OPTIMAL TAX POLICIES FOR SOCIAL MOBILITY WHEN WEALTH TRANSFERS AND EDUCATION INVESTMENTS MATTER

Pierre Pestieau, Maria Racionero

LIDAM Discussion Paper CORE
2025 / 13

CORE | Louvain Institute of Data Analysis and Modeling in economics and statistics

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/core-discussion-papers

Optimal Tax Policies for Social Mobility when Wealth Transfers and Education Investments Matter

Pierre Pestieau*and Maria Racionero†

June 2025

Abstract

We consider a society where social mobility is influenced by parental wealth transfers and education investments. Specifically, the educational investments capture the time parents devote to the education of their children. We show that, in the absence of government intervention, the market equilibrium results in a level of upward social mobility lower than that in an ideal first-best scenario. Given the challenge of observing individual characteristics, we characterize the second-best solution achievable through the implementation of non-linear taxation. We consider two alternative government objectives: a weighted utilitarian criterion and a Rawlsian criterion. Additionally, we explore the implications of two alternative informational assumptions: whether educational investments are observable or non-observable.

Keywords: social mobility, non-linear taxation, inheritance taxation

JEL classification: H21,H31,H52

1 Introduction

The notion that social mobility is increasingly hindered is now broadly recognized, as evidenced by OECD (2018), Markowitz (2019), and Sandel (2020). Social mobility has markedly slowed, and the phenomenon of socioeconomic inheritance, where

*HEC, ULiege and CORE

†Research School of Economics, ANU

socioeconomic status is essentially passed down from one generation to the next, is becoming increasingly pronounced globally. This issue appears to stem from two primary factors. Firstly, as Piketty (2013) thoroughly documents, wealth inequality is on the rise, with inherited wealth becoming a more significant portion of overall wealth. Secondly, the educational system is failing to provide equitable opportunities, leading to a substantial disparity in academic success between children from low-income and high-income families.

We examine the interplay between education and wealth and their impact on an individual's ability to improve their socioeconomic status. The relationship between these factors is complex; for instance, a family's wealth can enhance the educational opportunities available to their offspring, which in turn can have significant implications for social mobility. Additionally, attending a prestigious educational institution can help individuals make better use of their wealth.

Although both education and wealth can facilitate social mobility, education is often viewed as a more merit-based pathway because it typically requires personal effort and academic achievement. Nonetheless, access to quality education can still be heavily influenced by one's socioeconomic background. Conversely, wealth is primarily contingent upon the circumstances of one's birth. Thus, we posit that for families with limited resources, education plays a critical role in social mobility, whereas for affluent families, wealth is a crucial means of maintaining their social standing

We consider parental time investments in education as a key factor influencing social mobility. Parental time involvement, such as helping with homework and engaging in educational activities, can significantly enhance a child's academic performance and long-term socioeconomic outcomes.[1] This investment appears particularly important for families with limited financial means.

We show that, without governmental intervention, the market equilibrium results in a level of social mobility significantly lower than the optimal scenario. A key issue is that individuals often overlook the broader societal implications of their educational and wealth decisions. Achieving the optimal outcome theoretically requires information that is not readily available. Therefore, we propose a second-best progressive taxation solution to mitigate these informational challenges and foster a more equitable and mobile society.

The second-best results show that when educational investments are observable, the wealthy face no marginal labor tax but receive marginal subsidies on both wealth

[1]Parental time investment is widely viewed as a key determinant for children's future economic and social success and can be a source of the intergenerational transmission of human capital (Cunha et al. 2010; Del Boca et al. 2014; Hsin and Felfe, 2014).

transfers and educational investments. The poor are subject to positive marginal labor taxes, and their wealth transfers and educational investments may be taxed or subsidized at the margin. The tax policy must balance encouraging investments by the poor while preventing mimicking by the wealthy. When educational investments are not observable, both groups face positive marginal labor taxes. The wealthy still receive marginal subsidies on wealth transfers, while the tax treatment of wealth transfers for the poor remains ambiguous.

In Section 2, we introduce the basic model, in which individuals differ in wealth and productivity, and we characterize the laissez-faire market equilibrium and first-best optimal conditions, highlighting the under-provision of education and wealth transfers without government intervention. In Section 3, we explore Pareto efficient taxation, introducing individualized tax instruments and social weights, and we characterize the optimal second-best tax policies with both observable and non-observable educational investments. In Section 4, we apply the Rawlsian maximin criterion to evaluate tax policies that focus on the welfare of the poorest individuals. In Section 5, we interpret the second-best results by further analyzing the derived tax formulas, providing insights into the effectiveness of different tax policies in promoting upward mobility and reducing inequality. Finally, in Section 6, we conclude.

2 The model

We consider a model of successive generations. In each period, that corresponds to a generation, society comprises individuals that differ in two characteristics: their productivity w and their wealth ω. We assume that these each of these variables can take two values: $\underline{w} < \bar{w}$ and $\underline{\omega} < \bar{\omega}$. In the simple model we assume that there are only two types of individuals, denoted $i = \{1, 2\}$. Type-1 individuals are endowed with both low productivity and low wealth $(\underline{w}, \underline{\omega})$, while type-2 individuals are endowed with both high productivity and high wealth $(\overline{w}, \varpi)$. This simplifying assumption makes it easier to deal with the non-linear taxation approach.

Type-i individuals care for their own consumption c_i, their labor supply, l_i and the upward mobility of their (unique) child π_i. The probability of upward mobility of a type-i child is given by $\pi_i = \varphi^i(x_i, e_i)$, where x_i represents the parental wealth transfer and e_i represents the education investment of the type-i parent. By assumption, we have $0 < \varphi^i(x_i, e_i) < 1, \varphi_1^i > 0, \varphi_{11}^i < 0, \varphi_2^i > 0$ and $\varphi_{22}^i < 0$. [2]

[2] A particular example could be $\pi_i = G\left(x_i^{\alpha_i} e_i^{\beta_i}\right)$,where $G(.)$ is an increasing and concave function that ensures that the probability belongs to the interval $(0, 1)$. $\alpha^1 < \alpha^2$ and $\beta^1 > \beta^2$ would imply that wealthy high-productivity type-2 individuals have a comparative advantage in

All individuals have the same separable utility function:

$$U = u(c) + \varphi(e, x) - v(l + e) \tag{1}$$

where $u(.)$ and $v(.)$ are respectively strictly concave and convex: i.e., $u' > 0, u" < 0, v' > 0$ and $v" > 0$. Note that educational investment is represented by the amount of time e parents devote to their children's education.[3]

We denote the relative numbers of the two types by N_1 and N_2. The size of the population is normalized to 1: i.e., $N_1 + N_2 = 1$. We define the relative number of type-2 individuals as: $N = N_2 = 1 - N_1$. In each period t, that lasts the length of a generation, the fraction of wealthy high-productivity individuals is N_t. This implies that:

$$N_{t+1} = \pi_{2t} N_t + \pi_{1t} (1 - N_t). \tag{2}$$

Following Pestieau and Racionero (2023), it can be shown that:

$$1 > \frac{dN_{t+1}}{dN_t} = \pi_{2t} - \pi_{1t} > 0 \tag{3}$$

and the share of wealthy individuals N_t converges to N^*:

$$N^* = \frac{\pi_1}{1 + \pi_1 - \pi_2}. \tag{4}$$

In the optimization below, we take into account the effect of e and x on N^*. We thus obtain in the steady state:

$$N^* = \frac{\varphi^1(e_1, x_1)}{1 + \varphi^1(e_1, x_1) - \varphi^2(e_2, x_2)}. \tag{5}$$

We thus have:

$$\frac{dN^*}{de_1} = \frac{(1 - \varphi^2(e_2, x_2))\, \varphi_1^1(e_1, x_1)}{(1 + \varphi^1(e_1, x_1) - \varphi^2(e_2, x_2))^2}, \tag{6}$$

$$\frac{dN^*}{dx_1} = \frac{(1 - \varphi^2(e_2, x_2))\, \varphi_2^1(e_1, x_1)}{(1 + \varphi^1(e_1, x_1) - \varphi^2(e_2, x_2))^2}, \tag{7}$$

wealth transfers x over education investment e relative to type-1 individuals.

[3]Parental time investments in education are crucial for children's development, as highlighted by several studies. Cunha et al. (2010) emphasize that parental involvement, such as helping with homework and engaging in educational activities, significantly enhances children's cognitive and non-cognitive skills. Del Boca, Flinn, and Wiswall (2014) find that time spent by parents on educational activities positively impacts children's academic achievements and future economic success. Similarly, Hsin and Felfe (2014) show that increased parental time devoted to educational activities leads to better educational outcomes for children.

$$\frac{dN^*}{de_2} = \frac{\varphi^1(e_1,x_1)\varphi^2_1(e_2,x_2)}{\left(1+\varphi^1(e_1,x_1)-\varphi^2(e_2,x_2)\right)^2}, \tag{8}$$

$$\frac{dN^*}{dx_2} = \frac{\varphi^1(e_1,x_1)\varphi^2_2(e_2,x_2)}{\left(1+\varphi^1(e_1,x_1)-\varphi^2(e_2,x_2)\right)^2}. \tag{9}$$

The educational investment or wealth transfer by either the wealthy or the poor increase the relative number of wealthy in the steady state. Later, to ease the interpretation of our results, we assume that type-2 individuals have a comparative advantage in wealth transfers whereas type-1 individuals have a comparative advantaje in education investments.[4]

2.1 Laissez-faire

A type-i individual who belongs to generation t is assumed to choose three variables: his labor supply l, the level of educational time e, and the level of wealth transfer x. The individual maximizes the following utility:

$$U_{it} = u(w_{it}l_{it} + \omega_i - x_{it}) + \varphi^i(e_{it}, x_{it}) - v(l_{it} + e_{it}). \tag{10}$$

The first-order conditions (FOCs) yield:

$$u'(c_{it}) = \varphi^i_2(e_{it}, x_{it}) = \frac{v'(l_{it}+e_{it})}{w_{it}} = \frac{\varphi^i_1(e_{it}, x_{it})}{w_{it}}. \tag{11}$$

2.2 First best

We now turn to the first best utilitarian optimal conditions in a dynamic setting. In each period of time, we face the resource constraint:

$$F\left(K_t, L_t\right) = \sum N_{it}\left(c_{it} + x_{it}\right) + K_{t+1} \tag{12}$$

where K is the capital stock and $L = \sum N_i \varepsilon_i l_i$ the aggregate labor supply. We use the parameter $\varepsilon_i > 1$ for $i = 2$ and $\varepsilon_i = 1$ for $i = 1$ to measure the efficiency of workers. $F\left(K, L\right)$ is a constant returns of scale (CRS) production function and total

[4]Becker and Tomes (1979) highlight that wealthier type-2 families are more likely to transfer wealth, while less wealthy type-1 families invest more in education to improve their children's future prospects. This idea is further reinforced by Becker et al. (2018), who explore the dynamics of human capital investment and wealth transfers across generations, emphasizing that wealthier families focus on wealth transfers, whereas less wealthy families prioritize educational investments.

depreciation after one period is assumed to be 0. Note that because of the CRS production function, the above resource constraint amounts to:

$$\sum N_{it}(c_{it} + x_{it}) + K_{t+1} = F_{Lt}L_t + F_{Kt}K_t, \tag{13}$$

or

$$\sum N_{it}(c_{it} + x_{it} - F_{Lt}l_{it}\varepsilon_i) + K_{t+1} - F_{Kt}K_t = 0. \tag{14}$$

The optimal conditions are obtained by maximizing the following Lagrangian:

$$\begin{aligned} \mathcal{L} \;=\; & \sum \gamma^t \{ \sum N_{it} \left[u(c_{it}) + \varphi_t^i(e_{it}, x_{it}) - v(l_{it} + e_{it}) \right] \\ & + \mu \left[F\left(K_t, \sum N_{it}\varepsilon_i l_{it} \right) - \sum N_{it}\left(c_{it} + x_{it} \right) - K_{t+1} \right] \}, \end{aligned}$$

where the factor of social time preference is denoted by γ and the parameter μ is the Lagrange multiplier associated with the resource constraint.

The FOCs are:

$$u(c_{it}) = \mu, v'(l_{it} + e_{it}) = \mu F_{Lt}\varepsilon_i = w_{it}, F_{Kt} = \gamma^{-1} = R_t, \tag{15}$$

$$\varphi_{1t}^i(e_{it}, x_{it}) + \frac{\gamma^{-1}}{N_{it}} \sum V_t \frac{\partial N_{t+1}}{\partial e_{it}} - \mu w_t = 0, \tag{16}$$

$$\varphi_{2t}^i(e_{it}, x_{it}) + \frac{\gamma^{-1}}{N_{it}} \sum V_t \frac{\partial N_{t+1}}{\partial x_{it}} - \mu = 0. \tag{17}$$

where

$$V_i = \left[u(c_i) + \varphi^i(x_i, e_i) - v(l_i + e_i) \right] - \mu \left[c_i + x_i - w_i l_i \right] \tag{18}$$

denotes the lifetime utility of the type-i individual plus the utility value of the difference between his resources and his spending. One expects that $V_2 > V_1$. In what follows we denote the difference between those two values as $\Lambda = V_2 - V_1 > 0$.

Comparing the first-best equilibrium conditions with those of the laissez-faire scenario reveals underprovision of both educational investments and wealth transfers. The insufficient levels of educational investments and wealth transfers are due to the effects these actions have on the distribution of types in the next generation, which are not considered by individuals in a laissez-faire scenario.

Assuming that we are in the steady-state and that $\gamma = 1$, we have the Golden Rule $F_K = 1$ and

$$\frac{\varphi_1^i(e_i, x_i)}{u(c_i)} = w_i - \frac{\Lambda}{N_i u(c_i)} \frac{\partial N}{\partial e_i}, \tag{19}$$

$$\frac{\varphi_2^i(e_i, x_i)}{u(c_i)} = 1 - \frac{\Lambda}{N_i u(c_i)} \frac{\partial N}{\partial x_i}. \tag{20}$$

In what follows, we focus on the steady state and assume a small open economy with fixed factor prices $R = 1$ and $w_2 > w_1$. We first analyze the Pareto efficient tax policy and then turn to the Rawlsian objective.

3 Pareto efficient taxation

3.1 Tax instruments

In this paper we consider a number of tax instruments that can be used to achieve the social optimum. In particular, in line with the informational structure, we explore the role of individualised taxes or subsidies on labor income τ_i, on wealth transfers θ_x^i and educational investments θ_e^i. When expressed in terms of the tax instruments, the utility function becomes:

$$U_i = u\left(\omega_i + w_i l_i (1 - \tau_i) - x_i \left(1 + \theta_x^i\right)\right) + \varphi^i (x_i, e_i) - v\left(l_i + e_i \left(1 + \theta_e^i\right)\right) \tag{21}$$

and the FOCs of the laissez-faire problem become:

$$\frac{v'\left(l_i + e_i \left(1 + \theta_e^i\right)\right)}{u'(c_i)} = w_i (1 - \tau_i), \tag{22}$$

$$\frac{\varphi_1^i (e_i, x_i)}{u'(c_i)} = w_i \left(1 + \theta_e^i\right) (1 - \tau_i), \tag{23}$$

$$\frac{\varphi_2^i (e_i, x_i)}{u'(c_i)} = 1 + \theta_x^i. \tag{24}$$

3.2 First best

We introduce social weights β_i to derive efficient tax policies. By doing so, we encompass the utilitarian case where $\beta_1 = \beta_2$. The Lagrangian associated with the social planner maximization problem becomes:

$$\mathcal{L} = \sum N_i \beta_i \left[u(c_i) + \varphi^i (x_i, e_i) - v\left(\frac{y_i}{w_i} + e_i\right)\right] - \mu \sum N_i (c_i + x_i - y_i - \omega_i). \tag{25}$$

and the FOCs of the first best problem become:

$$\beta_i u'(c_i) = \mu, \frac{\beta_i v'\left(\frac{y_i}{w_i} + e_i\right)}{w_i} = \mu, \tag{26}$$

$$\beta_i \varphi_1^i N_i = \beta_i v'\left(\frac{y_i}{w_i} + e_i\right) N_i - \frac{dN}{de_i}\Lambda, \tag{27}$$

$$\beta_i \varphi_2^i N_i = \mu N_i - \frac{dN}{dx_i}\Lambda. \tag{28}$$

where $V_i = \beta_i\left[u(c_i) + \varphi^i(x_i, e_i) - v(l_i + e_i)\right] - \mu\left[c_i + x_i - \omega_i - w_i l_i\right]$ now. V_i denotes the lifetime utility of the type-i individual, now adjusted by the social weight β_i, plus the utility value of the difference between type-i individual's resources and spending.

These can be rewritten as:

$$\frac{v'(\frac{y_i}{w_i} + e_i)}{u'(c_i)} = w_i, \tag{29}$$

$$\frac{\varphi_1^i}{u'\left(c_i\right)} = w_i - \frac{\Lambda}{N_i \beta_i u'(c_i)} \frac{dN_i}{de_i} < w_i, \tag{30}$$

$$\frac{\varphi_2^i}{u'(c_i)} = 1 - \frac{\Lambda}{N_i \beta_i u'(c_i)} \frac{dN_i}{dx_i} < 1. \tag{31}$$

Decentralizing this solution would require redistributive lump-sum transfers T_i and a marginal subsidy on wealth transfers ($\theta_x^i < 0$), as well as educational investments ($\theta_e^i < 0$). There is no marginal tax on labor: i.e., $\tau_i = 0$. We now consider two settings for the second-best. In the first, we assume that educational investments are observable and can thus be taxed. In the second, education investments are not observable.

3.3 Second best with observable education

In a second-best framework with imperfect information, the social planner must account for self-selection constraints to ensure that each type is better off selecting the bundles intended for her rather than mimicking another type. We focus on the regime where only the downward self-selection constraint, associated with a wealthy high-productivity type mimicking a poor low-productivity type, needs to be considered.

The problem of the social planner amounts to maximizing the following Lagrangian with respect to c_i, y_i, e_i and x_i:

$$\begin{aligned}\mathcal{L} &= \sum N_i\beta_i \left[u(c_i) + \varphi^i(e_i, x_i) - v(\frac{y_i}{w_i} + e_i)\right] - \mu \sum N_i \left(c_i + x_i - y_i - \omega_i\right) \\ &\quad + \lambda \left[u(c_2) + \varphi^2(e_2, x_2) - v(\frac{y_2}{w_2} + e_2) - u(c_1) - \varphi^2(e_1, x_1) + v(\frac{y_1}{w_2} + e_1)\right],\end{aligned}$$

where λ is the Lagrange multiplier associated with the self-selection constraint.

These FOCs can be rewritten in terms of the individual's marginal rate of substitution (MRS). Those concerning the wealthy high-productivity type-2 individuals are

$$\frac{v'(\frac{y_2}{w_2} + e_2)}{u'(c_2)} = w_2, \frac{\varphi_1^2(e_2, x_2)}{u'(c_2)} = w_2 - \frac{dN}{de_2}, \frac{\lambda}{\mu N_2} \frac{\varphi_2^2(e_2, x_2)}{u'(c_2)} = 1 - \frac{dN}{dx_2} \frac{\lambda}{\mu N_2} \tag{32}$$

or, written in terms of tax rates,

$$\tau_2 = 0, \theta_e^2 = -\frac{dN}{de_2} \frac{\Lambda}{\mu N_2} < 0, \theta_x^2 = -\frac{dN}{dx_2} \frac{\Lambda}{\mu N_2} < 0. \tag{33}$$

Hence, for type-2 individuals there is no marginal income tax, but a marginal subsidy on both educational investments and wealth transfers.

Turning to the poor low-productivity type-1 individuals, we have:

$$\frac{v'(\frac{y_1}{w_1} + e_1)}{u'(c_1)} = w_1 - \frac{\lambda w_1}{\mu N_1} \left[\frac{v'(\frac{y_1}{w_1} + e_1)}{w_1} - \frac{v'(\frac{y_1}{w_2} + e_1)}{w_2}\right], \tag{34}$$

$$\frac{\varphi_1^1(e_1, x_1)}{u'(c_1)} = w_1 - \frac{dN}{de_1} \frac{\Lambda}{\mu N_1} + \frac{\lambda}{\mu N_1} \left[\varphi_1^2(e_1, x_1) - \varphi_1^1(e_1, x_1)\right] + \frac{\lambda}{\mu N_1} \left(\frac{w_1}{w_2} - 1\right), \tag{35}$$

$$\frac{\varphi_2^1(e_1, x_1)}{u'(c_1)} = 1 - \frac{dN}{dx_1} \frac{\Lambda}{\mu N_1} + \frac{\lambda}{\mu} \left[\varphi_2^2(e_1, x_1) - \varphi_2^1(e_1, x_1)\right]. \tag{36}$$

Written in terms of tax rates:

$$\tau_1 = -\frac{\lambda w_1}{\mu N_1} \left[\frac{v'(\frac{y_1}{w_1} + e_1)}{w_1} - \frac{v'(\frac{y_1}{w_2} + e_1)}{w_2}\right] > 0, \tag{37}$$

$$\theta_e^1 = -\frac{dN}{de_1} \frac{\Lambda}{\mu N_1} + \frac{\lambda}{\mu N_1} \left[\varphi_1^2(e_1, x_1) - \varphi_2^1(e_1, x_1)\right] + \frac{\lambda}{\mu N_1} \left(\frac{w_1}{w_2} - 1\right) \gtrless 0, \tag{38}$$

$$\theta_x^1 = -\frac{dN}{dx_1}\frac{\Lambda}{\mu N_1} + \frac{\lambda}{\mu}\left[\varphi_2^2(e_1, x_1) - \varphi_2^1(e_1, x_1)\right] \gtrless 0. \tag{39}$$

Hence, for type-1 individuals, there is a positive marginal income tax. However, the results regarding the optimal tax treatment of educational investments and wealth transfers are ambiguous. On the one hand, subsidizing educational investments and wealth transfers for type-1 individuals can enhance upward mobility by improving their children's prospects. On the other hand, if these marginal subsidies are too generous, wealthier type-2 individuals might pretend to be type-1 individuals. Therefore, the tax policy must balance encouraging investments by type-1 individuals while preventing mimicking by type-2 individuals.

Note that if the social planner is not concerned by the tax incidence on N (i.e., the social planner is myopic and takes the distribution of individual types as given) and $\varphi^2 = \varphi^1$ (i.e., the probability of upward mobility is independent of the parent's type, provided that the investments are of the same magnitude), then all marginal taxes are nil except $\tau_1 > 0$ and $\theta_e^1 < 0$. The tax policy is now exclusively designed to prevent type-2 individuals from mimicking type-1 individuals.

3.4 Second best with non-observable education

We now assume that the government does not observe parental education time investments e_i. These are chosen optimally by each type-i individual so that

$$\varphi_1^i(e_i^*, x_i) = v'(\frac{y_i}{w_i} + e_i^*). \tag{40}$$

From this FOC, we have a demand function $e_i = e(y_i, x_i)$ with $e_1^i(y_i, x_i) < 0$ and $e_2^i(y_i, x_i) > 0$, granted that $\varphi_{12}^i > 0$. We thus have the following Lagrangian:

$$\begin{aligned} \mathcal{L} &= \sum N_i\beta_i\left[u(c_i) + \varphi^i(e_i, x_i) - v(\frac{y_i}{w_i} + e_i)\right] - \mu\sum N_i\left(c_i + x_i - y_i - \omega_i\right) \\ &\quad + \lambda\left[u(c_2) + \varphi^2(e_2^*, x_2) - v(\frac{y_2}{w_2} + e_2^*) - u(c_1) - \varphi^2(\hat{e}_1^*, x_1) + v(\frac{y_1}{w_2} + \hat{e}_1^*)\right], \end{aligned}$$

where $\hat{e}_1^*$ is the level of education chosen by a type-2 mimicker, which satisfies

$$\varphi_1^2(\hat{e}_1^*, x_1) = v'(\frac{y_1}{w_2} + \hat{e}_1^*). \tag{41}$$

The FOCs can be rewritten in terms of the individual's MRS. Focusing first on type-2 individuals, we have:

$$\frac{v'(\frac{y_2}{w_2}+e_2)}{u'(c_2)} = w_2\left[1+\frac{dN}{de_2}\frac{e_1^2(y_2,x_2)\Lambda}{u'(c_2)\,(N\beta_2+\lambda)}\right], \tag{42}$$

$$\frac{\varphi_2^2}{u'(c_2)} = 1-\left(\frac{dN}{dx_2}+\frac{dN}{de_2}e_2^2(y_2,x_2)\right)\frac{\Lambda}{u'(c_2)\,(N_2\beta_2+\lambda)}, \tag{43}$$

or, expressed in terms of marginal tax rates,

$$\tau_2 = -\frac{dN}{de_2}\frac{e_1^2(y_2,x_2)\Lambda}{u'(c_2)\,(N\beta_2+\lambda)} > 0, \tag{44}$$

$$\theta_x^2 = -\left(\frac{dN}{dx_2}+\frac{dN}{de_2}e_2^2(y_2,x_2)\right)\frac{\Lambda}{u'(c_2)\,(N\beta_2+\lambda)} < 0. \tag{45}$$

Hence, for type-2 individuals there is a positive marginal income tax and a marginal subsidy on wealth transfers.

Now turning to type-1 individuals, we have:

$$\frac{v'(\frac{y_1}{w_1}+e_1)}{u'(c_1)} = w_1\left[1+\frac{dN}{de_1}\frac{e_1^1(y_1,x_1)\Lambda}{N_1\beta_1-\lambda}\right]-\frac{\lambda w_1}{u'(c_1)\,(N_1\beta_1-\lambda)}\left[\frac{v'(\frac{y_1}{w_1}+e_1)}{w_1}-\frac{v'(\frac{y_1}{w_2}+\hat{e}_1^*)}{w_2}\right], \tag{46}$$

$$\frac{\varphi_2^1(e_1,x_1)}{u'(c_1)} = 1-\left(\frac{dN}{dx_1}+\frac{dN}{de_1}e_2^1(y_1,x_1)\right)\frac{\Lambda}{(N_1\beta_1-\lambda)}+\frac{\lambda}{u'(c_1)\,(N_1\beta_1-\lambda)}\left[\varphi_2^2(\hat{e}_1^*,x_1)-\varphi_2^1(e_1,x_1)\right], \tag{47}$$

or, expressed in terms of marginal tax rates,

$$\tau_1 = -\frac{dN}{de_1}\frac{e_1^1(y_1,x_1)\Lambda}{N_1\beta_1-\lambda}+\frac{\lambda}{u'(c_1)\,(N_1\beta_1-\lambda)}\left[\frac{v'(\frac{y_1}{w_1}+e_1)}{w_1}-\frac{v'(\frac{y_1}{w_2}+\hat{e}_1^*)}{w_2}\right] > 0, \tag{48}$$

$$\theta_x^1 = -\left(\frac{dN}{dx_1}+\frac{dN}{de_1}e_2^1(y_1,x_1)\right)\frac{\Lambda}{N_1\beta_1-\lambda}+\frac{\lambda}{u'(c_1)\,((N_1\beta_1-\lambda))}\left[\varphi_2^2(\hat{e}_1^*,x_1)-\varphi_2^1(e_1,x_1)\right] \gtreqless 0. \tag{49}$$

Hence, for type-1 individuals, there is a positive marginal income tax, while the optimal tax treatment of wealth transfers is ambiguous. As before, subsidizing wealth transfers for type-1 individuals can enhance upward mobility of their children. However, if the marginal subsidy is too generous, wealthier type-2 individuals might pretend to be type-1 individuals.

4 Rawlsian criterion

4.1 First best

The problem of the Rawlsian social planner can be represented by the following Lagrangian:

$$\begin{aligned}\mathcal{L} &= u(c_1) + \varphi^1(e_1, x_1) - v(\frac{y_1}{w_i} + e_1) - \mu \sum N_i\left(c_i + x_i - y_i - \omega_i\right) \\ &\quad + \zeta\left[u(c_2) + \varphi^2(e_2, x_2) - v(\frac{y_2}{w_2} + e_2) - u(c_1) - \varphi^1(e_1, x_1) + v(\frac{y_1}{w_1} + e_1)\right],\end{aligned}$$

where ζ represents the Lagrange multiplier associated with the Rawlsian constraint, ensuring equality between the utilities of the two types of individuals.

The FOCs can be rewritten as:

$$\frac{v'(\frac{y_i}{w_i} + e_i)}{u'(c_i)} = w_i, \tag{50}$$

$$\frac{\varphi_1^i}{u'\left(c_i\right)} = w_i - \frac{dN_i}{de_i}\frac{\Omega}{u'(c_i)\zeta_i}, \tag{51}$$

$$\frac{\varphi_2^i}{u'(c_i)} = 1 - \frac{dN_i}{dx_i}\frac{\Omega}{u'(c_i)\zeta_i}, \tag{52}$$

where $\zeta_1 = 1 - \zeta; \zeta_2 = \zeta$ and $\Omega = y_2 + \omega_2 - c_2 - x_2 - y_1 - \omega_1 + c_1 + x_1 > 0$.

4.2 Second best with non-observable education

We assume that the government does not observe parental time educational investments but that e^* is chosen optimally by each type-i individual according to (40). We thus have the following Lagrangian:

$$\begin{aligned}\mathcal{L} &= u(c_1) + \varphi^1(e_1^*, x_1) - v(\frac{y_1}{w_i} + e_1^*) - \mu \sum N_i\left(c_i + x_i - y_i - \omega_i\right) \\ &\quad + \lambda\left[u(c_2) + \varphi^2(e_2^*, x_2) - v(\frac{y_2}{w_2} + e_2^*) - u(c_1) - \varphi^2(\hat{e}_1^*, x_1) + v(\frac{y_1}{w_2} + \hat{e}_1^*)\right],\end{aligned}$$

where, as before, λ is the multiplier associated with the self-selection constraint and $\hat{e}_1^*$ is the level of education chosen by the type-2 mimicker.

The FOCs can be rewritten in terms of the individual's marginal rate of substitution (MRS). Focusing first on type-2 individuals, we have:

$$\frac{v'(\frac{y_2}{w_2}+e_2)}{u'(c_2)} = w_2\left[1+\frac{dN}{de_2}\frac{e_1^2(y_2,x_2)\Omega}{u'(c_2)\,(N_2+\lambda)}\right], \tag{53}$$

$$\frac{\varphi_2^2}{u'(c_2)} = 1-\left(\frac{dN}{dx_2}+\frac{dN}{de_2}e_2^2(y_2,x_2)\right)\frac{\Omega}{u'(c_2)\,(N_2+\lambda)}, \tag{54}$$

or, expressed in terms of marginal tax rates,

$$\tau_2 = -\frac{dN}{de_2}\frac{e_1^2(y_2,x_2)\Omega}{u'(c_2)\lambda} > 0, \tag{55}$$

$$\theta_x^2 = -\left(\frac{dN}{dx_2}+\frac{dN}{de_2}e_2^2(y_2,x_2)\right)\frac{\Omega}{u'(c_2)\lambda} < 0. \tag{56}$$

Hence, for type-2 individuals there is a positive marginal income tax and a marginal subsidy on wealth transfers.

Now turning to type-1 individuals, we have:

$$\frac{v'(\frac{y_1}{w_1}+e_1)}{u'(c_1)} = w_1\left[1+\frac{dN}{de_1}\frac{e_1^1(y_1,x_1)\Omega}{u'(c_1)\,(1-\lambda)}\right]-\frac{\lambda w_1}{u'(c_1)\,(1-\lambda)}\left[\frac{v'(\frac{y_1}{w_1}+e_1)}{w_1}-\frac{v'(\frac{y_1}{w_2}+\hat{e}_1^*)}{w_2}\right], \tag{57}$$

$$\frac{\varphi_2^1(e_1,x_1)}{u'(c_1)} = 1-\left(\frac{dN}{dx_1}+\frac{dN}{de_1}e_2^1(y_1,x_1)\right)\frac{\Omega}{u'(c_1)\,(1-\lambda)}+\frac{\lambda}{u'(c_1)\,(1-\lambda)}\left[\varphi_2^2(\hat{e}_1^*,x_1)-\varphi_2^1(e_1,x_1)\right], \tag{58}$$

or, expressed in terms of marginal tax rates,

$$\tau_1 = -\frac{dN}{de_1}\frac{e_1^1(y_1,x_1)\Omega}{u'(c_1)\,(1-\lambda)}+\frac{\lambda}{u'(c_1)\,(1-\lambda)}\left[\frac{v'(\frac{y_1}{w_1}+e_1)}{w_1}-\frac{v'(\frac{y_1}{w_2}+\hat{e}_1^*)}{w_2}\right] > 0 \tag{59}$$

$$\theta_x^1 = -\left(\frac{dN}{dx_1}+\frac{dN}{de_1}e_2^1(y_1,x_1)\right)\frac{\Omega}{u'(c_1)\,(1-\lambda)}+\frac{\lambda}{u'(c_1)\,(1-\lambda)}\left[\varphi_2^2(\hat{e}_1^*,x_1)-\varphi_2^1(e_1,x_1)\right] \lesseqgtr 0. \tag{60}$$

Hence, for type-1 individuals, there is a positive marginal income tax, while the optimal tax treatment of wealth transfers is ambiguous.

5 Interpretation

To interpret these second-best results, we focus on the four relevant tax formulas applicable when parental education time investments are not observable: i.e., (44)-(45) for type-2 individuals and (48)-(49) for type-1 individuals.

Irrespective of values of the social welfare weights β_i, it appears that the marginal income tax is positive for both types of individual. This is more so for the poor than for the wealthy, which is consistent with optimal non-linear income taxation results (e.g., Stiglitz (1982)). These results hinge on $e_1^i(y_i, x_i) < 0$. There is a marginal subsidy on the wealth transfers of the wealthy, while the tax treatment of wealth transfers of the poor is ambiguous. On the one hand, the desire to increase the number of wealthy high-productivity individuals in society leads to a subsidy. On the other hand, a marginal tax may be desirable to relax the self-selection constraint.

There is something paradoxical about subsidizing the intergenerational transfers of the wealthy, even when the aim is to maximize the utility of the poor. This is explained by the fact that the transfers of the wealthy high-productivity parents have a positive effect on the proportion of wealthy high-productivity individuals in the next generation. This increase makes it possible to finance greater redistribution through lump-sum transfers, benefiting poor low-productivity individuals. For the same reason, taxing labour earnings is desirable as it fosters educational investments.

Shifting from a utilitarian objective ($\beta_1 = \beta_2$) to Rawlsian objective ($\beta_1/\beta_2 \to \infty$) has the effect of decreasing the marginal income tax on the poor and increasing that one the wealthy. It has also the effect of decreasing the marginal subsidy on wealth transfers by the wealthy and either increasing the marginal subsidy or decreasing the marginal tax on wealth transfers by the poor.

If we assume that $dN/de_2 \to 0$, then $\tau_2 = 0$. The assumption implies that educational investments by wealthy parents have a negligible effect on the proportion of wealthy high-productivity individuals in the next generation, indicating a significant comparative advantage in wealth transfers over educational time investments. In this case the optimal marginal labour income tax rate on the wealthy becomes zero, suggesting that policies should focus on subsidizing wealth transfers to maximize their positive impact on the distribution and support redistribution efforts benefiting the poor.

Note, finally, that if the social planner is myopic and ignores the effect of its tax policies on the distribution of individual types, we obtain:

$$\tau_2 = \theta_x^2 = 0, \tag{61}$$

$$\tau_1 = \frac{\lambda}{u'(c_1)\,(N_1\beta_1 - \lambda)} \left[\frac{v'(\frac{y_1}{w_1} + e_1)}{w_1} - \frac{v'(\frac{y_1}{w_2} + \hat{e}_1^*)}{w_2} \right] > 0, \tag{62}$$

$$\theta_x^1 = \frac{\lambda}{u'(c_1)\,(N_1\beta_1 - \lambda)} \left[\varphi_2^2(\hat{e}_1^*, x_1) - \varphi_2^1(e_1, x_1) \right] > 0. \tag{63}$$

The optimal marginal tax rate on labor income and the marginal subsidy on wealth transfers for the wealthy become zero: i.e., we recover the traditional non-distortion at the top result. In contrast, the poor face positive marginal taxes on both labor income and wealth transfers, with both taxes designed to prevent wealthy individuals from mimicking poor individuals.

6 Conclusion

This paper explores the optimal tax mix policy for a government aiming to enhance social mobility and equity, focusing on a society where individuals value personal consumption, leisure, and the upward mobility of their offspring. Social mobility depends on parental educational time investments and wealth transfers. The model considers two groups: the wealthy and the poor, who differ in both labor productivity and initial wealth. We examine the steady-state implications of using non-linear labor income, education and wealth transfer taxes as policy instruments.

When parental time educational investments are observable, the optimal policy is characterized by no marginal labor income tax and marginal subsidies on both educational investments and wealth transfers for the wealthy. These subsidies seek to encourage wealthy parents to invest in their children, enhancing their children's prospects of becoming wealthy high-productivity individuals. Poor individuals face a positive marginal income tax, while the optimal tax treatment of their educational investments and wealth transfers is ambiguous. The optimal tax menu is designed to balance encouraging poor parents' investments to ensure the upward mobility of their children while preventing wealthier individuals from mimicking them.

When educational investments are not observable, the optimal policy involves positive marginal labor income taxes for both the wealthy and the poor. The wealthy still receive marginal subsidies on wealth transfers, while the optimal tax treatment of the wealth transfers by the poor remains ambiguous. These results hinge on the effect of the tax policy on the distribution of individual types. When the social planner disregards the impact of taxation on the proportion of wealthy high-productivity versus poor low-productivity individuals, the resulting optimal policy imposes zero marginal labor income or wealth taxes on the wealthy, while imposing positive marginal taxes on both the poor's labor income and wealth transfers.

Surprisingly, this analysis indicates that subsidizing the educational investments and wealth transfers of the wealthy could maximize the utility of the poor. This strategy is effective because it increases the proportion of wealthy high-productivity individuals, thereby expanding the tax base and enhancing the redistribution of resources to support the poor.

We acknowledge two main limitations: the binary nature of individual types and the exclusive focus on the steady-state scenario. In future research we plan to expand on these aspects, incorporating a broader range of individual types and examining the dynamic interactions within this model through numerical simulations, following further work by Cremer et al. (2003). This extension will aim to provide a more nuanced understanding of the policy implications across different societal structures and economic conditions.

At first glance, the conclusion of this paper might seem to vindicate the trickle-down theory, which states that favoring the upper tier of the economic spectrum will trickle down to those less fortunate in the form of higher income. Indeed, our model shows that marginally subsidizing the wealth transfers of the wealthy, and their educational investments when these are observable, allows for lump-sum transfers that improve the well-being of the poor. However, in the real world, trickle-down policies have tended to widen income inequality because those lump-sum transfers are typically missing. As an analogy, applying Mirrlees' (1971) conclusion that the marginal income taxation of the rich should be zero without lump-sum redistribution would be regressive and go against the spirit of optimal income tax theory.

References

[1] Becker, G. S., Kominers, S. D., Murphy, K. M., & Spenkuch, J. L. (2018). A Theory of Intergenerational Mobility. Journal of Political Economy, 126(S1), S7-S25.

[2] Becker, G. S., Tomes, N. (1979). An Equilibrum Theory of the Distribution of Income and Intergenerational Mobility, Journal of Political Economy, 87(6), 1153-89.

[3] Cremer, H., Pestieau, P., Rochet, J.-Ch., (2001). Direct versus indirect taxation. The design of the tax structure revisited. International Economic Review, 42, 781–799.

[4] Cremer, H., Pestieau, P., Rochet, J.-Ch., (2003). Capital income taxation when inherited wealth is not observable, Journal of Public Economics, 87, 2475-2490.

[5] Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. Econometrica, 78(3), 883-931. doi:10.3982/ECTA6551.

[6] Del Boca, D., Flinn, C., & Wiswall, M. (2014). Household Choices and Child Development. The Review of Economic Studies, 81(1), 137-185. doi:10.1093/restud/rdt026.

[7] Hsin, A., & Felfe, C. (2014). When Does Time Matter? Maternal Employment, Children's Time With Parents, and Child Development. Demography, 51(5), 1867-1894. doi:10.1007/s13524-014-0334-5.

[8] Markovits, D. (2019). The Meritocracy Trap: How America's Foundational Myth Feeds Inequality, Dismantles the Middle Class, and Devours the Elite, New York: Penguin Press.

[9] Mirrlees J. (1971). The Theory of Optimal Income Taxatio, Review of Economic Studies, 38(114), 175-208.

[10] OECD (2018). A Broken Social Elevator? How to Promote Social Mobility, OCDE Paris.

[11] Pestieau, P., Racionero, M. (2023). Mobility, education and redistribution, CORE DP 2023/24.

[12] Piketty, T. (2013). Le Capital au XXIe siecle. Paris: Editions du Seuil.

[13] Sandel, M. J. (2020). The Tyranny of Merit: What's Become of the Common Good? New York: Farrar, Straus and Giroux.

[14] Stiglitz, J.E. (1982). Self-selection and Pareto efficient taxation, Journal of Public Economics, 17(2), 213-240.