# Optimal Redistribution in the Presence of Signaling

Spencer Bastani[*] Tomer Blumkin [†] Luca Micheletto[‡]

July 21, 2021

**Abstract**

We analyze optimal redistribution in the presence of labor market signaling where innate productive ability is not only unobserved by the government, but also by prospective employers. Signaling in both one and two dimensions is considered, where in the latter case firms have an informational advantage vis-a-vis the government. The dual role of income taxation in redistributing income and affecting signaling incentives is analyzed, as well as extended tax systems that combine income taxation with direct instruments allowing the signals to be taxed. A key focus is the analysis of the feasibility and social desirability of redistribution through wage compression.

**Keywords:** optimal taxation, signaling, education, monitoring

**JEL classification:** D82, H21, H52, J31

[*]Institute for Evaluation of Labour Market and Education Policy (IFAU) and Research Institute of Industrial Economics (IFN); UCFS, UCLS, CESIfo, Germany. E-mail: spencer.bastani@ifau.uu.se.

[†]Department of Economics, Ben Gurion University; CESifo, Germany; IZA. E-mail: tomerblu@bgu.ac.il.

[‡]Department of Law, University of Milan, and Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University; UCFS; CESifo, Germany. E-mail: luca.micheletto@unimi.it.

# 1 Introduction

In the canonical framework of optimal income taxation, pioneered by Mirrlees (1971), asymmetric information between the government and private agents is the major constraint on public policy. The government desires to redistribute between individuals on the basis of their innate productive abilities, but as these abilities cannot be observed for tax purposes, the government has to rely on the taxation of income and other observable quantities, serving as proxies for the (unobserved) abilities. This results in a second-best problem where incentive compatibility considerations warrant the introduction of distortions, typically taking the form of positive marginal tax rates to mitigate binding incentive compatibility constraints (with extensive margin choices in place, such as migration and labor market participation, negative marginal tax rates can be desirable).

Since the seminal contributions by Spence (1973) and Akerlof (1976), economists have recognized that asymmetric information in the labor market has a major impact on the nature of interactions between employees and firms, and can be an important source of inefficiency in market outcomes. Several recent papers have revisited the Mirrlees setup and extended it by introducing a second source of asymmetric information, between workers and employers (see e.g., Stantcheva 2014, Bastani et al. 2015, Craig 2020).

The presence of a second source of asymmetric information embedded in the market structure implies that employers cannot observe the productivity of workers, and hence, even in a competitive labor market, workers are not necessarily compensated according to their marginal product. Instead, the wage distribution becomes endogenous, and is affected by the screening/signaling possibilities available to employers/workers. Notably, this happens even though the production technology is linear and skill types are perfect substitutes, as in Mirrlees (1971).[1]

The second source of asymmetric information may also make it optimal for the government to intervene on equity grounds, and use its available policy instruments to affect the wage distribution. Bastani et al. (2015) demonstrated how the structure of the income tax affects the transmission of information in the labor market and the incentives for employers to engage in screening. The key focus in their paper was to highlight how the income tax could be used to implement earned-income bunching (and possibly pooling) as a means to mitigate the information rents associated with workers' differences in productivities, thereby achieving enhanced redistribution.

The purpose of the current paper is to provide a framework to analyze optimal redistributive policy in the presence of (impure) signaling. In accordance with Mirrlees (1971),

[1] As discussed by Stantcheva (2014) and Bastani et al. (2015, 2019), an adverse selection problem may then arise, taking the form of a rate-race in which high-skilled workers choose to work longer hours than the efficient amount, or opt for a demanding career with an inefficiently low degree of workplace flexibility, so as to differentiate themselves from their lower skilled counterparts. These types of inefficiencies justify efficiency-enhancing policy interventions in the form of mandates on work hours (e.g., the 35-hours law in France which limits the work-time per week) and binding parental leave rules.

we assume that workers differ in their innate productive abilities which are unobservable to the government. However, unlike Mirrlees, we assume that these abilities are also unobservable to prospective employers and that workers must signal their type to firms through costly effort choices (hours of work or educational attainment).[2] To render such signaling feasible, we assume that workers differ in the cost of acquiring the signal(s). We analyze optimal labor income taxation as well as optimal tax systems that combine income taxation with taxation of the signals transmitted in the labor market, allowing for both separating and pooling equilibria.

The income tax system is an indirect instrument to tax the signaling activities that take place in the labor market. However, in many relevant circumstances, direct instruments affecting the transmission of information between workers and their actual (or prospective) employers can be implemented as well, namely, using instruments that tax the signals themselves. These direct instruments may take the form of education mandates, overtime regulation or worker monitoring technologies. In the standard Mirrlees setup, such direct instruments were ruled out so as to render the analysis non-trivial, as the possibility to flexibly tax labor hours (alongside income) in the standard Mirrless framework would imply that the first best allocation could be achieved. With a second source of asymmetric information, however, employers cannot distinguish between workers with different abilities, and need to rely on the observed signals strategically chosen by the workers. This implies that in a setting with two layers of asymmetric information, the first-best optimum cannot be achieved, even when direct instruments are available.

Our analysis begins by considering a framework with signaling in one dimension. We then consider the more realistic case in which employers, while still being uninformed about workers' productivities, possess better information than the government. We do so by assuming that workers' signaling has two dimensions: quantity (hours of work, years of schooling) which is universally observable (that is, both by the government and the employers) and quality (effort level) that is exclusively observed by the employers.

Our main results can be summarized as follows. With signaling confined to a single dimension, and when only income can be taxed, the social optimum is given by either a separating or a pooling allocation. However, when the income tax is supplemented by a direct tax on the signal, the optimum is always given by a separating allocation. The reason is that a direct tax on the signal enables to fully eliminate the information rents associated with differences in productivities, which eliminates the case for redistribution through wage compression (pooling). Due to the presence of a second layer of asymmetric information, however, a first-best is not necessarily attainable, as there remain information rents associated with the differences in the cost of signaling. The presence of binding incentive-compatibility constraints implies that it is optimal for the government to impose

[2] Given that the signals considered in our model are not a pure waste (since they also affect the output produced by workers), ours is, more properly, a model of impure signaling.

distortions on the labor/signaling choices of individuals, distortions that differ from the standard ones typically emphasized in optimal tax models. Interpreting the signal as educational attainment, the case for taxing the signal directly provides a normative justification for e.g., education mandates on redistributive grounds.

In the presence of bi-dimensional signaling, and when only income can be taxed, we show that a pooling equilibrium does not exist and the social optimum is necessarily given by a separating allocation (implying that redistribution through the wage channel is infeasible). However, we show that a pooling equilibrium can be supported by supplementing the income tax with a direct tax levied on the signal observable by the government by eliminating the possibility for cream-skimming. Furthermore, we show that a pooling equilibrium can be socially desirable since a tax on one of the two signals available to agents does not fully eliminate the information rents associated with differences in productivities, as workers still have another avenue to exploit in order to signal their productivity to firms. Thus, when one of the signals is taxed, redistribution through the wage channel is both feasible and sometimes also socially desirable. We illustrate the nature of the social optimum and the associated welfare gains depending on the parameters of the economy through several numerical examples. These examples highlight both the role of the difference in the inherent producitvities of agents, as well as the differences in the costs of signaling.

While we present our signaling analysis in a general way, opening up for several different interpretations of the information transmission that takes place between workers and firms, one important context is educational attainment. In the context of educational attainment, government intervention is often justified on efficiency grounds, serving to address common market failures such as alleviating credit constraints and internalizing externalities, and sometimes on redistributive grounds, viewing the increase in returns to education as a key contributor to the well documented rise in earnings inequality. A frequent feature in many studies (and in common practice) is the desirability of education subsidies.[3] Such subsidies typically serve a dual role of mitigating (or potentially offsetting) the dis-incentivizing effect of income taxation on human capital acquisition; and, enhancing redistribution. In our paper, we demonstrate that both education subsidies and taxes are warranted on redistributive grounds to mitigate binding incentive compatibility constraints that restrict the amount of achievable redistribution.[4]

[3]The optimal tax literature has analyzed education subsidies in conjunction with redistributive tax and transfer systems, focusing on the productivity-enhancing role of education (see, Sheshinski 1971, Ulph 1977, Tuomala 1986, Boadway and Marchand 1995, Brett and Weymark 2003, Bovenberg and Jacobs 2005, Maldonado 2008 and Findeisen and Sachs 2016 among many others).

[4]Blumkin and Sadka (2008) offer a different normative justification for the desirability of taxing education, relying on a positive correlation between the observed educational attainment and the unobserved innate productive ability. Unlike the current framework, they assume that skill levels are observed by firms. In our framework, in contrast, a tax on education serves to implement a separating allocation, via mitigating the incentives of the low-skilled workers to mimic their higher-skilled counterparts, when employers are unable to observe the true productivity of their employees.

The paper is organized as follows. In Section 2 we analyze optimal redistribution in the presence of signaling in one dimension. Section 3 considers a model with signaling in two dimensions, capturing the possibility for firms to have an informational advantage vis-a-vis the government. Section 4 presents further analytical and numerical characterization of the two-dimensional signaling model. Section 5 provides concluding remarks.

# 2 A model with signaling in one dimension

Consider an economy comprised of two types of workers: a low-skilled worker, denoted by $i = 1$, and a high-skilled worker denoted by $i = 2$, who differ in their innate productive ability. The fraction of type-$i$ workers in the population (normalized to a unit measure, with no loss of generality) is denoted by $0 < \gamma^i < 1$. The labor market is competitive, but the innate productive ability of a worker is assumed to be private information unavailable to the firm. Thus, we deviate from the standard Mirrleesian setup.[5]. Workers exert costly effort denoted by $e^i$ which serves a dual purpose: (i) increasing the output of the worker, and, (ii) signaling innate productive ability. The output of worker $i$ is given by:

$$z^i = e^i \theta^i, \tag{1}$$

where $\theta^i$ denotes the innate productive ability of type $i$, with $\theta^2 > \theta^1 > 0$. The utility of a type-$i$ worker is given by:

$$u^i(c^i, e^i) = c^i - g^i(e^i), \tag{2}$$

where $c^i$ denotes consumption and $g^i$ denotes the cost of exerting effort and is assumed to be strictly increasing and strictly convex. Moreover, the cost of acquiring effort (both marginal and total) decreases with respect to the skill level, i.e., $g^2(e) < g^1(e)$ and $\partial g^2(e)/\partial e < \partial g^1(e)/\partial e$. These are standard properties in the signaling literature and ensure that the single-crossing property, that is essential for the feasibility of a separating (fully revealing) equilibrium, is satisfied. To simplify notation, we let $g^2(e) \equiv g(e)$ and $g^1(e) \equiv kg(e)$ where we assume $k > 1$ so that low-skilled agents incur a higher effort cost, for any given effort level. The firms observe $e^i$, which serves as a signal for the unobserved innate productive ability of the worker.

## 2.1 Laissez-faire equilibrium

We define a Bayesian Nash Equilibrium for the signaling game in the absence of government intervention. The signaling game is comprised of two stages. In the first stage, workers choose their level of effort $e^i$, $i = 1, 2$. In the second stage, each firm offers a labor

[5]See Bastani et al. (2015, 2019) and Stantcheva (2014) for a similar approach.

contract which specifies the income level (which is also the consumption level with no taxes or transfers in place) as a function of the observed effort levels, namely, $y^i(e^i)$. Based on the observable effort levels, firms form their beliefs with respect to the workers' types. In equilibrium, choices are consistent in the sense that firms maximize their expected profits by choosing labor contracts, given their beliefs; and, workers maximize their utility by choosing their effort level given the labor contracts offered by firms.[6] Assuming a perfectly competitive labor market with free entry of firms implies that rents are fully dissipated, hence, $y^i = e^i\theta^i$ and the compensation of all agents is equal to their production.

Due to the asymmetric information between firms and workers (with respect to innate productive ability) a laissez-faire equilibrium has to satisfy incentive compatibility constraints. In particular, at a separating equilibrium the effort chosen by type-1 agents, $e^{1*}$, is given by

$$e^{1*} = \underset{e^1}{\operatorname{argmax}} \left\{ e^1\theta^1 - kg(e^1) \right\}, \tag{3}$$

whereas the effort chosen by type-2 agents, $e^{2*}$, is given by

$$e^{2*} = \underset{e^2}{\operatorname{argmax}} \left\{ e^2\theta^2 - g(e^2) \right\}, \tag{4}$$

subject to:

$$e^{1*}\theta^1 - kg(e^{1*}) \geq e^2\theta^2 - kg(e^2). \tag{5}$$

According to problem (3), type-1 agents exert their efficient effort level, which is implicitly characterized by the first order condition $g'\left(e^1\right) = \theta^1/k$. Type-2 agents, instead, have to exert an effort level that credibly signals their higher productivity to the firms. This requires that $e^{2*}$ must satisfy the incentive constraint (5).[7] The reason for this constraint is that low-skilled agents may find it desirable to replicate the effort of high-skilled agents in order to receive a higher remuneration (per unit of effort). Therefore, an adverse selection problem generally arises, implying that type-2 workers may be induced to choose an inefficiently high level of effort. As we will see, the possibility of a binding upward constraint carries over to the optimal income tax problem studied in the next subsection, which implies the possibility of negative optimal marginal tax rates at the top.

## 2.2 The income tax regime

Suppose that the government is seeking to design a nonlinear income tax system that serves redistributive purposes. In particular, assume that the government is invoking a max-min social welfare function and is hence interested in maximizing the utility of

[6] For a more formal treatment see, e.g., Fudenberg and Tirole (1991).

[7] The set of contracts $\{(e^{1*}, e^{1*}\theta^1), (e^{2*}, e^{2*}\theta^2)\}$ represents the unique Nash equilibrium pair of contracts which satisfies the intuitive criterion proposed by Cho and Kreps (1987). We further discuss this commonly applied refinement condition and its policy implications in subsection 2.2.1 below.

the least well-off workers (type-1) subject to a balanced budget constraint (we assume no exogenous revenue needs with no loss of generality).

For later purposes, define the wage rate earned by a given individual as the ratio between his/her pre-tax income and his/her effort. Due to the second layer of asymmetric information between the firms and the workers, the equilibrium wage distribution is affected by the tax-and-transfer system implemented by the government. This enables the government to use the wage channel as a supplementary tool to the standard income channel to attain enhanced redistribution. In particular, the government can, by a proper choice of the tax schedule, implement either a separating or a pooling equilibrium.

### 2.2.1 A separating equilibrium

In the spirit of the literature on optimal nonlinear income taxation, we assume the income tax is defined by a set of pre-tax/post-tax income bundles denoted by $(y^i, c^i)$ where the total tax (or transfer if negative) is defined by $t^i \equiv y^i - c^i$. A naive statement of the problem solved by a government seeking to implement a separating equilibrium is the following:

$$\max_{\{y^1,c^1,y^2,c^2\}} \left\{c^1 - kg(e^1)\right\} \tag{6}$$

subject to:

$$y^i = e^i\theta^i, \quad i = 1,2 \tag{7}$$

$$c^1 - kg(e^1) \geq c^2 - kg(e^2), \tag{8}$$

$$c^2 - g(e^2) \geq c^1 - g(e^1), \tag{9}$$

$$\sum_i \gamma^i(y^i - c^i) = 0. \tag{10}$$

Conditions (8)-(9) are the standard incentive compatibility constraints, which ensure no mimicking. Each type weakly prefers his bundle over the one associated with his counterpart. Condition (10) states the government revenue constraint which ensures a balanced budget. Notice that the incentive constraints (8)-(9) cannot both be binding at the same time by virtue of the single crossing property (reflected by the fact that $k > 1$).

The intuition for the presence of either a downward or upward binding constraint can be understood as follows. Redistribution from high- to low-skilled agents invites mimicking by high-skilled agents, who might be tempted to earn the income of low-skilled agents in order to qualify for a lower tax burden. In contrast, low-skilled agents might have an incentive to mimic high-skilled agents in order to qualify for a higher wage rate (i.e., for a higher remuneration per unit of effort), even though they would be subject to higher income taxation.

A limitation of the above statement of a separating equilibrium in the presence of an income tax is that it does not take into account the possibility for off-equilibrium deviations. In order to accommodate these potential stability threats, we will rely on the extension, provided by Grossman and Perry (1986), of the intuitive criterion proposed by Cho and Kreps (1987), which imposes some natural restrictions on out-of-equilibrium beliefs.[8]

Let $T$ denote the set of all types (in our case $T = \{1,2\}$) and let $S \subseteq T$ denote a subset of $T$. Suppose all types in the subset $S$ deviate by choosing an effort level, $\hat{e}$, different than those specified (for their respective types) in equilibrium. Three questions need to be asked:

a) Assuming that all firms believe that the deviating types come from the subset $S$ (and accordingly update their beliefs in a Bayesian fashion), what would be the maximum (gross) income that the firms could offer? Denote the maximum income level by $\hat{y}$. A deviation to $\hat{e}$ thus defines a new contract (which is not part of the presumably stable equilibrium) given by the triplet $(\hat{e}, \hat{y}, \hat{c})$, where $\hat{c}$ denotes the net income associated with $\hat{y}$.

b) Would any type $s \in S$ be strictly better off with $(\hat{e}, \hat{y}, \hat{c})$ than with his equilibrium bundle?

c) Would any type $s \notin S$ be weakly better off with the equilibrium bundle than with the contract $(\hat{e}, \hat{y}, \hat{c})$?

If the answer to both (b) and (c) is yes, then any type in $S$ finds it profitable to deviate, as he is able to credibly distinguish himself from types that are not in $S$ and thereby become strictly better off.[9]

Turning back to the separating equilibrium defined above, we acknowledge that the only possible deviations in the presence of the income tax are those associated with the income levels $y^1$ and $y^2$. Consider a deviation to the following bundle: $(\hat{e}, y^1, c^1)$, where $\hat{e} = \frac{y^1}{\sum_i \gamma^i \theta^i}$. As $\theta^2 > \theta^1$, we have that $\hat{e} < e^1 = \frac{y^1}{\theta^1}$. It thus follows that type-1 strictly prefers the bundle $(\hat{e}, y^1, c^1)$ over his equilibrium bundle $(e^1, y^1, c^1)$. Now, suppose that $c^1 - g(\hat{e}) > c^2 - g(e^2)$, namely that type 2 also strictly prefers the bundle $(\hat{e}, y^1, c^1)$ over his equilibrium bundle (which is $(e^2, y^2, c^2)$). Letting the subset $S$ be given by $\{1,2\}$, it follows that $y^1$ is the maximum income level associated with the effort level $\hat{e}$ in order for the firm not to make negative profits. Moreover the answer for question (b) is affirmative and condition (c) is vacuously satisfied (as there are no types $s$ which are not in $S$). Thus, our suggested equilibrium fails to satisfy the extended intuitive criterion.

It is straightforward to verify that given the income tax schedule, which confines the set of (gross) income levels to the pair $(y^1, y^2)$, the deviation analyzed above is the only threat that challenges the separating equilibrium defined in problem (6). To see this

[8]Our exposition of the Grossman and Perry (1986) criterion follows Riley (2001).

[9]Notice that in case $S$ is confined to be a singleton, we obtain the standard intuitive criterion.

notice that bunching at the higher (gross) income level $y^2$ would dictate, by the zero profit condition, an effort level higher than $e^2$, which would make it an unattractive deviation for both types of workers.

Modifying the definition of the separating equilibrium to accommodate the extended intuitive criterion yields the following problem to be solved by the government:

$$\max_{\{y^1,c^1,y^2,c^2\}} \left\{c^1 - kg(e^1)\right\} \tag{11}$$

subject to:

$$y^i = e^i\theta^i, \quad i = 1,2 \tag{12}$$

$$c^1 - kg(e^1) \geq c^2 - kg(e^2), \tag{13}$$

$$c^2 - g(e^2) \geq c^1 - g\left(\frac{y^1}{\sum_i \gamma^i\theta^i}\right), \tag{14}$$

$$\sum_i \gamma^i(y^i - c^i) = 0. \tag{15}$$

Three remarks are in order. First, notice that the incentive compatibility constraint given by (14) implies the weaker constraint given by (9) and hence renders the latter redundant. The fact that (14) is a tighter constraint than (9) reflects the presence of an information rent (obtained by type 2) which is associated with the difference in productivities. If the innate productivities were equal, there would be no difference between the two constraints.

Second, notice that the separating allocation obtained under the extended intuitive criterion coincides with the Rothschild and Stiglitz (1976) separating equilibrium in the screening game, in which the first movers are the firms rather than the workers (who are the first movers in the signaling game).

Finally, whereas the incentive compatibility constraints (8)-(9) satisfy the single-crossing property, thus ensuring that only one of these constraints (typically the one associated with the high-skilled worker) is binding at an optimum, the modified conditions in (13) and (14) may well both bind at the same time. This implies a potential violation of the standard zero marginal tax at the top property (Sadka 1976) which is due to a problem of adverse selection in the labor market.

To get an intuition why constraints (13) and (14) could both be binding under an optimal nonlinear income tax, recall that, under laissez-faire, low-skilled agents may find it attractive to replicate the effort of high-skilled agents in order to receive a higher remuneration (per unit of effort). Thus, absent any government intervention, an upward incentive-compatibility constraint might be binding (low-skilled agents might be tempted to mimic high-skilled agents) because of the problem of asymmetric information between firms and workers. In the presence of a government redistributing from high- to low-skilled workers, the incentives for low-skilled agents to mimic high-skilled agents are clearly weakened. Under a max-min social welfare function, the downward incentive-compatibility constraint, requiring high-skilled agents not to be tempted to mimic low-skilled agents, will be binding. However, if this constraint implies that high-skilled agents enjoy a sufficiently large information rent associated with differences in productivities, the upward incentive-compatibility constraint will not cease to be binding, implying that both the upward and downward constraint can be binding at the same time.[10]

### 2.2.2 A pooling equilibrium

By offering a single income level on the income tax schedule, denoted by $\hat{y}$, the government can implement a pooling equilibrium, as it can rule out deviations to other income levels than $\hat{y}$.[11] In a pooling equilibrium, all workers choose the same level of effort $\hat{e}$. Hence, total production is equal to $\hat{y} = \hat{e}\overline{\theta}$, where $\overline{\theta} = \sum_i \gamma^i \theta^i$ is the average productivity. A government seeking to implement a pooling equilibrium will then solve the following problem:

$$\max_{\hat{y}} \left\{ \hat{c} - kg(\hat{e}) \right\} \tag{16}$$

subject to:

$$\hat{y} = \hat{e}\overline{\theta}, \tag{17}$$

$$\hat{c} = \hat{y}. \tag{18}$$

It is worth noting that pooling is not subject to cream-skimming and is hence stable by virtue of the income tax that induces a full compression of the income distribution (on and off equilibrium). We now proceed to compare the separating and pooling equilibrium.

### 2.2.3 Comparison between separating and pooling equilibria

As we pointed out at the end of Section 2.2.1, there is an equivalence between the signaling model (subject to the extended intuitive criterion) and the competitive screening setup.[12] Because of this equivalence, the result stated in Proposition 1 of Bastani et al. (2015) also applies to our model with signaling in one dimension. In particular, the social optimum can either be a separating or a pooling equilibrium.

The desirability of a pooling equilibrium derives from the role played by the wage channel in realizing redistributive goals. A pooling equilibrium forces wage equalization

[10] Put differently, the violation of the single crossing property by conditions (13) and (14) stems from the fact that, whereas mimicking low-skilled agents are replicating the choices of their high-skilled counterparts, mimicking high-skilled agents are choosing an effort level strictly lower that that chosen by their low-skilled counterparts. Thus, we are not anymore requiring that two bundles chosen by both types in equilibrium lie along two indifference curves that intersect at a single point (a standard single-crossing argument).

[11] The possibility to implement a pooling allocation in a two-type optimal income tax model without adverse selection in the labor market was discussed by (Stiglitz 1982). However, in that setting, pooling is Pareto-inferior to the laissez-faire allocation.

[12] See also Riley (2001, pp. 445-46).

and thereby eliminates all the information rent (derived by high-skilled agents) associated with the difference in productivities (but not the information rent associated with the difference in the acquisition costs of the signal). Whether a pooling equilibrium constitutes the social optimum depends on how these equity gains compare to the efficiency properties of the pooling equilibrium relative to the separating equilibrium.

With respect to the efficiency properties notice that, under an optimal pooling equilibrium, the first order condition of the government's problem is $1 = kg'\left(\hat{y}/\overline{\theta}\right)/\overline{\theta}$. This implies that, while both types of agents choose the same effort, the effort choice of type-1 agents is distorted upwards and the effort choice of type-2 agents is distorted downwards. Under an optimal separating equilibrium, instead, the effort choice of type-1 agents is distorted downwards, whereas the effort choice of type-2 agents may either be left undistorted or distorted upwards. To see this, consider the problem solved by the government under a separating equilibrium (i.e., the problem given by (11)-(15)). Denoting by $\lambda^1$ the Lagrange multiplier attached to the constraint (13), by $\lambda^2$ the multiplier attached to the constraint (14) and by $\mu$ the multiplier attached to the constraint (15), from the first order conditions of the government's problem it is straightforward to obtain that

$$1 - \frac{kg'\left(\frac{y^1}{\theta^1}\right)}{\theta^1} = \frac{\lambda^2}{\mu\gamma^1}\left[\frac{kg'\left(\frac{y^1}{\theta^1}\right)}{\theta^1} - \frac{g'\left(\frac{y^1}{\overline{\theta}}\right)}{\overline{\theta}}\right], \tag{19}$$

and

$$1 - \frac{g'\left(\frac{y^2}{\theta^2}\right)}{\theta^2} = \frac{\lambda^1}{\mu\gamma^2}\frac{g'\left(\frac{y^2}{\theta^2}\right)}{\theta^2}(1-k). \tag{20}$$

Given our focus on a max-min social welfare function, constraint (14) is necessarily binding ($\lambda^2 > 0$), which implies that the right hand side of (19), which provides a measure of the distortion imposed on $y^1$, is unambiguously positive. On the other hand, the right hand side of (20), which provides a measure of the distortion imposed on $y^2$, will either be nil (if constraint (13) is slack) or it will be negative (if constraint (13) is binding, recalling that $k > 1$).

A more detailed characterization of the conditions under which the pooling equilibrium welfare-dominates the separating equilibrium is provided in Bastani et al. (2015) (see especially the discussion following Proposition 2).

## 2.3 The extended tax regime

We now extend our tax base and allow the government to tax each individual based on two observable characteristics: (i) earned income, (ii) the signal transmitted in the labor market. Thus, the tax function that we consider in this section is a nonlinear and nonseparable function of both $y$ and $e$. We maintain all our earlier assumptions and in particular the asymmetry in information between firms and workers with respect to the latter's innate productive ability. We focus on the separating equilibrium as the pooling equilibrium, in the extended tax regime, is identical to the one under the income tax regime.[13]

The problem of the government in the extended tax regime is the following:

$$\max_{\{y^1,c^1,y^2,c^2,e^1,e^2\}} \left\{c^1 - kg(e^1)\right\} \tag{21}$$

subject to:

$$y^i = e^i\theta^i, \quad i = 1,2, \tag{22}$$

$$c^1 - kg(e^1) \geq c^2 - kg(e^2), \tag{23}$$

$$c^2 - g(e^2) \geq c^1 - g(e^1), \tag{24}$$

$$\sum_i \gamma^i(y^i - c^i) = 0. \tag{25}$$

Comparing the constraints (22)-(25) with the corresponding constraints under the income tax regime, i.e. constraints (12)-(15), one can see that the difference between the two regimes is confined to the incentive compatibility constraint associated with the high-skill type. The possibility to tax $e$ prevents agents from choosing effort levels other than those specified in the bundles $\left(y^1,c^1,e^1\right)$ and $\left(y^1,c^2,e^2\right)$. This implies that the stability threat posed by the extended intuitive criterion is no longer feasible: a type-2 worker behaving as a mimicker, i.e. earning $y^1$ in order to benefit from a more lenient tax treatment, is forced to choose the same effort level prescribed in equilibrium for a type-1 worker. In contrast to what happened under the pure income tax regime considered in Section 2.2.1, high-skilled agents no longer enjoy an information rent associated with the innate difference in productivities. Under an extended tax regime, the amount of achievable redistribution is solely determined by the differences in the costs of acquiring the signal.

By alleviating one of the incentive constraints that limit the amount of feasible redistribution, the extended tax regime allows to increase social welfare relative to the case with only an income tax in place. This result, which has important policy implications, is formally stated in Proposition 1 below. Proposition 1 also provides two additional results. First, it shows that the possibility to tax the signal renders a pooling equilibrium socially undesirable for any difference in productivities between the two types of workers and independently of the difference in the acquisition cost of the signal. This result, which is in contrast to what was the case under the optimal income tax regime, is related to the fact that, by replacing the right hand side of (14) with the right hand side of (24), the single-crossing condition is restored (since constraints (23)-(24) cannot be binding at the

[13] In the extended tax regime, either an income tax or a tax on the signal can be used to implement a pooling equilibrium.

same time).[14] Second, the proposition shows that the optimal distortions imposed by the extended tax regime are unrelated to the differences in productivities.

**Proposition 1.** *In the extended tax regime, the following holds:*

*(i) the social optimum is always attained by a separating allocation;*

*(ii) social welfare is strictly higher than under the optimal income tax regime;*

*(iii) the optimal distortions imposed by the extended tax regime are unrelated to the differences in productivities.*

**Proof** See Appendix A. □

We now proceed to discuss some policy implications of Proposition 1.

*Monitoring cannot eliminate the distortions of the tax system*

By assuming that the innate productive abilities of workers are private information, unobserved by firms, we have deviated from the standard Mirrleesian setup (Stiglitz 1982). In the standard setup with only one layer of asymmetric information (between the government and the agents), the possibility to tax effort implies that the government effectively can tax the abilities of the workers, thereby attaining the first-best allocation.[15]

In our setup, with a second layer of asymmetric information, the first best is unattainable but, nonetheless, a tax on the signal can enhance welfare. In the case where the signal represents work effort, we have that even when work effort (not just nominal time spent at work) can be monitored and subject to taxation (which is not an excessively unrealistic assumption given recent technological advancements), there would still be a non-trivial trade-off between efficiency and equity. The desirability of introducing such monitoring technologies, which would typically entail some non-trivial costs, is beyond the scope of the current study. However, an important insight from our analysis is that monitoring cannot fully eliminate the distortions from the tax system in the presence of a second layer of asymmetric information (between the firms and the workers).[16]

[14] Notice that the set of constraints (22)-(25) is identical to the set of constraints (7)-(10) that characterized the naive statement of the government's problem under a pure income tax regime.

[15] If income is the product of skill and effort, $y = \theta e$, and the government observes $y$ and $e$, then clearly $\theta$ can be inferred. This implies that the government can assign tax burdens based on ability. The reason ability cannot be inferred in this way in presence of asymmetric information between workers and firms is that $y$ no longer needs to be equal to $\theta e$.

[16] Zoutman and Jacobs (2016) consider monitoring in an optimal income tax framework without asymmetric information between workers and firms where the government can learn the labor supply of a worker with a certain probability at a finite cost. Related monitoring ideas have appeared in the literature on optimal tax administration. Keen and Slemrod (2017) argue that the government can either design optimal tax policy taking the behavioral responses of agents as given, or design tax policy and enforcement together, implying that the behavioral responses are viewed as partially being under the control of the government. Slemrod and Kopczuk (2002) discuss similar ideas in the context of taxable income elasticities and Kleven and Kopczuk (2011) in the context of welfare programs.

*Mandates*

An alternative to levying a fully nonlinear tax on the signal (which in the case of work effort, would correspond to a perfect and costless monitoring technology) would be to set a binding mandate at the level of effort associated with type-1 workers under the separating optimal allocation. By doing so, type-2 mimickers would be denied the information rent associated with the difference in productivities between the two types of agents. The mandate essentially blocks the possibility, for type-2 mimickers, to credibly signal their superior productive ability to the firm. This serves to enhance redistribution. Our analysis therefore suggests a novel normative justification for commonly used mandates, such as education mandates, often warranted on efficiency grounds (internalization of positive spillovers, mitigation of imperfections in capital markets etc.), see e.g., Balestrino et al. (2016).

# 3 A model with signaling in two dimensions

We now extend the setting analyzed in section 2 by considering the availability of two signals. The first signal is denoted by $e_s$, and represents the quantity of effort. The second signal is denoted by $e_q$, and represents the intensity of effort. In the context of education, the variables $e_s$ and $e_q$ would capture, respectively, the quantity (for example, the time spent in acquiring vocational training and/or academic degrees) and quality (for instance, GPA, reputation of certifying institute, etc.) dimensions of educational attainment. In the context of labor supply, $e_s$ and $e_q$ would respectively capture the time and intensity components of labor effort. We assume that $e_s$ is observed by both the government and the firms, whereas $e_q$ is only observed by the firms (or prohibitively costly to observe by the government).

The output of a type-$i$ worker is given by the production function:

$$z^i = h(e_s^i, e_q^i)\theta^i, \tag{26}$$

where $h(\cdot)$ is jointly concave and strictly increasing in both arguments.

The utility function is given by:

$$u^i(c^i, e_s^i.e_q^i) = c^i - R^i(e_s^i, e_q^i), \tag{27}$$

where $c^i$ denotes consumption and

$$R^i(e_s^i, e_q^i) = p_s^i e_s^i + p_q^i e_q^i \tag{28}$$

is the cost function for type-$i$ agents, with $p_s^i$ and $p_q^i$ denoting, respectively, the unitary price of $e_s$ and the unitary price of $e_q$ for an agent of type $i$. Regarding the prices $p_s^i$ and $p_q^i$ we will make hereafter the following assumptions:

$$p_s^1 = p_s^2 \equiv p_s, \tag{29}$$

and

$$p_q^1 > p_q^2, \tag{30}$$

which jointly imply that type-2 agents have a comparative advantage in the quality signal $e_q$. Before turning to the government problem, we define the laissez-faire market equilibrium.

## 3.1 Laissez-faire equilibrium

The two-stage signaling game in the presence of two-dimensional signals works as follows. In the first stage, workers choose their levels of effort (both quality and quantity components), $(e_s^i, e_q^i)$; $i = 1,2$. In the second stage each firm offers a labor contract which specifies the income level (which is also the consumption level with no taxes/transfers in place) as a function of the observed signals, namely, $y(e_s, e_q)$. Based on the observed signals, firms form their beliefs with respect to the workers' types. In equilibrium, choices are consistent in the sense that firms maximize their expected profits by choosing the labor contracts, given their beliefs; and, workers maximize their utility by choosing their signal effort levels given the labor contracts offered by the firms. Assuming a perfectly competitive labor market with free entry of firms implies that rents are fully dissipated, hence, $y^i = h(e_s^i, e_q^i)\theta^i$.

To illustrate the framework, consider the following laissez-faire allocation given by the pairs $(e_s^{i*}, e_q^{i*})$, $i = 1,2$, satisfying the following conditions:

$$(e_s^{1*}, e_q^{1*}) = \underset{e_s^1, e_q^1}{\operatorname{argmax}} \left\{ h(e_s^1, e_q^1)\theta^1 - R^1(e_s^1, e_q^1) \right\}, \tag{31}$$

and

$$(e_s^{2*}, e_q^{2*}) = \underset{e_s^2, e_q^2}{\operatorname{argmax}} \left\{ h(e_s^2, e_q^2)\theta^2 - R^2(e_s^2, e_q^2) \right\} \tag{32}$$

subject to:

$$h(e_s^{1*}, e_q^{1*})\theta^1 - R^1(e_s^{1*}, e_q^{1*}) \geq h(e_s^2, e_q^2)\theta^2 - R^1(e_s^2, e_q^2), \tag{33}$$

which has a similar structure to the problem illustrated in section 2.1. The above incentive constraint reflects the fact that type-1 workers might have an incentive to mimic their higher-skilled counterparts in order to be remunerated according to the (higher) productivity of type-2 agents. If the constraint is binding, an adverse selection problem arises, implying that type-2 workers are induced to over-invest in the quality-signal $e_q$ (due to their comparative advantage in this dimension of signaling).

## 3.2 The income tax regime

We first consider the benchmark setup where an individual's tax liability is just a function of his/her earned income. We assume, like we did for the case of signaling in one dimension, that the government is invoking a max-min social welfare function. The income tax is, as before, defined by a set of pre-tax/post-tax income bundles denoted by $(y^i, c^i)$ where the total tax (or transfer if negative) is defined by $t^i \equiv y^i - c^i$. We define the wage rate earned by a given individual as the ratio between his/her pre-tax income $y$ and the value of the $h$-function evaluated at the effort vector chosen by the individual, thereby generalizing the definition of wage rate provided in Section 2.2.

### 3.2.1 A separating equilibrium

Under a separating equilibrium, the tax schedule is designed in such a way to induce type-1 agents to select a bundle $(y^1, c^1)$ with associated tax payment $t^1 = y^1 - c^1 < 0$, and to induce type-2 agents to select a bundle $(y^2, c^2)$ with associated tax payment $t^2 = y^2 - c^2 > 0$. Type-$i$ agents choosing the bundle intended for them on the tax schedule would choose an efficient mix of $e_s$ and $e_q$, denoted by $(e^i_s(y^i), e^i_q(y^i)), i = 1, 2$, with an associated cost given by:

$$R^i(y^i) = \min_{e_s, e_q} R^i(e_s, e_q) \quad \text{subject to} \quad h(e_s, e_q)\theta^i = y^i. \tag{34}$$

Notice that efficiency in the choice of the effort mix means that $e^i_s(y^i)$ and $e^i_q(y^i)$ satisfy the condition

$$\frac{\partial h\left(e^i_s(y^i), e^i_q(y^i)\right)/\partial e^i_s}{\partial h\left(e^i_s(y^i), e^i_q(y^i)\right)/\partial e^i_q} = \frac{p_s}{p^i_q}, \tag{35}$$

which equates the marginal rate of technical transformation (MRTS) to the price ratio. Moreover, notice that under a separating equilibrium, agents are paid by the firm according to their true productivity, in other words, an agent of type $i$ is paid a wage rate equal to $\theta^i$.

To implement a given separating equilibrium, the government has to guard against various deviating strategies available to agents, i.e. the government has to make sure that no agent has an incentive to deviate from the behavior expected from him/her. There are in principle three deviating strategies that may be adopted by an agent of type $i$ choosing to earn the income $y^j$ intended for the other type. The agent can choose an effort vector that enables him/her to get remunerated according to: (i) the productivity of the other type, (ii) the average productivity, or, (iii) his/her true productivity. We consider these three deviating strategies in more detail below. Given that a deviating agent is someone who earns an amount of income which is intended for some other type of agents, we will follow the common practice of using the word "mimicker" to refer to a deviating agent in all three cases.

A first deviating strategy is for type-$i$ agents to earn the income level $y^j$ by choosing the effort mix $(e_s^j(y^j), e_q^j(y^j))$ chosen in equilibrium by type-$j$ agents. Behaving in this way, a type-$i$ mimicker would be paid a wage rate $\theta^j$ (i.e., according to the productivity of the type being mimicked) and would incur the following cost:

$$\breve{R}^i(y^j) = p_s^i \breve{e}_s^i(y^j) + p_q^i \breve{e}_q^i(y^j), \tag{36}$$

where $(\breve{e}_s^i(y^j), \breve{e}_q^i\left(y^j\right)) = (e_s^j(y^j), e_q^j(y^j))$ denotes the effort mix of a type $i$ mimicker, which is identical to the effort mix chosen in equilibrium by agents of type $j$. Notice that $(\breve{e}_s^i(y^j), \breve{e}_q^i\left(y^j\right))$ is, from the perspective of type-$i$ agents, a distorted effort mix, in the sense that it does not satisfy the condition $\frac{\partial h(e_s,e_q)/\partial e_s}{\partial h(e_s,e_q)/\partial e_q} = \frac{p_s}{p_q^i}$.

Besides the deviating strategy described above, which involves a type-$i$ mimicker choosing the effort vector selected in equilibrium by agents of type $j \neq i$, there are also deviating strategies that involve the choice, by a mimicker, of an off-equilibrium effort vector.

The first of such strategies is the possibility for a type-$i$ agent to earn the income level $y^j$ by choosing an effort vector which is at the same time: i) different from the one chosen in equilibrium by type-$j$ agents, ii) attractive also for type-$j$ agents, and iii) sufficient to allow firms to make non-negative profits when remunerating agents according to the average productivity $\bar{\theta}$ $(= \gamma^1\theta^1 + \gamma^2\theta^2)$. For a type-$i$ mimicker, the most attractive of this kind of deviating strategies is the one with associated cost given by:

$$\hat{R}^i(y^j) = \min_{(e_s,e_q) \neq (e_s^j(y^j), e_q^j(y^j))} R^i(e_s, e_q) \tag{37}$$

subject to:

$$R^j(e_s, e_q) \leq R^j(y^j), \tag{38}$$

$$y^j \leq h(e_s, e_q)\bar{\theta}. \tag{39}$$

The first constraint (38) captures the fact that the deviating strategy is feasible insofar as it induces also type-$j$ agents to change their effort vector. The second constraint (39) ensures that the effort vector is enough to deliver a non-negative profit for the hiring firm in a pooling equilibrium where both agents are paid according to the average productivity $\bar{\theta}$.

The following Lemma shows that this kind of off-equilibrium deviation is only feasible for type 2 agents.

**Lemma 1.** *Type-1 agents cannot succeed in earning $y^2$ while being remunerated according* *to the average productivity $\bar{\theta}$.*

**Proof** Under the suggested deviating strategy both types of workers would earn $y^2$ while being paid according to the average productivity $\bar{\theta}$ and exerting the same effort vector $(e_s, e_q)$ satisfying $h(e_s, e_q)\bar{\theta} \geq y^2$. However, recalling that the equilibrium effort vector chosen by type-2 agents at the income level $y^2$ is efficient, type 2 agents cannot be induced to prefer such an allocation. The reason is that, since they would be remunerated according to $\bar{\theta}$ rather than according to their true productivity $\theta^2$, they would necessarily be forced to adopt a more costly effort vector.[17] $\square$

The next Lemma shows that the off-equilibrium strategy with cost $\hat{R}^i(y^j)$ is always superior (in the sense of being less costly) for type-2 agents relative to the mimicking strategy of replicating the effort vector chosen in equilibrium by agents of type 1.

**Lemma 2.** $\hat{R}^2(y^1) < \breve{R}^2(y^1)$.

**Proof** Let $(e_s^1(y^1), e_q^1(y^1))$ denote the effort vector chosen by type-1 agents at the bundle intended for them by the government, and let $\bar{e}_s = e_s^1(y^1) - \epsilon$ and $\bar{e}_q = e_q^1(y^1) - \epsilon$, for small $\epsilon > 0$, represent a candidate effort vector for a type-2 mimicker. As $\bar{\theta} > \theta^1$ and $y^1 = h(e_s^1(y^1), e_q^1(y^1)) \cdot \theta^1$, it follows by continuity that $h(\bar{e}_s, \bar{e}_q) \cdot \bar{\theta} > y^1$. Hence, the suggested effort vector does not violate the constraint requiring firms to make non-negative profits. By construction, $R^2(\bar{e}_s, \bar{e}_q) < \breve{R}^2(y^1)$ and $R^1(\bar{e}_q, \bar{e}_s) < R^1(y^1)$, so the candidate effort vector is preferred by both types of workers and induces pooling. Moreover, by virtue of the fact that $\hat{R}^2(y^1)$ represents the minimal cost for type 2 under a pooling equilibrium, we have that $\hat{R}^2(y^1) \leq R^2(\bar{e}_s, \bar{e}_q)$. Thus, it follows that $\hat{R}^2(y^1) < \breve{R}^2(y^1)$. This completes the proof. $\square$

Notice that Lemma 2 admits a quite intuitive interpretation. It states that for a type-2 agent it is always more attractive to earn $y^1$, while being remunerated according to the average productivity $\bar{\theta}$, compared to earning $y^1$, while being remunerated according to the productivity $\theta^1 < \bar{\theta}$.

The other deviating strategy that involves the choice of an off-equilibrium effort vector is the one where type-$i$ agents replicate the earned income $y^j$ of type-$j$ agents, but invest in the signals in such a way so as to separate themselves from type $j$ agents, thereby succeeding in being remunerated by firms according to their true productivity $\theta^i$. For a type-$i$ mimicker, the most attractive such deviating strategy is the one with associated

[17]Notice that, if the equilibrium effort vector chosen by type-2 agents at the income level $y^2$ were not efficient, one could no longer rule out the possibility that type-1 agents succeed in earning $y^2$ while being remunerated according to the average productivity $\bar{\theta}$.

cost given by:

$$\tilde{R}^i(y^j) = \min_{(e_s,e_q)\neq(e_s^j(y^j),e_q^j(y^j))} R^i(e_s,e_q) \tag{40}$$

subject to:

$$R^j(e_s,e_q) \geq R^j(y^j), \tag{41}$$

$$y^j \leq h(e_s,e_q)\theta^i. \tag{42}$$

In the problem above, constraint (41) ensures that the effort vector chosen by type-$i$ mimickers is not attractive for type-$j$ agents, thereby allowing type-$i$ mimickers to separate themselves from their type-$j$ counterparts. The constraint (42) ensures instead that the effort vector chosen by type-$i$ mimickers is sufficient to produce $y^j$.

Notice that since $\theta^2 > \theta^1$, and given our assumptions that $p_s^2 = p_s^2 \equiv p_s$ and $p_q^1 > p_q^2$, it necessarily follows that $\tilde{R}^2(y^1) < R^1(y^1)$ and $\tilde{R}^1(y^1) > R^2(y^2)$. Notice also that there are two possible scenarios in which type-$i$ agents succeed in separating themselves from type-$j$ agents at the income level $y^j$: one in which constraint (41) is binding, and another in which it is slack. In the former case, the agent behaving as a mimicker will employ a distorted effort mix (i.e., an effort mix which does not satisfy the condition $\frac{\partial h(e_s,e_q)/\partial e_s}{\partial h(e_s,e_q)/\partial e_q} = \frac{p_s}{p_q^i}$); in the latter case the effort mix chosen by the mimicker will be undistorted.

Using the definitions of $R^i, \breve{R}^i, \hat{R}^i$ and $\tilde{R}^i, i = 1,2$ above, and Lemmas 1 and 2, we can state the optimal tax problem in the presence of two dimensions of signaling:

**Proposition 2.** *In a model with two dimensions of signaling, with only an income tax in place, the optimal tax problem of the government is to choose* $(y^i, c^i), i = 1,2$ *in order to maximize:*

$$u^1 = c^1 - R^1(y^1) \tag{43}$$

*subject to the government budget constraint:*

$$\sum_i \gamma^i (y^i - c^i) = 0, \tag{44}$$

*and the incentive constraints:*

$$c^2 - R^2(y^2) \geq c^1 - \min\left\{\tilde{R}^2(y^1), \hat{R}^2(y^1)\right\}, \tag{45}$$

$$c^1 - R^1(y^1) \geq c^2 - \min\left\{\breve{R}^1(y^2), \tilde{R}^1(y^2)\right\}. \tag{46}$$

In Proposition 2, the expressions for $R^i(y^i), i = 1,2$, are given by (34), the expression for $\breve{R}^1(y^2)$ is given by (36), the expressions for $\hat{R}^2(y^1)$ is given by (37), and the expression for $\tilde{R}^1(y^2)$ is given by (40). A type-$i$ mimicker naturally opts for the least costly choice of the possibilities in curly brackets [as reflected by the min[·] argument on the right hand side of (45) and (46)]. Notice that under both the relevant deviating strategies a type $i$ mimicker obtains a consumption equal to $c^j$. Hence, determining which one of the two deviating strategies dominates requires only comparing the sustained effort cost under each of the two strategies. Notice also that the incentive compatibility constraint for type 2 in (45) is naturally binding in the optimum, by virtue of the strictly egalitarian preferences exhibited by the government.

A final remark is in order. If the difference $p_q^1 - p_q^2$ is sufficiently large, we have that $\widetilde{R}^2(y^1) < \hat{R}^2(y^1)$. This can be seen by considering the limiting case where $p_q^1 \to \infty$ and $p_q^2 \to 0$, in which type-2 agents can separate themselves from type-1 agents (conditional on earning the same level of gross income $y^1$) by investing heavily in the quality component and incurring arbitrarily small costs. In this scenario, type-1 workers would entail arbitrarily large costs by following the same strategy, and hence would find it unattractive. Moreover, as compared to adopting a deviating strategy that is attractive to both types of agents, type 2 agents incur a lower effort cost when they replicate the earned income of type 1 agents, but invest in signals so as to distinguish themselves from their type 1 counterparts. This observation highlights a crucial difference between the current two-signal model and the previous one-signal model. The fact that there are two signals that can be used by type-2 workers, combined with the fact that type-2 workers have a comparative advantage (in the quality dimension) makes it possible (and under certain parametric assumptions optimal) for type-2 workers to separate themselves from their type-1 counterparts (conditional on choosing a given level of income). This could not be done with a single signal in place, as any reduction in the level of investment in the signal would be desirable for type-1 as well.[18]

### 3.2.2 A pooling equilibrium

As an alternative to a separating equilibrium the government might try to implement a pooling equilibrium, in which both types are offered the same (pre-tax) income - consumption bundle. Given our assumption that there are no exogenous public revenue needs, under a pooling equilibrium, the income tax system offers the same pre-tax income $\hat{y}$ to both types of agents, which also coincides with the net income/consumption, denoted by $\hat{c}$.

In the model with one dimension of signaling discussed in Section 2, pooling on income necessarily implied pooling on the signal. With two-dimensional signaling, instead, it is possible to have pooling on income without pooling in terms of the effort vectors chosen

[18]Here we assume that the quantity price is the same for both types whereas the quality price faced by type-2 workers is lower. One could alternatively assume that both prices are lower for type-2 and satisfy: $p_n^1 = kp_n^2$, where $k > 1$ and $n = s, q$. In such a scenario, type-2 would have an absolute advantage in acquiring the signal (thereby ensuring the single crossing property, which is essential for separation). However, no comparative advantage emerges, as both types are faced with the same quality/quantity price ratio. In such a case separation (conditional on choosing a given level of income) is infeasible.

by the two agents. The following Proposition shows that such an equilibrium will never constitute the social optimum.

**Proposition 3.** *In the model with signaling in two dimensions, and only an income tax in place, pooling on income without pooling on the effort signals observed by firms can never constitute the social optimum.*

**Proof** We begin by noting that pooling on income without pooling on the effort signals observed by firms entails no redistribution as both workers would have the same pre-tax income and would be remunerated according to the wage rates commensurate with their true ability. This implies that redistribution is carried out neither through the income channel nor through the wage channel. If the laissez-faire equilibrium is efficient, which happens when the incentive constraint of type 1 agents in the laissez-faire is non-binding, pooling on income without pooling on signals would be Pareto-dominated by the laissez-faire allocation. The reason is that such a pooling allocation distorts the effort/consumption bundle of at least one of the types without gaining anything on the equity side.[19] If, on the other hand, the incentive constraint of type 1 agents is binding at the laissez-faire equilibrium, such that the effort choices of type 2 agents are distorted in the laissez-faire, it is possible that pooling on income could mitigate this distortion. However, since pooling on income without pooling on signals cannot achieve any redistribution, this efficiency gain would only benefit type 2 agents, and would hence not contribute to social welfare given our focus on the max-min social welfare function.[20] □

Given the result stated in Proposition 3, one can restrict attention to pooling equilibria where all agents choose the same effort vector. In this case, workers become indistinguishable from the perspective of the firm and redistribution is accomplished through the wage channel (as both workers receive a wage rate equal to the average productivity $\bar{\theta}$). However, as Proposition 4 below shows, such a pooling equilibrium does not exist.

**Proposition 4.** *In the model with two dimensions of signaling, with only an income tax in place, a pooling equilibrium where both workers choose the same effort vector does not exist.*

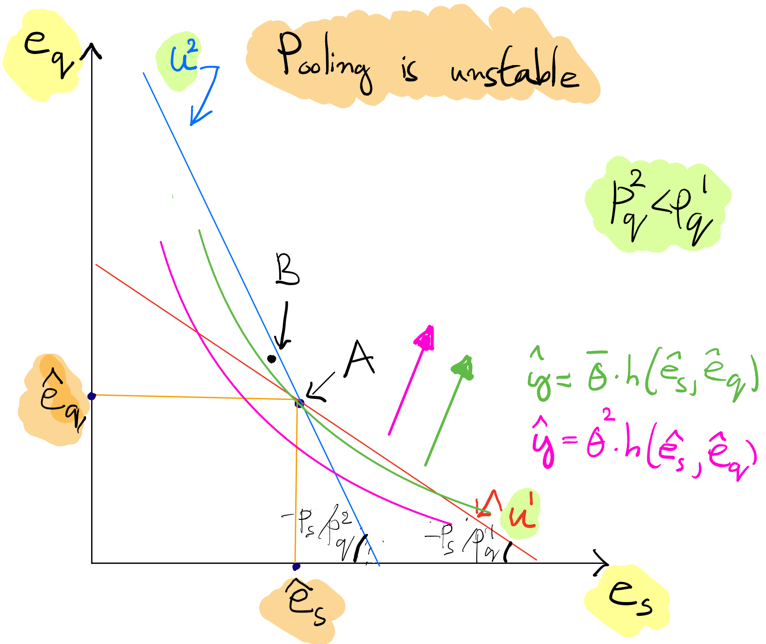
**Proof** Consider a candidate pooling allocation $(\hat{y}, \hat{c})$. By virtue of Proposition 3, we can restrict attention to situations where both workers choose the same effort mix, given by the tuple $(\hat{e}_s, \hat{e}_q)$. By construction, we have that $\hat{c} = \hat{y} = \bar{\theta} h(\hat{e}_s, \hat{e}_q)$, with $\bar{\theta} \equiv \sum_i \gamma^i \theta^i$. Let $\hat{u}^i = u^i(\hat{c}, \hat{e}_s, \hat{e}_q)$. Then, $\hat{c} - (p_s e_s^i + p_q^i e_q^i) = \hat{u}^i$; $i = 1,2$, will describe the indifference curves,

[19] The argument bears similarity to the standard argument why bunching is never optimal in the two-type case with a standard Mirrleesian setup without asymmetric information between firms and workers, see e.g., Stiglitz (1982).

[20] With a social welfare function that attaches a positive weight to type 2 agents, the potential efficiency gain accruing to type 2 agents would have to be weighed against the distortions imposed by the pooling allocation on the choices of type 1 agents. In this case, pooling on income without pooling on the observable effort signal can be ruled out by continuity provided the weight on type 2 agents is sufficiently small.

in the $(e_s, e_q)$ plane, going through the point $(\hat{e}_s, \hat{e}_q)$ indicated by $A$ in figure 1. By virtue of our parametric assumptions, the indifference curve associated with type-2 workers is steeper than that associated with their type-1 counterparts. The intersection of the two downward sloping indifference curves creates a fork-shaped area to the north-west of point $A$ in the figure. Now consider a slight shift from point $A$ to $B$ which lies within the aforementioned fork-shaped area. By construction, point $B$ is preferred by type-2 workers to the pooling allocation (it lies below their indifference curve), whereas type-1 workers strictly prefer the pooling allocation $A$ over the perturbed allocation, $B$ (it lies above his indifference curve). Thus, by deviating to point $B$, type-2 workers credibly reveal their productivity to the firm and gets remunerated accordingly. As $\theta^2 > \bar{\theta}$, it follows, by continuity, that the total output produced by a deviating type-2 worker would strictly exceed $\hat{y}$. Thus, the firm would find it profitable to hire the deviating type-2 worker. We have therefore established that the candidate pooling equilibrium is unstable.

Figure 1: Illustration of Proposition 4 and the non-existence of a pooling equilibrium




□

An immediate corollary of Proposition 3 and 4 is that, under a pure income tax regime, the optimal solution is given by a separating equilibrium in which types 1 and 2 earn different levels of income. This represents a crucial difference between the current model with signaling in two dimensions and the model with one signal considered in section 2.

**Corollary 1.** *In the model with two dimensions of signaling, with only an income tax in place, the social optimum is always given by a separating equilibrium.*

## 3.3 The extended tax regime

We now extend our tax base and allow the government to tax each individual based on two observable characteristics: (i) earned income, $y$; (ii) the quantity of effort (e.g., years of education, hours physically spent at work), $e_s$. Recall that by assumption, the intensity (quality) dimension of effort $e_q$ is only observable by the firms, and hence, cannot be subject to taxation. We maintain all our earlier assumptions, and in particular, the asymmetry in information between the firms and the workers with respect to the latter's innate productive ability. An immediate and interesting implication of allowing to tax $e_s$ is the existence of a pooling equilibrium. By conditioning the tax on both $y$ and $e_s$, the government effectively confines signaling to one dimension, implying that pooling is rendered feasible, and a setting similar to the one-dimensional signaling case analyzed in section 2.[21]

**The pooling equilibrium in the extended tax regime** A government seeking to implement a pooling equilibrium in the extended tax regime chooses $(y, e_s)$ in order to maximize:

$$u^1 = y - R^1(e_s, \widehat{e}_q(y, e_s)), \tag{47}$$

where $\widehat{e}_q(y, e_s)$ is defined as the value of $e_q$ which solves the equation

$$y = h(e_s, e_q)\bar{\theta}. \tag{48}$$

Notice that the equity gains of pooling crucially depend on the difference in productivities between the two types of workers whereas the efficiency properties of a pooling equilibrium crucially depend on the differences in the costs of acquiring the signal $e_q$.

**The separating equilibrium in the extended tax regime** Let us now turn to the separating equilibrium. The first thing to notice is that with a tax on $e_s$, the deviations associated with $\tilde{R}$ in equations (45) and (46) are rendered infeasible. For type-2 agents, the best deviating strategy will be to pool with type 1 agents, obtaining the wage rate $\bar{\theta}$. However, due to the tax on $e_s$, a type-2 mimicker is forced to replicate the quantity-signal $e_s$ of type-1 agents ($e_s^1$), making mimicking less attractive. For type-1 agents, the tax on $e_s$ implies that the only feasible deviating strategy is to replicate the effort choices of type-2 agents. Formally, a government seeking to implement a separating equilibrium in the extended tax regime chooses $(y^i, c^i, e_s^i), i = 1,2$ in order to maximize:

$$u^1 = c^1 - R^1(e_s^1, e_q^{1*})$$

[21] In the current setting, type-2 workers have a *comparative* advantage in $e_q$, which can be exploited by type-2 workers to separate themselves from their type-1, lower-skilled, counterparts. In the setting considered in section 2, high-skilled workers had an *absolute* advantage in acquiring the (single) signal $e$.

subject to the government budget constraint:

$$\sum_i (y^i - c^i)\gamma^i = 0, \tag{49}$$

and the incentive constraints

$$c^2 - R^2(e_s^2, e_q^{2*}) \geq c^1 - R^2(e_s^1, \hat{e}_q^2) \tag{50}$$
$$c^1 - R^1(e_s^1, e_q^{1*}) \geq c^2 - R^1(e_s^2, e_q^{2*}) \tag{51}$$

where

$$\hat{e}_q^2 \quad \text{is the solution to} \quad y^1 = h(e_s^1, e_q^2)\bar{\theta} \tag{52}$$
$$e_q^{i*} \quad \text{is the solution to} \quad y^i = h(e_s^i, e_q^i)\theta^i, i = 1,2. \tag{53}$$

The incentive constraint (50) and equation (52) reflect that, for type-2 mimickers, the optimal deviating strategy is the off-equilibrium deviation where they pool with type-1 agents at the income level $y^1$. This is a feasible strategy since the effort vector $(e_s^1, \hat{e}_q^2)$ would certainly be attractive to type-1 agents due to the fact that, since $\bar{\theta} > \theta^1$, $\hat{e}_q^2 < e_q^{1*}$. The incentive constraint (51) reflects that the only feasible deviating strategy for type-1 mimickers is for them to earn $y^2$ by replicating the optimal effort level $e_q^{2*}$ of type-2 agents (given that $e_s$ is observable by the government).

**The social optimum** The problem solved by a max-min government is to maximize the well being of type-1 workers. In the presence of a tax on the quantity-signal $e_s$, this is done by implementing either the optimal separating or the optimal pooling equilibrium, depending on which equilibrium configuration yields the highest utility to type-1 agents. We thus have the following:

**Proposition 5.** *In the model with two dimensions of signaling, under the extended tax regime, the social optimum can be either a separating or a pooling equilibrium.*

A formal proof is skipped for brevity purposes, as the problem solved in the presence of a tax on $e_s$ is similar to the government's problem under a pure income tax regime and with a single dimension of signaling (i.e., the problem that was discussed in Section 2.2).

Once again notice that, by implementing a pooling equilibrium, the government forces wage equalization and thereby eliminates all the information rent (derived by high-skilled agents) associated with the difference in productivities (but not the information rent associated with the difference in the acquisition costs of the signal). However, whether or not a pooling equilibrium represents the social optimum will also depend on a comparison between the efficiency properties of the two types of equilibria.

With respect to the efficiency properties of an optimal pooling equilibrium, notice that the first order conditions of the government's problem are given by:

$$1-\frac{\partial R^1}{\partial \widehat{e}_q}\frac{\partial \widehat{e}_q}{\partial y} = 0, \tag{54}$$

$$-\frac{\partial R^1}{\partial e_s}-\frac{\partial R^1}{\partial \widehat{e}_q}\frac{\partial \widehat{e}_q}{\partial e_s} = 0, \tag{55}$$

which can, respectively, be rewritten as

$$1-p_q^1\frac{\partial \widehat{e}_q}{\partial y} = 0, \tag{56}$$

$$-p_s+p_q^1\frac{\partial h\left(e_s,\widehat{e}_q\right)/\partial e_s}{\partial h\left(e_s,\widehat{e}_q\right)/\partial e_q} = 0. \tag{57}$$

Eq. (57) shows that, at the optimal pooling equilibrium, type-1 agents choose an undistorted effort mix. Given that $p_q^1 > p_q^2$, eq. (57) also implies that at the optimal pooling equilibrium, type-2 agents are forced to choose a distorted effort mix, with $e_q^2 = \widehat{e}_q$ being distorted downwards.[22] Instead, as we show in appendix B, at an optimal *separating* equilibrium the following conditions hold:

$$p_s-p_q^1\frac{\partial h\left(e_s^1,e_q^{1*}\right)/\partial e_s}{\partial h\left(e_s^1,e_q^{1*}\right)/\partial e_q} > 0, \tag{58}$$

$$p_s-p_q^2\frac{\partial h\left(e_s^2,e_q^{2*}\right)/\partial e_s}{\partial h\left(e_s^2,e_q^{2*}\right)/\partial e_q} \leq 0. \tag{59}$$

According to inequality (58), the effort mix chosen by type-1 agents is distorted towards $e_s$, whereas according to (59) the effort mix chosen by type-2 agents is either left undistorted (which happens when the constraint (51) is slack) or is distorted towards $e_q$ (when the constraint (51) is binding). Keeping in mind that type-2 agents have a comparative advantage in the effort/signal $e_q$ (and, therefore, type-1 agents have a comparative advantage in the effort/signal $e_s$), this pattern of distortions is coherent with the goal of relaxing the binding incentive-compatibility constraints. Given that constraint (50) is necessarily binding under a max-min objective, distorting the effort mix of type-1 agents towards the component $e_s$ is instrumental in discouraging mimicking by type-2 agents. Conversely, distorting the effort mix of type-2 agents towards the component $e_q$ is instrumental in discouraging mimicking by type-1 agents. However, given that constraint (51) may either be binding or not, distorting the effort mix of type-2 agents is only warranted when

[22] Notice that eq. (56) can also be equivalently rewritten as $1-\frac{p_q^1}{\bar{\theta}\partial h(e_s,e_q)/\partial e_q}=0$. Since an undistorted level of income would satisfy the condition $1-\frac{p_q^1}{\theta^1\partial h(e_s,e_q)/\partial e_q}=0$ for type-1 agents and $1-\frac{p_q^2}{\theta^2\partial h(e_s,e_q)/\partial e_q}=0$ for type-2 agents, it also follows that at the optimal pooling equilibrium $y$ is distorted upwards for type-1 agents and downwards for type-2 agents.

mimicking by type-1 agents is a relevant threat.[23]

As a final remark, we would like to emphasize a crucial difference between Proposition 5 and Proposition 1. The latter, which we presented in section 2, showed that in the extended tax regime, when signaling is confined to one dimension, the optimum is always given by a separating equilibrium. Proposition 5 shows, in contrast, that when signaling is carried out along two dimensions, the optimum in the extended tax regime can be either a separating or a pooling equilibrium. The reason for this difference is that, when there are multiple signaling possibilities, the information rents associated with the difference in productivities are not fully eliminated under the separating equilibrium in the extended tax regime.

# 4 Numerical characterization

In this section, we make a functional form assumption. This assumption allows us to shed more light on which incentive-compatibility constraints that are relevant in the bi-dimensional signaling case when the government aims at implementing a separating equilibrium. In particular, we assume that:

$$h(e_s, e_q) = \left(e_s e_q\right)^{\beta} A, \tag{60}$$

where $A$ is a positive constant and $0 < \beta < 1/2$. Assumption (60) will also be used to provide a numerical example of the results delivered by our model.

## 4.1 (In)efficiency of the laissez-faire

Consider what would be an efficient laissez-faire outcome if agents were remunerated based on their true productivities. A type $i$ agent would then solve the following problem:

$$\max_{e_q, e_s} \left\{ (e_q e_s)^{\beta} A\theta^i - p_s e_s - p_q^i e_q \right\}. \tag{61}$$

[23]In the Appendix we also show that an optimal separating equilibrium satisfies the conditions $1 - \frac{p_q^1}{\theta^1 \partial h(e_s^1, e_q^{1*})/\partial e_q} > 0$ and $1 - \frac{p_q^2}{\theta^2 \partial h(e_s^2, e_q^{2*})/\partial e_q} \leq 0$. The first inequality states that the income earned by type-1 agents, i.e. $y^1$, is distorted downwards (to deter mimicking by type-2 agents). The second inequality states that the income earned by type-2 agents, i.e. $y^2$, is either left undistorted (when the constraint (51) is slack) or is distorted upwards (the constraint (51) is binding).

Denoting by $e_q^i$ and $e_s^i$ the optimal value for the choice variables, from the first order conditions of the problem above one would get that

$$e_q^i = (p_s)^{-\frac{\beta}{1-2\beta}} \left(p_q^i\right)^{\frac{(\beta-1)}{1-2\beta}} \left(\beta A\theta^i\right)^{\frac{1}{1-2\beta}}, \tag{62}$$

$$e_s^i = (p_s)^{-\frac{1-\beta}{1-2\beta}} \left(p_q^i\right)^{-\frac{\beta}{1-2\beta}} \left(\beta A\theta^i\right)^{\frac{1}{1-2\beta}}, \tag{63}$$

implying that

$$U^i = \left(p_s p_q^i\right)^{-\frac{\beta}{1-2\beta}} \left(A\theta^i\right)^{\frac{1}{1-2\beta}} \beta^{\frac{2\beta}{1-2\beta}} \left(1-2\beta\right). \tag{64}$$

However, to assess whether or not the laissez-faire equilibrium will indeed fulfil the efficiency conditions (62)-(63), we need to take a closer look at the incentives faced by type-1 agents. In particular, we need to check that type-1 agents would not be better off by replicating the effort choices made by type-2 agents in order to get remunerated according to the productivity of their higher-skilled counterparts. By behaving as mimickers, type-1 agents would obtain a utility given by:

$$\widehat{U}^1 = (e_q^2 e_s^2)^{\beta} A\theta^2 - p_s e_s^2 - p_q^1 e_q^2. \tag{65}$$

Substituting for $e_q^2$ and $e_s^2$ in (65) the values provided, for $i = 2$, by (62)-(63) gives:

$$\begin{aligned}\widehat{U}^1 \quad = \quad & \left(p_s p_q^2\right)^{-\frac{\beta}{1-2\beta}} \left(A\theta^2\right)^{\frac{1}{1-2\beta}} \beta^{\frac{2\beta}{1-2\beta}} \\ & -(p_s)^{-\frac{\beta}{1-2\beta}} \left(p_q^2\right)^{-\frac{\beta}{1-2\beta}} \left(\beta A\theta^2\right)^{\frac{1}{1-2\beta}} \\ & -(p_s)^{-\frac{\beta}{1-2\beta}} p_q^1 \left(p_q^2\right)^{-\frac{1-\beta}{1-2\beta}} \left(\beta A\theta^2\right)^{\frac{1}{1-2\beta}}.\end{aligned} \tag{66}$$

Thus, a type-1 agent will have no incentive to mimic a type 2 agent, and the laissez-faire equilibrium will be efficient, provided that $U^1 \geq \widehat{U}^1$, an inequality that can be rewritten (using (64), for $i = 1$, and (66)) as

$$\left(\frac{\theta^1}{\theta^2}\right)^{\frac{1}{1-2\beta}} \geq \frac{1-\beta-\frac{p_q^1}{p_q^2}\beta}{1-2\beta} \left(\frac{p_q^1}{p_q^2}\right)^{\frac{\beta}{1-2\beta}}. \tag{67}$$

A sufficient condition for (67) to be satisfied is that $1-\beta-\frac{p_q^1}{p_q^2}\beta \leq 0$, i.e. $\frac{p_q^1}{p_q^2} \geq \frac{1-\beta}{\beta}$. This highlights that efficiency under laissez-faire requires that the difference in the cost of acquiring the quality signal between the two agents is sufficiently large.

## 4.2 The income tax regime

Exploiting assumption (60) allows us to obtain closed-form expressions for the incentive constraints in equations (45)-(46). To achieve this goal, we begin by deriving the effort cost sustained by agents who choose the point on the income tax schedule intended for them.

**Choices of a truthfully reporting agent** Consider a type $i$ agent who earns the income level $y^i$ intended for him/her by the government. This agent will choose an efficient mix of $e_s$ and $e_q$, and incur an effort cost equal to:

$$R^i(y^i) = \min_{e_s,e_q} R^i(e_s,e_q) \quad \text{subject to} \quad \left(e_s e_q\right)^\beta A\theta^i = y^i. \tag{68}$$

The optimal effort-choices are:

$$e_s^i\left(y^i\right) = \sqrt{\left(\frac{y^i}{A\theta^i}\right)^{1/\beta}\frac{p_q^i}{p_s}}, \tag{69}$$

$$e_q^i\left(y^i\right) = \sqrt{\left(\frac{y^i}{A\theta^i}\right)^{1/\beta}\frac{p_s}{p_q^i}}. \tag{70}$$

To see this, notice that under assumption (60) the cost-minimizing input mix satisfies the condition $e_q = \frac{p_s}{p_q^i}e_s$. This, combined with the equation $\left(e_s e_q\right)^\beta A\theta^i = y^i$, gives the two equations above. Insertion of (69)-(70) into the cost function yields:

$$R^i(y^i) = p_s\sqrt{\left(\frac{y^i}{A\theta^i}\right)^{1/\beta}\frac{p_q^i}{p_s}} + p_q^i\sqrt{\left(\frac{y^i}{A\theta^i}\right)^{1/\beta}\frac{p_s}{p_q^i}} = 2\sqrt{\left(\frac{y^i}{A\theta^i}\right)^{1/\beta}p_s p_q^i}, \quad i = 1,2. \tag{71}$$

**Optimal deviating strategies** In appendix C, we show that three cases need to be distinguished in order to determine which of the deviating strategies embedded in the incentive constraints (45)-(46) that are relevant.[24] In appendix C, we also show that the cases can be distinguished using conditions that depend on the ratio $\theta^2/\theta^1$, the proportions of agents $\gamma^1$ and $\gamma^2$, and a constant defined as:

$$\Omega \equiv \left[\frac{1}{2}\frac{p_q^2 + p_q^1}{\sqrt{p_q^2 p_q^1}}\right]^{2\beta}. \tag{72}$$

**Case 1: $\frac{\theta^2}{\theta^1} \le \Omega$** In this case, we need to take into account the following *downward* incentive-compatibility constraint associated with the cost $\widetilde{R}^2\left(y^1\right)$ (given by equation (C4)

[24]In our numerical example, we will vary parameters in such a way that all three cases are considered.

in appendix C):

$$c^2 - R^2(y^2) \geq c^1 - 2\sqrt{\left(\frac{y^1}{A\theta^2}\right)^{1/\beta} p_q^2 p_s}, \tag{73}$$

and the following *upward* constraint (associated with the cost $\widetilde{R}^1\left(y^2\right)$ in equation (C25) in appendix C):

$$c^1 - R^1(y^1) \geq c^2 - 2\sqrt{\left(\frac{y^2}{A\theta^1}\right)^{1/\beta}} \sqrt{p_q^1 p_s}. \tag{74}$$

**Case 2:** $\gamma^1 + \gamma^2 \frac{\theta^2}{\theta^1} < \Omega < \frac{\theta^2}{\theta^1}$ In this case, we need to take into account the following *downward* incentive-compatibility constraint (associated with the cost $\widetilde{R}^2\left(y^1\right)$ in equation (C9) in appendix C):

$$c^2 - R^2(y^2) \geq c^1 - \sqrt{p_s p_q^1} \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}} \left(\frac{1}{\theta^2}\right)^{1/\beta}}{\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}} \tag{75}$$

$$-p_q^2 \sqrt{\frac{p_s}{p_q^1}} \sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}} \left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}\right], \tag{76}$$

and the following *upward* incentive-compatibility constraint (associated with the cost $\breve{R}^1\left(y^2\right)$ in equation (C22) in appendix C):

$$c^1 - R^1(y^1) \geq c^2 - \sqrt{\left(\frac{y^2}{A\theta^2}\right)^{1/\beta}} \frac{p_q^2 + p_q^1}{\sqrt{p_q^2}} \sqrt{p_s}. \tag{77}$$

**Case 3:** $\gamma^1 + \gamma^2 \frac{\theta^2}{\theta^1} \geq \Omega$ In this case, we need to take into account two *downward* incentive-compatibility constraints. The first, associated with the cost $\widetilde{R}^2(y^1)$ provided by (C9) in appendix C, is given by:

$$c^2 - R^2(y^2) \geq c^1 - \sqrt{p_s p_q^1} \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}} \left(\frac{1}{\theta^2}\right)^{1/\beta}}{\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}}$$

$$-p_q^2 \sqrt{\frac{p_s}{p_q^1}} \sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}} \left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}\right]. \tag{78}$$

The second, associated with the cost $\hat{R}^2(y^1)$ provided by (C14) in appendix C), is given by:

$$c^2 - R^2(y^2) \geq c^1 - 2\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}}\sqrt{\left(\frac{1}{\gamma^1\theta^1+\gamma^2\theta^2}\right)^{1/\beta} p_s p_q^2}. \tag{79}$$

In addition, we need to take into account the following *upward* incentive-compatibility constraint (associated with the cost $\breve{R}^1\left(y^2\right)$ in (C22) in appendix C):

$$c^1 - R^1(y^1) \geq c^2 - \sqrt{\left(\frac{y^2}{A\theta^2}\right)^{1/\beta}}\frac{p_q^2+p_q^1}{\sqrt{p_q^2}}\sqrt{p_s}. \tag{80}$$

## 4.3 The problem of the government in the extended tax regime

In the extended tax regime, obtaining closed form solutions to the incentive constraints using the functional form assumption in (60) is more straightforward. We begin with the separating equilibrium. In this case, equations (52) and (53) take the form:

$$e_q^{i*} = \left(\frac{y^i}{A\theta^i}\right)^{1/\beta}\frac{1}{e_s^i}, \quad i = 1,2 \tag{81}$$

$$\hat{e}_q = \left(\frac{y^1}{A\bar{\theta}}\right)^{1/\beta}\frac{1}{e_s^1}. \tag{82}$$

Thus, the incentive constraints (50)-(51) can be written as follows:

$$c^1 - p_s^1e_s^1 - \left(\frac{y^1}{A\theta^1}\right)^{1/\beta}\frac{p_q^1}{e_s^1} \geq c^2 - p_s^1e_s^2 - \left(\frac{y^2}{A\theta^2}\right)^{1/\beta}\frac{p_q^1}{e_s^2}, \tag{83}$$

$$c^2 - p_s^2e_s^2 - \left(\frac{y^2}{A\theta^2}\right)^{1/\beta}\frac{p_q^2}{e_s^2} \geq c^1 - p_s^2e_s^1 - \left(\frac{y^1}{A\bar{\theta}}\right)^{1/\beta}\frac{p_q^2}{e_s^1}. \tag{84}$$

In a pooling equilibrium, following equations (47) and (48), the government chooses $(y, e_s)$ in order to maximize

$$u^1 = y - p_s^1e_s + p_q^1\widehat{e}_q(y,e_s), \tag{85}$$

where $\widehat{e}_q(e_s,y)$ is the value of $e_q$ which solves the equation $y = (e_se_q)^\beta A\bar{\theta}$.

When $h(e_s,e_q) = \left(e_se_q\right)^\beta A$, the first order conditions (56)-(57) can be respectively restated as follows:

$$1 - \frac{p_q^1}{\beta A\bar{\theta}e_s}\left(\frac{y}{A\bar{\theta}}\right)^{(1-\beta)/\beta} = 0, \tag{86}$$

$$-p_s + \frac{p_q^1}{(e_s)^2}\left(\frac{y}{A\bar{\theta}}\right)^{1/\beta} = 0. \tag{87}$$

Thus, combining (86)-(87) one obtains:

$$\frac{e_s}{e_q} = \frac{p_q^1}{p_s}, \tag{88}$$

$$y = \left( \frac{p_s p_q^1}{\left(A\bar{\theta}\right)^{\frac{1}{\beta}} \beta^2} \right)^{\frac{\beta}{2\beta - 1}}. \tag{89}$$

## 4.4 Results

We fix the productivity of type 2 to $\theta^2 = 100$ and compute the social welfare level, distortions, and income levels in the various tax regimes when $\theta^1$ varies between 1 and 100. In this way, we consider a wide range of values for the ratio $\theta^1/\theta^2$. We maintain our previous normalization $p_s^1 = p_s^2 = p_q^2 = 1$ and set $\beta = 0.25$, $A = 1$. We consider two scenarios regarding the difference in the cost of acquiring $e_q$ between the two types of agents. In the first scenario, we consider a small value of $p_q^1 - p_q^2$, letting $p_q^1 = 1.05$. In the second scenario we consider a large value of $p_q^1 - p_q^2$, letting $p_q^1 = 1.5$.[25]

In our simulations, we report the maximum achievable welfare gain. This maximum welfare gain is computed at the value of $\theta^1$ where the difference between the social welfare level in the extended regime and the social welfare level in the income tax regime is the largest, and we focus on an equivalent-variation type of welfare measure.[26]

[25]The qualitative features of the results are robust to changes in $p_q^1$ and $\beta$. Additional simulations are available upon request.

[26]The equivalent-variation type welfare gain is obtained by first calculating the wage level $\theta^1$ at which the vertical distance between the social welfare levels of the extended tax regime and the income tax regime is the largest. We then compute the minimal amount of resources that must be injected into the income tax regime in order to reach the social welfare level of the extended tax regime (repeatedly solving the government optimization program). We then divide this minimum amount of revenue by the aggregate output of the income tax regime in order to obtain a welfare gain measure expressed as a fraction of output.

Figure 2: Numerical illustration, small difference in $p_q^1 - p_q^2$

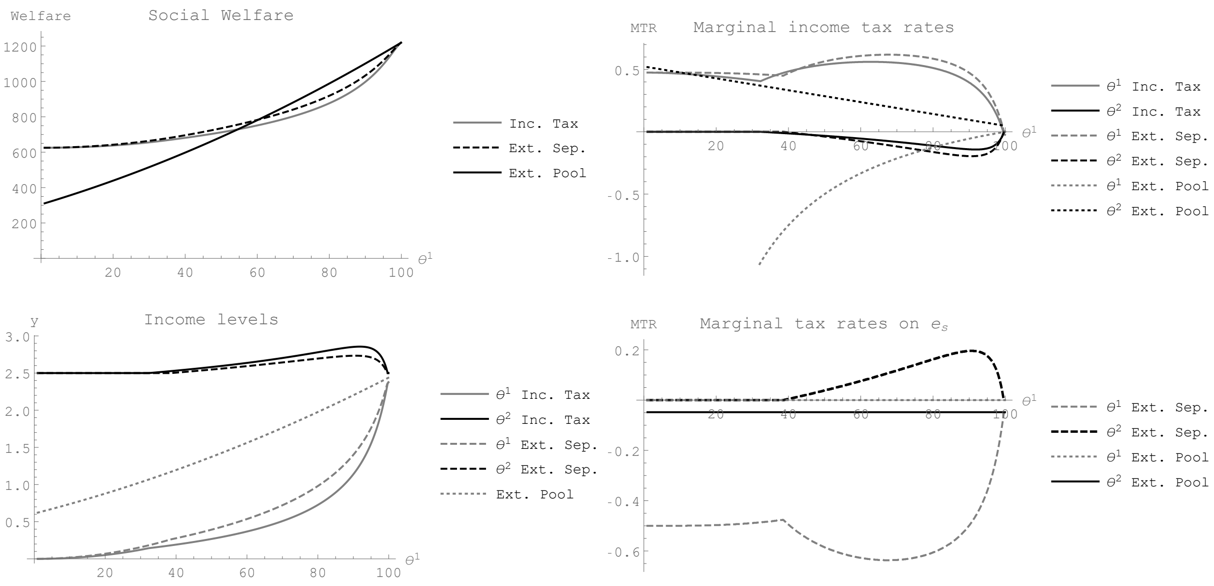



Figure 2 shows the results for the case when the difference $p_q^1 - p_q^2$ is small. As expected, we see from the top left panel, that the extended tax regime always welfare-dominates the income tax regime. More interestingly, we see that the it is optimal to implement a separating allocation when $\theta^1$ takes on low and intermediate values, whereas the pooling allocation dominates when $\theta^1$ is relatively close to $\theta^2$.[27] The maximum welfare gain of taxing the signal is obtained at $\theta^1 = 84.1$, amounts to 6.31% of aggregate output, and is associated with the implementation of a pooling allocation.

Turning to the top right panel of figure 2, we see that the income tax regime and the extended separating regime always imply a positive marginal tax rate on low-skilled agents. This is a standard property in optimal income tax models and serves to mitigate the binding downwards incentive compatibility constraint. The negative marginal tax rate is non-standard and is due the binding upwards incentive constraint. It can also be noted that the earned income distortions in the separating extended regime are more pronounced than in the income tax regime. In the extended pooling regime, a positive marginal tax rate is levied on high-skill agents and a negative marginal tax rate is levied on low-skill agents. This is due to the cross-subsidiziation between the two types in the pooling allocation. Moreover, since the degree of cross-subsidization decreases when $\theta^1$ approaches $\theta^2$, the marginal tax rates converge when $\theta^1$ approaches $\theta^2$.

The bottom-left panel shows the income levels in the different optimal tax regimes. Not surprisingly, the income levels associated with the pooling regime lie in between the income levels of the tax regimes where the labor market equilibrium is separating. Moreover, the income levels of the two agents converge as the two agents become more

[27]Notice that when $\theta^1$ is very close to 100, the separating allocation dominates, even though it is not visible in the figure. This is a knife-edge case of no practical relevance.

similar in terms of $\theta$.

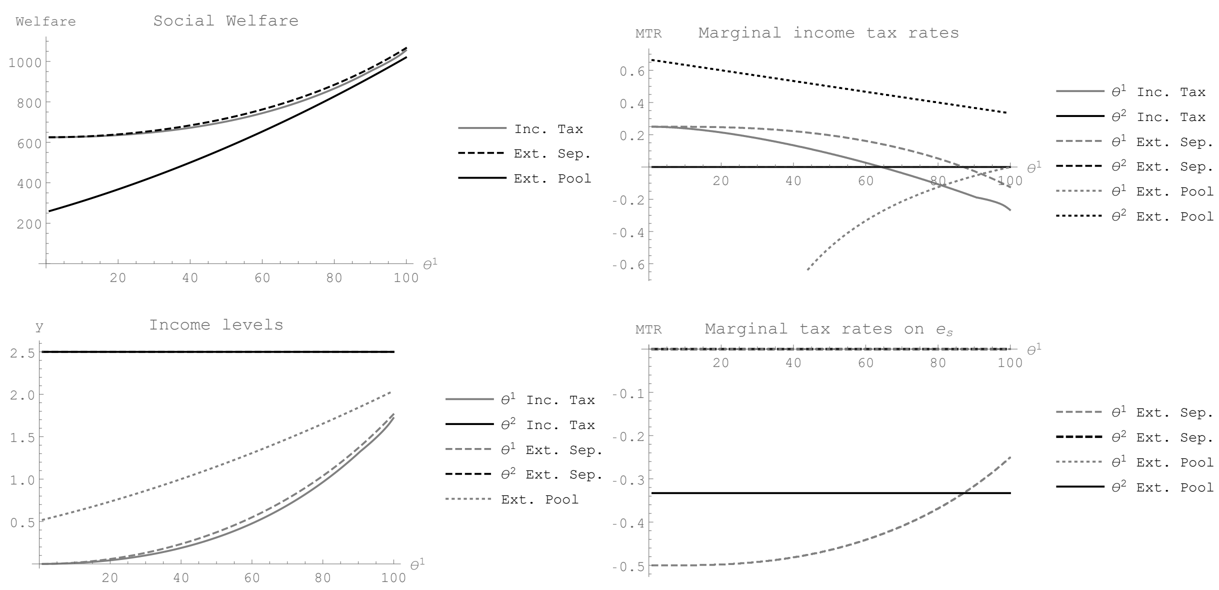
Finally, we turn to the bottom-right panel. Notice that, in this panel, we are only representing the graphs of the extended tax regime. The reason is that, by construction, the mix of effort inputs is always undistorted with only an income tax in place. In the separating extended tax regime, we note that it is always desirable to subsidize the observable dimension of education effort $e_s$ of type 1 agents (reflected in an upwards-distortion). The reason is that subsidizing the dimension of education effort in which the low-skill agent has a comparative advantage, mitigates the incentive constraint of the high-skill type. For the high-skill type, it is desirable to tax $e_s$ over the range of values of $\theta^1$ for which the upwards incentive constraint is binding. This can be understood from the fact that by taxing the quantity dimension, we are implicitly subsidizing the *quality* dimension in which the high-skill type has a comparative advantage. This serves to mitigate the binding incentive constraint of the low-skill type. In the pooling regime, there is by definition no binding incentive constraint. As we are invoking a Rawlsian welfare function, the optimum is simply to achieve the best effort mix from the perspective of the low skilled type. This implies an efficient effort mix for the low skilled type and, hence, an upwards distortion for the high skilled types, due to their comparative advantage in the quality dimension of effort.

The numerical example illustrates that both education subsidies and taxes are warranted on redistributive grounds to mitigate binding incentive compatibility constraints. Notably, and contrary to conventional wisdom, a tax on education is shown to be desirable. It serves to implement a separating allocation, via mitigating the incentives of the low-skilled workers to mimic their higher-skilled counterparts, when employers are unable to observe the true productivity of their employees.

Figure 3 shows the results for the case when the difference $p_q^1 - p_q^2$ is relatively large. Again, as expected, the extended tax regime always welfare-dominates the income tax regime. However, in contrast to the case when $p_q^1 - p_q^2$ is relatively small, we can see that pooling is always suboptimal. This reflects the fact that, as the difference in the costs of acquiring the signals becomes larger and larger, it becomes less likely that an optimal separating equilibrium features a double distortion with both a downward and an upward binding incentive-compatibility constraint. In particular, when the difference in the costs of acquiring the signals is sufficiently large, the optimal separating equilibrium will preserve efficiency at the top (and only entail a distortion on the behavior of low-skilled agents). In Figure 3 this is shown by the fact that both in the top-right- and bottom-right panel, the black dashed line coincides with the horizontal axis (in contrast to what happens in Figure 2). The maximum welfare gain of taxing the signal is obtained at $\theta^1 = 71.0$, amounts to 1.20% of aggregate output, and is associated with the implementation of a separating allocation.

As we highlighted in Section 3.3, where we compared the relative merits of pooling versus separating equilibria, the attractiveness of implementing a pooling equilibrium hinges on its equity gains, i.e. the fact that it fully eliminates all the information rent (derived by high-skilled agents) associated with the difference in productivities. Obviously, this gain is bigger the larger the information rent, associated with differences in productivities, that is enjoyed by high-skilled agents under an optimal separating equilibrium. On the other hand, whether this equity gain of pooling can be reaped cheaply or not (in efficiency terms), depends on the magnitude of the difference in the costs of acquiring the signals (in our case, the magnitude of the difference $p_q^1 - p_q^2$). This is due to the fact that an optimal pooling equilibrium always violates efficiency at the top. Instead, an optimal separating equilibrium preserves efficiency at the top when the difference $p_q^1 - p_q^2$ is sufficiently large.

Figure 3: Numerical illustration, large difference in $p_q^1 - p_q^2$




# 5 Discussion and Concluding Remarks

In this paper, we have analyzed optimal redistribution in the presence of signaling, introducing two realistic new features to the standard Mirrleesian framework: (i) a second layer of asymmetric information between employers and workers regarding the productive capacity of the workers, with the latter having the possibility of engaging in signaling to credibly convey this information to prospective employers; (ii) the possibility to levy a direct tax on the signals acquired by the workers (in addition to taxing income).

The combination of the two new features maintains the second-best nature of the government optimization problem, and the inherent trade-off between conflicting equity and efficiency considerations. For tractability reasons we have focused attention on a two-type model with high- and low-ability agents; regarding the dimension of signaling, we have considered both a setting with unidimensional signaling and a setting with bi-dimensional signaling. To ensure the possibility that agents credibly signal their true type to prospective employers, we have assumed that low- and high-skilled agents also differ in the cost of acquiring the signal(s).

The presence of asymmetric information between employees and firms implies that the government can influence the labor market equilibrium, implementing either a separating or a pooling equilibrium. It also implies that the incentive-compatibility constraints faced by the government are non-standard. In particular, agents of a given type $i$ often have more than one deviating strategy that allow them to earn the income level intended by the government for agents of type $j \neq i$. Depending on the deviating strategy that is adopted, mimickers may end up being remunerated by firms according to their true productivity, the average productivity, or the productivity of the type that they mimic. By supplementing income taxation with a tax on the signal(s), the government succeeds in shrinking the set of deviating strategies that are available to agents, thereby mitigating (or fully eliminating) the information rents associated with the differences in productivities between workers. This may lead either to extracting a higher fiscal surplus from the high skilled, and therefore enhanced redistribution via the income channel, or to a change in the wage structure, inducing cross-subsidization between skill levels.[28]

We have shown that, when signaling is confined to a single dimension, and the only policy instrument is a nonlinear income tax, the social optimum may either be a separating- or a pooling equilibrium. The desirability of a pooling equilibrium derives from the role played by policy-induced wage compression in realizing redistributive goals. A pooling equilibrium forces wage equalization; therefore, it eliminates all the information rent (derived by high-skilled agents) associated with the difference in productivities (but not the information rent associated with the difference in the costs of acquiring the signal). However, whether or not a pooling equilibrium constitutes the social optimum will also depend on a comparison between the efficiency properties of the two types of equilibria.[29]

With unidimensional signaling, supplementing the income tax with a direct tax levied on the signal implies that a pooling equilibrium is no longer necessary to fully eliminate the information rent associated with differences in productivities. In this case, the social optimum is always given by a separating equilibrium. A pooling equilibrium is suboptimal because its main advantage, from an equity point of view, is the elimination of the information rent associated with differences in productivities. However, this equity gain can more efficiently be obtained by taxing the signals and implementing a separating equilibrium.

[28]With more than two types, a combination of both could be desirable.

[29]A necessary condition for a pooling equilibrium to represent the social optimum is that the information rent associated with the differences in productivities cannot be fully eliminated by implementing a separating equilibrium. The larger this information rent, the more likely it is that a separating equilibrium will feature a double distortion with both a downward- and an upward binding incentive-compatibility constraint. Given that at a pooling equilibrium both high- and low-skilled agents face a distortion, when the same happens at a separating equilibrium, the efficiency properties of the two equilibria become more similar.

With bidimensional signaling, we have shown that, with only an income tax in place, there cannot be any redistribution though wage compression given that a pooling equilibrium fails to exist. Moreover, in contrast to what happens in the setting with unidimensional signaling, supplementing the income tax with a tax on the signal that is observed by the government is not enough to fully eliminate, under a separating equilibrium, the information rent associated with differences in productivities. However, by eliminating the possibility for cream-skimming by firms an extended tax regime can support a pooling equilibrium and, thereby, achieve redistribution through wage compression.

The above remarks highlight a crucial difference between the cases of unidimensional- and bidimensional signaling. In the former, taxing the signal observed by the government allows to fully eliminate the information rent associated with differences in productivities; in the latter, this information rent is only partially mitigated (unless a pooling equilibrium is implemented).[30]

From a policy perspective, our analysis calls for revisiting commonly applied policy tools, such as monitoring of labor hours and education mandates, often warranted on efficiency grounds to address market failures. We argue that these instruments effectively constitute forms of direct taxes levied on signals, viewing labor efforts and educational attainment as signaling devices used by high-skilled workers to separate themselves in the labor market. Our analysis alludes, therefore, to the potentially significant role played by such policy tools in realizing redistributive goals as a supplement to income taxation.

[30]More generally, if the tax liability cannot be conditioned on the entire set of signals, taxing the signals can only mitigate the information rent stemming from differences in productivities. In this general case, one might be tempted to argue that the implementation of a pooling equilibrium might still represent, based on equity considerations, an appealing option from the perspective of the government. Furthermore, one might also be tempted to argue that the case for implementing a pooling equilibrium becomes stronger the smaller the set of signals on which the tax liability is conditioned. The problem with this reasoning is that it overlooks the fact that a pooling equilibrium must necessarily be sustainable in order to be implementable. In general, denoting by $n$ the cardinality of the set of signals, pooling will be sustainable either when the (minimum) number of taxed signals is $n-1$, or when it is $n-j$ (with $1<j<n$) and the high-skilled individuals have no comparative advantage within the set of signals unobserved (and therefore untaxed) by the government.

## References


Akerlof, G. (1976). The economics of caste and of the rat race and other woeful tales. *The Quarterly Journal of Economics*, 90(4):599–617.

Balestrino, A., Grazzini, L., and Luporini, A. (2016). A normative justification of compulsory education. *Journal of Population Economics*, 30:537–567.

Bastani, S., Blumkin, T., and Micheletto, L. (2015). Optimal wage redistribution in the presence of adverse selection in the labor market. *Journal of Public Economics*, 131:41–57.

Bastani, S., Blumkin, T., and Micheletto, L. (2019). The Welfare-Enhancing Role of Parental Leave Mandates. *The Journal of Law, Economics, and Organization*, 35(1):77–126.

Blumkin, T. and Sadka, E. (2008). A case for taxing education. *International Tax and Public Finance*, 15:145–163.

Boadway, R. and Marchand, M. (1995). The use of public expenditures for redistributive purposes. *Oxford Economic Papers*, 47(1):45–59.

Bovenberg, A. L. and Jacobs, B. (2005). Redistribution and education subsidies are siamese twins. *Journal of Public Economics*, 89(11):2005–2035.

Brett, C. and Weymark, J. A. (2003). Financing education using optimal redistributive taxation. *Journal of Public Economics*, 87(11):2549–2569. ISPE conference on Public economics and redistribution.

Craig, A. C. (2020). Optimal income taxation with spillovers from employer learning. *Working Paper, University of Michigan*.

Findeisen, S. and Sachs, D. (2016). Education and optimal dynamic taxation: The role of income-contingent student loans. *Journal of Public Economics*, 138:1–21.

Fudenberg, D. and Tirole, J. (1991). Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory*, 53(2):236 – 260.

Keen, M. and Slemrod, J. (2017). Optimal tax administration. *Journal of Public Economics*, 152:133 – 142.

Kleven, H. J. and Kopczuk, W. (2011). Transfer program complexity and the take-up of social benefits. *American Economic Journal: Economic Policy*, 3(1):54–90.

Maldonado, D. (2008). Education policies and optimal taxation. *International Tax and Public Finance*, 15:131–143.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4):629–649.

Sadka, E. (1976). On Income Distribution, Incentive Effects and Optimal Income Taxation. *The Review of Economic Studies*, 43(2):261–267.

Sheshinski, E. (1971). On the theory of optimal income taxation. Discussion paper no. 172, Harvard Institute for Economic Research, Harvard University.

Slemrod, J. and Kopczuk, W. (2002). The Optimal Elasticity of Taxable Income. *Journal of Public Economics*, 84(1):91–112.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Stantcheva, S. (2014). Optimal Income Taxation with Adverse Selection in the Labor Market. *Review of Economic Studies*, 81:1296–1329.

Stiglitz, J. E. (1982). Self-selection and Pareto-efficient taxation. *Journal of Public Economics*, 17:213–240.

Tuomala, M. (1986). On the optimal income taxation and educational decisions. *Journal of Public Economics*, 30(2):183–198.

Ulph, D. (1977). On the optimal distribution of income and educational expenditure. *Journal of Public Economics*, 8(3):341–356.

Zoutman, F. T. and Jacobs, B. (2016). Optimal redistribution and monitoring of labor supply. *Journal of Public Economics*, 135:15 – 31.

# A Proof of Proposition 1

**Part (i)** Suppose by negation that for some $\theta^2 > \theta^1 > 0$ the social optimum is given by a pooling allocation, which satisfies, $\hat{c} = \hat{y} = \hat{e} \cdot \sum_i \gamma^i \theta^i$. To show that pooling is suboptimal, consider the following reform starting from an initial pooling equilibrium:

$$\tilde{e}^1 = \hat{e} - \epsilon, \quad \tilde{c}^1 = \hat{c} - \epsilon k g'(\hat{e}), \tag{A1}$$

$$\tilde{e}^2 = \hat{e} + \delta, \quad \tilde{c}^2 = \hat{c} + \delta g'(\hat{e}), \tag{A2}$$

where $\epsilon > 0$ and $\delta > 0$. By construction, the reform is welfare neutral for both types of workers since we have that $dU^1 = dc - (kg')\,de = -\epsilon kg'(\widehat{e}) + \epsilon kg'(\widehat{e}) = 0$ and $dU^2 = dc - (g')\,de = \delta g'(\widehat{e}) - \delta g'(\widehat{e}) = 0$ Notice also that, after the reform is implemented, no agent is tempted to behave as a mimicker. By behaving as a mimicker, the utility of a type-1 agent would change by $\delta g'(\widehat{e}) - \delta kg'(\widehat{e}) < 0$ (since $k > 1$). Similarly, by behaving as a mimicker, the utility of a type-2 agent would change by $-\epsilon kg'(\widehat{e}) + \epsilon g'(\widehat{e}) < 0$ (since $k > 1$). Let's now consider the effect of the proposed reform on the net revenue collected by the government. We know that net revenue is 0 in the initial equilibrium. Net revenue after the reform is equal to the difference between the variation in aggregate output and the variation in aggregate consumption:

$$dY - dC = \underbrace{\left[\gamma^2\delta\theta^2 - \gamma^1\epsilon\theta^1\right]}_{dY} - \underbrace{\left[\gamma^2\delta - \gamma^1\epsilon k\right]g'(\widehat{e})}_{dC} = \gamma^2\delta\left[\theta^2 - g'(\widehat{e})\right] - \gamma^1\epsilon\left[\theta^1 - kg'(\widehat{e})\right]. \quad \text{(A3)}$$

Notice that $dY - dC$ will be necessarily positive when at the initial pooling equilibrium we have $\theta^2 - g'(\widehat{e}) > 0$ and $\theta^1 - kg'(\widehat{e}) < 0$. Notice also that it is not possible that at the initial pooling equilibrium we have at the same time that $\theta^2 - g'(\widehat{e}) < 0$ and $\theta^1 - kg'(\widehat{e}) > 0$ (this is due to the fact that $\theta^2 > \theta^1$ and $k > 1$). Suppose then that at the initial pooling equilibrium we have that $\theta^2 - g'(\widehat{e}) > 0$ and $\theta^1 - kg'(\widehat{e}) > 0$.

Thus, net revenue increases whenever $\delta$ is set in accordance with the following condition:

$$\delta > \epsilon\frac{\gamma^1}{\gamma^2}\frac{\theta^1 - kg'(\widehat{e})}{\theta^2 - g'(\widehat{e})}. \quad \text{(A4)}$$

In other words, there exists a reform to the presumed optimal allocation that yields a fiscal surplus. Since this surplus can be rebated to the workers in a lump-sum fashion without violating the incentive compatibility constraints (due to the quasi-linearity of preferences), a Pareto improvement can be obtained. This gives us the desired contradiction.

**Part (ii)** In order to prove the second part it suffices to show that the welfare level associated with a separating equilibrium would be higher under the extended tax regime than under the pure income tax regime. This follows from the fact that the pooling equilibrium yields the same welfare level under the two tax regimes. Thus, in case the optimal solution under a pure income tax regime calls for implementing a pooling equilibrium, it is clearly dominated by the optimal solution under the extended tax regime, by virtue of part (i) of the proposition.

In the optimal separating equilibrium under an income tax regime, the incentive compatibility constraint associated with type-2 workers [given by condition (14)] is binding. Otherwise, by continuity considerations one could transfer units of consumption from type-2 to type-1 workers and enhance the latter's level of utility, without violating the incentive compatibility constraint associated with type-2 workers (and clearly without violating the revenue constraint).

Under an extended tax regime, the incentive constraint in (14) is replaced by the weaker constraint given by (24). Thus, the optimal separating equilibrium under the pure income tax regime satisfies the reformulated incentive constraint [given in (24)] as a strict inequality.[31] This slack allows the government to enhance the utility of type-1 workers. In particular, the government can slightly increase the income tax levied on type-2 workers and correspondingly increase the transfer given to type-1 workers, maintaining the budget balanced, and without violating type-2 workers' incentive constraint (by continuity considerations). The welfare level attained under the optimal separating allocation with a tax on the signal in place is therefore strictly higher than the corresponding allocation under the income tax regime. This establishes part (ii).

**Part (iii)** Consider the government's problem presented in the beginning of Section 2.3. Notice that constraints (23) and (24) cannot be binding at the same time. To see this, assume by contradiction that both constraints are binding. From (23) we can then establish that

$$c^1 = c^2 + k\left[g(e^1) - g(e^2)\right]. \tag{A5}$$

Substituting for $c^1$ in (24) the right hand side of the equation above gives:

$$c^2 - g(e^2) = c^2 + k\left[g(e^1) - g(e^2)\right] - g(e^1). \tag{A6}$$

Simplifying terms one would then obtain

$$(1-k)\left[g(e^1) - g(e^2)\right] = 0, \tag{A7}$$

a result that cannot hold, given that $k \neq 1$, under a separating equilibrium (where $e^1 \neq e^2$).

Having established that constraints (23) and (24) cannot be binding at the same time, and given the fact that the social welfare function maximized by the government is of the max-min type, we will hereafter assume that constraint (23) is slack.

Denote by $\delta^i$, for $i = 1, 2$, the Lagrange multiplier associated with the constraint (22), by $\lambda$ the Lagramge multiplier associated with the constraint (24), and by $\mu$ the Lagrange multiplier associated with the constraint (25). The first order conditions of the

[31]The only (knife-edge) cases in which no slack exists are when either $\theta^1 = 0$, or, $\theta^1 = \theta^2$.

government's problem (with respect to, respectively, $y^2$, $e^2$, $c^2$, $y^1$, $e^1$, $c^1$) are given by:

$$\delta^2 = \mu\gamma^2 \tag{A8}$$

$$\delta^2\theta^2 - \lambda g'\left(e^2\right) = 0 \tag{A9}$$

$$\lambda = \mu\gamma^2 \tag{A10}$$

$$\delta^1 = \mu\gamma^1 \tag{A11}$$

$$\delta^1\theta^1 - kg'\left(e^1\right) + \lambda g'\left(e^1\right) = 0 \tag{A12}$$

$$1-\lambda = \mu\gamma^1 \tag{A13}$$

Using (A10) to substitute $\mu\gamma^2$ for $\lambda$ in (A13) gives the result that $\mu = 1$. From (A8) we then obtain that $\delta^2 = \gamma^2$, and from (A10) we obtain that $\lambda = \gamma^2$. Thus, one can rewrite (A9) as $\gamma^2\left[\theta^2 - g'\left(e^2\right)\right] = 0$, which implies that $1 = g'\left(e^2\right)/\theta^2$, i.e. no distortion is imposed on type-2 agents.

Exploiting the fact that $\mu = 1$ and therefore, from (A11), $\delta^1 = \gamma^1$, we can rewrite (A12) as

$$kg'\left(e^1\right) = \lambda g'\left(e^1\right) + \gamma^1\theta^1. \tag{A14}$$

Since (A13) can be equivalently rewritten (taking into account that $\mu = 1$) as $1 = \lambda + \gamma^1$, combining (A14) with (A13) allows obtaining that

$$kg'\left(e^1\right)\left[\lambda + \gamma^1\right] = \lambda g'\left(e^1\right) + \gamma^1\theta^1, \tag{A15}$$

from which one gets

$$1 - \frac{kg'\left(e^1\right)}{\theta^1} = \frac{\lambda}{\gamma^1}\frac{g'\left(e^1\right)}{\theta^1}(k-1), \tag{A16}$$

or, equivalently (since $\lambda = \gamma^2$):

$$1 - \frac{kg'\left(e^1\right)}{\theta^1} = \frac{\gamma^2}{\gamma^1}\frac{g'\left(e^1\right)}{\theta^1}(k-1). \tag{A17}$$

The right hand side of (A17) provides a measure of the distortion imposed on type-1 agents for self-selection considerations.[32] Given that the right hand side of (A17) is strictly positive, the effort provided by type-1 agents is downward distorted. Notice also that, since from (A14) we also obtain (taking into account that $\lambda = \gamma^2$) that $g'\left(e^1\right)/\theta^1 = \gamma^1/\left(k-\gamma^2\right)$, we can equivalently restate (A17) as:

$$1 - \frac{kg'\left(e^1\right)}{\theta^1} = \frac{\gamma^2}{k-\gamma^2}(k-1). \tag{A18}$$

As we can see from (A18) the distortion imposed on type-1 agents only depends on the

[32]Notice that, when type-1 agents are remunerated according to their true productivity $\theta^1$, an undistorted choice for $e^1$ would fulfil the condition $1 = kg'\left(e^1\right)/\theta^1$.

differences in the costs of acquiring the signal (and on the relative proportion of the two groups of agents). The fact that the difference in productivities does not affect the magnitude of the distortion shows that an extended tax regime allows to fully eliminate any information rent related to this dimension of heterogeneity.[33]

# B Optimal distortions in the two-dimensional case

Consider the problem solved by a government under an extended tax regime and denote $\mu$ the Lagrange multiplier attached to the constraint (49), by $\lambda^2$ the multiplier attached to the constraint (50) and by $\lambda^1$ the multiplier attached to the constraint (51). The first order conditions for $y^1$, $e_s^1$, $c^1$, $y^2$, $e_s^2$, $c^2$ are respectively given by:

$$\left(1+\lambda^1\right)\frac{\partial R^1\left(e_s^1,e_q^{1*}\right)}{\partial e_q^{1*}\left(y^1,e_s^1,\theta^1\right)}\frac{\partial e_q^{1*}}{\partial y^1}=\lambda^2\frac{\partial R^2\left(e_s^1,\hat{e}_q^2\right)}{\partial \hat{e}_q^2\left(y^1,e_s^1,\overline{\theta}\right)}\frac{\partial \hat{e}_q^2}{\partial y^1}+\mu\gamma^1, \tag{B1}$$

$$\begin{aligned}&\left(1+\lambda^1\right)\left[\frac{\partial R^1\left(e_s^1,e_q^{1*}\right)}{\partial e_s^1}+\frac{\partial R^1\left(e_s^1,e_q^{1*}\right)}{\partial e_q^{1*}\left(y^1,e_s^1,\theta^1\right)}\frac{\partial e_q^{1*}}{\partial e_s^1}\right]\\ = \;&\lambda^2\left[\frac{\partial R^2\left(e_s^1,\hat{e}_q^2\right)}{\partial e_s^1}+\frac{\partial R^2\left(e_s^1,\hat{e}_q^2\right)}{\partial \hat{e}_q^2\left(y^1,e_s^1,\overline{\theta}\right)}\frac{\partial \hat{e}_q^2}{\partial e_s^1}\right],\end{aligned} \tag{B2}$$

$$1+\lambda^1=\lambda^2+\mu\gamma^1, \tag{B3}$$

$$\lambda^2\frac{\partial R^2\left(e_s^2,e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2,e_s^2,\theta^2\right)}\frac{\partial e_q^{2*}}{\partial y^2}=\lambda^1\frac{\partial R^1\left(e_s^2,e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2,e_s^2,\theta^2\right)}\frac{\partial e_q^{2*}}{\partial y^2}+\mu\gamma^2, \tag{B4}$$

$$\begin{aligned}&\lambda^2\left[\frac{\partial R^2\left(e_s^2,e_q^{2*}\right)}{\partial e_s^2}+\frac{\partial R^2\left(e_s^2,e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2,e_s^2,\theta^2\right)}\frac{\partial e_q^{2*}}{\partial e_s^2}\right]\\ = \;&\lambda^1\left[\frac{\partial R^1\left(e_s^2,e_q^{2*}\right)}{\partial e_s^2}+\frac{\partial R^1\left(e_s^2,e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2,e_s^2,\theta^2\right)}\frac{\partial e_q^{2*}}{\partial e_s^2}\right],\end{aligned} \tag{B5}$$

$$\lambda^2=\lambda^1+\mu\gamma^2. \tag{B6}$$

[33]One can easily establish that, for a setting with $J>2$ types, the generalized version of (A18) would be given by:

$$1-\frac{k^i g'\left(e^i\right)}{\theta^i}=\frac{\left(k^i-k^{i+1}\right)\sum_{j\geq i+1}\gamma^j}{\left[\left(k^i-k^{i+1}\right)\sum_{j\geq i+1}\gamma^j\right]+\gamma^i k^i},$$

where $\gamma^i$ denotes the proportion of agents of type $i$ and $k^i$ denotes the cost of acquiring the signal for agents of type $i$. The formula above applies for $i\in\{1,2,...,J-1\}$. For $i=J$ the no distortion at the top result applies, i.e. $1-k^J g'\left(e^J\right)/\theta^J=0$.

Dividing (B2) by (B3), and multiplying both sides by the RHS of (B3) gives:

$$\begin{aligned}
&\left[\frac{\partial R^1\left(e_s^1, e_q^{1*}\right)}{\partial e_s^1}+\frac{\partial R^1\left(e_s^1, e_q^{1*}\right)}{\partial e_q^{1*}\left(y^1, e_s^1, \theta^1\right)} \frac{\partial e_q^{1*}}{\partial e_s^1}\right]\left(\lambda^2+\mu \gamma^1\right) \\
= \;& \lambda^2\left[\frac{\partial R^2\left(e_s^1, \hat{e}_q^2\right)}{\partial e_s^1}+\frac{\partial R^2\left(e_s^1, \hat{e}_q^2\right)}{\partial \hat{e}_q^2\left(y^1, e_s^1, \bar{\theta}\right)} \frac{\partial \hat{e}_q^2}{\partial e_s^1}\right],
\end{aligned} \tag{B7}$$

which can be equivalently rewritten as

$$\left[p_s+p_q^1 \frac{\partial e_q^{1*}}{\partial e_s^1}\right]\left(\lambda^2+\mu \gamma^1\right)=\lambda^2\left[p_s+p_q^2 \frac{\partial \hat{e}_q^2}{\partial e_s^1}\right],$$

from which one obtains

$$p_s+p_q^1 \frac{\partial e_q^{1*}}{\partial e_s^1}=\frac{\lambda^2}{\mu \gamma^1}\left[\left(p_s+p_q^2 \frac{\partial \hat{e}_q^2}{\partial e_s^1}\right)-\left(p_s+p_q^1 \frac{\partial e_q^{1*}}{\partial e_s^1}\right)\right], \tag{B8}$$

and therefore, simplifying terms,

$$p_s+p_q^1 \frac{\partial e_q^{1*}}{\partial e_s^1}=\frac{\lambda^2}{\mu \gamma^1}\left[p_q^2 \frac{\partial \hat{e}_q^2}{\partial e_s^1}-p_q^1 \frac{\partial e_q^{1*}}{\partial e_s^1}\right]. \tag{B9}$$

Notice that we have

$$\frac{\partial e_q^{1*}}{\partial e_s^1}=-\frac{\partial h\left(e_s^1, e_q^{1*}\right) / \partial e_s}{\partial h\left(e_s^1, e_q^{1*}\right) / \partial e_q} \quad \text{and} \quad \frac{\partial \hat{e}_q^2}{\partial e_s^1}=-\frac{\partial h\left(e_s^1, \hat{e}_q^2\right) / \partial e_s}{\partial h\left(e_s^1, \hat{e}_q^2\right) / \partial e_q}. \tag{B10}$$

Moreover, since $e_q^{1*}=e_q^{1*}\left(y^1, e_s^1, \theta^1\right)$ and $\hat{e}_q^2=\left(y^1, e_s^1, \bar{\theta}\right)$, it follows that $e_q^{1*}>\hat{e}_q^2>0$ and therefore $\frac{\partial e_q^{1*}}{\partial e_s^1}<\frac{\partial \hat{e}_q^2}{\partial e_s^1}<0$. Thus, the RHS of (B9) is positive (given that our max-min objective implies that $\lambda^2>0$). Since $p_s+p_q^1 \frac{\partial e_q^{1*}}{\partial e_s^1}=p_s-p_q^1 \frac{\partial h\left(e_s^1, e_q^{1*}\right) / \partial e_s}{\partial h\left(e_s^1, e_q^{1*}\right) / \partial e_q}$, we can conclude that at an optimal separating equilibrium

$$p_s-p_q^1 \frac{\partial h\left(e_s^1, e_q^{1*}\right) / \partial e_s}{\partial h\left(e_s^1, e_q^{1*}\right) / \partial e_q}>0. \tag{B11}$$

Dividing (B5) by (B6), and multiplying both sides by the RHS of (B6) gives:

$$\begin{aligned}
&\left[\frac{\partial R^2\left(e_s^2, e_q^{2*}\right)}{\partial e_s^2}+\frac{\partial R^2\left(e_s^2, e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2, e_s^2, \theta^2\right)} \frac{\partial e_q^{2*}}{\partial e_s^2}\right]\left(\lambda^1+\mu \gamma^2\right) \\
= \;& \lambda^1\left[\frac{\partial R^1\left(e_s^2, e_q^{2*}\right)}{\partial e_s^2}+\frac{\partial R^1\left(e_s^2, e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2, e_s^2, \theta^2\right)} \frac{\partial e_q^{2*}}{\partial e_s^2}\right],
\end{aligned} \tag{B12}$$

which can be equivalently rewritten as

$$\left[p_s + p_q^2 \frac{\partial e_q^{2*}}{\partial e_s^2}\right]\left(\lambda^1 + \mu\gamma^2\right) = \lambda^1 \left[p_s + p_q^1 \frac{\partial e_q^{2*}}{\partial e_s^2}\right], \tag{B13}$$

from which one obtains

$$p_s + p_q^2 \frac{\partial e_q^{2*}}{\partial e_s^2} = \frac{\lambda^1}{\mu\gamma^2}\left[\left(p_s + p_q^1 \frac{\partial e_q^{2*}}{\partial e_s^2}\right) - \left(p_s + p_q^2 \frac{\partial e_q^{2*}}{\partial e_s^2}\right)\right], \tag{B14}$$

and therefore, simplifying terms:

$$p_s + p_q^2 \frac{\partial e_q^{2*}}{\partial e_s^2} = \frac{\lambda^1}{\mu\gamma^2}\frac{\partial e_q^{2*}}{\partial e_s^2}\left(p_q^1 - p_q^2\right). \tag{B15}$$

Since we have that $p_q^1 - p_q^2 > 0$ and

$$\frac{\partial e_q^{2*}}{\partial e_s^2} = -\frac{\partial h\left(e_s^2, e_q^{2*}\right)/\partial e_s}{\partial h\left(e_s^2, e_q^{2*}\right)/\partial e_q} < 0, \tag{B16}$$

it follows that the RHS of (B15) is either negative (when $\lambda^1 > 0$) or zero (when $\lambda^1 = 0$). Since $p_s + p_q^2 \frac{\partial e_q^{2*}}{\partial e_s^2} = p_s - p_q^2 \frac{\partial h\left(e_s^2, e_q^{2*}\right)/\partial e_s}{\partial h\left(e_s^2, e_q^{2*}\right)/\partial e_q}$, we can conclude that at an optimal separating equilibrium

$$p_s - p_q^2 \frac{\partial h\left(e_s^2, e_q^{2*}\right)/\partial e_s}{\partial h\left(e_s^2, e_q^{2*}\right)/\partial e_q} \leq 0. \tag{B17}$$

Dividing (B1) by (B3), and multiplying both sides by the RHS of (B3) gives:

$$\frac{\partial R^1\left(e_s^1, e_q^{1*}\right)}{\partial e_q^{1*}\left(y^1, e_s^1, \theta^1\right)}\frac{\partial e_q^{1*}}{\partial y^1}\left(\lambda^2 + \mu\gamma^1\right) = \lambda^2 \frac{\partial R^2\left(e_s^1, \hat{e}_q^2\right)}{\partial \hat{e}_q^2\left(y^1, e_s^1, \overline{\theta}\right)}\frac{\partial \hat{e}_q^2}{\partial y^1} + \mu\gamma^1, \tag{B18}$$

which can be equivalently rewritten as

$$p_q^1 \frac{\partial e_q^{1*}}{\partial y^1}\left(\lambda^2 + \mu\gamma^1\right) = \lambda^2 p_q^2 \frac{\partial \hat{e}_q^2}{\partial y^1} + \mu\gamma^1, \tag{B19}$$

from which one obtains

$$1 - p_q^1 \frac{\partial e_q^{1*}}{\partial y^1} = \frac{\lambda^2}{\mu\gamma^1}\left[p_q^1 \frac{\partial e_q^{1*}}{\partial y^1} - p_q^2 \frac{\partial \hat{e}_q^2}{\partial y^1}\right]. \tag{B20}$$

Since we have that

$$\frac{\partial e_q^{1*}}{\partial y^1} = \frac{1}{\theta^1 \partial h\left(e_s^1, e_q^{1*}\right)/\partial e_q} \quad \text{and} \quad \frac{\partial \hat{e}_q^2}{\partial y^1} = \frac{1}{\overline{\theta}\partial h\left(e_s^1, \hat{e}_q^2\right)/\partial e_q}, \tag{B21}$$

it follows that $\frac{\partial e_q^{1*}}{\partial y^1} > \frac{\partial \hat{e}_q^2}{\partial y^1}$ (taking into account that $\theta^1 < \overline{\theta}$ and $e_q^{1*} > \hat{e}_q^2 > 0$). Thus, the RHS of (B20) is strictly positive (given that our max-min objective implies that $\lambda^2 > 0$). Since $1 - p_q^1 \frac{\partial e_q^{1*}}{\partial y^1} = 1 - \frac{p_q^1}{\theta^1 \partial h(e_s^1, e_q^{1*})/\partial e_q}$, it follows that at an optimal separating equilibrium

$$1 - \frac{p_q^1}{\theta^1 \partial h\left(e_s^1, e_q^{1*}\right)/\partial e_q} > 0. \tag{B22}$$

Dividing (B4) by (B6), and multiplying both sides by the RHS of (B6) gives:

$$\frac{\partial R^2\left(e_s^2, e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2, e_s^2, \theta^2\right)} \frac{\partial e_q^{2*}}{\partial y^2}\left(\lambda^1 + \mu\gamma^2\right) = \lambda^1 \frac{\partial R^1\left(e_s^2, e_q^{2*}\right)}{\partial e_q^{2*}\left(y^2, e_s^2, \theta^2\right)} \frac{\partial e_q^{2*}}{\partial y^2} + \mu\gamma^2, \tag{B23}$$

which can be equivalently rewritten as

$$p_q^2 \frac{\partial e_q^{2*}}{\partial y^2}\left(\lambda^1 + \mu\gamma^2\right) = \lambda^1 p_q^1 \frac{\partial e_q^{2*}}{\partial y^2} + \mu\gamma^2, \tag{B24}$$

from which one obtains

$$1 - p_q^2 \frac{\partial e_q^{2*}}{\partial y^2} = \frac{\lambda^1}{\mu\gamma^2} \frac{\partial e_q^{2*}}{\partial y^2}\left(p_q^2 - p_q^1\right). \tag{B25}$$

Given that $p_q^2 - p_q^1 < 0$, it follows that the RHS of (B25) is either negative (when $\lambda^1 > 0$) or zero (when $\lambda^1 = 0$). Since $1 - p_q^2 \frac{\partial e_q^{2*}}{\partial y^2} = 1 - \frac{p_q^2}{\theta^2 \partial h(e_s^2, e_q^{2*})/\partial e_q}$, it follows that at an optimal separating equilibrium

$$1 - \frac{p_q^2}{\theta^2 \partial h\left(e_s^2, e_q^{2*}\right)/\partial e_q} \leq 0. \tag{B26}$$

# C Derivation of optimal deviating strategies

In this appendix, under the assumption provided by (60), we derive closed form solutions for the cost functions that appear on the right hand side of the incentive-compatibility constraints (45)–(46) in section 3.

## C.1 The two deviating strategies for type 2 agents

**Type 2 agents earn $y^1$ but separate themselves from type 1 agents** Consider problem (40) for $i = 2$ and $j = 1$. If one could neglect constraint (41), it is obvious that the cost-minimizing effort mix for a type-2 mimicker would be given by:

$$e_s^{2*} = \sqrt{\left(\frac{y^1}{A\theta^2}\right)^{1/\beta} \frac{p_q^2}{p_s}}, \tag{C1}$$

$$e_q^{2*} = \sqrt{\left(\frac{y^1}{A\theta^2}\right)^{1/\beta} \frac{p_s}{p_q^2}}. \tag{C2}$$

To assess whether (C1)-(C2) represent a valid characterization of a mimicker's behavior, we need to check this (undistorted) effort mix allows type 2 agents to separate themselves from their lower skilled counterparts. In other words, we need to check whether the incentive constraint for type 1 agents, $p_s e_s^{2*} + p_q^1 e_q^{2*} \geq R^1\left(y^1\right)$, is satisfied or not. Given that $R^1\left(y^1\right) = 2\sqrt{\left(\frac{y^1}{A\theta^1}\right)^{1/\beta} p_q^1 p_s}$ (see (71) for $i = 1$), it can be easily shown that $p_s e_s^{2*} + p_q^1 e_q^{2*} \geq R^1\left(y^1\right)$ when

$$\frac{\theta^2}{\theta^1} \leq \Omega, \tag{C3}$$

where $\Omega$ was defined in (72). Thus, if condition (C3) is satisfied, the minimal cost incurred by a type-2 mimicker is given by:

$$\widetilde{R}^2\left(y^1\right) = p_s e_s^{2*} + p_q^2 e_q^{2*} = 2\sqrt{\left(\frac{y^1}{A\theta^2}\right)^{1/\beta} p_q^2 p_s}. \tag{C4}$$

If instead condition (C3) is violated, a type 2 mimicker will need to adopt a distorted effort mix, and therefore incur a higher cost, in order to achieve separation. In particular, when condition (C3) is violated, the effort mix chosen by a type-2 mimicker is obtained by solving the following system:

$$\begin{aligned}\left(e_s^2 e_q^2\right)^{\beta} A\theta^2 &= y^1, & \text{(C5)}\\ R^1(y^1) &= p_s e_s^2 + p_q^1 e_q^2. & \text{(C6)}\end{aligned}$$

The first equation is the binding output constraint. The second equation restricts attention to signals that cause the incentive constraint of type 1 to bind. From (C5) we obtain $e_s^2 = \left(\frac{y^1}{A\theta^2}\right)^{1/\beta} \frac{1}{e_q^2}$ which, substituted in (C6 ) gives (after some algebraic manipulation):

$$e_s^{2*} = \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}\left(\frac{1}{\theta^2}\right)^{1/\beta} p_q^1}}{\sqrt{p_s p_q^1}\left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}\right]}, \tag{C7}$$

$$e_q^{2*} = \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta} p_s p_q^1}\left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}\right]}{p_q^1}, \tag{C8}$$

implying

$$\widetilde{R}^2(y_1) = p_s e_s^{2*} + p_q^2 e_q^{2*} = p_s \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}\left(\frac{1}{\theta^2}\right)^{1/\beta} p_q^1}}{\sqrt{p_s p_q^1}\left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}\right]}$$

$$+ p_q^2 \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}}\sqrt{p_s p_q^1}\left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\theta^2}\right)^{1/\beta}}\right]}{p_q^1}. \tag{C9}$$

Thus, $\widetilde{R}^2(y_1)$ on the RHS of the incentive-compatibility constraint (45) is given by (C4) if (C3) is satisfied and it is given by (C9) if (C3) is violated.

**Type 2 agents earn $y^1$ without achieving separation from type 1 agents** Consider problem (37) for $i = 2$ and $j = 1$. If one could neglect constraint (38), it is obvious that the cost-minimizing effort mix for a type-2 mimicker would be given by:

$$e_s^{2*} = \sqrt{\left(\frac{y^1}{\overline{\theta}A}\right)^{1/\beta}\frac{p_q^2}{p_s}}, \tag{C10}$$

$$e_q^{2*} = \sqrt{\left(\frac{y^1}{\overline{\theta}A}\right)^{1/\beta}\frac{p_s}{p_q^2}}. \tag{C11}$$

To assess whether (C10)-(C11) represent a valid characterization of a mimicker's behavior, we need to whether this effort mix is also attractive to type 1 agents. Put differently, we need to check whether it is the case that

$$p_s e_s^{2*} + p_q^1 e_q^{2*} \leq R^1\left(y^1\right), \tag{C12}$$

where $R^1\left(y^1\right)$ is defined in (71). Through some simple algebraic manipulations, one can show that (C12) can be rewritten as

$$\gamma^1 + \gamma^2\frac{\theta^2}{\theta^1} \geq \Omega. \tag{C13}$$

Thus, if condition (C13) is satisfied, the minimal cost incurred by a type-2 mimicker is given by:

$$\widehat{R}^2\left(y^1\right) = 2\sqrt{\left(\frac{y^1}{\overline{\theta}A}\right)^{1/\beta} p_s p_q^2}. \tag{C14}$$

If instead condition (C13) is violated, a type 2 mimicker will need to choose a different effort mix, and therefore incur a higher cost, in order to induce also type-1 agents to choose the same effort vector. In particular, if condition (C13) is violated, the effort mix chosen by a type-2 mimicker is obtained by solving the following system:

$$\left(e_s^2 e_q^2\right)^\beta \overline{\theta} A = y^1, \tag{C15}$$

$$p_s e_s^2 + p_q^1 e_q^2 = R^1(y^1). \tag{C16}$$

From (C15) we obtain $e_s^2 = \left(\frac{y^1}{\overline{\theta}A}\right)^{1/\beta} \frac{1}{e_q^2}$ which, substituted in (C16) allows us to conclude (details are straightforward but tedious, hence omitted):

$$e_s^{2*} = \left(\frac{1}{\overline{\theta}}\right)^{1/\beta} \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}}\sqrt{\frac{p_q^1}{p_s}}}{\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\overline{\theta}}\right)^{1/\beta}}}, \tag{C17}$$

$$e_q^{2*} = \left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\overline{\theta}}\right)^{1/\beta}}\right] \sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}} \sqrt{\frac{p_s}{p_q^1}}, \tag{C18}$$

implying

$$\begin{aligned} \widehat{R}^2\left(y^1\right) &= p_s e_s^{2*} + p_q^2 e_q^{2*} = \left(\frac{1}{\overline{\theta}}\right)^{1/\beta} \frac{\sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}}\sqrt{\frac{p_q^1}{p_s}}p_s}{\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\overline{\theta}}\right)^{1/\beta}}} \\ &+ \left[\sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{1/\beta} - \left(\frac{1}{\overline{\theta}}\right)^{1/\beta}}\right] \sqrt{\left(\frac{y^1}{A}\right)^{1/\beta}} \sqrt{\frac{p_s}{p_q^1}} p_q^2. \end{aligned} \tag{C19}$$

**Comparing the two deviating strategies for type 2 agents** Based on the above discussion, we conclude that the value of $\widetilde{R}^2\left(y^1\right)$ depends on whether $\frac{\theta^2}{\theta^1} \le \Omega$ is satisfied or not, and the value of $\hat{R}^2\left(y^1\right)$ depends on whether $\gamma^1 + \gamma^2 \frac{\theta^2}{\theta^1} \ge \Omega$ is satisfied or not.

Consider first the case when (C3) is satisfied. In this case, we have that (C13) is necessarily violated. Thus, the relevant comparison is between (C19) and (C4), and it can be verified that $\widetilde{R}^2\left(y^1\right) < \widehat{R}^2\left(y^1\right)$. Intuitively, this happens since: i) under the deviating strategy associated with the cost $\widetilde{R}^2\left(y^1\right)$, the effort mix chosen by the mimicker is undistorted (i.e., it satisfies the condition $e_q/e_s = p_s/p_q^2$) and the mimicker is remunerated according to his/her true productivity $\theta^2$; ii) under the deviating strategy associated with the cost $\widehat{R}^2\left(y^1\right)$, the effort mix chosen by the mimicker is distorted (i.e., it does not satisfy the condition $e_q/e_s = p_s/p_q^2$) and the mimicker is paid according to the average productivity $\overline{\theta}$.

Consider now the case when (C3) is violated. In this case, condition (C13) can either be satisfied or violated. Therefore, two possibilities need to be considered.

1. If $\gamma^1 + \gamma^2 \frac{\theta^2}{\theta^1} < \Omega < \frac{\theta^2}{\theta^1}$, in which case (C13) is violated, then the relevant comparison is between the cost $\tilde{R}^2(y^1)$ given by (C9) and the cost $\hat{R}^2(y^1)$ given by (C19). In this

case, since $\overline{\theta} < \theta^2$ and $\frac{p_s}{p_q^2} > \frac{p_s}{p_q^1}$ it follows that $\tilde{R}^2(y^1) < \hat{R}^2(y^1)$.

2. If $\Omega \leq \gamma^1 + \gamma^2 \frac{\theta^2}{\theta^1} < \frac{\theta^2}{\theta^1}$, in which case (C13) is satisfied, then the relevant comparison is between $\widetilde{R}^2(y^1)$ given by (C9) and $\hat{R}^2(y^1)$ given by (C14) . In this case, one cannot unambiguously rank $\widetilde{R}^2\left(y^1\right)$ and $\widehat{R}^2\left(y^1\right)$. Intuitively, this happens since: i) under the deviating strategy associated with the cost $\widetilde{R}^2\left(y^1\right)$, the effort mix chosen by the mimicker is distorted, but the mimicker is paid according to his/her true productivity $\theta^2$; ii) under the deviating strategy associated with the cost $\widehat{R}^2\left(y^1\right)$, the effort mix chosen by the mimicker is undistorted, but the mimicker is paid according to the average productivity $\overline{\theta}$. Thus, in this case, there are two downwards incentive constraints that need to be taken into account.

## C.2 The two deviating strategies for type 1 agents

**Type 1 agents earn $y^2$ and choose the same effort mix as type 2 agents** Under this deviating strategy, a type 1 mimicker earn $y^2$ by selecting the effort mix $(e_s^2(y^2), e_q^2(y^2))$ chosen in equilibrium by type 2 agents. Thus, denoting by $(\breve{e}_s^1\left(y^2\right), \breve{e}_q^1\left(y^2\right))$ the effort mix selected by a type 1 mimicker, we have that:

$$\breve{e}_s^1\left(y^2\right) = e_s^2\left(y^2\right) = \sqrt{\left(\frac{y^2}{A\theta^2}\right)^{1/\beta} \frac{p_q^2}{p_s}}, \tag{C20}$$

$$\breve{e}_q^1\left(y^2\right) = e_q^2\left(y^2\right) = \sqrt{\left(\frac{y^2}{A\theta^2}\right)^{1/\beta} \frac{p_s}{p_q^2}}, \tag{C21}$$

and the cost sustained by a type 1 mimicker is given by:

$$\breve{R}^1\left(y^2\right) = p_s \breve{e}_s^1\left(y^2\right) + p_q^1 \breve{e}_q^1\left(y^2\right) = \sqrt{\left(\frac{y^2}{A\theta^2}\right)^{1/\beta}} \frac{p_q^2 + p_q^1}{\sqrt{p_q^2}} \sqrt{p_s}. \tag{C22}$$

Notice that (C20)-(C21) represent distorted effort choices for a type 1 mimicker, since they violate the condition $e_q/e_s = p_s/p_q^1$ (instead, they fulfil the condition $e_q/e_s = p_s/p_q^2$).

**Type 1 agents earn $y^2$ but separate themselves from type 2 agents** Consider now problem (40) for $i = 1$ and $j = 2$. In this case, a type 1 mimicker chooses the undistorted effort mix:

$$e_s^{1*}\left(y^2\right) = \sqrt{\left(\frac{y^2}{A\theta^1}\right)^{1/\beta} \frac{p_q^1}{p_s}}, \tag{C23}$$

$$e_q^{1*}\left(y^2\right) = \sqrt{\left(\frac{y^2}{A\theta^1}\right)^{1/\beta} \frac{p_s}{p_q^1}}. \tag{C24}$$

Therefore, the cost sustained by a type 1 mimicker is given by:

$$\widetilde{R}^1\left(y^2\right) = p_s e_s^{1*}\left(y^2\right) + p_q^1 e_q^{1*}\left(y^2\right) = 2\sqrt{\left(\frac{y^2}{A\theta^1}\right)^{1/\beta}}\sqrt{p_q^1 p_s}. \tag{C25}$$

**Comparing the two deviating strategies for type 1 agents** We have that $\widetilde{R}^1\left(y^2\right) > \breve{R}^1\left(y^2\right)$ when the following condition holds:

$$2\sqrt{\left(\frac{y^2}{A\theta^1}\right)^{1/\beta}}\sqrt{p_q^1 p_s} > \sqrt{\left(\frac{y^2}{A\theta^2}\right)^{1/\beta}}\frac{p_q^2 + p_q^1}{\sqrt{p_q^2}}\sqrt{p_s}. \tag{C26}$$

Notice that the condition above can be rewritten as

$$\frac{\theta^2}{\theta^1} > \left[\frac{1}{2}\frac{p_q^2 + p_q^1}{\sqrt{p_q^2 p_q^1}}\right]^{2\beta} = \Omega, \tag{C27}$$

where (C27) is the reverse of inequality (C3).[34]

Thus, we can conclude that if $\frac{\theta^2}{\theta^1} > \Omega$, the preferred (cost-minimizing) deviating strategy for a type 1 mimicker is to replicate the signals of a type 2 agent, thereby incurring the cost $\breve{R}^1(y^2)$. If, on the other hand, $\frac{\theta^2}{\theta^1} \leq \Omega$, a type 1 mimicker will choose an effort mix that enables him/her to produce $y^2$, while separating himself/herself from type 2 agents, incurring the cost $\widetilde{R}^1\left(y^2\right)$.

[34]Thus, we may notice that when a type 2 mimicker has to choose the distorted effort mix given by eqs. (C7)-(C8) to achieve separation, the preferred deviating strategy for a type 1 agent is to replicate the effort mix chosen by a type 2 non-mimicker (i.e. to choose $e_s$ and $e_q$ according to (C20)-(C21)). On the other hand, when a type 2 mimicker does not need to choose a distorted effort mix to achieve separation (i.e. his/her effort mix is given by eqs. (C1)-( C2)), the preferred deviating strategy for a type 1 agent is to separate himself/herself and choose $e_s$ and $e_q$ according to (C23)-(C24).