

Blomquist, Sören; Micheletto, Luca

**Working Paper**

## Redistribution, In-Kind Transfers and Matching Grants when the Federal Government Lacks Information on Local Costs

Working Paper, No. 2004:15

**Provided in Cooperation with:**

Department of Economics, Uppsala University

*Suggested Citation:* Blomquist, Sören; Micheletto, Luca (2004) : Redistribution, In-Kind Transfers and Matching Grants when the Federal Government Lacks Information on Local Costs, Working Paper, No. 2004:15, Uppsala University, Department of Economics, Uppsala

This Version is available at:

<http://hdl.handle.net/10419/82761>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



UPPSALA  
UNIVERSITET

Working Paper 2004:15

*Department of Economics*

Redistribution, In-Kind Transfers and  
Matching Grants when the Federal  
Government Lacks Information on  
Local Costs

Sören Blomquist and Luca Micheletto

Department of Economics  
Uppsala University  
P.O. Box 513  
SE-751 20 Uppsala  
Sweden  
Fax: +46 18 471 14 78

Working paper 2004:15  
October 2004  
ISSN 0284-2904

REDISTRIBUTION, IN-KIND TRANSFERS AND MATCHING GRANTS WHEN THE FEDERAL  
GOVERNMENT LACKS INFORMATION ON LOCAL COSTS

SÖREN BLOMQUIST AND LUCA MICHELETTO

Papers in the Working Paper Series are published  
on internet in PDF formats.  
Download from <http://www.nek.uu.se>  
or from S-WoPEC <http://swopec.hhs.se/uunewp/>

# Redistribution, in-kind transfers and matching grants when the federal government lacks information on local costs\*

Sören Blomquist<sup>†</sup> and Luca Micheletto<sup>‡</sup>

October 18, 2004

## Abstract

Recent literature on optimal nonlinear taxation has shown that in models with a single level of government public provision of private goods can help redistribution by mitigating self-selection constraints. The aim of the present paper is to extend the analysis to a fiscal federalism setting with two levels of government. To accomplish this goal we start by explicitly modelling the informational asymmetry that in our framework motivates decentralization in the first place: the lower level is able to observe the local costs of production, which vary across localities, but the central level does not. Then, using an extended version of the optimal income taxation model with two ability types, we show that even though it is the lower level that is responsible for the public provision of private goods, in-kind transfers can still help the federal level to redistribute between high- and low skilled individuals. Finally, we characterize the optimal marginal tax rates, which take a different form from that of a unitary model, and the optimal matching grants. The latter, in particular, have a very different structure than the one derived in previous fiscal federalism studies.

We also find that it is vital to model informational asymmetries between the central and the local level explicitly. Models where the informational asymmetry is not explicit might have very little to say about decentralization in economies where the local level has an informational advantage over the central level.

Keywords: Optimal taxation; in-kind transfers; public provision of day care; fiscal federalism.

JEL Classification: H21; H42; H77.

---

\*We would like to thank Thomas Aronsson, seminar participants at University of Uppsala, University of Brescia, PET 2004 in Beijing and IIPF 2004 in Milan for their comments. Usual disclaimer applies.

<sup>†</sup>Uppsala University, Department of Economics, Box 513 SE-751 20 Uppsala, Sweden.

<sup>‡</sup>Uppsala University, Department of Economics, Box 513 SE-751 20 Uppsala, Sweden.

# 1 Introduction

Public expenditures are large in all developed countries. In many European countries they are close to 50% of GNP. These expenditures occur both at the federal and the local level. As a rough characterization one can say that federal level expenses are in the form of cash transfers whereas local level expenditures take the form of locally publicly provided goods and services. In Sweden, as an example, in year 2001 local level expenditures, almost exclusively in the form of publicly provided goods and services, amounted to 28% of GNP. Given the large quantitative magnitude of both federal and local level expenditures it is of great importance to understand how the federal and local levels interact, the principles that should govern the responsibility for different types of expenditures and the structure of intergovernmental grants. Since long there exists a large literature studying these issues.

In order to understand the interplay between the federal and lower levels we need a stylized model. However, it is essential that the model corresponds to the real world. In our view many previous models use a misleading stylization of the economy. Hence, we will begin this paper by briefly describing what we believe are the essential facts about the actual fiscal federalism structure.

1. There is income heterogeneity among individuals. The high taxes and large public expenditures that we observe in all developed countries can only be understood if we believe redistribution between heterogeneous individuals is an important goal.

2. There is heterogeneity among local communities in at least three important respects: costs of production and preferences vary across communities; the income distribution varies across communities; the demographic composition varies across communities.

3. To achieve redistribution the federal level uses nonlinear income taxation, not a proportional tax.

4. Local communities have an informational advantage in terms of local preferences and production conditions.

5. Local communities provide mainly *private* goods, not public goods.

Previous studies have not taken these stylized facts into account in a satisfactory way. The generic fiscal federalism model is a model with a representative agent in each local community; both the federal level and the local level use a proportional income tax and the local level is responsible for provision of a local public good. Asymmetric information is rarely explicitly modelled. In our view, this type of model does not capture the important aspects of real economies.

If we want to study whether a certain good should be provided by the local or the federal level and how a grant system should be designed we believe it is essential to understand, in the first place, why the good is publicly provided. What function does public provision serve? In almost all previous

theoretical fiscal federalism studies one has assumed that the local communities provide a local public good. The reason for public provision is then taken for granted and not explicitly modelled. Hundred years ago it was maybe the case that public goods constituted the major part of local public expenditures. Nowadays, however, local public expenditures are only to a very small part of the public goods type. The lions part are private goods with rivalness in consumption, like education, health care, care of elderly and day care. There exist several explanations for why public provision of certain types of private goods can be beneficial. All these explanations are provided in models with just one level in the public sector. In this paper we focus on the reason pointed out in papers by among others Boadway and Marchand (1995), Blomquist and Christiansen (1995) and Cremer and Gahvari (1997). The motivation for public provision in those papers is that public provision of a certain type of private goods helps to mitigate the self-selection constraint in a model where redistribution is an important goal. Day care is probably the good in question that best fits the explanation provided in those papers. Hence we will in this paper use a model where day care is publicly provided, the reason being that this helps redistribution. Since we believe redistribution, not public good provision, is the most important reason for the type of tax/transfer systems we observe in the developed countries, we assume the federal level wants to achieve income redistribution via a nonlinear income tax.

The decision making in the local communities is of crucial importance for the workings of the model, the form of vertical fiscal externality and the design of the grant system. The stylization used in many previous models with a representative agent in each community and each jurisdiction deciding on how much of a local public good to provide gives a simple structure, but is not a valid stylization of the local decision making. It is a fact that in each community there is heterogeneity in the population, which makes it plausible to consider a more complex decision making where also the local authority is interested in redistribution.

It is widely recognized that an informational advantage is one important reason for decentralizing some decisions to the local level. However, this informational advantage is rarely modelled explicitly. In this paper we explicitly model the informational advantage that might motivate decentralization. We assume that the costs of production vary across communities. Local politicians know the local production function, but the federal level does not. The production costs for, say, day care might vary across localities because of the presence of radon in the ground in some places, presence of air pollution requiring extra cleaning costs, etc.<sup>1</sup>

---

<sup>1</sup>We assume that the cost conditions in all other respects are equal across communities. It is, of course, true that the cost of living can vary between regions and that this might have implications for how taxes and grants should be designed. Such differences in living costs would be observable and would give rise to other issues than those that we study

The instrument set available at the local level is also vital. As will be shown, letting the local level have access to a linear tax can seriously hamper the federal level in pursuing its goals. We will therefore assume the local level can only use a proportional tax. The local level will also be responsible for the provision of day care and a local public good. Since the local level can not make cash transfers, the local policy maker has to use provision of day care and the local public good to affect the utility distribution in the community.

The paper contributes to two strands of literature. First, it contributes to the fiscal federalism literature in the sense that we investigate if standard results “survive” when we consider a setting that corresponds better to reality, namely that there is provision of a private good that can help mitigate a self-selection constraint. Second, it contributes to the literature on publicly provided goods. Earlier analyses of how public provision of private goods can help to soften the self-selection constraints in an optimal taxation problem have all been performed in unitary models. Here we investigate how results are affected in a setting where the incentive-compatibility constraints thwarts the redistributive attempts of the central level but the provision of the good that weakens these constraints is done at the local level.

The fiscal federalism literature is vast and contains sub-literatures focusing on different aspects of how the federal and local levels can be interconnected. For example, risk sharing and horizontal competition due to mobility of some economic factor/agent constitute two important topics.<sup>2</sup> To simplify, we do not deal at all with these issues in this paper. The literature more related to the present study is concerned with vertical interconnections arising from some commonality of tax base between different level of government.<sup>3</sup> In Boadway and Keen (1996) the analysis is concerned solely with efficiency issues (they abstract from differences in fiscal capacity across states and also assume all individuals to be identical) and the emphasis is on the optimal fiscal gap (the mismatch between expenditure and revenue-raising responsibilities) in a federation. They find that the optimal federal government policy involves a negative tax rate and that this can imply in reasonable circumstances a negative fiscal gap. In Dahlby (1996) the focus is on state commodity taxes and tax bases that are interrelated across states. In each state the decisions are made by a representative agent. The federal level can use an income tax and a set of grants. Dahlby characterizes the set of optimal grants and shows that the unitary solution can be implemented by a suitable set of grants. Dahlby and Wilson (2003) examine vertical fiscal externalities in a model where state governments provide a productivity

---

in this paper, namely that the local level has better information on local production conditions than the one the federal level has access to.

<sup>2</sup>See e.g. Lockwood (1999), Cornes and Silva (2000), Aronsson and Wikström (2003), and Wellish (2000).

<sup>3</sup>See Keen (1998) for a general discussion of the topic.

enhancing public input and the states and the central government share the tax base for both a labor income tax and a profits tax. All states are identical and local public policies are in each state made by a representative agent. Dahlby and Wilson (2003) find that matching grants are needed to correct for the vertical externality.<sup>4</sup> The matching grant should reflect the marginal tax revenue that accrues to the federal level when the public expenditure increases by one unit. The unitary nation solution can be implemented by a suitable set of grants.

Boadway et al. (1998) consider a federation where both the federal and local levels have access to a linear income tax and the local levels provide a local public good. All states are alike. An interesting feature is that in each state there is a population with heterogeneous skills and that both the state and the federal policy makers have a redistributive motive; they both pursue policies in order to maximize utility for the worst off individual (the Rawlsian criterion). The federal level can also use lump-sum transfers/taxes to the local level. As in most earlier papers there are no informational asymmetries. They find that the unitary nation solution can be decentralized. In this decentralized solution the local level pursues all redistributive taxation and transfers income to the federal level.

Our model has strong similarities with Boadway et al. (1998), Dahlby (1996) and Dahlby and Wilson (2003), but there are also important differences. In our model the local level provides day care. In a model with a federal proportional income tax this would have generated a tax base effect, as in Blomquist and Bergstrom (1996). This externality could be corrected by a matching grant of the form described in Dahlby and Wilson (2003). However, since we assume that the federal level uses an optimal nonlinear income tax, the important effect from the day care provision is no longer the tax base effect but the effect on the self-selection constraints. We believe our formulation of the federal level tax system is more realistic and also in line with how optimal income tax systems are modelled in unitary economies. Another important difference is that we explicitly model the asymmetry of information between the federal and local level. In our model the costs to produce day care differ across communities. If the federal level can observe these costs the unitary nation solution can also be implemented as a decentralized solution. However, if the federal level can not observe the day care costs, but local communities can, then there exists no unitary solution (involving public provision) and the decentralized optimal solution differs from the solution obtained under the assumption that the federal level has full information.

The paper is organized as follows. Section 2 gives a brief outline of our model. In section 3 we analyze the unitary solution with observability of

---

<sup>4</sup>A very interesting contribution of Dahlby and Wilson (2003) is that they show that the fiscal vertical externality can be negative.



costs. Section 4 describes the decentralized solution for the cases with and without observability at the federal level of local costs. For the latter case we characterize the federal income tax and the optimal grant system for supporting the local communities. Section 5 concludes.

## 2 The model

The model builds on the two-types version of the optimal income taxation model developed by Stern (1982) and Stiglitz (1982). There is a federal level of government and two local communities. The two local communities have an identical number of citizens. To simplify, mobility of households is ruled out.<sup>5</sup> In each local community the population is made up by low skilled and high skilled agents. Both local communities have the same proportion of high- and low skilled individuals.<sup>6</sup> Moreover, we assume for simplicity that in each local community the number of low skilled workers is equal to the number of high skilled workers and we normalize it to one.

Agents of type 1 are assumed to be the low skilled ones: they earn a unitary wage  $w^1$ . Agents of type 2 are assumed to be the high skilled ones: they earn a unitary wage  $w^2 > w^1$ . The two local communities differ with respect to the producer price of providing day care. Local community  $h$  is characterized by a high cost parameter  $p_h$ , whereas local community  $l$  is characterized by a low cost parameter  $p_l < p_h$ .

Households supply labor in the market and gross income is taxed non-linearly by the federal authority. The disposable income net of expenditures on purchases of day care services is used to finance consumption of the single consumption good produced in this economy. This consumption good is also used as numeraire. Following Blomquist and Christiansen (2003) the demand for day care services is modelled in a very simple way: day care services demanded are always equal to labor supplied.<sup>7</sup> Day care can be publicly provided and households can top up public provision with private purchases. Households have (identical) preferences represented by the concave utility function  $U = u(c, L) + v(G)$ , where  $c$  is the consumption good,  $L$  is labor supply and  $G$  a local public good. The federal level maximizes a social welfare function given by a weighted sum of the utilities achieved by

---

<sup>5</sup>As long as mobility is not perfect the qualitative conclusions of our analysis would not change if some degree of costly mobility were introduced.

<sup>6</sup>This assumption does not affect the qualitative results of the model. See Aronsson and Blomquist (2004) for a model of optimal redistributive taxation in a fiscal federalism setting where the distribution of ability types in the population differs across local communities.

<sup>7</sup>Even if some parents maybe would leave the children in day care when they go to the dentist or go shopping we believe that parents enjoy being with their kids and that it is a reasonable stylization to let the demand for day care hours be equal to the hours of work. To let the demand be a positive monotone function of hours of work would not change the content of the analysis but create a slightly more complex notation.

low- and high skilled agents in the two local communities. As regards the local policy makers, we assume that each one of them chooses the value of its policy instruments so as to maximize a weighted sum of the utilities of both low- and high skill living in its own jurisdiction, using the same weights as the federal level.<sup>8</sup>

We start by considering the centralized solution where the federal level does not observe ability types but knows in which local community the cost parameter takes the value  $p_l$  and in which it takes the value  $p_h$ . We then consider two decentralized solutions where the local communities can impose a proportional tax on their residents and are responsible for the public provision of day care and local public goods. We first analyze the case where the federal level knows “which local community is which” and look at how the federal level can use a matching grant policy to implement the unitary solution previously derived. Then, we analyze the same problem for the case where the federal level cannot observe which is the low cost and which is the high cost community.

### 3 Unitary solution with observability of costs

We model the local public good as additively separable in the utility function. This will imply that in the unitary solution with observability of costs, the first order conditions with respect to the other policy instruments will not be affected by the level of public good provision. In this section we therefore do not include the public good in the problem formulation. It will play a role in the decentralized solution.

Letting  $T(Y)$  denote the federal nonlinear income tax and  $\bar{x}$  the maximum amount of day care that is publicly provided free of charge to anyone who demands it, the maximization problem for an individual will take the form:  $Max\ u(c, Y/w) \quad s.t. \quad c = Y - T(Y) - p \max(0, Y/w - \bar{x})$ . Substituting the budget constraint into the utility function we obtain:

$$Max\ u(Y - T(Y) - p \max\{0, Y/w - \bar{x}\}, Y/w)$$

Suppressing the dependence on  $w$  in the second argument and denoting by  $MRS_{cY}$  the individual’s marginal rate of substitution between gross income and consumption ( $MRS_{cY} = -\frac{\partial u}{\partial Y} / \frac{\partial u}{\partial c}$ ), we obtain from the first order condition of the individual’s maximization problem that, for an agent that is not supplementing the publicly provided day care ration, the marginal income tax rate  $T'(Y)$  is implicitly given by:

$$T'(Y) = 1 - MRS_{cY}, \tag{1}$$

---

<sup>8</sup>This does not affect much the final results but allows using the envelope theorem to a greater extent.

whereas for an agent that is supplementing it is:

$$T'(Y) = 1 - MRS_{cY} - p/w. \quad (2)$$

Since we assume no mobility the federal level can, if costs are observable, impose community specific taxes. We let  $Y_i^k$  and  $B_i^k$ ,  $k = 1, 2$ ;  $i = l, h$ , denote the before tax income and after tax incomes for type  $k$  in community  $i$ . The federal level government's problem is to choose type and community specific taxes and also the optimal provision level of day care in each community to maximize a weighted sum of agents' utilities subject to a budget constraint and a set of incentive-compatibility constraints. These are needed to make sure that agents truthfully reveal their type and do not mimic, namely do not choose a bundle that is intended by the federal government for another group of agents. Assuming redistribution goes from the high- to the low skilled group, the federal government has only to take into account that high skilled agents do not mimic. As regards the day care, it is easy to show that the optimal level of the public provision is not uniquely determined and that in particular all values above the one demanded by the low skilled agents are equivalent (see Blomquist and Christiansen, 2003).<sup>9</sup> To obtain a unique determination we impose the assumption that the ration is set at the lowest of all equivalent amounts. The public provision of day care,  $\bar{x}_i$ , in community  $i$ , will then be equal to  $Y_i^1/w^1$ . We will also make the simplifying assumption that at the optimum in each local community the high skilled agents supply a larger amount of labor than the low skilled ones. The federal government's problem takes the following form:

$$\begin{aligned} & \max_{Y_l^1, Y_l^2, B_l^1, B_l^2, Y_h^1, Y_h^2, B_h^1, B_h^2} \alpha^1 u \left( B_l^1, \frac{Y_l^1}{w^1} \right) + \alpha^1 u \left( B_h^1, \frac{Y_h^1}{w^1} \right) + \\ & + \alpha^2 u \left( B_l^2 - p_l \left\{ \frac{Y_l^2}{w^2} - \frac{Y_l^1}{w^1} \right\}, \frac{Y_l^2}{w^2} \right) + \alpha^2 u \left( B_h^2 - p_h \left\{ \frac{Y_h^2}{w^2} - \frac{Y_h^1}{w^1} \right\}, \frac{Y_h^2}{w^2} \right) \end{aligned}$$

subject to

$$u \left( B_l^2 - p_l \left( \frac{Y_l^2}{w^2} - \frac{Y_l^1}{w^1} \right), \frac{Y_l^2}{w^2} \right) \geq u \left( B_l^1, \frac{Y_l^1}{w^1} \right); \quad (\lambda_l)$$

$$u \left( B_h^2 - p_h \left( \frac{Y_h^2}{w^2} - \frac{Y_h^1}{w^1} \right), \frac{Y_h^2}{w^2} \right) \geq u \left( B_h^1, \frac{Y_h^1}{w^1} \right); \quad (\lambda_h)$$

---

<sup>9</sup>Since Blomquist and Christiansen (2003) show how public provision of day care can mitigate self-selection problems and yield a strict pareto improvement we do not show this here, but refer the interested reader to their work.

$$(Y_l^1 - B_l^1) + (Y_h^1 - B_h^1) + (Y_l^2 - B_l^2) + (Y_h^2 - B_h^2) \geq 2 \left( p_l \frac{Y_l^1}{w^1} + p_h \frac{Y_h^1}{w^1} \right), \quad (\gamma)$$

where Lagrange multipliers are within parentheses, the  $\lambda_i$ -constraints ( $i = l, h$ ) are the self-selection constraints and the  $\gamma$ -constraint represents the government budget constraint. Denoting by a “hat” a variable when referred to a mimicker and writing  $u_i^1$ ,  $u_i^2$  and  $\hat{u}_i$  as a short for respectively  $u \left( B_i^1, \frac{Y_i^1}{w^1} \right)$ ,  $u \left( B_i^2 - p_i \left\{ \frac{Y_i^2}{w^2} - \frac{Y_i^1}{w^1} \right\}, \frac{Y_i^2}{w^2} \right)$  and  $u \left( B_i^1, \frac{Y_i^1}{w^2} \right)$ , the first order conditions to this problem are:

$$B_i^1 : \quad \alpha^1 \frac{\partial u_i^1}{\partial c} - \lambda_i \frac{\partial \hat{u}_i}{\partial c} - \gamma = 0 \quad (3)$$

$$Y_i^1 : \quad \alpha^1 \frac{\partial u_i^1}{\partial Y_i^1} + (\alpha^2 + \lambda_i) \frac{\partial u_i^2}{\partial c} \frac{p_i}{w^1} - \lambda_i \frac{\partial \hat{u}_i}{\partial Y_i^1} + \gamma \left( 1 - 2 \frac{p_i}{w^1} \right) = 0 \quad (4)$$

$$B_i^2 : \quad (\alpha^2 + \lambda_i) \frac{\partial u_i^2}{\partial c} - \gamma = 0 \quad (5)$$

$$Y_i^2 : \quad (\alpha^2 + \lambda_i) \left[ \frac{\partial u_i^2}{\partial Y_i^2} - \frac{\partial u_i^2}{\partial c} \frac{p_i}{w^2} \right] + \gamma = 0. \quad (6)$$

For the high skill individuals we obtain from eqs. (5) and (6) that  $1 - MRS_{cY}^2 - p_i/w^2 = 0$ . By assumption the high skill individuals supplement. Hence, from eq. (2) we obtain that the high skill individuals in both localities face a zero marginal tax.

For a low skill individual, manipulating the first order conditions (3), (4) and (5), we obtain

$$1 - MRS_{cY}^1 = \lambda^* (MRS_{cY}^1 - \widehat{MRS}_{cY}) + p_i/w^1,$$

where  $\lambda^* = (\lambda/\gamma) (\partial \hat{u}_i / \partial c)$ .

By construction the low skill individuals do not supplement and the expression for their marginal tax is given by (1). Hence the marginal tax for a low skill individual will be

$$T'(Y_i^1) = \lambda_i^* (MRS_{cY}^1 - \widehat{MRS}_{cY}) + p_i/w^1. \quad (7)$$

If the public provision of day care is so welfare enhancing that the self-selection constraint is not binding at the solution to the federal government's problem then the low skilled agents will face marginal taxes equal to:  $T'(Y_i^1) = p_i/w^1$ . If the low skilled agents live in a local community where the self-selection constraint is still binding at the optimum with public provision of day care, then they will face a positive marginal tax rate equal to a self-selection term plus the term reflecting the real cost of providing day care per extra unit of income earned.

These results are similar to, but slightly different from, the results obtained in Blomquist and Christiansen (2003). The reason why the low skill persons face positive marginal tax rates is that they should face the real cost of day care. The reason why the high skill should not face this cost via the marginal tax is that in our model they top up and pay by themselves for the marginal hours of day care.

## 4 Decentralized solution

We now assume that the local communities are responsible for the provision of day care and local public goods. For this to make sense the local level must have a reason and incentives for providing day care. In our view local politicians, as those active at the federal level, have an ambition to redistribute. However, we also know that usually local level authorities are not allowed to perform income distribution by use of cash transfers. Our model provides a rationale for this. If local politicians were allowed to use cash transfers they would not provide day care. However, not being allowed to use cash transfers, they will use public provision of day care in order to achieve their redistributive goals.

As we already anticipated, we will assume that the local policy maker's objective is to maximize a weighted sum of the utilities for the high- and low skill groups and that the weights used are the same as those used by the policy maker at the federal level (see footnote 8).

Public provision of a local *public* good is assumed partly for realism and partly for analytical convenience. It allows us to build a model where in equilibrium there aren't any surpluses in the budgets of the local jurisdictions. However, in the following we will not focus on the issue of local supply of public goods when presenting the results.

### 4.1 Decentralization with observability of local costs

In a setting where local communities have the power to impose a (positive) proportional tax on the income of their residents, if the federal level can observe which local community has a low producer price of day care and which has instead a high producer price, there is a straightforward way to implement the unitary solution obtained in the previous Section. Remember

that the optimal level of the public provision was not uniquely determined and that all values above the one demanded by the low skilled agents were equivalent. This means that an equivalent optimal solution could be found with the ration set at the level of day care demanded by the agents with the largest labor supply (by assumption the high skilled individuals). Then, to implement the unitary solution in a decentralized setting, it is sufficient to tell the local communities that all their expenditures on public provision of day care will be refunded by the federal government. In a framework without local public goods provision, if the federal government chooses a matching grant at a 100% rate, the local communities are not constrained in their budget decisions. This however does not mean that the local public expenditure will be infinite, since, as we set up the model, the households do not demand day care in excess of their labor supply. Thus, each local community will provide agents with as much day care as they demand without imposing any additional tax on them; in particular in the  $i$ -th local community the level of public provision will be set at  $\max \{Y_i^1/w^1, Y_i^2/w^2\}$  ( $= Y_i^2/w^2$  by assumption). If the local community is also responsible for the provision of a local public good this must (partly) be financed by a proportional local income tax. However, the federal level can adapt its income tax schedule so that the net effect for the individuals of the federal and local level taxes are the same as in the unitary solution. The level of the public good provision could be controlled by community specific grants. These grants can be of the matching grant type, but lump-sum grants would also work. Hence, if costs are observable the unitary solution can be decentralized.

## 4.2 Decentralization without observability of local costs

The really interesting case is, of course, that where the information set of the federal level is inferior to that available at the local level. Hence in this subsection we consider the case where the federal level can observe neither the per unit cost of producing day care nor the total day care costs in a community. However, we assume that *the hours of day care publicly provided* in a community can be observed and that the grant system can use this information. Thus, the federal government observes the level of local public provision but it does not observe the per unit cost of day care. It only knows that half of the local communities are characterized by  $p_h$  and half are characterized by  $p_l$ . Now the federal government must take into account the possibility that mimicking occurs also between agents living in different communities since it cannot provide each local community with its own (community specific) nonlinear tax schedule. There is only one federal income tax schedule applicable to everyone. This schedule must be designed in such a way that people self-select properly.

The federal government chooses a nonlinear income tax schedule  $T(Y)$  and each local community chooses its proportional tax rate  $t_i$  together with

the level of the ration  $\bar{x}_i$  and the level of the local public good  $G_i$ . The cost of the local public goods is uniform across communities and normalized to one. This value is known to the federal government. There will also be a grant system from the federal to the local level. Below we will both study the case of a linear grant and that of a nonlinear grant.

#### 4.2.1 Individuals' behavior

An individual pays both federal and local income tax. Income left after the payment of the federal income tax is denoted by  $B$ :  $B \equiv Y - T(Y)$ . An agent of type  $k$  living in local community  $i$  chooses optimally labor supply  $L_i^k = Y_i^k/w^k$  by solving the following problem:

$$\max_Y u \left( Y(1 - t_i) - T(Y) - p_i \max \left\{ 0, \frac{Y}{w^k} - \bar{x}_i \right\}, \frac{Y}{w^k} \right) + v(G_i). \quad (8)$$

Defining the indicator function  $I_i^k$  as:

$$\begin{aligned} I_i^k &= 1 \text{ if } \min \left\{ \frac{Y_i^k}{w^k}, \bar{x}_i \right\} = \bar{x}_i \\ &= 0 \text{ otherwise,} \end{aligned}$$

from the first order condition of problem (8) we get:

$$\left( 1 - t_i - T' - I_i^k p_i/w^k \right) u'_c = -u'_L/w^k. \quad (9)$$

(9) allows us to implicitly define the federal marginal tax rate and the total marginal tax rate  $\theta = t_i + T'$  faced by the individual as:

$$-\frac{u'_L}{u'_c} \frac{1}{w^k} = 1 - t_i - T' - I_i^k p_i/w^k = 1 - \theta - I_i^k p_i/w^k \quad (10)$$

The optimal  $Y_i^k$ , solution to (8), gives the indirect utility

$$V_i^k \equiv V \left( Y_i^k, B_i^k, t_i, \bar{x}_i; p_i, w^k \right) + v(G_i) \quad (11)$$

of an agent of type  $k$  living in local community  $i$ . With respect to  $V_i^k$  the following holds:  $\partial V_i^k / \partial t_i = -Y_i^k (\partial V_i^k / \partial B_i^k)$  and  $\partial V_i^k / \partial \bar{x}_i = I_i^k p_i (\partial V_i^k / \partial B_i^k)$ .

### 4.2.2 Local governments' behavior

We will study two types of grant systems; one linear and one nonlinear. When studying the nonlinear grant system we will linearize the local governments budget constraints around their optimum points. Hence, here it is sufficient to describe local governments' behavior when a linear grant system is in place.

Under a linear grant system the total grants paid out to a local community will be:  $p^*D_i - R$  where  $p^*$  is a matching grant per unit of day care provided by the community and  $R$  is a lump sum tax (or transfer if it takes a negative value). The federal government is assumed to be a Stackelberg leader in the tax setting game. It knows that the local community  $i$  chooses  $t_i$ ,  $\bar{x}_i$  and  $G_i$  solving the following problem:

$$\begin{aligned} & \max_{t_i, \bar{x}_i, G_i} \sum_{k=1}^2 \alpha^k V_i^k + v(G_i) \\ \text{subject to: } & t_i \sum_{k=1}^2 Y_i^k \geq (p_i - p^*) D_i + R + G_i, \end{aligned}$$

where  $D_i \equiv \sum_{k=1}^2 \min \left\{ \frac{Y_i^k}{w^k}, \bar{x}_i \right\}$  is the total amount of units of day care publicly provided free of charge by the local community  $i$ .

The solution to this problem will give the local tax rate  $t_i$ , the provision  $D_i$  and the provision  $G_i$  as functions of the federal level instruments  $Y_i^k, B_i^k, p^*$  and  $R$ . These functions must satisfy the local government's budget constraint. Differentiating the budget constraint of local community  $i$ , we get:

$$\begin{aligned} \frac{\partial t_i}{\partial \bar{x}_i} &= \frac{\partial t_i}{\partial D_i} \frac{\partial D_i}{\partial \bar{x}_i} = \frac{(p_i - p^*)}{\sum_{k=1}^2 Y_i^k} \frac{\partial D_i}{\partial \bar{x}_i} = \frac{(p_i - p^*) \sum_{k=1}^2 I_i^k}{\sum_{k=1}^2 Y_i^k}, \\ \frac{\partial t_i}{\partial G_i} &= \frac{\partial t_i}{\partial R} = \left( \sum_{k=1}^2 Y_i^k \right)^{-1}, \quad \frac{\partial t_i}{\partial p^*} = - \frac{\sum_{k=1}^2 \min \left\{ \frac{Y_i^k}{w^k}, \bar{x}_i \right\}}{\sum_{k=1}^2 Y_i^k}. \end{aligned}$$

Moreover, we have also:



$$\begin{aligned}\frac{dt_i}{dp^*} &= \frac{\partial t_i}{\partial p^*} + \frac{\partial t_i}{\partial \bar{x}_i} \frac{\partial \bar{x}_i}{\partial p^*} + \frac{\partial t_i}{\partial G_i} \frac{\partial G_i}{\partial p^*}; \\ \frac{dt_i}{dR} &= \frac{\partial t_i}{\partial R} + \frac{\partial t_i}{\partial \bar{x}_i} \frac{\partial \bar{x}_i}{\partial R} + \frac{\partial t_i}{\partial G_i} \frac{\partial G_i}{\partial R}.\end{aligned}$$

### 4.2.3 The federal government's optimal policy

The federal government's optimal policy can be found as the solution to the problem of maximizing a weighted sum of utilities subject to self-selection constraints and a budget constraint.

If we think of the agents' indifference curves as defined by the indirect utility function (11), we could have two possible outcomes with respect to the ranking of the slopes of these curves. For sure, the low skilled living in community  $h$  will have the steepest indifference curves in this space and the high skilled living in community  $l$  the flattest ones. However, one cannot tell a priori how the slopes of the indifference curves of the low skilled living in community  $l$  and of the high skilled living in community  $h$  are relatively ranked. If the difference in local technologies for day care provision is sufficiently large relatively to difference in unitary wage rates for low- and high skilled workers, then we will have that the indifference curves for the high skilled living in community  $h$  are steeper than the indifference curves for low skilled living in community  $l$ . The opposite case will occur if the effect of the difference in marginal productivities for low- and high skilled workers dominates the effect of differences in local technologies for day care.<sup>10</sup>

From the ranking of the slopes of indifference curves it can be determined which self-selection constraints should be taken into account in the federal government's problem. The possible binding self-selection constraints are those running downwards and linking pair of adjacent types. Assuming that the pattern of relevant binding self-selection constraints is such that the high skilled in  $l$  must be induced not to mimic the high skilled in  $h$ , that the high skilled in  $h$  must be induced not to mimic the low skilled in  $l$  and finally that the low skilled in  $l$  must be induced not to mimic the low skilled in  $h$ , the mimickers' utilities are defined as:

$$\begin{aligned}\widehat{V}_l^2 &\equiv V(Y_h^2, B_h^2, t_l, \bar{x}_l; p_l, w^2) + v(G_l) \\ \widehat{V}_h^2 &\equiv V(Y_l^1, B_l^1, t_h, \bar{x}_h; p_h, w^2) + v(G_h) \\ \widehat{V}_l^1 &\equiv V(Y_h^1, B_h^1, t_l, \bar{x}_l; p_l, w^1),\end{aligned}$$

with properties

---

<sup>10</sup>However, it cannot in principle be ruled out the case where the indifference curves for these two groups of agents are equally sloped.

$$\begin{aligned}
\frac{\partial \widehat{V}_l^2}{\partial t_l} &= -Y_h^2 \frac{\partial \widehat{V}_l^2}{\partial B_h^2}, & \frac{\partial \widehat{V}_l^2}{\partial \bar{x}_l} &= \widehat{I}_l^2 p_l \frac{\partial \widehat{V}_l^2}{\partial B_h^2}, \\
\frac{\partial \widehat{V}_h^2}{\partial t_h} &= -Y_l^1 \frac{\partial \widehat{V}_h^2}{\partial B_l^1}, & \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} &= \widehat{I}_h^2 p_h \frac{\partial \widehat{V}_h^2}{\partial B_l^1}, \\
\frac{\partial \widehat{V}_l^1}{\partial t_l} &= -Y_h^1 \frac{\partial \widehat{V}_l^1}{\partial B_h^1}, & \frac{\partial \widehat{V}_l^1}{\partial \bar{x}_l} &= \widehat{I}_l^1 p_l \frac{\partial \widehat{V}_l^1}{\partial B_h^1},
\end{aligned}$$

where, using  $\widehat{Y}_i^k$  to denote the gross income earned by a mimicker of ability type  $k$  living in local community  $i$ , the indicator functions  $\widehat{I}_i^k$  have been defined as:

$$\begin{aligned}
\widehat{I}_i^k &= 1 \text{ if } \min \left\{ \frac{\widehat{Y}_i^k}{w^k}, \bar{x}_i \right\} = \bar{x}_i \\
&= 0 \text{ otherwise.}
\end{aligned}$$

The pattern of binding self-selection constraints assumed above reflects the case where the difference in local technologies for day care provision is sufficiently small relatively to the difference between high- and low skilled productivities in the market. It is interesting to note that in this case all the binding self-selection constraints involve agents living in different local communities. This feature contributes to make public provision of private goods especially effective, even more than in the standard models with just one level of government. The reason is that here it can serve for the federal government as a partial substitute for the impossibility to impose community-specific nonlinear income tax schedules. Whereas in the framework where it was originally considered public provision of private goods had no welfare effect as long as the level of the ration was inframarginal for all agents (including the mimickers) in the economy since all agents could have access to it, this is no longer true in a context where mimickers and mimicked live in different jurisdictions. In fact, in this latter case public provision of private goods turns out to be welfare-enhancing even when a marginal increase is evaluated starting from a zero initial level, the reason being that not all agents can take advantage from its provision.

**Linear grant system** With a linear grant system and the assumed pattern of binding self-selection constraints we can now formulate the federal government's problem as:

$$\max_{Y_l^2, B_l^2, Y_h^2, B_h^2, Y_l^1, B_l^1, Y_h^1, B_h^1, R, p^*} \sum_{k=1}^2 \sum_{i=l, h} \alpha^k V_i^k + \sum_{i=l, h} v(G_i)$$

subject to:

$$\begin{aligned} V_l^2 &\geq \widehat{V}_l^2; & (\lambda_l^2) \\ V_h^2 &\geq \widehat{V}_h^2; & (\lambda_h^2) \\ V_l^1 &\geq \widehat{V}_l^1; & (\lambda_l^1) \end{aligned}$$

$$\sum_{k=1}^2 \sum_{i=l,h} (Y_i^k - B_i^k) + 2R \geq p^* \sum_{i=l,h} D_i, \quad (\gamma)$$

where Lagrange multipliers are within parentheses, the  $\lambda_i^k$ -constraints represent the self-selection constraints and the  $\gamma$ -constraint is the federal government's budget constraint.

Forming the Lagrangian  $\Lambda$  of the federal government's problem, we have:

$$\begin{aligned} \Lambda = & \sum_{k=1}^2 \sum_{i=l,h} \alpha^k V_i^k + \sum_{i=l,h} v(G_i) + \sum_{k=1}^2 \lambda_l^k (V_l^k - \widehat{V}_l^k) + \lambda_h^2 (V_h^2 - \widehat{V}_h^2) + \\ & + \gamma \left[ \sum_{k=1}^2 \sum_{i=l,h} (Y_i^k - B_i^k) + 2R - p^* \sum_{i=l,h} D_i \right] \end{aligned}$$

The first order conditions of the federal government's problem are:

$$(\alpha^2 + \lambda_l^2) \frac{\partial V_l^2}{\partial Y_l^2} = -\gamma \left( 1 - p^* \frac{\partial D_l}{\partial Y_l^2} \right) - \delta_{Y_l^2} \quad (12)$$

$$(\alpha^2 + \lambda_l^2) \frac{\partial V_l^2}{\partial B_l^2} = \gamma \left( 1 + p^* \frac{\partial D_l}{\partial B_l^2} \right) - \delta_{B_l^2} \quad (13)$$

$$(\alpha^2 + \lambda_h^2) \frac{\partial V_h^2}{\partial Y_h^2} = \lambda_l^2 \frac{\partial \widehat{V}_l^2}{\partial Y_h^2} - \gamma \left( 1 - p^* \frac{\partial D_h}{\partial Y_h^2} \right) - \delta_{Y_h^2} \quad (14)$$

$$(\alpha^2 + \lambda_h^2) \frac{\partial V_h^2}{\partial B_h^2} = \lambda_l^2 \frac{\partial \widehat{V}_l^2}{\partial B_h^2} + \gamma \left( 1 + p^* \frac{\partial D_h}{\partial B_h^2} \right) - \delta_{B_h^2} \quad (15)$$

$$(\alpha^1 + \lambda_l^1) \frac{\partial V_l^1}{\partial Y_l^1} = \lambda_h^2 \frac{\partial \widehat{V}_h^2}{\partial Y_l^1} - \gamma \left( 1 - p^* \frac{\partial D_l}{\partial Y_l^1} \right) - \delta_{Y_l^1} \quad (16)$$

$$(\alpha^1 + \lambda_l^1) \frac{\partial V_l^1}{\partial B_l^1} = \lambda_h^2 \frac{\partial \widehat{V}_h^2}{\partial B_l^1} + \gamma \left( 1 + p^* \frac{\partial D_l}{\partial B_l^1} \right) - \delta_{B_l^1} \quad (17)$$

$$\alpha^1 \frac{\partial V_h^1}{\partial Y_h^1} = \lambda_l^1 \frac{\partial \widehat{V}_l^1}{\partial Y_h^1} - \gamma \left( 1 - p^* \frac{\partial D_h}{\partial Y_h^1} \right) - \delta_{Y_h^1} \quad (18)$$

$$\alpha^1 \frac{\partial V_h^1}{\partial B_h^1} = \lambda_l^1 \frac{\partial \widehat{V}_l^1}{\partial B_h^1} + \gamma \left( 1 + p^* \frac{\partial D_h}{\partial B_h^1} \right) - \delta_{B_h^1} \quad (19)$$

$$\delta_R = -2\gamma \quad (20)$$

$$\delta_{p^*} = \gamma \sum_{i=l,h} \sum_{k=1}^2 \min \left\{ \frac{Y_i^k}{w^k}, \bar{x}_i \right\} = \gamma \sum_{i=l,h} D_i, \quad (21)$$

where  $\delta_{Y_l^2}$ ,  $\delta_{B_l^2}$ ,  $\delta_{Y_h^2}$ ,  $\delta_{B_h^2}$ ,  $\delta_{Y_l^1}$ ,  $\delta_{B_l^1}$ ,  $\delta_{Y_h^1}$ ,  $\delta_{B_h^1}$ ,  $\delta_R$  and  $\delta_{p^*}$  represent the indirect effects of the federal government's policy variables via the adjustments in the values of the decision variables chosen by the local authorities. These terms are defined in the Appendix.

We are now ready to characterize the optimal distortions imposed at the margin on the labor/leisure choice of each type of individuals.

**Proposition 1** Denoting by  $MRS_i^k$  the marginal rate of substitution between earned income and consumption for an agent of type  $k$  living in local community  $i$  ( $MRS_i^k \equiv (u'_c I_i^k p_i - u'_L) / w^k u'_c = -\frac{\partial V_i^k}{\partial Y_i^k} / \frac{\partial V_i^k}{\partial B_i^k}$ ), the total marginal tax rates faced by different agents in the federation are:<sup>11</sup>

$$\theta_l^2 = p^* \underbrace{\left( \frac{\partial D_l}{\partial Y_l^2} + MRS_l^2 \frac{\partial D_l}{\partial B_l^2} \right)}_{\Psi'} - \underbrace{\left( \frac{\delta_{Y_l^2}}{\gamma} + MRS_l^2 \frac{\delta_{B_l^2}}{\gamma} \right)}_{\Psi''}, \quad (22)$$

$$\theta_h^2 = \frac{\lambda_l^2}{\gamma} \frac{\partial \widehat{V}_l^2}{\partial B_h^2} \left( \frac{\partial \widehat{V}_l^2}{\partial Y_h^2} - \frac{\partial V_h^2}{\partial Y_h^2} \right) + p^* \left( \frac{\partial D_h}{\partial Y_h^2} + MRS_h^2 \frac{\partial D_h}{\partial B_h^2} \right) - \left( \frac{\delta_{Y_h^2}}{\gamma} + MRS_h^2 \frac{\delta_{B_h^2}}{\gamma} \right), \quad (23)$$

$$\theta_l^1 = \frac{\lambda_h^2}{\gamma} \frac{\partial \widehat{V}_h^2}{\partial B_l^1} \left( \frac{\partial \widehat{V}_h^2}{\partial Y_l^1} - \frac{\partial V_l^1}{\partial Y_l^1} \right) + p^* \left( \frac{\partial D_l}{\partial Y_l^1} + MRS_l^1 \frac{\partial D_l}{\partial B_l^1} \right) - \left( \frac{\delta_{Y_l^1}}{\gamma} + MRS_l^1 \frac{\delta_{B_l^1}}{\gamma} \right), \quad (24)$$

$$\theta_h^1 = \frac{\lambda_l^1}{\gamma} \frac{\partial \widehat{V}_l^1}{\partial B_h^1} \left( \frac{\partial \widehat{V}_l^1}{\partial Y_h^1} - \frac{\partial V_h^1}{\partial Y_h^1} \right) + p^* \left( \frac{\partial D_h}{\partial Y_h^1} + MRS_h^1 \frac{\partial D_h}{\partial B_h^1} \right) - \left( \frac{\delta_{Y_h^1}}{\gamma} + MRS_h^1 \frac{\delta_{B_h^1}}{\gamma} \right). \quad (25)$$

<sup>11</sup>Notice that the definition of  $MRS$  that we are using here is slightly different from the one we have been using in Section 3.

**Proof.** See the Appendix. ■

In Section 3 we found that in the unitary solution with observability of local costs all the high skilled agents, both those living in local community  $l$  and those living in local community  $h$ , faced a zero marginal tax rate. On the other hand, if the central government cannot observe “which local community is which”, we would get that in the unitary solution only the high skilled agents living in local community  $l$  should be undistorted at the margin, the reason being that, when community specific income tax schedules cannot be used, the high skilled living in local community  $h$  should face a non-zero marginal tax rate in order to discourage mimicking from the high skilled living in local community  $l$ . According to (22) we get instead the result that in the decentralized solution (under the assumption that the central authority cannot tell “which local community is which”) also the (total) marginal tax rate faced by the high skilled agents living in local community  $l$  is in general different from zero. The term labelled  $\Psi'$  is a budget term that reflects the change in the value of the net transfer of the federal government to the local communities when the high skilled agents living in local community  $l$  are induced to (work more and) marginally increase their earned income. It is due to the change in the optimal level of public provision of day care chosen by the local community  $l$ . The term labelled  $\Psi''$  is instead a welfare term that reflects how the well-being of agents living in local community  $l$  is affected as a consequence of the changes in the optimal values of the variables chosen at the local level ( $t_l$ ,  $\bar{x}_l$  and  $G_l$ ) when the high skilled living there are induced to marginally increase their earned income. Reflecting the effects on the well-being of agents living in local community  $l$ , it also incorporates the effects on the binding self-selection constraints involving agents living in local community  $l$  as potential mimickers.

It is immediate to see how from (22) one can easily obtain the optimal marginal tax rate of the centralized solution, for in this case we would have that  $\Psi'' = 0$  and  $\Psi' = p_l \left( \frac{\partial D_l}{\partial Y_l^2} + 0 \right) = (1 - I_l^2) \frac{p_l}{w^2}$ , thus getting  $\theta_l^2 = T'(Y_l^2) = (1 - I_l^2) \frac{p_l}{w^2}$ .

As regards the high skilled living in local community  $h$ , the (total) marginal tax rate provided by (23) is made up of three terms. The first term is a fairly standard self-selection terms that relies on the difference in the slopes of the indifference curves of a mimicker and of the agent being mimicked at the bundle intended for the latter.<sup>12</sup> In this particular case, this term reflects the distortion that has to be imposed on the bundle of-

---

<sup>12</sup>Notice however that in this particular case the mimicker and the mimicked are agents with the same market ability (they both are high skilled). Nevertheless, since they live in different local communities, they are facing different local proportional tax rates and therefore, even if their labor supply is the same, the consumption level they have access to is different.

ferred to the high skilled agents living in local community  $h$  in order to deter mimicking from the high skilled agents living in local community  $l$ . As for the second and third term of (23), they are similar to the ones appearing in (22) and their interpretation proceeds along the same lines.

Finally, the interpretation of (24) and (25) follows the one provided for (23).

Before considering the optimal value of  $p^*$  it is necessary to define in this context the notion of “compensated” increase in the rate of subsidy. By this expression we mean a marginal increase in  $p^*$  accompanied by a reduction in the value of the lump-sum transfer (or increase in the value of the lump-sum tax)  $dR$  such that the value of the Lagrangian of the local policy makers remains unaffected. To denote the effects of such a change in the federal government’s transfer policy, a “tilde” will be used over the local governments’ choice variables. It will therefore be:

$$\begin{aligned}\frac{\partial \tilde{x}_i}{\partial p^*} &= \frac{\partial \bar{x}_i}{\partial p^*} + \sum_{k=1}^2 \min \left\{ \frac{Y_i^k}{w^k}, \bar{x}_i \right\} \frac{\partial \bar{x}_i}{\partial R} \\ &= \frac{\partial \bar{x}_i}{\partial p^*} + D_i \frac{\partial \bar{x}_i}{\partial R};\end{aligned}\tag{26}$$

$$\begin{aligned}\frac{\partial \tilde{G}_i}{\partial p^*} &= \frac{\partial G_i}{\partial p^*} + \sum_{k=1}^2 \min \left\{ \frac{Y_i^k}{w^k}, \bar{x}_i \right\} \frac{\partial G_i}{\partial R} \\ &= \frac{\partial G_i}{\partial p^*} + D_i \frac{\partial G_i}{\partial R}.\end{aligned}\tag{27}$$

We are now ready to characterize the efficient value of the matching rate  $p^*$ .

**Proposition 2** *Denoting by  $\Theta_i$  the federal government’s net marginal valuation of a transfer to the local community  $i$  ( $\Theta_i \equiv -\delta_{R_i}/\gamma$ ), the optimal  $p^*$  is given by:*

$$p^* = \underbrace{\frac{\text{cov}(D_i, \Theta_i)}{E\left(\frac{\partial D_i}{\partial \bar{x}_i} \frac{\partial \tilde{x}_i}{\partial p^*}\right)}}_{\Omega'} + \underbrace{\frac{\sum_{i=l,h} \left[ \frac{\partial \tilde{x}_i}{\partial p^*} \sum_{k=1}^2 \lambda_i^k \left( \frac{\partial V_i^k}{\partial \bar{x}_i} - \frac{\partial \widehat{V}_i^k}{\partial \bar{x}_i} \right) + \frac{d\tilde{t}_i}{dp^*} \sum_{k=1}^2 \lambda_i^k \left( \frac{\partial V_i^k}{\partial t_i} - \frac{\partial \widehat{V}_i^k}{\partial t_i} \right) \right]}{\gamma \sum_{i=l,h} \frac{\partial D_i}{\partial \bar{x}_i} \frac{\partial \tilde{x}_i}{\partial p^*}}}{\Omega''},\tag{28}$$

where  $E\left(\frac{\partial D_i}{\partial \bar{x}_i} \frac{\partial \tilde{x}_i}{\partial p^*}\right) \equiv \frac{1}{2} \sum_{i=l,h} \frac{\partial D_i}{\partial \bar{x}_i} \frac{\partial \tilde{x}_i}{\partial p^*}$ ,  $\lambda_h^1 = 0$  and  $\frac{d\tilde{t}_i}{dp^*} \equiv \frac{\partial t_i}{\partial \bar{x}_i} \frac{\partial \tilde{x}_i}{\partial p^*} + \frac{\partial t_i}{\partial G_i} \frac{\partial \tilde{G}_i}{\partial p^*}$ .

**Proof.** See the Appendix. ■

According to (28) the optimal specific subsidy chosen by the federal government depends on two terms. The term labelled  $\Omega'$  highlights an equity-efficiency trade-off similar in structure to the one that characterizes the formulas for the optimal marginal tax rate obtained in the literature on optimal linear income taxation (see e.g. Atkinson and Stiglitz, 1980, pp. 407-408). The equity element is represented by the numerator and the efficiency element by the denominator. The latter is a measure of the distortion induced by the subsidy and it is definitely positive. The former, which determines the sign of  $\Omega'$ , reflects the way in which the social marginal utility of a transfer to a local community varies with the total amount of unities of day care publicly provided by that local community. Intuitively, if such a covariance is negative (positive), then the value of the optimal  $p^*$  would be pushed down (up) for equity reasons since it would mostly benefit the less (more) deserving local community, namely the one with the smaller (greater) value of  $\Theta$ .

The (numerator of) term labelled  $\Omega''$  reflects instead how the binding self-selection constraints of the federal government are affected by a compensated marginal increase in the value of  $p^*$ . Such a change would indirectly affect the binding self-selection constraints altering the attractiveness of the mimicking option for the potential mimickers through adjustments in both the local proportional tax rates  $t_i$  and the level of publicly provided day care  $\bar{x}_i$ .<sup>13</sup> Since the denominator of  $\Omega''$  is positive,  $\Omega''$  has the same sign as its numerator. A positive (negative) numerator would mean that a marginal compensated increase in the subsidy  $p^*$  has an overall beneficial (detrimental) effect in terms of alleviating the self-selection problems faced by the federal government in pursuing its redistributive goals. For this reason, it would push upwards (downwards) the optimal value of the subsidy  $p^*$ .

**Nonlinear grant system** Now we analyze the case when the federal government uses a nonlinear transfer scheme in its relations with the local communities. Through a nonlinear transfer scheme the federal government has a better control over the local level of public provision of day care. However, it has also to take into account an additional self-selection constraint, the one requiring that the local community with the lower cost of day care provision does not have incentive to mimic the high cost local community level. The local community  $l$  might in fact be tempted to provide the same total amount of day care that is provided by the local community  $h$  in order to receive the same transfer as the one intended only for the local community  $h$ .

---

<sup>13</sup>Notice however that the tightness of the binding self-selection constraints is not directly affected through variations in the level of local public good provision since it has been assumed separability between the local public good and other goods in individuals' preferences.

Our approach will be to linearize for each local community the transfer schedule around the equilibrium point.  $p_i^*$  and  $R_i$  ( $i = l, h$ ) will denote the virtual transfer parameters for the local community  $i$ . The federal government's problem can be written as:

$$\max_{Y_l^2, B_l^2, Y_h^2, B_h^2, Y_l^1, B_l^1, Y_h^1, B_h^1, p_l^*, p_h^*, R_l, R_h} \sum_{k=1}^2 \sum_{i=l, h} \alpha^k V_i^k + \sum_{i=l, h} v(G_i)$$

subject to:

$$\begin{aligned} V_l^2 &\geq \widehat{V}_l^2; & (\lambda_l^2) \\ V_h^2 &\geq \widehat{V}_h^2; & (\lambda_h^2) \\ V_l^1 &\geq \widehat{V}_l^1; & (\lambda_l^1) \end{aligned}$$

$$\begin{aligned} &\sum_{k=1}^2 \alpha^k V_l^k + v(G_l) \\ \geq &\alpha^1 V(Y_l^1, B_l^1, \widehat{t}_l, \widehat{x}_l; p_l, w^1) + \alpha^2 V(Y_l^2, B_l^2, \widehat{t}_l, \widehat{x}_l; p_l, w^2) + v(\widehat{G}_l) \quad (\lambda) \end{aligned}$$

$$\sum_{k=1}^2 \sum_{i=l, h} (Y_i^k - B_i^k) + \sum_{i=l, h} R_i \geq \sum_{i=l, h} p_i^* D_i. \quad (\gamma)$$

As compared to the previous problem there is now an additional self-selection constraint (the  $\lambda$ -constraint) reflecting that the low cost community should not mimic the high cost community.

The optimal (total) marginal tax rates faced by different agents are still given by (22), (23), (24) and (25). The only difference is that now  $p_l^*$  replaces  $p^*$  in (22) and (24), and that  $p_h^*$  replaces  $p^*$  in (23) and (25).

With respect to the optimal marginal matching rates faced by different local communities, Proposition 3 provides the main result for the low cost community.

**Proposition 3** *Defining  $\frac{d\widehat{t}_l}{dp_l^*} \equiv \frac{\partial t_l}{\partial \bar{x}_l} \frac{\partial \widehat{x}_l}{\partial p_l^*} + \frac{\partial t_l}{\partial G_l} \frac{\partial \widehat{G}_l}{\partial p_l^*}$ , the optimal marginal matching rate faced by the local community  $l$  is given by:*

$$p_l^* = \frac{\sum_{k=1}^2 \lambda_l^k \left[ \left( \frac{\partial V_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} \right) \frac{\partial \widehat{x}_l}{\partial p_l^*} + \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{d\widehat{t}_l}{dp_l^*} \right]}{\gamma \frac{\partial D_l}{\partial \bar{x}_l} \frac{\partial \widehat{x}_l}{\partial p_l^*}}. \quad (29)$$



**Proof.** See the Appendix. ■

According to (29) the only reason to distort the marginal cost of day care provision faced by the local community  $l$  is related to the effects that the induced changes in the local community's choice of  $\bar{x}_l$  and  $t_l$  exert on the (federal government's) binding self-selection constraints involving agents living in local community  $l$  as potential mimickers. Multiplying both sides of (29) by  $\frac{\partial D_l}{\partial \bar{x}_l} \frac{\partial \tilde{x}_l}{\partial p_l^*}$  one would obtain:

$$p_l^* \frac{\partial D_l}{\partial \bar{x}_l} \frac{\partial \tilde{x}_l}{\partial p_l^*} = \sum_{k=1}^2 \lambda_l^k \left[ \left( \frac{\partial V_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} \right) \frac{\partial \tilde{x}_l}{\partial p_l^*} + \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{d\tilde{t}_l}{dp_l^*} \right],$$

an equation that can be interpreted in terms of equality at the margin between costs and gains for the federal government from a compensated increase in  $p_l^*$ . In fact, the left hand side provides an evaluation of the dead-weight loss from departing the actual cost of marginal day care provision perceived by the local community  $l$  from the real resource cost  $p_l$ . The right hand side measures instead the net marginal benefit (due to the effect on the binding self-selection constraints  $\lambda_l^1$  and  $\lambda_l^2$ ) of the induced change in the values of the policy variables locally chosen.

Finally, Proposition 4 states the main result for the high cost local community.

**Proposition 4** *Defining  $\frac{d\tilde{t}_h}{dp_h^*} \equiv \frac{\partial t_h}{\partial \bar{x}_h} \frac{\partial \tilde{x}_h}{\partial p_h^*} + \frac{\partial t_h}{\partial G_h} \frac{\partial \widehat{G}_h}{\partial p_h^*}$ , the optimal marginal matching rate faced by the local community  $h$  is given by:*

$$\begin{aligned} p_h^* = & -\frac{\lambda}{\gamma \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \tilde{x}_h}{\partial p_h^*}} \sum_{k=1}^2 \alpha^k \left[ \frac{\partial V^k}{\partial \hat{t}_l} \frac{d\hat{t}_l}{dp_h^*} + D_h \frac{\partial V^k}{\partial \hat{t}_l} \frac{d\hat{t}_l}{dR_h} + \frac{\partial V^k}{\partial \hat{x}_l} \frac{d\hat{x}_l}{dp_h^*} + D_h \frac{\partial V^k}{\partial \hat{x}_l} \frac{d\hat{x}_l}{dR_h} \right] + \\ & -\frac{\lambda}{\gamma \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \tilde{x}_h}{\partial p_h^*}} \left( \frac{d\widehat{G}_l}{dp_h^*} + D_h \frac{d\widehat{G}_l}{dR_h} \right) v' + \\ & + \frac{\lambda_h^2 \left[ \left( \frac{\partial V_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} \right) \frac{\partial \tilde{x}_h}{\partial p_h^*} + \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{d\tilde{t}_h}{dp_h^*} \right]}{\gamma \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \tilde{x}_h}{\partial p_h^*}}. \end{aligned} \quad (30)$$

**Proof.** See the Appendix. ■

According to (30) the optimal marginal matching rate faced by the local community  $h$  is made up of three terms, all related to self-selection problems. As compared to the formula providing the optimal marginal matching rate faced by the local community  $l$  (eq. (29)), the main difference is represented by the first two lines of (30). The terms appearing there reflect how a compensated marginal increase in  $p_h^*$  affects the binding self-selection constraint

requiring that local communities truthfully reveal their type. They are intended to specifically characterize the distortion that has to be imposed at the margin in order to discourage the local community  $l$  from mimicking the local community  $h$ . Since the local policy makers choose among alternatives comparing the welfare of their residents, the mimicking option is made less attractive when a reform to the system of transfers to local communities induces changes in the policy variables chosen by the mimicking local community (in our case  $\widehat{dt}_l$ ,  $\widehat{d\bar{x}}_l$ ,  $\widehat{d\hat{G}}_l$ ) that are welfare-worsening from the local community viewpoint.<sup>14,15</sup>

The interpretation of the last line of (30) is instead similar to the one provided for (29). The only difference is that in this case it looks at how the (federal government's) binding self-selection constraint involving agents living in local community  $h$  as potential mimickers is affected by the changes in  $\bar{x}_h$  and  $t_h$  induced by a compensated marginal increase in  $p_h^*$ .

Similarly to what we did for eq. (29), also for (30) if we multiply both sides by  $\frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \bar{x}_h}{\partial p_h^*}$  we can rewrite the formula for the optimal marginal matching rate in terms of equality at the margin between marginal costs (left hand side) and gains (right hand side) for the federal government from a compensated increase in  $p_h^*$ :

$$\begin{aligned}
p_h^* \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \bar{x}_h}{\partial p_h^*} &= -\frac{\lambda}{\gamma} \sum_{k=1}^2 \alpha^k \left[ \frac{\partial V^k}{\partial \widehat{t}_l} \frac{d\widehat{t}_l}{dp_h^*} + D_h \frac{\partial V^k}{\partial \widehat{t}_l} \frac{d\widehat{t}_l}{dR_h} + \frac{\partial V^k}{\partial \widehat{x}_l} \frac{d\widehat{x}_l}{dp_h^*} + D_h \frac{\partial V^k}{\partial \widehat{x}_l} \frac{d\widehat{x}_l}{dR_h} \right] + \\
&\quad -\frac{\lambda}{\gamma} \left( \frac{d\widehat{G}_l}{dp_h^*} + D_h \frac{d\widehat{G}_l}{dR_h} \right) v' + \\
&\quad + \frac{\lambda_h^2}{\gamma} \left[ \left( \frac{\partial V_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial p_h^*} + \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{dt_h}{dp_h^*} \right].
\end{aligned} \tag{31}$$

Notice that both (29) and (30) make clear that in our model the only scope for a non-zero marginal matching grant for whichever local community has to be related to the effects on the binding self-selection constraints thwarting the federal government in its redistributive attempts.

<sup>14</sup>The fact that the changes in the policy variables chosen by the mimicking local community might be regarded as welfare-worsening by the local community itself is of course a consequence of the fact that a mimicking local community does not have full freedom of choice over its policy variables. At least some of them are in fact constrained to replicate the behavior of the mimicked local community.

<sup>15</sup>Notice that the values taken by  $\frac{d\widehat{t}_l}{dp_h^*}$  and  $\frac{d\widehat{t}_l}{dR_h}$  (and therefore, through the local community budget constraint, also the values taken by  $\frac{d\widehat{G}_l}{dp_h^*}$  and  $\frac{d\widehat{G}_l}{dR_h}$ ) are affected by the assumptions that are made on the possibility for the federal government to observe  $t_i$ . In particular, assuming that  $t_i$  is observable by the federal government restricts further the freedom of choice of a mimicking local community, for in order to be successful in masquerading its cost-type it has also to choose the same proportional tax rate chosen by the mimicked local community.

## 5 Summary

In this paper we have tried to integrate the literature on vertical fiscal externalities with the literature on publicly provided private goods as a means to mitigate the self-selection problems in the optimal income taxation model. We also explicitly modelled the informational asymmetry between local and federal level that often has been suggested as a motive for decentralization.

As a benchmark we have characterized the unitary solution under the assumption that the federal level has access to information about the cost structure in local communities. Under this assumption community specific nonlinear income taxes can be used. The results are similar to those given in Blomquist and Christiansen (2003). The high skill persons' choice between leisure and consumption are undistorted and they face a zero marginal income tax. Low skill individuals face a positive marginal income tax which consists of two parts. There is the usual "mimicking term", well known from many previous studies. However, since the low skill persons get day care for free from the public sector there is also a term reflecting the real cost of providing day care. Under the assumption that the cost structure is observable the unitary solution can be decentralized in a straightforward way.

If the costs of producing day care are not known to the federal level, a decentralized solution must be used. A vertical fiscal externality arises from co-occupancy of a common tax base (labor income) by central and local governments and the fact that local policy makers disregard the binding incentive-compatibility constraints that thwart the redistributive efforts of the central government. We have characterized the total marginal income tax rates (the sum of the federal and local marginal income taxes) for the different types of individuals. For each type of agent there is still the usual "mimicking" term, except for the high skilled in the low cost community, since, given the redistributive objectives of the federal government, no other agent can be tempted to misrepresent his type by mimicking them. Then for each agent there are two new types of terms. One reflects the increased subsidy cost for the federal level as an individual earns an income unit more. The other captures the effects on self selection constraints and individuals' well being that go through changes in the locally decided policy variables as an individual is induced to earn one more income unit. Hence, the pattern of total marginal income taxes in the decentralized solution will differ from that obtained in the unitary solution with observable costs.

Much of the focus of the paper is on the form of the matching grant. According to which criteria should the federal grant to the local levels be designed? We have studied both a linear and a nonlinear grant system. Here we concentrate on the nonlinear scheme and contrast our formulas with those obtained in previous studies as, for example, Dahlby and Wilson (2003). In our model the local level provides a private good — day care. The federal

level has a strong interest in local public provision of day care since this can mitigate the self-selection constraints in the federal level's optimal income taxation problem. Why would local politicians be interested in local public provision of day care? Leaving aside the incentives related to the subsidization of this kind of expenditure coming from the federal government, in our model the reason is that local policy makers are interested in redistribution and that public provision of day care is an instrument for them to pursue their redistributive goals. This implies that it is essential that the local level can not perform cash transfers! If the local level politicians had access to a linear income tax, they would pursue their redistribution via the income tax, and would not provide day care.

The federal level can influence the level of the local public provision by using matching grants. In Dahlby and Wilson (2003), which consider a productive good that affects the tax base for the federal level, the matching grant should be constructed in such a way that the grant reflects the marginal tax revenue that accrues to the federal level when the local public expenditures increase by one unit. If in our model the federal level used a linear income tax we would arrive at a similar result. However, since the federal level can use a nonlinear income tax it wants to govern the local level provision so that mimicking is deterred. In the nonlinear grant system the matching grants are set in such a way that the local communities self-select to the proper points on the nonlinear schedule. The matching grant for the low cost community should be set such that on the margin the deadweight loss from the subsidy exactly balances the marginal gains. These are related to the effects on the self-selection constraints involving as potential mimickers agents living in the low cost community. Since the low cost community must be discouraged from trying to pick the grant intended for the high cost community, the matching grant for the high cost community has a slightly more complicated form. The marginal cost of increasing it has the same form as for the low cost community. The gains are of two different types. On one hand there are gains in terms of slackening the self-selection constraint that requires the low cost community not to mimic the high cost community. On the other hand there are gains in terms of slackening the constraint requiring the high skill person in the high cost community not to mimic the low skill person in the low cost community. Hence, the gains of the matching grant are in our model of a totally different kind than in the Dahlby and Wilson (2003) framework.

One of the aims of our study was to see if public provision of private goods can serve to mitigate self-selection constraints in a model with several levels in the public sector. We have found, as explained on p. 14, that the role of public provision of private goods can in fact be even more important in a fiscal federalism setting than in a unitary model.

In common with many earlier fiscal federalism studies we find that if costs are observable, then the unitary solution can be decentralized. When costs

are not observable the unitary solution can not be decentralized; the optimal decentralized solution can not implement the unitary solution. Hence, it makes an important difference whether costs are observable or not. This illustrates that it is vital not only to argue that decentralization is needed because of an informational advantage at the local level. It is also important to model the informational asymmetry explicitly. Results on how to decentralize a unitary solution in a model where the informational asymmetry is not explicit might have very little to say about decentralization where the local level has more information than the central level.

## 6 Appendix

The quantities  $\delta_{Y_l^2}$ ,  $\delta_{B_l^2}$ ,  $\delta_{Y_h^2}$ ,  $\delta_{B_h^2}$ ,  $\delta_{Y_l^1}$ ,  $\delta_{B_l^1}$ ,  $\delta_{Y_h^1}$ ,  $\delta_{B_h^1}$  are defined as follows:

$$\begin{aligned} \delta_{Y_l^2} \equiv & \sum_{k=1}^2 \alpha^k \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial Y_l^2} + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial Y_l^2} + \\ & + \sum_{k=1}^2 \lambda_l^k \left[ \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial Y_l^2} + \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial Y_l^2} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial Y_l^2} - \left( \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} + \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial Y_l^2} \right] \end{aligned}$$

$$\begin{aligned} \delta_{B_l^2} \equiv & \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial B_l^2} + \\ & + \sum_{k=1}^2 \lambda_l^k \left[ \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial B_l^2} + \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial B_l^2} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial B_l^2} - \left( \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} + \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial B_l^2} \right] \end{aligned}$$

$$\begin{aligned} \delta_{Y_h^2} \equiv & \sum_{k=1}^2 \alpha^k \frac{\partial V_h^k}{\partial t_h} \frac{\partial t_h}{\partial Y_h^2} + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial Y_h^2} + \\ & + \lambda_h^2 \left[ \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial Y_h^2} + \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial Y_h^2} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial Y_h^2} - \left( \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} + \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial Y_h^2} \right] \end{aligned}$$

$$\begin{aligned} \delta_{B_h^2} \equiv & \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial B_h^2} + \\ & + \lambda_h^2 \left[ \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial B_h^2} + \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial B_h^2} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial B_h^2} - \left( \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} + \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial B_h^2} \right] \end{aligned}$$

$$\begin{aligned}\delta_{Y_l^1} &\equiv \sum_{k=1}^2 \alpha^k \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial Y_l^1} + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial Y_l^1} + \\ &+ \sum_{k=1}^2 \lambda_l^k \left[ \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial Y_l^1} + \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial Y_l^1} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial Y_l^1} - \left( \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} + \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial Y_l^1} \right]\end{aligned}$$

$$\begin{aligned}\delta_{B_l^1} &\equiv \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial B_l^1} + \\ &+ \sum_{k=1}^2 \lambda_l^k \left[ \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial B_l^1} + \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial B_l^1} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial B_l^1} - \left( \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} + \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \frac{\partial \bar{x}_l}{\partial B_l^1} \right]\end{aligned}$$

$$\begin{aligned}\delta_{Y_h^1} &\equiv \sum_{k=1}^2 \alpha^k \frac{\partial V_h^k}{\partial t_h} \frac{\partial t_h}{\partial Y_h^1} + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial Y_h^1} + \\ &+ \lambda_h^2 \left[ \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial Y_h^1} + \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial Y_h^1} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial Y_h^1} - \left( \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} + \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial Y_h^1} \right]\end{aligned}$$

$$\begin{aligned}\delta_{B_h^1} &\equiv \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial B_h^1} + \\ &+ \lambda_h^2 \left[ \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial B_h^1} + \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial B_h^1} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial B_h^1} - \left( \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} + \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \frac{\partial \bar{x}_h}{\partial B_h^1} \right]\end{aligned}$$

$$\begin{aligned}\delta_R &\equiv \sum_{i=l,h} \alpha^1 \frac{\partial V_i^1}{\partial t_i} \frac{\partial t_i}{\partial R} + \sum_{i=l,h} \alpha^2 \frac{\partial V_i^2}{\partial t_i} \frac{\partial t_i}{\partial R} + \\ &+ \sum_{k=1}^2 \lambda_l^k \frac{\partial \bar{x}_l}{\partial R} \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) + \\ &+ \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial R} + \sum_{k=1}^2 \lambda_l^k \frac{\partial t_l}{\partial R} \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) + \\ &+ \lambda_h^2 \frac{\partial \bar{x}_h}{\partial R} \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) + \\ &+ \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial R} + \lambda_h^2 \frac{\partial t_h}{\partial R} \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) - \gamma P^* \sum_{i=l,h} \frac{\partial D_i}{\partial \bar{x}_i} \frac{\partial \bar{x}_i}{\partial R}\end{aligned}$$

$$\begin{aligned}
\delta_{p^*} \equiv & \sum_{i=l,h} \alpha^1 \frac{\partial V_i^1}{\partial t_i} \frac{\partial t_i}{\partial p^*} + \sum_{i=l,h} \alpha^2 \frac{\partial V_i^2}{\partial t_i} \frac{\partial t_i}{\partial p^*} + \\
& + \sum_{k=1}^2 \lambda_l^k \frac{\partial \bar{x}_l}{\partial p^*} \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) + \\
& + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial p^*} + \sum_{k=1}^2 \lambda_l^k \frac{\partial t_l}{\partial p^*} \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) + \\
& + \lambda_h^2 \frac{\partial \bar{x}_h}{\partial p^*} \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) + \\
& + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial p^*} + \lambda_h^2 \frac{\partial t_h}{\partial p^*} \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) - \gamma p^* \sum_{i=l,h} \frac{\partial D_i}{\partial \bar{x}_i} \frac{\partial \bar{x}_i}{\partial p^*}.
\end{aligned}$$

## 6.1 Proof of Proposition 1

Dividing (12) by (13), multiplying the result by the r.h.s. of (13) and rearranging terms gives:

$$1 + \frac{\partial V_l^2}{\partial Y_l^2} / \frac{\partial V_l^2}{\partial B_l^2} = p^* \left( \frac{\partial D_l}{\partial Y_l^2} + MRS \frac{\partial D_l}{\partial B_l^2} \right) - \frac{\delta_{Y_l^2}}{\gamma} - MRS \frac{\delta_{B_l^2}}{\gamma}.$$

Exploiting the expression for the total marginal tax rate provided by (10) gives (22).

Dividing (14) by (15), multiplying the result by the r.h.s. of (15) and rearranging terms gives:

$$\begin{aligned}
1 + \frac{\partial V_h^2}{\partial Y_h^2} / \frac{\partial V_h^2}{\partial B_h^2} &= \frac{\lambda_l^2}{\gamma} \frac{\partial \widehat{V}_l^2}{\partial B_h^2} \left( \frac{\partial \widehat{V}_l^2}{\partial Y_h^2} - \frac{\partial V_h^2}{\partial Y_h^2} \right) + \\
& + p^* \left( \frac{\partial D_h}{\partial Y_h^2} + MRS_h^2 \frac{\partial D_h}{\partial B_h^2} \right) - \frac{\delta_{Y_h^2}}{\gamma} - MRS_h^2 \frac{\delta_{B_h^2}}{\gamma}.
\end{aligned}$$

Exploiting the expression for the total marginal tax rate provided by (10) gives (23).

Dividing (16) by (17), multiplying the result by the r.h.s. of (17) and rearranging terms gives:

$$\begin{aligned}
1 + \frac{\partial V_l^1}{\partial Y_l^1} / \frac{\partial V_l^1}{\partial B_l^1} &= \frac{\lambda_h^2}{\gamma} \frac{\partial \widehat{V}_h^2}{\partial B_l^1} \left( \frac{\partial \widehat{V}_h^2}{\partial Y_l^1} - \frac{\partial V_l^1}{\partial Y_l^1} \right) + \\
&+ p^* \left( \frac{\partial D_l}{\partial Y_l^1} + MRS_l^1 \frac{\partial D_l}{\partial B_l^1} \right) - \frac{\delta_{Y_l^1}}{\gamma} - MRS_l^1 \frac{\delta_{B_l^1}}{\gamma}.
\end{aligned}$$

Exploiting the expression for the total marginal tax rate provided by (10) gives (24).

Dividing (18) by (19), multiplying the result by the r.h.s. of (19) and rearranging terms gives:

$$\begin{aligned}
1 + \frac{\partial V_h^1}{\partial Y_h^1} / \frac{\partial V_h^1}{\partial B_h^1} &= \frac{\lambda_l^1}{\gamma} \frac{\partial \widehat{V}_l^1}{\partial B_h^1} \left( \frac{\partial \widehat{V}_l^1}{\partial Y_h^1} - \frac{\partial V_h^1}{\partial Y_h^1} \right) + \\
&+ p^* \left( \frac{\partial D_h}{\partial Y_h^1} + MRS_h^1 \frac{\partial D_h}{\partial B_h^1} \right) - \frac{\delta_{Y_h^1}}{\gamma} - \frac{\delta_{B_h^1}}{\gamma} MRS_h^1.
\end{aligned}$$

Exploiting the expression for the total marginal tax rate provided by (10) gives (25).

## 6.2 Proof of Proposition 2

Using (26) and (27), and since  $\frac{\partial t_i}{\partial p^*} = -\frac{D_i}{2}$ , f.o.c. (21) can be rewritten

$$\sum_{k=1}^2 Y_i^k$$

as:



$$\begin{aligned}
& -\alpha^1 \left( \frac{\partial V_h^1}{\partial t_h} \frac{D_h}{\sum_{k=1}^2 Y_h^k} + \frac{\partial V_l^1}{\partial t_l} \frac{D_l}{\sum_{k=1}^2 Y_l^k} \right) - \alpha^2 \left( \frac{\partial V_h^2}{\partial t_h} \frac{D_h}{\sum_{k=1}^2 Y_h^k} + \frac{\partial V_l^2}{\partial t_l} \frac{D_l}{\sum_{k=1}^2 Y_l^k} \right) + \\
& + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) \left( \frac{\partial \tilde{x}_l}{\partial p^*} - D_l \frac{\partial \bar{x}_l}{\partial R} \right) + \\
& + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \left( \frac{\partial \tilde{G}_l}{\partial p^*} - D_l \frac{\partial G_l}{\partial R} \right) - \frac{D_l}{\sum_{k=1}^2 Y_l^k} \left[ \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \right] + \\
& + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) \left( \frac{\partial \tilde{x}_h}{\partial p^*} - D_h \frac{\partial \bar{x}_h}{\partial R} \right) + \tag{32} \\
& + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \left( \frac{\partial \tilde{G}_h}{\partial p^*} - D_h \frac{\partial G_h}{\partial R} \right) - \lambda_h^2 \frac{D_h}{\sum_{k=1}^2 Y_h^k} \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) + \\
& - \gamma p^* \sum_{i=l,h} \frac{\partial D_i}{\partial \bar{x}_i} \left( \frac{\partial \tilde{x}_i}{\partial p^*} - D_i \frac{\partial \bar{x}_i}{\partial R} \right) \\
= & \gamma \sum_{i=l,h} D_i
\end{aligned}$$

Defining  $\delta_{\tilde{p}^*}$  as:

$$\begin{aligned}
\delta_{\tilde{p}^*} \equiv & \frac{\partial \tilde{x}_l}{\partial p^*} \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) + \\
& + \frac{\partial \tilde{G}_l}{\partial p^*} \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} + \\
& + \frac{\partial \tilde{x}_h}{\partial p^*} \lambda_h^2 \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) + \\
& + \frac{\partial \tilde{G}_h}{\partial p^*} \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} - \gamma p^* \sum_{i=l,h} \frac{\partial D_i}{\partial \bar{x}_i} \frac{\partial \tilde{x}_i}{\partial p^*},
\end{aligned}$$

using the definition of  $\Theta_i$  ( $\Theta_i \equiv -\delta_{R_i}/\gamma$ ) and given that  $\frac{\partial t_i}{\partial R} = \left( \sum_{k=1}^2 Y_i^k \right)^{-1}$ ,

f.o.c. (32) can be expressed as:

$$\delta_{\tilde{p}^*} + 2\gamma \text{cov}(D_i, \Theta_i) + 2\gamma E(D_i) E(\Theta_i) = \gamma \sum_{i=l,h} D_i. \quad (33)$$

$$\text{Since from (20) we have that } \sum_{i=l,h} \delta_{R_i} = -2\gamma \implies E(\Theta_i) = 1, \quad (33)$$

reduces to

$$\delta_{\tilde{p}^*} + 2\gamma \text{cov}(D_i, \Theta_i) = 0. \quad (34)$$

Substituting in (34) the definition of  $\delta_{\tilde{p}^*}$  and rearranging terms gives (28).

### 6.3 Proof of Proposition 3

The f.o.c. with respect to  $R_l$  and  $p_l^*$  are:

$$\delta_{R_l} = -\gamma \quad (35)$$

$$\delta_{p_l^*} = \gamma \sum_{k=1}^2 \min \left\{ \frac{Y_l^k}{w^k}, \bar{x}_l \right\} = \gamma D_l, \quad (36)$$

where:

$$\begin{aligned} \delta_{R_l} \equiv & (1 + \lambda) \sum_{k=1}^2 \alpha^k \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial R_l} + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial R_l} + \\ & + \sum_{k=1}^2 \lambda_l^k \frac{\partial \bar{x}_l}{\partial R_l} \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) + \\ & + \sum_{k=1}^2 \lambda_l^k \frac{\partial t_l}{\partial R_l} \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) - \gamma p_l^* \frac{\partial D_l}{\partial \bar{x}_l} \frac{\partial \bar{x}_l}{\partial R_l}; \end{aligned} \quad (37)$$

$$\begin{aligned} \delta_{p_l^*} \equiv & (1 + \lambda) \sum_{k=1}^2 \alpha^k \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial p_l^*} + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial p_l^*} + \\ & + \sum_{k=1}^2 \lambda_l^k \frac{\partial \bar{x}_l}{\partial p_l^*} \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) + \\ & + \sum_{k=1}^2 \lambda_l^k \frac{\partial t_l}{\partial p_l^*} \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) - \gamma p_l^* \frac{\partial D_l}{\partial \bar{x}_l} \frac{\partial \bar{x}_l}{\partial p_l^*}. \end{aligned} \quad (38)$$

Multiplying (35) by  $-D_l$  and using the definition of  $\delta_{R_l}$  provided by (37) gives:

$$\begin{aligned}
\gamma D_l &= -(1 + \lambda) D_l \sum_{k=1}^2 \alpha^k \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial R_l} - D_l \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial R_l} + \\
&\quad - D_l \sum_{k=1}^2 \lambda_l^k \frac{\partial \bar{x}_l}{\partial R_l} \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) + \\
&\quad - D_l \sum_{k=1}^2 \lambda_l^k \frac{\partial t_l}{\partial R_l} \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) + \gamma D_l p_l^* \frac{\partial D_l}{\partial \bar{x}_l} \frac{\partial \bar{x}_l}{\partial R_l}. \tag{39}
\end{aligned}$$

From (36) and (38) we have:

$$\begin{aligned}
\gamma D_l &= (1 + \lambda) \sum_{k=1}^2 \alpha^k \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial p_l^*} + \sum_{k=1}^2 \lambda_l^k \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) \frac{\partial t_l}{\partial G_l} \frac{\partial G_l}{\partial p_l^*} + \\
&\quad + \sum_{k=1}^2 \lambda_l^k \frac{\partial \bar{x}_l}{\partial p_l^*} \left( \frac{\partial V_l^k}{\partial \bar{x}_l} + \frac{\partial V_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial \bar{x}_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \frac{\partial t_l}{\partial \bar{x}_l} \right) + \\
&\quad + \sum_{k=1}^2 \lambda_l^k \frac{\partial t_l}{\partial p_l^*} \left( \frac{\partial V_l^k}{\partial t_l} - \frac{\partial \widehat{V}_l^k}{\partial t_l} \right) - \gamma p_l^* \frac{\partial D_l}{\partial \bar{x}_l} \frac{\partial \bar{x}_l}{\partial p_l^*}. \tag{40}
\end{aligned}$$

Finally, combining (39) and (40) and using  $\frac{\partial \bar{x}_l}{\partial p^*} = \frac{\partial \widehat{\bar{x}}_l}{\partial p^*} - D_l \frac{\partial \bar{x}_l}{\partial R_l}$  and  $\frac{\partial G_l}{\partial p^*} = \frac{\partial \widehat{G}_l}{\partial p^*} - D_l \frac{\partial G_l}{\partial R_l}$  gives (29).

#### 6.4 Proof of Proposition 4

The f.o.c. with respect to  $R_h$  and  $p_h^*$  are:

$$\delta_{R_h} = -\gamma \tag{41}$$

$$\delta_{p_h^*} = \gamma \sum_{k=1}^2 \min \left\{ \frac{Y_h^k}{w^k}, \bar{x}_h \right\} = \gamma D_h, \tag{42}$$

where:

$$\begin{aligned}
\delta_{R_h} &\equiv \sum_{k=1}^2 \alpha^k \frac{\partial V_h^k}{\partial t_h} \frac{\partial t_h}{\partial R_h} + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial R_h} + \\
&+ \lambda_h^2 \frac{\partial \bar{x}_h}{\partial R_h} \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) + \\
&+ \lambda_h^2 \frac{\partial t_h}{\partial R_h} \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) - \gamma p_h^* \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \bar{x}_h}{\partial R_h} + \\
&- \lambda \sum_{k=1}^2 \alpha^k \left[ \frac{\partial V^k}{\partial \widehat{t}_l} \frac{d\widehat{t}_l}{dR_h} + \frac{\partial V^k}{\partial \widehat{x}_l} \frac{d\widehat{x}_l}{dR_h} \right] - \lambda \frac{d\widehat{G}_l}{dR_h} v'; \tag{43}
\end{aligned}$$

$$\begin{aligned}
\delta_{p_h^*} &\equiv \sum_{k=1}^2 \alpha^k \frac{\partial V_h^k}{\partial t_h} \frac{\partial t_h}{\partial p_h^*} + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial p_h^*} + \\
&+ \lambda_h^2 \frac{\partial \bar{x}_h}{\partial p_h^*} \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) + \\
&+ \lambda_h^2 \frac{\partial t_h}{\partial p_h^*} \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) - \gamma p_h^* \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \bar{x}_h}{\partial p_h^*} + \\
&- \lambda \sum_{k=1}^2 \alpha^k \left[ \frac{\partial V^k}{\partial \widehat{t}_l} \frac{d\widehat{t}_l}{dp_h^*} + \frac{\partial V^k}{\partial \widehat{x}_l} \frac{d\widehat{x}_l}{dp_h^*} \right] - \lambda \frac{d\widehat{G}_l}{dp_h^*} v'. \tag{44}
\end{aligned}$$

Multiplying (41) by  $-D_h$  and using the definition of  $\delta_{R_h}$  provided by (43) gives:

$$\begin{aligned}
\gamma D_h &= -D_h \sum_{k=1}^2 \alpha^k \frac{\partial V_h^k}{\partial t_h} \frac{\partial t_h}{\partial R_h} - \lambda_h^2 D_h \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial R_h} + \\
&- \lambda_h^2 D_h \frac{\partial \bar{x}_h}{\partial R_h} \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) + \\
&- \lambda_h^2 \frac{\partial t_h}{\partial R_h} D_h \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) + \gamma D_h p_h^* \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \bar{x}_h}{\partial R_h} + \\
&+ \lambda D_h \sum_{k=1}^2 \alpha^k \left[ \frac{\partial V^k}{\partial \widehat{t}_l} \frac{d\widehat{t}_l}{dR_h} + \frac{\partial V^k}{\partial \widehat{x}_l} \frac{d\widehat{x}_l}{dR_h} \right] + \lambda D_h \frac{d\widehat{G}_l}{dR_h} v'. \tag{45}
\end{aligned}$$

From (42) and (44) we have:

$$\begin{aligned}
\gamma D_h = & \sum_{k=1}^2 \alpha^k \frac{\partial V_h^k}{\partial t_h} \frac{\partial t_h}{\partial p_h^*} + \lambda_h^2 \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) \frac{\partial t_h}{\partial G_h} \frac{\partial G_h}{\partial p_h^*} + \\
& + \lambda_h^2 \frac{\partial \bar{x}_h}{\partial p_h^*} \left( \frac{\partial V_h^2}{\partial \bar{x}_h} + \frac{\partial V_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial \bar{x}_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \frac{\partial t_h}{\partial \bar{x}_h} \right) + \\
& + \lambda_h^2 \frac{\partial t_h}{\partial p_h^*} \left( \frac{\partial V_h^2}{\partial t_h} - \frac{\partial \widehat{V}_h^2}{\partial t_h} \right) - \gamma p_h^* \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \bar{x}_h}{\partial p_h^*} + \\
& - \lambda \sum_{k=1}^2 \alpha^k \left[ \frac{\partial V^k}{\partial \widehat{t}_l} \frac{d\widehat{t}_l}{dp_h^*} + \frac{\partial V^k}{\partial \widehat{x}_l} \frac{d\widehat{x}_l}{dp_h^*} \right] - \lambda \frac{d\widehat{G}_l}{dp_h^*} v'. \tag{46}
\end{aligned}$$

Finally, combining (45) and (46) and using  $\frac{\partial \bar{x}_h}{\partial p^*} = \frac{\partial \widehat{\bar{x}}_h}{\partial p^*} - D_h \frac{\partial \bar{x}_h}{\partial R_h}$  and  $\frac{\partial G_h}{\partial p^*} = \frac{\partial \widehat{G}_h}{\partial p^*} - D_h \frac{\partial G_h}{\partial R_h}$  gives (30), where for instance  $\frac{d\widehat{x}_l}{dp_h^*} \equiv \frac{\partial \widehat{x}_l}{\partial D_h} \frac{\partial D_h}{\partial \bar{x}_h} \frac{\partial \bar{x}_h}{\partial p_h^*}$ .

## References

- [1] Aronsson, T., and M. Wikström (2003), Optimal taxation and risk-sharing arrangements in an economic federation, *Oxford Economic Papers* 55, 104-120.
- [2] Aronsson, T., and S. Blomquist (2004), Redistribution and provision of public goods in an economic federation, Working Paper No. 2004:4, Department of Economics, University of Uppsala (also downloadable at the web site [http://www.nek.uu.se/pdf/wp2004\\_4.pdf](http://www.nek.uu.se/pdf/wp2004_4.pdf)).
- [3] Atkinson, A. B., and J. E. Stiglitz (1980), *Lectures on public economics*, New York: McGraw-Hill.
- [4] Bergstrom, T., and S. Blomquist (1996), The political economy of subsidized day care, *European Journal of Political Economy* 12, 443-458.
- [5] Blomquist, S., and V. Christiansen (1995), Public provision of private goods as a redistributive device in an optimum income tax model, *Scandinavian Journal of Economics* 97, 547-567.
- [6] Blomquist, S., and V. Christiansen (2003), Is there a case for public provision of private goods if preferences are heterogeneous? An example with day care, CES-ifo working paper no. 938.
- [7] Boadway, R., and M. Marchand (1995), The use of public expenditure for redistributive purposes, *Oxford Economic Papers* 47, 45-59.

- [8] Boadway, R., and M. J. Keen (1996), Efficiency and the optimal direction of federal-state transfers, *International Tax and Public Finance* 3, 137-155.
- [9] Boadway, R., M. Marchand and M. Vigneault (1998), The consequences of overlapping tax bases for redistribution and public spending in a federation, *Journal of Public Economics* 68, 453-478.
- [10] Cornes, R., and E. D. C. Silva (2000), Local public goods, risk sharing and private information in federal systems, *Journal of Urban Economics* 47, 39-60.
- [11] Cremer, H., and F. Gahvari (1997), In-kind transfers, self-selection and optimal tax policy, *European Economic Review* 41, 97-114.
- [12] Dahlby, B. (1996), Fiscal externalities and the design of intergovernmental grants, *International Tax and Public Finance* 3, 397-412.
- [13] Dahlby, B., and L. S. Wilson (2003), Vertical fiscal externalities in a federation, *Journal of Public Economics* 87, 917-930.
- [14] Keen, M. J. (1998), Vertical tax externalities in the theory of fiscal federalism, *IMF Staff Papers* 45 (3), 454-485.
- [15] Lockwood, B. (1999), Inter-regional insurance, *Journal of Public Economics* 72, 1-37.
- [16] Stern, N. H. (1982), Optimum taxation with errors in administration, *Journal of Public Economics* 17, 181-211.
- [17] Stiglitz, J. E. (1982), Self-selection and Pareto-efficient taxation, *Journal of Public Economics* 17, 213-240.
- [18] Wellish, D. (2000), *Theory of public finance in a federal state*, New York: Cambridge University Press.

## WORKING PAPERS\*

Editor: Nils Gottfries

- 2003:14 Karin Edmark, The Effects of Unemployment on Property Crime: Evidence from a Period of Unusually Large Swings in the Business Cycle. 27 pp.
- 2003:15 Patrik Hesselius, Does Sick Absence Increase the Risk of Unemployment? 32 pp.
- 2003:16 Mohammad Taslimi, Optimal Earnings-Related Unemployment Benefits. 30 pp.
- 2003:17 Peter Fredriksson and Bertil Holmlund, Optimal Unemployment Insurance Design: Time Limits, Monitoring, or Workfare? 25 pp.
- 2003:18 Yvonne Åberg, Peter Hedström and Ann-Sofie Kolm, Social Interactions and Unemployment. 36 pp.
- 2003:19 Tobias Lindhe, Jan Södersten and Ann Öberg, Economic Effects of Taxing Different Organizational Forms under a Dual Income Tax. 22 pp.
- 2003:20 Pär Österholm, The Taylor Rule – A Spurious Regression? 28 pp.
- 2003:21 Pär Österholm, Testing for Cointegration in Misspecified Systems – A Monte Carlo Study of Size Distortions. 32 pp.
- 2003:22 Ann-Sofie Kolm and Birthe Larsen, Does Tax Evasion Affect Unemployment and Educational Choice? 36 pp.
- 2003:23 Daniel Hallberg, A Description of Routes Out of the Labor Force for Workers in Sweden. 50 pp.
- 2003:24 N. Anders Klevmarken, On Household Wealth Trends in Sweden over the 1990s. 20 pp.
- 2003:25 Mats A. Bergman, When Should an Incumbent Be Obligated to Share its Infrastructure with an Entrant Under the General Competition Rules? 21 pp.
- 2003:26 Niclas Berggren and Henrik Jordahl, Does Free Trade Really Reduce Growth? Further Testing Using the Economic Freedom Index. 19 pp.
- 2003:27 Eleni Savvidou, The Relationship Between Skilled Labor and Technical Change. 44 pp.
- 2003:28 Per Pettersson-Lidbom and Matz Dahlberg, An Empirical Approach for Evaluating Soft Budget Constraints. 31 pp.
- 2003:29 Nils Gottfries, Booms and Busts in EMU. 34 pp.

---

\* A list of papers in this series from earlier years will be sent on request by the department.

- 2004:1 Iida Häkkinen, Working while enrolled in a university: Does it pay? 37 pp.
- 2004:2 Matz Dahlberg, Eva Mörk and Hanna Ågren, Do Politicians' Preferences Correspond to those of the Voters? An Investigation of Political Representation. 34 pp.
- 2004:3 Lars Lindvall, Does Public Spending on Youths Affect Crime Rates? 40 pp.
- 2004:4 Thomas Aronsson and Sören Blomquist, Redistribution and Provision of Public Goods in an Economic Federation. 23 pp.
- 2004:5 Matias Eklöf and Daniel Hallberg, Private Alternatives and Early Retirement Programs. 30 pp.
- 2004:6 Bertil Holmlund, Sickness Absence and Search Unemployment. 38 pp.
- 2004:7 Magnus Lundin, Nils Gottfries and Tomas Lindström, Price and Investment Dynamics: An Empirical Analysis of Plant Level Data. 41 pp.
- 2004:8 Maria Vredin Johansson, Allocation and Ex Ante Cost Efficiency of a Swedish Subsidy for Environmental Sustainability: The Local Investment Program. 26 pp.
- 2004:9 Sören Blomquist and Vidar Christiansen, Taxation and Heterogeneous Preferences. 29 pp.
- 2004:10 Magnus Gustavsson, Changes in Educational Wage Premiums in Sweden: 1992-2001. 36 pp.
- 2004:11 Magnus Gustavsson, Trends in the Transitory Variance of Earnings: Evidence from Sweden 1960-1990 and a Comparison with the United States. 63 pp.
- 2004:12 Annika Alexius, Far Out on the Yield Curve. 41 pp.
- 2004:13 Pär Österholm, Estimating the Relationship between Age Structure and GDP in the OECD Using Panel Cointegration Methods. 32 pp.
- 2004:14 Per-Anders Edin and Magnus Gustavsson, Time Out of Work and Skill Depreciation. 29 pp.
- 2004:15 Sören Blomquist and Luca Micheletto, Redistribution, In-Kind Transfers and Matching Grants when the Federal Government Lacks Information on Local Costs. 34 pp.

See also working papers published by the Office of Labour Market Policy Evaluation  
<http://www.ifau.se/>