UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche ed Aziendali "Marco Fanno"

OPTIMAL TAXATION AND PRODUCTIVE SOCIAL EXPENDITURE

THOMAS BASSETTI
University of Padova

LUCIANO GRECO
University of Padova

April 2015

"MARCO FANNO" WORKING PAPER N.196

Optimal Taxation and Productive Social Expenditure

Thomas Bassetti[*] Luciano Greco[†]

April 2015

Abstract

This paper characterizes the optimal tax and expenditure policies in economies where households' unobservable gross earnings depend on exogenous (or inherited) capabilities and input investments. In a two-class economy, optimal redistribution relies on non-linear income taxation and input public provision only if the poor households demand less input than the rich. In a multi-class economy, optimal redistribution is implemented by usual-shape, non-linear income taxation and uniform public provision of input, if inherited capability and input are economic substitutes. But, when capability and input are complements, optimal redistribution relies only on non-linear income taxation. Numerical analyses show that, when individual productivity is separable in input and capability, these factors are economic substitutes (or complements) if preferences take into account (or not) the income effects.

Keywords: In-kind redistribution; Non-linear income tax; Public provision of private goods; Opting out; Topping up; Numerical simulations

JEL classification: H42, H21

[*]Department of Economics and Management, University of Padova

[†]Corresponding author, Department of Economics and Management, University of Padova, via del Santo 33 - 35123 Padova (Italy), luciano.greco@unipd.it

1 Introduction

Several indicators show a long-run trend toward more unequal distribution of wealth, income, and economic opportunities among households in OECD and emerging countries. Though consensus is hard to find on such issues, a number of studies have documented the role of public policies in shaping wealth and income distribution at different times. Lower tax rates on top incomes and on wealth ownership and transmission have strongly contributed to the growth of inequality (Piketty, 2013). But the redistributive impact of public expenditure on social services has partly compensated the described trend (OECD, 2011).

The political and economic debate has recently focused on the possible causality between rising income and wealth inequality and reduced intergenerational social mobility (Corak, 2013). In particular, lingering low wages tend to reduce the capacity of poorer households to invest in inputs (e.g., education, health-care, child-care) that increase offsprings' earnings (Cingano, 2014). These findings strengthen the case for public expenditure on "productive" social services (i.e., inputs having a relevant impact on households' earnings[1]), and motivate investigations on the appropriate design of the redistributive policy mix, including taxation, monetary social benefits, and in-kind transfers (OECD, 2014).

The theoretical literature on taxation and public provision of private goods offers interesting lessons to address the issue of how to design optimal redistributive policies, trading off distortions (induced by taxes and social expenditure) and redistributive goals.[2] A first general lesson is that public programs for the provision of social services affect the way marginal tax rates distort households' choices and the optimal tax schedule (Blomquist et al., 2010). In particular, they may reduce the efficiency cost of redistribution in two

[1]Child-care, education and health-care affect human capital accumulation; also, child- and elderly-care influence households' (market) productive capacity, by reducing the need for informal work within the household.

[2]This literature had mainly been developed during the last two decades of the XX century. For a survey, see Balestrino (1999, 2000).

ways. First, the public provision may alleviate tax distortion on individual choices by forcing households to use more of the publicly-provided good (e.g., Boadway and Marchand (1995); Cremer and Gahvari (1997)). Second, if in-kind transfers can be targeted to poor households more effectively than cash, they improve the redistribution capacity of public policies by reinforcing self-selection mechanisms (e.g., Blomquist and Christiansen (1995)).

Another important lesson is that the precise mechanism implementing the public provision of the considered services matters (Blomquist and Christiansen, 1998a,b; Greco, 2011). Any public provision mechanism corresponds to one of two basic schemes (also called pillars) or is a combination of them. The first is the *topping-up* scheme, where government can support private expenditure on social services using conditional transfers - such as, vouchers or tax allowances - to buy a given quality or quantity of the considered good, and all households can top the basic public provision up. The second is the *opting-out* scheme, where government can offer a publicly-provided service as an alternative to market ones; in this case, households are free to choose private or public services, but once the latter is chosen private supplementing is legally forbidden or technologically unfeasible (e.g., a child can attend only one school).[3]

In this paper, we investigate the optimal redistribution policies that take into account both non-linear taxes on labor income and, possibly, the public provision of a productive social service. The latter is an input affecting the productivity of the household. In our analysis, we explicitly consider that part of each household's productivity is determined by inherited, exogenous productive capability (e.g., human capital, social network, wealth). If such capability has a substantial impact on the household's wage, the considered economy may also feature low social mobility.

The cornerstone of our analysis is the relationship between exogenous capability and

[3] Mixed schemes, where topping-up and opting-out mechanisms coexist, can be observed in the real world: e.g., households opting for a public school automatically give up tax credits for private schools; hence, savings on tax credits implicitly finance public school expenditure; also, an increase of tax allowances for households' expenditure on private schools implicitly reduces the additional transfer underlying free public schools.

endogenous investment in input. Broadly speaking, input and capability may be complements or substitutes in the terms of production technology that is represented by the wage function. This depends on the intrinsic features of the economy that we take as exogenous. However, whether high-capability (say, richer) households end up investing more (or less) in the productive input depends also on economic considerations, such as the impact of the income effects on individual choices.

A first contribution we make to the literature is that we include the choice of the optimal public provision scheme in a standard model of non-linear income taxation. To our knowledge, such extension is a novelty both as regards the theoretical analysis and the numerical simulation algorithm. In the case of a simple distribution with two types of household (i.e., high- and low-capability), we find that: the public provision of input supplements non-linear taxation only if low-capability households demand less input than high-capability ones; if households' investments in input differ substantially, the optimal public provision corresponds to a pure opting-out scheme and the marginal tax schedule has the usual shape.

A second innovative contribution of the paper is the extension of the analysis to the case of a continuum of households' capability levels. In this case, we find that: the optimal redistribution scheme is based on uniform public provision and on usual-shape, non-linear income tax only if higher-capability households demand less input than lower-capability ones; otherwise, the optimal redistribution policy is based only on monetary transfers.

A third contribution of the paper is the characterization of the assumptions on technology and preferences that are conducive to alternative optimal redistribution regimes. To this end we numerically simulate two standard optimal taxation models. We show that, if the household's wage function is separable in its arguments, the income effects change the relationship between input and capability, influencing the choice of the optimal public provision scheme.

The paper is organized as follows. Section 2 presents the model. In Section 3 the optimal policies in the two-class economy are analyzed. Section 4 extends the model to the case of a multi-class economy and numerically characterizes the results. Finally, Section 5 concludes.

2 The Model

The economy is populated by a continuum of households with utility function, $\mathcal{U}(c, l)$, that is strictly increasing in private consumption, c, strictly decreasing in labor supply, l, and concave in both arguments. Consumption and leisure are normal goods. The production technology features constant returns-to-scale, and the labor market is competitive. Thus, for each household, the gross wage, $w(\theta, q)$, is equal to the individual productivity. In particular, $w(.,.)$ is strictly increasing and twice differentiable in household's inherited capability, θ, and input, q, and concave in the latter argument. The prices of investment input and consumption good are normalized to one.

The government maximizes the sum of households' utilities under the budget and incentive constraints. In particular, it can only observe the gross income, $y = w(\theta, q) \cdot l$, and the net-of-tax income, x; while it cannot observe the gross wage, the household's labor supply, the private consumption, and the private investment in input, q^m. The input is also publicly provided through a two-pillar scheme. The first pillar is based on a topping-up mechanism in which the government supplies a uniform quantity, q^f, independent of households' consumption and investment choices. Then, a supplementary quantity of input, q^s, is provided to individuals opting for a second public pillar, and accepting not to privately top up the public provision. Individuals opting-out of the second pillar can privately supplement the first pillar input provision buying on the market q^m.

The timing of the model reads as follows. The government determines the non-linear income tax and the first- and second-pillar public provision of input; then the households choose gross- and net-of-tax incomes and whether to opt out of the second-pillar provision (and, if they do, the amount of q^m).

If the household with capability θ decides to opt out, it also chooses the private investment in input:

$$Q^m(x, y, q^f, \theta) \equiv Arg \max_{q^m \geq 0} \mathcal{U}(x - q^m, \frac{y}{w(\theta, q^f + q^m)}) \tag{1}$$

where $c = x - q^m$ and $l = \frac{y}{w}$; the indirect utility function can be written as:

$$V(x, y, q^f, \theta) \equiv \max_{q^m \geq 0} \mathcal{U}(x - q^m, \frac{y}{w(\theta, q^f + q^m)}). \tag{2}$$

By usual comparative statics, it is possible to show that $Q^m(.)$ is increasing in net and gross incomes, and it is decreasing in the first-pillar public provision, though the crowding out is not complete: $Q^m_{q^f}(.) \in (-1, 0)$. The sign of $Q^m_\theta(.)$ depends on the nature of technological complementarities between input and capability.[4] We assume that the Single Crossing Property (SCP) is satisfied: $\frac{d}{d\theta}\frac{dx}{dy} \mid_V < 0$, where $\frac{dx}{dy} \mid_V = -\frac{V_y}{V_x}$.

An opting-in household consumes all its net income, thus its indirect utility function is

$$U(x, y, q^f + q^s, \theta) \equiv \mathcal{U}(x, \frac{y}{w(\theta, q^f + q^s)}). \tag{3}$$

The minimum second-pillar provision inducing the household with capability θ to opt in is:

$$Q^s(x, y, q^f, \theta) \equiv \{q' \in \mathcal{R}_+ \mid U(x, y, q^f + q', \theta) = V(x, y, q^f, \theta)\}. \tag{4}$$

It can be easily proved that $Q^s(.)$ is increasing in x and y, and decreasing in q^f, with $Q^s_{q^f}(.) \in (-1, 0)$. Also, $Q^s_\theta > 0$ (respectively, $Q^s_\theta < 0$) if capability and input are strong

[4] If q and θ are strong technologic complements (respectively, weak technologic complements or technologic substitutes), i.e., there is some $w^+_{q\theta} > 0$ such that $w_{q\theta} > w^+_{q\theta}$ (respectively, $w_{q\theta} < w^+_{q\theta}$), then input and capability are economic complements, $Q^m_\theta > 0$, (respectively, economic substitutes, $Q^m_\theta < 0$). It is worth to remark that, *ceteris paribus*, to observe economic complementarity - that is higher-capability households demand less input - strong technologic complementarity is required. The intuition is that the income effects of capability growth lead to higher demand for consumption, thus reducing the scope of private demand for input. The results presented in this section, as well as other minor results we refer to throughout the paper, are proven in a separate technical appendix available from the authors upon request.

technologic complements (respectively, weak technologic complements or technologic substitutes), i.e., there is $w^*_{q\theta} > 0$ such that $w_{q\theta} > w^*_{q\theta}$ (respectively, $w_{q\theta} < w^*_{q\theta}$). The SCP (i.e., $\frac{d}{d\theta}\frac{dx}{dy}\mid_U < 0$, where $\frac{dx}{dy}\mid_U = -\frac{U_y}{U_x}$) is always satisfied for opting-in households.

If government could rely on first best redistribution (hence, $\frac{dx}{dy}\mid_V = 1$ for all θ), cash redistribution would always be superior to in-kind transfers. However, the first best lump sum transfers may be incentive-incompatible when individual capabilities are not observable. In fact, although poorer households have no incentive to mimic richer households, the reverse may occur. In the following, we assume that this is the case.

3 Two-Class Economy

In the simplest possible case, we assume that a fraction $\lambda \in (0, 1)$ of households have low capability ($\underline{\theta}$), while the others have high capability ($\overline{\theta}$). Depending on government policies and structural parameters, four possible regimes may arise (Greco, 2011):

PT in the pure taxation regime, q^f and q^s are such that no individual is constrained by the first pillar and no individual opts for the second pillar public provision; hence, $q^s < \min\{\underline{Q}^s, \overline{Q}^s\}$, where $\underline{Q}^s \equiv Q^s(\underline{x}, \underline{y}, q^f, \underline{\theta})$ and $\overline{Q}^s \equiv Q^s(\overline{x}, \overline{y}, q^f, \overline{\theta})$, with $\{\underline{x}, \underline{y}\}$ and $\{\overline{x}, \overline{y}\}$ the net and gross incomes of low-capability and high-capability households, respectively;

IN in the inclusive regime, q^f and q^s are such that either all individuals are constrained by the first pillar or all of them opt for the second pillar public provision; hence, $q^s > \max\{\underline{Q}^s, \overline{Q}^s\}$;

we may also have two discriminating regimes, depending on the type (high- or low-capability) of household that opts for the second pillar:

DL when $\underline{Q}^s \leq q^s \leq \overline{Q}^s$, low-capability households opt for the second-pillar provision while high-capability households opt out;

DH when $\underline{Q}^s \geq q^s \geq \overline{Q}^s$, high-capability households opt for the second-pillar provision while low-capability households opt out.

Passing from one regime to the other introduces discontinuities in the structure of government objective and constraints. Therefore, we first find optimal solutions within each policy regime, then we discuss global optima.

It is easy to check that in the PT and IN regimes, the optimal marginal tax schedule performs the usual shape, since it is not influenced by individual choices regarding input investment. In the PT case, this is so because input investment is completely private, and it is affected by marginal taxation only indirectly (through the wage function). Under the IN regime, input investment is uniformly determined by the government.

3.1 Optimal Redistribution in Discriminating Regimes

First, we consider the case in which only low-capability households opt in (i.e., DL regime). In this case, also the high-capability households who try to mimick the low-capability are forced to opt for the second pillar provision. The maximization problem of the government is

$$\max_{\{\underline{x},\underline{y},\overline{x},\overline{y},q^f,q^s\}} \lambda \cdot U(\underline{x},\underline{y},q^f+q^s,\underline{\theta}) + (1-\lambda)\cdot V(\overline{x},\overline{y},q^f,\overline{\theta})$$

$$s.t.: \tag{5}$$

$$\lambda\cdot(\underline{y}-\underline{x}) + (1-\lambda)\cdot(\overline{y}-\overline{x}) - q^f - \lambda\cdot q^s \geq 0 \quad (\mu)$$

$$V(\overline{x},\overline{y},q^f,\overline{\theta}) \geq U(\underline{x},\underline{y},q^f+q^s,\overline{\theta}) \quad (\nu)$$

$$Q^s(\overline{x},\overline{y},q^f,\overline{\theta}) \geq q^s \quad (\overline{\varphi})$$

$$q^s \geq Q^s(\underline{x},\underline{y},q^f,\underline{\theta}) \quad (\underline{\varphi})$$

$$q^f \geq 0 \quad (\underline{\phi})$$

$$Q^s(\overline{x},\overline{y},q^f,\overline{\theta}) \geq 0 \quad (\overline{\phi})$$

where μ is the Lagrangian multiplier associated to the government's budget constraint, ν is associated to the incentive constraint, $\overline{\varphi}$ and $\underline{\varphi}$ are associated to the upper and lower constraints on q^s, while $\underline{\phi}$ and $\overline{\phi}$ are associated to the upper and lower constraints on q^f.

By the first order conditions of (5), we derive the optimal marginal tax rate for the high-capability households:

$$T'_{DL}(\overline{y}) = 1 - \frac{dx}{dy}\mid_{\overline{V}} = -\frac{\overline{\varphi}+\overline{\phi}}{1-\lambda+\nu}\cdot\frac{\overline{Q}^s_x+\overline{Q}^s_y}{\overline{V}_x} \leq T'_{PT}(\overline{y}) = 0, \tag{6}$$

where $\frac{dx}{dy}\mid_{\underline{V}} \equiv -\frac{\underline{V}_y}{\underline{V}_x}$ and $\overline{V} \equiv V(\overline{x},\overline{y},q^f,\overline{\theta})$; and for the low-capability households:

$$T'_{DL}(\underline{y}) = 1 - \frac{dx}{dy}\mid_{\underline{U}} = \frac{\frac{\nu}{\lambda}\cdot\frac{\widehat{U}_x}{\underline{U}_x}\cdot\left(\frac{dx}{dy}\mid_{\underline{U}} - \frac{dx}{dy}\mid_{\widehat{U}}\right) + \frac{\underline{\varphi}}{\lambda\cdot\underline{U}_x}\cdot(\underline{Q}^s_x+\underline{Q}^s_y)}{1-\frac{\nu}{\lambda}\cdot\frac{\widehat{U}_x}{\underline{U}_x}} \geq T'_{IN}(\underline{y}), \tag{7}$$

where $\frac{dx}{dy}\mid_{\underline{U}} \equiv -\frac{\underline{U}_y}{\underline{U}_x}$, $\frac{dx}{dy}\mid_{\widehat{U}} \equiv -\frac{\widehat{U}_y}{\widehat{U}_x}$, $\underline{U} \equiv U(\underline{x},\underline{y},q^f+q^s,\underline{\theta})$, and $\widehat{U} \equiv U(\underline{x},\underline{y},q^f+q^s,\overline{\theta})$.[5]

We now consider optimal public provision schemes, abstracting from corner solutions.

[5]Notice that, by the SCP, $\frac{dx}{dy}\mid_{\underline{U}} > \frac{dx}{dy}\mid_{\widehat{U}}$.

This is the relevant case when the heterogeneity between low- and high-capability households is sufficiently large that lower and upper constraints on q^s and q^f do not bind. In other words, there is enough scope for the public provision program to discriminate between different classes, providing each of them with a different (total) level of publicly-provided input (i.e., the rich will take only the first-pillar provision, while the poor will benefit of the sum of first- and second-pillar publicly-provided input). Thus, we have:

Proposition 1 *Under the DL regime and assuming away corner solutions, the optimal discriminating policy is such that:*

1. *the public provision mix is equivalent to pure opting-out (i.e.,* $q^{f*} = 0$ *and* $q^{s*} > 0$*) and such that low-capability households are constrained to over-invest (or under-invest) if inherited capability and input are economic substitutes (or complements);*

2. *the tax schedule performs the usual shape as under the PT regime.*

Proof. See Appendix A. ■

The intuition of the Proposition 1 is that any mixed scheme is equivalent to a pure opting-out one (i.e., only the second pillar is active) in terms of individual behaviors, government budget constraint and social welfare. In fact, in the considered setting, the first-pillar provision is equivalent to a uniform cash transfer. However, the level of (second-pillar) publicly provided input depends on complementarity or substitutability of input and capability. This is justified by the impact of the public provision on the incentive constraint of the rich (who tries to mimick the poor). When input and capability are economic complements, a reduction of the publicly-provided input harms the mimicker more than the poor. Conversely, when input and capability are economic substitutes, increasing the public provision beyond what low-capability households would choose on their own allows to redistribute more at reduced incentive costs.

The result that the optimal tax schedule is unaffected by the optimal public provision does not hold when we consider corner solutions. If the scope for the second-pillar program is limited by the aim to discriminate between the rich (that should opt out) and the poor (that should opt in), the optimal tax schedule has to partly offset the effect of these discriminating constraints. And, the optimal public provision scheme may be influenced as well.

When input and capability are economic substitutes (see Proposition 5 in Appendix B), the only possible corner solution is that the upper constraint on q^s binds. This result is due to the fact that, in a second-best scenario, it is optimal to overextend the public provision of the productive input. The corner solution arises because the optimal level of the public provision would be beyond the minimum level inducing the high-capability households to opt in. In this case, the government can redistribute more by a first-pillar public provision that is set at the level the rich households would have chosen on their own without any public provision. However, in this case, the provision scheme affects the optimal tax schedule. In particular, given that the optimal public provision is limited from above, an additional incentive has to be given to high-capability labor supply.

When input and capability are economic complements, the lower bound on the second-pillar public provision may be binding (see Proposition 6 in Appendix B); in other terms, the optimal publicly-provided input may be below the minimum level needed to convince low-capability households to opt in. Therefore, the marginal tax rate on the poor requires an additional term to discourage mimicking.

The minimum level of input inducing households to opt for the second-pillar public provision increases in gross and net incomes. However, if economic substitutability between θ and q is sufficiently strong, we may observe a DH regime, where high-capability households opt in and low-capability households opt out. In this peculiar case, the maximization problem of the government becomes:

$$\max_{\{\underline{x},\underline{y},\overline{x},\overline{y},q^f,q^s\}} \lambda \cdot V(\underline{x},\underline{y},q^f,\underline{\theta}) + (1-\lambda)\cdot U(\overline{x},\overline{y},q^f+q^s,\overline{\theta})$$

$$s.t.: \tag{8}$$

$$\lambda\cdot(\underline{y}-\underline{x}) + (1-\lambda)\cdot(\overline{y}-\overline{x}) - q^f - (1-\lambda)\cdot q^s \geq 0 \quad (\mu)$$

$$U(\overline{x},\overline{y},q^f+q^s,\overline{\theta}) \geq V(\underline{x},\underline{y},q^f,\overline{\theta}) \quad (\nu)$$

$$Q^s(\underline{x},\underline{y},q^f,\underline{\theta}) \geq q^s \quad (\underline{\varphi})$$

$$q^s \geq Q^s(\overline{x},\overline{y},q^f,\overline{\theta}) \quad (\overline{\varphi})$$

$$q^f \geq 0 \quad (\underline{\phi})$$

$$Q^s(\underline{x},\underline{y},q^f,\underline{\theta}) \geq 0 \quad (\overline{\phi})$$

By the first order conditions of (8), the optimal marginal tax rate for high-capability households is:

$$T'_{DH}(\overline{y}) = 1 - \frac{dx}{dy}\mid_{\overline{U}} = \overline{\varphi}\cdot\frac{\overline{Q}^s_x + \overline{Q}^s_x}{(1-\lambda+\nu)\cdot\overline{U}_x} \geq T'_{IN}(\overline{y}) = 0; \tag{9}$$

and for low-capability households is:

$$T'_{DH}(\underline{y}) = 1 - \frac{dx}{dy}\mid_{\underline{V}} = \frac{\frac{\nu}{\lambda}\cdot\frac{\widehat{V}_x}{\underline{V}_x}\cdot\left(\frac{dx}{dy}\mid_{\underline{V}} - \frac{dx}{dy}\mid_{\widehat{V}}\right) + \frac{\underline{\varphi}+\overline{\phi}}{\lambda\cdot\underline{V}_x}\cdot(\underline{Q}^s_x + \underline{Q}^s_y)}{1-\frac{\nu}{\lambda}\cdot\frac{\widehat{V}_x}{\underline{V}_x}} \geq T'_{PT}(\underline{y}). \tag{10}$$

As regards the optimal public provision, we have:

Proposition 2 *Under the DH regime, the optimal discriminating policy replicates the optimal policy under the PT regime. In particular, public provision is irrelevant.*

Proof. See Appendix A. ■

Interestingly, when the rich enters the (second-pillar) public provision scheme for lower levels of publicly provided input, the public provision program does not add anything to taxation as redistribution tool.

3.2 Optimal Redistribution in a Two-Class Economy: Discussion

Under the DH regime, the public provision is irrelevant and the optimal discriminating policy corresponds to a PT regime. In contrast, under the more intuitive DL regime, the optimal discriminating policy always relies on both public provision and taxation.

As regards the comparison between the IN regime and the DL (or PT) regime, an argument similar to Greco (2011) applies: whenever the preference for redistribution is not too strong and/or the scope for discriminating policies is sufficiently large, the IN regime - which is affected by the efficiency loss due to uniform investment in input across heterogeneous households - is dominated by the DL (or PT) regime.

This result disappears in the case of a multi-class economy where the scope for discriminating policy dramatically shrinks.

4 Multi-Class Economy

We now extend our model to the case in which capability may vary continuously across households: $\theta \in [\underline{\theta}, \overline{\theta}]$, with probability and density functions $F(\theta)$ and $f(\theta)$, respectively. Let $\tilde{\theta} \equiv \{\theta \mid U(x(\theta), y(\theta), q^f + q^s, \theta) = V(x(\theta), y(\theta), q^f, \theta)\}$ be the threshold capability such that the household is indifferent between opting-in and opting-out.

First, we analyze the case in which input and capability are economic complements, $\tilde{\theta}_{q^s} > 0$ (i.e., $U_{\theta q} > \max\left\{0, U_{\theta x} \cdot \frac{Q^m - q^s}{Q^m}\right\}$), hence households with higher capability demand more input. It is easy to see that, in this case, households with $\theta \leq \tilde{\theta}$ opt for the (second-pillar) public provision, while households with $\theta > \tilde{\theta}$ opt out. The government maximizes the sum of households' utilities,

$$\max \int_{\underline{\theta}}^{\tilde{\theta}} u(\theta) \cdot dF(\theta) + \int_{\tilde{\theta}}^{\overline{\theta}} v(\theta) \cdot dF(\theta), \tag{11}$$

under the public budget constraint,

$$\int_{\underline{\theta}}^{\overline{\theta}} [y(\theta) - x(\theta)] \cdot dF(\theta) - q^f - q^s \cdot F(\tilde{\theta}) \geq 0. \tag{12}$$

and the incentive constraints for opting-in households,

$$u(\theta) \geq U(x(\theta'), y(q^f + q^s, \theta)) \tag{13}$$

for all $\theta' \neq \theta$ and all $\theta \leq \tilde{\theta}$, and opting-out households,

$$v(\theta) \geq V(x(\theta'), y(q^f, \theta)) \tag{14}$$

for all $\theta' \neq \theta$ and all $\theta > \tilde{\theta}$. Remark that, by definition of $\tilde{\theta}$, $u(\tilde{\theta}) = v(\tilde{\theta})$, where $u(\theta) \equiv U(x(\theta), y(\theta), q^f + q^s, \theta)$ and $v(\theta) \equiv V(x(\theta), y(\theta), q^f, \theta)$.

Given $u(\theta)$ and $v(\theta)$, and the gross income $y(\theta)$, the consumption level $x(\theta)$ can be substituted as follows:

$$x^U \equiv \{x \mid u(\theta) = U(x, y(\theta), q^f + q^s, \theta)\} \tag{15}$$

for $\theta \leq \tilde{\theta}$ (where $x_y^U = \frac{dx}{dy} \mid_U = -\frac{U_y}{U_x}$; $x_u^U = \frac{1}{U_x}$; $x_{q^f}^U = x_{q^f}^U = -\frac{U_q}{U_x}$), and

$$x^V \equiv \{x \mid v(\theta) = V(x, y(\theta), q^f, \theta)\} \tag{16}$$

for $\theta > \tilde{\theta}$ (where $x_y^V = \frac{dx}{dy} \mid_V = -\frac{V_y}{V_x}$; $x_v^V = \frac{1}{V_x}$; $x_{q^f}^V = -\frac{V_{q^f}}{V_x}$).

As usual, we treat $u(\theta)$, $v(\theta)$, $y(\theta)$,

$$R(\theta) = \int_{\underline{\theta}}^{\theta} [y(\theta') - x^U - q^f - q^s \cdot F(\tilde{\theta})] \cdot dF(\theta') \tag{17}$$

(for $\theta \leq \tilde{\theta}$), and

$$R(\theta) = \int_{\underline{\theta}}^{\theta} [y(\theta') - x^V - q^f - q^s \cdot F(\tilde{\theta})] \cdot dF(\theta') \tag{18}$$

(for $\theta > \tilde{\theta}$) as state variables with $\nu(\theta)$, $\gamma(\theta)$, and $\mu(\theta)$, the corresponding co-state variables. The control variables are the change of gross income with respect to households' capability, $\eta(\theta)$, and the first- and second-pillar public provision, q^f and q^s respectively. The government's optimal control problem is:

$$\begin{aligned}
\max \int_{\underline{\theta}}^{\tilde{\theta}} u(\theta) \cdot dF(\theta) + \int_{\tilde{\theta}}^{\overline{\theta}} v(\theta) \cdot dF(\theta) \\
s.t.: \\
\frac{dR}{d\theta} = [y(\theta) - x^U - q^f - q^s \cdot F(\tilde{\theta})] \cdot f(\theta) \quad (\mu(\theta)) \quad \forall \theta \leq \tilde{\theta} \\
\frac{dR}{d\theta} = [y(\theta) - x^V - q^f - q^s \cdot F(\tilde{\theta})] \cdot f(\theta) \quad (\mu(\theta)) \quad \forall \theta > \tilde{\theta} \\
\frac{du}{d\theta} = U_\theta \quad (\nu(\theta)) \quad \forall \theta \leq \tilde{\theta} \\
\frac{dv}{d\theta} = V_\theta \quad (\nu(\theta)) \quad \forall \theta > \tilde{\theta} \\
\frac{dy}{d\theta} = \eta(\theta) \quad (\gamma(\theta)) \\
\frac{dy}{d\theta} = \eta(\theta) \quad (\gamma(\theta)) \\
\eta(\theta) \geq 0 \quad (\sigma(\theta)) \\
q^f \cdot f(\theta) \geq 0 \quad (\phi(\theta)) \\
q^s \cdot f(\theta) \geq 0 \quad (\varphi(\theta)) \\
R(\underline{\theta}) = 0 \quad R(\overline{\theta}) = 0 \quad \nu(\underline{\theta}) = 0 \quad \nu(\overline{\theta}) = 0 \quad \gamma(\underline{\theta}) = 0 \quad \gamma(\overline{\theta}) = 0
\end{aligned} \tag{19}$$

By the optimization of (19), in the no-bunching case (i.e., when the marginal tax rate is different for households with different capabilities), we find the usual-shape marginal tax rate:

$$T'(y(\theta)) = 1 - \frac{dx}{dy} = (\frac{E_\theta(\mathcal{U}_c)}{E(\mathcal{U}_c)} - F(\theta)) \cdot (-\frac{\partial}{\partial\theta}\frac{dx}{dy}) \geq 0 \tag{20}$$

for all $\theta \in [\underline{\theta}, \overline{\theta}]$.[6]

We also have:

Proposition 3 *If input and capability are economic complements, no public provision is socially desirable.*

Proof. See Appendix A. ■

The economic intuition behind Proposition 3 is the following. When θ and q are economic complements, they are also strong technological complements. Thus, the public provision of q would especially favor higher-capability households. Since the marginal utility of consumption is decreasing, the growth of the earning capacity of higher-income households would require larger redistributive taxation. But, this would increase the distortion of labor supply.

The continuity of households' inherited capabilities plays a crucial role in this result. In fact, when households are characterized by a limited number of types (as in the two-class economy we analyzed in Section 3), and there is enough scope for the second pillar public provision, the government can always discriminate between two contiguous types. In other words, the policy-maker can exploit the complementarity between θ and q to provide the productive input only to low-capability households without distorting the choices of the high-capability ones.

We now consider the case in which input and capability are economic substitutes, $\tilde{\theta}_{q^s} < 0$ (i.e., $U_{\theta q} < \min\left\{0, U_{\theta x} \cdot \frac{Q^m - q^s}{Q^m}\right\}$), hence households with lower capability demand

[6] Remark that: $\mathcal{U}_c = U_x$ if $\theta \leq \tilde{\theta}$ and $\mathcal{U}_c = V_x$ if $\theta > \tilde{\theta}$; and $E_\theta(\mathcal{U}_c) = \int_{\underline{\theta}}^{\theta} \mathcal{U}_c \cdot f(\theta') \cdot d\theta'$, hence $E(\mathcal{U}_c) = E_{\overline{\theta}}(\mathcal{U}_c)$. It can be shown that $\mu = \mu(\theta) = E(\mathcal{U}_c)$ for all θ, and $\nu(\theta) = -\frac{E(\mathcal{U}_c)}{\mathcal{U}_c} \cdot (\frac{E_\theta(\mathcal{U}_c)}{E(\mathcal{U}_c)} - F(\theta)) \leq 0$. Therefore, as usual, $T'(y(\underline{\theta})) = T'(y(\overline{\theta})) = 0$.

more input than those with higher capability. Following the same procedure we considered above, the government's optimal control problem can be written as:

$$
\begin{aligned}
\max \int_{\underline{\theta}}^{\tilde{\theta}} v(\theta) \cdot dF(\theta) + \int_{\tilde{\theta}}^{\overline{\theta}} u(\theta) \cdot dF(\theta) \\
s.t.: \\
\frac{dR}{d\theta} = [y(\theta) - x^U - q^f - q^s \cdot (1 - F(\tilde{\theta}))] \cdot f(\theta) \quad (\mu(\theta)) \quad \forall \theta \geq \tilde{\theta} \\
\frac{dR}{d\theta} = [y(\theta) - x^V - q^f - q^s \cdot (1 - F(\tilde{\theta}))] \cdot f(\theta) \quad (\mu(\theta)) \quad \forall \theta < \tilde{\theta} \\
\frac{du}{d\theta} = U_\theta \quad (\nu(\theta)) \quad \forall \theta \geq \tilde{\theta} \\
\frac{dv}{d\theta} = V_\theta \quad (\nu(\theta)) \quad \forall \theta < \tilde{\theta} \\
\frac{dy}{d\theta} = \eta(\theta) \quad (\gamma(\theta)) \\
\eta(\theta) \geq 0 \quad (\sigma(\theta)) \\
q^f \cdot f(\theta) \geq 0 \quad (\phi(\theta)) \\
q^s \cdot f(\theta) \geq 0 \quad (\varphi(\theta)) \\
R(\underline{\theta}) = 0 \quad R(\overline{\theta}) = 0 \quad \nu(\underline{\theta}) = 0 \quad \nu(\overline{\theta}) = 0 \quad \gamma(\underline{\theta}) = 0 \quad \gamma(\overline{\theta}) = 0
\end{aligned}
\tag{21}
$$

By the optimization conditions of this problem, we find the usual marginal tax schedule. Moreover, we have:

Proposition 4 *If input and capability are economic substitutes, an inclusive regime where all households opt for the second pillar is socially desirable.*

Proof. See Appendix A. ■

The intuition is the following. When θ and q are economic substitutes, the public provision of q would especially favor households with lower capability. Since the marginal utility of consumption is decreasing, the growth of the earning capacity of lower-income households would facilitate redistribution and reduce tax distortions. In this case, the public provision becomes an extremely effective redistributive tool.

4.1 Numerical Analysis

In this section, we first numerically characterize the assumptions about primitives (i.e., preferences and technology) that may determine complementarity or substitutability between input investment and households' inherited capability. Second, we numerically investigate the way public provision affects (the level of) marginal tax rates and overall redistribution (including both taxation and the publicly-provided input) across households with different inherited capability.

For the sake of comparability with the literature on numerical simulations of optimal tax schedules, we rely on additive utility functions. In particular, we consider two alternative specifications. A classic, linear-logarithmic function (Mirrlees, 1971; Tuomala, 1984),

$$\mathcal{U}(c,l) = \log(c) + \log(1-l); \tag{22}$$

and a quasi-linear function (Atkinson, 1990; Diamond, 1998),

$$\mathcal{U}(c,l) = G(c + \log(1-l)). \tag{23}$$

where $G(\cdot) = \log(\cdot)$ implies that the marginal utility of consumption is decreasing.[7] As elsewhere in the literature (Saez, 2001), comparing these specifications let us to investigate the role of income effects.[8]

For the purpose of our analysis, an appropriate wage function, $w(\theta, q)$, is technologically

[7] An alternative interpretation is that $G(\cdot)$ reflects the government's distributional preferences. The specification of the argument of $G(\cdot)$ is linear in consumption and strictly concave in leisure. Some studies have also used a specification of the utility that is linear in leisure (Lollivier and Rochet, 1983; Weymark, 1987). However, such functions may easily lead to bunching problems (Ebert, 1992).

[8] Both specifications imply that the government is averse to inequality, and both have their limitations. In particular, the linear-logarithmic function implies a unitary elasticity of substitution between leisure and consumption, while the latter removes any income effect on households' decisions.

separable. In particular, we adopt an additive wage function with diminishing marginal products of inputs: $w(\theta, q) = \theta^{\alpha} + q^{\beta}$, with α and β less than one. As baseline values, we assume $\alpha = \beta = 0.5$. The considered specification affords sufficient flexibility to generate both cases of economic complementarity and substitutability between input and capability, depending only on the presence or absence of the income effects in the utility function.

Inherited capabilities are largely unobservable, exogenous productive skills of households. In our setting we consider a specification such that higher capability implies higher gross income. Given that the lognormal distribution is a benchmark in the literature on income taxation (since it fits well the unimodality of empirical distributions), we keep the same assumption for exogenous capabilities. However, with respect to other studies, we adopt a truncated lognormal distribution, excluding the possibility of households characterized by infinite capability.[9] This approach has two advantages. First, it allows us to include top-income (i.e., top-capability) taxpayers in the public budget balance. Second, it guarantees that, under different policies, transversality conditions equally hold.[10]

Figure 1 shows the private demand of input in absence of any public provision.[11] The solid line is obtained by using utility (22), whereas the dashed line corresponds to the quasi-linear utility (23). Wee see that, given the wage function, when the income effects are taken into account (i.e., with linear-logarithmic utility), q and θ are economic substitutes, while if they are assumed away (i.e., in the quasi-linear case) q and θ become economic complements.

The Table 1 contrasts the numerical simulations of the household's choice variables - net transfer received, marginal tax rate (MRT), and household's utility - for each decile of $F(\theta)$ in the two cases of the (best) policy under PT regime (upper section of Table 1,

[9] Here, the logarithm of θ is normally distributed with mean equal to 2 and standard deviation equal to 1 in the support $[0, 100]$. This means that the 90 percent of the distribution lies below a value of θ equal to 26. In this way, we take into account the presence of outliers.

[10] This second argument is particularly relevant when we compare various tax schedules, since an imprecise calibration may lead to asymptotic violations of the transversality conditions.

[11] All calculations are made by Mathematica, version 8. The algorithm used to solve the maximization problem is described in the Appendix C.

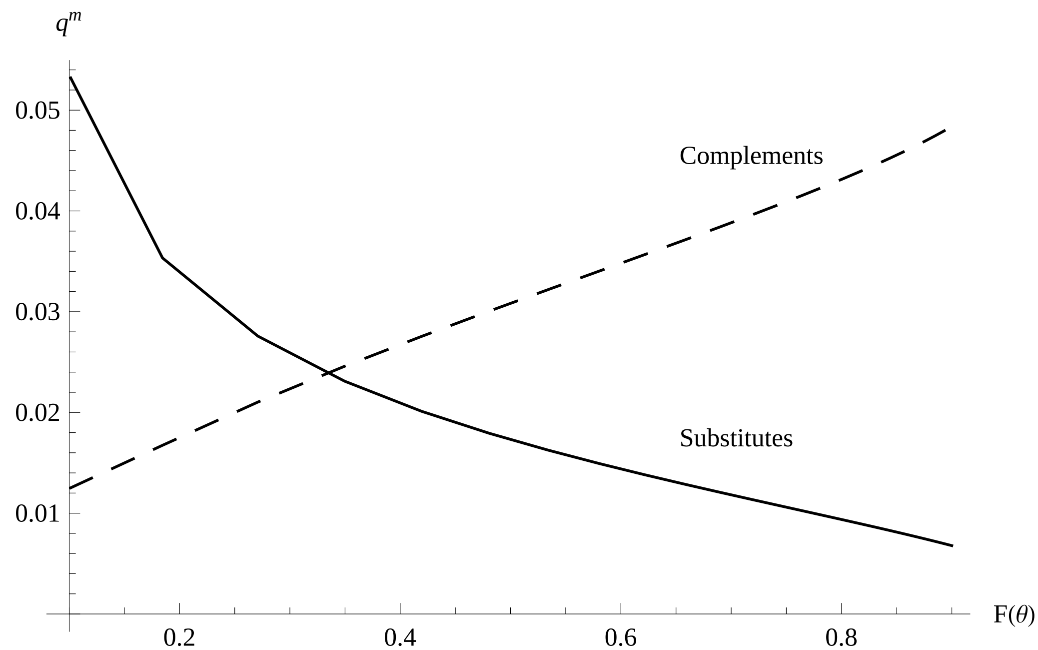

Figure 1: Private input investment depending on household's capability

PT_S) and of the (best) policy under IN regime (lower section of Table 1, IN_S), when the income effects are taken into account (i.e., input and capability are economic substitutes). Consistent to our theoretical prediction (see Proposition 4), the optimal redistribution scheme is the IN regime (i.e., $W_{IN_S} > W_{PT_S}$, the asterisk denotes the optimal policy). Moreover, all households are better off under such regime (see the last column of Table 1).

Table 1: Economic substitutes

Pure Taxation ($W_{PT_S}= -0.329$)							
$F(\theta)$	q^m	$l(\theta)$	$y(\theta)$	$x(\theta)$	$Net\ Tr.$	MTR	$utility$
0.1	0.053	0.431	0.710	0.805	0.096	0.198	−0.850
0.2	0.035	0.419	0.805	0.902	0.097	0.223	−0.686
0.3	0.028	0.418	0.906	0.992	0.086	0.234	−0.578
0.4	0.020	0.424	1.098	1.152	0.054	0.242	−0.427
0.5	0.018	0.427	1.186	1.224	0.038	0.243	−0.369
0.6	0.015	0.433	1.351	1.356	0.005	0.243	−0.273
0.7	0.012	0.440	1.573	1.532	−0.041	0.240	−0.161
0.8	0.009	0.449	1.896	1.787	−0.109	0.235	−0.021
0.9	0.007	0.460	2.386	2.174	−0.212	0.225	0.156
*Inclusive Regime ($W_{IN_S}= -0.321$ and $q^f+q^s= 0.054$)							
$F(\theta)$	q^m	$l(\theta)$	$y(\theta)$	$x(\theta)$	$Net\ Tr.$	MTR	$utility$
0.1	–	0.428	0.704	0.756	0.106	0.198	−0.838
0.2	–	0.423	0.831	0.881	0.103	0.223	−0.677
0.3	–	0.425	0.949	0.984	0.089	0.233	−0.569
0.4	–	0.432	1.159	1.158	0.054	0.239	−0.419
0.5	–	0.435	1.253	1.235	0.036	0.240	−0.361
0.6	–	0.441	1.427	1.373	0.001	0.240	−0.265
0.7	–	0.448	1.657	1.556	−0.048	0.237	−0.153
0.8	–	0.457	1.990	1.818	−0.118	0.231	−0.013
0.9	–	0.467	2.491	2.212	−0.224	0.221	0.164

$Net\ Tr.$ (Net Transfer): $x(\theta) - y(\theta) + (q^f + q^s)$

In both cases, the marginal tax rates increase (or do not decrease) with income and households receive a net transfer from the government for the first six deciles of the distribution of θ. Remark that under the IN regime, the net transfer is the sum of the cash transfer determined by the tax schedule and by the (monetary value) of the publicly-provided input. The first three deciles of the population receive a higher transfer under the IN regime, with respect to the PT regime. In contrast, the cost of the public provision is borne by the last six deciles of the population. Although the PT regime affords a higher disposable income to the poorest 30% of the population, part of this income has to be spent to purchase q^m. Hence, private consumption is always lower under the PT regime than under the IN regime.

Though richer households pay more taxes under the IN regime, marginal tax rates are lower for all households with respect to the PT regime (see Figure 2).[12] The economic intuition behind this result is the following. First, the public provision of q keeps high the household's wage and helps to control tax distortions on labor decisions (see Figure 3). Second, the income effects determined by the public provision tends to increase the sensitivity of the labor supply to marginal tax rates.

Consistent to our prediction (see Proposition 3), when the income effects are assumed away (hence, input and capability are economic complements), the best policy is the optimal one under the PT regime (upper section of Table 2), contrasted to the best possible policy under the IN regime (lower section of Table 2). However, in this case only households belonging to the first three deciles of the distribution are better off. As in Table 1, about 60% of the population receives a net transfer from the government. But, the net transfer and the marginal tax rates are now always larger under PT than under IN regime. Also, all households are forced to work more under IN than under PT regime. The intuition of why the IN regime is now sub-optimal is the following. Absent any income effect, the public provision of q pushes up the wage rate, the incentive to work, and the optimal marginal tax rates. And, this effect is particularly strong for the poorest households (also taking into account the substantial difference in the level of the public

[12] Our results are consistent with those obtained in previous studies (e.g., Mirrlees (1971, p. 202)): the marginal tax schedule is hump-shaped; in particular, it decreases for the upper tail of the distribution; as predicted by the theoretical analysis, marginal tax rates are zero for both at $\underline{\theta}$ and $\overline{\theta}$.

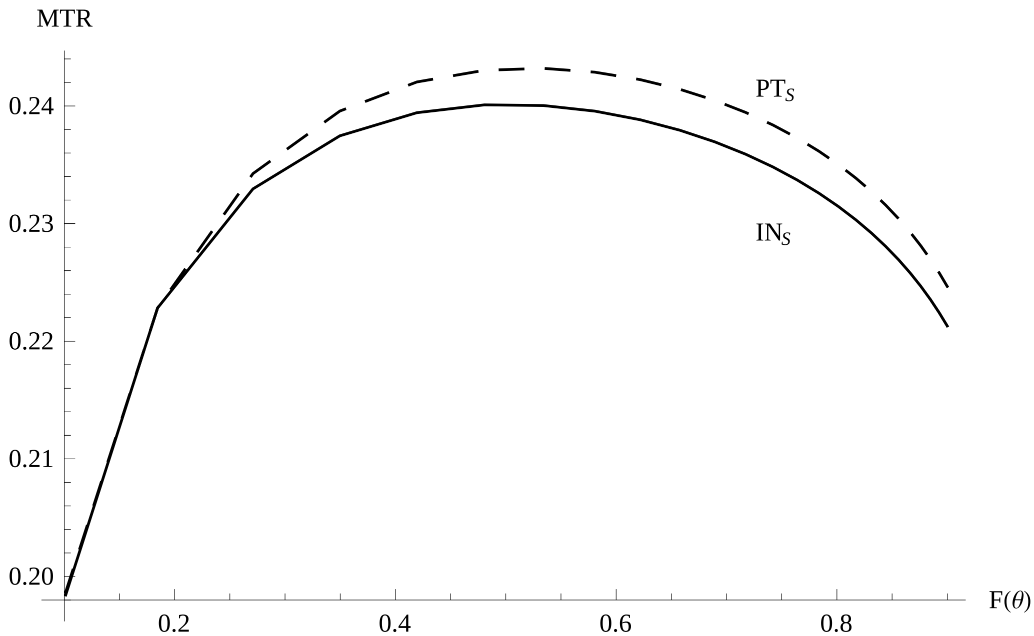

Figure 2: Marginal tax rates under PT_S (dashed line) and IN_S (solid line)

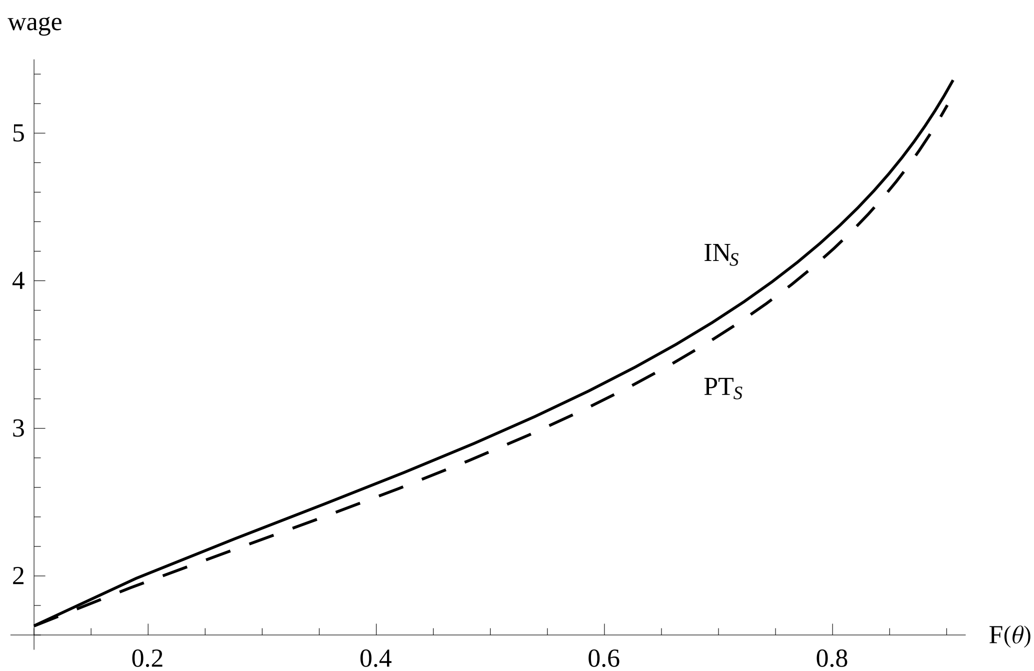

Figure 3: Wages under PT_S (dashed line) and IN_S (solid line)

provision in this case with respect to Table 1). In other terms, the most socially relevant part of the population is worse off under the IN regime, contrasted to PT regime.[13]

To complete our analysis, we test the robustness of the results against alternative specifications of the technology (i.e., the productivity parameters characterizing the wage function). Table 3 shows the optimal redistribution policy (i.e., marginal tax rates and, possibly, public provision of input) under the alternative specifications of household's preferences (i.e., IN_S regime, when input and capability are substitutes, and PT_C, when input and capability are complements). Starting from our baseline parametrization (i.e., $\alpha = 0.5$ and $\beta = 0.5$), we first reduce the productivity parameter of θ and then the productivity parameter of q (in both cases, the parameter of the other argument of the wage function is kept constant). For both specifications, the lower the productivity of θ the lower the marginal tax rates. In particular, when inputs are economic substitutes, a reduction of α leads to a significant increase in the public provision of q and to a flatter taxation. In other terms, when households' heterogeneity (i.e., inherited capability) plays a minor role in the determination of wages, the government tends to rely more on the public provision of input as redistributive device, contrasted to taxation. Similarly, when inputs are economic complements, a reduction of α tends to reduce inherited inequality and, therefore, the need for distorting taxation (column 5 of Table 3).

[13] In line with previous studies (e.g., Tuomala (2010)), the lack of income effects - hence, the stronger responsiveness of labor supply to wages and taxes - also implies increasing marginal tax rates over a wider range of capability levels (even with lognormal distribution of capabilities). Except for the first deciles of the distribution of capabilities, the marginal tax rates are higher than before, reaching a maximum value at the end of the distribution. In other terms, the marginal tax rate increases for a large fraction of the population even if the asymptotic result of a zero marginal tax rate still holds.

Table 2: Economic complements

*Pure Taxation (W_{PT_C}= -0.074)							
$F(\theta)$	q^m	$l(\theta)$	$y(\theta)$	$x(\theta)$	$Net\ Tr.$	MTR	$utility$
0.1	0.012	0.252	0.385	0.791	0.406	0.123	0.488
0.2	0.017	0.325	0.605	0.976	0.371	0.204	0.566
0.3	0.021	0.384	0.823	1.145	0.323	0.244	0.641
0.4	0.028	0.465	1.216	1.436	0.220	0.286	0.783
0.5	0.030	0.494	1.393	1.563	0.169	0.299	0.851
0.6	0.034	0.540	1.721	1.791	0.071	0.317	0.980
0.7	0.038	0.589	2.156	2.087	-0.069	0.335	1.159
0.8	0.043	0.643	2.784	2.502	-0.283	0.353	1.429
0.9	0.048	0.701	3.726	3.105	-0.621	0.372	1.851
Inclusive Regime (W_{IN_C}= -0.079 and q^f+q^s= 0.127)							
$F(\theta)$	q^m	$l(\theta)$	$y(\theta)$	$x(\theta)$	$Net\ Tr.$	MTR	$utility$
0.1	–	0.344	0.610	0.870	0.388	0.138	0.448
0.2	–	0.382	0.797	1.034	0.364	0.225	0.553
0.3	–	0.426	1.003	1.193	0.317	0.261	0.638
0.4	–	0.493	1.383	1.470	0.215	0.297	0.791
0.5	–	0.518	1.556	1.592	0.164	0.308	0.862
0.6	–	0.559	1.876	1.813	0.065	0.324	0.995
0.7	–	0.603	2.304	2.102	-0.075	0.340	1.178
0.8	–	0.653	2.924	2.508	-0.288	0.357	1.451
0.9	–	0.707	3.858	3.104	-0.626	0.374	1.876

$Net\ Tr.$ (Net Transfer): $x(\theta) - y(\theta) + (q^f + q^s)$

Columns 3 and 6 of Table 3 show the effects of the productivity of q on optimal policies. When the income effects are taken into account (i.e., θ and q are substitutes), a reduction of β reduces the marginal tax rates for all households. But, the marginal tax schedule (hence, the public provision of q) is less sensitive to β than to α (contrast of column 3 and 2 of Table 3). When the income effects are assumed away (i.e., θ and q are complements), an increase of β increases the marginal tax rates for low-capability households, and decreases them for high-capability households. Given that the productivity of input investment drops, it is less relevant for the government to control the tax distortion (on private input investment) that hits low-capability households.

Table 3: Sensitivity analysis on optimal policies (MTR)

	Economic substitutes (IN_S regime)			Economic complements (PT_C regime)		
	(1)	(2)	(3)	(4)	(5)	(6)
$F(\theta)$	*Benchmark*	$\alpha = 0.4$	$\beta = 0.4$	*Benchmark*	$\alpha = 0.4$	$\beta = 0.4$
0.1	0.198	0.161	0.189	0.123	0.105	0.126
0.2	0.223	0.174	0.214	0.204	0.173	0.208
0.3	0.233	0.179	0.224	0.244	0.204	0.246
0.4	0.239	0.181	0.231	0.286	0.234	0.286
0.5	0.240	0.181	0.232	0.299	0.243	0.298
0.6	0.240	0.180	0.232	0.317	0.255	0.316
0.7	0.237	0.177	0.230	0.335	0.265	0.333
0.8	0.231	0.172	0.226	0.353	0.275	0.351
0.9	0.221	0.164	0.216	0.372	0.283	0.370
$q^f + q^s$	0.054	0.081	0.058	–	–	–

5 Conclusions

In this paper we analyzed the optimal redistribution policies carried out by a benevolent government in economies where households' gross earnings are unobservable and driven by inherited capabilities and endogenous investment in input (e.g., education). The two policy tools we considered in our analysis are a non-linear income tax and a public program for the provision of input. Without loss of generality, we assumed that the latter program is made by two pillars. A first pillar provides a uniform amount of input to all households. Then, each household may decide to top up the basic public provision with additional input bought on the market or it may opt for a second pillar public provision.

In a two-class economy, the public provision is welfare improving (with respect to pure non-linear income taxation) only if the low-capability households opt for the public program for lower levels of the publicly-provided input than the high-capability. In particular, when the scope for the second-pillar provision is sufficiently wide, the optimal discriminating policy is constituted by a pure opting-out mechanism. In contrast, when the scope for the second pillar is limited, the opting-out mechanism is desirable only when the two factors of production are economic complements. In this case, high-capability households have no incentives to mimic low-type ones. Vice versa, when inherited capability and input are economic substitutes, a pure topping-up scheme is optimal.

In economies with a continuum of households' types (i.e., households' inherited capabilities), the optimal redistribution scheme is affected by the impossibility to discriminate among contiguous types (as we do in the two-class economy). In particular, when capability and input are economic substitutes, an inclusive regime in which all households opt for the publicly-provided input is optimal. And conversely when input and capability are economic complements, taxation is the only redistributive tool that a government should consider.

Finally, the numerical analysis of the economy with continuous capability levels showed that, when the wage function is assumed to be linear (hence, input and capability are not technologically complements nor substitutes), the existence of the income effects in the utility function specification drives the economic relationship between input and capability, hence the optimal policy mix. In particular, when the income effects are taken into consideration input and capability become economic substitutes and the optimal redistribution scheme relies both on taxation and the public provision of input.

References

Atkinson, A. (1990). Public Economics and the Economic Public. *European Economic Review*, 34(2-3):225–248.

Balestrino, A. (1999). The Desirability of In-Kind Transfers in the Presence of Distortionary Taxes. *Journal of Economic Surveys*, 13(4):333–354.

Balestrino, A. (2000). Public Provision of Private Goods and the Redistribution of Income. *International Tax and Public Finance*, 7:463–478.

Blomquist, S. and Christiansen, V. (1995). Public Provision of Private Goods as Redistributive Device in an Optimum Income Tax Mode. *Scandinavian Journal of Economics*, 97:547–567.

Blomquist, S. and Christiansen, V. (1998a). Price Subsidies versus Public Provision. *International Tax and Public Finance*, 5:283–306.

Blomquist, S. and Christiansen, V. (1998b). Topping up or Opting out? The Optimal Design of Public Provision Schemes. *International Economic Review*, 39:399–411.

Blomquist, S., Christiansen, V., and Micheletto, L. (2010). Public Provision of Private Goods and Nondistortionary Marginal Tax Rates. *American Economic Journal: Economic Policy*, 2:1–27.

Boadway, R. and Marchand, M. (1995). The Use of Public Expenditures for Redistributive Purposes. *Oxford Economic Papers*, 47:45–59.

Cingano, F. (2014). Trends in Income Inequality and its Impact on Economic Growth. OECD Social, Employment and Migration Working Paper no. 163.

Corak, M. (2013). Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives*, 27(3):79–102.

Cremer, H. and Gahvari, F. (1997). In-Kind Transfers, Self-Selection and Optimal Tax Policy. *European Economic Review*, 41:97–114.

Diamond, P. (1998). Optimal Income Taxation. *The American Economic Review*, 88(1):83–95.

Ebert, U. (1992). A Reexamination of the Optimal Nonlinear Income Tax. *Journal of Public Economics*, 49:47–73.

Greco, L. (2011). Optimal Redistribution with Productive Social Services. *Scandinavian Journal of Economics*, 113(1):55–73.

Lollivier, S. and Rochet, J.-C. (1983). Bunching and Second Order Conditions: A Note on Optimal Tax Theory. *Journal of Economic Theory*, 31:392–400.

Mirrlees, J. A. (1971). An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38(2):175–208.

OECD (2011). *Divided We Stand: Why Inequality Keeps Rising*. OECD Publishing.

OECD (2014). Focus on Inequality and Growth - December 2014. http://www.oecd.org/social/inequality-and-poverty.htm.

Piketty, T. (2013). *Le Capital au XXI^e Siècle*. Èdition du Seuil, Paris.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rate. *The Review of Economic Studies*, 68:205–229.

Tuomala, M. (1984). On the Optimal Income Taxation. *Journal of Public Economics*, 23:351–366.

Tuomala, M. (2010). On Optimal Non-linear Income Taxation: Numerical Results Revisited. *International Tax and Public Finance*, 17:259–270.

Weymark, J. (1987). Comparative Static Properties of Optimal Nonlinear Income Taxes. *Econometrica*, 55:1165–1185.

Appendix A: Proofs of the Main Propositions

Proof of Proposition 1. As regards (1), substituting the first order condition w.r.t. $\underline{x}$ and $\overline{\varphi} = \underline{\varphi} = 0$ in the first order condition w.r.t. q^s, we obtain $\underline{FOC} = \frac{\frac{\nu}{\lambda}}{1-\frac{\nu}{\lambda}} \cdot (\widehat{FOC} - \underline{FOC})$, where $\underline{FOC} = -\underline{U}_x + \underline{U}_q$ and $\widehat{FOC} = -\widehat{U}_x + \widehat{U}_q$. By comparative statics on household's optimization problem, $FOC_\theta < 0$ and $\widehat{FOC} - \underline{FOC} < 0$ (or $FOC_\theta > 0$ and $\widehat{FOC} - \underline{FOC} > 0$), when capability and input are substitutes (or complements). By the first order conditions w.r.t. $\underline{x}$ and $\underline{y}$, $\nu = \lambda \cdot \frac{\underline{U}_x + \underline{U}_y}{\widehat{U}_x + \widehat{U}_y}$. Since the private consumption is a normal good, $\underline{U}_x + \underline{U}_y < \widehat{U}_x + \widehat{U}_y$ and $\nu < \lambda$. Therefore, the (public) provision of input is above, i.e., $\underline{FOC} < 0$ (or below, i.e., $\underline{FOC} > 0$) the optimal level that low-capability households would choose on their own when capability and input are substitutes (or complements). Let $q^{f*} > 0$ and $q^{s*} \geq 0$ be an optimal public provision mix. Consider a marginal perturbation of it such that: the first pillar provision is decreased by $dq^f < 0$, the second-pillar provision is increased by $dq^s = -dq^f > 0$, and the net-of-tax income of high-capability households is increased by $d\overline{x} = -dq^f > 0$. Such a marginal perturbation leaves unaffected: the total amount of input that is publicly provided to low-capability households (hence, their welfare), the welfare and optimal behavior of high-capability households, and the government budget balance. By iterating this argument, any optimal public provision mix is equivalent (in social welfare terms) to a pure opting-out scheme.

The result (2) is easily obtained by inspection of conditions (6) and (7), contrasted with corresponding conditions obtained under the PT regime. ■

Proof of Proposition 2. At the optimum, $\overline{\varphi} = 0$. Assume conversely that $\overline{\varphi} > 0$ (hence $\underline{\varphi} = 0$), then necessarily $q^s = \overline{Q}^s$, hence $-\overline{U}_x + \overline{U}_q \geq 0$, by construction of $\overline{Q}^s$. By the first order conditions w.r.t. $\overline{x}$ and q^s,

$$-\overline{U}_x + \overline{U}_q = \frac{\underline{\varphi} - \overline{\varphi} \cdot (1 + \underline{Q}^s_x)}{1 - \lambda + \nu} < 0,$$

which implies a contradiction. Therefore, the marginal tax rate on high-capability households is zero. Moreover, at the optimum, also $\underline{\varphi} = 0$. Assume by contradiction that $\underline{\varphi} > 0$, then necessarily $q^s = \underline{Q}^s$. By the first order conditions w.r.t. $\overline{x}$ and q^s, this would imply that $-\overline{U}_x + \overline{U}_q > 0$. Let q^f and q^s be the optimal first- and second-pillar provision at the optimum such that $\underline{\varphi} > 0$. Such a policy mix cannot be optimal. Consider the following marginal policy reform: the first-pillar public provision increases by $dq^f > 0$; the tax on low-capability households is decreased lump-sum accordingly, $\frac{d\underline{x}}{dq^f} = -1$; the second-pillar public provision is modified to satisfy the condition $q^s = \underline{Q}^s$, hence the total (first- and second-pillar) provision for the high-capability households change is $1+\underline{Q}^s_{q^f} > 0$, and the tax on high-capability households is increased (lump sum) by the same amount (i.e., $\frac{d\overline{x}}{dq^f} = -(1 + \underline{Q}^s_{q^f}) < 0$). The considered policy reform keeps the equilibrium of government's budget, and increases the total welfare:

$$
\begin{aligned}
\lambda \cdot (\underline{V}_x \cdot \frac{d\underline{x}}{dq^f} + \underline{V}_{q^f}) + (1-\lambda) \cdot (\overline{U}_x \cdot \frac{d\overline{x}}{dq^f} + \overline{U}_q \cdot (1 + \underline{Q}^s_{q^f})) &= \\
= \lambda \cdot (\underbrace{-\underline{V}_x + \underline{V}_{q^f}}_{=0}) + (1-\lambda) \cdot (\underbrace{-\overline{U}_x + \overline{U}_q}_{>0}) \cdot (1 + \overline{Q}^s_{q^f}).
\end{aligned}
$$

Hence, we have a contradiction. The optimal policy mix necessarily requires $-\overline{U}_x+\overline{U}_q = 0$, then by the first order conditions w.r.t. $\underline{x}$, $\overline{x}$ and q^f we also have that

$$
(\underline{\varphi} + \overline{\phi}) \cdot (\underline{Q}^s_x - \underline{Q}^s_{q^f}) + \overline{\varphi} \cdot \overline{q}^s_{q^f} - \underline{\phi} = 0
$$

hence, by $\underline{\varphi} = \overline{\varphi} = 0$ necessarily $\underline{\phi} = \overline{\phi} = 0$. Therefore, the marginal tax wedge on low-capability households is under the PT regime.

Remark that the optimal public provision mix is equivalent to what high-capability households would do on their own in absence of any public provision. ■

Proof of Proposition 3. Assume that the optimal solution is such that $q^s > 0$. If

$\tilde{\theta} \in (\underline{\theta}, \overline{\theta}]$, the first order condition w.r.t. q^s for all $\theta \in (\tilde{\theta}, \overline{\theta}]$ would imply $-(F(\tilde{\theta}) + q^s \cdot f(\tilde{\theta}) \cdot \tilde{\theta}_{q^s}) = 0$; given that $\tilde{\theta}_{q^s} > 0$ when input and capability are economic complements, then necessarily $\tilde{\theta} = \underline{\theta}$ and $q^s = 0$ and a contradiction arises. If $\tilde{\theta} = \overline{\theta}$, there is an inclusive regime (i.e., all households opt for the second pillar provision); considering the transversality condition $\nu(\overline{\theta}) = 0$ and that $-x^U_{q^s} - 1 = \frac{-U_x + U_q}{Ux}$, the first order condition w.r.t. q^s for $\theta = \overline{\theta}$ becomes $E(\mathcal{U}_c) \cdot \frac{-U_x + U_q}{Ux} \cdot f(\overline{\theta}) = 0$, thus $q^s > 0$ implies that $-U_x + U_q = 0$ for $\theta = \overline{\theta}$ (i.e., $q^f + q^s$ is the optimal level of q for top-capability households). Given that, under complementarity between input and capability, the optimal level of q increases with θ, for $\theta < \overline{\theta}$, it necessarily follows that $-U_x + U_q < 0$. By complementarity between input and capability, i.e., $U_{\theta q} > 0$, we have a contradiction given that, by the first order condition w.r.t. q^s, $E(\mathcal{U}_c) \cdot \frac{-U_x + U_q}{U_q} \cdot f(\theta) + \nu(\theta) \cdot U_{\theta q} = 0$. Hence, at the optimum $q^s = 0$. Assume now that the optimal solution is such that $q^f > 0$. If the public provision is sufficiently high that $\tilde{\theta} = \overline{\theta}$, then the same argument we made for q^s bring us to show that there is a contradiction. If $\tilde{\theta} < \overline{\theta}$, for all $\theta > \tilde{\theta}$ (i.e., opting-out households) the first order condition w.r.t. q^f becomes $\nu(\theta) \cdot V_{\theta q^f} = 0$, that implies a contradiction unless $V_{\theta q^f} = 0$ for all θ, hence at the optimum also $q^f = 0$. ■

Proof of Proposition 4. Assume that the optimal public provision is not inclusive, i.e., $\tilde{\theta} < \overline{\theta}$. If $q^s > 0$, the first order condition w.r.t. q^s for all $\theta > \tilde{\theta}$ can be written as $-E(\mathcal{U}_c) \cdot (1 - F(\tilde{\theta}) - q^s \cdot f(\tilde{\theta}) \cdot \tilde{\theta}_{q^s}) \cdot f(\theta) = 0$. Remark that, when input and capability are economic substitutes, $\tilde{\theta}_{q^s} < 0$, then a contradiction arises, and necessarily $q^s = 0$. If $q^f \geq 0$, the first order condition w.r.t. q^f for all $\theta > \tilde{\theta}$ can be written as $-E(\mathcal{U}_c) \cdot (x_{q^f} + 1) \cdot f(\theta) + \nu(\theta) \cdot V_{\theta q^f} = 0$. Remark that for opting-out households $x_{q^f} + 1 = \frac{-V_x + V_{q^f}}{V_x} = 0$ and a contradiction arises, unless $V_{\theta q^f} = 0$ for all θ. Hence, no public provision pattern (including in particular $q^f = q^s = 0$) is compatible with $\tilde{\theta} < \overline{\theta}$. Thus, by inspection of optimization conditions, we conclude that at the optimum the public provision pattern implies $\tilde{\theta} = \overline{\theta}$. ■

Appendix B: Corner Solutions in a Two-Class Economy

Proposition 5 *Under the DL regime, if capability and input are economic substitutes, the corner solution of the problem (5) is such that:*

1. *the provision mix is pure topping-up (i.e., $q^{f*} > 0$ and $q^{s*} = 0$), and such that $\overline{Q}^s = \overline{Q}^m = 0$;*

2. *at the margin, the tax schedule subsidizes high-capability labor supply and taxes (with the usual shape of tax wedge) low-capability labor supply.*

Proof. By substituting the first order conditions w.r.t. $\overline{x}$ and q^f in the first order condition w.r.t. q^s, observing that $\overline{V}_x = \overline{V}_{q^f}$, and by definition of $\overline{Q}^s_x$ and $\overline{Q}^s_{q^f}$, we get

$$\overline{\varphi} \cdot \underbrace{(1 - \overline{Q}^s_x + \overline{Q}^s_{q^f})}_{>0} + \underline{\phi} = \underline{\varphi} \cdot \underbrace{(1 + \underline{Q}^s_{q^f})}_{>0} + \overline{\phi} \cdot \underbrace{(\overline{Q}^s_x - \overline{Q}^s_{q^f})}_{>0} \geq 0.$$

Hence, only two types of corner solutions are possible: either $\overline{\varphi} > 0$ (i.e., $q^{s*} = \overline{Q}^s$) and $\overline{\phi} > 0$ (i.e., $\overline{Q}^s = 0$); or $\underline{\varphi} > 0$ (i.e., $q^{s*} = \underline{Q}^s$) and $\underline{\phi} > 0$ (i.e., $q^{f*} = 0$). Assume that $\underline{\varphi} > 0$ and $\underline{\phi} > 0$ when capability and input are economic substitutes. The considered corner solution is such that: $q^{f*} > 0$ and $q^{s*} = \underline{Q}^s > 0$; thus, $\underline{FOC} \geq 0$. By economic substitutability between capability and input, $\widehat{FOC} - \underline{FOC} < 0$ (see the argument of Proposition 1). Thus, we have a contradiction: $0 \leq \underline{FOC} < \frac{\frac{\nu}{\lambda}}{1-\frac{\nu}{\lambda}} \cdot (\widehat{FOC} - \underline{FOC}) < 0$. The only corner solution compatible with economic substitutability between capability and input is $\overline{\varphi} > 0$ and $\overline{\phi} > 0$, hence $\overline{Q}^s = 0$. Notice that this also implies $\overline{Q}^m = 0$. Moreover, by inspection of (6) and (7), also the second result is proven. ■

Proposition 6 *Under the DL regime, if capability and input are economic complements, the corner solution of the problem (5) is such that:*

1. *the public provision mix is pure opting-out (i.e., $q^{f*} = 0$, and $q^{s*} > 0$);*

2. *the tax schedule does not distort high-capability labor supply, and distorts low-capability labor supply with a marginal tax featured by an additional term with respect to the PT regime.*

Proof. Assume that the corner solution featured by $\overline{\varphi} > 0$ and $\overline{\phi} > 0$ is compatible with economic complementarity between capability and input. Then, the optimal public provision mix is pure topping up, and such that $\overline{Q}^s = \overline{Q}^m = 0$. By economic complementarity between capability and input, q^{f*} is such that $\underline{FOC} < 0$. But given that $\overline{\varphi} > 0$ and economic complementarity, necessarily $\underline{FOC} = \frac{\frac{\nu}{\lambda}}{1-\frac{\nu}{\lambda}} \cdot (\widehat{FOC} - \underline{FOC}) > 0$. Thus, we have a contradiction and the only corner solution compatible with economic complementarity between capability and input is $\underline{\varphi} > 0$ and $\underline{\phi} > 0$, that imply $q^{s*} = \underline{Q}^s > 0$ and $q^{f*} = 0$. The second result derives by inspection of (6) and (7). ■

Appendix C: Numerical Analysis Algorithm

The algorithm we employ in our numerical analyses is an extension of the Mirrlees's procedure to the case of optimal public provision. Following Mirrlees (1971) and Tuomala (1984), we numerically maximize the social welfare function under feasibility and incentive compatibility constraints.

Pure Taxation Regime

Following Diamond (1998), we use the general additive function: $\mathcal{U}(x - q^m, l) = h(x - q^m) + z(1-l)$. Thus, the utility state variable is obtained by $\mathcal{U}(x-q^m, l) = v$, that implies $x = h^{-1}(v - z(1-l)) + q^m$. The corresponding Hamiltonian function is $H(l, q^m, v, \theta) = G(v)f(\theta) + \mu(wl - h^{-1}(v - z(1-l)) - q^f - q^s(1-F(\theta)))f(\theta) + \nu(\theta)\frac{dv}{d\theta}$. The numerical solution proceeds according to the following steps.[14]

1. The exercise requires an exogenous distribution of capabilities. As argued in the

[14] We use a four-digit approximation.

text, we rely on a truncated lognormal distribution over the interval $[0, 100]$, where $\log(\theta) \sim N(2, 1)$.

2. We impose $q^f = q^s = 0$ and, in line with previous studies, we choose a trial value of μ and $T(\underline{\theta})$ (the taxation for agents with capability $\underline{\theta}$).

3. We fix an initial value of $\underline{\theta} > 0$, considering that the transversality condition $\nu(\underline{\theta}) = 0$ may restrict the choice of $\underline{\theta}$.

4. Using $l(\theta) = 0$ for $\theta \leq \underline{\theta}$, we obtain the initial utility level from $u(\underline{\theta}) = h(-T(\underline{\theta}))$ and $\nu(\underline{\theta})$ from the costate equation $\frac{d\nu}{d\theta} = -\frac{\partial}{\partial v} H(l, q^m, v, \theta)$.

5. In order to find the labor supply and the private investment in q^m at each level of θ, we solve the following first order conditions: $\frac{\partial H}{\partial l} = 0$ and $\frac{\partial \mathcal{U}}{\partial q^m} = 0$. This solution comes from a Newton's method. This method attempts to construct a sequence of pairs $(l_n(\theta), q^m_n(\theta))$ from an initial guess $(l_0(\theta), q^m_0(\theta))$ that converges towards a stationary point $(l_*(\theta), q^m_*(\theta))$ satisfying the FOCs.

6. Given the sequence of stationary points, we calculate the paths of $v(\theta) = v(\underline{\theta}) + \sum_{\underline{\theta}}^{\theta} \frac{d}{d\theta} v(l_*(\theta), q^m_*(\theta))$ and $\nu(\theta) = \nu(\underline{\theta}) + \sum_{\underline{\theta}}^{\theta} \frac{d}{d\theta} \nu(l_*(\theta), q^m_*(\theta))$.

7. We check the following transversality conditions: $R(\overline{\theta}) = 0$ and $\nu(\overline{\theta}) = 0$. If one or both of them are not satisfied, we change the values of μ and $T(\underline{\theta})$ and repeat the procedure until a fixed-point is reached. In some cases, it might also be necessary to adjust the value of $\underline{\theta}$.

Inclusive Regime

Once the calculus of the pure taxation regime is concluded, we can repeat the exercise for the inclusive regime. The optimal inclusive regime is obtained by replacing q^m with $q^f + q^s$ and removing equation $\frac{\partial U}{\partial q^m} = 0$ from step 5. In particular, we set $q^f + q^s = \sup[0, \max(q^m)]$ and we repeat the algorithm. The choice of $q^f + q^s$ is based on the idea that an inclusive regime is costly and therefore the social planner wants to provide the minimum amount of q that induces all agents to opt in. However, after having calculated the solution for the inclusive regime, we suggest to check in a neighborhood of $q^f + q^s$ whether the solution represents a maximum. Obviously, we also verify the validity of the single crossing condition, but the exploration of the neighborhood represents an additional control for numerical approximations.