Sas, Willem

Conference Paper

Ad valorem taxation, equalisation and commuting in a federation: Horizontal and vertical tax externalities revisited

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Sas, Willem (2015) : Ad valorem taxation, equalisation and commuting in a federation: Horizontal and vertical tax externalities revisited, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/124684

Ad valorem taxation, equalisation and commuting in a federation: horizontal and vertical tax externalities revisited

Willem Sas*
Center for Economic Studies, KU Leuven

Abstract

In this paper, commuting is introduced to a federal economy where benevolent lower-level (state) governments levy an *ad valorem* tax on labour income. This results in inefficiently low levels of taxation, even when households as a whole do not migrate. Indeed, rather than attracting more workers by lowering taxes, states are out to boost labour supplied by own residents and impede work incentives of non-residents. When the tax base is co-occupied by the federal and state governments secondly, either under- or overtaxation occurs. We find that when taxation is levied ad valorem, undertaxation is more liable to occur than under unit taxation. For the same underlying reasons lastly, fiscal equalisation is expected to give less cause for overtaxation as commonly assumed.

JEL Classification: H71, H77, H23.

Keywords: Tax competition, Ad valorem taxation, Fiscal equalisation, Vertical and horizontal externalities, Decentralisation, Federalism.

*Center for Economic Studies (CES), KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium. Tel: +32 472 26 09 89. E-mail: Willem.Sas@kuleuven.be.

Introduction

In most federations today, both federal and state governments raise taxes on the same -or at least interdependent- tax bases. Inevitably then, tax policies of one government have an impact on revenues raised by other governments, as well as on the welfare of residents living in other states.[1] Now, whether strategically or unwittingly, governments often ignore such effects of own taxation on other parties. Misjudging the 'true' social cost of public provision, policy makers are then thrown off by *tax externalities* making their decisions, leading to an inefficient mix of public and private consumption. In this chapter we revisit these inefficiencies and their implications, which come in many forms.

As soon as tax bases can move freely between states for one, *horizontal* externalities bring about suboptimal levels of taxation and public provision. The textbook 'race to the bottom' scenario is often given as the example here,[2] although other outcomes are certainly possible.[3] If the tax base is co-occupied by the federal and state governments moreover, *vertical* externalities enter the fray. Here the externality works through the *shared* tax base, which may contract as states ratchet up tax levels. When the resulting negative effect on federal tax revenues is overlooked by the states, regional taxes as well as public provision come out inefficiently *high* compared to the second-best unitary optimum.[4] But when taxes are levied on an *ad valorem* basis, such vertical externalities can just as well have the opposite effect as shown by Dahlby and Wilson (2003).

Now, since both types of externalities often take effect simultaneously -and can work in opposite directions- the question becomes what their joint effect will be. Indeed, if horizontal externalities work against their vertical counterparts, the kind of welfare losses described above start wearing thin or could even fully cancel out. First tackled by Keen and Kotsogiannis (2002) with a focus on capital mobility, tax base elasticities and the relative size of the federal government are shown to be crucial factors in this trade-off.[5] Building on these results, Brülhart and Jametti (2006) employ a novel empirical specification confirming vertical externalities as observables, rather than "theoretical curiosities". Devereux et al. (2007) focus on excise taxation lastly, laying out the impact of cross-border shopping and transport costs. What remains is the labour market, marking a blind spot as yet overlooked in the literature: the effect of inter-state *commuting*, which is where our theoretical contribution comes in.

Unlike most work so far, which targets inter-state *migration* as the main driver of horizontal labour tax externalities,[6] we thus move beyond the standard approach geared towards Anglo-Saxon countries. Consider for example a situation with Belgian features,

[1] Throughout this paper, we focus on the regional tier within a federation: the *state* level. Our findings would also apply to the municipal level however, even within a two-tiered unitary state.

[2] Wilson (1986) and Zodrow and Mieszkowski (1986) provide the seminal formal derivations. For a survey of the empirical literature on horizontal interactions, see Brueckner (2003).

[3] Keen and Marchand (1997) discuss the overprovision of *productive* inputs. McLure (1967), Krelove (1992) or De Borger et al. (2007) consider tax exporting. See also Lockwood (2001) for a theoretical synthesis of commodity tax competition.

[4] See Dahlby (1996), Boadway and Keen (1996) or Boadway et al. (1998) for the theoretical outlay. Esteller-Moré and Solé-Ollé (2001) and Andersson et al. (2004) deliver empirical evidence.

[5] Wilson and Janeba (2005) add one more dimension, by endogenising the *degree* of decentralisation in Nash-competing federations.

[6] The last section in Boadway and Keen (1996) forms a notable exception, yet also here clearing *regional* labour markets drive wage formation.

where at least 10% of the workforce commutes between states but households as a whole rarely migrate.[7] Since there is no inter-state migration in this case, relying on the conventional framework would only leave us with vertical externalities to analyse.[8] In our model on the other hand, horizontal policy effects are re-introduced through commuting flows. Wages are set in an integrated, national labour market, so that changes in one state are felt throughout the entire federal system even when household *migration* does not occur. In other words, we translate the model of Keen and Kotsogiannis (2002) to a labour market setting. Yet by introducing an ad valorem, residence based labour tax, we also extend it. Indeed, when the commuting labour force is taxed *ad valorem*, vertical externalities can be of either sign which re-routes the interplay with their horizontal antagonists. This adds to the novelty of our results.

First, a fully decentralised labour tax brings about inefficiently low levels of taxation and public provision. An outcome which chimes well with the conventional -tax competition- reasoning. But in stead of attracting more workers by lowering taxes, the underlying political motives run differently here. Although states in our model are out to boost tax revenues as well, their attention limits itself to *own* workers. Since labour is taxed in the state of residence, revenues can only go up if own-residents are induced to work more, and work incentives of non-residents are impeded. A lower state tax rate, implying a higher net wage in the tax-raising state and lower net wages in other states, achieves exactly this. This for the simple reason that in an integrated labour market, local tax cuts pull the market clearing gross wage down. The fact that taxation is ad valorem only amplifies this mechanism.

When the tax base is co-occupied by the federal and state governments secondly, either under- or overtaxation can occur. Our model identifies conditions for states to overtax, i.e. for the overall fiscal externality to be negative. Low labour supply elasticities combined with highly responsive labour demand nation-wide turn out to be crucial here. However, the latter effect comes out less pronounced because taxation is ad valorem, which brings in the impact of the federation-wide gross wage increase on revenues. In other words, when taxation is levied ad valorem, undertaxation is more liable to occur than under unit taxation. For the same underlying reasons, fiscal equalisation is also expected to give less cause for overtaxation.

The remainder of this chapter is organized as follows. Section 1 provides some background on the literature and elaborates on the key ideas. Section 2 introduces the model. Employing a unitary country focus, the second-best optimum is characterised in section 3. The following section then allows for varying degrees of decentralisation, so that externalities enter the fray. In section 4.1, state governments have full discretion over fiscal policy, whilst in section 4.2 the federal and state governments share the labour tax base. Section 5 introduces a fiscal equalisation mechanism, after which section 6 concludes.

[7] More commuting examples; Switzerland, Germany,..? Case can be made for all densely populated federations.

[8] In fact, and to the best of our knowledge, only Boadway and Keen (1996) have considered *both* horizontal and vertical externalities caused by labour induced migration in a coherent framework.

I. Background

Making sense of locational inefficiencies in labour markets, theoretical approaches usually have one thing in common: workers only supply labour in their own state of residence, and not in other states.[9] Fitted to Anglo-Saxon conditions on the ground, such models then tend to favour inter-state *migration* over commuting as a primary focus of analysis. As argued above, this perspective has its limits in the absence of household migration, which is where our model fills in the blanks. If migration flows were to dry up entirely, commuting still allows for horizontal bias. Rather surprisingly moreover, our approach serves the exact same purpose at the other end of the mobility-spectrum, where migration is perfectly mobile.

Indeed, when migration is fully costless and unobstructed by other barriers, the kind of horizontal externalities described earlier are fully internalised, resulting in an optimal mix of public and private consumption. The only precondition here, as Boadway (1982) was first to point out, is that governments take the actual *migration responses* to policy shifts into account. The standard optimisation of own-residents' utility, *given* that residents themselves keep moving until their utility is equalised across all states, is what drives this result. State governments in fact anticipate the equalising effect of in- or outward migration on utility, so that every single worker in the federation pops up in their welfare function, and not just the ones living within state borders. In other words, fully mobile workers drive Nash-competing governments to self-interested cooperation. Because of this 'incentive equivalence' -we owe the term to Myers & Papageorgiou (1993)- each state indirectly satisfies its Samuelson condition after all, much like a social planner would.

Also under a scenario of perfect mobility consequently, where horizontal externalities would otherwise cancel out because of such 'incentive equivalence', does our model offer a wider scope. As was the case when migration didn't take place at all, the optimal mix of public and private consumption can still be thrown off by allowing for commuting. When household mobility is less than fully perfect however, the optimal outcome can be distorted by other factors as well. Mansoorian and Myers (1993) e.g. were first to analyse the effect of varying degrees of home attachment -leading to costly migration- on government decision making. Wellisch (2000) goes on to discuss alternative relaxations to the initial model, under which 'incentive equivalence' is expected to break down as well. In all of these cases however, commuting still has a potentially amplifying part to play among more conventional distortions.

An argument which holds just the same when it comes to other kinds of locational inefficiencies, due for example to the inefficient allocation of households across states. But since such welfare losses are directly caused by *individuals* -as described by Buchanan and Goetz (1972) or Flatters et al. (1974), and recently revisited by Albouy (2012)- rather than by inefficient government decision making, they are of lesser interest here.

[9] This work ranges from the seminal contributions of Buchanan and Goetz (1972), Flatters et al. (1974), and Boadway and Flatters (1982), to Boadway and Keen (1996), Wildasin (1991, 2000), Richter and Wellisch (1996), Wellisch and Wildasin (1996) and Sato (2000).

II. The Model

Our federal economy consists of a limited number $n > 1$ of states, where ad valorem taxes are levied on labour incomes of *immobile* households. Importantly, and although their *residence* is fixed as a result, members of each household are free to work in any other state of the federation. To simplify notation as much as possible, we normalize the mass of households in each state to unity.[10]

Output in each state is given by technology $F_i(L_{D_i})$, where L_{D_i} denotes the amount of labour demanded by firms in state i, with $F' > 0 > F''$.[11] The private sector maximises profits, given by

$$\pi_i = F_i(L_{D_i}(w_i)) - w_i L_{D_i}(w_i) \tag{1}$$

with w_i the gross wage in state i. As a result, labour demand L_{D_i} is implicitly defined by $F_i'(L_{D_i}(w_i)) = w_i$, with $L'_{D_i}(w_i) = \frac{1}{F''} < 0$. Production is used for private as well as public consumption, with a marginal rate of transformation of 1. Profit taxes $\theta_{(i)}$, when included, are exogenously fixed. Turning to the consumer side of our economy, the representative household of state i derives utility from private consumption, public provision, and leisure. That is,

$$U_i(C_i, L_{S_i}, G_i, G_i^F) = u_i(C_i, L_i) + \Gamma_i(G_i, G_i^F) \tag{2}$$

with C_i the consumption of a composite (numeraire) private good, L_{S_i} labour supply, and $G_i^{(F)}$ state and federal publicly provided *private* goods.[12] Sub-utility $u_i(C_i, L_{S_i})$ is concave, increasing in C_i and decreasing in L_i. $\Gamma_i(G_i, G_i^F)$ is concave and increasing in both G_i and G_i^F respectively. As in Kotsogiannis and Martínez (2008), public provision is financed by an ad valorem tax on labour income, which is levied according to the *residence* principle. Denoted by t_i for the states and T for the federal government, the consolidated labour tax for state i becomes $\tau_i = t_i + T$.

Each household then maximises (2) subject to its budget constraint: $C_i = \bar{w}_i L_i + \pi_i$, with $\bar{w}_i = (1 - \tau_i) w_i$ the net wage. As a result, labour supply $L_{S_i}(\bar{w}_i)$ is implicitly defined by $u_{C_i}(.)\bar{w}_i + u_{L_i} = 0$ and assumed increasing, so that $L'_{S_i}(\bar{w}_i) > 0$. Indirect utility is then given by

$$V_i(\bar{w}_i, \pi_i, G_i, G_i^F) = v_i(\bar{w}_i, \pi_i) + \Gamma_i(G_i, G_i^F) \tag{3}$$

Crucially, households in each state i are immobile, but are free to work in any state of their choosing. A common inter-state labour market thus allows for commuting, where labour itself is costlessly mobile and commuting flows equilibrate *gross* wages across all states. Since state populations are normalised to 1, this implies that states where labour supply L_{S_i} outweighs labour demand L_{D_i} are marked by commuting outflows, and vice versa.[13]

[10] A common assumption in the literature focusing on similar efficiency issues, see e.g. Keen and Kotsogiannis (2002, 2004), Lucas (2004), Brülhart and Jametti (2006), Aronsson and Blomquist (2008) or Kotsogiannis (2010).

[11] A subscript denotes the derivative of a function of several variables whereas a prime denotes the derivative of a function of one variable.

[12] Note that since we have normalised population to 1, these could just as well be pure public goods as in Kotsogiannis and Martínez (2008).

[13] Here, normalised state populations keep us from introducing a commuting cost. Since such a cost would not change our core results however, we can safely omit it. Also, commuting from one state to another will -at least in the kind of dense population settings we are targeting here- often be as costly as commuting from one city to another within the same state.

Denoting the n-vector of consolidated tax rates by $\boldsymbol{\tau} \equiv (\tau_1, ..., \tau_n)$, the gross wage $w(\boldsymbol{\tau})$ clearing the common labour market is implicitly defined by

$$\sum_i^n L_{S_i}(\bar{w}_i(\boldsymbol{\tau})) = \sum_i^n L_{D_i}(w(\boldsymbol{\tau})) \tag{4}$$

Taking the total differential with respect to τ_i of (4) yields[14]

$$\frac{\partial w}{\partial \tau_i} = \frac{w\eta_i \frac{L_{s_i}}{\bar{w}_i}}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)} > 0 \tag{5}$$

with $\eta_i > 0$ labour supply, and $\varepsilon_i < 0$ labour demand elasticities in state i. Throughout the analysis we confine most of our attention to symmetric equilibria, in which all states set the same tax rate ($t_i = t, \forall i$). The gross wage in such an equilibrium then becomes $w(\boldsymbol{\tau}) \equiv w(\tau, ..., \tau)$, with

$$w'(\boldsymbol{\tau}) = \frac{w\eta}{(1-\tau)(\eta - \varepsilon)} > 0 \tag{6}$$

So that, still in symmetric equilibrium and using (5), we also get

$$\frac{\partial w}{\partial \tau_i} = \frac{1}{n} w'(\boldsymbol{\tau}) \tag{7}$$

Logically, since marginally increasing the *common* state tax rate τ has a federation-wide impact, the gross wage will respond differently compared to a state-specific tax hike. Indeed, from (7) we find the latter to be smaller. As the gross wage starts to rise in the state raising its taxes, more and more workers from other states will flock to this region, mitigating the gross wage increase. When n were to go to infinity, the gross wage effect will be fully countered by the commuting response, as can be seen in expression (5) or (7). The marginal tax burden then falls entirely on the state in question through the drop in net wages, as $\frac{\partial \bar{w}_i}{\partial \tau_i} = (1-\tau_i)\frac{\partial w}{\partial \tau_i} - w = -w$.
The effects of taxation on net wages can then be written as

$$\bar{w}'(\boldsymbol{\tau}) = \frac{w\varepsilon}{(\eta - \varepsilon)} < 0 \tag{8}$$

and

$$\frac{\partial \bar{w}}{\partial \tau_i} = \frac{w\left(n\varepsilon - (n-1)\eta\right)}{n(\eta - \varepsilon)} < 0 \tag{9}$$

Lastly, the effect of marginally increased taxation on profits -which is the same for state as well as federal taxation- is given by

$$\frac{\partial \pi_i}{\partial \tau_{(i)}} = \frac{\partial\left(F_i(L_{D_i}) - wL_{D_i}\right)}{\partial \tau_{(i)}} = -L_{D_i}\frac{\partial w}{\partial \tau_{(i)}} < 0 \tag{10}$$

[14] All derivations of the wage effects used throughout the paper are given in appendix A.

III. Second-best optimum in a 'unitary' country

We start with the benchmark case of a unitary country, where states are given no taxing or spending powers and the federal level makes all the calls. Here, the federal government sets a uniform tax rate τ to finance consolidated public provision. Since in this case tax externalities will not arise, the second-best outcome under distortionary taxation is attained.

We furthermore assume the federal government can tailor public provision ($G_i^{(F)}$) to the preferences of the representative household living in each state i, thus ruling out inefficiencies at the federal level working through *policy* uniformity. Other arguments in favour of more decentralisation described in chapter 1, such as enhanced accountability, cost-effectiveness or innovation, are also omitted. This way, the second-best optimum as we describe it in what follows, will always be welfare superior to the decentralised outcomes discussed in the following sections. It hence serves as the ideal benchmark. Moreover, inter-state spill-overs of public provision are excluded as well, keeping the focus strictly on *tax* externalities as a source of lower-level inefficiency.

Keeping things simple, we will assume states to be perfectly symmetric. Also, profits accrue entirely to the representative household living in the state where rents are realised. In section VI of the paper both assumptions will in turn be relaxed. Since we do not deal with redistributional issues lastly, and consider just one identical representative household per state, the federal government simply maximises a Utilitarian welfare function given by

$$Max_{G_i,G_i^F,\tau} \sum_i^n V_i\left(\bar{w}_i(\boldsymbol{\tau}), \pi_i(\boldsymbol{\tau}), G_i, G_i^F\right) \tag{11}$$

subject to its budget constraint given by

$$\sum_i^n \left(G_i + G_i^F\right) = \tau \sum_i^n L_{S_i}\left(\bar{w}_i(\boldsymbol{\tau})\right) w(\boldsymbol{\tau}) \tag{12}$$

where the values of τ, G_i^F and G_i are chosen by the government. Using symmetry, (6), and (12), the first-order conditions readily reduce to[15]

$$\frac{\frac{\partial V}{\partial G^F}}{\lambda} = \frac{\frac{\partial V}{\partial G}}{\lambda} = \frac{1}{\left(1 - \frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon - \eta}\right)} = MCPF \quad \forall\, i = 1, .., n \tag{13}$$

with λ the marginal utility of income. Equations (13), together with the budget constraint (12), characterize the second-best optimum denoted by (τ^*, G^*, G^{F*}) for each state i. At the unitary optimum, τ is set such that the marginal rate of substitution (MRS) between both the publicly provided good and the private good must be equal to the Marginal Cost of Public Funds (MCPF). As is well known, the MCPF is the efficiency cost of raising revenue with a distortionary tax (Dahlby, 2008). Indeed, equations (13) are the usual optimality conditions for public provision in a distorted economy, being simplified versions of the standard Atkinson and Stern (1974) rule.[16]

[15] The optimisation problem is written out, and solved, in appendix B.

[16] The Atkinson-Stern rule includes an additional term in the numerator, capturing the effect of changes in public provision on tax revenues. Because utility is additively seperable, this relation does not bite in our model.

IV. Decentralised taxation

We now consider various degrees of decentralised fiscal decision making, setting the stage for tax externalities to come about. To fully understand the interplay of horizontal and vertical externalities, we first study the horizontal interaction in isolation. A setting where taxation is *fully* decentralised -and vertical externalities are as a result assumed away- then serves as the logical stepping stone to the more complex shared tax base case.

IV.1. Full tax autonomy

Here, fiscal decision making is fully decentralised to the state level, and each state i levies a labour tax t_i to finance the publicly provided goods G_i and G_i^F. As a result, only horizontal externalities can distort tax and spending decisions. We assume that all states take decisions made by other states as given, and thus behave as Nash competitors. The government of a given state i then maximises the indirect utility of the representative household living within state borders

$$Max_{G_i,G_i^F,\tau_i} V_i\left(\bar{w}_i(\tau_i,\boldsymbol{\tau}),\pi_i\left(w(\boldsymbol{\tau})\right),G_i,G_i^F\right) \tag{14}$$

subject to state i's budget constraint, given by

$$G_i + G_i^F = \tau_i L_{S_i}\left(\bar{w}_i(\tau_i,\boldsymbol{\tau})\right) w(\boldsymbol{\tau}) \tag{15}$$

Using (5) and (15), and evaluated in symmetric equilibrium, we obtain the following characterisation:

Lemma 1 *With households immobile, but workers commuting to other states, the MCPF of a state government with full tax autonomy is given by*

$$\frac{\frac{\partial V}{\partial G^F}}{\lambda} = \frac{\frac{\partial V}{\partial G}}{\lambda} = \frac{1}{\left(1-\frac{\tau\eta(n\varepsilon-(n-1)\eta)}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)} = MCPF_i \tag{16}$$

Equations (16), together with the budget constraint (15), define the decentralised Nash equilibrium denoted by (τ^*,G^*,G^F).[17]

Also in the fully decentralised case in other words, does public provision continue until the marginal rate of substitution between both the publicly provided good and the private good is equal to the marginal cost of public funds (MCPF). However, this efficiency cost overlooks all effects of own taxation on other states. The cost of raising revenue as it is perceived by the state government, therefore weighs in more *or* less than the socially relevant cost derived in section 3. To verify the sign of this bias, we compare the unitary MCPF which was unaffected by externalities, to the state MCPF captured by lemma 1. The resulting trade-off, evaluated at the symmetric Nash-equilibrium, then gives us

Proposition 1 *When state governments have full tax autonomy and workers commute, positive tax externalities lead to inefficiently low levels of taxation and public provision.*

[17] Proofs of both lemma 1 and proposition 1 are given in appendix C.

Or, to put it otherwise, the state MCPF will be biased upwards. Because the effect of state taxation on other states is not included in the state welfare cost, and because this effect is proven to be positive, state governments perceive the cost of distortionary taxation to be *higher* than it actually is. As a result, state taxation as well as public provision are set at inefficiently low levels compared to the second-best outcome derived in section III. Consequently, if the state government were to internalise the positive externality in its welfare cost calculations, the cost would weigh in precisely at its socially relevant level.[18]

Now, to understand all of the effects underpinning proposition 1, we re-write the marginal cost of public funds derived under lemma 1 in its most basic form

$$MCPF_i = -\frac{\frac{\partial V_i(\bar{w},\pi_i,G_i,G_i^F)}{\partial \tau_i}}{\left(\frac{\partial R_i(\tau_i,w)}{\partial \tau_i}\right)} \frac{1}{\lambda_i} \tag{17}$$

What this rewritten efficiency cost of state taxation shows, is that the positive externality comes two-sided. In the denominator of (17) the impact of a tax hike in region i on the tax revenues $\sum_{j\neq i}^{n} R_j$ of other states is left out. Whilst the numerator omits the *direct* effects on non-resident welfare $\sum_{j\neq i}^{n} V_j$ as well. Starting with the 'indirect' revenue effects in the denominator, these can be written as

$$\sum_{j\neq i}^{n} \frac{\partial R_j}{\partial \tau_i} = \sum_{j\neq i}^{n} \left(\overbrace{\tau_j L_{S_j} \frac{\partial w}{\partial \tau_i}}^{Wage\uparrow} + \overbrace{\tau_j w \frac{\partial L_{S_j}}{\partial \bar{w}_j} \frac{\partial \bar{w}_j}{\partial \tau_i}}^{Labour\ supply\uparrow} \right) > 0 \tag{18}$$

What we learn from (18), and contrary to a scenario employing unit taxation, is that state taxation has two effects on the budgets of other states. Due to the gross wage increase firstly, which is identical across states because of commuting, tax bases in other states rise alongside collected tax revenues. Second, since other states keep their tax rates constant, net wages in these states follow gross wage hikes so that more labour is supplied. This also fattens state coffers, and would be the only effect under unit taxation.

Turning to the denominator of (17), the omitted direct impact of own taxation on non-resident welfare is given by

$$\sum_{j\neq i}^{n} \frac{\partial V_j}{\partial \tau_i} = \sum_{j\neq i}^{n} \frac{1}{\lambda_j} \left(\overbrace{L_{S_j} \frac{\partial \bar{w}_j}{\partial \tau_i}}^{Consumption\uparrow} - \overbrace{L_{D_j} \frac{\partial w}{\partial \tau_i}}^{Profits\downarrow} \right) \tag{19}$$

Here we find that higher net wages not only improve non-resident welfare through increased public provision, as before in (18), but also simply because purchasing power comes out reinforced. Furthermore, all three effects combined appear to dominate the negative direct effect on non-resident welfare, which is the drop in collected profits due to the higher gross wage.

In addition, the higher the amount of regions in the federation, the stronger the positive externality and subsequent welfare losses. A result reminiscent of Hoyt (1991) or Keen and

[18] This point is illustrated formally in appendix C.

Kotsogiannis (2004). This can also easily be seen in (58), where a higher n relatively nudges up the left hand side of the expression.

Looking at these inefficiencies from a more strategic angle lastly, the political incentives behind proposition 1 are rather different from other interpretations in the literature. Indeed, by setting taxes at inefficiently low levels, state governments aren't out to attract more workers as they would in the standard competition case. Contrarily, their aim is to boost labour supplied by *own* residents, and impede work incentives of non-residents. Reducing state taxation fits that purpose: the net wage of own residents will increase and provide a boost to own labour supply. Whilst the net wage of non-residents goes down together with overall gross wages, undercutting non-resident willingness to work.

Nonetheless, since all states are identical and follow the same strategy, we arrive at a familiar outcome: a sub-optimal welfare level where all taxes are set too low.

IV.2. Shared tax base

We now move on to the case where both the federal and state governments levy taxes on labour, giving rise to horizontal *and* vertical tax externalities at the same time. The federal government now sets a uniform tax rate T to tailor public provision (G_i^F) to the preferences of each representative household, thus maximising (11) subject to:

$$\sum_i^n G_i^F = T \sum_i^n L_{S_i}\left(\bar{w}_i(\tau_i, \boldsymbol{\tau})\right) w(\boldsymbol{\tau}) \tag{20}$$

On top of this, each state raises additional taxes to finance own public provision (G_i). Following the literature, we assume the federal government sees through states' budget constraints, so there will be no 'top-down' vertical externality. We also continue to assume that all governments behave non-cooperatively -taking each other's policies as given- and that states are identical in every way. Since under these assumptions federal decision making is equivalent to the unitary outcome derived in section III, we can jump straight to the state level. The government of state i now maximises (14) subject to

$$G_i = t_i L_{S_i}\left(\bar{w}_i(\tau_i, \boldsymbol{\tau})\right) w(\boldsymbol{\tau}) \tag{21}$$

Using (7), 9 and (21), and evaluated in symmetric equilibrium, we arrive at the following characterisation:

Lemma 2 *With households immobile, but workers commuting to other states, the MCPF of a state government sharing its labour tax base with the federal government, is given by*

$$\frac{\frac{\partial V}{\partial G}}{\lambda} = \frac{1}{\left(1 - \frac{t\eta(n\varepsilon-(n-1)\eta)-T\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)} = MCPF_i \tag{22}$$

Expression (22), together with the budget constraint (21), defines the Nash equilibrium, denoted by (t^*, G^*).[19]

[19] Proofs of lemma 2 and proposition 2 are given in appendix D.

Again, public provision continues until the marginal rate of substitution between both the publicly provided good and the private good is equal to the marginal cost of public funds (MCPF). Moreover, not only does this efficiency cost overlook all (positive) horizontal effects of own taxation on other states, it also fails to internalise the *vertical* effects.

Now, as touched upon in the introduction, these vertical effects can be positive as well as negative. Negative, since a state tax hike induces a higher federation-wide gross wage, which curtails federation-wide labour demand. And since the labour market clears at the national level, this brings about lower federation-wide labour supply, and by consequence, shrinking federal tax receipts. However, and crucially, since the labour income tax base is taxed ad valorem, the same gross wage increase also boosts the tax base. This then has a positive effect on federal coffers.

In short, the positive horizontal externalities established in the previous section can be hollowed out, or reinforced, by their vertical counterparts. To find out which effect comes out on top, and thus to determine the sign of the *overall* externality, we again compare the unitary MCPF with the state efficiency cost. Evaluated in the symmetric Nash-equilibrium, using lemma 2 and (13), we obtain:

Proposition 2 *In a federation where workers can commute between identical states, where households as a whole do not migrate, and both the federal and state governments tax labour income, the overall externality:*

1. *Is positive when labour supply is highly responsive to changes in the net wage, and federation-wide labour demand is inelastic*
2. *Is negative when labour supply is relatively inelastic compared to labour demand, and the federal level accounts for the larger part of public provision*
3. *Disappears as soon as labour supply is unaffected by the net wage*

When the overall externality is negative, state over-taxation and over-provision ensue. In this case the upward pressure on the perceived state efficiency cost is more than undone by the vertical effects, which turn out to be negative. If, on the other hand, the state MCPF stays well above the second-best efficiency cost, the positive effects -horizontal as well as possibly vertical- on non-resident welfare will have prevailed. We then end up with under-provision and under-taxation, similar to the outcome under full tax autonomy in the previous section.

Earlier results by Keen and Kotsogiannis (2002) thus seem to carry through in our setting here, where vertical externalities can turn either way. Zoom in on the underlying factors driving proposition 2 however, and an important difference emerges. Although tax base elasticities play a crucial part in our setting as well, ad valorem taxation renders their interplay less clear cut, bringing about *less* overtaxation in general.

To shed more light on this mechanism, and using (1), (3), (4) and (21), we write the welfare of the representative household in state i as

$$W_i(t_i, T, \boldsymbol{\tau}) = v_i\left(w(\boldsymbol{\tau}), \tau_i, \pi_i(w(\boldsymbol{\tau}))\right) + \Gamma_i\left(t_i L_S((\tau_i, w(\boldsymbol{\tau}))w(\boldsymbol{\tau}), G_i^F(T)\right) \tag{23}$$

Reformulating state optimisation expressed by (14) and (21) above, state i then chooses its tax rate t_i to maximize (23), taking all other tax rates as given. The necessary first

order condition for this, evaluated in symmetric equilibrium and making use of (7) and (10), becomes

$$\frac{\partial W_i}{\partial t_i} = \lambda \left[L_S \left((1-\tau)\frac{w'}{n} - w \right) - L_D \frac{w'}{n} \right] + \Gamma_G \left[tL_S' \left((1-\tau)\frac{w'}{n} - w \right) w + tL_S \frac{w'}{n} \right] \quad (24)$$

Condition (24) is nothing more than lemma 2 in rewritten form, implicitly defining the equilibrium state tax rate t^*, again given T^* set by the federal government and the number of states n.[20] The first term of (24) reflects the *direct* utility loss incurred by the representative household because of increased state taxation, working through the decreasing net wage and falling profits. The second term, involving Γ_G, expresses the effect of a state tax hike on state government revenues. Crucially, since taxes are levied ad valorem, this could either inflict utility losses *or* gains. A tax hike depresses the net wage and consequently labour supply, but also pushes up the federation-wide gross wage, so the effect on revenues is ambiguous.

Now, to bring potential tax externalities to the surface by means of (24), we first write out welfare in symmetric equilibrium and under full information. Using (1), (3), (4), (20) and (21), we obtain[21]

$$W(t, T, \boldsymbol{\tau}) = v\left(w(\boldsymbol{\tau}), \tau, \pi(w(\boldsymbol{\tau}))\right) + \Gamma\left(tL_S((\tau, w(\boldsymbol{\tau}))w(\boldsymbol{\tau}), TL_D(w(\boldsymbol{\tau}))w(\boldsymbol{\tau})\right) \quad (25)$$

Differentiating (25) with respect to the common tax rate t, which is equivalent to a coordinated tax increase in all states, then yields

$$W_t = \lambda \left[L_S\left((1-\tau)w' - w\right) - L_D w' \right] + \Gamma_G \left[tL_S'\left((1-\tau)w' - w\right) w + tL_S w' \right] + \Gamma_{G^F} \left[TL_D' w' w + TL_D w' \right] \quad (26)$$

where the third term reflects the impact of state taxation on welfare through Γ_{G^F} and federal tax revenues. Now, since setting (26) equal to zero implicitly defines the socially optimal state tax rate, the sign of W_t evaluated at the Nash-equilibrium established in (24) is vital. If W_t turns out to be positive at this point, a slight increase in state tax rates would improve overall welfare. Hence, tax rates in the non-cooperative equilibrium defined by (24) -or lemma 2- were set inefficiently low from a social viewpoint. Conversely, when W_t is negative, state taxes were set too high. To investigate the sign of (26) at the Nash equilibrium (t^*, T^*) we subtract (24) from (26), so that

$$W_t = \overbrace{\left[-\lambda\tau L_D + t^*\Gamma_G \left(L_S'(1-\tau)w + \underbrace{L_S}_{(+)} \right) \right] \left(1 - \frac{1}{n} \right) w'}^{Horizontal\,externality} + \overbrace{T^*\Gamma_{G^F} \left[wL_D' + \underbrace{L_D}_{(+)} \right] w'}^{Vertical\,externality} \quad (27)$$

Since w' is unambiguously positive, the sign of (27) hinges on the terms between square brackets. On the left we find the effects of state taxation on other states, reflecting the horizontal externalities. As described above, these operate through non-resident purchasing

[20] Where the federal tax rate T^* comes in through the consolidated tax rate $\boldsymbol{\tau}$ defining $w(\boldsymbol{\tau})$.

[21] Note how, under symmetric equilibrium, the federal budget constraint (20) reduces to $G^F = TL_D(w(\boldsymbol{\tau})w(\boldsymbol{\tau})$, using the market clearing condition (4).

power, incurred rents and state tax revenues.[22] On the right of (27) on the other hand, the impact of a state tax hike on the federal budget is spelled out. As $L_D^{'} < 0$, shrinking federation-wide labour demand comes in via $L_D^{'}w$, whilst the positive twist of the gross wage increase is expressed by L_D.

Now, we see that (27) turns positive when labour supply becomes more responsive to changes in the net wage ($L_S^{'} \uparrow$), and labour demand less sensitive to gross wage movements ($L_D^{'} \downarrow$). More so, this latter effect is strengthened by a higher state tax rate t^*, in itself a result of a higher appreciation of state public provision Γ_G. Inversely, and coming full circle on points (1) and (2) of proposition 2, inelastic labour supply and highly responsive labour demand amplify the negative effects in (27), as will a higher utility share of federal public provision Γ_{G^F}. Setting $L_S^{'} = 0$ in (6) lastly, also reduces w' and by consequence the whole of (27) to zero, as captured by point (3) of proposition 2.

So far our results chime well with Keen and Kotsogiannis (2002), even though the strict demarcation between positive horizontal, and negative vertical externalities breaks down in our setting. Where (27) diverges significantly from earlier work however, is in the degree to which the negative effects are expected to bite. Take for example a scenario where federation-wide labour demand is highly elastic, whilst labour supply is not. As a result, a state tax hike reduces the federation-wide tax base considerably, whilst commuting flows are slow to arbitrage away income differentials. Under unit taxation, W_t can therefore easily be expected to turn negative in the Nash equilibrium, with overtaxation at the state level as a result. But when the tax is levied at valorem, the overall gross wage increase has a second part to play. Marked with a (+) sign in (27), we see how w' indeed boosts tax revenues at the state and federal level, aside from its interaction with labour supply and demand responsiveness. These latter wage effects would not arise if, rather than full labour income $L_S w$ here, only labour supply L_S had been taxed. We summarise this result in the following corollary:[23]

Corollary 1 *When different levels of government levy ad valorem taxes on a shared tax base, lower-level governments are less likely to overtax than under unit taxation.*

Corollary 1 then suggests that undertaxation, underprovision and -from a strategic perspective- tax competition will more often than not hold sway over other outcomes, even when the vertical dimension is taken up in the analysis. In the next section, we investigate whether the same friction between ad valorem and unit taxation carries over to equalisation grant mechanisms.

V. Revenue equalisation and ad valorem taxation

Suppose now an equalisation mechanism is in place in our federation. Such a mechanism corrects for differences in fiscal capacity across the various states, levelling out (a degree of) the divergence. The underlying principle here is mainly one of horizontal equity, ensuring each state has sufficient revenues at its disposal to provide a minimum level of public services, without having to rely on draconian levels of taxation.[24]

[22] We assume Γ_G is such that, in line with proposition 1, these sum up to be positive.
[23] We derive the sufficient conditions for corollary 1 to hold in general, in appendix D.
[24] Efficiency arguments can also be provided in favour of equalisation, see e.g. Boadway and Shah (2009).

We assume an equalisation mechanism of the conventional sort. Here, the per capita equalisation grant ω_i to each state i is given by the difference between its fiscal capacity and a benchmark fiscal capacity, multiplied by a standard -usually average- federal tax rate:[25]

$$\omega_i = \bar{t}(\bar{B} - B_i) \tag{28}$$

where $\bar{t} = \frac{\sum_i^n t_i L_{S_i} w}{\sum_i^n L_{S_i} w}$ is the average lower-level government tax rate. B_i is the fiscal capacity of state i, captured by its tax base $L_{S_i} w$, whilst $\bar{B} = \frac{\sum_i^n L_{S_i} w}{n}$ is the benchmark fiscal capacity, being the average federation-wide tax base. Deriving the equalisation grant w.r.t. t_i gives us

$$\frac{\partial \omega_i}{\partial t_i} = \frac{\partial \bar{t}}{\partial t_i}(\bar{B} - B_i) + \bar{t}(\frac{\partial \bar{B}}{\partial t_i} - \frac{\partial B_i}{\partial t_i}) \tag{29}$$

A tax hike in state i thus influences the equalisation grants via two channels: through the change in the average state tax rate (given fiscal capacity), and through a change in the actual fiscal capacities. In symmetric equilibrium, (29) becomes

$$\frac{\partial \omega_i}{\partial t_i} = -t_i \left(\frac{n-1}{n}\right) \left(\overbrace{w L_S' \frac{\partial \bar{w}}{\partial \tau_i}}^{(-)} + \overbrace{L_S \frac{\partial w}{\partial \tau_i}}^{(+)} \right) \lessgtr 0 \tag{30}$$

Now, whilst under unit taxation the equalisation grant would only compensate for the negative effect of a shrinking tax base on fiscal capacity, (30) again tells a different story. Indeed, when taxation is ad valorem, a state tax hike also ratchets up the fiscal capacity through the gross wage effect. The equalisation grant flowing back to the state increasing its tax rate is therefore stemmed.

VI. Concluding remarks

As one of the smallest federations in Europe, Belgium makes for an interesting case study. Counting only three sub-central entities (states), each with diverging preferences and productivity, it enjoys a very high degree of cross-state commuting. Zooming in on these commuting flows brings out a blind spot in the fiscal federalism literature on tax externalities, where models have workers supply labour *only* in their state of residence. For a federation with Belgian features, where at least 10% of the workforce commutes between regions, this assumption of non-integrated regional labour markets becomes difficult to maintain.

We therefore presented a theoretical model based on a *common* labour market, where wages are endogenously determined as commuting flows equilibrate wages across all states of a federation. Policy changes in one state will consequently be felt throughout the entire federal system, even when household *migration* does not occur. We thus model a situation where horizontal externalities are re-introduced to the analysis through commuting effects.

[25] Since the fiscal capacity differences are entirely equalised in the proposed scheme, we have assumed 'full equalisation' here. Partial equalisation however, where only a fraction of the fiscal capacity divide is bridged, yields similar -but logically less pronounced- results.

In other words, we translated the model of Keen and Kotsogiannis (2002) to a labour market setting. Yet by introducing an ad valorem, residence based labour tax, we also extend it. Indeed, when the commuting labour force is taxed *ad valorem*, vertical externalities can be of either sign which re-routes the interplay with their horizontal antagonists. This adds to the novelty of our results.

First, a fully decentralised labour tax brings about inefficiently low levels of taxation and public provision. An outcome which chimes well with the conventional -tax competition- reasoning. But in stead of attracting more workers by lowering taxes, the underlying political motives run differently here. Although states in our model are out to boost tax revenues as well, their attention limits itself to *own* workers. Since labour is taxed in the state of residence, revenues can only go up if own-residents are induced to work more, and work incentives of non-residents are impeded. A lower state tax rate, implying a higher net wage in the tax-raising state and lower net wages in other states, achieves exactly this. This for the simple reason that in an integrated labour market, local tax cuts pull the market clearing gross wage down. The fact that taxation is ad valorem only amplifies this mechanism.

When the tax base is co-occupied by the federal and state governments secondly, either under- or overtaxation can occur. Our model identifies conditions for states to overtax, i.e. for the overall fiscal externality to be negative. Low labour supply elasticities combined with highly responsive labour demand nation-wide turn out to be crucial here. However, the latter effect comes out less pronounced because taxation is ad valorem, which brings in the impact of the federation-wide gross wage increase on revenues. In other words, when taxation is levied ad valorem, undertaxation is more liable to occur than under unit taxation. For the same underlying reasons, fiscal equalisation is also expected to give less cause for overtaxation.

Appendix A. Wage effects

A1. State taxation

We obtain the tax shifting formula from the equilibrium in the labour market:

$$\sum_{i}^{n} L_{S_i}(\bar{w}_i(\boldsymbol{\tau})) = \sum_{i}^{n} L_{D_i}(w(\boldsymbol{\tau})) \tag{31}$$

Taking the total differential with respect to t_i of the labour market equilibrium condition then yields

$$\sum_{j\neq i}^{n}\left(\frac{\partial\left(L_{S_j}(\bar{w_j})\right)}{\partial\bar{w_j}}\frac{\partial\bar{w_j}}{\partial\tau_i}\right)+\frac{\partial L_{S_i}(w_i)}{\partial\bar{w_i}}\frac{\partial\bar{w_i}}{\partial\tau_i}=\sum_{i}^{n}\left(\frac{\partial\left(L_{D_i}(w)\right)}{\partial w}\right)\frac{\partial w}{\partial\tau_i}$$

keeping in mind that $\frac{\partial\tau_i}{\partial t_i}=1$ and $\frac{\partial\tau_j}{\partial t_i}=0\,\forall j\neq i$. Rewriting net wages in terms of the gross wage and solving further gives:

$$\sum_{j\neq i}^{n}\left(L'_{S_j}\frac{\partial\left((1-\tau_j)w\right)}{\partial\tau_i}\right)+L'_{S_i}\frac{\partial\left((1-\tau_i)w\right)}{\partial\tau_i}=\left(\sum_{i}^{n}L'_{D_i}\right)\frac{\partial w}{\partial\tau_i} \tag{32}$$

Now, since we know that:

$$\frac{\partial\left((1-\tau_i)w\right)}{\partial\tau_i}=(1-\tau_i)\frac{\partial w}{\partial\tau_i}-w$$

We also know that:

$$\partial\left((1-\tau_i)w\right)=(1-\tau_i)\partial w-w\partial\tau_i \tag{33}$$

And:

$$\frac{\partial\left((1-\tau_j)w\right)}{\partial\tau_i}=(1-\tau_j)\frac{\partial w}{\partial\tau_i}$$

So that:

$$\partial\left((1-\tau_j)w\right)=(1-\tau_j)\partial w \tag{34}$$

Plugging (33) and (34) into (32) then gives us:

$$\sum_{j\neq i}^{n}\left(L'_{S_j}\left((1-\tau_j)\partial w\right)\right)+L'_{S_i}\left((1-\tau_i)\partial w-w\partial\tau_i\right)=\left(\sum_{i}^{n}L'_{D_i}\right)\partial w$$

$$\left(\sum_{j\neq i}^{n}\left(L'_{S_j}(1-\tau_j)\right)+L'_{S_i}(1-\tau_i)-\sum_{i}^{n}L'_{D_i}\right)\partial w=L'_{S_i}w\partial\tau_i$$

$$\frac{\partial w}{\partial\tau_i}=\frac{wL'_{S_i}}{\left(\sum_i^n\left(L'_{S_i}(1-\tau_i)\right)-\left(\sum_i^n L'_{D_i}\right)\right)}>0 \tag{35}$$

Rewriting (36) in terms of regional labour supply elasticity $\eta_i=L'_{S_i}\frac{\bar{w}_i}{L'_{S_i}}$ and labour demand elasticity $\varepsilon_i=L'_{D_i}\frac{w}{L'_{D_i}}$ now yields:

$$\frac{\partial w}{\partial\tau_i}=\frac{w\eta_i\frac{L_{s_i}}{\bar{w}_i}}{\left(\sum_i^n\left((1-\tau_i)\eta_i\frac{L_{s_i}}{\bar{w}_i}\right)-\left(\sum_i^n\varepsilon_i\frac{L_{D_i}}{w}\right)\right)} \tag{36}$$

We know that, in equilibrium, $\frac{\partial \bar{w}_i}{\partial \tau_i} = (1-\tau_i)\frac{\partial w}{\partial \tau_i} - w$, so plugging in (36) then gives us:

$$\frac{\partial \bar{w}_i}{\partial \tau_i} = \frac{(1-\tau_i)\, w L'_{S_i} - w\left(\sum_i^n \left(L'_{S_i}(1-\tau_i)\right) - \left(\sum_i^n L'_{D_i}\right)\right)}{\left(\sum_i^n \left(L'_{S_i}(1-\tau_i)\right) - \left(\sum_i^n L'_{D_i}\right)\right)}$$

$$\frac{\partial \bar{w}_i}{\partial \tau_i} = \frac{w\left(\left(\sum_i^n L'_{D_i}\right) - \sum_{j\neq i}^n \left(L'_{S_j}(1-\tau_j)\right)\right)}{\left(\sum_i^n \left(L'_{S_i}(1-\tau_i)\right) - \left(\sum_i^n L'_{D_i}\right)\right)} < 0 \tag{37}$$

For later purposes, we rewrite (37) in terms of regional labour supply and demand:

$$\frac{\partial \bar{w}_i}{\partial \tau_i} = \frac{w\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w}_j}\right)\right)}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)} \tag{38}$$

Under the assumption of homogeneous regions, (36) and (38) reduce to:

$$\frac{\partial w}{\partial \tau_i} = \frac{w\eta}{n(1-\tau)\left(\eta - \varepsilon\right)} \tag{39}$$

$$\frac{\partial \bar{w}}{\partial \tau_i} = \frac{w\left(n\varepsilon - (n-1)\eta\right)}{n\left(\eta - \varepsilon\right)} \tag{40}$$

A2. Federal taxation

Taking the total differential of 31 with respect to the common, federal tax rate T yields:

$$\sum_i^n \left(\frac{\partial L_{S_i}(\bar{w})}{\partial(\bar{w})}\frac{\partial \bar{w}}{\partial \tau}\right) = \sum_i^n \left(\frac{\partial L_{D_i}(w)}{\partial w}\frac{\partial w}{\partial \tau}\right)$$

Keeping in mind that $\frac{\partial \tau_i}{\partial T} = 1\ \forall i$. Rewriting net wages in terms of the gross wage and solving further finally gives us:

$$\sum_i^n \left(L'_{S_i}\frac{\partial\left((1-\tau)w\right)}{\partial \tau}\right) = \left(\sum_i^n L'_{D_i}\right)\frac{\partial w}{\partial \tau} \tag{41}$$

Now, since we know that:

$$\frac{\partial\left((1-\tau)w\right)}{\partial \tau} = (1-\tau)\frac{\partial w}{\partial \tau} - w$$

We also know that:

$$\partial\left((1-\tau)w\right) = (1-\tau)\partial w - w\partial\tau \tag{42}$$

Plugging (42) into (41) then gives us:

$$\sum_i^n \left(L'_{S_i}\left((1-\tau)\partial w - w\partial\tau\right)\right) = \left(\sum_i^n L'_{D_i}\right)\partial w$$

$$\left(\sum_i^n \left(L'_{S_i}(1-\tau)\right) - \left(\sum_i^n L'_{D_i}\right)\right)\partial w = \sum_i^n \left(L'_{S_i} w\partial\tau\right)$$

$$\frac{\partial w}{\partial \tau} = \frac{w \sum_i^n L'_{S_i}}{\left(\sum_i^n \left(L'_{S_i}(1-\tau)\right) - \left(\sum_i^n L'_{D_i}\right)\right)} \tag{43}$$

For later purposes, we rewrite (43) in terms of regional labour supply and demand:

$$\frac{\partial w}{\partial \tau} = \frac{w \sum_i^n \eta_i \frac{L_{s_i}}{\bar{w}_i}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)} > 0 \tag{44}$$

Moving on to the effect on net wages, we know that, in equilibrium, $\frac{\partial \bar{w}}{\partial \tau} = (1-\tau)\frac{\partial w}{\partial \tau} - w$. Plugging in (44) then gives us:

$$\frac{\partial \bar{w}}{\partial \tau} = \frac{(1-\tau)\, w \sum_i^n L'_{S_i} - w\left((1-\tau)\sum_i^n L'_{S_i} - \sum_i^n L'_{D_i}\right)}{\left((1-\tau)\sum_i^n L'_{S_i} - \sum_i^n L'_{D_i}\right)}$$

$$\frac{\partial \bar{w}}{\partial \tau} = \frac{w \sum_i^n L'_{D_i}}{\left((1-\tau)\sum_i^n L'_{S_i} - \sum_i^n L'_{D_i}\right)} < 0 \tag{45}$$

For later purposes, we rewrite (46) in terms of regional labour supply and demand:

$$\frac{\partial \bar{w}}{\partial \tau} = \frac{w \sum_i^n \varepsilon_i \frac{L_{D_i}}{w}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)} \tag{46}$$

Under the assumption of homogeneous regions, (44) and (46) reduce to the well known expressions for a uniform, ad valorem tax rate τ:[26]

$$\frac{\partial w}{\partial \tau} = \frac{w\eta}{(1-\tau)(\eta - \varepsilon)} \tag{47}$$

$$\frac{\partial \bar{w}}{\partial \tau} = \frac{w\varepsilon}{(\eta - \varepsilon)} \tag{48}$$

Appendix B. Second-best optimum in a 'unitary' country

With states being homogeneous, the second best optimization problem can be expressed by the Lagrangian:

$$\mathcal{L} = n\left\{V_i(\bar{w}, \pi_i, G_i, G_i^F)\right\} - \gamma\left\{n\left(G_i + G_i^F\right) - \tau n L_i w\right\}$$

Leading to the following first order conditions:

$$\frac{\partial \mathcal{L}}{\partial \tau} = n\left\{\frac{\partial V_i}{\bar{w}}\frac{\partial \bar{w}}{\partial \tau} + \frac{\partial V_i}{\pi_i}\frac{\partial \pi_i}{\partial \tau}\right\} + \gamma\left(\frac{\partial(\tau n L_i w)}{\partial \tau}\right) = 0 \tag{49}$$

$$\frac{\partial \mathcal{L}}{\partial G_i} = \frac{\partial V_i}{\partial G_i} - \gamma = 0 \quad for\, i = 1, .., n$$

$$\frac{\partial \mathcal{L}}{\partial G_i^F} = \frac{\partial V_i}{\partial G_i^F} - \gamma = 0 \quad for\, i = 1, .., n$$

[26] See e.g. Dahlby and Wilson (2003), or Kotsogiannis and Martínez (2008).

Solving both conditions yields:

$$n\left\{\frac{\partial V_i}{\bar{w}}\frac{\partial \bar{w}}{\partial \tau}+\frac{\partial V_i}{\pi_i}\frac{\partial \pi_i}{\partial \tau}\right\}+\gamma\left(nL_iw+\tau nL_i\frac{\partial w}{\partial \tau}+\tau nw\frac{\partial L_i}{\partial \bar{w}}\frac{\partial \bar{w}}{\partial \tau}\right)=0 \ (\tau)$$

$$\frac{\partial V_i}{\partial G_i}-\gamma=0 \quad for\, i=1,..,n \quad (G_i)$$

$$\frac{\partial V_i}{\partial G_i^F}-\gamma=0 \quad for\, i=1,..,n \quad (G_i^F)$$

With $nL_iw+\tau nL_i\frac{\partial w}{\partial \tau}+\tau nw\frac{\partial L_i}{\partial \bar{w}}\frac{\partial \bar{w}}{\partial \tau}$ equal to $\frac{\partial R}{\partial \tau}$, being the marginal effect on federal revenue by raising the labour tax. Substituting for γ then yields:

$$\frac{\partial V_i}{\partial G_i^F}=\frac{\partial V_i}{\partial G_i}=-\frac{n\left\{\frac{\partial V_i}{\bar{w}}\frac{\partial \bar{w}}{\partial \tau}+\frac{\partial V_i}{\pi_i}\frac{\partial \pi_i}{\partial \tau}\right\}}{\frac{\partial R}{\partial \tau}} \quad for\, i=1,..,n \tag{50}$$

Public provision in each region will thus continue until its marginal benefits equal its marginal cost, and this in terms of welfare cost as well as actual provision cost (see also Dahlby (2008)). The RHS of the equation thus expresses the marginal welfare cost in utility terms of raising an additional euro of revenue to finance public provision in region i, multiplied by the marginal cost of actual provision (1 in our case). This expression can be reformulated to arrive at the conventional MCPF expression using Roy's identity:

$$\frac{\frac{\partial V_i}{\partial G_i}}{\lambda_i}=-\frac{\left(L_i\frac{\partial \bar{w}}{\partial \tau}+\frac{\partial \pi_i}{\partial \tau_i}\right)}{L_iw+\tau L_i\frac{\partial w}{\partial \tau}+\tau w\frac{\partial L_i}{\partial \bar{w}}\frac{\partial \bar{w}}{\partial \tau}} \quad for\, i=1,..,n \tag{51}$$

Expression (51) simply states that at the unitary optimum the ad valorem tax τ is set such that the marginal rate of substitution (MRS) between both the public and the private good must be equal to the MCPF. The conventional MCPF formula is thus extracted, now in monetary terms as is usual:

$$MCPF_u=-\frac{\left(L_i\frac{\partial \bar{w}}{\partial \tau}+\frac{\partial \pi_i}{\partial \tau_i}\right)}{L_iw+\tau L_i\frac{\partial w}{\partial \tau}+\tau w\frac{\partial L_i}{\partial \bar{w}}\frac{\partial \bar{w}}{\partial \tau}}$$

With subscript u standing for unitary case. Rewriting the third term of the RHS denominator as:

$$\tau w\frac{\partial L_i}{\partial \bar{w}}\frac{\partial \bar{w}}{\partial \tau}=\tau w\frac{\partial L_i}{\partial \bar{w}}\frac{\bar{w}}{L_i}\frac{L_i}{\bar{w}}\frac{\partial \bar{w}}{\partial \tau}=\tau w\eta_i\frac{L_i}{w(1-\tau_i)}\frac{\partial \bar{w}}{\partial \tau}=\frac{\tau L_i\eta_i}{(1-\tau_i)}\frac{\partial \bar{w}}{\partial \tau}$$

And plugging in the profit effect (10), we get:

$$MCPF_u=-\frac{\left(L_S\frac{\partial \bar{w}}{\partial \tau}-L_D\frac{\partial w}{\partial \tau}\right)}{\left(L_Sw+\tau L_S\frac{\partial w}{\partial \tau}+\tau w\frac{\partial L_S}{\partial \bar{w}}\frac{\partial \bar{w}}{\partial \tau}\right)}$$

Plugging in the wage effects (6) and (8), with wages and labour supply/demand dropping out (homogeneous regions), then gives us

$$MCPF_C=-\frac{\frac{\varepsilon}{(\eta-\varepsilon)}-\frac{\eta}{(1-\tau)(\eta-\varepsilon)}}{\left(1+\tau\frac{\eta}{(1-\tau)(\eta-\varepsilon)}+\frac{\tau\eta}{(1-\tau)}\frac{\varepsilon}{(\eta-\varepsilon)}\right)} \tag{52}$$

$$MCPF_u = -\frac{\frac{(1-\tau)\varepsilon-\eta}{(1-\tau)(\eta-\varepsilon)}}{\left(1+\tau\frac{\eta}{(1-\tau)(\eta-\varepsilon)}+\frac{\tau\eta}{(1-\tau)}\frac{\varepsilon}{(\eta-\varepsilon)}\right)}$$

$$MCPF_u = \frac{1}{\left(1+\frac{\tau\eta}{(1-\tau)\varepsilon-\eta}-\tau\frac{\eta}{(1-\tau)\varepsilon-\eta}-\frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon-\eta}\right)}$$

$$MCPF_u = \frac{1}{\left(1-\frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon-\eta}\right)}$$

So that:

$$\frac{\frac{\partial V}{\partial G^F}}{\lambda} = \frac{\frac{\partial V}{\partial G}}{\lambda} = \frac{1}{\left(1-\frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon-\eta}\right)} \tag{53}$$

Which gives us expression (13).

Appendix C. Fully decentralised case

C1. Proof of Lemma 1

The decentralised optimization problem is expressed by the Lagrangian:

$$\mathcal{L} = V_i(\bar{w}_i, \pi_i, G_i, G_i^F) - \gamma\left\{\left(G_i + G_i^F\right) - \tau L_i w\right\}$$

Giving us the following FOCs:

$$\frac{\partial \mathcal{L}}{\partial \tau_i} = \frac{\partial V_i}{\bar{w}_i}\frac{\partial \bar{w}_i}{\partial \tau_i} + \frac{\partial V_i}{\pi_i}\frac{\partial \pi_i}{\partial \tau_i} + \mu\left(L_{S_i} w + \tau_i L_{S_i}\frac{\partial w}{\partial \tau_i} + \tau_i w\frac{\partial L_{S_i}}{\partial \bar{w}_i}\frac{\partial \bar{w}_i}{\partial \tau_i}\right) = 0 \tag{54}$$

$$\frac{\partial \mathcal{L}}{\partial G_i} = \frac{\partial V_i}{\partial G_i} - \gamma = 0$$

$$\frac{\partial \mathcal{L}}{\partial G_i^F} = \frac{\partial V_i}{\partial G_i^F} - \gamma = 0$$

Therefore:

$$\frac{\partial V_i}{\partial G_i^F} = \frac{\partial V_i}{\partial G_i} = -\frac{\frac{\partial V_i}{\bar{w}_i}\frac{\partial \bar{w}_i}{\partial \tau_i} + \frac{\partial V_i}{\pi_i}\frac{\partial \pi_i}{\partial \tau_i}}{\left(L_{S_i} w + \tau_i L_{S_i}\frac{\partial w}{\partial \tau_i} + \tau_i w\frac{\partial L_{S_i}}{\partial \bar{w}_i}\frac{\partial \bar{w}_i}{\partial \tau_i}\right)}$$

Using Roy's identity, with λ_i marginal utility of income, and since $\frac{\partial V_i}{\pi_i} = \frac{\partial V_i}{\partial c_i}\frac{\partial c_i}{\pi_i} = \lambda_i \times 1$, we get:

$$MRS_i = \frac{\frac{\partial V_i}{\partial G_i}}{\lambda_i} = -\frac{L_{S_i}\frac{\partial \bar{w}_i}{\partial \tau_i} + \frac{\partial \pi_i}{\partial \tau_i}}{\left(L_{S_i} w + \tau_i L_{S_i}\frac{\partial w}{\partial \tau_i} + \tau_i w\frac{\partial L_{S_i}}{\partial \bar{w}_i}\frac{\partial \bar{w}_i}{\partial \tau_i}\right)}$$

Now, since in symmetric Nash equilibrium all identical states will set an identical tax rate τ, and after plugging in the profit effect (10), we obtain

$$MCPF_i = -\frac{\left(L_S\frac{\partial \bar{w}}{\partial \tau_i} - L_D\frac{\partial w}{\partial \tau_i}\right)}{\left(L_S w + \tau L_S\frac{\partial w}{\partial \tau_i} + \tau w\frac{\partial L_S}{\partial \bar{w}}\frac{\partial \bar{w}}{\partial \tau_i}\right)} \tag{55}$$

Rewriting as before, and plugging in the wage effects (5) and (9), with wages and labour supply/demand dropping out:

$$MCPF_i = -\frac{\left(\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)} - \frac{\eta}{n(1-\tau)(\eta-\varepsilon)}\right)}{\left(1+\tau\frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)}\right)}$$

Rewriting yields:

$$MCPF_i = -\frac{\frac{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}{n(1-\tau)(\eta-\varepsilon)}}{\left(1+\tau\frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)}\right)}$$

$$MCPF_i = \frac{1}{\left(\frac{n(1-\tau)(\varepsilon-\eta)}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta} - \frac{\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta} + \frac{\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta} - \tau\frac{\eta}{((1-\tau)(n\varepsilon-(n-1)\eta)-\eta)} - \frac{\tau\eta(n\varepsilon-(n-1)\eta)}{((1-\tau)(n\varepsilon-(n-1)\eta)-\eta)}\right)}$$

$$MCPF_i = \frac{1}{\left(1+\frac{-(1-\tau)\eta+\eta-\tau\eta-\tau\eta(n\varepsilon-(n-1)\eta)}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)}$$

$$MCPF_i = \frac{1}{\left(1-\frac{\tau\eta(n\varepsilon-(n-1)\eta)}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)}$$

So that:

$$\frac{\frac{\partial V}{\partial G^F}}{\lambda} = \frac{\frac{\partial V}{\partial G}}{\lambda} = \frac{1}{\left(1-\frac{\tau\eta(n\varepsilon-(n-1)\eta)}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)} \tag{56}$$

Which, together with (15) characterises the symmetric Nash equilibrium. ■

C2. Proof of proposition 1

To verify the sign of this bias, we compare the unitary $MCPF_u$ derived in section 3 which was unaffected by externalities, to the state $MCPF_i$ captured by lemma 1:

$$MCPF_i = \frac{1}{\left(1-\frac{\tau^*\eta(n\varepsilon-(n-1)\eta)}{(1-\tau^*)(n\varepsilon-(n-1)\eta)-\eta}\right)} \lessgtr \frac{1}{\left(1-\frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon-\eta}\right)} = MCPF_u \tag{57}$$

Evaluating at the Nash-equilibrium ($\tau^* = \tau$), and looking at the denominators in (57), the state $MCPF_i$ will be perceived as larger than the unitary $MCPF_u$ if

$$\frac{\tau\eta\left(n\varepsilon-(n-1)\eta\right)}{(1-\tau)\left(n\varepsilon-(n-1)\eta\right)-\eta} > \frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon-\eta} \tag{58}$$

From (58), this is easily seen to be the case in our simplified setting here. Keeping in mind that $\varepsilon < 0$ and $\eta > 0$, the term $(n\varepsilon-(n-1)\eta)$ on the LHS of expression (58) will be larger in absolute value compared to its counterpart in the RHS which is simply ε. It then suffices to take the derivative of the RHS with respect to $|\varepsilon|$, which is positive and equal to $\frac{\eta}{((1-\tau)\varepsilon-\eta)^2}$, to prove that the inequality always holds. ■

C3. Internalising the horizontal externalities

Expression (59) gives us the MCPF of region i when all externalities are internalised. The denominator thus has the effect of a tax increase on other regions' tax revenues incorporated, whilst the numerator integrates the effect on the welfare of non-residents:

$$MCPF_{I_i} = -\left(\frac{L_{S_i}\frac{\partial \bar{w}_i}{\partial \tau_i} + \frac{\partial \pi_i}{\partial \tau_i} + \sum_{j\neq i}^{n}\left(L_{S_j}\frac{\partial \bar{w}_j}{\partial \tau_i} + \frac{\partial \pi_i}{\partial \tau_i}\right)}{\frac{\partial R_i}{\partial \tau_i} + \sum_{j\neq i}^{n}\frac{\partial R_j}{\partial \tau_i}}\right) \tag{59}$$

Or, rewriting:

$$MCPF_{I_i} = -\left(\frac{L_{S_i}\frac{\partial \bar{w}_i}{\partial \tau_i} + \frac{\partial \pi_i}{\partial \tau_i} + \sum_{j\neq i}^{n}\left(L_{S_j}\frac{\partial \bar{w}_j}{\partial \tau_i} + \frac{\partial \pi_i}{\partial \tau_i}\right)}{L_{S_i}w + \tau_i L_{S_i}\frac{\partial w}{\partial \tau_i} + \tau_i w\frac{\partial L_{S_i}}{\partial \bar{w}_i}\frac{\partial \bar{w}_i}{\partial \tau_i} + \sum_{j\neq i}^{n}\left(\tau_j L_{S_j}\frac{\partial w}{\partial \tau_i} + \tau_j w\frac{\partial L_{S_j}}{\partial \bar{w}_j}\frac{\partial \bar{w}_j}{\partial \tau_i}\right)}\right)$$

Plugging in the profit effect (10) gives us:

$$MCPF_{I_i} = -\left(\frac{L_{S_i}\frac{\partial \bar{w}_i}{\partial \tau_i} - L_{D_i}\frac{\partial w}{\partial \tau_i}\eta + \sum_{j\neq i}^{n}\left(L_{S_j}\frac{\partial \bar{w}_j}{\partial \tau_i} - L_{D_j}\frac{\partial w}{\partial \tau_i}\right)}{L_{S_i}w + \tau_i L_{S_i}\frac{\partial w}{\partial \tau_i} + \tau_i w\frac{\partial L_{S_i}}{\partial \bar{w}_i}\frac{\partial \bar{w}_i}{\partial \tau_i} + \sum_{j\neq i}^{n}\left(\tau_j L_{S_j}\frac{\partial w}{\partial \tau_i} + \tau_j w\frac{\partial L_{S_j}}{\partial \bar{w}_j}\frac{\partial \bar{w}_j}{\partial \tau_i}\right)}\right)$$

Since $\frac{\partial \bar{w}_j}{\partial \tau_i} = \frac{\partial((1-\tau_j)w)}{\partial \tau_i} = (1-\tau_j)\frac{\partial w}{\partial \tau_i}$, and regions are homogeneous, this becomes:

$$MCPF_{I_i} = -\left(\frac{L\frac{\partial \bar{w}}{\partial \tau_i} - L\frac{\partial w}{\partial \tau_i} + (n-1)L\left((1-\tau)\frac{\partial w}{\partial \tau_i} - \frac{\partial w}{\partial \tau_i}\right)}{Lw + \tau L\frac{\partial w}{\partial \tau_i} + \tau w\frac{\partial L}{\partial \bar{w}}\frac{\partial \bar{w}_i}{\partial \tau} + (n-1)\left(\tau L\frac{\partial w}{\partial \tau_i} + \tau w\frac{\partial L}{\partial \bar{w}}(1-\tau)\frac{\partial w}{\partial \tau}\right)}\right) \tag{60}$$

Or, since $\tau w\frac{\partial L}{\partial \bar{w}} = \tau w\frac{\partial L}{\partial \bar{w}}\frac{\bar{w}}{L}\frac{L}{\bar{w}} = \tau w\eta\frac{L}{w(1-\tau)} = \frac{\tau L\eta}{(1-\tau)}$:

$$MCPF_{I_i} = -\left(\frac{L\frac{\partial \bar{w}}{\partial \tau_i} - L\frac{\partial w}{\partial \tau_i} + (n-1)L\left((1-\tau)\frac{\partial w}{\partial \tau_i} - \frac{\partial w}{\partial \tau_i}\right)}{Lw + \tau L\frac{\partial w}{\partial \tau_i} + \frac{\tau L\eta}{(1-\tau)}\frac{\partial \bar{w}_i}{\partial \tau} + (n-1)\left(\tau L\frac{\partial w}{\partial \tau_i} + \frac{\tau L\eta}{(1-\tau)}(1-\tau)\frac{\partial w}{\partial \tau}\right)}\right)$$

Plugging in the regional wage effects (5) and (9), and with labour supply and demand canceling out:

$$MCPF_{I_i} = -\left(\frac{\frac{w(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)} - \frac{w\eta}{n(1-\tau)(\eta-\varepsilon)} + (n-1)\left((1-\tau)\frac{w\eta}{n(1-\tau)(\eta-\varepsilon)} - \frac{w\eta}{n(1-\tau)(\eta-\varepsilon)}\right)}{w + \tau\frac{w\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{w(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)} + (n-1)\left(\tau\frac{w\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}(1-\tau)\frac{w\eta}{n(1-\tau)(\eta-\varepsilon)}\right)}\right)$$

Wages drop out as well:

$$MCPF_{I_i} = -\left(\frac{\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)} - \frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + (n-1)\left(\frac{\eta}{n(\eta-\varepsilon)} - \frac{\eta}{n(1-\tau)(\eta-\varepsilon)}\right)}{1 + \tau\frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)} + (n-1)\left(\tau\frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{\eta}{n(\eta-\varepsilon)}\right)}\right)$$

So that we get:

$$MCPF_{I_i} = -\left(\frac{n\frac{\varepsilon}{n(\eta-\varepsilon)} - n\frac{\eta}{n(1-\tau)(\eta-\varepsilon)}}{1 + n\tau\frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{n\varepsilon}{n(\eta-\varepsilon)}}\right)$$

Or:

$$MCPF_{I_i} = -\left(\frac{\frac{\varepsilon}{(\eta-\varepsilon)} - \frac{\eta}{(1-\tau)(\eta-\varepsilon)}}{1+\tau\frac{\eta}{(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{\varepsilon}{(\eta-\varepsilon)}}\right)$$

Which is exactly equal to the welfare cost (52) derived under the second-best unitary case:

$$MCPF_C = -\frac{\frac{\varepsilon}{(\eta-\varepsilon)} - \frac{\eta}{(1-\tau)(\eta-\varepsilon)}}{\left(1+\tau\frac{\eta}{(1-\tau)(\eta-\varepsilon)} + \frac{\tau\eta}{(1-\tau)}\frac{\varepsilon}{(\eta-\varepsilon)}\right)} \tag{61}$$

Appendix D. Shared tax base

D1. Proof of lemma 2

The optimisation problem yields the same marginal rate of substitution as before in (55). Since in symmetric Nash equilibrium all identical states will set an identical tax rate t, and after plugging in the profit effect (10), we thus obtain

$$MCPF_i = -\frac{\left(L_S\frac{\partial\bar{w}}{\partial\tau_i} - L_D\frac{\partial w}{\partial\tau_i}\right)}{\left(L_S w + tL_S\frac{\partial w}{\partial\tau_i} + tw\frac{\partial L_S}{\partial\bar{w}}\frac{\partial\bar{w}}{\partial\tau_i}\right)}$$

Plugging in the wage effects (5), (9) and profit effect (10), with wages and labour supply and demand dropping out as before:

$$MCPF_i = -\frac{\left(\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)} - \frac{\eta}{n(1-\tau)(\eta-\varepsilon)}\right)}{\left(1+t\frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{t\eta}{(1-\tau)}\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)}\right)}$$

Keeping in mind that $\tau = t + T$. Rewriting yields:

$$MCPF_i = -\frac{\frac{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}{n(1-\tau)(\eta-\varepsilon)}}{\left(1+t\frac{\eta}{n(1-\tau)(\eta-\varepsilon)} + \frac{t\eta}{(1-\tau)}\frac{(n\varepsilon-(n-1)\eta)}{n(\eta-\varepsilon)}\right)}$$

$$MCPF_i = \frac{1}{\left(\frac{n(1-\tau)(\varepsilon-\eta)}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta} - \frac{\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta} + \frac{\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta} - t\frac{\eta}{((1-\tau)(n\varepsilon-(n-1)\eta)-\eta)} - \frac{t\eta(n\varepsilon-(n-1)\eta)}{((1-\tau)(n\varepsilon-(n-1)\eta)-\eta)}\right)}$$

Hence,

$$MCPF_i = \frac{1}{\left(1+\frac{-(1-t-T)\eta+\eta-t\eta-t\eta(n\varepsilon-(n-1)\eta)}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)}$$

Or,

$$MCPF_i = \frac{1}{\left(1-\frac{t\eta(n\varepsilon-(n-1)\eta)-T\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)}$$

So that:

$$\frac{\frac{\partial V}{\partial G}}{\lambda} = \frac{1}{\left(1-\frac{t\eta(n\varepsilon-(n-1)\eta)-T\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)} = MCPF_i \tag{62}$$

Which, together with (21), then defines the symmetric Nash equilibrium. ■

D2. Proof of proposition 2

We compare the unitary $MCPF_u$ derived in section 3 which was unaffected by externalities, to the state $MCPF_i$ captured by lemma 2:

$$MCPF_i = \frac{1}{\left(1 - \frac{t^*\eta(n\varepsilon-(n-1)\eta)-T^*\eta}{(1-\tau)(n\varepsilon-(n-1)\eta)-\eta}\right)} \lessgtr \frac{1}{\left(1 - \frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon-\eta}\right)} = MCPF_u \tag{63}$$

Once more, as soon as $MCPF_i$ outweighs the unitary outcome $MCPF_u$, positive externalities will result in undertaxation. Looking at the denominators in (63), and evaluating in symmetric Nash equilibrium so that $t^* + T^* = \tau^* = \tau$, this happens when:

$$\frac{t^*\eta\left(n\varepsilon-(n-1)\eta\right)-T^*\eta}{(1-\tau)\left(n\varepsilon-(n-1)\eta\right)-\eta} > \frac{\tau\eta\varepsilon}{(1-\tau)\varepsilon-\eta}$$

Rewriting yields:

$$\frac{\tau\left(n\varepsilon-(n-1)\eta\right)-T^*\left(n\varepsilon-(n-1)\eta\right)-T^*}{(1-\tau)\left(n\varepsilon-(n-1)\eta\right)-\eta} > \frac{\tau\varepsilon}{(1-\tau)\varepsilon-\eta}$$

Or:

$$\frac{n\tau\varepsilon-(n-1)T^*\eta-(n-1)t^*\eta-T^*\left(n\varepsilon-(n-1)\eta\right)-T^*}{(1-\tau)\left(n\varepsilon-(n-1)\eta\right)-\eta} > \frac{n\tau\varepsilon}{n\left((1-\tau)\varepsilon-\eta\right)}$$

Which, after some more manipulation gives us

$$\frac{n\tau\varepsilon-t^*(n-1)\eta-T^*(1+n\varepsilon)}{n(1-\tau)\left(\varepsilon-\eta\right)+(1-\tau)\eta-\eta} > \frac{n\tau\varepsilon}{n\left((1-\tau)\varepsilon-\eta\right)}$$

Or,

$$\frac{n\tau\varepsilon-t^*(n-1)\eta-T^*(1+n\varepsilon)}{n\left((1-\tau)\varepsilon-\eta\right))+(n-1)\tau\eta} > \frac{n\tau\varepsilon}{n\left((1-\tau)\varepsilon-\eta\right)} \tag{64}$$

Now, Keeping in mind that $\varepsilon < 0$ and $\eta > 0$, the denominators and the numerators on both sides of (64) will always be negative. A higher labour supply elasticity η then brings the denominator on the LHS down in absolute value, and this through $(n-1)\tau\eta$, thus pushing up the bias. Turning to the numerators, we see that η brings about the same effect, but via $t^*(n-1)\eta$. The labour demand elasticity ε on the other hand, pulls the perceived MCPF down through $T^*(1+n\varepsilon)$. Lastly, as the state tax rate t^* accounts for a smaller share of the total tax rate τ^*, this latter effect comes out reinforced. Indeed, in this case we move away further from the fully decentralised case expressed by (56), where the bias was proven strictly positive. All four points combined gives us proposition 2. ■

D3. Proof of Corollary 1

It is straightforward to show that (27) reduces to the following expression under unit taxation, evaluated at the same Nash equilibrium (t^*, T^*):

$$W_t = t^*\Gamma_G\left[L_S^{'}\right]\left(1-\frac{1}{n}\right)w' + T^*\Gamma_{G^F}\left[L_D^{'}\right]w' \tag{65}$$

Now, given $(1-\tau^*)w \geq 1$ and $\lambda < (\Gamma_G + \Gamma_{G^F})$ (27), which are non-restrictive assumptions, the positive effects of ad valorem taxation marked out in (27) -and relative to (65)- come into play. ■

Appendix E. Appendix E: Heterogeneous regions

E1. Optimisation problem second-best optimum

The second best optimization problem is expressed by the Lagrangian:

$$\mathcal{L} = S - \gamma(\sum_{i}^{n} G_i - \tau \sum_{i}^{n} L_i w - \sum_{i}^{n} \pi_i) \tag{66}$$

With $\sum_i^n \{V_i(\bar{w}, G_i\} = S$. Leading to the following first order conditions:

$$\frac{\partial \mathcal{L}}{\partial \tau} = \frac{\partial S}{\partial \tau} + \gamma \left(\frac{\partial(\tau \sum_i^n L_i w + \sum_i^n \pi_i)}{\partial \tau} \right) = 0 \tag{67}$$

$$\frac{\partial \mathcal{L}}{\partial G_i} = \frac{\partial S}{\partial G_i} - \gamma = 0 \quad for\, i = 1, .., n \tag{68}$$

Solving both conditions yields:

$$\sum_{i}^{n} \frac{\partial V_i}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \tau} + \gamma \left(\sum_{i}^{n} L_i w + \tau \sum_{i}^{n} L_i \frac{\partial w}{\partial \tau} + \tau \sum_{i}^{n} w \frac{\partial L_i}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \tau} + \sum \frac{\partial \pi_i}{\partial \tau} \right) = 0 \ (\tau) \tag{69}$$

$$\frac{\partial V_i}{\partial G_i} - \gamma = 0 \quad for\, i = 1, .., n \quad (G_i) \tag{70}$$

With $\sum_i^n L_i w + \tau \sum_i^n L_i \frac{\partial w}{\partial \tau} + \tau \sum_i^n w \frac{\partial L_i}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \tau} + \sum_i^n \frac{\partial \pi_i}{\partial \tau}$ equal to $\frac{\partial R}{\partial \tau}$, being the marginal effect on federal revenue by raising the labour tax. Substituting for γ then yields:

$$\frac{\partial V_i}{\partial G_i} = -\frac{\sum_i^n \frac{\partial V_i}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \tau}}{\frac{\partial R}{\partial \tau}} \quad for\, i = 1, .., n \tag{71}$$

Public provision in each region will again continue until its marginal benefits equal its marginal cost, and this in terms of welfare cost as well as actual provision cost (see also Dahlby, 2008). This expression can be reformulated to arrive at the conventional MCPF expression using Roy's identity:

$$\frac{\partial V_i}{\partial G_i} = -\frac{\sum_i^n \left(L_i \lambda_i \frac{\partial \bar{w}}{\partial \tau} \right)}{\frac{\partial R}{\partial \tau}} \quad for\, i = 1, .., n \tag{72}$$

Which we can rewrite as:

$$\frac{\frac{\partial V_i}{\partial G_i}}{\sum_i^n (s_i \lambda_i)} = -\frac{L \frac{\partial \bar{w}}{\partial \tau}}{\frac{\partial R}{\partial \tau}} \quad for\, i = 1, .., n \tag{73}$$

With $L = \sum_i^n L_i$, λ_i the marginal utility of income of a household residing in region i, and s_i this household's share $(\frac{L_i}{L})$ in total labour supplied in the federation as a whole. The conventional MCPF formula is thus once again extracted:

$$MCPF_C = -\frac{L \frac{\partial \bar{w}}{\partial \tau}}{\sum_i^n L_i w + \tau \sum_i^n L_i \frac{\partial w}{\partial \tau} + \tau \sum_i^n w \frac{\partial L_i}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \tau} + \sum_i^n \frac{\partial \pi_i}{\partial \tau}} \tag{74}$$

With the third term of the denominator to be rewritten as:

$$\tau \sum_i^n w \frac{\partial L_i}{\partial \bar{w}} \frac{\partial \bar{w}}{\partial \tau} = \tau \sum_i^n w \frac{\partial L_i}{\partial \bar{w}} \frac{\bar{w}}{L_i} \frac{L_i}{\bar{w}} \frac{\partial \bar{w}}{\partial \tau} = \frac{w\tau}{\bar{w}} \frac{\partial \bar{w}}{\partial \tau} \sum_i^n \eta_i L_i = \frac{\tau}{(1-\tau)} \frac{\partial \bar{w}}{\partial \tau} \sum_i^n \eta_i L_i \quad (75)$$

Plugging in (44), (46) and (10), we get:
$MCPF_C =$

$$-\left(\frac{\frac{Lw \sum_i^n \varepsilon_i \frac{L_{D_i}}{w}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)}}{\sum_i^n L_i w + \frac{\tau \sum_i^n L_i w \sum_i^n \eta_i \frac{L_{s_i}}{\bar{w}_i}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)} + \frac{\frac{\tau}{(1-\tau)} Lw \sum_i^n \varepsilon_i \frac{L_{D_i}}{w} \sum_i^n \eta_i L_i}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)} - \frac{nL_{D_i} w \sum_i^n \eta_i \frac{L_{s_i}}{\bar{w}_i}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)}} \right)$$

Factoring out Lw and solving further gives us:

$$MCPF_C = -\left(\frac{\frac{\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)}}{1 + \frac{\tau \sum_i^n \eta_i \frac{L_{s_i}}{\bar{w}_i}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)} + \frac{\frac{\tau}{(1-\tau)} \sum_i^n \varepsilon_i \frac{L_{D_i}}{w} \sum_i^n \eta_i L_i}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)} - \frac{\sum_i^n \eta_i \frac{L_{s_i}}{\bar{w}_i}}{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)}} \right) \quad (76)$$

$$MCPF_C = -\left(\frac{1}{\frac{\left((1-\tau)\sum_i^n \left(\eta_i \frac{L_{s_i}}{\bar{w}_i}\right) - \sum_i^n \left(\varepsilon_i \frac{L_{D_i}}{w}\right)\right)}{\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}} + \tau \frac{\sum_i^n \eta_i \frac{L_{s_i}}{\bar{w}}}{\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}} + \frac{\tau}{L(1-\tau)} \sum_i^n \eta_i L_i - \frac{\sum_i^n \eta_i \frac{L_{s_i}}{\bar{w}}}{\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}}} \right) \quad (77)$$

$$MCPF_C = \left(\frac{1}{1 - \frac{\tau}{L(1-\tau)} \sum_i^n \eta_i L_i} \right) \quad (78)$$

So that:

$$\frac{\frac{\partial V_i}{\partial G_i}}{\sum_i^n (s_i \lambda_i)} = \left(\frac{1}{1 - \frac{\tau}{L(1-\tau)} \sum_i^n \eta_i L_i} \right) \quad for\, i = 1, .., n \quad (79)$$

Optimisation problem fully decentralised case

It is straightforward to derive the following condition describing the optimum of the problem:

$$\frac{\frac{\partial g_i}{\partial G_i}}{\lambda_i} = -\frac{L_i \frac{\partial \bar{w}_i}{\partial \tau_i}}{L_i w + \tau_i L_i \frac{\partial w}{\partial \tau_i} + \tau_i w \frac{\partial L_i}{\partial \bar{w}_i} \frac{\partial \bar{w}_i}{\partial \tau_i} + \frac{\partial \pi_i}{\partial \tau_i}} \quad (80)$$

Plugging in (36), (38) and rewriting as before, we get:

$$MCPF_i = -\frac{\left(L_i \frac{w\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}\right)}{L_i w_i + \tau_i L_i \left(\frac{w\eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}\right) + \frac{\tau_i L_i \eta_i}{(1-\tau_i)} \frac{w\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}} \tag{81}$$

$$-L_{D_i}\left(\frac{w\eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}\right)$$

Factoring out $L_i w$ and solving further gives us:

$$MCPF_i = -\frac{\left(\frac{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}\right)}{1 + \tau_i \left(\frac{\eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}\right) + \frac{\tau_i \eta_i}{(1-\tau_i)} \frac{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}} \tag{82}$$

$$-\frac{L_{D_i}}{L_{s_i}}\left(\frac{\eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}\right)$$

$$MCPF_i = -\frac{1}{\frac{\left(\sum_i^n \left((1-\tau_i)\eta_i \frac{L_{s_i}}{\bar{w_i}}\right) - \left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right)\right)}{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)} + \frac{\tau_i \eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)} + \frac{\tau_i \eta_i}{(1-\tau_i)} - \frac{\frac{L_D}{L_{s_i}} \eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)}} \tag{83}$$

$$MCPF_i = -\frac{1}{\frac{\left(\sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right) + \eta_i \frac{L_{s_i}}{\bar{w_i}} - \left(\sum_i^n \varepsilon \frac{L_D}{w}\right)\right)}{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)} + \frac{\tau_i \eta_i}{(1-\tau_i)} - \frac{\frac{L_D}{L_{s_i}} \eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)}} \tag{84}$$

$$MCPF_i = \frac{1}{1 - \frac{\tau_i \eta_i}{(1-\tau_i)} - \left(1 - \frac{L_{D_i}}{L_{s_i}}\right)\left(\frac{\eta_i \frac{L_{s_i}}{\bar{w_i}}}{\left(\left(\sum_i^n \varepsilon_i \frac{L_{D_i}}{w}\right) - \sum_{j\neq i}^n \left((1-\tau_j)\eta_j \frac{L_{s_j}}{\bar{w_j}}\right)\right)}\right)} \tag{85}$$

Acknowledgements

We would like to thank André Decoster, Stef Proost, Amedeo Piolatto, David Bartolini, and Laura Birg for their helfpul comments and suggestions, as well as participants at the IIPF annual conference (Dublin), ERSA annual conferences (Lisbon, Palermo) and LAGV conference (Aix-Marseille). We thankfully recognise the financial support of BELSPO (Contract TA/00/39), and the 'Steunpunt beleidsrelevant Onderzoek Fiscaliteit en Begroting' (2011-2015).

References

Albouy, D. (2012). Evaluating the efficiency and equity of federal fiscal equalization. *Journal of Public Economics*, 96(9):824–839.

Andersson, L., Aronsson, T., and Wikstrom, M. (2004). Testing for vertical fiscal externalities. *International Tax and Public Finance*, 11:243–263.

Aronsson, T. and Blomquist, S. (2008). Redistribution and provision of public goods in an economic federation. *Journal of Public Economic Theory*, 10(1):125–143.

Atkinson, A. B. and Stern, N. H. (1974). Pigou, taxation and public goods. *The Review of economic studies*, pages 119–128.

Boadway, R. (1982). On the method of taxation and the provision of local public goods: Comment. *The American Economic Review*, pages 846–851.

Boadway, R. and Flatters, F. (1982). Efficiency and equalization payments in a federal system of government: A synthesis and extension of recent results. *Canadian Journal of Economics*, pages 613–633.

Boadway, R. and Keen, M. (1996). Efficiency and the optimal direction of federal-state transfers. *International Tax and Public Finance*, 3(2):137–155.

Boadway, R., Marchand, M., and Vigneault, M. (1998). The consequences of overlapping tax bases for redistribution and public spending in a federation. *Journal of public Economics*, 68(3):453–478.

Boadway, R. and Shah, A. (2009). *Fiscal federalism: principles and practice of multiorder governance*. Cambridge University Press.

Brueckner, J. K. (2003). Strategic interaction among governments: An overview of empirical studies. *International Regional Science Review*, 26(2):175–188.

Brülhart, M. and Jametti, M. (2006). Vertical versus horizontal tax externalities: an empirical test. *Journal of Public Economics*, 90(10):2027–2062.

Buchanan, J. M. and Goetz, C. J. (1972). Efficiency limits of fiscal mobility: An assessment of the tiebout model. *Journal of Public Economics*, 1(1):25–43.

Dahlby, B. (1996). Fiscal externalities and the design of intergovernmental grants. *International tax and public finance*, 3(3):397–412.

Dahlby, B. (2008). The marginal cost of public funds: Theory and applications. *MIT Press Books*, 1.

Dahlby, B. and Wilson, L. S. (2003). Vertical fiscal externalities in a federation. *Journal of Public Economics*, 87(5-6):917–930.

De Borger, B., Dunkerley, F., and Proost, S. (2007). Strategic investment and pricing decisions in a congested transport corridor. *Journal of Urban Economics*, 62(2):294–316.

Devereux, M. P., Lockwood, B., and Redoano, M. (2007). Horizontal and vertical indirect tax competition: Theory and some evidence from the usa. *Journal of Public Economics*, 91(3):451–479.

Esteller-Moré, Á. and Solé-Ollé, A. (2001). Vertical income tax externalities and fiscal interdependence: evidence from the us. *Regional Science and Urban Economics*, 31(2):247–272.

Flatters, F., Henderson, V., and Mieszkowski, P. (1974). Public goods, efficiency, and regional fiscal equalization. *Journal of Public Economics*, 3(2):99–112.

Hoyt, W. H. (1991). Property taxation, nash equilibrium, and market power. *Journal of Urban Economics*, 30(1):123–131.

Keen, M. and Marchand, M. (1997). Fiscal competition and the pattern of public spending. *Journal of Public Economics*, 66(1):33–53.

Keen, M. J. and Kotsogiannis, C. (2002). Does federalism lead to excessively high taxes? *The American Economic Review*, 92(1):363–370.

Keen, M. J. and Kotsogiannis, C. (2004). Tax competition in federations and the welfare consequences of decentralization. *Journal of Urban Economics*, 56(3):397–407.

Kotsogiannis, C. (2010). Federal tax competition and the efficiency consequences for local taxation of revenue equalization. *International Tax and Public Finance*, 17(1):1–14.

Kotsogiannis, C. and Martínez, D. (2008). Ad valorem taxes and the fiscal gap in federations. *Economics Letters*, 99(3):431–434.

Krelove, R. (1992). Efficient tax exporting. *Canadian Journal of Economics*, pages 145–155.

Lockwood, B. (2001). Tax competition and tax co-ordination under destination and origin principles: a synthesis. *Journal of Public Economics*, 81(2):279–319.

Lucas, V. (2004). Cross-border shopping in a federal economy. *Regional Science and Urban Economics*, 34(4):365–385.

Mansoorian, A. and Myers, G. M. (1993). Attachment to home and efficient purchases of population in a fiscal externality economy. *Journal of public Economics*, 52(1):117–132.

McLure, C. E. (1967). The interstate exporting of state and local taxes: estimates for 1962. *National Tax Journal*, pages 49–77.

Richter, W. F. and Wellisch, D. (1996). The provision of local public goods and factors in the presence of firm and household mobility. *Journal of Public Economics*, 60(1):73–93.

Sato, M. (2000). Fiscal externalities and efficient transfers in a federation. *International Tax and Public Finance*, 7(2):119–139.

Wellisch, D. (2000). *Theory of public finance in a federal state*. Cambridge University Press.

Wellisch, D. and Wildasin, D. E. (1996). Decentralized income redistribution and immigration. *European Economic Review*, 40(1):187–217.

Wildasin, D. E. (1991). Income redistribution in a common labor market. *The American Economic Review*, pages 757–774.

Wildasin, D. E. (2000). Labor-market integration, investment in risky human capital, and fiscal competition. *American Economic Review*, pages 73–95.

Wilson, J. D. (1986). A theory of interregional tax competition. *Journal of Urban Economics*, 19(3):296–315.

Wilson, J. D. and Janeba, E. (2005). Decentralization and international tax competition. *Journal of Public Economics*, 89(7):1211–1229.

Zodrow, G. R. and Mieszkowski, P. (1986). Pigou, tiebout, property taxation, and the underprovision of local public goods. *Journal of Urban Economics*, 19(3):356–370.