

Where Should the Elderly Live and and Who Should Pay for their Care? A Study in Demographics and Geographical Economics

Thomas Aronsson, Sören Blomquist*and Luca Micheletto[†]
Umeå University, University of Uppsala
and University “L. Bocconi”, Milan

October 10, 2005

Abstract

There is a rich literature analyzing the problems that will arise as the share of elderly and retired in the population will rise in the near future. However, the locational decisions among the elderly as well as their implications in terms of taxes/transfers and the allocation of responsibilities or elderly care between the federal and local levels have not received much attention. In this paper we aim at investigating these issues.

We explore a model with congestion and agglomeration effects in production in the big city. This gives rise to an unbalanced demographic structure with a large proportion of young in the big city and a high proportion of elderly in small villages in the country side. We study and characterize the inefficiencies that arise because of individuals' free location choice. Within a fiscal federalism structure we also investigate what instruments that are needed in order to decentralize the second best policy.

*Nationalekonomiska institutionen, Ekonomikum, Uppsala University, Kyrkogårdsgatan 10, P.O. Box 513, SE-751 20, Uppsala, Sweden. E-mail address: soren.blomquist@nek.uu.se; tel.: ++46 18 4711102; fax: ++46 18 4711478.

[†]Corresponding author. Istituto di Economia Politica, Università “L. Bocconi”, via U. Gobbi 5, 20136 Milano, Italy. E-mail address: luca.micheletto@uni-bocconi.it; tel. ++39 02 58365324; fax ++39 02 58365318.

1 Introduction

In many developed countries, the share of elderly and retired in the population will rise considerably in the near future. This aging of the population poses several problems, and there is a rich literature analyzing them. However, the locational decisions among the elderly, as well as their implications in terms of taxes/transfers and the allocation of responsibilities for elderly care between the federal and local levels of government, have not received much attention so far. In this paper, we aim at investigating these issues.

A basic observation in several countries is the tendency towards a geographically unbalanced demographic structure, with the “villages” being populated mainly by elderly people, while the “big city” to a greater extent is populated by people in working ages. One reason for this may be a life-cycle pattern of mobility behavior. For instance, people may live in the small communities, where they were born, until they have finished high school. Then, they move to the “big cities” and work for many years. When becoming old, quite a few return back to the region from where they came or to a region where they consider the living conditions (climate) to be pleasurable. Another reason is that the mobility pattern is part of a long-term transition process, which has been going on for at least a hundred years, where the country side is depopulated and the big cities increase in size. Those who move away from the country side typically do so when young. This also contributes to a geographically unbalanced demographic structure.

Clearly, a geographically unbalanced demographic structure may be perfectly consistent with the notion of efficiency. The basic argument is that it may be cheaper to provide elderly care at the country side than in the big city. Elderly living in the big city contribute to congestion, higher land prices, etc., without adding to the production capacity. We will elaborate on this in more detail below. In countries with publicly provided elderly care, as in the Nordic countries, this unbalanced demographic structure can lead to financial problems for small villages, since the local taxable income might be too small to finance high quality care for the elderly. In fact, in many small Swedish municipalities, the taxable income is so small that they would not be able to provide such care without large transfers from the central government.¹

Transfers to the poor parts of a country are often motivated by income redistribution arguments, and individuals living in the big cities sometimes complain about these transfers. However, given that elderly care is publicly provided, and if it is cheaper to provide elderly care at the country side than in the big city, it may be better for those working in the big cities to transfer money to the country side, allowing good care for the elderly there, instead

¹Korpi (2003) describes a very unbalanced demographic structure in Sweden, where the population in the countryside largely consists of elderly people, while Stockholm has a population with a larger share of people in working ages.

of having the elderly living in the big cities, in which case they would still have to pay for their care via taxes. One can, therefore, provide efficiency arguments for transfers from the big city to the country side.

Throughout the paper, we assume that elderly care is publicly provided. Without going into a detailed argument of why this might be a good policy, we want to point to three motivations for considering publicly provided care for the elderly. First, it is a common phenomenon in many countries that some form of basic elderly care is publicly provided, often at the local level. Second, the same types of arguments as those proposed in Bergstrom and Blomquist (1996) for public subsidies to day care are also valid for elderly care, as are the arguments put forward by Boadway and Marchand (1995), Blomquist and Christiansen (1995) and Cremer and Ghavari (1997) for public provision of certain private goods. Third, some of the arguments used for explaining public pensions are also valid for public provision of elderly care.

There are two salient differences between the big city and the country side that we want to capture in our model. First, there is congestion in the big city and no congestion at the country side. This congestion can take several forms like bad air quality and/or scarce land resources leading to high costs for housing and long times for commuting. Retired and earners contribute in the same way towards the congestion. Second, there are agglomeration effects in the big city, in the sense that the marginal and average products of workers increase with the number of people working in the consumption goods industry. Workers contribute to the agglomeration effects, whereas the retired do not. An important aspect of our model is that the retired are characterized by an attachment to home, meaning that some of them have an attachment to the big city, whereas others have an attachment to a village.

Our paper also relates to a large, and growing, literature on fiscal federalism dealing with the interactions within the public sector, as well as how a federal government must act in order to implement efficiency aspects of a unitary resource allocation. Several earlier studies concentrate on fiscal external effects, and a distinction is commonly made between horizontal² and vertical³ external effects. Our paper, on the other hand, focuses the main attention on redistribution and population mobility. An important reference here is Boadway et al. (2003) dealing with redistribution and equalization in

²A standard reference here is Oates (1972). Wildasin (1991) shows how horizontal external effects associated with mobility can be internalized by means of a system of matching grants from the central to the local governments.

³Vertical external effects may arise from co-occupancy of a common tax base. Typically, the local authorities do not recognize that their policies affect the central authority's tax base. This was pointed out by e.g. Hansson and Stuart (1987) and Johnson (1988). Methods to internalize vertical fiscal external effects have been discussed by e.g. Boadway and Keen (1996), Dahlby (1996), Boadway et al. (1998), Sato (2000) and Aronsson and Wikström (2001).

the context of an economic federation.⁴ In their study, individuals differ in ability, and all ability-types are mobile across local jurisdictions. Their main contribution is to characterize the behavior of the local and central governments in the context of a federal decision-structure, and then compare the resulting outcome with a unitary optimum where all decisions about taxes and public expenditures are made by the central government.

The rest of the paper is organized as follows. In Section 2, we present the basic assumptions of our model and how individuals make their locational decisions. As a benchmark we in section 3 describe the first best, where the policy maker can decide on each individual’s consumption, the quality of the elderly care as well as each person’s location. One of the starting points for the literature on the effects of mobility, Tiebout (1956), made strong claims that free mobility leads to an efficient allocation. Within our framework we obtain the opposite result; free mobility leads to inefficiencies! Given, the popular view that free mobility of individuals enhance efficiency we believe it is of interest to see how free mobility among the individuals hampers the possibilities to achieve the first best, and how the incentives of the individuals differ from those of the social planner (due to the congestion and agglomeration effects). Thus, in Section 4, we consider how results are affected, if the policy maker is free to set all variables except those associated with the individuals’ locations.⁵ Section 5, finally, investigates to what extent decentralization of the decisions about elderly care results in a less efficient allocation of agents between the big city and the country side. Our focus will be on how the grant system can be used to achieve a better resource allocation.

2 General model and individuals’ location decisions

There is one big city and a “sufficiently” large number of small villages. There are two types of individuals (to have three or more types would only complicate matters without adding any new qualitative insights); productive people (e) that work and earn an income and retired people (r) who need care. Within each type, everyone is identical. Individuals can either live in the big city or in a village. We use the notations N_b^r , N_b^e , N_v^r and N_v^e to represent the number of retired living in the big city, the number of earners living in the big city, the number of retired living in a village and the number of earners living in a village, respectively. In addition, we define

⁴See also Boadway et al. (1998).

⁵We often study distortions that are due to the fact that individuals can choose their hours of work freely. Locational choices are then usually neglected. Here we do the reverse, we study the distortions that arise because individuals can choose where to live freely. To keep things simple we assume hours of work are fixed.

$N_b \equiv N_b^r + N_b^e$ and $N_v \equiv N_v^r + N_v^e$.

To keep the model tractable, we disregard the labor supply decision; each earner supplies one unit of labor inelastically. The utility of earners only depends on consumption, $u(c)$, and the only decision made by an earner is where to live. On the other hand, the retired both consume the consumption good c and elderly care. Their utility is given by $u(c) + \phi(q)$, where q is the quality of the care. The quality of elderly care is assumed to be directly proportional to the number of earners used to take care of each elderly.

In the city, there is a fixed number of (many) firms, each with a constant returns to scale production function. However, due to agglomeration effects, the marginal and average products of labor are increasing in the total number of individuals working in the consumption goods industry, i.e. the marginal product is given by $w = F(N_b^e - qN_b^r)$ with $F' > 0$. Here, qN_b^r denotes the number of earners that is needed to take care of the elderly living in the big city for a given level, q , of the quality of the service provided. Thus, the elderly and those taking care of the elderly do not contribute to the agglomeration effects. In the big city, there is also congestion given by the function $m(N_b^e + N_b^r)$, which gives the congestion imposed on each person living there; $m < 0$, $m' < 0$.

In the villages, there is neither congestion nor agglomeration effects. The labor required to take care of N_v^r retired people is qN_v^r . The production of the consumption good is given by the constant returns to scale technology, $Q_v = \theta(N_v^e - qN_v^r)$. Since there is no congestion, it will be cheaper to provide elderly care in a village than in the big city.

Much of the focus in this paper will be on the individuals' location decisions. For earners, the utility difference between living in a village and living in the big city is given by

$$\Delta^e = u(c_v^e) - u(c_b^e) - m(N_b). \quad (1)$$

Therefore, if $\Delta^e > 0$, the earner wants to live in a village; if $\Delta^e = 0$, he/she is indifferent; and if $\Delta^e < 0$, he/she wants to live in the big city.

We assume that retired people have an attachment to home, and that this attachment varies in strength among the retired. This is accomplished by following the approach of Wellisch (1994). Ranking the retired according to how strong their attachment is to a village in such a way that the individual with strongest attachment is numbered N_1^r , we obtain for individual j :

$$\Delta_j^r = u(c_v^r) + \phi(q_v) + h(N^r - N_j^r) - u(c_b^r) - \phi(q_b) - hN_j^r - m(N_b). \quad (2)$$

If $\Delta_j^r > 0$, the individual wants to live in a village; if $\Delta_j^r = 0$, he/she is indifferent; and if $\Delta_j^r < 0$, he/she wants to live in the big city.

3 First best analysis

The policy maker maximizes a utilitarian social welfare function. To facilitate comparisons with the second best analysis we perform later we use lump sum taxes T_v^e, T_b^e to control the consumptions c_v^e, c_b^e of earners in the village and the big city. Given, the use of these instruments the policy maker's maximization problem can be written as

$$\begin{aligned} & \max_{c_v^e, c_b^e, T_v^e, T_b^e, q_v, q_b, N_v^r, N_v^e} N_v^r [u(c_v^r) + \phi(q_v)] + (N^r - N_v^r) [u(c_b^r) + \phi(q_b)] + N_v^e u(\theta - T_v^e) + \\ & + (N^e - N_v^e) u \{F [N^e - N_v^e - q_b (N^r - N_v^r)] - T_b^e\} + \\ & + h \sum_{n=1}^{N_v^r} (N^r - n) + h \sum_{n=1+N_v^r}^{N^r} n + (N^e - N_v^e + N^r - N_v^r) m (N^e - N_v^e + N^r - N_v^r) \end{aligned}$$

subject to:

$$(N^e - N_v^e) T_b^e + N_v^e T_v^e - (N^r - N_v^r) \{c_b^r + q_b F [N^e - N_v^e - q_b (N^r - N_v^r)]\} - N_v^r (c_v^r + \theta q_v) \geq 0$$

$$N_v^e - q_v N_v^r \geq 0$$

Let μ denote the Lagrange multiplier associated with the government's budget constraint and γ the Lagrange multiplier associated with the other constraint, requiring that the number of earners in the village must at least be sufficient to take care of the elderly living there. We assume there exists an interior solution to this problem in the sense that there will be a nonzero population of earners and retired both in the representative village and in the big city. The first order conditions are presented in the Appendix.

Here, we summarize the main result characterizing the first best and give the basic intuition behind it.

Proposition 1 *The first best resource allocation is characterized by*

$$u'(c_b^r) = u'(c_v^r) = u'(c_b^e) = u'(c_v^e) = \mu, \quad (3)$$

$$\phi'(q_v) = \mu\theta + \gamma, \quad (4)$$

$$\phi'(q_b) = \mu [F + (N_b^e - q_b N_b^r) F'] \quad (5)$$

as well as by the following conditions for N_v^r and N_v^e , respectively:

$$\phi(q_v) - \phi(q_b) + h(N^r - 2N_v^r) - (N_b m' + m) = q_v(\gamma + \mu\theta) - \mu q_b [F + (N_b^e - q_b N_b^r) F'], \quad (6)$$

$$\gamma - (N_b m' + m) = \mu [F + (N_b^e - q_b N_b^r) F' - \theta]. \quad (7)$$

The proof of Proposition 1 is straightforward; see the Appendix.

The consumption good is perfectly transferable. Hence, it is no surprise that the first best is characterized by marginal utility of consumption being equal across all individuals(3). Elderly care is not transferable. What is produced in the village can not be transferred to the big city and vice versa. Conditions (4) and (5) imply that the quality of elderly care should be set so that the marginal utility of elderly care (left hand side of the eqs.) equals the marginal cost (right hand side of the eqs.). From the condition (7) it follows that $\mu [F + (N_b^e - q_b N_b^r) F'] > \mu\theta + \gamma$. The marginal cost is higher in the big city than in the village. Conditions (4) and (5) therefore require that the quality level of elderly care in the country side should be higher than in the big city: $q_v > q_b$.

Our focus in this paper is, to a large extent, on the allocation of earners and retired, respectively, between the big city and villages. Considering eq. (7) we see that the policy maker sets the number of earners in the village in such a way that the sum of the lagrange multiplier γ and the marginal reduction in congestion when a marginal earner is moved to a village exactly balances the net reduction in production. This implies that an individual earner and the policy maker evaluates the locational benefits and costs differently. The relative evaluation of living in a village versus in the big city is for an earner given by (1).

Since we in the first best have equalized consumption for earners in a village and the big city the only remaining term in (1) is the congestion term. Earners therefore would prefer to live in a village, given the allocation of consumption in first best. By combining equations (1) and (7), we see that the policy maker values the village versus the big city according to

$$\Delta_e^p = \Delta^e - N_b m' + \gamma - \mu [F + (N_b^e - q_b N_b^r) F' - \theta]. \quad (8)$$

From the point of view of the policy maker, therefore, there are three additional effects in comparison with the decision rule facing the earner. First, the policy maker also recognizes that an extra person in the village reduces congestion in the big city, $N_b m'$. This effect works in the direction of allocating more earners to the villages. Second, if in the country side the number of earners is just sufficient to take care of the elderly and no production of the consumption good c takes place locally, an additional earner is valuable also because it weakens the binding γ -constraint. Third, an increase in the number of persons living in a village reduces the benefits associated with agglomeration in the big city, which is captured by the term $F + (N_b^e - q_b N_b^r) F' - \theta$.⁶

⁶ $F - \theta$ is the productivity differential of workers between the big city and the country side. When an earner moves from the big city to the country side, the net wage rate of those who remain the big city is lowered due to the way agglomeration effects work. On the other hand, given that the earners' wage rate is reduced, also the cost for the

Equation (6) implicitly defines the optimal number of retired persons living in the village. The interpretation is that the benefit of moving an additional retiree from the big city to the village (the left hand side) exactly balances the cost (the right hand side) of such a reallocation. We can solve equation (6) to derive the optimal number of retired living in the village explicitly. This is given by:

$$N_v^r = \frac{N^r}{2} + \frac{\phi(q_v) - (q_b) - (N_b m' + m) + \mu q_b [F + (N_b^e - q_b N_b^r) F'] - q_v (\gamma + \mu \theta)}{2h}. \quad (9)$$

The second term on the right hand side of equation (9) measures the extent to which the optimal distribution of elderly people deviates from the one where exactly the same number of elderly lives in the big city and in the villages. The sign of the deviation depends on the sign of the numerator. If the numerator is positive (negative), relatively more (less) elderly should be moved to the village. Note also that the terms in the numerator reflect the welfare and productivity gains of a marginal reshuffling of elderly from the big city to the village, disregarding any cost related to the attachment to home component of the preferences of the elderly. This is the reason for discounting by $2h$; the greater the importance for people of the attachment to home element, the more respectful of this preference should be the optimal allocation in space of the elderly, despite any other kind of gains potentially achievable distorting it.

According to equation (6), which implicitly defines the allocation of retired people, the benefit of moving an additional retiree from the big city to the village (the left hand side of the equation) should exactly balance the cost (represented by the right hand side of the equation) of such a reallocation. The first group of terms ($\phi(q_v) - (q_b) + h(N^r - 2N_v^r)$) on the left hand side represents the net welfare gain in terms of increased utility of elderly care and attachment to home. Then there is also the gain in terms of reduced congestion in the city $-(N_b m' + m)$. On the right hand side we have the change in the cost of providing elderly care as the marginal retired is moved to a village. On one hand there is the cost of providing the marginal retired elderly care in the village, but there is also a decrease in the resource cost for the elderly care in the big city.

It may also be of interest to compare the policy maker's valuation of the marginal retired person moving from the big city to a village with the valuation made by the marginal individual himself/herself. We can write the policy maker's valuation as

government of provision of elderly care in the big city is reduced. Thus, $(N_b^e - q_b N_b^r) F'$ represents the net effect for the government budget of on one hand lowering the tax rate on earners in the big city to keep their disposable income unaffected ($N_b^e F'$), and on the other hand of the savings related to a smaller cost of provision of elderly care at the city level ($-q_b N_b^r F'$).

$$\Delta_r^p = \Delta^r - N_b m' - \gamma q_v + \underbrace{\mu [q_b F + q_b (N_b^e - q_b N_b^r) F' - q_v \theta]}_{\Xi}. \quad (10)$$

In comparison with the valuation made by the marginal individual himself/herself, the policy maker recognizes also in this case three additional effects. First, the congestion effect $-N_b m'$ (which contributes to reduce the population in the big city). Second, the cost $-\gamma q_v$ related to the need of also moving additional earners in the country side or of lowering the quality of elderly care provided there if the γ -constraint turns out to be binding. Third, the effect for the government's budget coming from the working of the agglomeration forces (the term labelled Ξ).⁷ number of retired in the as well as the relative cost of providing elderly care to the marginal retiree in a village and the big city. Depending on the relative size of q_b and q_v this second component can be either positive or negative. Therefore, we cannot in general sign Δ^r ; this would require additional assumptions. For instance, for sufficiently strong increases in the congestion effect, Δ^r would be negative and the marginal retiree would like to move from a village to the big city. However, it is in general possible that Δ^r is positive, implying that the marginal retired would like to move from the big city to a village.

To sum up, in the results corresponding to the first best the marginal utility of consumption should be equal across individuals. Due to congestion in the big city the marginal cost of elderly care will be lower in the village and the quality of elderly care will because of this be higher in the village than in the big city. The allocation of individuals will be such that earners would like to move to a village. In general there would be retired that would like to move, but we cannot tell whether there would be retired in the big city that would like to move to a village or vice versa.

4 Second best analysis

We have seen that the first best has an allocation of individuals such that at least part of the population would like to move. Such an outcome is of course not sustainable in a setting where agents can freely choose where to

⁷The interpretation of this term is similar to the one provided in footnote 5 for the corresponding term of eq. (8). The only difference is that in this case the effects on the wage rate of earners in the big city and on the cost of provision of elderly care (always in the big city) are those brought about by a marginal reduction in the number of retired living in the big city (rather than in the number of earners as it was the case in eq. (8)). Also, notice that an increase in N_v^r has for the government's budget opposite effects as compared to those implied by an increase in N_v^e discussed in footnote 5. On one hand, an increase in N_v^r allows the government to increase the tax on earners in the big city without affecting their well-being; on the other hand, it also determines an increase in the cost of provision of elderly care at the city level.

live. It is therefore appropriate that we now consider the consequences of free mobility of individuals for the optimal policy chosen by the government. However, we still assume that the policy maker has instruments so he can perfectly control consumption and the quality of elderly care. Thus, the only choice left to people is the choice of the place where to live. To deal with this case we continue to treat N_v^r and N_v^e as government's control variables but add the migration equilibrium conditions as constraints in its optimization problem.⁸

The problem solved by the policy maker becomes the following:

$$\begin{aligned} & \max_{c_b^r, c_b^e, T_b^e, T_b^r, q_v, q_b, N_v^r, N_v^e} N_v^r [u(c_v^r) + \phi(q_v)] + (N^r - N_v^r) [u(c_b^r) + \phi(q_b)] + N_v^e u(\theta - T_v^e) + \\ & + (N^e - N_v^e) u \{F [N^e - N_v^e - q_b (N^r - N_v^r)] - T_b^e\} + \\ & + h \sum_{n=1}^{N_v^r} (N^r - n) + h \sum_{n=1+N_v^r}^{N^r} n + (N^e - N_v^e + N^r - N_v^r) m (N^e - N_v^e + N^r - N_v^r) \end{aligned}$$

subject to:

$$(N^e - N_v^e) T_b^e + N_v^e T_v^e - (N^r - N_v^r) \{c_b^r + q_b F [N^e - N_v^e - q_b (N^r - N_v^r)]\} - N_v^r (c_v^r + \theta q_v) \geq 0 \quad (\mu)$$

$$N_v^e - q_v N_v^r \geq 0 \quad (\gamma)$$

$$u(F [N^e - N_v^e - q_b (N^r - N_v^r)] - T_b^e) + m (N^e - N_v^e + N^r - N_v^r) - u(\theta - T_v^e) \geq 0 \quad (\lambda_1)$$

$$u(c_b^r) + \phi(q_b) + m (N^e - N_v^e + N^r - N_v^r) - u(c_v^r) - \phi(q_v) - 2h (N^r - N_v^r) + h N^r \geq 0 \quad (\lambda_2),$$

where Lagrange multipliers are within parentheses. The first order conditions for the above problem are presented in the Appendix.

Proposition 2 provides a characterization of the second best optimum.

Proposition 2 *Denoting by ε_F the elasticity of the marginal product of earners in the big city ($\varepsilon_F = (N_b^e - q_b N_b^r) F' / F$), the second best resource allocation is characterized by the following conditions:*

$$u'(c_v^r) = \frac{\mu N_v^r}{(N_v^r - \lambda_2)}, \quad (11)$$

⁸This is also the approach followed by Sato (2000) and Boadway et al. (2003).

$$u'(c_b^r) = \frac{\mu N_b^r}{(N_b^r + \lambda_2)} \quad (12)$$

$$u'(c_v^e) = \frac{\mu N_v^e}{(N_v^e - \lambda_1)} \quad (13)$$

$$u'(c_b^e) = \frac{\mu N_b^e}{(N_b^e + \lambda_1)} \quad (14)$$

$$\phi'(q_v) = \frac{(\mu\theta + \gamma)N_v^r}{(N_v^r - \lambda_2)}, \quad (15)$$

$$\phi'(q_b) = \frac{\mu N_b^r (F - q_b N_b^r F') + N_b^r (N_b^e + \lambda_1) u'(c_b^e) F'}{(N_b^r + \lambda_2)} \quad (16)$$

$$c_v^r + \theta q_v - c_b^r - q_b F = q_b \varepsilon_F F - \frac{\gamma q_v}{\mu} - \frac{(N_b + \lambda_1) m'}{\mu} - \frac{\lambda_2 (m' - 2h)}{\mu}, \quad (17)$$

$$T_b^e - T_v^e = -\varepsilon_F F + \frac{\gamma}{\mu} - \frac{(N_b + \lambda_1) m'}{\mu} - \frac{\lambda_2 m'}{\mu}. \quad (18)$$

Proof. See the Appendix. ■

Eqs.(11) -(14) show that marginal utility of consumption is no longer equal across individuals. The binding constraints λ_1 and λ_2 imply that marginal utility of consumption is higher in the village than in the big city. Accordingly, the consumption levels are lower in the village than in the big city. From (15) we see that marginal utility of elderly care is no longer equal to marginal cost. In the village marginal utility is equal to marginal cost scaled by the factor $N_v^r/(N_v^r - \lambda_2)$. From(16) we see that there is a similar scaling of the marginal cost in the big city, but in the equilibrium condition there is also a term showing how consumption of earners, and indirectly the constraint λ_1 , is affected as q_b is increased.

Eqs. (17) and (18) are the two conditions characterizing the optimal population distribution. Starting with (17), on the left hand side we have the difference between the value of resources absorbed by a retired person in the village and in the big city. It measures the direct resource cost for the government of having an additional retired living in the village.

On the right hand side the second term $(-\frac{\gamma q_v}{\mu})$ is smaller or equal to zero depending on whether the γ -constraint is binding or not. It measures the cost, normalized by μ , of the tightening of the γ -constraint that comes from an increase in the number of retired in the village. Both the first ($q_b \varepsilon_F F$) and the third term $(-\frac{(N_b + \lambda_1) m'}{\mu})$ on the right hand side take instead a positive sign. The former provides an evaluation of the benefit for the government's budget associated with the agglomeration effects coming from a marginal

reallocation of elderly from the big city to the village; the latter evaluates such a reallocation in terms of the positive welfare effects for those living in the big city descending from a marginal reduction in congestion and in terms of slackening the binding λ_1 -constraint.⁹ The sign of the last term on the right hand side of (17) is instead ambiguous. In fact, whereas $m' - 2h < 0$, it is not possible in general to sign unambiguously the Lagrange multiplier λ_2 .¹⁰ If $\lambda_2 > 0$ a marginal increase in the number of elderly living in the village would entail an additional benefit from the government's perspective; if $\lambda_2 < 0$ it would entail an additional cost.

As regards (18), on the left hand side there is the difference between the lump-sum tax levied on an earner in the big city and the lump-sum tax levied on an earner in a village. It measures the direct net revenue cost for the government of having an additional worker living in the village.

On the right hand side the first term is negative and it represents the cost for the government's budget, due to the working of the agglomeration effects, of marginally reducing the number of active persons in the big city. The second and third term are non-negative. The former (≥ 0) measures the benefits for the government of softening the eventually binding γ -constraint. The latter (> 0) evaluates the gains from reduced congestion costs in the big city: these are given both by the direct effect on the government's objective function ($-N_b m'$) and by the effect on the binding λ_1 -constraint ($-\lambda_1 m'$). The sign of the last term on the right hand side of (18) is once again ambiguous. A positive value for λ_2 would imply an additional benefit associated to the marginal increase in the number of earners living in the village, whereas a negative value for λ_2 would imply an additional cost.

Finally, before turning our attention in Section 5 to the study of the problems posed by a system of hierarchical policy makers, Proposition 3 provides the value of the optimal transfer to the village (denoted by G_v) implicit in the solution to the government's problem and offers an interpretation of the marginal cost of public funds.

Proposition 3 *At a second best optimum the implicit value of the transfer to the village is:*

$$G_v = \frac{N_v^r}{N^r} [(N^e - N_v^e) T_b^e + N_v^e T_v^e] - \frac{N_v^r (N^r - N_v^r)}{N^r} \left[\frac{N_b m' + \gamma q_v + \lambda_1 m' + \lambda_2 (m' - 2h)}{\mu} - q_b F \varepsilon_F \right]. \quad (19)$$

⁹It is in fact possible to show that $\lambda_1 = \frac{N_b^e N_v^e [u'(c_v^e) - u'(c_b^e)]}{N_v^e u'(c_b^e) + N_b^e u'(c_v^e)} > 0$.

¹⁰From the first order conditions of the government's problem it can be shown that $\lambda_2 = \frac{N_b^r N_v^r [u'(c_v^r) - u'(c_b^r)]}{N_v^r u'(c_b^r) + N_b^r u'(c_v^r)} (> < 0)$. If at an optimum $\lambda_2 < (>) 0$, then the amount of pension granted to a retiree in the country side is larger than the amount of pension granted to a retiree in the big city.

Moreover,

$$\frac{1}{\mu} = \frac{N_v^r}{N^r u'(c_v^r)} + \frac{N^r - N_v^r}{N^r u'(c_b^r)} = \frac{N_v^e}{N^e u'(c_v^e)} + \frac{N^e - N_v^e}{N^e u'(c_b^e)}.$$

Proof. See the Appendix. ■

Denoting N_v^r/N^r and N_b^r/N^r by respectively n_v^r and n_b^r , from (19) we also get the result that the equalization entitlement of the country side E_v (defined as $E_v \equiv G_v - N_v^e T_v^e$) is given by:

$$E_v = n_v^r n_b^r \left[\frac{N_b^e T_b^e}{n_b^r} - \frac{N_v^e T_v^e}{n_v^r} \right] - \frac{N_v^r N_b^r}{N^r} \left[\frac{N_b m' + \gamma q_v + \lambda_1 m' + \lambda_2 (m' - 2h)}{\mu} - q_b F \varepsilon_F \right].$$

Having completed the analysis of the second best case with full centralization, we are now ready to consider the consequences of having a multi-level government structure.

5 The federal case

In this Section we abandon the assumption of full centralization and assume the existence of local policy makers that are responsible for the provision of elderly care and, at least partly, for the funding of it. Thus, we assume that each local government chooses the local tax rate τ on earners and the quality of elderly care q subject to local budget balance and taking into account individuals' migration responses. The budget balance of other governments is however ignored: from a local perspective the tax rate and the public expenditures of other governments are taken as given. Pensions are still directly decided upon (and paid) by the central government. The central government is assumed to be a Stackelberg leader and it incorporates the effects of its decision making on the Nash equilibrium. We start the analysis with the problem faced by the policy maker in the big city.

5.1 The problem solved by the policy maker in the big city

Denoting by G_b the value of a lump-sum transfer to the big city, the problem solved by the policy maker there can be written as:

$$\begin{aligned} & \max_{\tau_b^e, q_b, N_v^r, N_v^e} (N^r - N_v^r) [u(c_b^r) + \phi(q_b)] + \\ & + (N^e - N_v^e) u \{ F [N^e - N_v^e - q_b (N^r - N_v^r)] - \tau_b^e \} + \\ & + h \sum_{n=1+N_v^r}^{N^r} n + (N^e - N_v^e + N^r - N_v^r) m (N^e - N_v^e + N^r - N_v^r) \end{aligned}$$

subject to:

$$(N^e - N_v^e) \tau_b^e + G_b \geq q_b (N^r - N_v^r) F [N^e - N_v^e - q_b (N^r - N_v^r)] \quad (\mu_b)$$

$$u(F [N^e - N_v^e - q_b (N^r - N_v^r)] - \tau_b^e) + m(N^e - N_v^e + N^r - N_v^r) - u(\theta - \tau_b^e) \geq 0 \quad (\lambda_b^e)$$

$$u(c_b^r) + \phi(q_b) + m(N^e - N_v^e + N^r - N_v^r) - u(c_b^r) - \phi(q_b) - 2h(N^r - N_v^r) + hN^r \geq 0, \quad (\lambda_b^r)$$

where Lagrange multipliers are within parentheses. The associated first order conditions are:

$$\begin{aligned} & (N^r - N_v^r) \phi'(q_b) - (N^e - N_v^e) (N^r - N_v^r) u'(c_b^e) F' - \mu_b (N^r - N_v^r) [F - q_b (N^r - N_v^r) F'] \\ & - \lambda_b^e (N^r - N_v^r) u'(c_b^e) F' + \lambda_b^r \phi'(q_b) \\ = & 0 \end{aligned} \quad (q_b) \quad (20)$$

$$(N^e - N_v^e + \lambda_b^e) u'(c_b^e) = \mu_b (N^e - N_v^e) \quad (\tau_b^e) \quad (21)$$

$$\begin{aligned} & -[u(c_b^r) + \phi(q_b)] + (N^e - N_v^e) u'(c_b^e) q_b F' - hN_v^r \\ & - N_b m' - m + \mu_b [q_b F - (N^r - N_v^r) (q_b)^2 F'] + \lambda_b^e [u'(c_b^e) q_b F' - m'] + \lambda_b^r (2h - m') \\ = & 0 \end{aligned} \quad (N_v^r) \quad (22)$$

$$\begin{aligned} & -u(F - \tau_b^e) - (N^e - N_v^e) u'(c_b^e) F' - N_b m' - m + \\ & + \mu_b [-\tau_b^e + (N^r - N_v^r) q_b F'] - \lambda_b^e [u'(c_b^e) F' + m'] - \lambda_b^r m' \\ = & 0 \end{aligned} \quad (N_v^e) \quad (23)$$

The necessary conditions for a local optimum are:

$$\frac{(N^r - N_v^r + \lambda_b^r) \phi'(q_b)}{(N^e - N_v^e + \lambda_b^e) u'(c_b^e)} = \frac{N^r - N_v^r}{N^e - N_v^e} (1 + \varepsilon_F) F, \quad (24)$$

$$q_b F = \frac{(N_b + \lambda_b^e + \lambda_b^r) m' - 2\lambda_b^r h + u(c_b^r) + \phi(q_b) + hN_v^r + m}{\mu_b} - q_b \varepsilon_F F, \quad (25)$$

$$-\tau_b^e = \frac{(N_b + \lambda_b^e + \lambda_b^r) m' + u(F - \tau_b^e) + m}{\mu_b} + \varepsilon_F F, \quad (26)$$

where (25) and (26) are the conditions representing how individuals of different age-group should be allocated from the perspective of the policy maker in the big city.

5.2 The problem solved by the policy maker in the village

Denoting by G_v the value of a lump-sum transfer to the village, the problem solved by the policy maker there can be written as:

$$\max_{\tau_v^e, q_v, N_v^r, N_v^e} N_v^r [u(c_v^r) + \phi(q_v)] + N_v^e u(\theta - \tau_v^e) + h \sum_{n=1}^{N_v^r} (N^r - n)$$

subject to:

$$N_v^e \tau_v^e + G_v \geq N_v^r \theta q_v \quad (\mu_v)$$

$$N_v^e - q_v N_v^r \geq 0 \quad (\gamma_v)$$

$$u(F [N^e - N_v^e - q_b (N^r - N_v^r)] - \tau_b^e) + m (N^e - N_v^e + N^r - N_v^r) - u(\theta - \tau_v^e) \geq 0 \quad (\lambda_v^e)$$

$$u(c_b^r) + \phi(q_b) + m (N^e - N_v^e + N^r - N_v^r) - u(c_v^r) - \phi(q_v) - 2h (N^r - N_v^r) + h N^r \geq 0. \quad (\lambda_v^r)$$

As usual, Lagrange multipliers are within parentheses. The associated first order conditions are:

$$(N_v^r - \lambda_v^r) \phi'(q_v) - \mu_v \theta N_v^r - \gamma_v N_v^r = 0 \quad (q_v) \quad (27)$$

$$(N_v^e - \lambda_v^e) u'(c_v^e) = \mu_v N_v^e \quad (\tau_v^e) \quad (28)$$

$$u(c_v^r) + \phi(q_v) + h (N^r - N_v^r) - \mu_v \theta q_v - \gamma_v q_v + \lambda_v^e [u'(c_b^e) q_b F' - m'] + \lambda_v^r (2h - m') = 0 \quad (N_v^r) \quad (29)$$

$$u(\theta - \tau_v^e) + \mu_v \tau_v^e + \gamma_v - \lambda_v^e [u'(c_b^e) F' + m'] - \lambda_v^r m' = 0 \quad (N_v^e) \quad (30)$$

The necessary conditions for a local optimum are:

$$\frac{(N_v^r - \lambda_v^r) \phi'(q_v)}{(N_v^e - \lambda_v^e) u'(c_v^e)} = \frac{\mu_v \theta N_v^r + \gamma_v N_v^r}{\mu_v N_v^e}, \quad (31)$$

$$-\theta q_v = \frac{\gamma_v q_v - \lambda_v^e [u'(c_b^e) q_b F' - m'] - \lambda_v^r (2h - m') - u(c_v^r) - \phi(q_v) - h(N^r - N_v^r)}{\mu_v}, \quad (32)$$

$$\tau_v^e = \frac{\lambda_v^e [u'(c_b^e) F' + m'] + \lambda_v^r m' - u(\theta - \tau_v^e) - \gamma_v}{\mu_v}, \quad (33)$$

where (32) and (33) are the conditions representing how individuals of different age-group should be allocated from the perspective of the policy maker in the village.

5.3 The federal government's problem

From (25), (26), (32) and (33), it is apparent that without proper incentives the population purchase choice of the local policy makers would be suboptimal and that lump-sum transfers alone are not sufficient. In order to be able to implement in a decentralized solution the population distribution characterized in the previous Section, the federal government needs to be empowered with policy instruments that independently work on each of the two migration equilibrium constraints to the local policy makers' optimization problems. The reason for this is that such a set of instruments allows the federal government to affect the value of the shadow prices λ_v^r , λ_v^e , λ_b^r and λ_b^e which in turn determine the way in which the local policy makers would like to see agents allocated between the big city and the village. Thus, we will assume that the federal government can also transfer/collect resources from the local governments by means of a set of differentiated per-earner and per-elderly subsidies/taxes. In particular, s_v^r and s_b^r will respectively denote the per-elderly subsidy (or tax, if negative) in the village and the big city, whereas s_v^e and s_b^e will denote the per-earner subsidies (or taxes, if negative). In the remainder of the Section we will assume that the local policy makers are the only authorities allowed to impose taxes on earners.¹¹ Under this assumption, the problem solved by the federal government is:

¹¹In the context of our model it can be shown that differentiated lump-sum taxes on earners are irrelevant for the purpose of implementation of the second best in a decentralized framework.

$$\begin{aligned} & \max_{c_v^r, c_b^r, G_v, G_b, s_v^r, s_b^r, s_v^e, s_b^e} \\ & V_b(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e, \tau_v^e(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e), q_v(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e)) \\ & + V_v(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e, \tau_b^e(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e), q_b(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e)) \end{aligned}$$

subject to:

$$\begin{aligned} & -G_v - G_b - s_v^r N_v^r - s_v^e N_v^e - s_b^r (N^r - N_v^r) \\ & -s_b^e (N^e - N_v^e) - (N^r - N_v^r) c_b^r - N_v^r c_v^r \geq 0 \quad (\mu_c), \end{aligned}$$

where

$$\begin{aligned} & V_b(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e, \tau_v^e(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e), q_v(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e)) \\ & \equiv \\ & (N^r - N_v^r) [u(c_b^r) + \phi(q_b)] + (N^e - N_v^e) u \{F [N^e - N_v^e - q_b (N^r - N_v^r)] - \tau_b^e\} \\ & + h \sum_{n=1+N_v^r}^{N^r} n + (N^e - N_v^e + N^r - N_v^r) m (N^e - N_v^e + N^r - N_v^r), \end{aligned}$$

and

$$\begin{aligned} & V_v(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e, \tau_b^e(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e), q_b(c_v^r, c_b^r, G_v, G_b, s_v^r, s_v^e, s_b^r, s_b^e)) \\ & = \\ & N_v^r [u(c_v^r) + \phi(q_v)] + N_v^e u(\theta - \tau_v^e) + h \sum_{n=1}^{N_v^r} (N^r - n). \end{aligned}$$

Proposition 4 provides the necessary conditions that characterize the optimal values of the subsidies/taxes s .

Proposition 4 *The optimal values for s_b^r , s_v^r , s_b^e and s_v^e satisfy respectively the following conditions:*

$$\begin{aligned} \mu_c (N^r - N_v^r) & = \mu_b (N^r - N_v^r) + \frac{\lambda_v^r \phi'(q_b)}{F - q_b (N^r - N_v^r) F'} + \left[\lambda_b^e u'(c_v^e) - \lambda_b^r \phi'(q_v) \frac{N_v^e}{N_v^r \theta} \right] \frac{d\tau_v^e}{ds_b^r} \\ & + \left[\frac{\lambda_v^r \phi'(q_b) (N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} - \lambda_v^e u'(c_b^e) \right] \frac{d\tau_b^e}{ds_b^r} \\ & + H_1 \frac{dN_v^e}{ds_b^r} + H_2 \frac{dN_v^r}{ds_b^r}, \end{aligned} \quad (34)$$

$$\begin{aligned}
\mu_c N_v^r &= \mu_v N_v^r - \frac{\lambda_b^r \phi'(q_v)}{\theta} + \left[\lambda_b^e u'(c_v^e) - \lambda_b^r \phi'(q_v) \frac{N_v^e}{N_v^r \theta} \right] \frac{d\tau_v^e}{ds_v^r} \\
&+ \left[\frac{\lambda_b^r \phi'(q_b) (N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} - \lambda_b^e u'(c_b^e) \right] \frac{d\tau_b^e}{ds_v^r} \quad (35) \\
&+ H_1 \frac{dN_v^e}{ds_v^r} + H_2 \frac{dN_v^r}{ds_v^r},
\end{aligned}$$

$$\begin{aligned}
\mu_c (N^e - N_v^e) &= \mu_b (N^e - N_v^e) + \frac{\lambda_b^r \phi'(q_b) (N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \\
&+ \left[\lambda_b^e u'(c_v^e) - \lambda_b^r \phi'(q_v) \frac{N_v^e}{N_v^r \theta} \right] \frac{d\tau_v^e}{ds_b^e} \\
&+ \left[\frac{\lambda_b^r \phi'(q_b) (N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} - \lambda_b^e u'(c_b^e) \right] \frac{d\tau_b^e}{ds_b^e} \\
&+ H_1 \frac{dN_v^e}{ds_b^e} + H_2 \frac{dN_v^r}{ds_b^e}, \quad (36)
\end{aligned}$$

$$\begin{aligned}
\mu_c N_v^e &= \mu_v N_v^e - \frac{\lambda_b^r \phi'(q_v) N_v^e}{N_v^r \theta} + \left[\lambda_b^e u'(c_v^e) - \frac{\lambda_b^r \phi'(q_v) N_v^e}{N_v^r \theta} \right] \frac{d\tau_v^e}{ds_v^e} \\
&+ \left[\frac{\lambda_b^r \phi'(q_b) (N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} - \lambda_b^e u'(c_b^e) \right] \frac{d\tau_b^e}{ds_v^e} \quad (37) \\
&+ H_1 \frac{dN_v^e}{ds_v^e} + H_2 \frac{dN_v^r}{ds_v^e},
\end{aligned}$$

where H_1 and H_2 have been defined as:

$$\begin{aligned}
H_1 &= -\lambda_b^r \phi'(q_v) \frac{\tau_v^e + s_v^e}{N_v^r \theta} - \lambda_b^r \phi'(q_b) \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} - \mu_c (s_v^e - s_b^e), \\
H_2 &= \lambda_b^r \phi'(q_v) \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} - \lambda_b^r \phi'(q_b) \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \\
&\quad - \mu_c (s_v^r + c_v^r - s_b^r - c_b^r)
\end{aligned}$$

Proof. See the Appendix. ■

The first thing that emerges from (34)-(37) is that the values of the Lagrange multipliers associated with the budget constraints of different level of governments are not equalized.¹²

¹²It is straightforward to show that such an equalization was instead implicit in the solution to the problem considered in Section 4.

Conditions (34)-(37) admit all a similar interpretation. To save space, we will here only discuss eq. (34). A marginal increase in the subsidy s_b^r (or a marginal decrease in the tax, if $s_b^r < 0$) is worth $\mu_b (N^r - N_v^r)$ from the perspective of the policy maker in the big city; this additional resources granted to the big city allows the local policy maker to increase the quality of elderly care q_b and this increase, evaluated from the point of view of the policy maker in the village, is worth $\frac{\lambda_v^r \phi'(q_b)}{F - q_b (N^r - N_v^r) F'}$ given that it affects the tightness of the λ_v^r -migration-constraint. On the other hand, a marginal increase in s_b^r has not only a direct effect on the expenditure possibilities of the policy maker in the big city, but it has also an indirect effect working through the change in the optimal choice of the tax rate on earners τ_v^e . The effect of this change spills over to the other local government, affecting both the migration constraints λ_v^r and λ_v^e : this is captured by the second line of (34). However, given the structure of the game played between the federal and the local governments, a marginal increase in s_b^r also induces the policy maker in the village to revise its optimal choice of τ_v^e and this in turn has an effect on the policy maker in the big city, affecting the strength of the migration equilibrium constraints λ_b^e and λ_b^r . This is what is captured by the term in square brackets in the first line of (34). Finally, a marginal change in s_b^r also produces inter-community migration, which alters the tax base for both the local policy makers and the federal government. For each of the two local policy makers this means that there are two additional effects on the migration equilibrium constraint for retired people; for the federal government this implies an additional effect on the budget constraint.

6 Concluding remarks

To be written

7 Appendix

7.1 Proof of Proposition 1

The first order conditions for the first best are

$$u'(c_v^r) = u'(c_b^r) = u'(c_v^e) = u'(c_b^e) = \mu$$

$$\phi'(q_v) = \mu\theta + \gamma \quad (q_v)$$

$$\phi'(q_b) = (N^e - N_v^e) u'(c_b^e) F' + \mu [F - q_b (N^r - N_v^r) F'] \quad (q_b)$$

$$\begin{aligned}
& u(c_v^r) + \phi(q_v) - u(c_b^r) - \phi(q_b) + h(N^r - 2N_v^r) - (N_b m' + m) + (N^e - N_v^e) q_b u'(c_b^e) F' \\
= & \gamma q_v + \mu \left[\theta q_v + c_v^r - c_b^r - q_b F + (q_b)^2 N_b^r F' \right] \quad (N_v^r)
\end{aligned}$$

$$u(c_v^e) - u(c_b^e) - (N^e - N_v^e) u'(c_b^e) F' - (N_b m' + m) = -\gamma + \mu (-q_b N_b^r F' - T_v^e + T_b^e). \quad (N_v^e)$$

To get the result stated in Proposition 1 we have also exploited the identity $-T_v^e + T_b^e = F - \theta + c_v^e - c_b^e$, plus the fact that $u'(c_b^e) = \mu$ from the f.o.c. w.r.t. T_b^e .

7.2 Proof of Proposition 2

The first order conditions are:

$$(N_v^r - \lambda_2) u'(c_v^r) = \mu N_v^r \quad (c_v^r) \quad (38)$$

$$[(N^r - N_v^r) + \lambda_2] u'(c_b^r) = \mu (N^r - N_v^r) \quad (c_b^r) \quad (39)$$

$$(N_v^r - \lambda_2) \phi'(q_v) = (\mu\theta + \gamma) N_v^r \quad (q_v) \quad (40)$$

$$\begin{aligned}
[(N^r - N_v^r) + \lambda_2] \phi'(q_b) &= (N^r - N_v^r) [(N^e - N_v^e) + \lambda_1] u'(c_b^e) F' & (41) \\
&+ \mu (N^r - N_v^r) [F - q_b (N^r - N_v^r) F'] & (q_b)
\end{aligned}$$

$$(N_v^e - \lambda_1) u'(c_v^e) = \mu N_v^e \quad (T_v^e) \quad (42)$$

$$[(N^e - N_v^e) + \lambda_1] u'(c_b^e) = \mu (N^e - N_v^e) \quad (T_b^e) \quad (43)$$

To get (11), divide (38) by (39). For (15), first use (43) to rewrite (41) as:

$$[(N^r - N_v^r) + \lambda_2] \phi'(q_b) = \mu (N^r - N_v^r) (1 + \varepsilon_F) F. \quad (44)$$

Then divide (40) by (44) to get

$$\frac{(N_v^r - \lambda_2) \phi'(q_v)}{[(N^r - N_v^r) + \lambda_2] \phi'(q_b)} = \frac{\mu\theta + \gamma}{\mu(1 + \varepsilon_F) F} \frac{N_v^r}{N^r - N_v^r}. \quad (45)$$

To get (15) multiply both sides of (45) by $[(N^r - N_v^r) + \lambda_2] / (N_v^r - \lambda_2)$.

The first order conditions with respect to N_v^r and N_v^e are respectively given by:

$$\begin{aligned}
& u(c_v^r) + \phi(q_v) - [u(c_b^r) + \phi(q_b)] - m + h(N^r - N_v^r) - hN_v^r \\
& + (N^e - N_v^e) u'(c_b^e) q_b F' - N_b m' - \gamma q_v - \mu(N^r - N_v^r) (q_b)^2 F' \\
& + \lambda_1 [u'(c_b^e) q_b F' - m'] - \lambda_2 (m' - 2h) \\
= & -\mu(c_b^r + q_b F - c_v^r - \theta q_v) \quad (N_v^r) \quad (46)
\end{aligned}$$

$$\begin{aligned}
& u(\theta - T_v^e) - u\{F[N^e - N_v^e - q_b(N^r - N_v^r)] - T_b^e\} - m \\
& - (N^e - N_v^e) u'(c_b^e) F' - N_b m' + \gamma + \mu\{T_v^e - T_b^e + (N^r - N_v^r) q_b F'\} \\
& - \lambda_1 [u'(c_b^e) F' + m'] - \lambda_2 m' \\
= & 0 \quad (N_v^e) \quad (47)
\end{aligned}$$

Using the equilibrium migration condition $u(c_b^r) + \phi(q_b) + m(N^e - N_v^e + N^r - N_v^r) = u(c_v^r) + \phi(q_v) + 2h(N^r - N_v^r) - hN^r$, eq. (46) can be simplified to:

$$\begin{aligned}
& (N^e - N_v^e) u'(c_b^e) q_b F' - \mu(N^r - N_v^r) (q_b)^2 F' - N_b m' - \gamma q_v \\
& + \lambda_1 [u'(c_b^e) q_b F' - m'] - \lambda_2 (m' - 2h) \\
= & \mu(c_v^r + \theta q_v - c_b^r - q_b F).
\end{aligned}$$

Dividing by μ and using (43) gives (17).

Using the equilibrium migration condition $u(F[N^e - N_v^e - q_b(N^r - N_v^r)] - T_b^e) + m(N^e - N_v^e + N^r - N_v^r) = u(\theta - T_v^e)$, eq. (47) can be simplified to:

$$-(N^e - N_v^e) u'(c_b^e) F' - N_b m' + \gamma + \mu(N^r - N_v^r) q_b F' - \lambda_1 [u'(c_b^e) F' + m'] - \lambda_2 m' = \mu(T_b^e - T_v^e).$$

Dividing by μ and using (43) gives (18).

7.3 Proof of Proposition 3

The problem solved by the government can equivalently be stated as:

$$\begin{aligned}
& \max_{c_b^r, c_b^e, T_v^e, T_b^e, q_v, q_b, N_v^r, N_v^e, G_v, G_b} N_v^r [u(c_v^r) + \phi(q_v)] + (N^r - N_v^r) [u(c_b^r) + \phi(q_b)] + N_v^e u(\theta - T_v^e) + \\
& + (N^e - N_v^e) u\{F[N^e - N_v^e - q_b(N^r - N_v^r)] - T_b^e\} + \\
& + h \sum_{n=1}^{N_v^r} (N^r - n) + h \sum_{n=1+N_v^r}^{N^r} n + (N^e - N_v^e + N^r - N_v^r) m(N^e - N_v^e + N^r - N_v^r) \\
& (N^e - N_v^e) T_b^e + N_v^e T_v^e - G_v - G_b \geq 0 \quad (\mu) \quad (48)
\end{aligned}$$

$$G_b - (N^r - N_v^r) \{c_b^r + q_b F [N^e - N_v^e - q_b (N^r - N_v^r)]\} \geq 0 \quad (\mu_b) \quad (49)$$

$$G_v - N_v^r (c_v^r + \theta q_v) \geq 0 \quad (\mu_v) \quad (50)$$

$$N_v^e - q_v N_v^r \geq 0 \quad (\gamma)$$

$$u(F[N^e - N_v^e - q_b(N^r - N_v^r)] - T_b^e) + m(N^e - N_v^e + N^r - N_v^r) - u(\theta - T_v^e) \geq 0 \quad (\lambda_1)$$

$$u(c_b^r) + \phi(q_b) + m(N^e - N_v^e + N^r - N_v^r) - u(c_v^r) - \phi(q_v) - 2h(N^r - N_v^r) + hN^r \geq 0, \quad (\lambda_2)$$

where G_v and G_b are the transfers to the village and the big city respectively.

From (49) and (50) we have

$$\frac{G_b}{N^r - N_v^r} = c_b^r + q_b F [N^e - N_v^e - q_b (N^r - N_v^r)], \quad (51)$$

$$\frac{G_v}{N_v^r} = c_v^r + \theta q_v. \quad (52)$$

Subtracting (51) from (52) gives:

$$\frac{G_v}{N_v^r} - \frac{G_b}{N^r - N_v^r} = c_v^r + \theta q_v - c_b^r - q_b F [N^e - N_v^e - q_b (N^r - N_v^r)]. \quad (53)$$

Using (17), (53) can be written as:

$$\frac{G_v}{N_v^r} - \frac{G_b}{N^r - N_v^r} = -\frac{N_b m' + \gamma q_v + \lambda_1 m' + \lambda_2 (m' - 2h)}{\mu} + q_b F_{\varepsilon F}. \quad (54)$$

From (48) we have:

$$(N^e - N_v^e) T_b^e + N_v^e T_v^e - G_v = G_b. \quad (55)$$

Substituting (55) in (54) gives:

$$\frac{G_v}{N_v^r} + \frac{G_v}{N^r - N_v^r} = \frac{(N^e - N_v^e) T_b^e + N_v^e T_v^e}{N^r - N_v^r} - \frac{N_b m' + \gamma q_v + \lambda_1 m' + \lambda_2 (m' - 2h)}{\mu} + q_b F_{\varepsilon F}. \quad (56)$$

Finally, collecting terms on the left hand side of (56) gives the result stated in the first part of Proposition 3.

Combining the first order conditions on c_v^r and c_b^r , and those on T_v^e and T_b^e yields the interpretation of the marginal cost of public funds provided by the second part of Proposition 3.

7.4 Proof of Proposition 4

With subsidies (taxes) related to the number of earners and retirees, the budget constraints of the local policy makers in the big city and in the country side become respectively:

$$\frac{(N^e - N_v^e)(\tau_b^e + s_b^e) + G_b + s_b^r(N^r - N_v^r)}{(N^r - N_v^r)F[N^e - N_v^e - q_b(N^r - N_v^r)]} \geq q_b,$$

$$\frac{N_v^e(\tau_v^e + s_v^e) + G_v + s_v^r N_v^r}{N_v^r \theta} \geq q_v.$$

Thus, the effects of the per person subsidies (taxes) s on the local choices of elderly care quality are:

$$\frac{dq_b}{ds_b^e} = \frac{(N^e - N_v^e)}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \left(1 + \frac{d\tau_b^e}{ds_b^e}\right) - \frac{(\tau_b^e + s_b^e) - q_b(N^r - N_v^r)F'}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^e}{ds_b^e} - \frac{s_b^r - q_b[F - (N^r - N_v^r)q_b F']}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^r}{ds_b^e};$$

$$\frac{dq_b}{ds_b^r} = \frac{1}{F - q_b(N^r - N_v^r)F'} + \frac{(N^e - N_v^e)}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{d\tau_b^e}{ds_b^r} - \frac{(\tau_b^e + s_b^e) - q_b(N^r - N_v^r)F'}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^e}{ds_b^r} - \frac{s_b^r - q_b[F - (N^r - N_v^r)q_b F']}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^r}{ds_b^r};$$

$$\frac{dq_b}{ds_v^e} = \frac{(N^e - N_v^e)}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{d\tau_b^e}{ds_v^e} - \frac{(\tau_b^e + s_b^e) - q_b(N^r - N_v^r)F'}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^e}{ds_v^e} - \frac{s_b^r - q_b[F - (N^r - N_v^r)q_b F']}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^r}{ds_v^e};$$

$$\frac{dq_b}{ds_v^r} = \frac{(N^e - N_v^e)}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{d\tau_b^e}{ds_v^r} - \frac{(\tau_b^e + s_b^e) - q_b(N^r - N_v^r)F'}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^e}{ds_v^r} - \frac{s_b^r - q_b[F - (N^r - N_v^r)q_b F']}{(N^r - N_v^r)[F - q_b(N^r - N_v^r)F']} \frac{dN_v^r}{ds_v^r};$$

$$\frac{dq_v}{ds_v^e} = \frac{N_v^e}{N_v^r \theta} + \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_v^e} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_v^e} - \frac{N_v^e(\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_v^e};$$

$$\frac{dq_v}{ds_v^r} = \frac{1}{\theta} + \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_v^r} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_v^r} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_v^r},$$

$$\frac{dq_v}{ds_b^e} = \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_b^e} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_b^e} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_b^e},$$

$$\frac{dq_v}{ds_b^r} = \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_b^r} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_b^r} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_b^r}.$$

Denoting by L_v and L_b the Lagrangians of the local policy makers' problems, it is:

$$\begin{aligned} L_v = & N_v^r [u(c_v^r) + \phi(q_v)] + N_v^e u(\theta - \tau_v^e) + h \sum_{n=1}^{N_v^r} (N^r - n) \\ & + \mu_v \left[\frac{N_v^e (\tau_v^e + s_v^e) + G_v + s_v^r N_v^r}{N_v^r \theta} - q_v \right] + \gamma_v (N_v^e - q_v N_v^r) \\ & + \lambda_b^e \{u(F[N^e - N_v^e - q_b(N^r - N_v^r)] - \tau_b^e) + m(N^e - N_v^e + N^r - N_v^r) - u(\theta - \tau_v^e)\} \\ & + \lambda_b^r \{u(c_b^r) + \phi(q_b) + m(N^e - N_v^e + N^r - N_v^r) - u(c_v^r) - \phi(q_v) - 2h(N^r - N_v^r) + hN^r\} \end{aligned}$$

and

$$\begin{aligned} L_b = & (N^r - N_v^r) [u(c_b^r) + \phi(q_b)] + (N^e - N_v^e) u\{F[N^e - N_v^e - q_b(N^r - N_v^r)] - \tau_b^e\} \\ & + h \sum_{n=1+N_v^r}^{N^r} n + (N^e - N_v^e + N^r - N_v^r) m(N^e - N_v^e + N^r - N_v^r) \\ & + \mu_b \{(N^e - N_v^e) (\tau_b^e + s_b^e) + G_b + s_b^r (N^r - N_v^r) - q_b (N^r - N_v^r) F[N^e - N_v^e - q_b(N^r - N_v^r)]\} \\ & + \lambda_b^e \{u(F[N^e - N_v^e - q_b(N^r - N_v^r)] - \tau_b^e) + m(N^e - N_v^e + N^r - N_v^r) - u(\theta - \tau_v^e)\} \\ & + \lambda_b^r \{u(c_b^r) + \phi(q_b) + m(N^e - N_v^e + N^r - N_v^r) - u(c_v^r) - \phi(q_v) - 2h(N^r - N_v^r) + hN^r\}. \end{aligned}$$

Thus, we have that:

$$\begin{aligned} \frac{dV_b}{ds_b^e} &= \frac{\partial L_b}{\partial s_b^e} + \frac{\partial L_b}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_b^e} + \frac{\partial L_b}{\partial q_b} \frac{dq_b}{ds_b^e} \\ \frac{dV_v}{ds_b^e} &= \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_b^e} + \frac{\partial L_v}{\partial q_b} \frac{dq_b}{ds_b^e} \end{aligned}$$

$$\begin{aligned} \frac{dV_b}{ds_b^r} &= \frac{\partial L_b}{\partial s_b^r} + \frac{\partial L_b}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_b^r} + \frac{\partial L_b}{\partial q_b} \frac{dq_b}{ds_b^r} \\ \frac{dV_v}{ds_b^r} &= \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_b^r} + \frac{\partial L_v}{\partial q_b} \frac{dq_b}{ds_b^r} \end{aligned}$$

$$\begin{aligned}\frac{dV_b}{ds_v^e} &= \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_v^e} + \frac{\partial L_b}{\partial q_v} \frac{dq_v}{ds_v^e} \\ \frac{dV_v}{ds_v^e} &= \frac{\partial L_v}{\partial s_v^e} + \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_v^e} + \frac{\partial L_v}{\partial q_b} \frac{dq_b}{ds_v^e}\end{aligned}$$

$$\begin{aligned}\frac{dV_b}{ds_v^r} &= \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_v^r} + \frac{\partial L_b}{\partial q_v} \frac{dq_v}{ds_v^r} \\ \frac{dV_v}{ds_v^r} &= \frac{\partial L_v}{\partial s_v^r} + \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_v^r} + \frac{\partial L_v}{\partial q_b} \frac{dq_b}{ds_v^r}.\end{aligned}$$

Eqs. (34)-(37) are straightforwardly obtained taking into account that:

$$\begin{aligned}\frac{dV_b}{ds_b^r} &= \frac{\partial L_b}{\partial s_b^r} + \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_b^r} + \frac{\partial L_b}{\partial q_v} \left[\frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_b^r} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_b^r} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_b^r} \right] \\ &= \mu_b (N^r - N_v^r) + \lambda_b^e u' (c_v^e) \frac{d\tau_v^e}{ds_b^r} - \lambda_b^r \phi' (q_v) \left[\frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_b^r} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_b^r} \right] \\ &\quad + \lambda_b^r \phi' (q_v) \left[\frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_b^r} \right]\end{aligned}$$

$$\begin{aligned}\frac{dV_v}{ds_b^r} &= \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_b^r} + \frac{\partial L_v}{\partial q_b} \left[\frac{1}{F - q_b (N^r - N_v^r) F'} + \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{d\tau_b^e}{ds_b^r} \right] \\ &\quad - \frac{\partial L_v}{\partial q_b} \left[\frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_b^r} + \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_b^r} \right] \\ &= -\lambda_v^e u' (c_b^e) \frac{d\tau_b^e}{ds_b^r} + \lambda_v^r \phi' (q_b) \frac{1}{F - q_b (N^r - N_v^r) F'} \\ &\quad + \lambda_v^r \phi' (q_b) \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{d\tau_b^e}{ds_b^r} \\ &\quad - \lambda_v^r \phi' (q_b) \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_b^r} \\ &\quad - \lambda_v^r \phi' (q_b) \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_b^r}\end{aligned}$$

$$\begin{aligned}\frac{dV_b}{ds_v^r} &= \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_v^r} + \frac{\partial L_b}{\partial q_v} \frac{dq_v}{ds_v^r} \\ &= \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_v^r} + \frac{\partial L_b}{\partial q_v} \left[\frac{1}{\theta} + \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_v^r} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_v^r} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_v^r} \right] \\ &= \lambda_b^e u' (c_v^e) \frac{d\tau_v^e}{ds_v^r} - \lambda_b^r \phi' (q_v) \left[\frac{1}{\theta} + \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_v^r} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_v^r} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_v^r} \right]\end{aligned}$$

$$\begin{aligned}
\frac{dV_v}{ds_v^r} &= \frac{\partial L_v}{\partial s_v^r} + \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_v^r} + \frac{\partial L_v}{\partial q_b} \frac{dq_b}{ds_v^r} \\
&= \frac{\partial L_v}{\partial s_v^r} + \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_v^r} + \frac{\partial L_v}{\partial q_b} \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{d\tau_b^e}{ds_v^r} \\
&\quad - \frac{\partial L_v}{\partial q_b} \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_v^r} \\
&\quad - \frac{\partial L_v}{\partial q_b} \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_v^r} \\
&= \mu_v N_v^r - \lambda_v^e u'(c_b^e) \frac{d\tau_b^e}{ds_v^r} + \lambda_v^r \phi'(q_b) \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{d\tau_b^e}{ds_v^r} \\
&\quad - \lambda_v^r \phi'(q_b) \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_v^r} \\
&\quad - \lambda_v^r \phi'(q_b) \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_v^r}
\end{aligned}$$

$$\begin{aligned}
\frac{dV_b}{ds_b^e} &= \frac{\partial L_b}{\partial s_b^e} + \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_b^e} + \frac{\partial L_b}{\partial q_v} \frac{dq_v}{ds_b^e} \\
&= \frac{\partial L_b}{\partial s_b^e} + \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_b^e} + \frac{\partial L_b}{\partial q_v} \left[\frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_b^e} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_b^e} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_b^e} \right] \\
&= \mu_b (N^e - N_v^e) + \lambda_b^e u'(c_v^e) \frac{d\tau_v^e}{ds_b^e} - \lambda_b^r \phi'(q_v) \left[\frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_b^e} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_b^e} \right] \\
&\quad + \lambda_b^r \phi'(q_v) \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_b^e}
\end{aligned}$$

$$\begin{aligned}
\frac{dV_v}{ds_b^e} &= \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_b^e} + \frac{\partial L_v}{\partial q_b} \frac{dq_b}{ds_b^e} \\
&= \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_b^e} + \frac{\partial L_v}{\partial q_b} \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \left(1 + \frac{d\tau_b^e}{ds_b^e} \right) \\
&\quad - \frac{\partial L_v}{\partial q_b} \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_b^e} \\
&\quad - \frac{\partial L_v}{\partial q_b} \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_b^e} \\
&= -\lambda_v^e u'(c_b^e) \frac{d\tau_b^e}{ds_b^e} + \lambda_v^r \phi'(q_b) \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \left(1 + \frac{d\tau_b^e}{ds_b^e} \right) \\
&\quad - \lambda_v^r \phi'(q_b) \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_b^e} \\
&\quad - \lambda_v^r \phi'(q_b) \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_b^e}
\end{aligned}$$

$$\begin{aligned}
\frac{dV_b}{ds_v^e} &= \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_v^e} + \frac{\partial L_b}{\partial q_v} \frac{dq_v}{ds_v^e} \\
&= \frac{\partial L_b}{\partial \tau_v^e} \frac{d\tau_v^e}{ds_v^e} + \frac{\partial L_b}{\partial q_v} \left[\frac{N_v^e}{N_v^r \theta} + \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_v^e} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_v^e} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_v^e} \right] \\
&= \lambda_b^e u'(c_v^e) \frac{d\tau_v^e}{ds_v^e} - \lambda_b^r \phi'(q_v) \left[\frac{N_v^e}{N_v^r \theta} + \frac{N_v^e}{N_v^r \theta} \frac{d\tau_v^e}{ds_v^e} + \frac{\tau_v^e + s_v^e}{N_v^r \theta} \frac{dN_v^e}{ds_v^e} - \frac{N_v^e (\tau_v^e + s_v^e) + G_v}{(N_v^r)^2 \theta} \frac{dN_v^r}{ds_v^e} \right]
\end{aligned}$$

$$\begin{aligned}
\frac{dV_v}{ds_v^e} &= \frac{\partial L_v}{\partial s_v^e} + \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_v^e} + \frac{\partial L_v}{\partial q_b} \frac{dq_b}{ds_v^e} \\
&= \frac{\partial L_v}{\partial s_v^e} + \frac{\partial L_v}{\partial \tau_b^e} \frac{d\tau_b^e}{ds_v^e} + \frac{\partial L_v}{\partial q_b} \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{d\tau_b^e}{ds_v^e} \\
&\quad - \frac{\partial L_v}{\partial q_b} \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_v^e} \\
&\quad - \frac{\partial L_v}{\partial q_b} \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_v^e} \\
&= \mu_v N_v^e - \lambda_v^e u'(c_b^e) \frac{d\tau_b^e}{ds_v^e} + \lambda_v^r \phi'(q_b) \frac{(N^e - N_v^e)}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{d\tau_b^e}{ds_v^e} \\
&\quad - \lambda_v^r \phi'(q_b) \frac{(\tau_b^e + s_b^e) - q_b (N^r - N_v^r) F'}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^e}{ds_v^e} \\
&\quad - \lambda_v^r \phi'(q_b) \frac{s_b^r - q_b [F - (N^r - N_v^r) q_b F']}{(N^r - N_v^r) [F - q_b (N^r - N_v^r) F']} \frac{dN_v^r}{ds_v^e}
\end{aligned}$$

References

- [1] Aronsson, T. and M. Wikström (2001). Optimal Taxes and Transfers in a Multilevel Public Sector. *FinanzArchiv* 58, 158-166.
- [2] Bergstrom T. and S. Blomquist (1996). The Political Economy of Subsidized Day Care. *European Journal of Political Economy* 12, 443-457.
- [3] 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,
- [4] Boadway, R. and M. Marchand (1995). The Use of Public Expenditures for Redistributive Purposes. *Oxford Economic Papers* 47, 45-59.
- [5] Boadway, R. and M. Keen (1996). Efficiency and the Optimal Direction of Federal-State Transfers. *International Tax and Public Finance* 3, 137-55.

- [6] Boadway, R., Marchand, M. and M. Vigneault (1998). The consequences of Overlapping Tax Bases for Redistribution and Public Spending in a Federation. *Journal of Public Economics* 68, 453-78.
- [7] Boadway, R., Cuff, K and M. Marchand (2003). Equalization and the Decentralization of Revenue-Raising in a Federation. *Journal of Public Economic Theory* 5, 201-228.
- [8] Cremer, H. and F. Gahvari (1997). In-Kind Transfers, Self-Selection and Optimal Tax Policy. *European Economic Review* 41, 97-114.
- [9] Dahlby, B. (1996). Fiscal Externalities and the Design of Intergovernmental Grants. *International Tax and Public Finance* 3, 397-412.
- [10] Hansson, I. and C. Stuart (1987). The Suboptimality of Local Taxation under Two-Tier Fiscal Federalism. *European Journal of Political Economy* 3, 407-411.
- [11] Johnson, W.R. (1988). Income Redistribution in a Federal System. *American Economic Review* 78, 570-73.
- [12] Korpi, M. (2003). Regional obalanser- ett demografiskt perspektiv. Working paper 2003:1, Institute For Futures Studies.
- [13] Mansoorian, A. and G. M. Myers (1997). On the Consequences of Government Objectives for Economies with Mobile Populations. *Journal of Public Economics* 63, 265-281.
- [14] Mirrlees, J. A. (1972). The Optimum Town. *Swedish Journal of Economics* 74, 114-135.
- [15] Oates, W. (1972). *Fiscal Federalism*. Harcourt Brace, Jovanovich, New York.