_j} V(p_j, t_j, g_j; y) \quad \text{s.t.} \quad g_j = \frac{t_j Y_j + \sigma \sum_{i \neq j} G_i}{\left(N_j + \nu \sum_{i \neq j} N_i\right)^{\rho}}. \tag{6}$$

The median voter y_j^m in community j is implicitly defined by $\int_{\underline{y_j}}^{y_j^m} f(y) \mathrm{d}y = \frac{1}{2} \int_{\underline{y_j}}^{\overline{y_j}} f(y) \mathrm{d}y$. Due to the monotonicity of preferences, the median income households are the only households that maximize utility. All others do not get their optimal tax rate, although they (weakly) prefer the triple (p_j, t_j, g_j) in their community to that in other communities.

Substituting the community's budget constraint, the first-order condition with respect to t_j is

$$\frac{\partial V_j}{\partial t_j} + \frac{\partial V_j}{\partial g_j} \frac{Y_j}{\left(N_j + \nu \sum_{i \neq j} N_i\right)^{\rho}} = 0, \tag{7}$$

[14]Note that even for the case without spillovers, the Samuelson-condition will (generally) not be satisfied since the publicly provided good supply follows from utility maximization of the median rather than the average household in terms of income. Whether this leads to under- or over-provision depends on the distribution of preferences. Calabrese et al. (2012) argue that this source of inefficiency is likely to be small.

[15]There is an ongoing debate in the literature concerning voter myopia. Some authors (among them Schmidheiny 2006*b*, Calabrese, Epple, Romer & Sieg 2006) call such a behavior "myopic", since, as the argument goes, voters do not consider migration responses and housing demand reactions to a change in the communities tax rate. However, since in equilibrium nobody has an incentive to move or to change housing consumption by definition, this implies that even when voters are labeled "myopic" they in fact perfectly foresee equilibrium household distribution and housing prices. Outside an equilibrium their behavior is different, however. This implies that we look at *different* equilibria, whenever we make different assumptions about voter sophistication, even though the structure of the equilibrium conditions remains the same. See Epple, Romer & Sieg (2001) for an empirical analysis of the subject.

which is solved by $t_j = t(p_j, Y_j, N_j, y_j^m, \{N_i\}_{i \neq j}, \{G_i\}_{i \neq j})$ for every community j. This implicitly defines the tax rate for every community for given contributions to the publicly provided good of the other communities. A voting outcome in our local income tax competition model constitutes a Nash equilibrium such that, for every j, $G_j = t_j Y_j$ is a best reply to the corresponding provisions $G_i \ \forall\, i \neq j$. This allows us to formulate the last of the necessary conditions:

Condition 3. *Majority voting equilibrium: The tax rate in every community is chosen such that equation* (7) *holds, and the corresponding provisions of the publicly provided good are corresponding best replies.*

Conditions 1 to 3 identify the set of necessary equilibrium conditions as a system of $3J - 1$ equations (note that there are only $J - 1$ indifferent households) and $3J - 1$ unknowns (housing prices, public provision levels and border households). Neither existence nor uniqueness can be established, as is usually the case in tax competition problems. Without further assumptions on preferences it is not possible to analytically determine the effect of spillovers and imperfect rivalry on equilibrium outcomes. For the case of Stone-Geary preferences, however, we solve the model for various combinations of σ and ρ numerically in section 3 and quantify the inefficiencies in section 4.

2.2 Efficiency conditions in first-best

In this subsection we address the efficiency consequences of decentralization for the case of an (im-) perfectly congested publicly provided good that may have spillovers to other communities. To do this, we need to formulate a first-best version of the model of the preceding section.

The social planner maximizes a social welfare function by assigning the population across communities and choosing the efficient level of public provision and taxation. Furthermore, by allowing income-dependent transfers $r(y)$ in addition to uniform head taxes in every community T_j, the social planner effectively has access to individualized lump-sum taxation, which renders the problem first-best. Successively more constrained versions of the model can be derived by denying the planner access to transfers or head taxes, by allowing households to choose their preferred community, or by majority vote on the level of public provision. Such second-best versions are considered in section 4.3 when decomposing the inefficiency from full decentralization in its individual components.[16]

[16]Because we do not allow for preference heterogeneity and take the view of a social planner, decentralization neces-

The indirect utility function of a household with exogenous income y that lives in community j is a generalized version of (4) and is given by

$$V(p_j, t_j, T_j, g_j; y, r(y)) = \max_{x,h} U(x, h, g_j) \quad \text{s.t.} \quad y(1 - t_j) - T_j + r(y) = x + p_j h, \tag{8}$$

where $y(1 - t_j) - T_j + r(y)$ is the after-tax budget of the household. For brevity, we define $V^j(y) \equiv V(p_j, t_j, T_j, g_j; y, r(y))$ where the superscript j refers to the set of policy variables in j. Solving the above problem leads to demand functions for housing and numeraire consumption, which we label by $h^j(y)$ and $x^j(y)$, respectively. For comparison purposes with constrained versions of the model, we include a linear income tax t_j in the first-best model, even though it is clear that it will not be used when individualized lump-sum taxation is available. Borrowing notation from Calabrese et al. (2012), we write the first-best problem as

$$\max_{a_j(y), r(y), R, t_j, T_j, p_j, g_j} \quad \sum_{j=1}^{J} \left\{ \int_{\underline{y}}^{\overline{y}} \omega(y) V^j(y)\, a_j(y) f(y) \mathrm{d}y + \omega_R \left(\frac{R}{J} + \int_0^{p_j} H^S(z) \mathrm{d}z \right) \right\} \tag{9}$$

$$\text{s.t.} \quad g_j = \frac{t_j Y_j + T_j N_j + \sigma \sum_{i \neq j} (t_i Y_i + T_i N_i)}{\left(N_j + \nu \sum_{i \neq j} N_i \right)^{\rho}} \quad \forall\, j \tag{10}$$

$$H^S(p_j) = \int_{\underline{y}}^{\overline{y}} h^j(y)\, a_j(y) f(y) \mathrm{d}y \quad \forall\, j \tag{11}$$

$$R + \int_{\underline{y}}^{\overline{y}} r(y) f(y) \mathrm{d}y = 0, \tag{12}$$

where $\omega(y)$ is the welfare weight placed on households with income y, and ω_R is the welfare weight that relates the monetary transfer R to the absentee landlords (which may be negative) and their economic rent into social welfare. The share of the population with income y that resides in community j is given by $a_j(y) \in [0, 1]$, with $\sum_j a_j(y) = 1$. For all income segregating equilibria, $a_j(y) = 1$ if $y \in [\underline{y_j}, \overline{y_j}]$ and $a_{i \neq j}(y) = 0$.

The budget constraint of the public sector includes head taxes and is given by (10). Constraint (11) ensures housing market clearing and (12) mandates that the sum of transfer payments is zero, the purpose of which is to allow for a clean separation between taxes used to finance the publicly

sarily leads to a welfare loss. The social planner could always duplicate the decentralized solution, but is free to choose an equilibrium that achieves a higher social welfare.

provided good, and revenue-neutral redistribution. To derive the efficiency conditions, we assign the following Lagrange multipliers: λ_j for (10), η_j for (11), and Ω for (12).

The derivative of the Lagrangian with respect to $a_j(y)f(y)$ reflects the marginal social value $MSV_j(y)$ of adding a unit measure of the population with income y to community j, and is given by

$$
\begin{aligned}
MSV_j(y) =& \omega(y)V^j(y) \\
&+ \Big(\lambda_j + \sigma \sum_{i \neq j} \lambda_i\Big)\Big[t_j y + T_j\Big] \\
&- \rho \Big\{ \lambda_j g_j \Big(N_j + \nu \sum_{i\neq j} N_i\Big)^{\rho-1} + \sigma \sum_{i \neq j} \lambda_i g_i \Big(N_i + \nu \sum_{k \neq i} N_k\Big)^{\rho-1} \Big\} \\
&+ \eta_j h^j(y).
\end{aligned} \tag{13}
$$

For income-segregating equilibria, on which we focus in this paper, the social planner chooses the border household $y_{i,j}^{border}$ between any two adjacent communities i and j such that $MSV_i(y_{i,j}^{border}) = MSV_j(y_{i,j}^{border})$. All poorer (richer) households have a strictly higher marginal social welfare in the poorer (richer) community when compared to the other. Note that in this case the first-order condition (13) is not zero because $a_j(y)$ is either zero or one, such that an interior solution does not exist.

Suppressing the argument in $a_j(y)$ and denoting partial derivatives by subscripts, the FONCs with respect to $r(y)f(y)$ and R can be combined to

$$
\sum_{j=1}^{J} \omega(y)V_y^j(y)a_j + \sum_{j=1}^{J} \eta_j h_y^j(y) a_j = \omega_R \tag{14}
$$

and the remaining FONCs are

$$
L_{T_j} = 0 = \int_{\underline{y}}^{\overline{y}} \omega(y)V_T^j(y)a_j f(y)\mathrm{d}y + N_j\Big(\lambda_j + \sigma\sum_{i\neq j}\lambda_i\Big) + \eta_j \int_{\underline{y}}^{\overline{y}} h_T^j(y)a_j f(y)\mathrm{d}y, \tag{15}
$$

$$
L_{t_j} = 0 = \int_{\underline{y}}^{\overline{y}} \omega(y)V_t^j(y)a_j f(y)\mathrm{d}y + Y_j\Big(\lambda_j + \sigma\sum_{i\neq j}\lambda_i\Big) + \eta_j \int_{\underline{y}}^{\overline{y}} h_t^j(y)a_j f(y)\mathrm{d}y, \tag{16}
$$

$$
L_{p_j} = 0 = \int_{\underline{y}}^{\overline{y}} \omega(y)V_p^j(y)a_j f(y)\mathrm{d}y + \omega_R H^S(p_j) + \eta_j \Big[\int_{\underline{y}}^{\overline{y}} h_p^j(y)a_j f(y)\mathrm{d}y - H_p^S(p_j)\Big], \tag{17}
$$

$$
L_{g_j} = 0 = \int_{\underline{y}}^{\overline{y}} \omega(y)V_g^j(y)a_j f(y)\mathrm{d}y - \lambda_j\Big(N_j + \nu\sum_{i\neq j}N_i\Big)^{\rho} + \eta_j \int_{\underline{y}}^{\overline{y}} h_g^j(y)a_j f(y)\mathrm{d}y. \tag{18}
$$

Together with the constraints (10)–(12), conditions (13)–(18) determine the social optimum, conditional on the set of choice variables available to the social planner. For example, if the planner does not have access to individualized transfers, (14) drops out; if the income tax system is further restricted to be linear such that $T_j = 0$, then (15) is removed as well. On the other hand, with individualized transfers and head taxes, $t_j = 0$ and (16) drops out.

We can now discuss the efficiency loss from decentralization in more detail. There are three sources of inefficiency: First, not having access to lump-sum redistribution but being restricted to linear income taxation leads to inefficient redistribution. Second, free household mobility leads to a jurisdictional choice externality (JCE) due to an inefficient distribution of households; we will refer to this as *intra-community free-riding*, as households seek to consume more of the public good than what they contribute in the form of income taxes. Finally, we find underprovision of the publicly provided good due to decentralized choice of G_i. This is the classical free-riding on the other communities' provision (i.e. *inter-community free-riding*). We now explore the properties of the last two inefficiencies in more detail.

To obtain some intuition about (13), we can divide it into a private and a social component. The private component increases welfare by $\omega(y)V^j(y)$. The social component consists of the remaining three terms, and in keeping with the language of Calabrese et al. (2012) we refer to this as the jurisdictional choice externality (JCE). It describes the external cost or benefit that the locational choice by a household with income y conveys on the other households, and which the household does not consider when making the locational decision.

The first term of the JCE is the social value of the tax payment by this type of household, which accrues not only to community j but also to all other communities due to the spillovers of the publicly provided good. The second term reflects the aggregate congestion costs arising in all communities; with zero congestion, this term drops out. The last term measures the social cost or benefit of adding another household to the housing market. As we show below, in the presence of individualized lump-sum taxation, $\eta_j = 0$ and this term drops out. In a more realistic, second-best environment, η_j can be positive or negative. To gain intuition about the nature of the externality associated with the housing market, we solve (17) for η_j:

$$\eta_j = \frac{\int_{\underline{y}}^{\overline{y}} \omega(y) V_p^j(y) a_j f(y) \mathrm{d}y + \omega_R H^S(p_j)}{H_p^S(p_j) - \int_y h_p^j(y) a_j f(y) dy}. \tag{19}$$

Under regular assumptions about demand and supply of housing, the denominator is positive, whereas the two terms in the numerator have opposite signs. If the welfare weight that is placed on the absentee landlords is sufficiently small, it follows that $\eta_j < 0$, and vice versa. Intuitively, if the landlords do not count much (or not at all) for social welfare, adding population to community j increases the housing price for all residents, which is only partially offset by social benefits in the form of higher economic rents.

For the special case of no spillovers ($\sigma = 0$), no access of neighbors ($\nu = 0$), and full congestion ($\rho = 1$), the JCE associated with adding a household of type y to community j is[17]

$$JCE_j(y) = \lambda_j t_j \big(y - Y_j/N_j\big) + \eta_j h^j(y). \tag{20}$$

The first term measures the fiscal externality of the household on all other households in community j. For households that are richer than the average, this fiscal externality is positive, because they contribute more to the publicly provided good than what they receive in terms of public consumption, and vice versa.

Proposition 1. *Jurisdictional choice externality (JCE) in first-best*

(i) In a first-best environment where the social planner has access to individualized transfers, chooses the level of local public provision and assigns the population across communities, there is no inefficiency associated with the housing market such that $\eta_j = 0\ \forall j$.

Proof. See Appendix. □

(ii) For the special case of no spillovers ($\sigma = 0$), no access of neighboring communities to domestic public consumption ($\nu = 0$), and full congestion ($\rho = 1$), the JCE is zero in first-best.

Proof. With access to individualized transfers, the social planner does not employ an income tax such that $t_j = 0$. Together with *(i)*, it follows immediately that (20) is zero. □

The assumption of a publicly provided private good is rather special case (although one that has been the focus of the tax competition literature). Whenever $\sigma > 0$, some of the fiscal effect of adding a household to community j spills over to other communities (i.e., the terms in (13) that contain σ). And for any $\rho < 1$ there are economies of scale which affect the consumption level of the publicly

[17] To derive this expression, substitute $\rho = 1, \sigma = \nu = 0$ and $g_j = (t_j Y_j + N_j T_j)/N_j$ into (13).

provided good for all households within the community. Furthermore, if individualized lump-sum taxation is not available, there will generally be a nonzero JCE via the housing market because $\eta_j \neq 0$. Finally, if revenue is raised via an income tax, the JCE also contains a fiscal externality in the sense that households pay more or less in taxes than they receive in the form of public provision. This means that there generally will be a nonzero JCE even in first-best.

Next, we turn to the efficiency condition for public provision, which can be simplified according to the following proposition:

Proposition 2. *Samuelson condition:*

(i) The generalized Samuelson condition that allows for varying degrees of spillovers across communities and rivalry in consumption can be derived as

$$\text{(21)} \quad \int_{\underline{y}}^{\overline{y}} \omega(y) V_g^j(y) a_j f(y) dy = \frac{\omega_R \left(N_j + \nu \sum_{i \neq j} N_i\right)^{\rho}}{\left(1 + \sigma(J-1)\right)(1-\sigma)} \left(\frac{1 + \sigma(J-2)}{1 - t_j} - \sigma \sum_{i \neq j} \frac{1}{1 - t_i} \right) - \eta_j \int_{\underline{y}}^{\overline{y}} h_g^j(y) a_j f(y) dy.$$

(ii) In first-best with $t_j = \eta_j = 0$, this condition simplifies to

$$\text{(22)} \quad \int_{\underline{y}}^{\overline{y}} \frac{V_g^j(y)}{V_y^j(y)} a_j f(y) dy = \frac{\left(N_j + \nu \sum_{i \neq j} N_i\right)^{\rho}}{\left(1 + \sigma(J-1)\right) N_j}.$$

(iii) In first-best, the marginal value of public funds is equalized across communities:

$$\text{(23)} \quad \lambda_j = \lambda = \frac{\omega_R}{1 + \sigma(J-1)} \quad \forall\, j.$$

Proof. See Appendix. □

Condition (21) states that the aggregate social marginal utility from an additional unit of g_j (left-hand side) needs to be equal to the sum of total fiscal costs (conditional on the degree of spillovers and congestion) and the housing market externality.

In first-best, i.e. with $t_j = \eta_j = 0$, this condition can be simplified and becomes more intuitive. The LHS of (22) is the aggregate marginal rate of substitution between public consumption and the numeraire, and the RHS is the marginal rate of transformation between g_j and the numeraire, adjusted for spillovers and congestion. In contrast, the median voter in second-best determines the amount of public provision from (7) such that $\frac{V_S^j(y_j^m)}{V_y^j(y_j^m)} = \frac{1-t_j}{y_j^m}\frac{\left(N_j+\nu\sum_{i\neq j}N_i\right)^\rho}{Y_j}$ and thus ignores the willingness to pay of the other households. For the special case of a publicly provided private good (i.e. if $\sigma = \nu = 0$ and $\rho = 1$), the Samuelson-condition simplifies to the familiar form of

$$\int_{\underline{y}}^{\overline{y}} \frac{V_g^j(y)}{V_y^j(y)} a_j f(y)\mathrm{d}y = 1, \tag{24}$$

where 1 equals the marginal rate of transformation of the publicly provided good and the numeraire.

The shadow price λ_j describes the marginal value of investing in the publicly provided good. With concave preferences, λ_j has to decrease in g_j. Condition (23) therefore implies that in first-best, the marginal value of public funds is equalized among communities, and that the greater the inter-community spillovers, the more resources are optimally used for the production of the publicly provided good (such that λ becomes smaller). This is intuitive, since spillovers increase the benefit of public provision without increasing its cost, leading to a higher optimal level of public provision, all else equal. Note also that in the absence of spillovers, it follows that $\lambda = \omega_R$.

3 Numerical implementation

In this section we introduce specific functional forms, which we use to describe decentralized second-best equilibria for varying degrees of spillovers and rivalry in a numeric application.

3.1 Functional forms

To comply with the necessary conditions for income segregation, preferences must be non-homothetic and support a linear expenditure system, which is the case for the Stone-Geary utility function (Stone 1954, Geary 1950). Specifically, we define preferences by

$$U_j(x,h,g_j) = \alpha \cdot \ln(g_j - \beta_g) + \gamma \cdot \ln(h - \beta_h) + (1-\alpha-\gamma)\cdot \ln(x - \beta_x), \tag{25}$$

where β_g, β_h and β_x are subsistence levels and α, γ and $1-\alpha-\gamma$ (all of which are between zero and one) are preference parameters for g_j, h and x, respectively. Solving the consumer's problem leads to a demand function for housing and the numeraire good of $h^j(y) = \gamma\frac{y(1-t_j)+r(y)-T_j-p_j\beta_h-\beta_x}{(1-\alpha)p_j} + \beta_h$ and

$x^j(y) = y(1-t_j) + r(y) - T_j - p_j h^j(y)$, respectively, where $r(y)$ is the individual lump-sum transfer for household y and T_j is the community specific lump-sum head tax to finance public provision. Both instruments are not available in second-best. The indirect utility function follows as

$$\begin{aligned} V(p_j, T_j, t_j, g_j; y) =& (1-\alpha) \cdot \ln[y(1-t_j) + r(y) - T_j - p_j\beta_h - \beta_x] \\ & + \alpha \cdot \ln(g_j - \beta_g) - \gamma \cdot \ln(p_j) + c, \end{aligned} \tag{26}$$

with $c \equiv (1-\alpha-\gamma) \cdot \ln\left(\frac{1-\alpha-\gamma}{1-\alpha}\right) + \gamma \cdot \ln\left(\frac{\gamma}{1-\alpha}\right)$.

The housing market clearing condition (3) is given by

$$L_j p_j^\theta - \int_{\underline{y_j}}^{\overline{y_j}} \left(\gamma \frac{y(1-t_j) + r(y) - T_j - p_j\beta_h - \beta_x}{(1-\alpha)p_j} + \beta_h\right) f(y)\mathrm{d}y = 0, \tag{27}$$

where $L_j p_j^\theta$ is the housing supply function, L_j is the relative land size of community j (that is, we normalize aggregate land to 1) and θ is the price elasticity of housing supply.

The set of first-best efficiency conditions for the functional forms described above is summarized in Table A1 in the Appendix. In the fully decentralized second-best equilibrium, the indifferent household between two adjacent communities j and $j-1$ is defined according to the migration condition (5) as

$$\widetilde{y}_{j-1,j} = \frac{\left[(p_j\beta_h + \beta_x) p_{j-1}^{\frac{\gamma}{1-\alpha}} (g_j - \beta_g)^{\frac{\alpha}{1-\alpha}} - (p_{j-1}\beta_h + \beta_x) p_j^{\frac{\gamma}{1-\alpha}} (g_{j-1} - \beta_g)^{\frac{\alpha}{1-\alpha}}\right]}{\left[(1-t_j) p_{j-1}^{\frac{\gamma}{1-\alpha}} (g_j - \beta_g)^{\frac{\alpha}{1-\alpha}} - (1-t_{j-1}) p_j^{\frac{\gamma}{1-\alpha}} (g_{j-1} - \beta_g)^{\frac{\alpha}{1-\alpha}}\right]}, \tag{28}$$

where we already considered that $r(y) = T_j = 0$.

For the median income household in community j, y_j^m, the first-order condition (7) derived from the voter's maximization problem (6) becomes[18]

$$\frac{(1-\alpha)y_j^m}{y_j^m(1-t_j) - p_j\beta_h - \beta_x} = \frac{\alpha Y_j}{t_j Y_j + \sigma \sum_{i\neq j} G_i - \left(N_j + \nu \sum_{i\neq j} N_i\right)^\rho \beta_g}. \tag{29}$$

Local production of the public good is given by $G_j = t_j Y_j$. Solving (29) for t_j and multiplying

[18] It is easy to show that the second order condition is negative, indicating that an equilibrium indeed corresponds to utility maximum, or in the case of multiple equilibria, to local utility maxima.

with Y_j leads to an optimal contribution to the publicly provided good of

$$G_j = \max \left\{ \left(N_j + \nu \sum_{i \neq j} N_i \right)^{\rho} \beta_g + \alpha Y_j B_j - (1-\alpha)\sigma \sum_{i \neq j} G_i,\ 0 \right\} \tag{30}$$

The first term in (30) is the subsistence level of the publicly provided good β_g, multiplied by the population with access to the domestic market ($N_j + \nu \sum_{i \neq j} N_i$) and corrected for congestion (ρ). The expression $B_j \equiv 1 - \frac{p_j \beta_h + \beta_x}{y_j^m} - \frac{\left(N_j + \nu \sum_{i \neq j} N_i\right)^{\rho} \beta_g}{Y_j}$ is the share of income that the median voter can spend on (x, h, g_j) after paying for subsistence levels, and can be interpreted as measure of disposable income. The second term therefore consists of the share α of this disposable income, multiplied by the tax base Y_j. This is the part of public provision chosen "freely" by the median voter, i.e. the part that exceeds the subsistence level. Finally, G_j is reduced by what spills over from the other communities. In the absence of spillovers, this last component drops out and the optimal contribution to the publicly provided good is independent of the supply in other communities.[19]

The contribution to the publicly provided good increases in the preference parameter: The higher is α, the more is spent on the publicly provided good and the smaller becomes the influence of the other communities. If $\alpha = 1$, the last term drops out and the community will spend the entire disposable income on local public provision G_j. In contrast, if α is small, spillovers are sufficiently large and the aggregate income of the community is sufficiently low, the median voter would in fact like to choose a negative public provision (corresponding to a negative income tax rate); because this is infeasible, the resulting G_j is zero. In this case, the community completely free-rides on the public provision of the other communities.

Note that the optimal contribution to the publicly provided good varies in income only due to the relative size of private subsistence levels (the $\frac{p_j \beta_h + \beta_x}{y_j^m}$ part in B_j). This is the reason why income tax competition models require non-homothetic preferences to comply with income sorting of households; otherwise, all households would choose the same tax rate and no sorting by income would occur, regardless of the degree of spillovers. Thus, even with perfect spillovers and (and thus only two dimensions according to which communities differ), the conditions identified in Schmidheiny (2002) must hold for sorting by income to take place.

[19]Even in the case without spillovers, the supply of the publicly provided good depends indirectly on the other communities' choices, which affect the population distribution and therefore determine N_j and Y_j along with the housing price p_j and the median voter y_j^m.

A decentralized equilibrium requires that housing markets clear (J housing prices solve J equations (27)), that indifferent households are at the communities' borders ($J-1$ indifferent housholds are determined by $J-1$ equations (28)) and that public provision complies to a majority voting equilibrium (J tax rates solve J equations (30)). Note that even for the most simple setup, this set of equilibrium conditions could not be solved analytically.

3.2 Second-best: Decentralized equilibria

In this section we provide quantitative results from a numerical implementation of the decentralized equilibrium for varying values of spillovers and congestion. There are $J = 2$ communities, and household income is uniformly distributed between $\underline{y} = 1$ and $\overline{y} = 2$. We assume that households have access to the other communities' public provision to the same degree as there are spillovers, i.e. we set $\nu = \sigma$. Aggregate land is normalized to 1 and both communities are of equal size. Without loss of generality, we assume that community 1 is inhabited by the poor and community 2 by the rich households. Thus, $\widetilde{y}_{1,2} = \overline{y_1} = \underline{y_2}$ is the indifferent household.

Table 1: Parameter values for the baseline calibration

Parameter	α	β_x	β_h	β_g	γ	ν	θ	L_1	$\underline{y}$	$\overline{y}$
Value	0.2	0.2	0.2	0	0.1	σ	3	0.5	1	2

Definition of parameters: α: preference for publicly provided good; β_x: subsistence level (SL) of the numeraire; β_h: SL of housing; β_g: SL of the publicly provided good; γ: housing preference; ν: neighborhood parameter (access to public provision in other community), set to equal σ; θ: price elasticity of housing supply; L_1: (relative) land size of community 1 ($L_2 = 1 - L_1$); $\underline{y}$: lower bound of income distribution; $\overline{y}$: upper bound of income distribution.

The values of the remaining parameters are given in Table 1. The parameters imply that subsistence levels account for 40% of the poorest household's income if the housing price were 1 (the price of the numeraire good is unity by construction). We chose moderate levels of the relative preference parameters for the publicly provided good and housing ($\alpha = 0.2$ and $\gamma = 0.1$, respectively), which implies that households optimally spend most of their disposable income on the numeraire good. $\theta = 3$ is a common value of the housing price elasticity (see e.g. Schmidheiny 2006*b*).

Figure 1 shows the indifferent border household $\widetilde{y}_{1,2}$ as a function of σ and ρ. With a uniform distribution of income, the population share is given by the difference between $\widetilde{y}_{1,2}$ and $\underline{y} = 1$ and $\overline{y} = 2$ for community 1 and 2, respectively. Figure 2 shows the resulting housing prices, tax rates, public provision and public consumption. Equilibrium values of selected variables for specific values of σ and ρ are given in Table A2 in the Appendix.

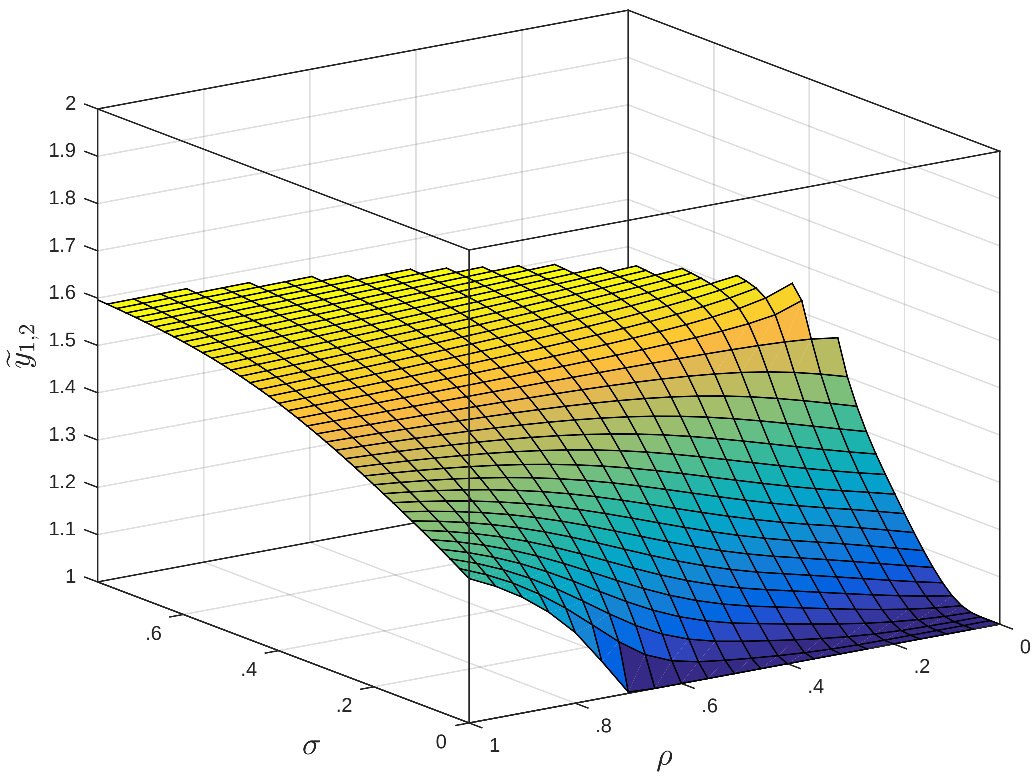

Figure 1: Border household in the decentralized equilibrium, for various combinations of spillovers (σ) and congestion (ρ).

The combination of $\sigma = 0$ and $\rho = 1$ corresponds to the publicly provided private good, which is the routine assumption in existing models of residence-based tax competition. For this special case, the poorest 30% of the population live in community 1, whereas the richest 70% of the population reside in community 2. The rich community exhibits higher public consumption (0.259 vs. 0.167) in exchange for higher housing prices (0.796 vs. 0.593) and higher income taxes (15.7% vs. 14.5%).[20]

In general, decreasing the degree of congestion leads to a more uneven distribution of the population among communities due to economies of scale. Intuitively, with a low degree of rivalry, the increasing returns to scale from public provision dominate the moving decision, since a larger tax base translates to a higher consumption level for everybody.[21]

[20]For fixed public provision and progressive cantonal tax functions, Schmidheiny (2006*a*) shows that richer households self-select into Swiss communities with lower tax multipliers and higher property prices. Thus our model does not fully capture empirically observed patterns of income segregation. The reason is that we consider equally sized communities, a uniform distribution of income, and linear tax rates.

[21]Consider combinations of $\sigma = 0$ and $\rho < 1$, which corresponds to the surface in front and to the right in Figure 1.

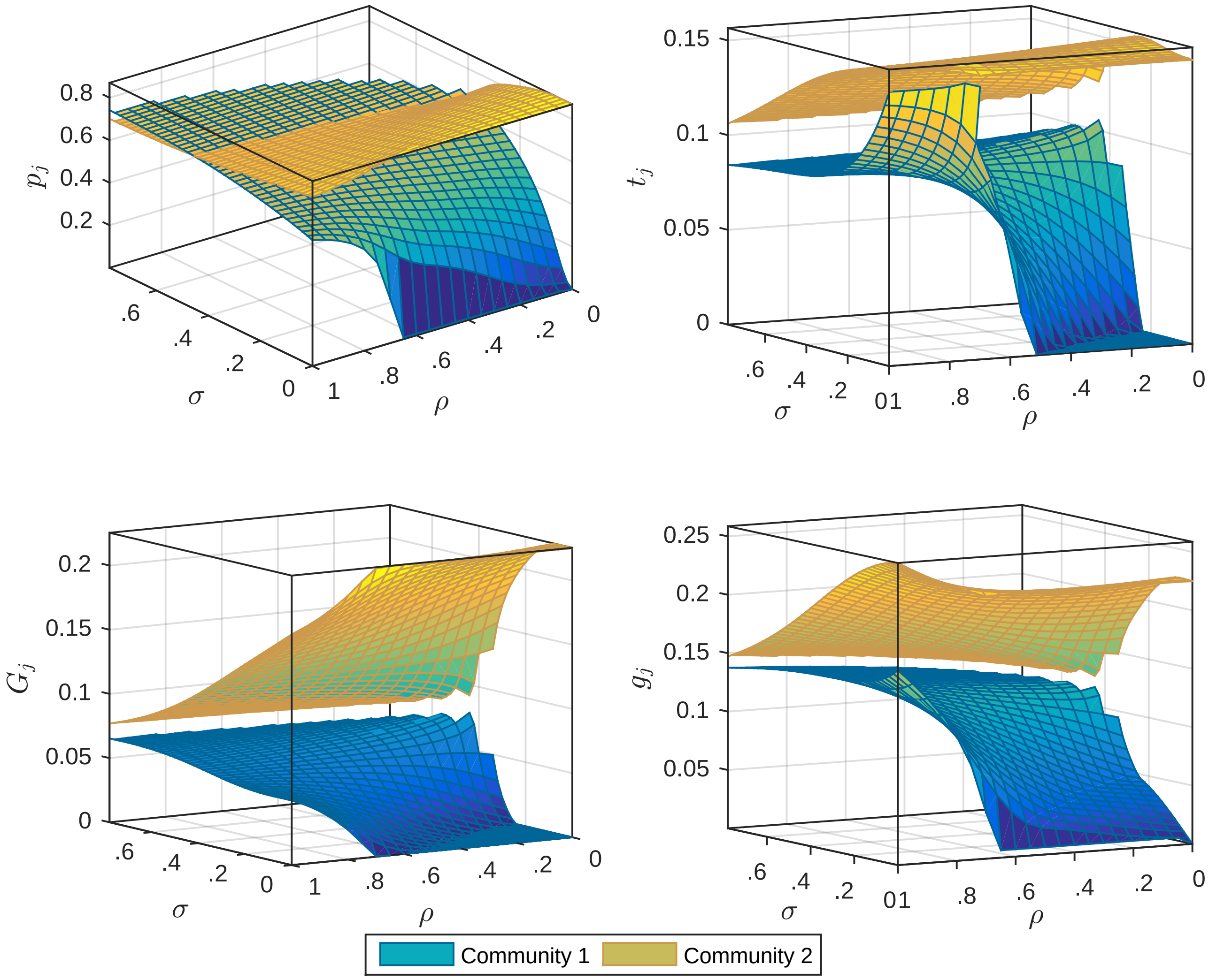

Figure 2: Housing price p_j, tax rate t_j, production G_j, and consumption g_j of the publicly provided good in the decentralized equilibrium, for various combinations of spillovers (σ) and rivalry in consumption (ρ).

An increase in the spillover parameter σ results in a more equal distribution of community income. Spillovers therefore reduce (rich) households' incentive to segregate according to income. Figure 2 shows that both production and consumption levels of the publicly provided goods are converging as spillovers increase (bottom two panels), but at a lower level of consumption for the rich community 2 than in the absence of spillovers. This is the free-riding effect: As spillovers increase, the incentive to provide the publicly provided good decrease.

If rivalry in consumption is low enough (in our case around 0.65), households choose to reside in only one community, leaving the other empty. These "empty community cases" are equilibria that comply with sorting, though with a rather unusual form of income segregation. With sufficiently low spillovers, nobody wants to live in the community with the small tax base. Naturally, the empty community cases rely on a sufficiently low price elasticity of housing supply and for sufficiently high values of θ, the negative effect of higher housing prices outweighs the positive effect of a larger tax base in the populated community and for all combinations of σ and ρ both communities will be inhabited.

As shown in the top left panel in Figure 2, housing prices are higher in the rich community for spillover levels of up to 0.5, whereas for larger levels of spillovers, prices are higher in the poor community. Income tax rates (top right panel), however, are higher in the rich community for all combinations of spillovers and congestion. There is a range of parameter combinations for which the tax rate in the poor community is zero, but the community is not empty. In these cases, the poor community completely free-rides on the public provision of the rich community.

We find no stable equilibria if spillovers are high and/or congestion is low enough, and we were not able to identify income separating equilibria for $\sigma \geq 0.8$ independent of the value of ρ. High spillovers and low levels of congestion reduce the degree of tax competition in the sense that the income differences between communities become less pronounced. Once spillovers are large enough, the difference in income between the communities vanishes, and we no longer find an equilibrium that is characterized by income segregation. This does not necessarily mean that an equilibrium does not exist; however, any such equilibrium would not be characterized by income segregation.

These results show that spillovers and congestion matter in the context of decentralized provision and financing of public goods, and that focusing on the special case of publicly provided private goods provides only a very limited picture of the implications of tax competition for the provision and financing of public goods.

4 Welfare analysis

In this section we focus on the welfare implications of decentralization for different degrees of spillovers and congestion of the publicly provided good. Because we abstract from preference heterogeneity and political economy concerns, decentralization leads to a welfare loss by construction in our model. The question we seek to address is how the overall welfare loss and its individual components depend on the degree of spillovers and congestion.

4.1 Efficiency benchmark

Our efficiency benchmark is based on the theoretical model as described in section 2.2. The planner maximizes a utilitarian social welfare function (SWF) with welfare weights for all households set to unity. We restrict our attention to income-segregating equilibria.[22] Our approach differs from Calabrese et al. (2012), who choose the weights of their welfare function implicitly such that the individual transfers are optimally zero.[23]

To simplify matters, we assume that the social planner cannot enforce transfers from or to the landlords, which implies that $R = 0$. We fix $\omega_R = 0.444$ such that in the "semi-public" good case of $\sigma = 0.3$ and $\rho = 0.5$ (see below) it is optimal from the planner's perspective to choose no transfer from or to the landlords.[24] First-best results for various combinations of σ and ρ are displayed in Table A3 in the Appendix.

The objective function is (in general) not single-peaked. To find the global maximum, we therefore use a grid search to determine the border household instead of letting the algorithm chose it. Our analysis reveals that for almost all combinations of σ and ρ there are indeed three saddle points that comply with the first-order conditions derived in section 2.2: A welfare minimum, one local welfare maximum and one global maximum. The shape of the SWF for varying levels of y^{border}, σ and ρ is displayed in Figures A1 and A2 in the Appendix.

4.2 Welfare loss from decentralization

Figure 3 shows the difference between the value of the SWF in first-best and in the fully decentralized equilibrium, for different combinations of σ and ρ. The welfare loss from decentralization increases with the degree of spillovers and decreases with rivalry in consumption. The underlying reason is that making the good more public in nature leads to a greater scope for inefficiency from decentralization, due to free-riding (spillovers) and an inefficient allocation of the population (rivalry).

[22]This implies that the distribution of households boils down to a one-dimensional decision for the planner: By determining the border household, all poorer (richer) households are located in community 1 (2). Formally, using the notation established in section 2.2, $a_1(y) = 1 \; \forall \; y < y^{border}$ and $a_2(y) = 1 \; \forall \; y > y^{border}$, whereas all other as are zero.

[23]Their implicit SWF places substantially more weight on the rich than on the poor and therefore lies outside the range defined by the Utilitarian and Rawlsian welfare functions, which are usually considered the set of politically acceptable welfare criteria. The drawback of choosing a socially more acceptable SWF, however, is that the first-best case relies on individual lump-sum transfers, which do not take place in practice.

[24]Note that setting $\omega_R = R = 0$ results in a model that is not closed, which is not suitable for normative analysis. Another option would be to redistribute housing profits (equally) among households and thus drop the absentee landlords altogether. Whereas this would pose no problem for first-best, it would require a redistribution scheme for housing profits even in second-best.

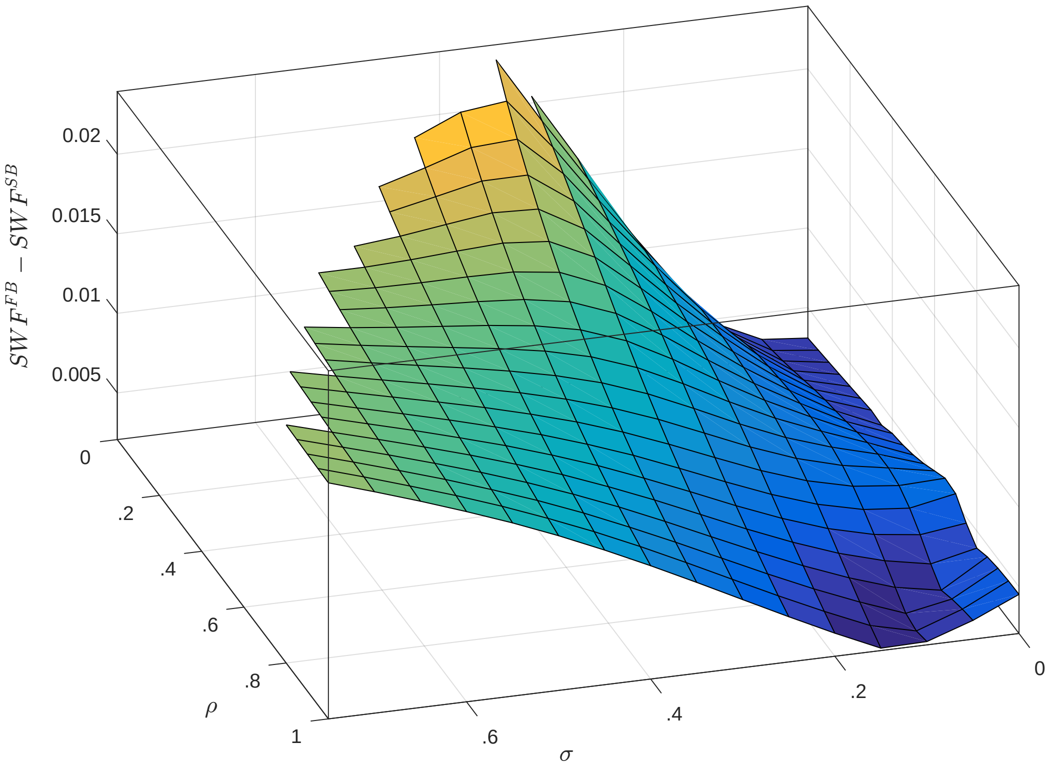

Figure 3: Difference in the value of the Social Welfare Function (SWF) between first-best and the fully decentralized equilibrium, as a function of σ and ρ

Note: For $\sigma > 0.78$ and combinations of ρ sufficiently small and σ sufficiently large we could not establish decentralized second-best income-segregating equilibria, as discussed in section 3.2. Therefore, we could not compute the difference of the SWF in first- and second-best for all combinations of σ and ρ, indicated by the "border" of the surface in the northwest. Note finally, that results depicted for $\sigma = 0$ are those for $\sigma = 0.01$, since we had computational problems for some combinations with $\sigma = 0$.

As a cardinal measure of welfare loss, we additionally compute households' compensating variation (CV), which we define as the money that a household needs to receive in second-best to maintain his first-best utility level.[25] Households with a positive CV are better off in first-best, whereas those with a negative CV are better off in the constrained solution. The aggregate monetized welfare loss of decentralization is the integral of the households' compensating variations (CV^{agg}). Figure A3 in the Appendix shows the aggregate compensating variation as a function of spillovers and rivalry.

To gain some intuition about the forces at work, we focus on two specific combinations of σ and ρ and identify the various inefficiencies that give rise to the welfare loss. For the case of a publicly provided private good ($\sigma = 0$, $\rho = 1$), there are no externalities and no returns to scale associated with the publicly provided good. The corresponding community and household characteristics of first- and second-best are presented in Table 2. More detailed household characteristics are shown in

[25]Technically, $CV(y)$ solves $V^{SB}(p_j^{SB}, t_j^{SB}, g_j^{SB}; [y + CV(y)]) = V^{FB}(p_j^{FB}, T_j^{FB}, t_j^{FB}, g_j^{FB}; y)$.

Table 2: First-best vs. second-best: Publicly provided private good

		First-best		Second-best	
Community characteristics		Community		Community	
		1	2	1	2
Income tax rate	t_j	0	0	0.145	0.157
Uniform lump-sum tax	T_j	0.225	0.239	*n.a.*	*n.a.*
Public consumption	g_j	0.225	0.239	0.167	0.259
Housing price	p_j	0.712	0.713	0.593	0.796
Border household	y^{border}	1.500		1.305	
Household characteristics					
Individual transfer	$r_{poorest}$	0.467		*n.a.*	
	r_{median}	-0.003		*n.a.*	
	$r_{richest}$	-0.466		*n.a.*	
Disposable income	$y^{disp}_{poorest}$	0.900		0.537	
	y^{disp}_{median}	0.952		0.906	
	$y^{disp}_{richest}$	0.952		1.328	
$V(y)$	$V(y)_{poorest}$	0.522		0.331	
	$V(y)_{median}$	0.522		0.534	
	$V(y)_{richest}$	0.552		0.725	
Welfare measures					
Value of Social Welfare Function		0.5657		0.5607	
CV^{agg} in % of total income		0.725			

Results are for $\sigma = 0$ and $\rho = 1$. Households are uniformly distributed between $\underline{y} = 1$ and $\overline{y} = 2$. Integrals are approximated by modeling 101 household types with income $y = 1, 1.01, 1.02, \ldots, 2$.
Note that the utilitarian planner would equalize all households' net incomes and distribute households equally such that everybody is equally well off. That we do not see a precise equalization is a result of relying on a finite number of households to solve our model numerically.

Figure A4 in the Appendix. The first-best equilibrium is characterized by equalization of incomes after taxes and transfers, housing prices, and public consumption. This equalization stems from the underlying utilitarian social welfare function in combination with concave utility functions and the fact that there are no externalities associated with the publicly provided good.[26] In the decentralized equilibrium, public and private consumption differs across the communities. Households with a pre-tax income that is above the mean benefit from decentralization, whereas the poorer households are better off in first-best. The monetized welfare loss due to decentralization amounts to 0.725% of total income, when measured as aggregate compensating variation.

Next, we consider the case of a "semi-public" good characterized by $\sigma = 0.3$, $\rho = 0.5$. The

[26]Note that this also applies to the jurisdictional choice externality (JCE), which is a fiscal rather than a technical externality. In first-best, the social planner uses head taxes which cover exactly the benefits received, such that there is no JCE, as shown by Proposition 1.

Table 3: First-best vs. second-best: The case of a "semi-public" good

		First-best		Second-best	
Community characteristics		Community		Community	
		1	2	1	2
Income tax rate	t_j	0	0	0.086	0.147
Uniform lump-sum tax	T_j	0.000	0.290	*n.a.*	*n.a.*
Public consumption	g_j	0.105	0.250	0.114	0.185
Housing price	p_j	0.516	0.825	0.680	0.759
Border household	y^{border}	1.200		1.457	
Household characteristics					
Individual transfer	$r_{poorest}$	-0.157		*n.a.*	
	r_{median}	0.173		*n.a.*	
	$r_{richest}$	-0.327		*n.a.*	
Disposable income	$y^{disp}_{poorest}$	0.539		0.578	
	y^{disp}_{median}	1.018		0.929	
	$y^{disp}_{richest}$	1.018		1.356	
$V(y)$	$V(y)_{poorest}$	0.307		0.321	
	$V(y)_{median}$	0.580		0.512	
	$V(y)_{richest}$	0.580		0.693	
Welfare measures					
Value of Social Welfare Function		0.553		0.541	
CV^{agg} in % of total income		2.238			

Results are for $\sigma = 0.3$ and $\rho = 0.5$. Households are uniformly distributed between $\underline{y} = 1$ and $\overline{y} = 2$. Integrals are approximated by modeling 101 household types with income $y = 1, 1.01, 1.02, \ldots, 2$.

community characteristics and central household characteristics for this case are presented in Table 3, and Figure 4 shows individualized transfers, disposable income, indirect utility and the CV as a function of income. The decentralized equilibrium is associated with an aggregate monetized welfare loss equal to more than 2.2% of total income, which is substantially more than in the case of the publicly provided private good.

In the decentralized solution, the utility level is monotonically increasing in y. The border household is $y = 1.457$, such that the poorest 46% of the population self-select into community 1, whereas the richer households choose to live in community 2. In contrast, the social planner distributes households into one larger rich and one smaller poor community ($y^{border} = 1.2$). He imposes a community-specific linear transfer scheme, in which the inhabitants of the poor community, along with the richest households, pay to subsidize the poorer inhabitants of the rich community (see upper left panel in Figure 4). Within each community, after taxes and transfers, everybody is equally well off. Due to much lower housing price in the less crowded community, the poor households' utility is relatively close to

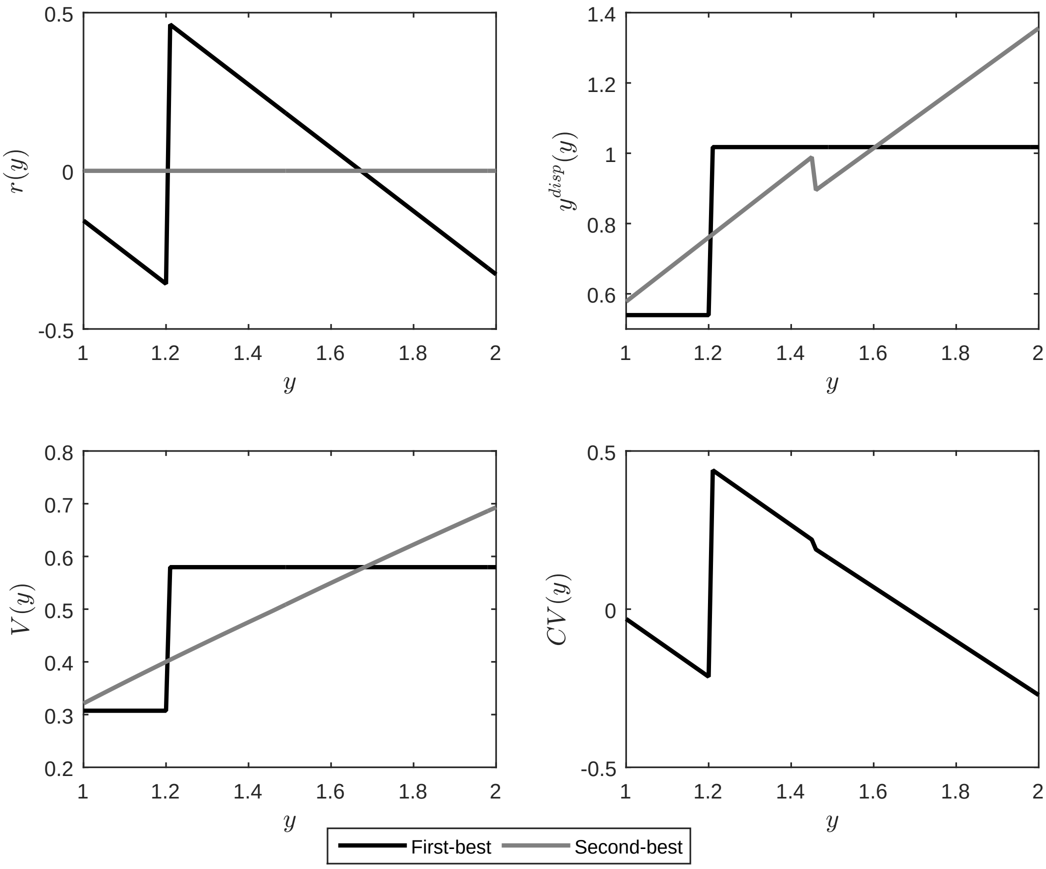

Figure 4: Household characteristics in FB and SB for the "semi-public" good ($\sigma = 0.3, \rho = 0.5$): Individual transfer scheme $r(y)$, disposable income $y^{disp}(y)$, utility level $V(y)$, and compensating variation $CV(y)$.

their second-best values, whereas the poorer households in the rich community are much better off. Thus, the social planner improves the outcome for middle-income households at the expense of the poor and the rich. The corresponding CV mirrors this result.[27]

The presence of spillovers leads to a substantial under-provision of the publicly provided good in the decentralized equilibrium due to free-riding. The population-weighted average of public consumption is lower by 31%. However, decentralized public consumption is higher in the poor community, mainly due to the fact that the average income is higher (compared with the first-best equilibrium) due to a different distribution of households across communities.

[27]This is a result of the utilitarian welfare function. If we were to place a higher welfare weight on poorer households, the results would likely change. What remains constant, however, is that the social planner increases aggregate welfare relative to the decentralized solution, but not by making everybody better off, but by redistributing utility from some households to others.

4.3 Welfare decomposition

The total welfare loss associated with decentralization is the combination of several inefficiencies present in second-best: imperfect tax instruments, decentralized tax setting, and free mobility, which, respectively, cause imperfect redistribution, inter-, and intra-community free-riding (cf. section 2.2). To quantify the relative size of these different inefficiencies, we design different "constrained first-best versions" of the model, in which we impose increasingly tighter restrictions on the instruments available to the social planner. Table 4 shows community characteristics and corresponding welfare measures for the sequence of equilibria.[28] The relative contribution of the individual inefficiencies to the overall welfare loss of decentralization can be inferred by comparing the aggregated CV across outcomes. We start out by holding σ and ρ fixed at 0.3 and 0.5, respectively, before examining the effect of varying these parameters.

Table 4: Decomposing the welfare loss of decentralization

	First-best		**I** $r(y) = 0$		**II** (I) & $T = 0$		**III** (II) & voting		**Second-best**	
	Community		Community		Community		Community		Community	
	1	2	1	2	1	2	1	2	1	2
t_j	0	0	1.000	0.602	0.000	0.175	0.006	0.156	0.086	0.147
T_j	0.000	0.290	-1.080	-0.679	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*	*n.a.*
g_j	0.105	0.250	0.107	0.243	0.105	0.244	0.081	0.206	0.114	0.185
p_j	0.516	0.825	0.509	0.829	0.526	0.824	0.597	0.801	0.680	0.749
y^{border}	1.200		1.160		1.180		1.280		1.457	
Value of SWF	0.5526		0.5502		0.5498		0.5456		0.5405	
CV^{agg} (% total income)	2.238		1.509		1.394		0.711			
Relative size of welfare loss			32.6% No between- community redistribution		5.1% No within- community redistribution		30.5% Inter-comm'ty free-riding		31.8% JCE	

Results are for $\sigma = 0.3$ and $\rho = 0.5$. Households are uniformly distributed between $\underline{y} = 1$ and $\overline{y} = 2$. Integrals are approximated by modeling 101 household types with income $y = 1, 1.01, 1.02, \ldots, 2$.

In intermediate step I, we set $r(y) = 0$ and thus restrict the social planner's ability to redistribute income only within communities (by combining a negative head tax with a positive income tax rate), but not across. In step II, we further remove head taxes as a policy instrument and thus also eliminate

[28] Naturally, there exist other sequences to move from first-best to the decentralized equilibrium. We chose this particular sequence because this order in which the social planner loses access to policy instruments appears intuitive to us. Any other path will generally result in a (slightly) different quantification of the respective welfare loss.

the possibility of within-community redistribution. In this step, public provision is financed entirely by income taxes. Since households living in the same community consume the same amount of the public good but pay different taxes, redistribution takes place in the form of public provision financed predominantly by the rich. Importantly, public provision (and thus the level of the income tax rate) and the distribution of households across communities remains determined by the social planner. The welfare loss resulting from eliminating between-community redistribution amounts to $(2.238 - 1.509 =) 0.73\%$ of total income, whereas eliminating within-community redistribution adds an additional loss of about 0.12%. In combination, the welfare loss from not being able to redistribute money across people and/or communities thus amounts to roughly a third of the total welfare loss associated with decentralization.

In step III, the communities are allowed to choose their own income tax rates and thus their preferred level of public provision by majority vote, whereas the role of the social planner is confined to allocating the population across communities. This opens the door to free-riding: Because some of the local public production is consumed in the other community, the utility-maximizing median voter does not consider the full benefits associated with local income taxes, and will thus choose a tax rate that is below the social optimum. At the same time, he free-rides on the public provision from outside the community borders. The levels of public consumption significantly drops in both communities. The inefficiency due to inter-community free-riding accounts for roughly another third of the total overall welfare loss.

Finally, letting households self-select into communities leads to the fully decentralized solution discussed above. This last step captures the jurisdictional choice externality: When selecting their preferred community, households do not consider the fiscal burden they impose on other community members via income tax payments, public consumption and the housing market. The JCE is responsible for the remaining third of the overall welfare loss.

Figure 5 displays household-specific utility levels and thus identifies winners and losers of each decomposition step. The utility levels in first-best (black) and second-best (gray) correspond to those in the bottom left graph of Figure 4 and show that decentralization implies an increase in utility by the rich and the poor, at the expense of middle incomes. Households with an income greater than 1.2 (almost) achieve their second-best utility already at decentralization step II, i.e. after removing inter-

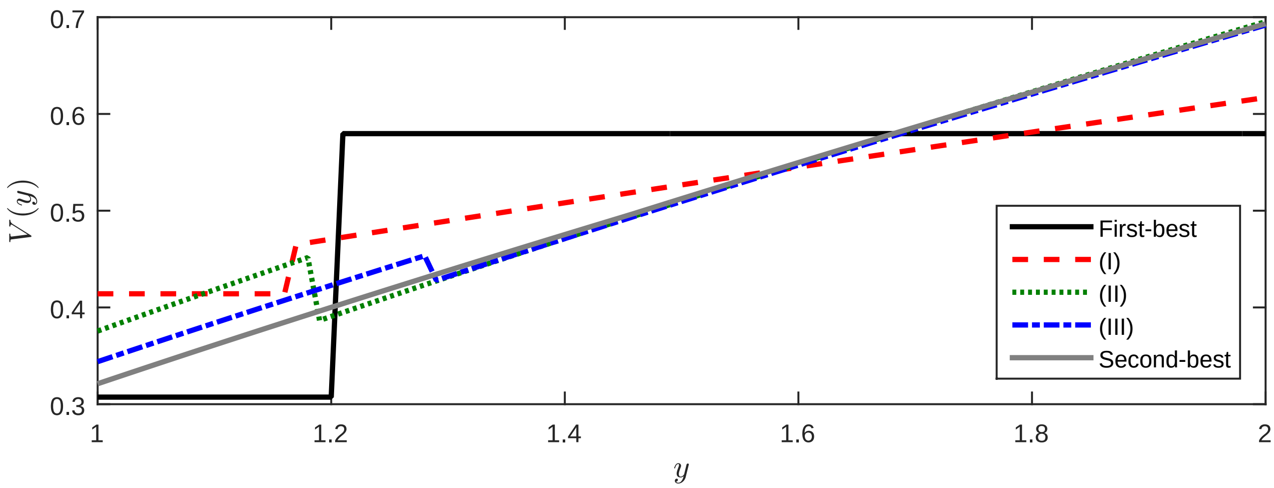

Figure 5: Utility as a function of income for different degrees of decentralization ($\sigma = 0.3, \rho = 0.5$).

and intra-community redistribution of income.[29] Inefficient public provision due to inter-community free-riding (II to III) mainly hurts the poor, for whom public consumption constitutes a larger share of total consumption. Likewise, the JCE reduces the utility of the poorest third of the population, while keeping the remainder of the population equally well off.

The JCE is sometimes discussed under the label of "the poor chasing the rich". A different way to state the same mechanism, and which more directly corresponds with our results, is that decentralization allows rich households to reduce the extent to which they have to share their income via tax-expenditure programs, by grouping themselves into communities with other rich households. In the extreme, i.e. with perfect sorting among extremely narrow intervals, no redistribution takes place via tax-expenditure programs at all, because everybody (approximately) pays the same taxes.

Varying the degree of spillovers

The above decomposition of the welfare loss was computed for a fixed combination of σ and ρ. Figure 6 shows the welfare loss at different decentralization steps as a function of σ, for ρ held fixed at 0.5. The top panel shows the value of the SWF, whereas the lower panel depicts the monetized welfare loss relative to the fully decentralized equilibrium in terms of the aggregate CV.

[29] Suppressing individualized lump-sum transfers as in Calabrese et al. (2012) would restrict the welfare loss associated with decentralization to the move from I to the fully decentralized solution (red dashed to gray solid line). In such a setting, there is no expropriation of the poorest and the richest at the expense of the middle class in the centralized solution, and as a consequence, decentralization benefits the rich at the expense of the poor. As discussed above, we chose to allow for individualized transfers in order to avoid the non-utilitarian SWF implicit in Calabrese et al. (2012).

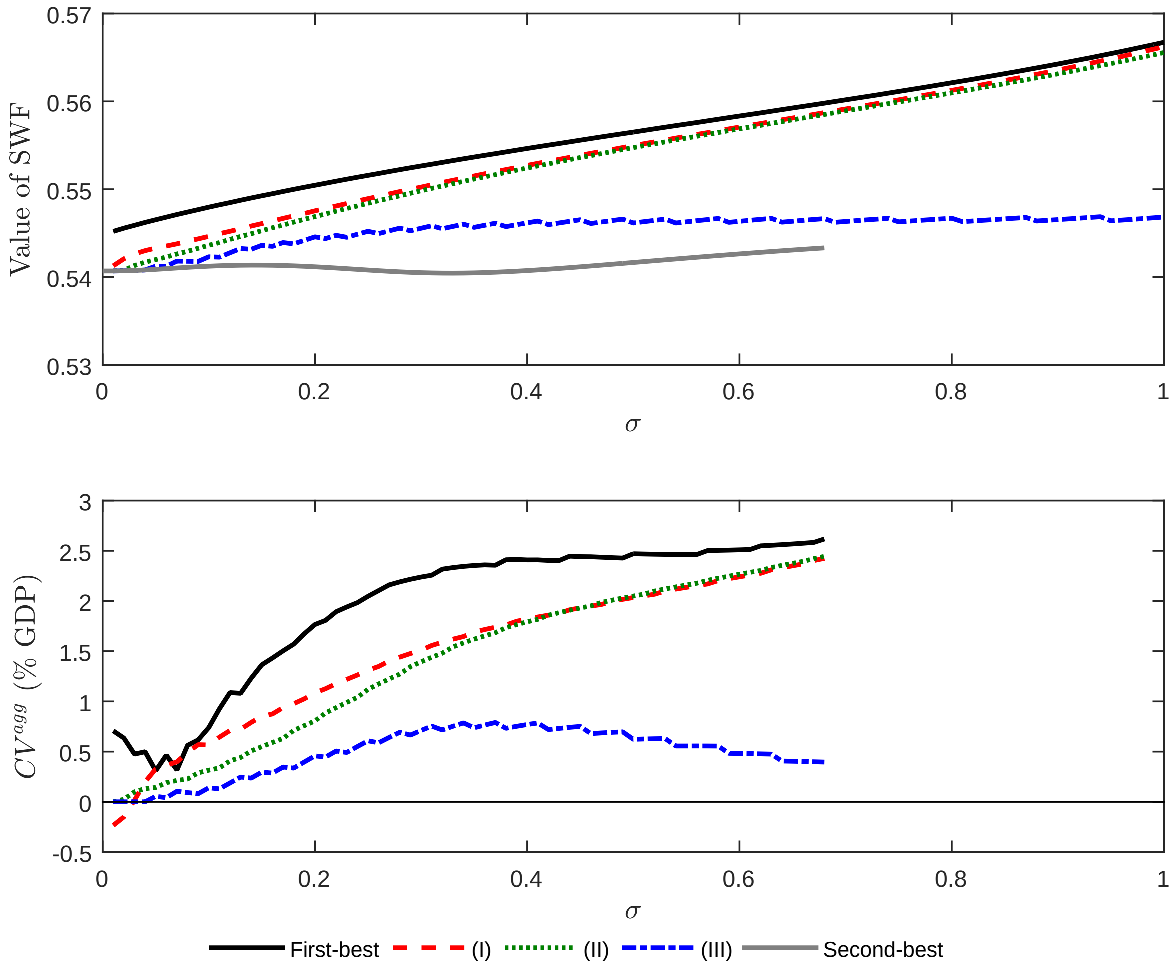

Figure 6: Welfare decomposition for varying σ ($\rho = 0.5$): Value of the social welfare function (SWF) and aggregate compensating variation (CV^{agg}) in % of total income.

Note: For combinations of $\rho = 0.5$ and $\sigma > 0.69$ we could not establish fully decentralized equilibria with income-segregation.

The value of the SWF is increasing in σ for all equilibria that are not fully decentralized, due to the larger benefits associated with producing a unit of the public good, at constant costs. The inefficiencies associated with spillovers and the JCE neutralize this increase in "technical" benefits almost completely, such that the SWF of the fully decentralized equilibrium is relatively insensitive to the value of σ. The welfare loss that is directly attributable to free-riding (the difference between the green and blue lines) is increasing in σ, which is intuitive: The larger the spillovers, the higher the inefficiencies related to free-riding. The drop in public consumption due to free-riding can be seen clearly in the bottom panel of Figure 7. This figure also shows that the positive association between welfare and spillovers is mainly due to an increase in private consumption, as public consumption does not significantly increase with the level of spillovers in first-best.

The welfare loss associated with the JCE is the difference between the blue and gray lines in the

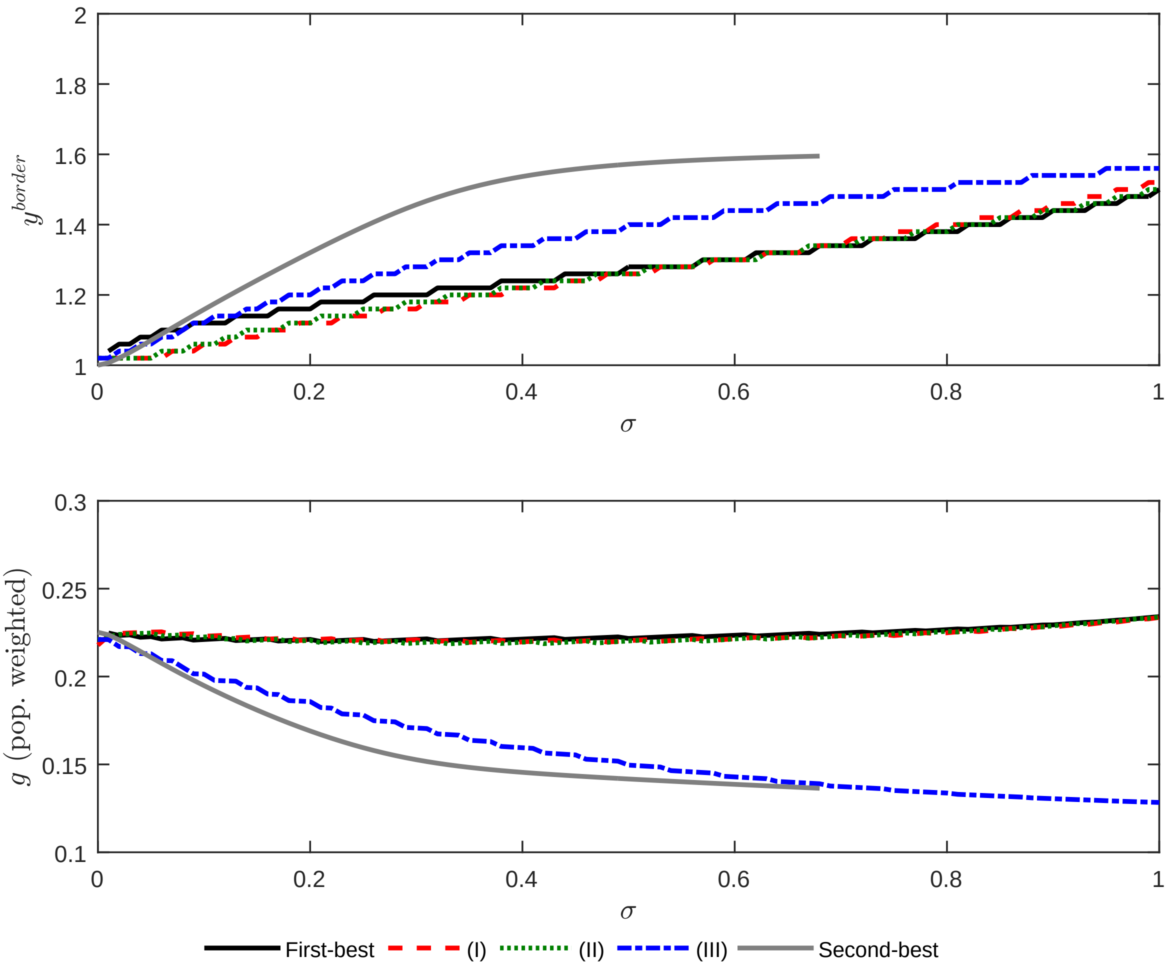

Figure 7: Welfare decomposition for varying σ ($\rho = 0.5$): Border household y^{border} and population weighted public consumption g.

Note: Population weighted public consumption means average public consumption level, i.e. $g \equiv \sum_j N_j g_j$.

top panel of Figure 6, or the blue line in the bottom panel in terms of monetized welfare. Our results show an inverse U-shaped relationship between the level of spillovers and the magnitude of the JCE, which can be explained by the presence of two effects of opposite sign.

Allowing households to choose results in segregation by income in the sense that the size of the rich community becomes smaller, relative to the situation where the population is allocated centrally.[30] Spillovers increase this segregation by income. This goes along with a decrease in

[30]Note that for most combinations of σ and ρ, the rich community is in fact larger than the poor community; even so, it remains true that for most values of σ, letting households choose their preferred location results in a smaller rich community (comparing the blue and gray lines in the top panel of Figure 7). For $0 < \sigma < 0.05$, the opposite occurs. However, as can be seen in Figures 1 and 2, the fully decentralized equilibrium at $\rho = 0.5$ and σ near zero consists in a poor community that is (almost) empty, with housing prices and tax rates equal to zero. This is clearly a degenerate case

population-weighted public consumption, which is the chief reason for the decrease in welfare due to the JCE.

At the same time, spillovers lead to more redistribution (holding public production constant), as poor households can consume some of the public good that is produced in the rich community. This non-monetary pathway of income redistribution lowers the inefficiency inherent in the JCE. Our results suggest that for $\rho = 0.5$, the former effect dominates until $\sigma \approx 0.38$, whereas the latter dominates afterward.

The relative welfare losses due to free-riding and the JCE, as well as their dependency on σ, are qualitatively similar when measured with the SWF or the CV. However, not having access to individualized transfers (the difference between the solid black and the dashed red lines) has a much larger welfare effect in terms of the CV than in terms of the value of the SWF, and this effect is highly nonlinear. The underlying reason is that the CV for a given change in utility depends on a household's income level. We have seen in Figure 5 that the social planner increases the utility of the middle incomes at the expense of everybody in the poor community and the very rich, and the same qualitative finding holds at different levels of spillovers. As spillovers increase from zero, two things happen: First, the size of the poor community increases (see top panel in Figure 7), and thus the relative share of the poor among the expropriated part of the population. This leads to a decrease in the CV, even though the value of the SWF increases (which is the basis for the optimization). Second, the utility level of the population in the poor community increases, which implies a larger CV relative to the decentralized equilibrium. As it turns out, the first effect dominates at very low levels of σ, whereas the latter dominates at higher levels.

Varying the degree of rivalry

Last, we allow the degree of rivalry to vary, while holding σ fixed at 0.3. The welfare loss of the decentralized second-best is displayed in Figure 8 as a function of ρ, again in terms of the value of the SWF and the CV associated with moving to the fully decentralized equilibrium; the corresponding border households and levels of public consumption are shown in Figure A5 in the Appendix.

The welfare loss from decentralization decreases with rivalry in consumption.[31] This is mainly

without much economic relevance.

[31]The fact that the level of welfare decreases with less rivalry may appear counter-intuitive, and is due to an artifact of our numerical model. We normalize the population to 1. Taking the derivative of the Lagrangian associated with (9)-(12)

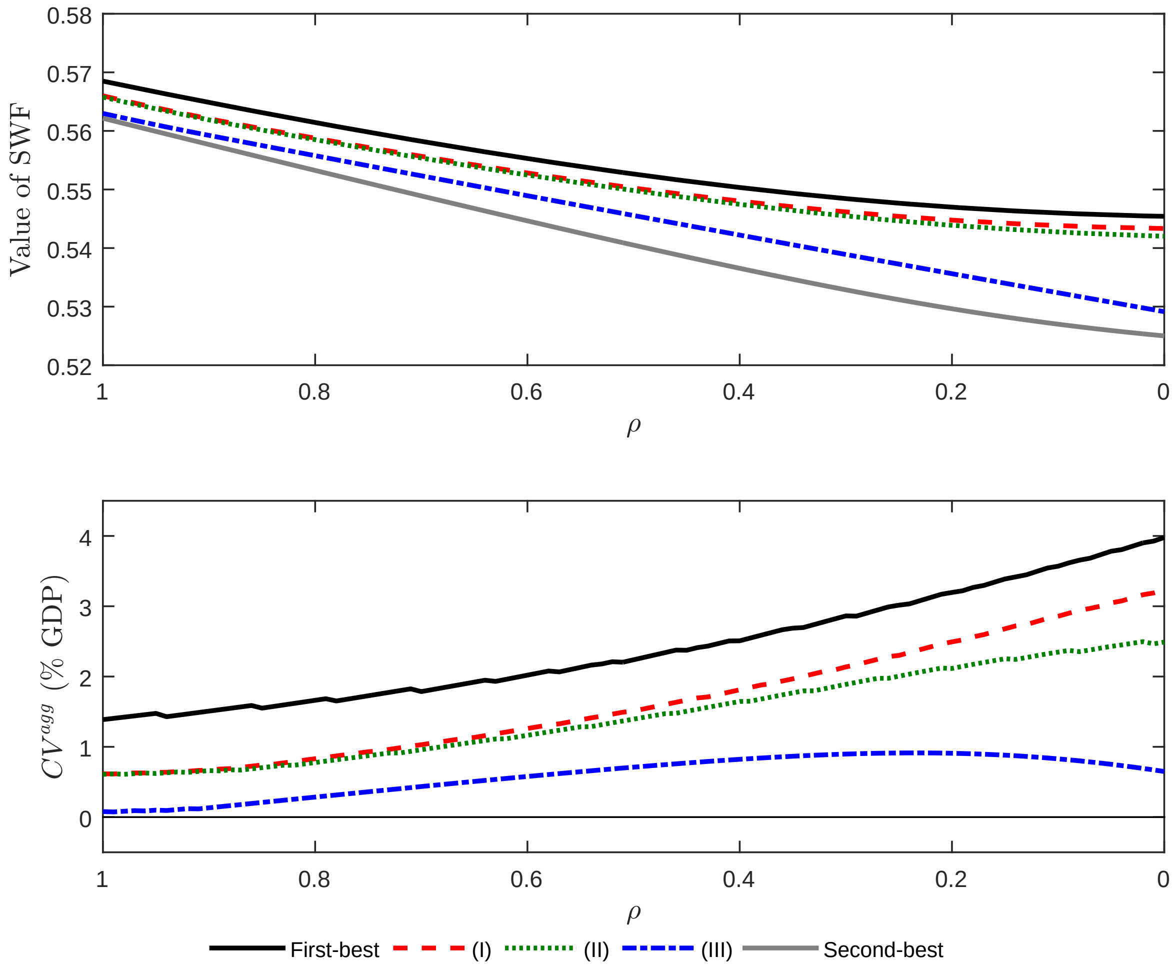

Figure 8: Welfare decomposition for varying ρ ($\sigma = 0.3$): Value of the social welfare function (SWF) and aggregate compensating variation (CV^{agg}) in % of total income.

due to an increase in free-riding (difference between green and blue lines): Decreasing the rivalry of the public good provides an incentive to produce even less of it from the point of view of an individual community, which exacerbates under-provision. Again, we find that the JCE is first increasing and then decreasing, due to a similar mechanism as before: Whereas for high levels of rivalry, segregation according to income is an effective way to prevent poorer households from sharing the public good, this becomes increasingly difficult as the good becomes less rival in consumption. Note that the figure is based on $\sigma = 0.3$, such that some of the public good produced in the rich community spills over to the poor community, where it can be shared effectively if ρ approaches zero.

with respect to ρ (and for simplicity, setting $\sigma = \nu = 1$) yields $-\sum_{j=1}^{J} \mu_j \sum_{i=1}^{J} (t_i Y_i + T_i N_i) \cdot \ln(N)/N^{\rho}$. For $N < 1$, this expression is positive, which is the level effect we observe in Figure 8. For N>1, however, welfare decreases with rivalry, which is more intuitive. The effect of ρ on the *change* in welfare from decentralization, however, is independent of the level.

5 Conclusions

The literature about environmental federalism is concerned with the decentralized regulation of externalities and publicly provided goods in the presence of inter-jurisdictional spillovers, but it abstracts from non-benefit taxation which is arguably an important issue both from the point of view of regional governments and for households when they choose the community they wish to live in. In contrast, the literature about tax competition focuses on the inefficiencies associated with decentralized non-benefit taxation, but by restricting the focus on publicly provided private goods it abstracts from the problem of inefficient public provision. This neglects the fact that many locally or regionally provided goods and services affect other jurisdictions as well. Examples include public facilities and environmental regulation. In this paper, we take a first step to merge these different strands of the literature. We do this by introducing spillovers and imperfect congestion into an income tax competition model. In addition to a theoretical model and a numerical implementation, we also carry out a normative analysis to assess the various inefficiencies inherent in decentralization. These inefficiencies have to be compared to the gains of decentralization e.g. in the form of better control of government agents or an increased sense of citizenship due to participation, from which we abstract throughout our analysis.

Our analsis shows that the case of a publicly provided private good, on which tax competition models usually focus, is very special in terms of the resulting equilibrium. The presence of spillovers and/or imperfect rivalry in consumption introduces additional terms to the equilibrium conditions, which render the utility of households in all communities a function of the public provision in all other communities. This interdependency does not arise with strictly local public provision.

As would be expected, spillovers cause communities to free-ride on the provision of others, and the presence of economies of scale exacerbates this issue. A direct policy implication based on our analysis is therefore that (relatively) pure public goods should be provided, or at least regulated, by the central government. This is in line with Oates' (1972) decentralization theorem.

However, spillovers also constitute a pathway of redistribution that limits the adverse effects of tax competition. We find that an increase in spillovers leads to a reduction in the difference between income levels, tax rates, and public consumption of the communities. If spillovers are sufficiently large, tax competition breaks down in the sense that households no longer segregate by income. This implies that the adverse effects of tax competition crucially depend on the characteristics of the publicly provided goods.

In this paper, we have focused on income tax competition. However, our model could also be applied to property tax competition, which is empirically relevant especially in the USA. Other possible extensions of our model include the use of more realistic distributions of income, progressive rather than linear taxation, jurisdictions of different size, asymmetric spillovers between communities (e.g., in the case of river management), and adding a vertical dimension of the federal system. We leave this for future research.

Appendix

Proof for Proposition 1

(i). In the first-best environment where the social planner has access to individualized transfers, $\eta_j = 0$ for any combination of welfare weights. To show this, note that for income-segregated equilibria (on which we concentrate), the summation terms in (14) can be dropped because $a_i = 0$ for all communities other than community j, which contains the income level y. Multiplying both sides by $h^j(y)a_j f(y)$, integrating over the support of the income distribution and substituting the budget constraint leads to

$$\int_{\underline{y}}^{\overline{y}} \omega(y) V_y^j(y) h^j(y) a_j f(y) \mathrm{d}y + \eta_j \int_{\underline{y}}^{\overline{y}} h_y^j(y) h^j(y) a_j f(y) \mathrm{d}y = \omega_R H^S(p_j).$$

Substituting into (17), dropping y as an argument and rearranging gives

$$\text{(A1)} \qquad \int_{\underline{y}}^{\overline{y}} \omega \left[V_p^j + V_y^j h^j \right] a_j f(y) \mathrm{d}y = -\eta_j \left(\int_{\underline{y}}^{\overline{y}} \left[h_p^j + h_y^j h^j \right] a_j f(y) \mathrm{d}y - H_p^S(p_j) \right).$$

The LHS is zero by Roy's identity. The parenthesis on the RHS consists of the slope of the compensated demand function, minus the slope of the supply function for housing, and is thus negative. In order for the equation to hold, η_j therefore has to be zero. □

Proofs for Proposition 2

(i). Multiplying (14) by $a_j f(y)$ and integrating leads to

$$\text{(A2)} \qquad \int_{\underline{y}}^{\overline{y}} \omega(y) V_y^j(y) a_j f(y) \mathrm{d}y + \eta_j \int_{\underline{y}}^{\overline{y}} h_y^j(y) a_j f(y) \mathrm{d}y = \omega_R N_j.$$

Substituting this expression into (15) while recognizing that $V_T(1-t_j) = -V_y$ and $h_T(1-t_j) = -h_y$ and simplifying leads to $\lambda_j + \sigma \sum_{i \neq j} \lambda_i = \frac{\omega_R}{1-t_j}$, which can be solved for

$$\text{(A3)} \qquad \lambda_j = \frac{\omega_R}{(1+\sigma(J-1))(1-\sigma)} \left(\frac{1+\sigma(J-2)}{1-t_j} - \sigma \sum_{i \neq j} \frac{1}{1-t_i} \right).$$

Substituting (A3) into (18) leads to the generalized Samuelson rule in part *(i)* of the proposition. □

(ii). Substitute $t_j = 0$ into (21). Divide the LHS by $\int_{\underline{y}}^{\overline{y}} \omega(y) V_y^j(y) a_j f(y) \mathrm{d}y$, and the RHS by $\omega_R N_j$, which by (A2) are equal (recall that in first-best, $\eta_j = 0$), to get

$$\text{(A4)} \qquad \frac{\int_{\underline{y}}^{\overline{y}} \omega(y) V_g^j(y) a_j f(y) \mathrm{d}y}{\int_{\underline{y}}^{\overline{y}} \omega(y) V_y^j(y) a_j f(y) \mathrm{d}y} = \frac{\lambda \left(N_j + \sum_{i \neq j} N_i \right)^{\rho}}{\omega_R N_j}$$

From (14), it follows that $\omega(y) V_y^j(y) = \omega_R$. Because this is a constant for all y, we can take it inside the integral in the numerator of the LHS; the integral of the denominator is one and drops out. Last, we cancel the $\omega(y)$ to arrive at (22). $\square$

(iii). This part follows directly from substituting $t_j = 0$ into (A3). $\square$

Table A1: First-best efficiency conditions for our functional forms

Choice variables	Efficiency conditions
t_j	$0 = \int_{\underline{y}}^{\overline{y}} \omega(y) V_t^j(y) a_j f(y) \mathrm{d}y + \eta_j \int_{\underline{y}}^{\overline{y}} h_t^j(y) a_j f(y) \mathrm{d}y + Y_j \Big(\lambda_j + \sigma \sum_{i \neq j} \lambda_i\Big)$
T_j	$0 = \int_{\underline{y}}^{\overline{y}} \omega(y) V_T^j(y) a_j f(y) \mathrm{d}y + \eta_j \int_{\underline{y}}^{\overline{y}} h_T^j(y) a_j f(y) \mathrm{d}y + N_j \Big(\lambda_j + \sigma \sum_{i \neq j} \lambda_i\Big)$
p_j	$0 = \int_{\underline{y}}^{\overline{y}} \omega(y) V_p^j(y) a_j f(y) \mathrm{d}y + \eta_j \Big(\int_{\underline{y}}^{\overline{y}} h_p^j(y) a_j f(y) \mathrm{d}y - H_p^S(p_j)\Big) + \omega_R H^S(p_j)$
g_j	$0 = \int_{\underline{y}}^{\overline{y}} \omega(y) V_g^j(y) a_j f(y) \mathrm{d}y + \eta_j \int_{\underline{y}}^{\overline{y}} h_g^j(y) a_j f(y) \mathrm{d}y - \lambda_j \Big(N_j + \nu \sum_{i \neq j} N_i\Big)^\rho$
$r(y)$	$0 = \omega(y) V_r^j(y) a_j f(y) + \eta_j h_r^j(y) a_j f(y) + \Omega$
R	$0 = \omega_R + \Omega$
$y_{i,j}^{border}$	$0 = MSV_i(y_{i,j}^{border}) - MSV_j(y_{i,j}^{border})$
Shadow variables	**Feasibility constraints**
λ_j	$0 = t_j Y_j + T_j N_j + \sigma \sum_{i \neq j}(t_i Y_i + T_i N_i) - g_j \cdot \Big(N_j + \nu \sum_{i \neq j} N_i\Big)^\rho$
η_j	$0 = \int_{\underline{y}}^{\overline{y}} h^j(y) a_j(y) f(y) \mathrm{d}y - H^S(p_j)$
Ω	$0 = R + \int_{\underline{y}}^{\overline{y}} r(y) f(y) \mathrm{d}y$
First derivatives	**Functional forms**
V_t^j	$0 = -\frac{(1-\alpha) y V^j(y)}{y(1-t_j) - T_j + r(y) - p_j \beta_h - \beta_x} - V_t^j$
V_T^j	$0 = -\frac{(1-\alpha) V^j(y)}{y(1-t_j) - T_j + r(y) - p_j \beta_h - \beta_x} - V_T^j$
V_p^j	$0 = -\frac{(1-\alpha) \beta_h V^j(y)}{y(1-t_j) - T_j + r(y) - p_j \beta_h - \beta_x} - \frac{\gamma V^j(y)}{p_j} - V_p^j$
V_g^j	$0 = \frac{\alpha V^j(y)}{g_j - \beta_g} - V_g^j$
V_r^j	$0 = \frac{(1-\alpha) V^j(y)}{y(1-t_j) - T_j + r(y) - p_j \beta_h - \beta_x} - V_r^j$
h_t^j	$0 = -\frac{\gamma y}{(1-\alpha) p_j} - h_t^j$
h_T^j	$0 = -\frac{\gamma}{(1-\alpha) p_j} - h_T^j$
h_p^j	$0 = -\gamma \frac{y(1-t_j) - T_j + r(y) - \beta_x}{(1-\alpha) p_j^2} - h_p^j$
h_g^j	$0 = h_g^j$
h_r^j	$0 = \frac{\gamma}{(1-\alpha) p_j} - h_r^j$
$H_{p_j}^S$	$0 = \theta L_j p_j^{\theta - 1} - H_{p_j}^S$

$H^S(p_j) = L_j p_j^\theta$, $MSV_j(y_{i,j}^{border})$ is given by (13) for $y = y_{i,j}^{border}$ and $V^j(y)$ by (26).

Table A2: Decentralized equilibrium results for different values of σ and ρ.

			$\overline{y_j}$	Y_j	p_j	t_j	G_j	g_j
$\sigma = 0$								
	$\rho = 1$	Community 1	1.305	0.352	0.593	0.145	0.051	0.167
		Community 2	2	1.148	0.796	0.157	0.180	0.259
	$\rho \leq 0.65$	Community 1	1	0	0.007	0.158	0	0.006
		Community 2	2	1.500	0.870	0.150	0.225	0.225
$\sigma = 0.1$								
	$\rho = 1$	Community 1	1.369	0.437	0.632	0.115	0.050	0.154
		Community 2	2	1.063	0.777	0.154	0.164	0.253
	$\rho = 0.5$	Community 1	1.159	0.171	0.495	0.050	0.008	0.058
		Community 2	2	1.329	0.835	0.153	0.203	0.221
	$\rho = 0$	Community 1	1.022	0.022	0.285	0	0	0.022
		Community 2	2	1.478	0.865	0.151	0.224	0.224
$\sigma = 0.3$								
	$\rho = 1$	Community 1	1.479	0.594	0.690	0.092	0.055	0.148
		Community 2	2	0.906	0.742	0.145	0.132	0.223
	$\rho = 0.5$	Community 1	1.457	0.561	0.680	0.086	0.048	0.114
		Community 2	2	0.939	0.749	0.147	0.138	0.185
	$\rho = 0$	Community 1	1.346	0.406	0.628	0.045	0.018	0.069
		Community 2	2	1.094	0.784	0.153	0.168	0.173
$\sigma = 0.5$								
	$\rho = 1$	Community 1	1.555	0.709	0.724	0.088	0.062	0.146
		Community 2	2	0.791	0.715	0.130	0.103	0.185
	$\rho = 0.5$	Community 1	1.572	0.736	0.730	0.094	0.069	0.132
		Community 2	2	0.764	0.708	0.126	0.096	0.155
	$\rho = 0.15$	Community 1	1.592	0.768	0.738	0.101	0.077	0.125
		Community 2	2	0.732	0.700	0.120	0.088	0.133
	$\rho \leq 0.1$	No income separating equilibrium found.						
$\sigma = 0.7$								
	$\rho = 1$	Community 1	1.590	0.764	0.739	0.086	0.065	0.141
		Community 2	2	0.736	0.703	0.112	0.083	0.156
	$\rho = 0.6$	Community 1	1.595	0.772	0.740	0.088	0.068	0.134
		Community 2	2	0.728	0.701	0.110	0.080	0.144
	$\rho \leq 0.55$	No income separating equilibrium found.						
$\sigma = 0.78$								
	$\rho = 1$	Community 1	1.596	0.774	0.741	0.084	0.065	0.138
		Community 2	2	0.726	0.701	0.106	0.077	0.147
	$\rho \leq 0.95$	No income separating equilibrium found.						
$\sigma \geq 0.8$		No income separating equilibrium found.						

Definition of parameters: σ: spillover parameter (0: no spillover, 1: perfect spillover; ρ: parameter for rivalry in consumption (0: no rivalry, 1: perfect rivalry). *Definition of variables:* $\overline{y_j}$: highest income household in j (community 1 inhabits poor, community 2 rich households; $\overline{y_1} = \underline{y_2}$ is the indifferent household); Y_j: aggregate income; p_j: housing price; t_j: income tax rate; G_j: contribution to publicly provided good; g_j: consumption of publicly provided good.

Table A3: First-best equilibrium results for different values of σ and ρ.

			$\overline{y_j}$	Y_j	p_j	T_j	g_j	CV^{agg} (% of total income)
$\sigma = 0$								
	$\rho = 1$	Community 1	1.5	0.625	0.712	0.225	0.225	0.725
		Community 2	2	0.875	0.713	0.239	0.239	
	$\rho = 0.5$	Community 1	1.06	0.062	0.323	0	0.016	1.062
		Community 2	2	1.438	0.859	0.244	0.237	
	$\rho = 0$	Community 1	1.02	0.02	0.235	0	0.005	0.457
		Community 2	2	1.48	0.866	0.233	0.229	
$\sigma = 0.1$								
	$\rho = 1$	Community 1	1.48	0.595	0.699	0.205	0.21	0.438
		Community 2	2	0.905	0.726	0.257	0.252	
	$\rho = 0.5$	Community 1	1.12	0.127	0.42		0.051	0.739
		Community 2	2	1.373	0.847	0.262	0.244	
	$\rho = 0$	Community 1	1.02	0.02	0.256	0	0.023	0.090
		Community 2	2	1.48	0.865	0.233	0.229	
$\sigma = 0.3$								
	$\rho = 1$	Community 1	1.34	0.398	0.601	0	0.13	1.387
		Community 2	2	1.102	0.79	0.354	0.307	
	$\rho = 0.5$	Community 1	1.2	0.22	0.516	0	0.105	2.238
		Community 2	2	1.28	0.825	0.29	0.25	
	$\rho = 0$	Community 1	1.02	0.02	0.289	0	0.069	3.978
		Community 2	2	1.48	0.864	0.233	0.229	
$\sigma = 0.5$								
	$\rho = 1$	Community 1	1.36	0.425	0.629	0	0.172	2.018
		Community 2	2	1.075	0.776	0.365	0.285	
	$\rho = 0.5$	Community 1	1.28	0.319	0.584	0	0.146	2.470
		Community 2	2	1.181	0.8	0.324	0.251	
	$\rho = 0$	Community 1	1.12	0.127	0.459	0	0.115	*n.a.*
		Community 2	2	1.373	0.841	0.262	0.231	
$\sigma = 0.7$								
	$\rho = 1$	Community 1	1.38	0.452	0.651	0	0.201	2.545
		Community 2	2	1.048	0.763	0.377	0.264	
	$\rho = 0.5$	Community 1	1.34	0.398	0.63	0	0.183	2.618
		Community 2	2	1.102	0.777	0.354	0.246	
	$\rho = 0$	Community 1	1.26	0.294	0.583	0	0.163	*n.a.*
		Community 2	2	1.206	0.801	0.315	0.233	
$\sigma = 1$								
	$\rho = 1$	Community 1	1.5	0.625	0.715	0	0.234	*n.a.*
		Community 2	2	0.875	0.711	0.468	0.234	
	$\rho = 0.5$	Community 1	1.5	0.625	0.715	0	0.234	*n.a.*
		Community 2	2	0.875	0.711	0.468	0.234	
	$\rho = 0$	Community 1	1.5	0.625	0.715	0	0.234	*n.a.*
		Community 2	2	0.875	0.711	0.468	0.234	

Definition of parameters: σ: spillover parameter (0: no spillover, 1: perfect spillover; ρ: parameter for rivalry in consumption (0: no rivalry, 1: perfect rivalry). *Definition of variables:* $\overline{y_j}$: highest income household in j (community 1 inhabits poor, community 2 rich households; $\overline{y_1} = \underline{y_2}$ is the border household); y_j^m: median income (for uniform distribution equal to mean income); Y_j: aggregate income; p_j: housing price; T_j: lump-sum tax; g_j: consumption of publicly provided good; CV^{agg} in % of total income: Aggregate compensating variation in % of aggregate income (welfare loss of decentralization).
Note: CV for ($\sigma = 0.7$, $\rho = 0.5$) is for $\sigma = 0.68$ because for sigma=0.7 no SB equilibrium was found.

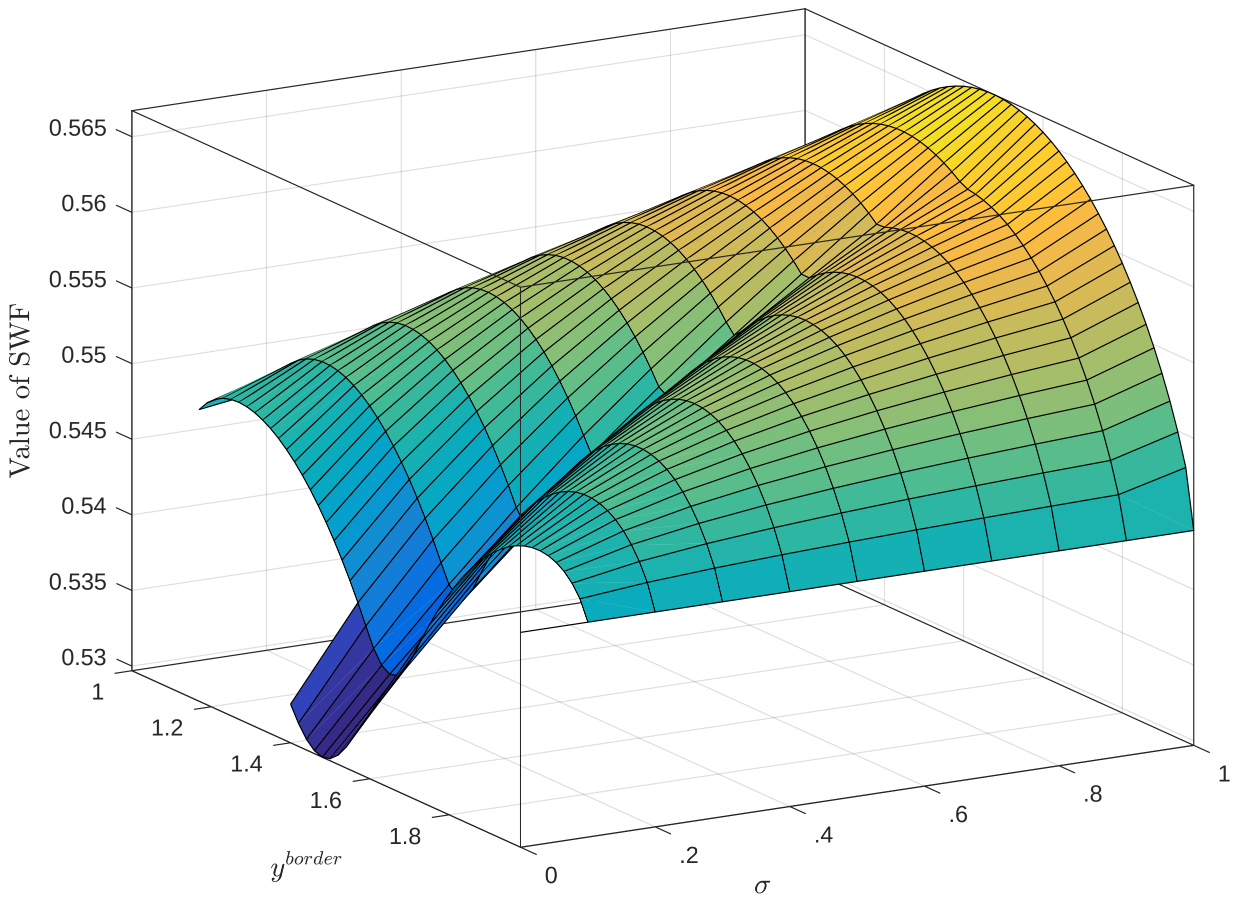

Figure A1: First-best value of the social welfare function (SWF) for variations in the level of spillovers σ and the border household y^{border}.

This surface plot describes the value of the SWF with $\rho = 0.5$. The degree of spillovers increases from left to right. From the front to the back the border household (y^{border}) between both communities varies. There are three saddle-points of the SWF: a minimum at $y^{border} = 1.5$, a local maximum that would imply a small, rich community and a global maximum with a small, poor community. The path along this saddle point is reported in Table A3 (for $\rho = 0.5$). Note that with perfect spillovers (at the right), dividing the population equally between both communities turns out to be optimal.

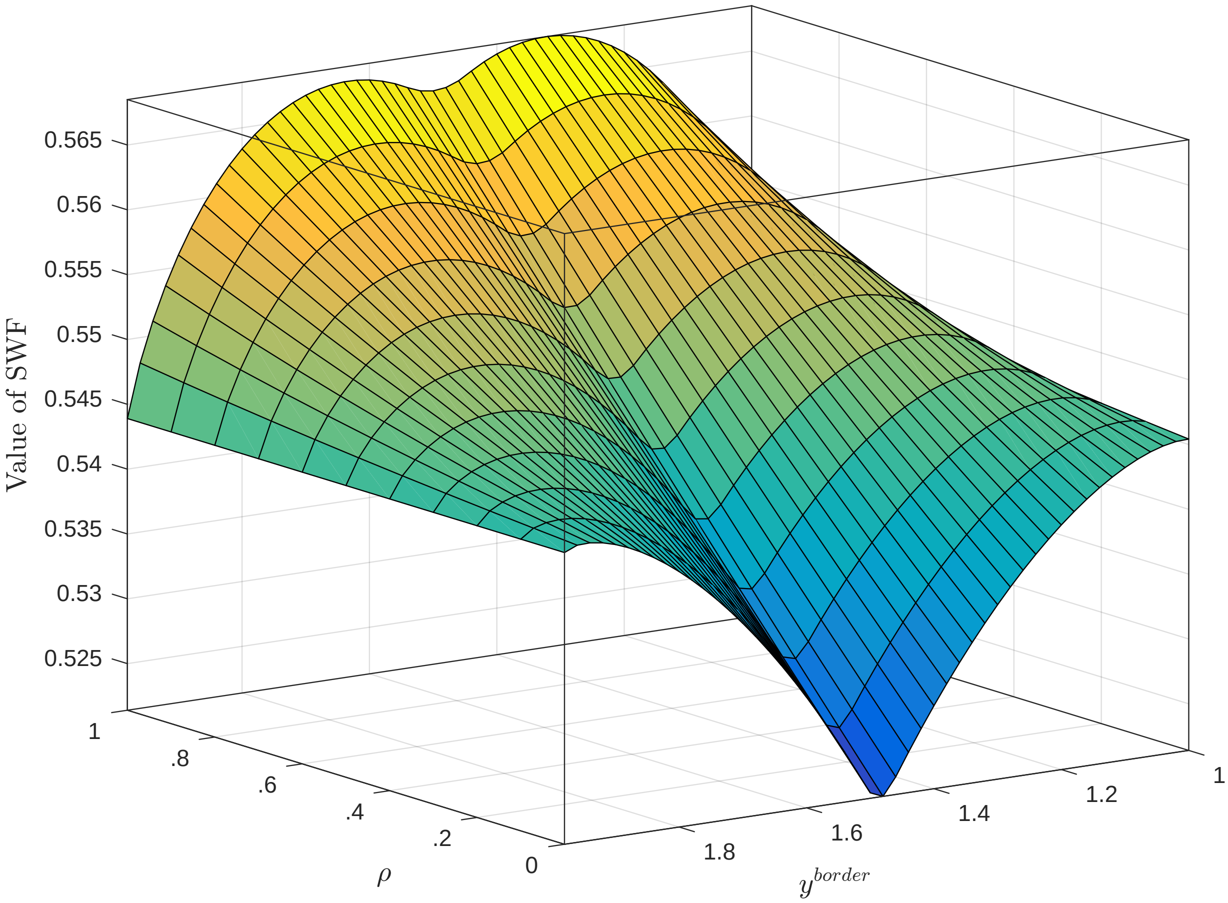

Figure A2: First-best value of the social welfare function (SWF) for variations in the level of rivalry ρ and the border household y^{border}.

This surface plot describes the achievable value of the SWF with $\sigma = 0.3$. The degree of rivalry decreases from the back to the front. From left to right the border household (y^{border}) varies. There are three saddle-points of the SWF: a minimum at $y^{border} = 1.5$, a local maximum that would imply a small, rich community and the global maximum with a small, poor community. The path along the latter is reported in Table A3 (for $\sigma = 0.3$).

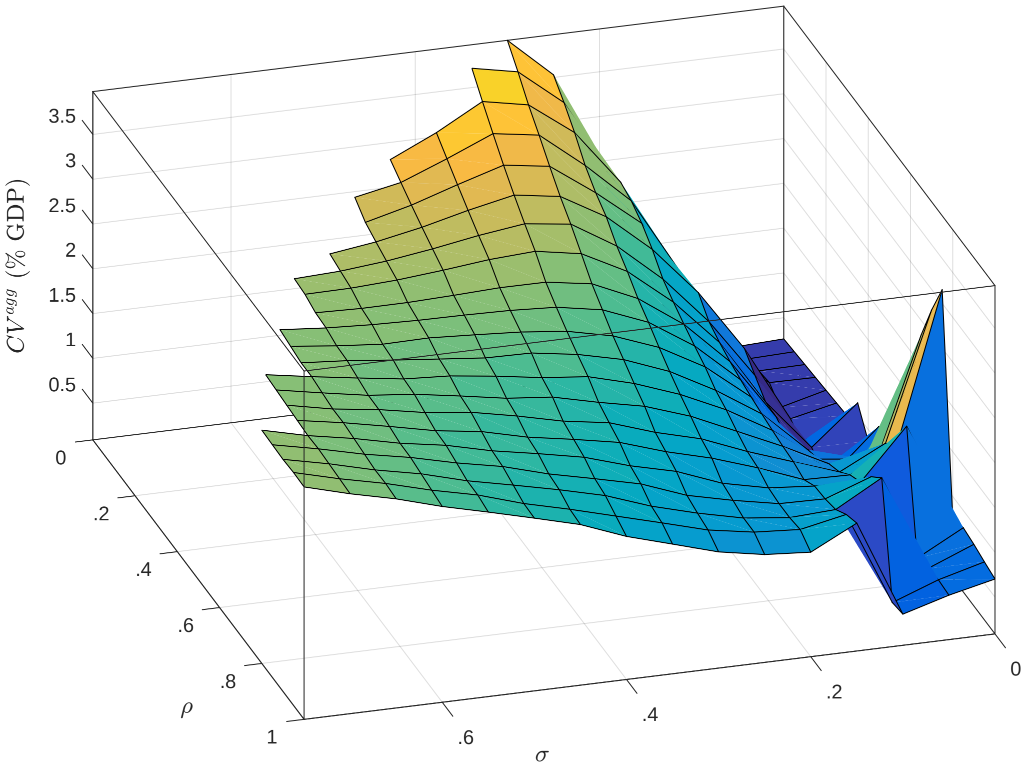

Figure A3: First-best vs. second-best: CV^{agg} in % of total income for σ and ρ variations

Note: For $\sigma > 0.78$ and combinations of ρ sufficiently small and σ sufficiently large, we could not establish decentralized income-segregating equilibria, as discussed in section 3.2. Therefore, we could not compute aggregate compensating variation levels for all combinations of σ and ρ, indicated by the "border" of the surface in the upper left of the figure. Note finally, that results depicted for $\sigma = 0$ are those for $\sigma = 0.01$, since we had computational problems for some combinations with $\sigma = 0$.

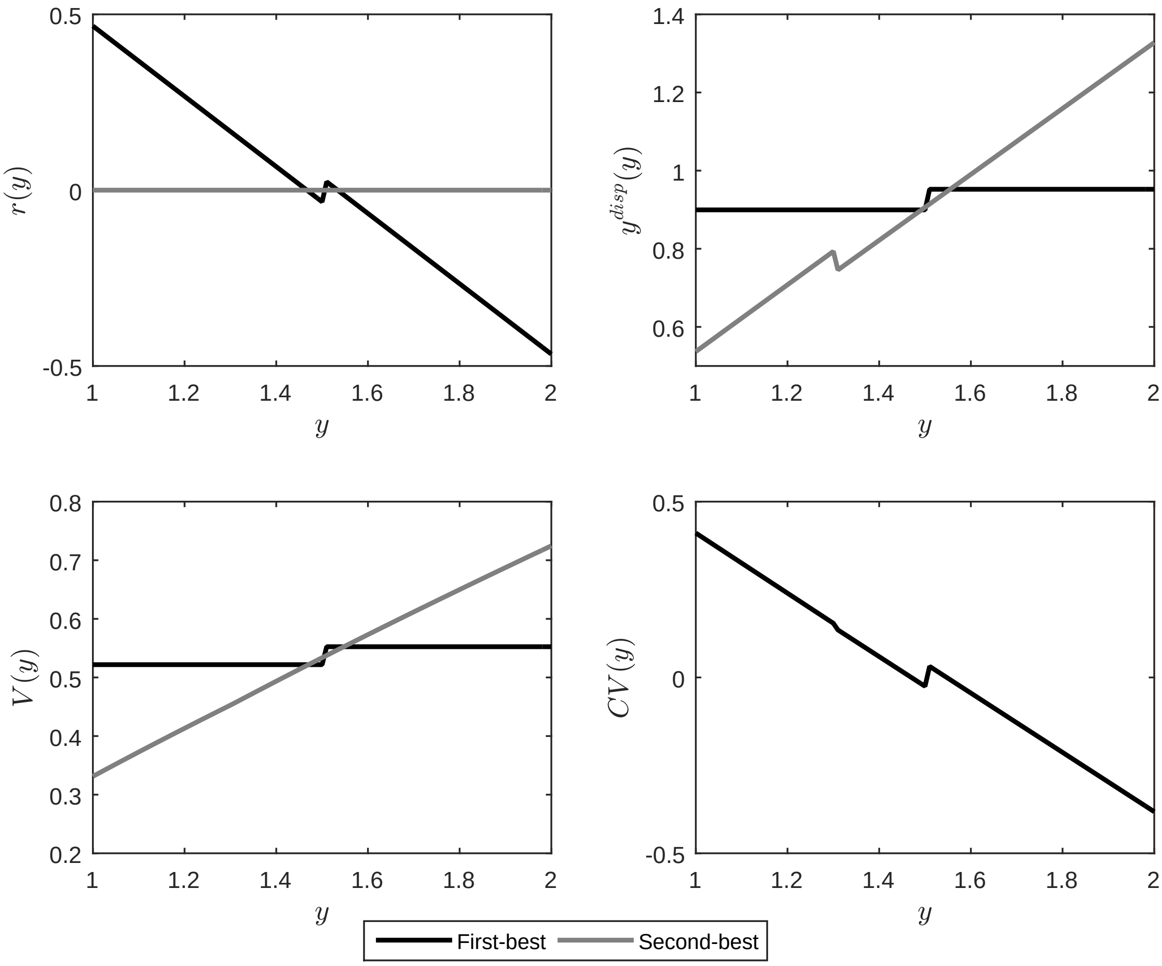

Figure A4: Household characteristics in FB and SB for a publicly provided private good ($\sigma = 0, \rho = 1$): Individual transfer scheme $r(y)$, disposable income $y^{disp}(y)$, utility level $V(y)$, and compensating variation $CV(y)$.

Note: The small kinks in the first-best (black) lines above are the result of relying on a finite number of households (101 in our case) to solve our model numerically. However, we do not split up the mass of households at the border household, but assign them to one of the two communities, which implies that both communities are not exactly equally inhabited. Increasing the number of modeled household types would smooth out these kinks.

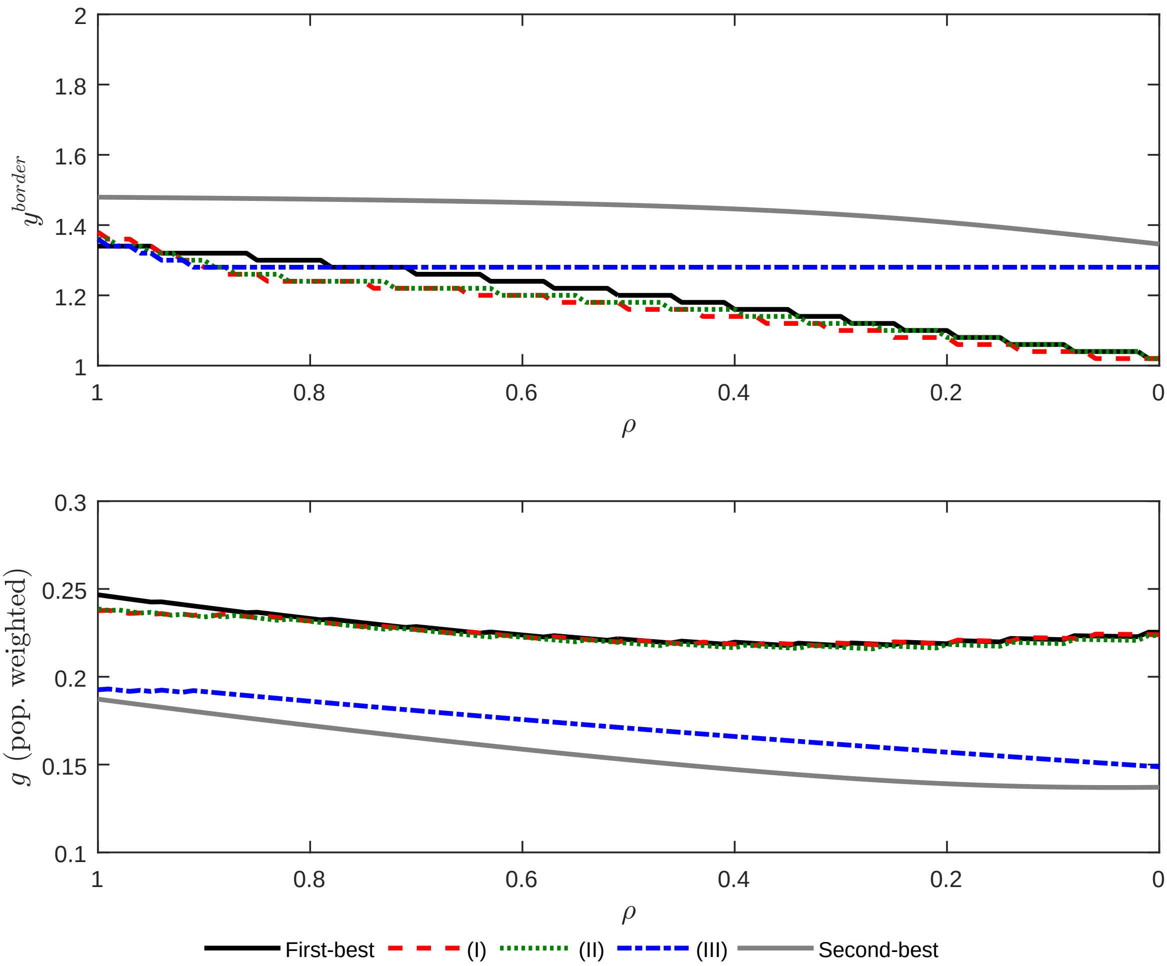

Figure A5: Welfare decomposition for varying ρ ($\sigma = 0.3$): Border household y^{border} and population weighted public consumption g.

Note: Population weighted public consumption means average public consumption level, i.e. $g \equiv \sum_j N_j g_j$.

References

Bjorvatn, K. & Schjelderup, G. (2002), 'Tax competition and international public goods.', *International Tax and Public Finance* **9**(2), 111 – 120.

Boadway, R. & Tremblay, J.-F. (2012), 'Reassessment of the tiebout model.', *Journal of Public Economics* **96**(11-12), 1063 – 1078.

Bruelhart, M., Bucovetsky, S. & Schmidheiny, K. (2015), Taxes in cities, Vol. 5 of *Handbook of Regional and Urban Economics*, Elsevier, pp. 1123 – 1196.

Bucovetsky, S. (2011), 'Incentive equivalence with fixed migration costs.', *Journal of Public Economics* **95**(11-12), 1292 – 1301.

Calabrese, S., Epple, D., Romer, T. & Sieg, H. (2006), 'Local public good provision: Voting, peer effects, and mobility.', *Journal of Public Economics* **90**(6-7), 959 – 981.

Calabrese, S. M., Epple, D. N. & Romano, R. E. (2012), 'Inefficiencies from metropolitan political and fiscal decentralization: Failures of tiebout competition', *The Review of Economic Studies* **79**(3), 1081 – 1111.

Dembour, C. & Wauthy, X. (2009), 'Investment in public infrastructure with spillovers and tax competition between contiguous regions.', *Regional Science and Urban Economics* **39**(6), 679 – 687.

Eichner, T. & Runkel, M. (2012), 'Interjurisdictional spillovers, decentralized policymaking, and the elasticity of capital supply.', *American Economic Review* **102**(5), 2349 – 2357.

Epple, D. & Platt, G. (1998), 'Equilibrium and local redistribution in an urban economy when households differ in both preferences and incomes', *Journal of Urban Economics* **43**(1), 23–51.

Epple, D., Romer, T. & Sieg, H. (2001), 'Interjurisdictional sorting and majority rule: An empirical analysis', *Econometrica* **69**(6), 1437–1465.

Feld, L. P. & Kirchgaessner, G. (2001), 'Income tax competition at the state and local level in Switzerland', *Regional Science and Urban Economics* **31**(2-3), 181–213.

Fernandez, R. & Rogerson, R. (1996), 'Income distribution, communities, and the quality of public education', *The Quarterly Journal of Economics* **111**(1), 135–164.

Geary, R. C. (1950), 'A note on "a constant-utility index of the cost of living"', *The Review of Economic Studies* **18**(1), pp. 65–66.

Hoel, M. (2004), 'Interregional interactions and population mobility.', *Journal of Economic Behavior and Organization* **55**(3), 419 – 433.

Hoel, M. & Shapiro, P. (2003), 'Population mobility and transboundary environmental problems', *Journal of Public Economics* **87**(5-6), 1013 – 1024.

Hoel, M. & Shapiro, P. (2004), 'Transboundary environmental problems with mobile but heterogeneous populations.', *Environmental and Resource Economics* **27**(3), 265 – 271.

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

Oates, W. E. (1999), 'An essay on fiscal federalism.', *Journal of Economic Literature* **37**(3), 1120 – 1149.

Oddou, R. (2011), 'The effect of spillovers and congestion on the segregative properties of endogenous jurisdiction structure formation', *FEEM Working Paper* (46-2011).

Oddou, R. (2016), 'The effect of spillovers and congestion on the endogenous formation of jurisdictions', *Journal of Public Economic Theory* **18**(1), 67–83.

Ogawa, H. & Wildasin, D. E. (2009), 'Think locally, act locally: Spillovers, spillbacks, and efficient decentralized policymaking.', *American Economic Review* **99**(4), 1206 – 1217.

Ross, S. & Yinger, J. (1999), 'Sorting and voting: A review of the literature on urban public finance', *Handbook of Regional and Urban Economics* **3**, 2001–2060.

Schmidheiny, K. (2002), Community Choice and Local Income Taxation, PhD thesis, University of Berne.

Schmidheiny, K. (2006*a*), 'Income segregation and local progressive taxation: Empirical evidence from Switzerland', *Journal of Public Economics* **90**(3), 429–458.

Schmidheiny, K. (2006*b*), 'Income segregation from local income taxation when households differ in both preferences and incomes.', *Regional Science and Urban Economics* **36**(2), 270 – 299.

Stone, R. (1954), 'Linear expenditure systems and demand analysis: an application to the pattern of British demand', *The Economic Journal* pp. 511–527.

Tiebout, C. M. (1956), 'A pure theory of local expenditures.', *Journal of Political Economy* **64**, 416 – 424.

Wellisch, D. (1994), 'Interregional spillovers in the presence of perfect and imperfect household mobility.', *Journal of Public Economics* **55**(2), 167 – 184.

Wilson, J. D. (1999), 'Theories of tax competition', *National tax journal* **52**, 269–304.

Wilson, J. D. & Wildasin, D. E. (2004), 'Capital tax competition: Bane or boon.', *Journal of Public Economics* **88**(6), 1065 – 1091.