Does mobility of educated workers undermine decentralized education policies?

Christiane Schuppert *

January 2007

The present paper studies a multi-jurisdictional framework, in which, from a federal perspective, educational subsidies turn out to be efficiency enhancing. However, in the presence of mobile high-skilled labor, local jurisdictions might try to free-ride on other regions education policies and abstain from subsidizing education. Social mobility is introduced as an additional dimension of labor mobility. Thus, educational subsidies not only affect the size of the mobile high-skilled, but furthermore the size of the immobile low-skilled workforce. Using this framework, it is shown that local governments abide to the optimal decision rule for subsidizing education and decentralized education policies remain to be efficient, although high-skilled workers are perfectly mobile. Only if one allows for high- and low-skilled mobility, local incentives to promote education vanish.

Keywords: optimal taxation, migration, education

JEL-Classification: I22, F22, H21

*University of Dortmund, Economics Department, 44227 Dortmund, Germany, email: c.schuppert@wiso.uni-dortmund.de, phone: +49-231-7554374.
I would like to thank Christian Bayer, Ben Lockwood and Wolfram Richter for helpful comments and suggestions. Financial support by the German Research Foundation DFG is gratefully acknowledged.

1. Introduction

In the course of economic integration, mobility of the labor force has increased considerably due to a rapid decline in transaction costs. Administrative or legal barriers have been reduced in highly integrated regions such as the European Union, while language or cultural obstacles to migration are of minor importance in federal states. Moreover, the mobility of labor increases with the skill level as specialized skills of highly educated workers may be locally less demanded and, hence, require searching a geographically larger labor market. In view of the high mobility especially of skilled labor, benefits resulting from local investments in education do not necessarily accrue to the region of origin. Therefore, public incentives to promote mobility enhancing education might be limited. Rather, it seems to be tempting to attract skilled labor from abroad using a favorable tax system, and free-ride on other regions' education policies. Hence, high-skilled mobility can be problematic for decentralized fiscal policies.

The present paper analyzes this free-rider problem at the level of local jurisdictions in a setup that allows for social mobility across skill types. Individual underinvestment in education provides a rationale for educational subsidies from a federal point of view.Local governments, however, might refrain from subsidizing education in the presence of high-skilled migration. Since the size of the educated workforce available in the respective jurisdiction is no longer determined by domestic educational investments but by migration flows, local governments might lose the incentive to correct private underinvestment. This suggests that decentralized education policies will be inefficient, and educational subsidies have to be provided at the federal or supranational level. Yet, local governments do engage in education policies, and subsidize especially early education substantially. The present paper provides an explanation for this phenomenon and shows, that local governments do not free-ride on other regions' education policies even though high-skilled workers are assumed to be perfectly mobile. Most previous studies stress an efficiency enhancing aspect of migration that goes ahead with a reduced need for public policy. Surprisingly, the recent literature on the brain drain has emphasized this point in various studies. The welfare improving impact of migration in new approaches to the brain drain rests on the observation that an exogenous increase in the migration probability of high-skilled labor fosters private educational investments as in Stark et al. (1998). Starting from a situation of educational underinvestment that is induced by externalities, increasing the probability to emigrate raises private incentives to accumulate human capital and, hence, welfare.[1] Since educational underinvestment provides a rational for public intervention, migration can even function as a substitute for subsidies in education (Stark and Wang, 2002).

Similarly, migration can increase efficiency in a setup in which private underinvestment results from uncertainty. Wildasin (2000) considers a framework in which high-skilled workers possess industry-specific human capital. This implies, that qualified labor is intersectorally immobile and, therefore, exposed to earning risks. If education is privately financed and wage risks are uninsurable,[2] globalization that raises the geographical mobility of high-skilled labor provides full insurance of the involved income risks. As a consequence, investments in human capital increase to an efficient level.[3] At the same time, migration can restore efficiency of local tax policies in a setup where educational subsidies are used to overcome a hold-up problem of time-inconsistent taxation. Andersson and Konrad (2003a,b) point out that in such a framework allowing for mobility of labor reintroduces the elasticity of the tax base and serves as a commitment device for low tax rates. Against this background they determine the welfare effects of globalization when education is risky.

These studies suggest that an increase in labor mobility should be accompanied by a decline in public investments due to a reduced need for fiscal policies. The present paper does not intend to cast doubt on this efficiency enhancing effect of migration. Yet, it is presumed that mobility of high-skilled workers will most likely not entirely restore efficiency. If migration fails to exactly offset the inefficiencies, private underinvestment persists. This calls for public policies to correct the underlying market failure. However, from point of view of a small jurisdiction facing migration of highly

[1]The possibility of an efficiency enhancing brain drain on the basis of a probabilistic approach to migration has been discussed in models with underinvestment due to production externalities (Stark (2004); Mountford (1997)) as well as intergenerational transmission of knowledge (Beine et al. (2001); Vidal (1998)).

[2]Wildasin (2000) analyzes the effect of high-skilled labor mobility for both, the case of private and public investments in education.

[3]In a comparable approach, Poutvaara (2000, 2001) model wage-tax financed educational transfers to students that can insure against regions-specific shocks. On this background, the trade-off between the efficiency enhancing effect of mobility and the possible erosion of local tax policies is assessed.

educated workers, the size of the high-skilled workforce available depends on migration flows and not on previous local investments in education. Therefore, the possibility to free-ride on other regions' educational investments arises, and decentralized policies in the presence of labor mobility might not be efficient anymore. As Sinn (1997) puts it, fiscal competition reduces the incentives of governments to correct market failures.

The study by Wildasin (2000) cited above additionally considers the case of public investment in human capital. It is shown, that if education is financed publicly, tax competition for high-skilled workers results in public underinvestment. Hence, decentralized education policies turn out to be inefficient. Similarly, Justman and Thisse (1997, 2000) show that mobility of high-skilled labor induces underprovision if regions interact strategically. They consider a model in which the supply of skilled workers is determined solely by public instead of private investments. Since migrants are assumed to respond to wage differences across regions, increased mobility results in underprovision of public education. To restore local incentives to provide public education, a system of interjurisdictional transfers based on migration flows is suggested.

The cited studies on education policies in the presence of migration conclude that government activity will decline if mobility rises, either due to a reduced need for fiscal intervention or due to diminished incentives to correct market failures. Hence, they fail to explain why education is still subsidized substantially even at the local level. The focus of the present paper lies entirely on the incentives of local jurisdictions to correct underinvestment in education, and ignores positive efficiency enhancing effects of migration at the private level. It is shown that despite of the mobility of educated workers, local governments abide by the optimal decision rule for subsidizing education. This is due to the fact that social mobility has been introduced as an additional dimension of labor mobility. As a consequence, lacking governmental support of education can not be compensated by immigration as it additionally affects the number of immobile low-skilled workers in a respective region. Thus, local governments correct private underinvestment in education even in the presence of perfect mobility of high-skilled workers, and decentralized decision making remains to be efficient. Only if one allows for high- and low-skilled mobility, local incentives to correct the underinvestment problem vanish as education policies neither affect the size of the domestic high- nor of the low-skilled workforce. This implies, that private underinvestment in education persists. To correct the market failure and establish the social optimum in a decentralized setting, a federal matching grant can be used. However, such a federal matching grant effectively assigns the education policy to the federal or supranational level.

The paper is organized as follows. In Section 2, the basic model setup is described with a special emphasis on the educational process and the migration dynamics. Section 3 and 4 derive the optimal fiscal policies a unitary state government chooses and discuss what policy instruments are needed to decentralize the welfare optimum. First, the case of perfect mobility of the educated workforce is considered, and then migration of high- as well as low-skilled labor is allowed for. The last section concludes.

2. The Model

Consider a federal economy with an infinite number of small local jurisdictions i.[4] Each of these jurisdictions represents a local tax authority that can raise lump-sum taxes to finance educational subsidies, given the fiscal policy decided by the central government. Initially and before any migration takes place, the population of region i consists of a given number of high-skilled ($N^H_{0,i}$) and low-skilled households ($N^L_{0,i}$). The current working population, $N^H_{t,i}$ and $N^L_{t,i}$, is determined by factor flows at the beginning of every period $t \geq 1$. The overall size of the population is fixed over time with $\bar{N} = \sum_{i=1}^{\infty} \left(N^H_{t,i} + N^L_{t,i} \right) \forall t \geq 0$.

Regions are populated by successive generations. Every household or parent has one offspring and invests an amount e^n_t in the education of this child, where $n = H, L$ indicates the respective type of the parent. This educational investment determines the child's probability of becoming high-skilled. Thus, while the overall federal population is given, the evolution of types depends on the regional investments in education.

The timing of events is as follows: The central government moves first, deciding on the sequence of fiscal policies that maximizes the welfare of the aggregate federal population. The local governments move next, behaving like small open economies

[4]This assumption is needed to abstract from any strategic interaction between regions.

when choosing the sequence of regional tax rates, followed by private agents, who take all tax rates as given. Note, that time-inconsistency of fiscal policies is not an issue here, because households are not optimizing intertemporally. Thus, the possibility of a future revision of the initially announced path of optimal policy will not influence the decisions of the working population as these are only affected by current tax rates.[5]

The basic model setup follows Boadway et al. (2003), who analyze fiscal equalization in a static model with two types of mobile labor. Their work is extended to a dynamic model with successive generations and an endogenous human capital formation process to study the consequences of labor mobility for optimal education policies. The production of human capital is based on a setup suggested by Cremer and Pestieau (2004). In their framework the educational success is determined by an endogenously derived probability of becoming high-skilled. In the present paper, this is interpreted as social mobility which reflects the fact that children of both low- and high-skilled parents face a positive probability to become high-skilled themselves. The approach makes it possible to analyze fiscal policies that not only affect the allocation of the mobile factor across regions, but additionally determine the endowment with the factor itself.

2.1. Regional Production

In every period $t \geq 1$, firms produce a single aggregate good that can be used for consumption and investments in education. Labor is the only input factor with high-skilled ($N_{t,i}^H$) and low-skilled households ($N_{t,i}^L$) being perfect substitutes. Households supply z^n efficiency units of labor inelastically, with $n = H, L$ denoting the respective skill group, and it is assumed that high-skilled labor is more productive ($z^H > z^L$). Aggregate effective labor supply can be written as $Z_{t,i} = z^H N_{t,i}^H + z^L N_{t,i}^L$. Firms in every region have access to the production function $F(Z_{t,i})$ with $F'(Z_{t,i}) > 0 >$

[5] As Kydland and Prescott (1977) point out, time inconsistencies arise solely in situations in which the current optimization behaviour of agents is influenced by their expectations of future fiscal policies.

$F''(Z_{t,i})$. As an example one can think of the following production technology

$$F(Z_{t,i}) = (Z_{t,i})^{\alpha} \tag{1}$$

where $0 < \alpha < 1$ denotes the production elasticity of labor. Labor markets are competitive, therefore the wage rate equals the marginal product of labor.

As the production function exhibits decreasing returns to scale, local rents arise. The rent income of region i is given by $R(Z_{t,i}) = F(Z_{t,i}) - Z_{t,i}F'(Z_{t,i})$ with $R'(Z_{t,i}) = -Z_{t,i}F''(Z_{t,i}) > 0$. It is assumed that these rents are entirely accrued by the regional government.[6] Since regions have access to the same production technology, that is fiscal capacities of local jurisdictions do not differ, there is not need for federal equalization. This implies that it is irrelevant whether rents go to the regional or central budget. Contrary, with private ownership of the fixed factor, source income arises that induces an additional fiscal externality of migration. Since the focus lies on the impact of migration on optimal education policies, inefficiencies resulting from rent income are not considered, and it is assumed that rents go to the local governments.

2.2. Central and Local Governments

Both the central and the regional governments are benevolent in the sense that they maximize the sum of their residents' utility, discounted over all periods. Put differently, governments care about the aggregate welfare of each parent generation living in the respective region. The central government has access to a federal lump-sum tax $\theta^n_{t,i}$, which can be differentiated both across regions and across skill-types. Furthermore, it can choose a skill-specific matching grant $\theta^n_{et,i}$ paid to regions for every unit invested locally in education. The budget constraint of the central government can be written as

$$\sum_{i=1}^{\infty} \sum_{n} N^n_{t,i} \left(\theta^n_{t,i} - \theta^n_{et,i} e^n_{t,i} \right) = 0. \tag{2}$$

[6]This can be justified by the implicit assumption that either the local jurisdiction is the owner of the fixed factor who generates the rents, or that it has access to a rent tax to fully appropriate the pure profit.

The set of available tax instruments at the local level of government is restricted to a personal head tax on residents (τ_t^n) and a skill-specific educational subsidy or tax (τ_{et}^n). The budget of the local government of region i additionally includes the regional rent,

$$\sum_n N_t^n \left(\tau_t^n - \left[\tau_{et}^n - \theta_{et,i}^n \right] e_t^n \right) + R(Z_t) = 0. \tag{3}$$

2.3. Household Behavior

Following Cremer and Pestieau (2004),[7] successive generations of two types of labor, low-skilled (N_t^L) and high-skilled (N_t^H) are assumed. Each of these workers has one offspring and is, hence, also referred to as a parent. Parents invest an amount e^n in the education of their children, and thereby determine the probability of their child to become high-skilled, $h(e^n)$ with $h_e^n > 0$ and $h_{ee}^n \geq 0$.[8] While young, children undergo education, but only enter the model explicitly when old, that is, once they have completed their education and start working as either high- or low-skilled. Note, that individuals do not decide on their own education, but only on the amount invested in their children. Thus, the model rather depicts basic or early education as compared to college or university education.

The probability to become high-skilled, and, thus, high-productive, is derived endogenously as it is a function of the different educational investments of the respective type of parent: The probability to become high-skilled is $h(e^H)$ if parents are high-skilled, and $h(e^L)$ if they are low-skilled. The amount invested by parents is determined by the net earnings realized by the respective type. Hence, social mobility is contingent on the productivity, or ultimately, on the educational background of parents.

Since both children of high- and low-skilled parents face a positive probability of becoming high-skilled themselves, the model allows for social mobility across skill types. Yet, in the absence of taxation, children of high-skilled, that is high-productive parents face a higher probability of becoming high-skilled themselves. This is due to the fact that children differ in their acquired ability whenever high-skilled parents

[7] Cremer and Pestieau (2004) consider the case of an immobile workforce and study optimal education policies when private investment is not observable.

[8] The region and time index is suppressed for the moment.

spent more resources on education than low-skilled parents do. This will be the case in the absence of any fiscal policy, as the high-skilled type is more productive, and earns a higher wage income. Therefore, the model additionally captures intergenerational earnings persistence, which can, however, be reduced using educational subsidies.

The framework replicates basic findings of the recent literature that stress the importance of early investments in shaping the cognitive ability of children that in turn determines their later educational success as well as their income prospects. Restuccia and Urrutia (2004) calibrate a model in which innate ability, acquired ability, based on parental investments in early education, and college education determine the probability of successful college graduation. They show that parental investments in education, especially early education, account for nearly one-half of the observed intergenerational earnings persistence. This evidence suggests that social mobility can be increased substantially by the provision of educational subsidies on private investments in early education. Additionally, Carneiro and Heckman (2002) stress the importance of long-run factors to explain the positive correlation between college enrolment and family income. They argue that children from high-income families have better access to resources that provide them with higher quality of education early in life leading to superior cognitive ability in the long-run.

Based on parent's educational investment in period $t-1$, the size of the aggregate high-skilled labor force in period t can be derived as

$$N_t^H = N_{t-1}^H \cdot h(e_{t-1}^H) + N_{t-1}^L \cdot h(e_{t-1}^L). \tag{4}$$

Analogously, the number of low-skilled workers can be deduced,

$$N_t^L = N_{t-1}^H \cdot (1 - h(e_{t-1}^H)) + N_{t-1}^L \cdot (1 - h(e_{t-1}^L)). \tag{5}$$

After the educational process determined the respective type, the child enters the working period as an either high- or low-skilled. Every worker or household is endowed with z^n efficiency units of labor that she supplies inelastically to firms in the region of residence. Labor income is spend on consumption, tax payments as well as net investment in the education of children. Therefore, the household's budget constraint can be expressed as

$$z^n F'(Z) - \tau^n - \theta^n = c^n + (1 - \tau_e^n)e^n, \tag{6}$$

where c^n denotes consumption and e^n investment in education of children. Note, that a consumption tax is not allowed for at neither level of government, although such a tax is a perfect substitute for an educational subsidy as it increases the opportunity cost of investments in education. Yet, a consumption tax is typically not differentiable across skill-types, and can thus not mimic type-specific education policies.[9]

Parents are altruistic in the sense that they experience a joy of giving when supporting their children's education (warm glow altruism). Preferences of different skill types are identical and separable between consumption and educational investments. They can be expressed by the strictly quasi concave utility function

$$U(c^n, e^n) = u(c^n) + \upsilon(e^n). \tag{7}$$

Households of each type maximize utility, taking both federal and regional tax rates as given. The optimization leads to the marginal rate of substitution between consumptions and investments in education, that equals the private cost of education in the optimum,[10]

$$\frac{\upsilon_e^n}{u_c^n} = (1 - \tau_e^n). \tag{8}$$

While equations (4) and (5) illustrate, that private investments in education have a crucial impact on the composition of the future workforce, parents optimize solely their own utility not taking into account the effect on future generations. Put differently, they do not fully anticipate their child's benefits resulting from these educational investments. As will be shown, this intergenerational externality leads to inefficient levels of investment in education that can be corrected using an educational subsidy.

[9] While a consumption tax is inefficient ex ante as it can not discriminate between types, it will turn out to be efficient ex post. However, this is due to the fact that the assumed utilitarian welfare function entails redistribution among types. This implies a symmetric equilibrium ex post, in which a consumption tax and an educational subsidy are indeed perfect substitutes.

[10] In the absence of any tax or subsidy, the private cost of education in terms of consumption equals one. This reflects the underlying assumption that consumption and education are produced with the same production technology.

2.4. Migration

Two different scenarios of migration are considered: mobility of the high-skilled workforce, and mobility of both the high- and the low-skilled workforce. For simplicity, the analysis abstracts from any migrations costs such as language barriers, moving costs or attachment to the home country.[11] Migration takes place at the beginning of every period, before households decide on consumption and educational spending. Hence, migration flows determine the workforce of a respective region in period t.

A potential migrant is indifferent between migrating or staying whenever utility is equalized across regions. Thus, a migration equilibrium between any regions $i \neq j$ and an arbitrarily chosen reference region j is characterized by identical utility levels,

$$u(c^n_{t,i}) + \upsilon(e^n_{t,i}) = u(c^n_{t,j}) + \upsilon(e^n_{t,j}). \tag{9}$$

Equation (9) implicitly defines the quantity of mobile labor allocated in a particular region after a migration equilibrium has been reached.

Introducing mobility of households affects the constraints for fiscal policies both at the federal and at the local level. While in a closed economy the number of low- and high-skilled workers available in a particular region is fully determined by investments in education, this is no longer true if one considers migration. Rather, migration flows are crucial for the allocation across regions. Yet, the size of the mobile population group in the federal state as a whole is still contingent on regional investments in education and is, hence, restricted. The federal government will take this into account, while the government of a small open region will view the supply of the mobile factor as infinite. Note that, as long as solely mobility of the high-skilled is considered, the number of the immobile low-skilled workforce in a particular region still depends on local human capital formation. With mobility of high- and low-skilled, however, migration flows alone determine the allocation of types across regions.

[11] As will be discussed below, this assumption influences the equilibrium attained. However, it has no impact on the results concerning the efficiency of regional education policies.

3. Optimal Education Policies with High-Skilled Migration

In the following section, optimal education policies at both the federal and regional level of government are studied. In the first subsection, it is shown that parents underinvest in education. This inefficiency justifies federal education policies on efficiency grounds. Yet, if one consider the possibility of migration, a free-rider problem at the level of local governments might arise. Since, from a local perspective, the size of the mobile workforce no longer depends on educational investments but on migration incentives, local governments might substitute efficiency enhancing education subsidies by fiscal policies that aim at attracting migrants. In the absence of federal education policies, private underinvestment in education might, therefore, persist. The second subsection analyzes whether this regional free-rider problem indeed occurs, and federal education policies are needed to correct the intergenerational inefficiency. To this aim, the unitary state optimum is derived as a benchmark case. Then, regional policies as well as federal tax instruments that are needed to decentralize the first best optimum are discussed.

3.1. Unitary State Optimum

Consider a unitary state government that is benevolent in the sense that it maximizes a utilitarian welfare function over generations, discounted by the social rate of time preference, $\beta < 1$. To characterize the central planning solution as a benchmark case, assume that the unitary state government does not only decide on federal taxes, but also optimally chooses the tax instruments available to regions. Note that the household's first-order condition has been used to eliminate $\tau^n_{et,i}$ from the central planner's optimization problem.

The unitary state government optimizes social welfare by choosing the tax rates $\left\{\tau^n_{t,i}, \theta^n_{t,i}, \theta^n_{et,i}\right\}$ as well as the allocations $\left\{c^n_{t,i}, e^n_{t,i}, N^n_{t,i}\right\}$. The population $N^n_{t,i}$ is treated as an artificial control variable, since the human capital formation as well as the migration equilibrium have been added as constraints to the optimization problem. Alternatively, one can determine the respective population group as an endogenous variable using equations (5) and (9).

The optimization problem is to solve

$$
\begin{aligned}
&\max \sum_{t=1}^{\infty} \beta^{t} \Big\{ \ \sum_{i=1}^{\infty} \sum_{n} N_{t,i}^{n} \left[u\left(c_{t,i}^{n}\right) + \upsilon(e_{t,i}^{n}) \right] \\
&+ \sum_{i=1}^{\infty} \sum_{n} \kappa_{t,i}^{n} \left[z^{n} F'(Z_{t,i}) - \tau_{t,i}^{n} - \theta_{t,i}^{n} - c_{t,i}^{n} - \frac{\upsilon_{et,i}^{n}}{u_{ct,i}^{n}} e_{t,i}^{n} \right] \\
&+ \lambda_{t} \sum_{i=1}^{\infty} \sum_{n} N_{t,i}^{n} \left(\theta_{t,i}^{n} - \theta_{et,i}^{n} e_{t,i}^{n} \right) \\
&+ \sum_{i=1}^{\infty} \lambda_{t,i} \left[\sum_{n} N_{t,i}^{n} (\tau_{t,i}^{n} + \frac{\upsilon_{et,i}^{n}}{u_{ct,i}^{n}} e_{t,i}^{n} - \left(1 - \theta_{et,i}^{n}\right) e_{t,i}^{n}) + R(Z_{t,i}) \right] \\
&+ \sum_{i=1}^{\infty} \mu_{t,i}^{L} \left[N_{t+1,i}^{L} - \sum_{n} N_{t,i}^{n} \left[1 - h(e_{t,i}^{n}) \right] \right] \\
&+ \mu_{t} \left[\bar{N} - \sum_{i=1}^{\infty} \sum_{n} N_{t,i}^{n} \right] \\
&+ \sum_{i \neq j} \varphi_{t,i}^{H} \left[u(c_{t,j}^{H}) + \upsilon(e_{t,j}^{H}) - u\left(c_{t,i}^{H}\right) - \upsilon(e_{t,i}^{H}) \right] \ \Big\},
\end{aligned}
\tag{10}
$$

where the variables $\kappa_{t,i}^{n}, \lambda_t, \lambda_{t,i}, \mu_{t,i}^{L}, \mu_t$ and $\varphi_{t,i}^{H}$ denote the Lagrange-multipliers for the respective optimization constraints. The first constraint guarantees that the household's budget is balanced. The next two constraints refer to the federal and regional budget that are distinguished to keep the solution comparable to the regional optimization discussed below. The fourth constraint reflects the fact that the size of a regions' immobile low-skilled population in period $t+1$ is contingent on local educational investments in period t. The remaining constraints illustrate that the aggregate labor force $\bar{N}_t$ has to allocate in one of the regions, and that the mobility of high-skilled workers requires that the migration constraint is met. Hence, the last constraint determines the allocation of educated workers, given that the utility level in region i equals the utility attainable in an arbitrarily chosen reference region j. The first-order conditions of the social planning problem are stated in the Appendix.

Optimal Policy Rule

The first-order conditions with respect to $\tau^n_{t,i}$, $\theta^n_{t,i}$ and $\theta^n_{et,i}$ reveal that $\kappa^n_{t,i} = \lambda_{t,i} N^n_{t,i}$ and $\lambda_{t,i} = \lambda_t$. This reduces the first-order conditions on consumption and educational spending for the low-skilled residing in region i to

$$u^L_{ct,i} = \lambda_t \qquad \text{and} \qquad \upsilon^L_{et,i} + \mu^L_{t,i} h^L_{et,i} = \lambda_t. \tag{11}$$

Equations (11) define the equilibrium values of $c^L_{t,i}$ and $e^L_{t,i}$, and can be used to derive the social marginal rate of substitution between consumption and investments in education,

$$\frac{\upsilon^L_{et,i}}{u^L_{ct,i}} = 1 - \frac{\mu^L_{t,i}}{\lambda_t} h^L_{et,i}. \tag{12}$$

In the welfare optimum, the social marginal rate of substitution has to equal the private cost of education minus the term $\frac{\mu^L_{t,i}}{\lambda_t} h^L_{et,i}$ that captures a social impact of education. This social benefit of education consists of the welfare gain of a declining number of low-skilled in $t+1$ ($\mu^L_{t,i}$), weighted by the marginal productivity of low-skilled educational investments ($h^L_{et,i} > 0$) and discounted by the opportunity cost, the marginal utility of consumption ($\lambda_t > 0$). If the contribution to aggregate welfare of a high-skilled worker exceeds the benefit of her being low-skilled, the social impact of education is positive. This implies that investments in education should take place until the marginal rate of substitution exceeds the private costs of one to additionally account for the social benefit of human capital formation.

Since in any decentralized market equilibrium this optimal marginal rate of substitution coincides with the private marginal rate of substitution (equation (8)), the optimal policy rule to determine the educational subsidy or tax for the low-skilled in region i can be deduced,

$$1 - \frac{\mu^L_{t,i}}{\lambda_t} h^L_{et,i} \underset{(12)}{=} \frac{\upsilon^L_{et,i}}{u^L_{ct,i}} \underset{(8)}{=} (1 - \tau^L_{et,i}) \quad \Leftrightarrow \quad \tau^L_{et,i} = \frac{\mu^L_{t,i}}{\lambda_t} h^L_{et,i}. \tag{13}$$

In the absence of any education policy, households adjust their private marginal rate of substitution to a marginal cost of educational investment of one (equation (8)). Hence, they do not take into account the social benefit of education and private investment will be inefficient. The educational subsidy or tax is a means to internalize this social impact of education and establish the welfare optimum. As long as the social impact of education is positive ($\mu^L_{t,i} > 0$), it is optimal to subsidize educational investments.

Analogously, one can rewrite the first-order conditions with respect to consumption and educational investment for the high-skilled type in region i to

$$\frac{\left(N^H_{t,i} - \varphi^H_{t,i}\right)}{N^H_{t,i}} u^H_{ct,i} = \lambda_t \qquad \text{and} \qquad \frac{\left(N^H_{t,i} - \varphi^H_{t,i}\right)}{N^H_{t,i}} \upsilon^H_{et,i} + \mu^L_{t,i} h^H_{et,i} = \lambda_t. \tag{14}$$

Again, equations (14) define the equilibrium values of $c^H_{t,i}$ and $e^H_{t,i}$. Moreover, the social marginal rate of substitution between consumption and educational investments involves the social impact of education,

$$\frac{\upsilon^H_{et,i}}{u^H_{ct,i}} = 1 - \frac{\mu^L_{t,i}}{\lambda_t} h^H_{et,i}. \tag{15}$$

This implies that - similarly to the low-skilled case - a subsidy or tax on educational investments should be used to internalize the social effect of education. Following the above procedure, the optimal educational subsidy or tax on high-skilled investments can be obtained as

$$\tau^H_{et,i} = \frac{\mu^L_{t,i}}{\lambda_t} h^H_{et,i}. \tag{16}$$

One can further simplify the educational subsidy using the first-order conditions on consumption and educational investments (equations (11) and (14)) to substitute out the Lagrange-multipliers λ_t and $\mu^L_{t,i}$,

$$\tau^n_{et,i} = \frac{\mu^L_{t,i}}{\lambda_t} h^n_{et,i} = \frac{u^n_{ct,i} - \upsilon^n_{et,i}}{u^n_{ct,i}}. \tag{17}$$

Equation (17) reveals that the optimal policy rule for subsidizing education is only dependent on the realized equilibrium values of consumption and education. Therefore, it is not influenced by assumptions on neither migration dynamics nor the underlying production technology.[12]

[12]While the policy rule remains unaffected, the assumptions certainly influence the equilibrium obtained.

Symmetric Intraregional Equilibrium

Consider now a situation in which the migration equilibrium constraint is not binding ($\varphi^H_{t,i} = 0$).[13] In this case, the first-order conditions as stated in equations (11) and (14) become

$$u^n_{ct,i} = \lambda_t \qquad \text{and} \qquad v^n_{et,i} + \mu^L_{t,i} h^n_{et,i} = \lambda_t. \tag{18}$$

From equations (18) it follows that consumption and educational investments in a particular region are type-independent ($c^n_{t,i} = c_{t,i}$ and $e^n_{t,i} = e_{t,i}$). Since preferences are the same, it follows that utility levels across types will be equalized. This is due to the fact that the unitary state governments aims at maximizing a utilitarian welfare function. To see this, consider a situation in which the level of educational spending of different types is identical, but high-skilled consumption exceeds low-skilled consumption. In this case, the marginal utility of additional consumption is higher for the low-skilled type. Therefore, redistribution that increases low-skilled consumption at the expense of high-skilled consumption generates a welfare gain.

Moreover, consumption is independent of the place of residence ($c_{t,i} = c_{t,j}$), while this need not be the case for educational investments. As long as the Lagrange-multiplier on the human capital formation constraint, $\mu^L_{t,i}$, is not identical across regions, the level of spending on education might differ ($e_{t,i} \gtreqless e_{t,j}$).

Identical consumption and investment levels in a respective region yield the result that the optimal subsidy or tax on education is independent of the respective type,

$$\tau_{et,i} = \frac{\mu^L_{t,i}}{\lambda_t} h_{et,i}. \tag{19}$$

This implies that both types spend an equal amount of resources on consumption and educational investments. According to the households budget constraint, identical spending patterns can only be achieved if the net income is equalized across types,

$$z^H F'(Z_{t,i}) - \tau^H_{t,i} - \theta^H_{t,i} = c_{t,i} + (1 - \tau_{et.,i}) e_{t,i} = z^L F'(Z_{t,i}) - \tau^L_{t,i} - \theta^L_{t,i} \tag{20}$$

[13]The solution to the optimization problem is derived ignoring the migration constraint. It is shown that the equilibrium deduced satisfies the required utility equalization as well. Hence, the obtained solution coincides with the one for the fully constrained optimization problem.

Since productivity differs, the aggregate head tax on high-skilled has to exceed the one on low-skilled workers,

$$\left(z^H - z^L\right) F'(Z_{t,i}) = \left(\tau^H_{t,i} + \theta^H_{t,i}\right) - \left(\tau^L_{t,i} + \theta^L_{t,i}\right) > 0. \tag{21}$$

Lump-sum taxes that account for the productivity difference between the two types combined with type-independent educational subsidies guarantee that the net income is identical. Such a way of redistribution between high- and low-skilled workers ensures that consumption and educational investments are the same, which entails that realized utility levels are type-independent and aggregate welfare is maximized.

Proposition 1 *A benevolent unitary state government redistributes income from high- to low-skilled labor to achieve a symmetric intraregional equilibrium.*

Still, it is not clear whether education should be subsidized or taxed. To determine the sign of the educational subsidy, one needs to evaluate the shadow price $\mu^L_{t,i}$ associated with the probability of being low-productive. Using the first-order conditions on $\tau^n_{t,i}$, $\theta^n_{t,i}$ and $\theta^n_{et,i}$ as well as the fact that $R'(Z_{t,i}) = -Z_{t,i}F''(Z_{t,i})$ and $\frac{v^n_{et,i}}{u^n_{ct,i}} = (1 - \tau^n_{et,i})$, one can simplify the first-order condition with respect to $N^H_{t,i}$ to

$$\mu_t = U^H_{t,i}(\cdot) + \lambda_t \left(\theta^H_{t,i} + \tau^H_{t,i} - \tau^H_{et,i} e^H_{t,i}\right) - \mu^L_{t,i}\left[1 - h(e^H_{t,i})\right]. \tag{22}$$

Equation (22) captures the contribution to social welfare of an additional high-skilled. This net benefit of being high-skilled comprises the attained level of utility and the fiscal revenue raised, minus the probability that the respective type's child will later become low-skilled.

Similarly, one can rewrite the first-order condition with respect to $N^L_{t,i}$ to

$$\mu_t = U^L_{t,i}(\cdot) + \lambda_t \left(\theta^L_{t,i} + \tau^L_{t,i} - \tau^L_{et,i} e^L_{t,i}\right) - \mu^L_{t,i}\left[1 - h(e^L_{t,i})\right] + \beta^{-1}\mu^L_{t-1,i}. \tag{23}$$

Substituting (22) into (23) and rearranging yields the shadow price $\mu^L_{t-1,i}$ that captures the social benefit of turning a child in period $t-1$ into a high-skilled instead of a low-skilled worker,

$$\beta^{-1}\mu^L_{t-1,i} = U^H_{t,i} - U^L_{t,i} + \lambda_t \left[\left(\theta^H_{t,i} + \tau^H_{t,i} - \tau^H_{et,i}e^H_{t,i}\right) - \left(\theta^L_{t,i} + \tau^L_{t,i} - \tau^L_{et,i}e^L_{t,i}\right) \right] + \mu^L_{t,i}\left[h(e^H_{t,i}) - h(e^L_{t,i})\right]. \quad (24)$$

This social benefit of education consists of the differences in each types contribution to social welfare, that is the differences with respect to the utility levels achieved, the net fiscal revenue raised as well as the impact on human capital formation.

According to the optimal policy rule (equation (17)), it is efficient to subsidize education as long as the contribution to social welfare of an additional high-skilled exceeds the contribution of a low-skilled worker ($\mu^L_{t,i} > 0$). Given the symmetric equilibrium outcomes derived above and accounting for intraregional redistribution policies (equation (29)), the net social benefit of education reduces to the difference in productivity between the two types,

$$\beta^{-1}\mu^L_{t-1,i} = \lambda_t \left[\left(\theta^H_{t,i} + \tau^H_{t,i}\right) - \left(\theta^L_{t,i} + \tau^L_{t,i}\right) \right] \underset{(21)}{=} \lambda_t \left(z^H - z^L\right) F'(Z_{t,i}) > 0. \quad (25)$$

The analysis reveals that the social benefit of turning a child into a high- instead of a low-skilled individual is strictly positive.This is due to the fact that high-skilled workers are more productive, and, hence, contribute more in terms of tax payments than low-skilled do. Since individual households do not take this positive social impact into account, they underinvest in education. This explains why the optimal educational subsidy is strictly positive with

$$\tau_{et,i} = \frac{\mu^L_{t,i}}{\lambda_{t,i}} h_{et,i} > 0. \quad (26)$$

Proposition 2 *To correct private underinvestment in education, a unitary state government subsidizes educational investments.*

Note that this result hinges on the assumption that the types are perfect substitutes in production with $z^H > z^L$. With complements, the net social benefit of education would be positive for low levels of human capital intensity and a subsidy would be optimal ($\mu^L_{t,i} > 0$). With a rising fraction of high-skilled workers, however, the net social benefit would decrease and eventually a tax on private educational investments would be efficient ($\mu^L_{t,i} < 0$). Recall, however, that the optimal decision rule for subsidizing education (equation (17)) is independent of assumptions concerning the production technology.

Symmetric Interregional Equilibrium

Since high-skilled workers are mobile across regions, the welfare optimum requires that the migration constraint is met. Hence, utility levels of mobile high-skilled workers have to equalize. As was already pointed out, with a non-binding migration constraint consumption is not only type-independent, but additionally identical across regions, $c^n_{t,i} = c_t$. This implies that utility resulting from consumption are the same, independent of the place of residence, $u\left(c^H_{t,i}\right) = u\left(c^H_t\right)$. Given the migration constraint, this entails that educational investments have to be equalized across regions,

$$u\left(c^H_t\right) + \upsilon(e^H_{t,i}) = u(c^H_t) + \upsilon(e^H_{t,j}) \Leftrightarrow \upsilon(e^H_{t,i}) = \upsilon(e^H_{t,j}) \Leftrightarrow e^H_{t,i} = e^H_t \Leftrightarrow e^n_{t,i} = e_t. \quad (27)$$

Hence, the solution to the optimization problem is entirely symmetric with $c^n_{t,i} = c_t$ and $e^n_{t,i} = e_t$.[14] Since the realized equilibrium values of consumption and investments in education are identical, it follows immediately from the optimal policy rule (equation (17)) that the optimal subsidy is type- and region independent with $\tau^n_{et,i} = \tau_{et}$. Furthermore, according to equation (18), identical educational spending across regions implies that the Lagrange-multiplier on the human capital formation constraint is independent of the respective region, $\mu^L_{t,i} = \mu^L_t$.

An efficient allocation of labor across regions requires that the contribution to social welfare of an additional high-skilled immigrant (equation ((22)) is the same for all regions. To determine the optimality condition for the allocation of mobile high-skilled

[14]Contrary, in the presence of migration costs, the equilibrium will no longer be symmetric as an interregional equalization of utility would violate the migration equilibrium.

workers, equate (22) across regions,

$$\lambda_t \left(\theta^H_{t,i} + \tau^H_{t,i} - \tau^H_{et,i} e^H_{t,i}\right) - \mu^L_{t,i} \left[1 - h(e^H_{t,i})\right] = \mu_t$$
$$= \lambda_t \left(\theta^H_{t,j} + \tau^H_{t,j} - \tau^H_{et,j} e^H_{t,j}\right) - \mu^L_{t,j} \left[1 - h(e^H_{t,j})\right] \qquad (28)$$

The left-hand side of equation (28) can be interpreted as the net social benefit of migration that consists of the contribution of a high-skilled immigrant in terms of net tax payments minus the social cost that arises if her child becomes low- instead of high-skilled ($\mu^L_{t,i}$), weighted by the probability to become low-skilled $(1 - h(e^H_{t,i}))$. The optimal allocation of high-skilled between regions i and j is attained when the net fiscal externality, that is the difference in the net social benefit between regions, vanishes. Evaluating in equilibrium reduces equation (28) to

$$\theta^H_{t,i} + \tau^H_{t,i} = \theta^H_{t,j} + \tau^H_{t,j}. \qquad (29)$$

In the welfare optimum, the central planner imposes federal lump-sum taxes such that the aggregate head tax of the high-skilled is independent of the region of residence.

Evaluating the households budget constraint in equilibrium, one can show that an efficient allocation of labor according to equation (29) requires that the marginal product of labor is equalized across regions,

$$z^H F'(Z_{t,i}) - \left(\tau^H_{t,i} + \theta^H_{t,i}\right) = c_t + (1 - \tau_{et}) e_t = z^H F'(Z_{t,j}) - \left(\tau^H_{t,j} + \theta^H_{t,j}\right) \qquad (30)$$
$$\Leftrightarrow \quad F'(Z_{t,i}) = F'(Z_{t,j}) \Leftrightarrow Z_{t,i} = Z_{t,j} \qquad (31)$$

Given identical aggregate head taxes, labor movements correspond to productivity differences. Thus, labor is allocated efficiently once productivity is equalized across regions, that is once production efficiency holds. This entails that the aggregate effective labor supply is the same for any region, $Z_{t,i} = Z_t$.

Proposition 3 *A benevolent unitary state government uses lump-sum taxes to equate the net social benefit of migration between regions. This ensures an efficient allocation of mobile labor (production efficiency).*

The starting point of the analysis was a non-binding migration constraint. In this case, the solution to the optimization problem is entirely symmetric with consumption, educational investments, as well as educational subsidies equalized both across regions and types. The central planner aims at redistributing income across types using type-specific head taxes. Contrary, lump-sum taxation is independent of the place of residence to ensure that the net social benefit of high-skilled migration is the same across regions. As a consequence, production efficiency holds and the aggregate effective labor supply is identical in all regions. Given this symmetric solution, the migration equilibrium is fulfilled as well. This entails that the solution that was derived ignoring the migration constraint coincides with the optimum of the fully constrained problem. Furthermore, the solution of the central planning problem is a unique global maximum. This is due to the fact that utility is strictly quasi-concave and the constraints are quasi-convex. If these conditions are fulfilled, and a local maximum exists, this local maximum is a unique global maximum.

The central planning solution reveals that it is optimal to subsidize education to overcome the underinvestment problem at the private level. Furthermore, the optimal subsidy will be type-independent and identical across regions. The following section focuses on the optimal fiscal policy a small open regions opts for. A local government disregards the fact, that the overall size of the high-skilled population in the federal economy is contingent on private investments in education. Hence, it might choose not to subsidize education, but try to attract high-skilled workers from other regions. Such free-riding of regional governments induces an educational underinvestment problem at the federal level, that can be corrected using an educational matching grant.

3.2. Decentralization of the Unitary State Optimum

Next, the optimal fiscal policy a local government chooses is derived, and policy instruments at the federal state level are deduced, that are needed to decentralize the unitary state optimum. Recall that the first-best solution requires that (1) consumption and educational spending are equalized across types, (2) educational investments are subsidized to prevent private underinvestment, and (3) production efficiency holds.

From point of view of a small open region, high-skilled households are in infinite supply. Thus, a regional government will not take into account the fact that the overall number of high-skilled available in the federal state is restricted, but only consider the migration constraint. The amount of low-skilled workers in any region i, however, is still contingent on local human capital formation. Therefore, a regional jurisdiction considers the local human capital formation constraint, and respects its own as well as the household's budget constraint, and the migration equilibrium constraint. The utility level attainable for mobile high-skilled workers in case of emigration is exogenous with $u(\overline{c}_t^H) + \upsilon(\overline{e}_t^H)$, where $\overline{c}_t^H$ and $\overline{e}_t^H$ denote the amount of consumption and educational investments realized outside of a small open region i. Furthermore, the regional government takes the federal head tax and the matching grant as given, since the central government is assumed to be the Stackelberg leader who moves first. Again, the household's first-order condition is used to eliminate τ_{et}^n from the optimization problem.

The local government maximizes the social welfare of it's resident population, choosing $\{\tau_t^n, c_t^n, e_t^n, N_t^n\}$,

$$\begin{aligned} \max \sum_{t=0}^{\infty} \beta^t \Big\{ & \sum_n N_t^n \left[u(c_t^n) + \upsilon(e_t^n)\right] \\ & + \sum_n \kappa_t^n \left[z^n F'(Z_t) - \tau_t^n - \theta_t^n - c_t^n - \frac{\upsilon_{et}^n}{u_{ct}^n} e_t^n \right] \\ & + \lambda_t \left[\sum_n N_t^n (\tau_t^n + \frac{\upsilon_{et}^n}{u_{ct}^n} e_t^n - (1 - \theta_{et}^n)\, e_t^n) + R(Z_t) \right] \\ & + \mu_t^L \left[N_{t+1}^L - \sum_n N_t^n \left[1 - h(e_t^n)\right] \right] \\ & + \varphi_t^H \left[u(\overline{c}_t^H) + \upsilon(\overline{e}_t^H) - u(c_t^H) - \upsilon(e_t^H) \right] \Big\} . \end{aligned} \tag{32}$$

Again, $\{\kappa_t^n, \lambda_t, \mu_t^L, \varphi_t^H\}$ denotes the set of Lagrange-multipliers. The first-order conditions are left to the Appendix.

Optimal Policy Rule

To derive the regionally optimal marginal rate of substitution between consumption and educational investments, use the first-order condition on τ_t^n to simplify the first-order conditions on consumption and investments for the low-skilled,

$$u_{ct}^L = \lambda_t \qquad \text{and} \qquad \upsilon_{et}^L + \mu_t^L h_{et}^L = \lambda_t \left(1 - \theta_{et}^L\right), \tag{33}$$

and for the high-skilled type,

$$\frac{\left(N_t^H - \varphi_t^H\right)}{N_t^H} u_{ct}^H = \lambda_t \qquad \text{and} \qquad \frac{\left(N_t^H - \varphi_t^H\right)}{N_t^H} \upsilon_{et}^H + \mu_t^L h_{et}^H = \lambda_t \left(1 - \theta_{et}^H\right). \tag{34}$$

Rearranging yields the optimal marginal rate of substitution from point of view of a small local jurisdiction,

$$\frac{\upsilon_{et}^n}{u_{ct}^n} = (1 - \theta_{et}^n) - \frac{\mu_t^L}{\lambda_t} h_{et}^n. \tag{35}$$

Equating the regionally optimal and the private marginal rate of substitution immediately reveals the optimal educational subsidy or tax a local government chooses,

$$(1 - \theta_{et}^n) - \frac{\mu_t^L}{\lambda_t} h_{et}^n = \frac{\upsilon_{et}^n}{u_{ct}^n} = (1 - \tau_{et}^n) \quad \Leftrightarrow \quad \tau_{et}^n - \theta_{et}^n = \frac{\mu_t^L}{\lambda_t} h_{et}^n. \tag{36}$$

The local jurisdiction, taking the federal grant as given, lowers the educational subsidy to meet the above optimality condition. On the other hand, the federal grant becomes redundant as long as the region chooses the educational subsidy according to the optimal policy rule (equation (17)). Obviously, the federal matching grant (θ_{et}^n) is a perfect substitute for regional education policies (τ_{et}^n). This is true, irrespective of whether the federal grant is paid directly to local governments or as a federal subsidy to households.[15] Hence, using the federal matching grant is equivalent to assigning the competencies on the field of education policy partly to the federal level.

To derive the fiscal policy a regional government chooses, suppose, there is no matching grant provided by the central government ($\theta_{et}^n = 0$). In this case, the local gov-

[15]One can easily reveal that both approaches result in the exact same optimization problem.

ernment imposes a tax rate according to the rule,

$$\tau_{et}^{n} = \frac{\mu_t^L}{\lambda_t} h_{et}^{n} = \frac{u_{ct}^{n} - \upsilon_{et}^{n}}{u_{ct}^{n}}. \tag{37}$$

Equation (37) reveals that in the absence of federal education policies, a benevolent government of a small region will stick to the optimal educational policy rule that supports the first-best, even though high-skilled workers are perfectly mobile.

Proposition 4 *Local jurisdictions abide to the optimal decision rule for subsidizing education. Thus, decentralized education policies are efficient, although high-skilled workers are perfectly mobile.*

This is due to the fact that the size of the immobile low-skilled workforce is still determined by regional investments in education. Thus, regional governments continue to respect the human capital formation constraint as well as the associated impact of education on future generation. Since parents do not take the intergenerational externality into account, local jurisdictions intend to correct this market failure and subsidize education, irrespective of the fact that high-skilled workers are perfectly mobile. Hence, introducing social mobility forces local governments to adhere to the optimal decision rule for education policies. Put differently, perfect mobility of high-skilled labor does not destroy local governments incentives to correct educational underinvestment. Only the size and sign of τ_{et}^{n}, that is determined by the level of consumption and education chosen by a regional government, might deviate from the optimal subsidy a unitary state government imposes. Thus, while regions stick to the optimal policy rule, it is not clear whether the implemented subsidy replicates first-best optimum, as the regional welfare optimum might entail different levels of consumption and educational investments than the unitary state outcome.

Moreover, the efficiency of decentralized education policies is not dependent on assumptions concerning the production technology or the migration dynamics, since it results solely from difference in the private and social marginal rate of substitution in consumption. While these assumptions affect the equilibrium obtained, this is true at both levels of government. For example, if migration costs are introduced, the solution will no longer be symmetric across countries as an interregional equalization of utility would violate the migration equilibrium. This holds for the unitary state optimization as it does at the local level. Yet, the efficiency of decentralized education policies remains unaffected, as local jurisdictions abide to the optimal decision rule for subsidizing education irrespective of the realized equilibrium.

Symmetric Intraregional Equilibrium

In the following, conditions that replicate the welfare maximum of the unitary state scenario are analyzed. To this aim, assume that regions are given enough resources by the central government to ensure that the migration constraint is not binding in the local optimum ($\varphi_t^H = 0$). In this case, optimal levels of consumption as well as educational investments in any region i are type-independent with $c_t^n = c_t$ and $e_t^n = e_t$ as can be revealed from the first-order conditions on consumption and educational investments (equations (33) and (34)). This, in turn, implies that utility levels are equalized in the optimum, and the local educational subsidy is identical for the different types, $\tau_{et}^n = \tau_{et}$. Following the above procedure, one can make use of the household's budget constraint to determine the optimal redistribution policy at the local level,

$$z^H F'(Z_t) - \tau_t^H - \theta_t^H = c_t + (1 - \tau_{et})e_t = z^L F'(Z_t) - \tau_t^L - \theta_t^L \tag{38}$$

$$\Leftrightarrow \left(z^H - z^L\right) F'(Z_t) = \left(\tau_t^H + \theta_t^H\right) - \left(\tau_t^L + \theta_t^L\right) > 0. \tag{39}$$

Since productivity across types differs, but consumption as well as educational investments are equalized in the regional optimum, the aggregate head tax on high-skilled households has to exceed the one on low-skilled. Given the federal head taxes already decided on by the central government who moves first, a local jurisdiction levies type-specific lump-sum taxes such that the productivity difference between types is fully taxed away. As in the unitary state optimum, a benevolent regional government redistributes income from high- to low-skilled workers to raise the utility of the low-skilled and maximize utilitarian welfare.

Proposition 5 *A benevolent local government redistributes income from mobile high-*

to immobile low-skilled labor to achieve a symmetric intraregional equilibrium.

Similarly to the case of educational tax instruments, federal and local head taxes are perfect substitutes to establish intraregional redistribution. As long as the central government levies type-independent head taxes ($\theta_t^H = \theta_t^L$), local jurisdictions will appropriate the whole income difference between the two types ($\tau_t^H > \tau_t^L$). Such a federal tax policy leaves local redistribution unaffected. In order to focus on local policies that replicate the first-best, the set of available federal tax instruments is restricted to type-independent head taxes, while the central government does not provide an educational grant to regions.

To figure out under which conditions the local subsidy or tax on education coincides with the educational subsidy that supports the first-best optimum (26), one has to determine the value of the shadow price μ_t^L associated with the probability of being high-skilled. Again, use the first-order conditions with respect to τ_t^n as well as the fact that $R'(Z_t) = -Z_t F''(Z_t)$ and $\frac{v_{et}^n}{u_{ct}^n} = (1 - \tau_{et}^n)$ to simplify the first-order conditions on high-skilled workers,

$$U_t^H(\cdot) + \lambda_t \left(\tau_t^H - \tau_{et}^H e_t^H\right) - \mu_t^L \left[1 - h(e_t^H)\right] = 0, \tag{40}$$

and on low-skilled workers,

$$0 = U_t^L(\cdot) + \lambda_t \left(\tau_t^L - \tau_{et}^L e_t^L\right) - \mu_t^L \left[1 - h(e_t^L)\right] + \beta^{-1} \mu_{t-1}^L. \tag{41}$$

Equating (40) and (41) yields the shadow price associated with the probability of being high-productive,

$$\beta^{-1} \mu_{t-1}^L = U_t^H(\cdot) - U_t^L(\cdot) + \lambda_t \left[\left(\tau_t^H - \tau_{et}^H e_t^H\right) - \left(\tau_t^L - \tau_{et}^L e_t^L\right)\right] \\ - \mu_t^L \left[h(e_t^H) - h(e_t^L)\right]. \tag{42}$$

This shadow price consists of the difference between future high- and low-skilled workers with respect to the utility levels, the tax payments they contribute to the regional budget, and their children's probability of becoming high-skilled. As long as the social benefit of an additional high-skilled exceeds the benefit of a low-skilled worker, the shadow price is positive and it is optimal to subsidize education. Evaluating in equilibrium and accounting for local redistribution (equation (39)) further reduces equation (42) to

$$\beta^{-1}\mu_{t-1}^{L} = \lambda_t\left(\tau_t^H - \tau_t^L\right) \underset{(39)}{=} \lambda_t\left[\left(z^H - z^L\right)F'(Z_t) - \theta_t^H - \theta_t^L\right]. \tag{43}$$

If the central government uses type-independent head taxes ($\theta_t^H = \theta_t^L$), the net social benefit of migration reduces to the difference in productivities between the two types,

$$\beta^{-1}\mu_{t-1}^{L} = \lambda_t\left(\tau_t^H - \tau_t^L\right) = \lambda_t\left(z^H - z^L\right)F'(Z_t) > 0. \tag{44}$$

The social benefit of turning a child into a high- instead of a low-skilled worker is strictly positive, although high-skilled labor is perfectly mobile across regions. This is due to the fact that high-skilled workers are assumed to be more productive than low-skilled workers and contribute more to social welfare in terms of higher tax payments. Put differently, the social benefit of education is entirely determined by differences in productivity among the two types, weighted by the marginal utility of consumption (λ_t). It follows, that the social benefit of eduction from point of view of a local government coincides with the social benefit of education as it is perceived by a federal government (equation (25)): Private underinvestment in education involves a welfare cost in terms of an increased number of less-productive low-skilled workers that is correctly accounted for at the local level. This explains why a small region continues to subsidize education and intends to correct private underinvestment,

$$\tau_{et} = \frac{\mu_t^L}{\lambda_t}h_{et} > 0. \tag{45}$$

Proposition 6 *Local jurisdictions subsidize education to correct private underinvestment.*

Yet, while a regional government will optimally subsidize education, it is not clear whether the first-best optimum can be replicated, since the regional government uses the available fiscal policy instruments only to equate consumption and investments in education across types, but not across regions. This implies that the net social benefit of migration need not be equalized between regions. Therefore, labor might be misallocated and production efficiency violated.

Symmetric Interregional Equilibrium

To attain the first-best resulting from the unitary state optimization, the central government has to insure that consumption and educational investments are not only equalized intraregionally, but also across regions. To derive the optimal federal redistribution policy, suppose that the central government only intends to equate consumption between regions,

$$\begin{aligned} z^H F'(Z_{t,i}) - \left(\tau^H_{t,i} + \theta^H_{t,i}\right) - (1 - \tau^H_{et,i})e^H_{t,i} = c^H_t \\ = z^H F'(Z_{t,j}) - \left(\tau^H_{t,j} + \theta^H_{t,j}\right) - (1 - \tau^H_{et,j})e^H_{t,j}. \quad (46) \end{aligned}$$

Then, the migration equilibrium constraint requires that educational investments are identical across regions as well,

$$u\left(c^H_t\right) + \upsilon(e^H_{t,i}) = u(c^H_t) + \upsilon(e^H_{t,j}) \Leftrightarrow \upsilon(e^H_{t,i}) = \upsilon(e^H_{t,j}) \Leftrightarrow e^H_{t,i} = e^H_t \Leftrightarrow e^n_{t,i} = e_t. \quad (47)$$

Given region-independent spending levels, the local decision rule for education policy (equation (37)) immediately yields that the subsidy provided by the local jurisdictions is independent of the respective region ($\tau^H_{et,i} = \tau^H_{et}$). Applying these results to equation (46) reveals that since households in any region face the same marginal rate of substitution, they consequently choose the same spending levels if they are confronted with an identical net income,

$$z^H F'(Z_{t,i}) - \left(\tau^H_{t,i} + \theta^H_{t,i}\right) = z^H F'(Z_{t,j}) - \left(\tau^H_{t,j} + \theta^H_{t,j}\right). \quad (48)$$

To establish a symmetric interregional equilibrium, the central government has to impose federal head taxes such that the net income of high-skilled households is equalized across regions.

To further determine the federal redistribution policy, the relation of local head taxes is derived. To this aim, use the equivalence of high-skilled utility that characterizes any migration equilibrium to equate the social benefit of qualified labor (equation (40)) across regions,

$$\lambda_{t,i}\left(\tau^H_{t,i}-\tau^H_{et,i}e^H_{t,i}\right)-\mu^L_{t,i}\left[1-h(e^H_{t,i})\right]=\lambda_{t,j}\left(\tau^H_{t,j}-\tau^H_{et,j}e^H_{t,j}\right)-\mu^L_{t,j}\left[1-h(e^H_{t,j})\right]. \quad (49)$$

Moreover, observe that according to the first-order condition on consumption (equations (34)) and the optimal local subsidy (37)), the symmetric equilibrium entails that the Lagrange-multipliers $\lambda_{t,i}$ and $\mu^L_{t,i}$ are identical across regions. Evaluating equation (49) in equilibrium then yields the result, that local head taxes on high-skilled labor are independent of the place of residence,

$$\tau^H_{t,i}=\tau^H_{t,j}. \quad (50)$$

Identical local head taxes on mobile labor have an important implication for the allocation of workers across regions. Since the Lagrange-multipliers $\lambda_{t,i}$ and $\mu^L_{t,i}$ are identical in all regions, and, hence, the social benefit of education as defined in equation (44) has to be the same, one can equate equation (44) across regions,

$$\lambda_t\left(z^H-z^L\right)F'(Z_{t,i}) \;=\; \beta^{-1}\mu^L_{t-1}=\lambda_t\left(z^H-z^L\right)F'(Z_{t,j}) \quad (51)$$

From equation (51) one can infer that production efficiency holds in the decentralized equilibrium, as the aggregate effective labor supply $Z_{t,i}$ will be equalized,

$$F'(Z_{t,i})=F'(Z_{t,j})\Leftrightarrow Z_{t,i}=Z_{t,j} \quad (52)$$

Applying these results to the net income differences as stated in equation (48) indicates that the federal head taxes on high-skilled workers are independent of the region of residence,

$$\theta^H_{t,i}=\theta^H_{t,j}. \quad (53)$$

Given local as well as federal head taxes that are identical for all regions, migration flows solely corresponds to productivity differences. Therefore, labor is allocated efficiently and production efficiency holds.

Since local governments adhere to the optimal decision rule for education policies and use lump-sum taxes to achieve a symmetric intraregional equilibrium, the central government can replicate the first-best by equating the net income of high-skilled households across regions using type-independent federal head taxes.[16] This guarantees that consumption and educational expenditures are equalized across regions. As a consequence the migration constraint is non-binding, which was the prerequisite for efficient regional redistribution policies. Furthermore, such a federal policy of interregional redistribution ensures that the aggregate effective labor supply is equalized across regions ($Z_{t,i} = Z_{t,j}$), and production efficiency holds.

4. Optimal Education Policies with High- and Low-Skilled Migration

The preceding analysis reveals that local jurisdictions abide by the optimal decision rule and use subsidies to correct private underinvestment in education even in the presence of high-skilled mobility. However, the result hinges on the assumption that low-skilled households are immobile. As the size of the low-skilled is determined by local investments in education, regional governments respect the human capital formation constraint, and efficiently correct private underinvestment in education. In the following, this assumption is relaxed, and mobility of both the high- and the low-skilled population is considered. Again, the unitary state optimum is derived as a benchmark case, and then decentralized policies of local governments are discussed.

4.1. Unitary State Optimum

Introducing mobility of high- and low-skilled households imposes an additional constraint on the optimization problem of the unitary state government, namely the migration equilibrium for the low-skilled type. This equilibrium condition states analogously to high-skilled migration, that in any migration equilibrium the utility levels of low-skilled workers have to be equalized across regions. In case of high- and low-

[16] To ensure that intra-regional redistribution is not violated, federal head taxes have to be type-independent.

skilled mobility, the unitary state government solves

$$
\begin{aligned}
\max \sum_{t=0}^{\infty} \beta^t \Big\{ & \sum_{i=0}^{\infty} \sum_{n} N_{t,i}^{n} \left[u\left(c_{t,i}^{n}\right) + \upsilon(e_{t,i}^{n}) \right] \\
& + \sum_{i=0}^{\infty} \sum_{n} \kappa_{t,i}^{n} \left[z^n F'(Z_{t,i}) - \tau_{t,i}^{n} - \theta_{t,i}^{n} - c_{t,i}^{n} - \frac{\upsilon_{et,i}^{n}}{u_{ct,i}^{n}} e_{t,i}^{n} \right] \\
& + \lambda_t \sum_{i=0}^{\infty} \sum_{n} N_{t,i}^{n} \left(\theta_{t,i}^{n} - \theta_{et,i}^{n} e_{t,i}^{n} \right) \\
& + \sum_{i=0}^{\infty} \lambda_{t,i} \left[\sum_{n} N_{t,i}^{n} (\tau_{t,i}^{n} + \frac{\upsilon_{et,i}^{n}}{u_{ct,i}^{n}} e_{t,i}^{n} - \left(1 - \theta_{et,i}^{n}\right) e_{t,i}^{n}) + R(Z_{t,i}) \right] \\
& + \mu_t^L \left[N_{t+1}^{L} - \sum_{i=0}^{\infty} \sum_{n} N_{t,i}^{n} \left[1 - h(e_{t,i}^{n})\right] \right] \\
& + \mu_t \left[\bar{N}_t - \sum_{i=0}^{\infty} \sum_{n} N_{t,i}^{n} \right] \\
& + \sum_{i \neq j} \sum_{n} \varphi_{t,i}^{n} \left[u(c_{t,j}^{n}) + \upsilon(e_{t,j}^{n}) - u\left(c_{t,i}^{n}\right) - \upsilon(e_{t,i}^{n}) \right] \Big\},
\end{aligned}
\tag{54}
$$

with the Lagrange-multipliers denoted by $\kappa_{t,i}^{n}, \lambda_t, \lambda_{t,i}, \mu_t^L, \mu_t$ and $\varphi_{t,i}^{n}$.

Note that in case of low-skilled mobility, the allocation both types of workers across regions is entirely determined by migration flows. Thus, the low-skilled workforce available in any region i is no longer contingent on local educational investments, only the aggregate federal endowment with low-skilled tomorrow depends on parental investments in education today. In contrast to the case of high-skilled mobility, the Lagrange-multiplier on the human capital formation constraint (μ_t^L) is therefore independent of the respective region i. The first-order conditions to the optimization problem are stated in the Appendix.

Following the above procedure, one can reveal that the unitary state government chooses the same decision rule for subsidizing education, irrespective of whether one or all types of labor are mobile across regions. Furthermore, the first-best optimum in the presence of high- and low-skilled migration coincides with the solution derived above for the case of high-skilled mobility. This is sensible, since the migration equilibrium constraints are not binding in the optimum. Hence, introducing mobility of high- and low-skilled labor does neither change the welfare maximum nor the optimal educational subsidy a unitary state government chooses.

Local jurisdictions perceive the supply of the mobile factor as infinite. In the presence of high- and low-skilled mobility, regional governments might refrain from subsidizing education, since both the number of high- and low-skilled workers allocated in the respective region are independent of regional investments in education. In the following section, optimal local fiscal policies are studied, and the need for federal education policies if high- and low-skilled workers can migrate across regions is discussed.

4.2. Decentralization of the Unitary State Optimum

Perfect mobility of both high- and low-skilled workers changes the optimization constraints local governments face substantially. Since from point of view of a small open region the mobile factor is in perfectly elastic supply, a regional government perceives the size of the respective type of worker as only dependent on migration incentives. It therefore disregards the fact that the evolution of the different population groups is contingent on regional investments in education, and no longer respect the human capital formation constraint. Hence, a local government chooses the regional tax rate $\{\tau_t^n\}$ and the allocations $\{c_t^n, e_t^n, N_t^n\}$ to solve the following problem,

$$
\begin{aligned}
&\max \sum_{t=0}^{\infty} \beta^t \Big\{ \sum_n N_t^n \left[u(c_t^n) + \upsilon(e_t^n) \right] \\
&+ \sum_n \kappa_t^n \left[z^n F'(Z_t) - \tau_t^n - \theta_t^n - c_t^n - \frac{\upsilon_{et}^n}{u_{ct}^n} e_t^n \right] \\
&+ \lambda_t \left[\sum_n N_t^n (\tau_t^n + \frac{\upsilon_{et}^n}{u_{ct}^n} e_t^n - (1 - \theta_{et}^n)\, e_t^n) + R(Z_t) \right] \\
&+ \sum_n \varphi_t^n \left[u(\overline{c}_t^n) + \upsilon(\overline{e}_t^n) - u(c_t^n) + \upsilon(e_t^n) \right] \Big\},
\end{aligned} \tag{55}
$$

where κ_t^n, λ_t and φ_t^n denote the Lagrange-multipliers. The utility level attainable for mobile workers outside region i is given with $u(\overline{c}_t^n) + \upsilon(\overline{e}_t^n)$. The first-order conditions are presented in the Appendix.

Optimal Policy Rule

Again, one can simplify the first-order conditions on consumption and investments using the first-order conditions on τ_t^n, c_t^n and e_t^n,

$$\frac{(N_t^n - \varphi_t^n)}{N_t^n} u_{ct}^n = \lambda_t \qquad \text{and} \qquad \frac{(N_t^n - \varphi_t^n)}{N_t^n} \upsilon_{et}^n = \lambda_t \left(1 - \theta_{et}^n\right). \tag{56}$$

Using (56), region i's marginal rate of substitution between consumption and educational investments can be derived,

$$\frac{\upsilon_e^n}{u_c^n} = \left(1 - \theta_{et}^n\right). \tag{57}$$

Equating the regionally optimal and the private marginal rate of substitution immediately reveals the optimal educational subsidy a local government chooses,

$$(1 - \theta_{et}^n) = \frac{\upsilon_e^n}{u_c^n} = (1 - \tau_{et}^n) \quad \Leftrightarrow \quad \tau_{et}^n = \theta_{et}^n. \tag{58}$$

The optimal regional policy in presence of high- and low-skilled mobility is not to promote education at all: Local governments pass on the federal grant θ_{et}^n to households, but refrain from subsidizing education themeselves.[17] If there is no federal grant ($\theta_{et}^n = 0$), a local jurisdiction does not subsidize education ($\tau_{et}^n = 0$).

The inefficiency of local education policies can be explained by the fact that from point of view of a small open region, the available size of both the high- and the low-skilled workforce is entirely determined by migration flows. Previous local investments in education have no impact on the composition of the future workforce. Hence, educational underinvestment is not corrected for at the local level. This implies an inefficient evolution of the composition of the federal population, since private underinvestment persists.

Since local governments disregard this inefficiency, the central government has to correct the intergenerational externality. According to equation (19), the central

[17] Again, this finding is robust, irrespective of whether the grant is paid to regional governments or directly to households.

government intends to set $\tau_{et} = \frac{\mu_t^L}{\lambda_t} h_e$ to establish the first-best optimum. The local decision rule (equation (58)) then requires a federal matching grant of $\theta_{et} = \frac{\mu_t^L}{\lambda_t} h_e$. As the local government passes on the optimal federal grant to it's residents, it will be forced to implement the optimal educational subsidy. Therefore, the federal matching grant is an efficient policy tool to decentralize the first-best optimum. Still, since the federal grant replaces local education policies, the first-best optimum can effectively only be established by assigning education policies to the federal level.

Proposition 7 *If high- and low-skilled workers are mobile across regions, decentralized education policies are inefficient as incentives of local jurisdictions to subsidize education vanish completely. To correct educational underinvestment, education policies have to be assigned effectively to the federal or supranational level.*

As in the case of high-skilled mobility only, the result concerning the inefficiency of decentralized education policies carries over to a scenario in which migration costs are imposed. Only if mobile households face a certain probability to emigrate, in which case part of the workforce remains immobile and, thus, affected by local education policies with a positive probability, local governments continue to respect the human capital formation constraint, and choose an efficient education policy.

Symmetric Intraregional Equilibrium

In the following, the optimal local policy that supports the first-best optimum is deduced. To this aim, it is assumed that regions are given enough resources by the central government to ensure that the migration constraint is not binding in the local optimum ($\varphi_t^H = 0$). According to the first-order conditions on consumption and education (equation (56)), this entails that optimal levels of consumption as well as educational investments in any region i are identical irrespective of the type of household, $c_t^n = c_t$ and $e_{et}^n = c_t$. Recall that intraregional symmetry in consumption and education guarantees that utility levels between types are identical in the optimum, and utilitarian welfare is maximized.

Since the optimal federal grant is type-independent, one can equate the household's budget constraint and derive the optimal redistribution policy,

$$\begin{aligned} z^H F'(Z_t) - \tau_t^H - \theta_t^L &= c_t + (1-\theta_{et})e_t = z^L F'(Z_t) - \tau_t^L - \theta_t^L & (59)\\ \Leftrightarrow &\ \left(z^H - z^L\right) F'(Z_t) = \left(\tau_t^H + \theta_t^H\right) - \left(\tau_t^L + \theta_t^L\right) > 0. & (60)\end{aligned}$$

Intraregional equalization of consumption and educational investments requires that the net income of the different types is equalized. To achieve this, the aggregate head taxes of the high-skilled workers have to exceed the taxes on low-skilled.

Local head taxes, however, are independent of the respective type in a decentralized optimum. To see this, use the first-order condition on τ_t^n and the fact that $R'(Z_t) = -Z_t F''(Z_t)$, and equate the first-order condition with respect to the population groups,

$$\begin{aligned} U_t^H(\cdot) + \lambda_t \tau_t^H &= 0 = U_t^L(\cdot) + \lambda_t \tau_t^L & (61)\\ \Leftrightarrow &\ \tau_t^H = \tau_t^L = \tau_t. & (62)\end{aligned}$$

Equation (61) can be interpreted as the net social benefit of an additional high- or low-skilled worker from point of view of a small open region. Immigration in case of high- and low-skilled mobility does no longer involve any educational benefits, but only contributions in terms of the utility and tax payments of the respective type. If the central government ensures that the migration constraint is not binding, consumption and education are identical for both types. This implies identical utility levels, and explains why local governments have no interest in redistributing income between the different types: High- and low-skilled workers are perfectly homogenous from point of view of a local, welfare maximizing government. The optimal policy at the local level is therefore to redistribute profits evenly among the two types, $\tau_t^n = \tau_t$.[18]

Given type-independent local head taxes, the central government has to achieve the symmetric intraregional equilibrium. One can use the local distribution policy to

[18]Note, that in the absence of a local educational subsidy, the only purpose of local fiscal policy is to distribute profits. Local lump-sum taxes are therefore negative.

rewrite equation (60), and deduce the optimal federal head taxes,

$$\left(z^H - z^L\right) F'(Z_t) = \theta_t^H - \theta_t^L > 0. \tag{63}$$

The central government levies type-specific federal head taxes to ensure that expenditure levels are equalized across types. This guarantees that a symmetric intraregional equilibrium is attained. However, it is not clear whether this replicates the first-best optimum that additionally entails interregional redistribution.

Symmetric Interregional Equilibrium

In the first-best equilibrium resulting from the unitary state optimization, consumption and educational investments are equalized interregionally. Yet, following the above procedure, it is assumed that the central government only intends to equate consumption between regions,

$$\begin{aligned} z^H F'(Z_{t,i}) - \left(\tau_{t,i}^H + \theta_{t,i}^H\right) - (1-\tau_{et,i}^H)e_{t,i}^H = c_t^H \\ = z^H F'(Z_{t,j}) - \left(\tau_{t,j}^H + \theta_{t,j}^H\right) - (1-\tau_{et,j}^H)e_{t,j}^H. \end{aligned} \tag{64}$$

Again, the migration equilibrium constraint requires that educational investments are identical across regions as well,

$$u\left(c_t^H\right) + \upsilon(e_{t,i}^H) = u(c_t^H) + \upsilon(e_{t,j}^H) \Leftrightarrow \upsilon(e_{t,i}^H) = \upsilon(e_{t,j}^H) \Leftrightarrow e_{t,i}^H = e_t^H \Leftrightarrow e_{t,i}^n = e_t. \tag{65}$$

Given identical spending levels across regions, and accounting for the fact that the central government provides the optimal type- and region-independent educational grant, one can simplify equation (64),

$$z^H F'(Z_{t,i}) - \left(\tau_{t,i}^H + \theta_{t,i}^H\right) = z^H F'(Z_{t,j}) - \left(\tau_{t,j}^H + \theta_{t,j}^H\right). \tag{66}$$

To establish a symmetric interregional equilibrium, the central government has to impose federal head taxes such that the net income of high-skilled households is equalized across regions.

Yet, it is not proven that production efficiency holds in the optimum: Observe that according to the first-order condition on consumption (equation (56)) and the optimal grant provided (equation (19)), the symmetric equilibrium entails that the Lagrange-multipliers $\lambda_{t,i}$ and $\mu^L_{t,i}$ are identical across regions. This implies that the social benefit of education as accounted for by the central government (equation (25)) is the same in all regions,

$$\lambda_{t,i}\left(z^H - z^L\right)F'(Z_{t,i}) \quad = \quad \beta^{-1}\mu^L_{t-1,i} = \lambda_{t,j}\left(z^H - z^L\right)F'(Z_{t,j}) \tag{67}$$

$$\Leftrightarrow \quad F'(Z_{t,i}) = F'(Z_{t,j}). \tag{68}$$

The analysis reveals that federal policies can replicate the first-best optimum in which production efficiency holds. This, however, requires that the central government provides the optimal educational grant, and - given type-independent local lump-sum taxation - imposes a type-specific head taxes to redistribute income. Furthermore, federal lump-sum taxation has to be used in a way to ensure that the aggregate head tax on mobile high-skilled labor is identical across regions. To see this, apply the result of production efficiency to equation (66),

$$\tau^H_{t,i} + \theta^H_{t,i} = \tau^H_{t,j} + \theta^H_{t,j}. \tag{69}$$

Such interregional redistribution ensures the efficient allocation of labor according to equation (29) as migration flows are driven by productivity differences alone, and the optimum is characterized by a situation in which production efficiency holds.

Instead of supporting education by subsidizing private educational investments, local governments refrain from using corrective tax instruments in the presence of high- and low-skilled mobility. Hence, private underinvestment in education persists unless it is accounted for at the federal or supranational level. A small region ignores this inefficiency as it regards the available amount of high- and low-skilled workers as infinite (perfect elasticity of supply). To correct the underinvestment problem, the central government can use a federal matching grant. However, such a grant is a perfect substitute for subsidies on education paid directly to households. Therefore, decentralization in a setting with high- and low-skilled mobility requires that the fiscal authority on the field of education policy has to be effectively assigned to the federal or supra-national level. Furthermore, federal head taxes have to be imposed to establish the intra- as well as interregionally symmetric equilibrium.

5. Conclusion

In the view of increased labor mobility, local governments face the option to free-ride on other regions' education policies instead of correcting educational underinvestment themselves. Still, we observe that education is financed substantially by local governments. The present paper provides an explanation for this phenomenon based on the effect of social mobility across different skill-types. An intergenerational externality is introduced that leads to a situation of underinvestment in education and, hence, provides a justification for a corrective educational subsidy from a federal perspective. Against this background, the efficiency of decentralized decision in the presence of perfect, high-skilled mobility making is discussed.

It is shown, that small regions abide to the optimal decision rule for subsidizing education although high-skilled workers are perfectly mobile across regions. This is surprising, since private underinvestment in human capital can, from point of view of a small region, fully be compensated by high-skilled immigration from other regions. However, in a setup with social mobility, underinvestment in education not only affects the size of the high-skilled workers, but additionally determines the number of immobile low-skilled workers in the respective region. Consequently, the incentive to correct educational underinvestment prevails although the high-skilled workforce is perfectly mobile. Thus, decentralized education policies remain to be efficient. Social mobility therefore provides an explanation of why local governments continue to subsidize education even though the high-skilled can migrate.

However, the efficiency of decentralized education policies does not persist in a scenario in which both types of labor are mobile. In such a setup, local governments lose any incentive to subsidize education since both the size of the high- and low-skilled workforce are solely determined by migration flows. This explains why decentralized policies turn out to be inefficient. To prevent underinvestment in this case, education policies have to be assigned to the federal or supra-national level.

Though two polar cases - perfect high-skilled mobility with immobile low-skilled workers versus perfect high- and low-skilled mobility - are considered, the results carry over to more realistic migration scenarios. Empirical studies suggest that the mobility of qualified labor is relatively high compared to low-skilled workers (see, e.g. Carrington and Detragiache (1998)). Moreover, one can conjecture that at least for some workers barriers to migration are prohibitively high. Hence, part of the low-skilled will most likely be immobile. As long as only a fraction of the low-skilled workforce remains immobile, the findings derived in the present paper indicate that decentralized education policies remain to be efficient. The results of the present paper, therefore, imply that local governments facing factor mobility do not necessarily lose their scope to correct market failures, at least not with respect to underinvestment in education.

A. Appendix

Unitary State Optimum: Migration of High-Skilled

The first-order conditions for the central planning problem are as follows,

$$\frac{\partial}{\partial \tau^n_{t,i}} : \kappa^n_{t,i} = \lambda_{t,i} N^n_{t,i} \tag{A.1}$$

$$\frac{\partial}{\partial \theta^n_{t,i}} : \kappa^n_{t,i} = \lambda_t N^n_{t,i} \tag{A.2}$$

$$\frac{\partial}{\partial \theta^n_{et,i}} : \lambda_t N^n_{t,i} e^n_{t,i} = \lambda_{t,i} N^n_{t,i} e^n_{t,i} \quad \Leftrightarrow \quad \lambda_t = \lambda_{t,i} \tag{A.3}$$

$$\frac{\partial}{\partial c^L_{t,i}} : u^L_{ct,i} N^L_{t,i} = \kappa^L_{t,i} \left(1 - \frac{\upsilon^L_{et,i}}{\left(u^L_{cct,i}\right)^2} e^L_{t,i} \right) + \lambda_{t,i} N^L_{t,i} \frac{\upsilon^L_{et,i}}{\left(u^L_{cct,i}\right)^2} e^L_{t,i} \tag{A.4}$$

$$\frac{\partial}{\partial c^H_{t,i}} : \left(N^H_{t,i} - \varphi^H_{t,i}\right) u^H_{ct,i} = \kappa^H_{t,i} \left(1 - \frac{\upsilon^H_{et,i}}{\left(u^H_{cct,i}\right)^2} e^H_{t,i} \right) + \lambda_{t,i} N^H_{t,i} \frac{\upsilon^H_{et,i}}{\left(u^H_{cct,i}\right)^2} e^H_{t,i} \tag{A.5}$$

$$\begin{aligned} \frac{\partial}{\partial e^L_{t,i}} &: N^L_{t,i} \upsilon^L_{et,i} - \kappa^L_{t,i} \left(\frac{\upsilon^L_{eet,i}}{u^L_{ct,i}} e^L_{t,i} + \frac{\upsilon^L_{et,i}}{u^L_{ct,i}} \right) - \lambda_t N^L_{t,i} \theta^L_{et,i} \\ &+ \lambda_{t,i} N^L_{t,i} \left(\frac{\upsilon^L_{eet,i}}{u^L_{ct,i}} e^L_{t,i} + \frac{\upsilon^L_{et,i}}{u^L_{ct,i}} - \left(1 - \theta^L_{et,i}\right) \right) + \mu^L_{t,i} N^L_{t,i} h^L_{et,i} = 0 \end{aligned} \tag{A.6}$$

$$\begin{aligned} \frac{\partial}{\partial e^H_{t}} &: \left(N^H_{t,i} - \varphi^H_{t,i}\right) \upsilon^H_{et,i} - \kappa^H_{t,i} \left(\frac{\upsilon^H_{eet,i}}{u^H_{ct,i}} e^H_{t,i} + \frac{\upsilon^H_{et,i}}{u^H_{ct,i}} \right) - \lambda_t N^H_{t,i} \theta^H_{et,i} \\ &+ \lambda_{t,i} N^H_{t,i} \left(\frac{\upsilon^H_{eet,i}}{u^H_{ct,i}} e^H_{t,i} + \frac{\upsilon^H_{et,i}}{u^H_{ct,i}} - \left(1 - \theta^H_{et,i}\right) \right) + \mu^L_{t,i} N^H_{t,i} h^H_{et,i} = 0 \end{aligned} \tag{A.7}$$

$$\begin{aligned} \frac{\partial}{\partial N^L_{t,i}} &: U^L_{t,i}(\cdot) + \left(\kappa^L_{t,i} z^L + \kappa^H_{t,i} z^H\right) z^L F''(Z_{t,i}) + \lambda_t \left[\theta^L_{t,i} - \theta^L_{et,i} e^L_{t,i}\right] + \lambda_{t,i} (\tau^L_{t,i} + \frac{\upsilon^L_{et,i}}{u^L_{ct,i}} e^L_{t,i}) \\ &- \lambda_{t,i} \left[\left(1 - \theta^L_{et,i}\right) e^L_{t,i} - z^L R'(Z_{t,i})\right] - \mu^L_{t,i} \left[1 - h(e^L_{t,i})\right] + \beta^{-1} \mu^L_{t-1,i} = \mu_t \end{aligned} \tag{A.8}$$

$$\begin{aligned} \frac{\partial}{\partial N^H_{t,i}} &: U^H_{t,i}(\cdot) + \left(\kappa^L_{t,i} z^L + \kappa^H_{t,i} z^H\right) z^H F''(Z_{t,i}) + \lambda_t \left[\theta^H_{t,i} - \theta^H_{et,i} e^H_{t,i}\right] + \lambda_{t,i} (\tau^H_{t,i} + \frac{\upsilon^H_{et,i}}{u^H_{ct,i}} e^H_{t,i}) \\ &- \lambda_{t,i} \left[\left(1 - \theta^H_{et,i}\right) e^H_{t,i} - z^H R'(Z_{t,i})\right] - \mu^L_{t,i} \left[1 - h(e^H_{t,i})\right] = \mu_t \end{aligned} \tag{A.9}$$

To gain symmetric first-order conditions, define the multiplier on the migration constraint for the arbitrarily chosen reference region j as $\varphi^H_{t,j} = -\varphi^H_{t,i}$.

Local Optimum: Migration of High-Skilled

The first-order conditions for the local optimization problem are

$$\frac{\partial}{\partial \tau_t^n} : \kappa_t^n = \lambda_t N_t^n \tag{A.10}$$

$$\frac{\partial}{\partial c_{t,i}^L} : N_t^L u_{ct}^L = \kappa_t^L \left(1 - \frac{\upsilon_{et}^L}{\left(u_{cct}^L\right)^2} e_t^L\right) + \lambda_t N_t^L \frac{\upsilon_{et}^L}{\left(u_{cct}^L\right)^2} e_t^L \tag{A.11}$$

$$\frac{\partial}{\partial c_{t,i}^H} : \left(N_t^H - \varphi_t^H\right) u_c^H = \kappa_t^H \left(1 - \frac{\upsilon_{et}^H}{\left(u_{cct}^H\right)^2} e_t^H\right) + \lambda_t N_t^H \frac{\upsilon_{et}^H}{\left(u_{cct}^H\right)^2} e_t^H \tag{A.12}$$

$$\begin{aligned}\frac{\partial}{\partial e_t^L} : & N_t^L \upsilon_{et}^L - \kappa_{t,i}^L \left(\frac{\upsilon_{eet}^L}{u_{ct}^L} e_t^L + \frac{\upsilon_{et}^L}{u_{ct}^L}\right) + \lambda_t N_t^L \left(\frac{\upsilon_{eet}^L}{u_{ct}^L} e_t^L + \frac{\upsilon_{et}^L}{u_{ct}^L} - \left(1 - \theta_{et}^L\right)\right) \\ & + \mu_t^L h_{et}^L N_t^L = 0\end{aligned} \tag{A.13}$$

$$\begin{aligned}\frac{\partial}{\partial e_t^H} : & \left(N_t^H - \varphi_t^H\right) \upsilon_{et}^H - \kappa_{t,i}^H \left(\frac{\upsilon_{eet}^H}{u_{ct}^H} e_t^H + \frac{\upsilon_{et}^H}{u_{ct}^H}\right) \\ & + \lambda_t N_t^H \left(\frac{\upsilon_{eet}^H}{u_{ct}^H} e_t^H + \frac{\upsilon_{et}^H}{u_{ct}^H} - \left(1 - \theta_{et}^H\right)\right) + \mu_t^L N_t^H h_e^H = 0\end{aligned} \tag{A.14}$$

$$\begin{aligned}\frac{\partial}{\partial N_t^L} : & U_t^L + \left(\kappa_t^L z^L + \kappa_t^H z^H\right) z^L F''(Z_t) + \lambda_t \left[\tau_t^L + \frac{\upsilon_{et}^L}{u_{ct}^L} e_t^L - \left(1 - \theta_{et}^L\right) e_t^L + z^L R'(Z_t)\right] \\ & - \mu_t^L \left[1 - h(e_t^H)\right] + \beta^{-1} \mu_{t-1}^L = 0\end{aligned} \tag{A.15}$$

$$\begin{aligned}\frac{\partial}{\partial N_t^H} : & U_t^H + \left(\kappa_t^L z^L + \kappa_t^H z^H\right) z^H F''(Z_{t,i}) \\ & + \lambda_t \left[\tau_t^H + \frac{\upsilon_{et}^H}{u_{ct}^H} e_t^H - \left(1 - \theta_{et}^H\right) e_t^H + z^H R'(Z_t)\right] - \mu_t^L \left[1 - h(e_t^H)\right] = 0.\end{aligned} \tag{A.16}$$

Unitary State Optimum: Migration of High- and Low-Skilled

Optimizations of the unitary state government in case of high- and low-skilled mobility yields the following first-order conditions

$$\frac{\partial}{\partial \tau_{t,i}^n} : \kappa_{t,i}^n = \lambda_{t,i} N_{t,i}^n \tag{A.17}$$

$$\frac{\partial}{\partial \theta_{t,i}^n} : \kappa_{t,i}^n = \lambda_t N_{t,i}^n \tag{A.18}$$

$$\frac{\partial}{\partial \theta_{et,i}^n} : \lambda_t N_{t,i}^n e_{t,i}^n = \lambda_{t,i} N_{t,i}^n e_{t,i}^n \quad \Leftrightarrow \quad \lambda_t = \lambda_{t,i} \tag{A.19}$$

$$\frac{\partial}{\partial c^n_{t,i}} : \left(N^n_{t,i} - \varphi^n_{t,i}\right) u^n_{ct,i} = \kappa^n_{t,i} \left(1 - \frac{\upsilon^n_{et,i}}{\left(u^n_{cct,i}\right)^2} e^n_{t,i}\right) + \lambda_{t,i} N^n_{t,i} \frac{\upsilon^n_{et,i}}{\left(u^n_{cct,i}\right)^2} e^n_{t,i} \tag{A.20}$$

$$\begin{aligned}\frac{\partial}{\partial e^n_t} &: \left(N^n_{t,i} - \varphi^n_{t,j}\right) \upsilon^n_{et,i} - \kappa^n_{t,i} \left(\frac{\upsilon^n_{eet,i}}{u^n_{ct,i}} e^n_{t,i} + \frac{\upsilon^n_{et,i}}{u^n_{ct,i}}\right) - \lambda_t N^n_{t,i} \theta^n_{et,i} \\ &+ \lambda_{t,i} N^n_{t,i} \left(\frac{\upsilon^n_{eet,i}}{u^n_{ct,i}} e^n_{t,i} + \frac{\upsilon^n_{et,i}}{u^n_{ct,i}} - \left(1 - \theta^n_{et,i}\right)\right) + \mu^L_t N^n_{t,i} h^n_{et,i} = 0\end{aligned} \tag{A.21}$$

$$\begin{aligned}\frac{\partial}{\partial N^L_{t,i}} &: U^L_{t,i}(\cdot) + \left(\kappa^L_{t,i} z^L + \kappa^H_{t,i} z^H\right) z^L F''(Z_{t,i}) + \lambda_t \left[\theta^L_{t,i} - \theta^L_{et,i} e^L_{t,i}\right] + \lambda_{t,i}(\tau^L_{t,i} + \frac{\upsilon^L_{et,i}}{u^L_{ct,i}} e^L_{t,i}) \\ &- \lambda_{t,i} \left[\left(1 - \theta^L_{et,i}\right) e^L_{t,i} - z^L R'(Z_{t,i})\right] \\ &- \mu^L_t \left[1 - h(e^L_{t,i})\right] + \beta^{-1} \mu^L_{t-1} = \mu_t\end{aligned} \tag{A.22}$$

$$\begin{aligned}\frac{\partial}{\partial N^H_{t,i}} &: U^H_{t,i}(\cdot) + \left(\kappa^L_{t,i} z^L + \kappa^H_{t,i} z^H\right) z^H F''(Z_{t,i}) + \lambda_t \left[\theta^H_{t,i} - \theta^H_{et,i} e^H_{t,i}\right] + \lambda_{t,i}(\tau^H_{t,i} + \frac{\upsilon^H_{et,i}}{u^H_{ct,i}} e^H_{t,i}) \\ &- \lambda_{t,i} \left[\left(1 - \theta^H_{et,i}\right) e^H_{t,i} - z^H R'(Z_{t,i})\right] - \mu^L_{t,i} \left[1 - h(e^H_{t,i})\right] = \mu_t.\end{aligned} \tag{A.23}$$

Local Optimum: Migration of High- and Low-Skilled

The first-order conditions of the regional optimization when high- and low-skilled workers are mobile, are as follows,

$$\frac{\partial}{\partial \tau^n_t} : \kappa^n_t = \lambda_t N^n_t \tag{A.24}$$

$$\frac{\partial}{\partial c^n_t} : \left(N^n_t - \varphi^n_t\right) u^n_{ct} = \kappa^n_t \left(1 - \frac{\upsilon^n_{et}}{\left(u^n_{cct}\right)^2} e^n_t\right) + \lambda_t N^n_t \frac{\upsilon^n_{et}}{\left(u^n_{cct}\right)^2} e^n_t \tag{A.25}$$

$$\begin{aligned}\frac{\partial}{\partial e^n_t} &: \left(N^n_t - \varphi^n_t\right) \upsilon^n_{et} - \kappa^n_t \left(\frac{\upsilon^n_{eet}}{u^n_{ct}} e^n_t + \frac{\upsilon^n_{et}}{u^n_{ct}}\right) \\ &- \lambda_t N^n_t \left(\frac{\upsilon^n_{eet}}{u^n_{ct}} e^n_t + \frac{\upsilon^n_{et}}{u^n_{ct}} - \left(1 - \theta^n_{et}\right)\right) = 0\end{aligned} \tag{A.26}$$

$$\begin{aligned}\frac{\partial}{\partial N^n_t} &: U^n_t(\cdot) + \left(\kappa^L_t z^L + \kappa^H_t z^H\right) z^n F''(Z_t) \\ &+ \lambda_t \left[\tau^n_t + \frac{\upsilon^n_{et}}{u^n_{ct}} e^n_t - \left(1 - \theta^n_{et}\right) e^n_t + z^n R'(Z_t)\right] = 0.\end{aligned} \tag{A.27}$$

References

Andersson, Frederik and Kai A. Konrad (2003a): Globalization and Risky Human-Capital Investment, *International Tax and Public Finance*, Vol. 10, No. 3, pp. 211–228.

Andersson, Frederik and Kai A. Konrad (2006): Taxation and Education Investment in the Tertiary Sector, in: Torben Andersen and Per Molander (Eds.), *Alternatives for Welfare Policy*, Chapter 13, Cambridge University Press.

Andersson, Fredrik and Kai A. Konrad (2003b): Human Capital Investment and Globalization in Extortionary States, *Journal of Public Economics*, Vol. 87, pp. 1539–1555.

Beine, Michel, Frederic Docquier and Hillel Rapoport (2001): Brain drain and economic growth: theory and evidence, *Journal of Development Economics*, Vol. 64, No. 1, pp. 275–289.

Boadway, Robin (1982): On the Method of Taxation and the Provision of Local Public Goods: Comment, *American Economic Review*, Vol. 72, No. 4, pp. 846–851.

Boadway, Robin, Katherine Cuff and Maurice Marchand (2003): Equalization and the Dezentralization of Revenue-Raising in a Federation, *Journal of Public Economic Theory*, Vol. 5, No. 2, pp. 201–228.

Boadway, Robin, Nicolas Marceau and Maurice Marchand (1996): Investment in Education and Time Inconsistency of Redistributive Tax Policy, *Economica*, Vol. 63, pp. 171–189.

Carneiro, Pedro and James J. Heckman (2002): The evidence on credit constraints in post-secondary schooling, *The Economic Journal*, Vol. 112, pp. 705–734.

Carrington, William J. and Enrica Detragiache (1998): How big is the brain drain?, IMF Working Paper.

Commander, Simon, Mari Kangasniemi and L. Alan Winters (2003): The Brain

Drain: Curse or Boon?, IZA Discussion Paper.

Cremer, Helmuth and Pierre Pestieau (2004): Intergenerational Transfer of Human Capital and Optimal Education Policy, CEPR Discussion Paper.

Flatters, Frank, Vernon Henderson and Peter Mieszkowski (1974): Public goods, efficiency, and regional fiscal equalization, *Journal of Public Economics*, Vol. 3, No. 2, pp. 99–112.

Justman, Moshe and Jacques-Francois Thisse (1997): Implications of the mobility of skilled labor for local public funding of higher education, *Economics Letters*, Vol. 55, No. 3, pp. 409–412.

Justman, Moshe and Jacques-Francois Thisse (2000): Local Public Funding of Higher Education when Skilled Labor is Imperfectly Mobile, *International Tax and Public Finance*, Vol. 7, No. 3, pp. 247–258.

Kydland, Finn E. and Edward C. Prescott (1977): Rules rather than discretion: The inconsistency of Optimal Plans, *Journal of Political Economy*, Vol. 85, No. 3, pp. 473–491.

Mas-Colell, Andreu, Michael Dennis Whinston and Jerry R. Green (1995): *Microeconomic theory*, New York NY: Oxford Univ. Press.

Mountford, Andrew (1997): Can a brain drain be good for growth in the source economy?, *Journal of Development Economics*, Vol. 53, No. 2, pp. 287–303.

Poutvaara, Panu (2000): Education, Mobility of Labour and Tax Competition, *International Tax and Public Finance*, Vol. 7, No. 6, pp. 699–719.

Poutvaara, Panu (2001): Alternative tax constitutions and risky education in a federation, *Regional Science and Urban Economics*, Vol. 31, No. 2-3, pp. 355–377.

Poutvaara, Panu and Vesa Kanniainen (2000): Why Invest in Your Neighbor? Social Contract on Educational Investment, *International Tax and Public Finance*, Vol. 7, No. 4-5, pp. 864–888.

Restuccia, Diego and Carlos Urrutia (2004): Intergenerational Persistence of Earnings: The Role of Early and College Education, *American Economic Review*, Vol. 94, No. 5, pp. 1354–1378.

Rey, E. Del (2001): Economic Integration and Public Provision of Education, *Empirica*, Vol. 28, pp. 203–218(16).

Sinn, Hans-Werner (1997): The selection principle and market failure in systems competition, *Journal of Public Economics*, Vol. 66, pp. 247–274.

Stark, Oded (2004): Rethinking the Brain Drain, *World Development*, Vol. 32, No. 1, pp. 15–22, hier: ZEF-Discussion Papers on Development Policy No. 17, 2003.

Stark, Oded, Christian Helmenstein and Alexia Prskawetz (1997): A brain gain with a brain drain, *Economics Letters*, Vol. 55, No. 2, pp. 227–234.

Stark, Oded, Christian Helmenstein and Alexia Prskawetz (1998): Human capital depletion, human capital formation, and migration: "'a blessing or a curse"', *Economics Letters*, Vol. 60, pp. 363–367.

Stark, Oded and Yong Wang (2002): Inducing human capital formation: migration as a substitute for subsidies, *Journal of Public Economics*, Vol. 86, No. 1, pp. 39–46.

Thum, Claudio and Silke Uebelmesser (2003): Mobility and The Role of Education as a Commitment Device, *International Tax and Public Finance*, Vol. 10, pp. 549–564.

Vidal, Jean-Pierre (1998): The effect of emigration on human capital formation, *Journal of Population Economics*, , No. 11, pp. 589–600.

Wellisch, Dietmar (1995): *Dezentrale Finanzpolitik bei hoher Mobilität*, Tübingen: Mohr (Paul Siebeck).

Wildasin, David E. (2000): Labor-Market Integration, Investment in Risky Human Capital, and Fiscal Competition, *American Economic Review*, Vol. 90, No. 1, pp. 73–95.