Optimal Redistribution and Education Signaling*

By Spencer Bastani, Tomer Blumkin, and Luca Micheletto

Institute for Evaluation of Labour Market and Education Policy (IFAU) and Uppsala University, Sweden; Ben Gurion University, Israel; University of Milan, Italy.

Abstract

We develop a theory of optimal income and education taxation under asymmetric information between firms and workers. Our results show that a max-min optimal tax code can achieve predistribution by pooling wages across ability levels, conditional on income. We identify conditions under which the optimal solution leads to pooling or separating equilibria, highlighting bidirectional incentive constraints. Implementation requires nonlinear income taxes coupled with education subsidies or mandates. Predistribution is only feasible when income taxes are complemented by policies that restrict signaling opportunities. Our framework provides new insights into reducing wage inequality through optimal tax policy and labor market information management.

Keywords: nonlinear taxation, education, asymmetric information, human capital, predistribution, wage inequality

JEL classification: D82, H21, H52, J31

Short title: (for running head): Optimal Redistribution and Signaling

*Manuscript received August 2022; revised February 2024. We are grateful to the editor, Dirk Krueger, and three anonymous reviewers for very helpful comments. We are also grateful to Håkan Selin and Anna Sjögren for providing valuable comments on an earlier draft of the paper. Spencer Bastani gratefully acknowledges financial support from Jan Wallanders och Tom Hedelius stiftelse (grant P2018-162). Please address correspondence to: Spencer Bastani, Institute for Evaluation of Labour Market and Education Policy (IFAU) and Department of Economics, Uppsala University; Research Institute of Industrial Economics (IFN); UCFS, UCLS, CESifo. E-mail: spencer.bastani@ifau.uu.se.

1 Introduction

In the canonical framework of optimal income taxation originally developed by Mirrlees (1971), the primary challenge facing tax policy stems from the presence of asymmetric information between the government and private individuals. The government's goal is to redistribute resources based on the innate productive abilities of individuals. However, since these abilities remain unobservable for tax purposes, the government resorts to taxing income and other observable measures that can serve as proxies for these unobserved abilities. This leads to the introduction of second-best solutions, where incentive compatibility considerations justify the introduction of distortions, often in the form of positive marginal tax rates. These distortions facilitate targeted transfers to low-income individuals while providing incentives for high-income individuals to exert labor effort.

The prevailing optimal tax literature has largely overlooked a crucial aspect of tax policy design: in addition to the standard information asymmetry between the government and private agents emphasized in traditional optimal tax theory, there is a second layer of information asymmetry between workers and employers. As economists have recognized since the seminal contributions of Spence (1973) and Akerlof (1976), asymmetric information in the labor market profoundly shapes the dynamics of interactions between workers and firms and can contribute significantly to market inefficiencies. This asymmetry implies that employers cannot accurately gauge the productivity of workers, and as a result, even in a competitive labor market, workers may not receive compensation commensurate with their marginal productivity. Instead, the wage distribution becomes endogenous, influenced by the screening and signaling methods available to employers and workers alike.

Two recent papers extend Mirrlees' framework by introducing a second layer of asymmetric information between workers and employers, focusing on how firms screen workers based on their choices about working hours. Stantcheva (2014) explores the implications of adverse selection in the labor market for the optimal design of income taxes, showing that firms' use of hours and compensation as screening tools can help governments achieve redistributive goals by counteracting the adverse responses of high-skilled workers to progressive taxation. Bastani et al. (2015), using a similar screening framework, examines how progressive income tax schedules can affect wage distribution by promoting bunching or pooling across worker types.

This paper develops a framework for evaluating optimal redistributive policies in the presence of multidimensional educational signaling, where workers signal their productivity through both the quantity and quality of their education. Our contributions are fourfold: (i) we provide a theory of optimal redistribution that addresses the complexities of multidimensional educational signaling; (ii) we show that a max-min optimal tax code can achieve *predistribution* by pooling the wages of workers with different skill levels conditional on income; (iii) we derive sufficient conditions under which the max-min optimum leads to either pooling or separating equilibria, highlighting that in a separating equilibrium incentive constraints between two types can bind in both directions simultaneously, an aspect that has been underexplored in the literature; and (iv) we explore the policy instruments necessary to implement these results, focusing on a nonlinear income tax together with a piecewise linear education subsidy schedule. A key insight is that achieving predistribution requires complementing the income tax with policies that limit signaling opportunities and prevent high-skilled individuals from fully separating from their low-skilled counterparts.

While the current paper shares the feature of a second layer of asymmetric information with the two studies mentioned above, it differs from them in at least five ways. First, we focus on worker signaling through investment in education. Despite its central role in economics, its prominence in economics curricula around the world, and its relevance in policy discussions (e.g., Caplan 2018), it is surprising that signaling has been addressed in only a few papers in the vast literature on optimal tax design since the seminal work of Mirrlees (1971).[1] Second, we consider multidimensional signaling in the context of

[1] Two early papers discussing signaling in the context of taxation are Spence (1974) and Manoli (2006). More recently, Craig (2023) studies signaling in the context of human capital investment and the design of optimal income taxation in a different setting where employers make Bayesian inferences about workers' productivity and the equilibrium wage is a weighted average of the worker's own productivity and the productivity of other similar workers. Sztutman (2024) studies optimal taxation in a dynamic job

taxation, which allows us to retain the realistic Mirrleesian feature that firms are more informed about workers than the government. Third, we consider tax systems that depend not only on income but also on the signals that the government can observe in the labor market.[2] Fourth, in line with Bastani et al. (2015) but in contrast to Stantcheva (2014), we emphasize the important role of redistribution through the wage (as opposed to the income) channel.[3] Finally, in contrast to Stantcheva (2014), we show that the presence of adverse selection due to asymmetric information between firms and workers does not necessarily lead to a higher level of welfare in the social optimum than that achieved in a Mirrleesian setup where worker types are observable by firms.

The details of our analysis are as follows. Consistent with the prevailing literature on optimal income taxation, we assume that workers differ in their intrinsic productive capabilities, which are unobservable to the government. However, unlike most studies in this area, and in line with the two studies discussed above, we extend this unobservability to potential employers. The distinguishing feature of our analysis is that workers must signal their productivity to firms by making costly effort decisions, allowing for information transmission between workers and firms along two dimensions: the quantity (e.g., years of schooling) and the quality (e.g., the difficulty or intensity of a particular educational pathway) of their education. While quantity is observable to both the government and employers, quality is only observable to employers, reflecting an environment in which employers have better information than the government. To make signaling feasible, we assume that workers differ not only in their innate productive abilities but also in their costs of signaling (e.g., the cost of obtaining education), with signals representing

signaling model where the career profile of labor supply conveys information about worker productivity.

[2]The taxation of signals has received surprisingly little attention in the optimal income tax literature. The only previous paper that we are aware of that explicitly discusses the taxation of signals is Andersson (1996).

[3]Notably, predistribution can occur even when production technology is linear and skill types are perfect substitutes, as in Mirrlees (1971). This differs from models where redistribution through the wage channel arises from sectoral reallocation of labor in general equilibrium contexts (see, for example, Stiglitz 1982, Rothschild and Scheuer 2013, and Sachs et al. 2020).

components of educational effort that realistically also increase human capital.[4]

Adopting a framework that captures the equity-efficiency tradeoff, similar to the two-type Stiglitz (1982) version of the Mirrlees (1971) optimal income tax model, we analyze constrained efficient (max-min) allocations that combine taxes on both income and observable signaling activity. By invoking the revelation principle, we solve for the optimal direct revelation mechanism and characterize feasible and incentive-compatible allocations. In most of our analysis, we assume that the signal observable to the government is the one in which the low type has a comparative advantage. However, we also discuss what happens in situations where neither signal is observable, where both signals are observable, and where the signal in which the high type has a comparative advantage is observable. We also briefly discuss some extensions of our analysis, such as the cases of more than two signals and more than two types.

We begin by defining the Perfect Bayesian Equilibria (PBE) of the signaling game in the presence of a general tax function, including laissez-faire as a special case. The PBE consists of strategies for workers (educational choices) and employers (wage offers), along with employers' beliefs about workers' productivity, which are updated in a Bayesian-consistent manner based on observed signals. We then apply equilibrium refinement along the lines of Grossman and Perry (1986) and characterize the labor market equilibrium in the presence of taxes, recognizing that it can be given by either a *separating tax equilibrium* (STE), where workers earn different levels of income and exert different levels of educational effort, or a *pooling tax equilibrium* (PTE), where all workers earn the same income and exert the same observable level of effort. We recognize that the richness of the tax function plays a key role in supporting the existence of equilibrium and in determining whether a predistributive PTE is achievable.

We then characterize the constrained efficient allocation assuming a max-min social objective, called the max-min optimum (MMO), and show that it is given by either an STE or a PTE, depending on which equilibrium configuration produces the highest level

[4]Our model is thus related to the literature on optimal income taxation in the presence of human capital investment and learning-by-doing, see, for example, Stantcheva (2017).

of social welfare. We derive necessary and sufficient conditions for the MMO to feature predistribution, emphasizing the role of both differences in agents' innate productivities and differences in the costs of signaling. Note that in our setting, incentive constraints can flow from low- to high types as well as from high- to low types. Low types may have an incentive to invest more in signaling in order to qualify for higher compensation, while high types may have an incentive to mimic low types in order to qualify for a more lenient tax treatment.

Our study highlights a key policy insight: when workers signal their productivity through their educational choices, the government can use a complementary wage channel for redistribution, namely predistribution. Crucially, achieving predistribution requires augmenting the income tax system with additional policy instruments that directly regulate the flow of information between workers and firms and prevent high-skilled individuals from separating themselves from their low-skilled counterparts. Our formal analysis, detailed in Online Appendix H, shows that, in our setting, predistribution cannot be achieved by an income tax system in isolation.

The policy framework required to implement the MMO (whether provided by an STE or a PTE) can take several forms. We propose two simple implementation schemes that combine a nonlinear income tax with income-tested education subsidies or mandates. A nonlinear income tax system—in practice often piecewise linear with multiple brackets—encourages individuals to locate at targeted income levels. Income-tested subsidies and mandates, on the other hand, ensure that higher-ability individuals are not incentivized to deviate from their lower-ability counterparts by choosing lower levels of education (i.e., lower quantity effort) conditional on income level. The design of income-tested education subsidies and mandates is driven by the need to provide the right incentives locally—at a given income level—without distorting the incentives to acquire education at other income levels.[5]

[5]Education subsidies have traditionally been used to correct market failures and redistribute income. In the optimal tax literature, they serve two primary functions: i) to mitigate the negative effects of income taxation on human capital formation, and ii) to enhance redistribution. See, for example, Ulph (1977), Tuomala (1986), Boadway and Marchand (1995), Brett and Weymark (2003), Bovenberg and

If the MMO is implemented as an STE, workers earn different income levels and exert different levels of educational effort. In this scenario, type-2 mimickers are pooled with low-skilled workers off the equilibrium path. A means-tested subsidy on the observable dimension of educational effort, in which low-skilled workers have a comparative advantage, serves to discourage high-skilled mimickers from separating themselves from their low-skilled counterparts at the lower income level, while avoiding distorting effort choices at the higher income level. This logic parallels models of optimal mixed taxation (combining income and goods taxes) where low-skilled and high-skilled workers have different consumption preferences (see, for example, Blomquist and Christiansen 2008). Using income taxes to finance subsidies for goods favored by low-skilled workers can achieve redistribution at a lower efficiency cost than income taxes alone, because it allows distinguishing between truly low-skilled and high-skilled workers, conditional on income.

If the MMO is implemented as a PTE, all workers earn the same income and exert the same observable level of effort, aligning their choices along the equilibrium path. Here, implementation requires a kink in the income tax schedule. Since there is no redistribution through income taxation in a PTE, the role of the tax schedule is to support the pooling equilibrium and thereby help achieve predistribution. The role of means-tested education subsidies in this context is to discourage high-skilled workers from differentiating themselves from low-skilled counterparts by opting for lower levels of educational effort off the equilibrium path. This is achieved by subsidizing effort levels below the common equilibrium effort, which effectively imposes a marginal tax on downward deviations. We propose the simplest way to deter such deviations by high-skilled mimickers, namely, to complement the nonlinear income tax system with a binding education mandate that sets a lower bound on educational effort.

While it is well known that kinks in the income tax schedule can bunch individuals with

Jacobs (2005), and Maldonado (2008). Some studies, such as Blumkin and Sadka (2008), also explore the possibility of education taxes due to the positive correlation between education and unobserved ability. More recently, Findeisen and Sachs (2016) examines income-contingent student loans and suggests that it may be optimal for very high-income individuals to repay more than the value of their loans, effectively creating an education tax.

different labor productivity at the same pre-tax income, resulting in identical after-tax incomes (which can sometimes serve redistributive purposes, see, e.g., Ebert 1992), our study emphasizes that combining these kinks with education mandates can also induce bunching at the education choice. This, in turn, induces wage bunching conditional on income and achieves redistribution through wage compression. Although a pooling equilibrium compresses all income levels into a single outcome, the broader insight extends to more complex scenarios involving multiple types and equilibria with partial bunching or full separation. At each income level along the equilibrium path, income-tested education subsidies or mandates can be used to enforce bunching both on and off the equilibrium path.

The paper is organized as follows. In section 2, we outline the structure of the game, the equilibrium concept that we use, and the role of government in the economy. We then define the STE and PTE in the presence of a general tax function, and describe the government optimization problem and the concept of MMO. In section 3, we characterize the optimal wedges associated with the MMO. In section 4, we discuss how these wedges can be implemented using means-tested education subsidies or mandates. In section 5, we discuss alternative observational assumptions and some robustness and extensions of the basic setup. Section 6 concludes. Most of our formal derivations and proofs are relegated to the Online Appendix.

2 The model

Consider an economy with a competitive labor market consisting of two types of workers: low-skilled, denoted by $i = 1$, and high-skilled, denoted by $i = 2$, who differ in their innate ability. Let $0 < \gamma^i < 1$ denote the proportion of workers of type i in the population (normalized to a unit measure, without loss of generality).

We build on the basic insights of the Mirrlees (1971) framework, which examines how a planner designs a nonlinear tax schedule $T(y)$ based on observed income y. A widely accepted interpretation of the Mirrlees model is that income directly equals output, justified by the assumption of a competitive labor market in which firms perfectly observe workers' productivity and compensate them accordingly. Our paper departs from this standard interpretation by relaxing the assumption that worker productivity is perfectly observed and compensated by firms. More broadly, we challenge the equivalence between income and worker output, motivated by scenarios where firms cannot directly observe or contract with workers based on their actual output. The central innovation of our approach is the introduction and analysis of two layers of asymmetric information: one between the government and private agents, and another between workers and firms.[6]

Workers exert costly effort that serves the dual purpose of (i) increasing worker productivity and (ii) signaling innate ability. Our model is general, but for concreteness we focus on educational attainment, which is interpreted as educational effort prior to entering the labor market. In line with this interpretation, workers are first movers in the interaction with firms.

We consider educational attainment along two dimensions. The first is denoted by e_s and represents the quantity of effort. The second dimension is denoted by e_q and represents the intensity of effort. For example, in the context of education, the variables e_s and e_q would capture the quantity (e.g., time spent acquiring vocational training and/or academic degrees) and quality (e.g., GPA, reputation of certifying institution, interviews, and letters of recommendation) dimensions of educational attainment, respectively. Our main focus will be on the case where e_s is observed by both the government and the firms, while e_q is only observed by the firms (or is prohibitively costly for the government to observe). However, in subsections 5.1-5.3 we will also briefly discuss the implications of other observability assumptions.

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

$$z^i = h(e_s^i, e_q^i)\theta^i, \tag{1}$$

where $h(\cdot)$ is jointly strictly concave and strictly increasing in both arguments and represents the acquired human capital; and θ^i denotes the innate productive ability of

[6]Other papers exploring two layers of asymmetric information in optimal policy design include Stantcheva (2014), Bastani et al. (2015, 2019), Craig (2023), and Sztutman (2024).

type i, where $\theta^2 > \theta^1$. In addition, we define $\bar{\theta} = \gamma^1\theta^1 + \gamma^2\theta^2$ as the average productivity of workers. We further assume that the Inada conditions are satisfied, i.e., $\lim_{e_s \to 0^+} \frac{\partial h}{\partial e_s} = \lim_{e_q \to 0^+} \frac{\partial h}{\partial e_q} = \infty$ and $\lim_{e_s \to \infty} \frac{\partial h}{\partial e_s} = \lim_{e_q \to \infty} \frac{\partial h}{\partial e_q} = 0$. We define the *wage rate* earned by a given individual as the ratio of pre-tax income, denoted by y, and the value of the h function evaluated at the effort vector chosen by the individual. We will also denote by h_1 and h_2 the first derivative with respect to the first and second arguments of h, respectively. The utility function is

$$u^i(c, e_s.e_q) = c - R^i(e_s, e_q), \tag{2}$$

where c is consumption and

$$R^i(e_s, e_q) = p_s^i e_s + p_q^i e_q, \tag{3}$$

is the cost function for agents of type i, where p_s^i and p_q^i denote the unitary marginal cost of e_s and the unitary marginal cost of e_q, respectively, for an agent of type i. The linear cost specification is used for tractability, and the qualitative features of our results could be obtained under more general specifications. We henceforth make the following assumptions:

$$p_s^1 = p_s^2 \equiv p_s \quad \text{and} \quad p_q^1 > p_q^2, \tag{4}$$

which together imply that type-2 agents have a (weak) absolute advantage in signaling through each channel, and a comparative advantage in the quality signal e_q.

Note that without being overly unrealistic, and in order to simplify the exposition and make the setup more tractable, we assume that labor supply is inelastic and normalized to a unit of time. We discuss the case of endogenous labor supply in subsection 5.4 below, where we argue that endogenous labor supply can be viewed as a special case of adding another signal.

2.1 Labor market equilibrium with taxes

We analyze a two-stage signaling game involving workers and firms. In the first stage, workers choose their effort levels along two dimensions: quality and quantity, denoted as (e_s^i, e_q^i), for types $i = 1, 2$. These effort levels act as signals of their productivity, which firms then observe. In the second stage, firms make wage offers based on these observed signals. Wage offers reflect firms' beliefs about workers' productivity, which are formed based on the signals received.

Perfect Bayesian Equilibrium As is standard in the literature, we focus on Perfect Bayesian Equilibria (PBE) of the signaling game, restricting our analysis to pure strategies.[7] Firms hold beliefs $\mu(e_s, e_q) \in [0, 1]$ about the probability that a worker has high productivity (θ^2), based on the observed signals (e_s, e_q). These beliefs are updated according to Bayes' rule. Firms make wage offers simultaneously based on their beliefs, and these offers reflect the worker's expected productivity, similar to Bertrand competition. We denote the wage offer function as $\Theta(e_s, e_q) = \mu(e_s, e_q)\theta^2 + (1 - \mu(e_s, e_q))\theta^1$. Along the equilibrium path, firms maximize expected profits by setting a wage policy based on their beliefs, while workers maximize their utility by choosing effort levels (e_s, e_q) in response to these wage offers and the relevant tax schedule. A pure strategy PBE under the general tax function $T(y, e_s, e_q)$ can be represented as a set of *equilibrium allocations* $(y^{i*}, e_s^{i*}, e_q^{i*})$ for $i = 1, 2$, where:

$$\left(y^{1*}, e_s^{1*}, e_q^{1*}\right) = \underset{y^1, e_s^1, e_q^1}{\operatorname{argmax}} \left\{y^1 - T\left(y^1, e_s^1, e_q^1\right) - p_s e_s^1 - p_q^1 e_q^1\right\} \quad \text{subject to}$$

$$y^1 \leq \Theta(e_s^1, e_q^1) \cdot h(e_s^1, e_q^1). \tag{5}$$

$$\left(y^{2*}, e_s^{2*}, e_q^{2*}\right) = \underset{y^2, e_s^2, e_q^2}{\operatorname{argmax}} \left\{y^2 - T\left(y^2, e_s^2, e_q^2\right) - p_s e_s^2 - p_q^2 e_q^2\right\} \quad \text{subject to}$$

$$y^2 \leq \Theta(e_s^2, e_q^2) \cdot h(e_s^2, e_q^2), \tag{6}$$

[7] For a formal treatment of PBE, see Fudenberg and Tirole (1991).

and the government's revenue constraint holds:

$$\sum_{i=1,2} \gamma^i \cdot T\left(y^{i*}, e_s^{i*}, e_q^{i*}\right) \geq 0. \tag{7}$$

The wage offer function $\Theta(e_s, e_q)$ is updated as firms learn from the observed signals provided by workers. The equilibrium wage function, $\Theta^*(e_s, e_q)$, represents the wage that emerges when firms' beliefs are consistent with the observed equilibrium behavior of workers. An equilibrium is a stable point, in the sense that beliefs are correct given strategies, and strategies are sequentially rational given beliefs.

In defining PBE, we assume that firms earn non-negative profits, rather than imposing a zero-profit condition, as shown in equations (5) and (6). While the zero-profit condition typically holds in competitive labor markets—due to competition among firms and the properties of the human capital production function, $h(e_s, e_q)$—this may not always be the case when considering a general tax function $T(y, e_s, e_q)$.[8] In addition, we focus on equilibria where the government's budget constraint (7) is satisfied. Throughout the analysis, we assume that the government cannot run a deficit. Since our primary interest is in taxation as a redistributive tool, we assume—without loss of generality —that the government has no revenue needs.

As we focus on a PBE with pure strategies, the equilibrium can be either separating or pooling. In a separating equilibrium, workers with different productivity types choose different effort levels (e_s^{i*}, e_q^{i*}), allowing firms to perfectly infer each worker's productivity: $\Theta(e_s^{i*}, e_q^{i*}) = \theta^i$ for $i = 1, 2$. Accordingly, firms have equilibrium beliefs $\mu^*(e_s^{2*}, e_q^{2*}) = 1$ for high productivity workers and $\mu^*(e_s^{1}, e_q^{1*}) = 0$ for low productivity workers. In a pooling equilibrium, all workers choose identical effort levels, so firms cannot distinguish between high- and low-productivity types based on observed effort. Instead, firms form beliefs about productivity based on the average population distribution, i.e. $\mu^* = \gamma^2$, where γ^2 is the proportion of high productivity workers. Consequently, the wage offered is based on average productivity: $\Theta(e_s^{i*}, e_q^{i*}) = \bar{\theta} = \gamma^2\theta^2 + \gamma^1\theta^1$.

[8]For example, $\partial T/\partial y$ could exceed 100% in certain income ranges.

Beliefs off the equilibrium path As defined earlier, firms form beliefs about a worker's productivity, denoted by $\mu(e_s, e_q) \in [0, 1]$, which represent the probability that a worker has high productivity (θ^2). Along the equilibrium path, these beliefs are updated using Bayes' rule to ensure that firms' wage offers reflect the expected productivity based on the observed signals. However, situations arise when a worker chooses an unexpected effort level —one that deviates from the equilibrium path. In such cases, firms must form off-equilibrium path beliefs to interpret these unexpected signals.

In our analysis, we adopt the extended intuitive criterion, an equilibrium refinement introduced by Grossman and Perry (1986), to restrict the possible beliefs that firms can hold in response to unexpected worker actions. This criterion allows firms to distinguish between credible and non-credible deviations, refining the set of equilibria to those consistent with plausible behavior. Unlike the standard intuitive criterion of Cho and Kreps (1987), which considers only unilateral deviations, the extended version is more flexible, allowing for deviations by subsets of types. Specifically, this refinement states that when a deviation occurs, firms should update their beliefs assuming that it was made by a subset of worker types for whom the deviation is most profitable, provided that such a deviation is credible. This method is more restrictive than considering strictly unilateral deviations, and thus refines the possible equilibrium outcomes.

These modeling choices are consistent with Riley (2001), who shows that under this refinement, a pooling equilibrium cannot be maintained in a no-tax, laissez-faire regime, and that a separating equilibrium exists only if the fraction of low-skilled workers is sufficiently large. In our context, however, these results change depending on the design of the tax function. As we discuss in more detail below, depending on what is observable and thus taxable, a separating equilibrium may always exist, and pooling equilibria may also become sustainable.

The equilibria derived from the application of the extended intuitive criterion are called refined PBEs. Specifically, we distinguish two types of refined PBEs under different tax regimes: the separating tax equilibrium (STE) and the pooling tax equilibrium (PTE). Our concept of tax equilibria accommodates a wide range of potential tax systems, from laissez-faire with no taxation to a fully flexible tax framework that taxes both income and the two educational signals.

We begin by characterizing the STE.

Lemma 1 (Separating Tax Equilibrium, STE)**.** *Suppose that, given the general tax function $T(y, e_s, e_q)$, the allocations*

$$\left(y^{1*}, e_s^{1*}, e_q^{1*}\right) \quad \textit{and} \quad \left(y^{2*}, e_s^{2*}, e_q^{2*}\right),$$

with $(e_s^{1*}, e_q^{1*}) \neq (e_s^{2*}, e_q^{2*})$, are the strategies played in a pure strategy PBE. Furthermore, let $\mu(e_s, e_q) : \mathbb{R}^2 \to [0, 1]$ be the belief function, where $\mu(e_s, e_q)$ represents the probability that a worker has high productivity (θ^2) given the observed signals (e_s, e_q). The belief system is such that on the equilibrium path, beliefs are updated using Bayes' rule such that $\mu(e_s^{1*}, e_q^{1*}) = 0$ and $\mu(e_s^{2*}, e_q^{2*}) = 1$. Off the equilibrium path, beliefs are refined using the extended intuitive criterion (Grossman and Perry, 1986).

The equilibrium allocations satisfy the following conditions:

(a) Optimality for type 1 workers

$$\left(y^{1*}, e_s^{1*}, e_q^{1*}\right) = \underset{y^1, e_s^1, e_q^1}{\operatorname{argmax}} \left\{ y^1 - T(y^1, e_s^1, e_q^1) - p_s e_s^1 - p_q^1 e_q^1 \right\},$$

subject to:

(8) $\quad y^{1*} \leq \theta^1 h(e_s^1, e_q^1),$

(9) $\quad y^{2*} - T(y^{2*}, e_s^{2*}, e_q^{2*}) - p_s e_s^{2*} - p_q^2 e_q^{2*} \geq y^1 - T(y^1, e_s^1, e_q^1) - p_s e_s^1 - p_q^2 e_q^1.$

(b) Optimality for type 2 workers

$$\left(y^{2*}, e_s^{2*}, e_q^{2*}\right) = \underset{y^2, e_s^2, e_q^2}{\operatorname{argmax}} \left\{ y^2 - T(y^2, e_s^2, e_q^2) - p_s e_s^2 - p_q^2 e_q^2 \right\},$$

subject to:

$$(10) \qquad y^{2*} \leq \theta^2 h(e_s^2, e_q^2),$$

$$(11) \qquad y^{1*} - T(y^{1*}, e_s^{1*}, e_q^{1*}) - p_s e_s^{1*} - p_q^1 e_q^{1*} \geq y^2 - T(y^2, e_s^2, e_q^2) - p_s e_s^2 - p_q^1 e_q^2.$$

(c) No profitable deviations off the equilibrium path

No allocation $(y, e_s, e_q) \in \mathbb{R}_+^3$ *satisfies:*

$$(12) \qquad y \leq \bar{\theta} h(e_s, e_q),$$

$$(13) \qquad y - T(y, e_s, e_q) - p_s e_s - p_q^1 e_q > y^{1*} - T(y^{1*}, e_s^{1*}, e_q^{1*}) - p_s e_s^{1*} - p_q^1 e_q^{1*},$$

$$(14) \qquad y - T(y, e_s, e_q) - p_s e_s - p_q^2 e_q > y^{2*} - T(y^{2*}, e_s^{2*}, e_q^{2*}) - p_s e_s^{2*} - p_q^2 e_q^{2*}.$$

Proof. A separating tax equilibrium is a PBE that is immune to strictly profitable deviations both on and off the equilibrium path, the latter evaluated by the extended intuitive criterion. On the equilibrium path, it must be the case that neither type can strictly profit from deviating to an allocation that separates them from the other type while still allowing the firm to make non-negative profits. Condition (a) captures that the choices made by *low-skilled workers* in equilibrium maximize their utility subject to the condition that the firm makes non-negative profits and that the incentive compatibility constraint associated with a high-skilled worker who might mimic their behavior is satisfied. Similarly, Condition (b) ensures that the equilibrium choices of *high-skilled workers* maximize their utility, subject to the condition that the firm must still make non-negative profits and the incentive compatibility constraint associated with the potential mimicking of a low-skilled worker. Finally, Condition (c) guarantee that neither type can strictly gain by jointly deviating to a pooling allocation off the equilibrium path, while the firm remains profitable. □

Lemma 1 characterizes an STE allocation when it exists. The three conditions in Lemma 1 ensure that workers cannot profitably deviate along the equilibrium path by mimicking the choices of their counterparts, nor can they deviate off the equilibrium path by choosing different levels of effort, either by separating (a unilateral deviation) or by pooling (a joint deviation).

We turn next to characterize the pooling refined PBE.

Lemma 2 (Pooling Tax Equilibrium, PTE)**.** *Suppose that, given the general tax function $T(y, e_s, e_q)$, the singleton allocation*

$$\left(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*\right),$$

forms a pure-strategy refined pooling PBE. Furthermore, let $\mu(e_s, e_q) : \mathbb{R}^2 \to [0,1]$ be the belief function, where $\mu(e_s, e_q)$ represents the probability that a worker has high productivity (θ^2) given observed signals (e_s, e_q). The belief system is such that on the equilibrium path $\mu(e_s^*, e_q^*) = \gamma^2$, the prior probability (share) of high-skill workers. Off the equilibrium path beliefs are refined using the extended intuitive criterion (Grossman and Perry, 1986).

The equilibrium allocation satisfies the following conditions:

(a) No profitable deviations for type 1 workers

There is no $(y^1, e_s^1, e_q^1) \in \mathbb{R}^3_+$ satisfying $y^1 \leq \theta^1 h(e_s^1, e_q^1)$ such that:

$$y^1 - T(y^1, e_s^1, e_q^1) - p_s e_s^1 - p_q^1 e_q^1 > \widehat{y}^* - T(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*) - p_s \widehat{e}_s^* - p_q^1 \widehat{e}_q^*, \tag{15}$$

$$\widehat{y}^* - T(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*) - p_s \widehat{e}_s^* - p_q^2 \widehat{e}_q^* \geq y^1 - T(y^1, e_s^1, e_q^1) - p_s e_s^1 - p_q^2 e_q^1. \tag{16}$$

(b) No profitable deviations for type 2 workers

There is no $(y^2, e_s^2, e_q^2) \in \mathbb{R}^3_+$ satisfying $y^2 \leq \theta^2 h(e_s^2, e_q^2)$ such that:

$$\widehat{y}^* - T(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*) - p_s \widehat{e}_s^* - p_q^1 \widehat{e}_q^* \geq y^2 - T(y^2, e_s^2, e_q^2) - p_s e_s^2 - p_q^1 e_q^2, \tag{17}$$

$$y^2 - T(y^2, e_s^2, e_q^2) - p_s e_s^2 - p_q^2 e_q^2 > \widehat{y}^* - T(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*) - p_s \widehat{e}_s^* - p_q^2 \widehat{e}_q^*. \tag{18}$$

(c) No joint deviations to a new pooling allocation

There is no $(\widehat{y}, \widehat{e}_s, \widehat{e}_q) \in \mathbb{R}^3_+ \setminus \{\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*\}$ satisfying $\widehat{y} \leq \bar{\theta} h(\widehat{e}_s, \widehat{e}_q)$ such that, for both

$i = 1, 2,$

$$(19) \qquad \widehat{y} - T(\widehat{y}, \widehat{e}_s, \widehat{e}_q) - p_s\widehat{e}_s - p_q^i\widehat{e}_q > \widehat{y}^* - T(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*) - p_s\widehat{e}_s^* - p_q^i\widehat{e}_q^*.$$

Proof. A pooling tax equilibrium is a PBE that is immune to strictly profitable deviations both on and off the equilibrium path, the latter evaluated by the extended intuitive criterion. Since there are no possible deviations along the equilibrium path, the only possible deviation is a deviation to a separating allocation off the equilibrium path. Condition (a) prevents low-skilled workers from benefiting by deviating from an allocation that separates them from high-skilled workers, while ensuring that firms continue to earn non-negative profits. Similarly, Condition (b) prevents high-skilled workers from benefiting by deviating from a separating allocation, with the same requirement on firm profitability. Finally, Condition (c) prevents both types of workers from jointly benefiting from deviating to an alternative pooling allocation, again while ensuring that firms remain profitable. □

Lemma 2 characterizes a PTE allocation, if it exists. By conditions (a)–(c), Lemma 2 ensures that the PTE is immune to strictly profitable deviations off the equilibrium path.

Before turning to the government problem, it is important to note that the tax function plays a crucial role in supporting the existence of an equilibrium, regardless of the relative size of the two groups of agents or the magnitude of the difference $p_q^1 - p_q^2$—a notable contrast to the typical results when the refinements of Grossman and Perry (1986) are applied. However, an equilibrium does not exist for every possible tax configuration. For example, a pooling equilibrium does not exist under a laissez-faire regime (a result well established in the literature) or under an income-only tax regime (as formally demonstrated in Online Appendix H). Intuitively, to maintain a pooling equilibrium, policy instruments must be sufficiently comprehensive to prevent type 2 workers from using their comparative advantage in an effort dimension to separate themselves from less skilled counterparts. In terms of the conditions in Lemma 2, a pooling equilibrium does not exist under laissez-faire or with only an income tax in place because condition b) is necessarily violated. Moreover, according to Lemma 1, a separating equilibrium may also not exist under laissez-faire, since it is possible to satisfy conditions (12)–(14) if the proportion of low-skilled workers in the population is small enough.

2.2 The government problem

We now turn to describe the optimal tax problem solved by the government. In line with the informational assumptions described at the beginning of Section 2, we focus on a setting where the (quality) signal e_q is observed only by firms, and thus an individual's tax liability can be conditioned only on labor income y and the (quantity) signal e_s.[9] In accordance with most of the literature on optimal taxation, instead of directly optimizing the tax function $T(y, e_s)$, we will follow a mechanism design (self-selection) approach, first characterizing a constrained efficient allocation and then, in a separate section, considering the properties of the implementing tax function.

Definition 1 (Max-Min Optimum, MMO)**.** *A Max-Min Optimum (MMO) is given by a solution to:*

$$\left\{\left(c^1, e_s^1, e_q^1\right), \left(c^2, e_s^2, e_q^2\right)\right\} \in \underset{c^1, e_s^1, e_q^1, c^2, e_s^2, e_q^2}{\arg\max} \; c^1 - R^1\left(e_s^1, e_q^1\right), \tag{20}$$

subject to the government revenue constraint

$$\sum_{i=1,2} \gamma^i [y^i - c^i] = \sum_{i=1,2} \gamma^i \left[h(e_s^i, e_q^i)\Theta^i - c^i\right] = 0, \tag{21}$$

where the wage rate is given by

$$\Theta^i = \begin{cases} \theta^i, & \textit{for all } \left(e_s^1, e_q^1\right) \neq \left(e_s^2, e_q^2\right) \\ \overline{\theta}, & \textit{for all } \left(e_s^1, e_q^1\right) = \left(e_s^2, e_q^2\right), \end{cases} \tag{22}$$

[9] In Section 5 we discuss how our results would change under alternative observational assumptions.

and the incentive-compatibility (IC) constraints are

$$c^2 - R^2\left(e_s^2, e_q^2\right) \geq c^1 - R^2\left(e_s^1, \widehat{e}_q^2\right), \tag{23}$$

$$c^1 - R^1\left(e_s^1, e_q^1\right) \geq c^2 - R^1\left(e_s^2, e_q^2\right), \tag{24}$$

where

$$\widehat{e}_q^2 = \begin{cases} e_q \text{ which solves } y^1 = h\left(e_s^1, e_q\right)\overline{\theta}, & \text{for all } \left(e_s^1, e_q^1\right) \neq \left(e_s^2, e_q^2\right) \\ e_q^1, & \text{for all } \left(e_s^1, e_q^1\right) = \left(e_s^2, e_q^2\right). \end{cases} \tag{25}$$

The MMO in Definition 1 implicitly defines the tax code that induces the best pure strategy PBE and it can correspond to either an STE or a PTE. In Online Appendix A, we show that the feasible set includes both equilibrium configurations and that the maximum is well defined. We postpone the welfare comparison of the two configurations to section 2.3.[10]

In the case of an STE, each of the two groups of agents is induced to choose a type-specific pair (e_s, e_q), and workers are compensated by firms based on their true productivity. Redistribution to type-1 agents occurs through the traditional ex-post tax/transfer channel, with type-2 agents paying a tax that finances a transfer to type-1 agents. In contrast, in the case of a PTE, all agents are induced to choose the same pair (e_s, e_q), and are compensated by firms according to their average productivity $\bar{\theta}$, thereby earning the common income level $\bar{\theta}h\left(e_s, e_q\right)$. In this scenario, since we assume no exogenous revenue requirement for the government, everyone pays the same tax, which is zero. Redistribution in this case occurs not through the traditional income channel—where high-income earners pay taxes to finance transfers to low-income earners—but through the wage channel, by suppressing wage inequality.

To distinguish between these two channels of redistribution, we use the term predis-

[10]Note that by defining the MMO as the best equilibrium from the class of pure strategy equilibria, we implicitly punt on the question of whether there are tax codes that induce mixed strategy equilibria that are even better.

tribution to refer specifically to redistribution that operates through the wage channel. According to the definition of the wage rate Θ^i, $i = 1, 2$, predistribution occurs when the MMO leads to a PTE, but not when it leads to an STE. In our framework, predistribution manifests itself as wage pooling, where workers are paid according to the average productivity rather than according to their marginal productivity.[11] Moreover, in our setup, wage pooling is synonymous with income pooling. This contrasts with the standard Mirrlees framework, where income pooling does not imply wage pooling. This distinction underscores that predistribution in our model operates through wage compression conditional on income, a mechanism that is not possible in the standard Mirrlees model.

The objective (20) reflects that the social welfare function is of the max-min type, focusing on a specific point on the second-best Pareto frontier. To relax the assumption of a max-min social objective, an additional constraint could be added to the maximization problem, requiring that the utility achieved by type-2 agents is weakly greater than a pre-specified target level $\overline{V}$. By varying $\overline{V}$ and repeatedly solving the government's optimization problem, all points on the second-best Pareto frontier could be obtained.

Equation (21) represents the government's budget constraint. We assume that the zero-profit condition holds for both workers and that the government's revenue constraint is binding. Relaxing either condition would allow the government to modify the tax function and increase redistribution.[12]

Equations (23)–(24) are the two incentive compatibility (IC) constraints. Equation

[11] We recognize that there may be other, broader definitions of predistribution that involve wage compression that does not manifest itself as wage pooling. For example, a government-mandated minimum wage could reduce wage inequality by raising the wages of the lowest-paid workers, thereby compressing the wage distribution without directly pooling wages across the workforce.

[12] For example, if the revenue constraint is slack (a budget surplus), the government could offer a small lump sum transfer to both types. Continuity would ensure that the revenue constraint is not violated. Incentive compatibility would be maintained by the linearity of utility in consumption. If the firm makes positive profits, the government could slightly increase the compensation level, y, which would maintain non-negative profits due to continuity. This would create a fiscal surplus that could be refunded as a lump sum transfer.

(23) is relevant because the government seeks to redistribute from type-2 agents to their type-1 counterparts, which implies that type-2 agents may have an incentive to mimic type-1 agents in order to qualify for more favorable tax treatment (i.e., to receive a fiscal transfer instead of paying a tax). Equation (24) is relevant because, due to asymmetric information in the labor market, type-1 agents may have an incentive to mimic type-2 agents in order to receive compensation based on a productivity higher than their real one.[13] Note that under a PTE, the equations (23)–(24) are trivially satisfied.

Let us now describe the incentive constraints in more detail. Equation (24) indicates that for a type-1 agent to qualify for a higher wage, he/she must replicate both effort dimensions of type-2 agents, since the firm observes *both* education dimensions. For type-2 agents, the situation is more complex because of possible off-equilibrium deviations. In order to qualify for the low-skilled tax treatment, they must replicate the pre-tax income level and effort e_s of type-1 agents. Type-2 agents might also replicate the quality of effort e_q of type-1 agents, which would make the two types indistinguishable to the firm, leading the firm to treat both as low-skilled types. To prevent such a deviation, the social planner must pay type-2 agents an information rent, since type-2 agents can earn the same income (y^1) as the low-skilled type while incurring lower costs due to $p_q^2 < p_q^1$.

Although firms do not directly observe worker productivity, and type-2 agents cannot identify themselves as high-productivity types while mimicking type-1 agents, there is a potential off-equilibrium deviation that is even more profitable for type-2 workers than simply replicating type-1 choices. Specifically, if type-2 agents choose lower quality effort while type-1 agents do the same, the firm will be unable to distinguish between them and will pay both the average wage, rationally expecting to hire both types. For type-2 agents, this off-equilibrium deviation is particularly attractive because the level of effort required to earn y^1 when paid the average wage is lower than that required to earn y^1 when paid θ^1.

[13] In our setting, the presence of two IC constraints, often both binding, is a key feature. In the standard setting, without the second layer of asymmetric information between firms and workers, typically only the downward IC constraint (associated with a mimicking high-skill type) is binding in the optimal solution.

Two important observations about the incentive constraint (23) are worth noting. First, violating this constraint would violate the extended intuitive criterion (discussed in section 2.1) —since both types would find it strictly profitable to deviate to the pooling allocation associated with y^1. Second, the constraint (23) reflects the information rent that accrues to high ability workers due to the productivity difference between types (a type-2 agent mimicking type-1 behavior is rewarded based on average productivity, not low productivity as would be the case if both low type signals were replicated). However, this information rent is smaller than in the standard Mirrleesian framework, where a type-2 mimicker would be rewarded according to true productivity. Thus, asymmetric information between firms and workers may make it less attractive for high-skill types to mimic low-skill types, potentially increasing redistribution relative to the standard setup (see also Stantcheva 2014). However, this is not a general result because, as our analysis shows, the potentially binding upward IC constraint must also be considered relative to the standard Mirrlees model. We return to this issue in section 2.4.

2.3 When is predistribution optimal?

Let us now analyze the social optimality of both STE and PTE configurations. In an STE, the government typically cannot fully eliminate the information rents that arise from productivity differences between workers. In contrast, in a PTE, the government fully eliminates these information rents by enforcing full wage compression, although the PTE generally has less desirable efficiency properties. The equity-efficiency tradeoff between pooling and separation depends critically on the magnitude of the productivity differences, and of the heterogeneity in the costs associated with acquiring the quality signal e_q, between the two types of workers.[14] Proposition 1 below summarizes the main results.

Proposition 1 (Optimality of Predistribution)**.** *Let the ability advantage of type-2 agents be denoted by* $\epsilon = \theta^2 - \theta^1 \geq 0$*, and the cost disadvantage of type-1 agents by* $\delta = p_q^1 - p_q^2 \geq 0$*.*

[14] In a standard Mirrleesian framework with two types of agents, pooling is never optimal and is in fact Pareto-dominated by the laissez-faire allocation.

The MMO can be characterized as follows:

(a) *There is a non-empty set of parameters in the* (ϵ, δ)*-space for which the MMO is given by a PTE (and thus features predistribution).*

(b) *For any* $\epsilon > 0$*, there exists a threshold* $\delta^*(\epsilon) \geq 0$ *such that the MMO is given by an STE for* $\delta > \delta^*$ *and a PTE for* $\delta < \delta^*$*.*

(c) *There exists some cutoff* $\varepsilon^* > 0$ *such that* $\delta^*(\varepsilon) = 0$ *for any* $\varepsilon > \varepsilon^*$ *(and thus the MMO is an STE for all* δ*), while* $\delta^*(\varepsilon) > 0$ *for all* $\varepsilon < \varepsilon^*$ *(so the MMO is either an STE or a PTE, depending on the value of* δ*).*

Proof. See Online Appendix B. □

While augmenting the income tax system with taxes or subsidies on education may improve redistribution under separation —by alleviating the binding IC constraints faced by the government —Proposition 1 outlines cases where taxing or subsidizing education, by enabling the implementation of a PTE, increases social welfare *beyond* what is achievable under an STE.

Part (a) of Proposition 1 establishes the case for predistribution by identifying a non-empty set of parameters where pooling increases welfare relative to separation. Part (b) shows that pooling is socially desirable when the cost difference of obtaining the quality signal between the two types is moderate. In this scenario, type-1 workers —who typically invest more effort in the quality dimension to qualify for higher wages —are more inclined to engage in mimicking. In contrast to the standard Mirrlees model, in this case, both IC constraints are binding, and the efficiency gains from separation are limited. Part (c) shows that pooling is preferred when the productivity gap between the two types is moderate, meaning that the efficiency loss from wage compression is relatively small.

2.4 The ambiguous welfare effects of asymmetric information

A key result in Stantcheva (2014) is that adverse selection in the labor market can increase welfare by reducing the information rent that high-skilled workers can earn by mimicking low-skilled workers. In other words, adverse selection makes it more costly for high-skilled workers to underinvest in human capital and pretend to be low-skilled. This result is somewhat surprising, as it suggests that the presence of asymmetric information, which is typically viewed as a market friction that reduces welfare, can have a positive welfare effect. However, we show that this result does not necessarily hold in our setting if both types of workers have an incentive to mimic each other, depending on the relative productivity and cost of acquiring human capital.

Our analysis suggests (see the proof of Proposition 1 in Online Appendix B) that when the MMO is given by an STE, it may well be the case that both IC constraints bind in the optimal solution for the government's optimization program. This will happen when the comparative advantage of type 2 workers in the quality dimension of education is modest (δ is small) and the difference in productivity between types is significant (ε is large). The former makes mimicking by type-1 workers (who want to be paid as if they had high productivity) more attractive. The latter makes the STE superior to a PTE because of the disincentives to human capital acquisition associated with a pooling equilibrium. That welfare may be lower in such a setting than in a "Mirrleesian" setting where firms observe workers' productivity is shown formally in Proposition 2.

Proposition 2. *If δ is sufficiently small and $\varepsilon > 0$ is sufficiently large, the MMO is given by an STE and the welfare level is lower than in a scenario where firms observe the productivity of workers.*

Proof. See Online Appendix C. □

Thus, our findings highlight that the impact of asymmetric information on welfare is context-dependent and generally ambiguous: while it can improve welfare under certain conditions (as in Stantcheva 2014), it can also reduce welfare when the conditions favor both types of workers having incentives to mimic each other.

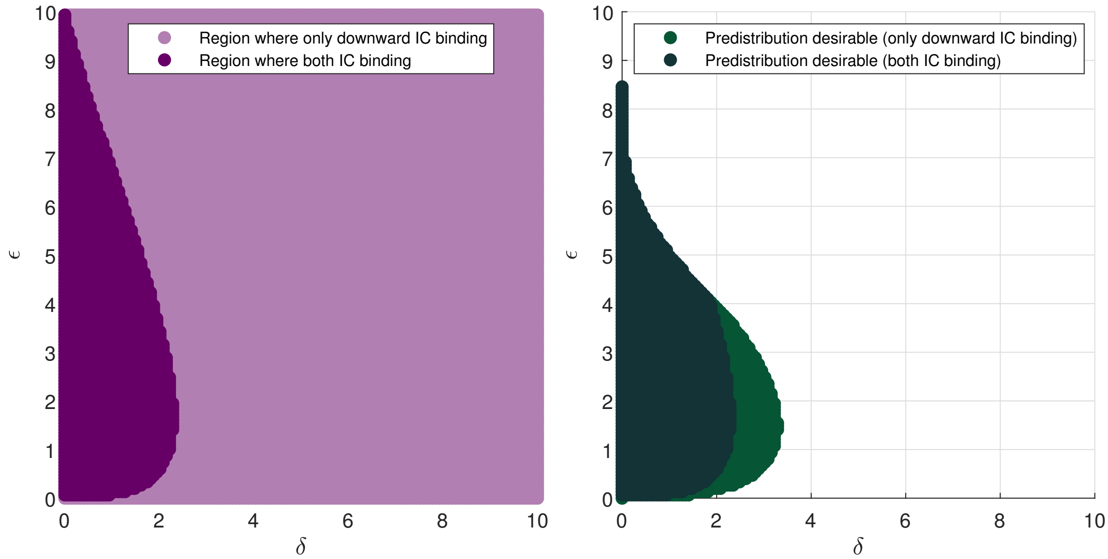
2.5 A parametric example

Given a specific functional form of the human capital production function, we can analytically identify the combinations of ϵ and δ for which predistribution is favorable. Specifically, we assume the following production function:

$$h(e_s, e_q) = (e_s e_q)^{\beta}\,, \tag{26}$$

where $0 < \beta < 1/2$, implying strict concavity. To evaluate the results, we take the following approach: for each (ϵ, δ) combination, based on Definition 1, we compute both the *optimal* STE, which maximizes welfare for type-1 agents, and the *optimal* PTE, which does the same. We then compare these results to determine which one yields higher welfare. A graphical illustration is provided to show the parameter regions in which predistribution (PTE) constitutes the social optimum, and how these regions depend on the set of binding incentive constraints in the optimal PTE. The analytical inequalities defining these regions are derived in Online Appendix D.2 and summarized in Online Appendix D.3. Figure 1 evaluates these regions using the parameters $\beta = 0.10$, $\gamma^1 = \gamma^2 = 0.5$, $p_q^1 = 10$, and $\theta^2 = 10$, where δ ranges from 0 to p_q^1 and ϵ ranges from 0 to θ^2.

Figure 1: Illustration of the case for predistribution and the pattern of binding IC constraints.



Left panel: Dark purple region is where both IC constraints are binding under the optimal STE. Light purple region is where only the downward constraint is binding under such an equilibrium. **Right panel:** Dark green region is the subregion of dark purple where the optimal PTE welfare dominates the optimal STE. Light green region is the subregion of the light purple region where the optimal PTE welfare-dominates the optimal STE.

Figure 1 shows that the region where PTE dominates STE (the dark green area in the right panel) largely overlaps with the region where both IC constraints are binding in STE (the dark purple area in the left panel). However, for moderate values of ϵ, there are also cases (indicated by the light green area in the right panel) where the PTE is welfare superior to the STE, even though only the downward IC constraint is binding in the latter. Online Appendix D.1 explores the reasoning behind the shape of these regions, distinguishing between cases where $\delta = 0$, $\delta > 0$ but small, and $\delta > 0$ and large, while also explaining the role of γ_1 and β. In addition, Online Appendix K provides further analysis based on the same functional form, quantifying the welfare gains from predistribution.

3 Wedges in the constrained efficient allocation

We turn next to the characterization of the optimal wedges, denoted by Ω and defined as the differences, at the MMO, between the marginal rates of transformation and the marginal rates of substitution among the variables entering individuals' utility functions. Proposition 3 summarizes the main results.

Proposition 3. *(a) If the MMO is a PTE $(\hat{c}, \hat{y}, \hat{e}_s, \hat{e}_q)$, then it satisfies:*

$$(27) \qquad \widehat{\Omega}^1_{e_s,e_q} \equiv \widehat{MRTS}^1 - \frac{p_s}{p^1_q} = 0 \quad \text{and} \quad \widehat{\Omega}^2_{e_s,e_q} \equiv \widehat{MRTS}^2 - \frac{p_s}{p^2_q} < 0,$$

$$(28) \qquad \widehat{\Omega}^1_{e_s,c} \equiv 1 - \frac{p_s}{\theta^1 h_1\left(\widehat{e}_s, \widehat{e}_q\right)} = \widehat{\Omega}^1_{e_q,c} = 1 - \frac{p^1_q}{\theta^1 h_2\left(\widehat{e}_s, \widehat{e}_q\right)} < 0,$$

$$(29) \qquad \widehat{\Omega}^2_{e_q,c} \equiv 1 - \frac{p^2_q}{\theta^2 h_2\left(\widehat{e}_s, \widehat{e}_q\right)} > \widehat{\Omega}^2_{e_s,c} = 1 - \frac{p_s}{\theta^2 h_1\left(\widehat{e}_s, \widehat{e}_q\right)} > 0,$$

where $\widehat{MRTS}^1 = \widehat{MRTS}^2 \equiv \frac{h_1(\widehat{e}_s,\widehat{e}_q)}{h_2(\widehat{e}_s,\widehat{e}_q)}$.

(b) If the MMO is an STE $\{(c^1, y^1, e^1_s, e^1_q), (c^2, y^2, e^2_s, e^2_q)\}$, then it satisfies:

$$(30) \qquad \Omega^1_{e_s,e_q} \equiv MRTS^1 - \frac{p_s}{p^1_q} = \frac{\lambda^2}{\gamma^1}\left(MRTS^{21}\frac{p^2_q}{p^1_q} - MRTS^1\right) < 0,$$

$$(31) \qquad \Omega^1_{e_q,c} \equiv 1 - \frac{p^1_q}{\theta^1 h_2\left(e^1_s, e^1_q\right)} = \frac{\lambda^2}{\gamma^1}\left(\frac{p^1_q}{\theta^1 h_2\left(e^1_s, e^1_q\right)} - \frac{p^2_q}{\overline{\theta} h_2\left(e^1_s, \widehat{e}^2_q\right)}\right) > 0,$$

$$(32) \qquad \Omega^1_{e_s,c} \equiv 1 - \frac{p_s}{\theta^1 h_1\left(e^1_s, e^1_q\right)} = \frac{\lambda^2}{\gamma^1}\left(\frac{h_1\left(e^1_s, \widehat{e}^2_q\right)}{h_1\left(e^1_s, e^1_q\right)}\frac{1}{\theta^1} - \frac{1}{\overline{\theta}}\right)\frac{p^2_q}{h_2\left(e^1_s, \widehat{e}^2_q\right)},$$

$$(33) \qquad \Omega^2_{e_s,e_q} \equiv MRTS^2 - \frac{p_s}{p^2_q} = \frac{\lambda^1}{\gamma^2}\left(\frac{p^1_q}{p^2_q} - 1\right) \cdot MRTS^2 \geq 0,$$

$$(34) \qquad \Omega^2_{e_q,c} \equiv 1 - \frac{p^2_q}{\theta^2 h_2\left(e^2_s, e^2_q\right)} = \frac{\lambda^1}{\gamma^2}\frac{p^2_q - p^1_q}{\theta^2 h_2\left(e^2_s, e^2_q\right)} \leq 0,$$

$$(35) \qquad \Omega^2_{e_s,c} \equiv 1 - \frac{p_s}{\theta^2 h_1\left(e^2_s, e^2_q\right)} = 0,$$

where λ^2 and λ^1 denote the Lagrange multipliers associated with constraint (23) and constraint (24), respectively, $MRTS^i \equiv \frac{h_1\left(e^i_s,e^i_q\right)}{h_2\left(e^i_s,e^i_q\right)}$ and $MRTS^{21} \equiv \frac{h_1\left(e^1_s,\widehat{e}^2_q\right)}{h_2\left(e^1_s,\widehat{e}^2_q\right)}$, and $\hat{e}^2_q$ is the quality effort chosen by a type 2 mimicker when pooling with type 1 agents at income level y^1, as defined by (25).

Proof. See Online Appendix E. □

Starting with part (a), condition (27) implies that in a PTE, the effort mix of type-1 agents is undistorted, while the effort mix of type-2 agents is distorted in the direction of e_s. Both results are driven by our assumption that the social objective is to maximize the welfare of type-1 agents, together with the fact that in a PTE all agents choose a common effort mix $(\hat{e}_s, \hat{e}_q)$. This is thus chosen to minimize the cost incurred by type-1 agents to earn $\hat{y}$. However, since type-2 individuals have a comparative advantage in the e_q dimension, they would be better off with a higher e_q and a lower e_s, implying that $(\hat{e}_s, \hat{e}_q)$ entails a distortion towards e_s for them. Since h represents the acquired human capital, condition (27) also implies that in a PTE the acquired human capital of type-1 agents is distorted upward (as stated in eq. (28)), while the acquired human capital of type-2 agents is distorted downward (as stated in eq. (29)).

Now consider part (b) of Proposition 3, which refers to the case of a separating tax equilibrium. Condition (30) implies that the effort-mix of type-1 agents is distorted in the direction of e_s.[15] An intuition for this result comes from the observation that, starting from an initial situation where type-1 agents are induced to choose an undistorted effort mix, the introduction of a small distortion in the direction of e_s has only a second-order welfare effect on type-1 agents, while it has a first-order detrimental effect on the welfare of type-2 mimickers. This in turn allows relaxing the binding incentive compatibility constraint (23).[16]

[15] Recall that our focus on a max-min social objective implies that the downward IC constraint (23) is necessarily binding, i.e. $\lambda^2 > 0$.

[16] Suppose that, on the isoquant $\theta^1 h(e_s, e_q) = y^1$, type-1 agents were initially induced to choose the effort mix (e_s^1, e_q^1) that satisfies the no-distortion condition $\frac{h_1(e_s^1, e_q^1)}{h_2(e_s^1, e_q^1)} = \frac{p_s}{p_q^1}$. Since the government can observe e_s, the type-2 mimickers must also choose e_s^1, and so their effort choice is given by $(e_s, e_q) = (e_s^1, \widehat{e}_q^2)$, where $\widehat{e}_q^2$ satisfies the equation $\overline{\theta} h(e_s^1, \widehat{e}_q^2) = y^1$, so $\widehat{e}_q^2 < e_q^1$. Since $p_q^2 < p_q^1$ and $\frac{h_1(e_s^1, \widehat{e}_q^2)}{h_2(e_s^1, \widehat{e}_q^2)} < \frac{h_1(e_s^1, e_q^1)}{h_2(e_s^1, e_q^1)}$, it follows that $MRTS^{21} = \frac{h_1(e_s^1, \widehat{e}_q^2)}{h_2(e_s^1, \widehat{e}_q^2)} < \frac{p_s}{p_q^2}$, meaning that type-2 mimickers would be forced to choose an effort mix biased toward e_s. Now consider the effect of a perturbation that induces type-1 agents to choose the effort mix $(e_s^1 + de_s, e_q^1 - \frac{h_1(e_s^1, e_q^1)}{h_2(e_s^1, e_q^1)} de_s)$, where de_s is positive and small. By construction, the new effort mix still belongs to the isoquant $\theta^1 h(e_s, e_q) = y^1$. Moreover, it entails only a second-order

Eqs. (31)–(32) shed light on the distortion of each given dimension of effort relative to consumption. According to (31), e_q^1 is distorted downward relative to consumption. This happens for two reasons. On the one hand, type-1 agents incur higher costs to acquire e_q compared to type-2 agents (since $p_q^1 > p_q^2$), and thus also compared to type-2 as mimickers. On the other hand, the marginal productivity of e_q is lower for type-1 agents compared to type-2 mimickers (due to the fact that $\overline{\theta} > \theta^1$ implies $\widehat{e}_q^2 < e_q^1$ and thus $h_2\left(e_s^1, e_q^1\right) < h_2\left(e_s^1, \widehat{e}_q^2\right)$). Taken together, these two circumstances imply that the additional cost that type-1 agents would incur in raising e_q^1 to the extent necessary to earn an additional dollar exceeds the corresponding cost for type-2 agents acting as imitators.

Eq. (32) tells us that in general one cannot determine the direction of the optimal distortion of e_s^1 (relative to consumption). This is due to the fact that one cannot unambiguously assess whether the marginal productivity of e_s is higher or lower for a type-1 agent compared to a type-2 mimicker. On the one hand, the fact that type-2 agents are more productive suggests that the marginal productivity of e_s should be lower for type-1 agents than for type-2 mimickers; this provides a motive to bias e_s^1 downward. On the other hand, the higher productivity of type-2 agents also implies that $\widehat{e}_q^2 < e_q^1$, which in turn implies (assuming a positive cross derivative h_{12}) that $h_1\left(e_s^1, e_q^1\right) > h_1\left(e_s^1, \widehat{e}_q^2\right)$; this represents a motive to distort e_s^1 upwards. Note that since $p_s^1 = p_s^2 = p_s$, price considerations play no role in determining the direction of the distortion. Note also that, at least for the case where the h function is additively separable in e_s and e_q, one can clearly conclude that e_s^1 is distorted downward relative to consumption.

Now consider the equations (33)–(35), which provide expressions for the wedges characterizing the allocation obtained by type-2 agents, and notice that λ^1 can be either positive (the upward IC constraint (24) is binding) or zero (the upward IC constraint (24) is slack).[17]

increase in the total costs borne by type-1 agents (assuming that the pre-reform effort mix satisfied the condition $\frac{h_1(e_s^1, e_q^1)}{h_2(e_s^1, e_q^1)} = \frac{p_s}{p_q^1}$). However, by exacerbating the initial distortion that characterizes the effort mix of type-2 imitators, the proposed reform would have a negative first-order effect on them.

[17] A necessary but not sufficient condition for $\lambda^1 > 0$ is that the upward IC constraint associated with the low-skilled workers is binding under *laissez-faire*. This is because the redistribution in favor of type-1

When $\lambda^1 = 0$, all wedges are zero in the allocation received by type-2 agents. When $\lambda^1 > 0$, eq. (33) tells us that the effort mix of type-2 agents is distorted in the direction of e_q (i.e., e_q^2 is distorted upward relative to e_s^2). The reason is that this is the dimension of effort in which type-2 agents have a comparative advantage over their type-1 counterparts. Thus, by distorting the effort mix of type-2 agents in the direction of e_q, one can make imitation by type-1 agents less attractive. The intuition behind this result can again be captured by a perturbation argument. For a given isoquant $\theta^2 h(e_s, e_q) = y^2$, suppose that type-2 agents are induced to choose the effort mix (e_s^2, e_q^2) that satisfies the no-distortion condition $MRTS^2 = \frac{p_s}{p_q^2}$. From constraint (24) we know that type-1 agents, when acting as mimickers, replicate the effort choices of type-2 agents. Given that $p_q^2 < p_q^1$, it follows that type-1 agents, when acting as mimickers, are forced to choose an effort mix that is distorted toward e_q. Now suppose that instead of letting type-2 agents satisfy the condition $MRTS^2 = \frac{p_s}{p_q^2}$, they are induced to choose an effort mix that is slightly distorted toward e_q. If the distortion is small, it will only have a second order effect on their total cost $p_s e_s^2 + p_q^2 e_q^2$; however, by increasing $p_s e_s^2 + p_q^1 e_q^2$, it will have a first order negative effect on type 1 mimickers.

According to (34), when $\lambda^1 > 0$, e_q^2 is unambiguously distorted upward relative to consumption. This happens because, compared to type-1 agents, type-2 agents incur a lower cost to acquire e_q ($p_q^2 < p_q^1$). Thus, the additional cost that type-2 agents would incur to raise e_q^2 to the extent necessary to earn an additional dollar is less than the corresponding cost for type-1 agents acting as mimickers.

Finally, looking at (35), we can see that e_s^2 is always undistorted relative to consumption. The reason for this is a combination of two circumstances. First, the marginal cost of acquiring e_s is the same for all agents. Second, when acting as mimickers, type-1 agents replicate the effort choices of type-2 agents. Taken together, these two circumstances imply that the additional cost that type-2 agents would incur if they were to raise e_s^2 to the extent necessary to earn an additional dollar is the same as for type-1 agents acting as

agents that occurs through the tax system necessarily reduces the incentive for type-1 agents to mimic type-2 agents.

mimickers.

4 Implementation through means-tested education subsidies or mandates

We now turn to discuss how the wedges given in Proposition 3 translate into properties of the implementing tax structure. We start with the case where the MMO is given by an STE.

4.1 Implementation of the STE

If the MMO is given by an STE, the implementation can be achieved by combining a nonlinear income tax with an income-contingent subsidy scheme for education. In particular, one can obtain the following result.

Proposition 4. *Let σ^* and $\widehat{\sigma}$ be defined as*

$$\sigma^* \equiv \frac{\lambda^2}{\gamma^1}\left(MRTS^1 - MRTS^{21}\frac{p_q^2}{p_q^1}\right)\frac{p_q^1}{p_s} > 0, \tag{36}$$

$$\widehat{\sigma} \equiv \left(1 + \frac{\lambda^2}{\gamma^1}\right)\left(MRTS^1 - MRTS^{21}\frac{p_q^2}{p_q^1}\right)\frac{p_q^1}{p_s} > \sigma^*. \tag{37}$$

Moreover, denote by (e_s^{int}, e_q^{int}) the intersection point between the two isocost lines:

$$(1-\widehat{\sigma})p_s e_s + p_q^1 e_q = (1-\widehat{\sigma})p_s e_s^1 + p_q^1 e_q^1 \tag{38}$$

$$(1-\widehat{\sigma})p_s e_s + p_q^2 e_q = (1-\widehat{\sigma})p_s e_s^1 + p_q^2 \widehat{e}_q^2. \tag{39}$$

Suppose $\theta^2 h(e_s^{int}, e_q^{int}) \leq y^1$. Implementation can then be achieved by an income-dependent subsidy scheme for e_s, denoted by $S(e_s, y)$, and a nonlinear income tax function $T(y)$, *satisfying:*

$$S(e_s, y) = \begin{cases} \widehat{\sigma} p_s e_s, & \text{if } 0 \leq e_s \leq e_s^1 \text{ and } y = y^1, \\ \left[e_s^1 \widehat{\sigma} + (e_s - e_s^1)\sigma^*\right] p_s, & \text{if } e_s > e_s^1 \text{ and } y = y^1, \\ 0, & \text{otherwise}, \end{cases} \tag{40}$$

$$T(y) = \begin{cases} y^1 - c^1 + \widehat{\sigma} p_s e_s^1, & \text{if } y = y^1, \\ y^2 - c^2, & \text{if } y \neq y^1. \end{cases} \tag{41}$$

Suppose instead that $\theta^2 h(e_s^{int}, e_q^{int}) > y^1$. *Implementation can then be achieved by:*

$$S(e_s, y) = \begin{cases} p_s e_s, & \text{if } 0 \leq e_s \leq e_s^{int} \text{ and } y = y^1, \\ \left[e_s^{int} + (e_s - e_s^{int})\widehat{\sigma}\right] p_s, & \text{if } e_s^{int} < e_s \leq e_s^1 \text{ and } y = y^1, \\ \left[e_s^{int} + (e_s^1 - e_s^{int})\widehat{\sigma} + (e_s - e_s^1)\sigma^*\right] p_s, & \text{if } e_s > e_s^1 \text{ and } y = y^1, \\ 0, & \text{otherwise}, \end{cases} \tag{42}$$

$$T(y) = \begin{cases} y^1 - c^1 + \left[e_s^{int} + (e_s^1 - e_s^{int})\widehat{\sigma}\right] p_s, & \text{if } y = y^1, \\ y^2 - c^2, & \text{if } y \neq y^1. \end{cases} \tag{43}$$

Proof. See Online Appendix F. □

The implementation scheme described in Proposition 4, which may look a bit overwhelming, is actually quite simple. It is based on a nonlinear income tax supplemented by an income-dependent subsidy schedule for e_s that is piecewise linear and follows a declining scale. In equilibrium, the education subsidy is provided exclusively to low-skilled workers (who produce a low level of income), which serves to distort their effort mix (toward the quantity dimension) in order to make mimicking more costly for high-skilled workers (whose effort mix remains undistorted).

The kinks characterizing the schedule $S(e_s, y)$ are required to ensure that any agent earning y^1 has the incentive to choose e_s^1, i.e., the constrained efficient level of e_s associated with type-1 agents. Note that a proportional subsidy set at the rate σ^*, as defined by (36), would be sufficient to guarantee that type-1 agents are incentivized to choose e_s^1 on the isoquant $\theta^1 h(e_s, e_q) = y^1$; this is because type-1 agents satisfy their first-order condition:

$$\frac{h_1(e_s, e_q)}{h_2(e_s, e_q)} = \frac{(1-\sigma^*) p_s}{p_q^1} = \frac{p_s}{p_q^1} - \frac{\sigma^* p_s}{p_q^1}.$$

The subsidy rate σ^* produces the wedge provided by (30). However, such a proportional subsidy would not be sufficient for implementation purposes. The reason is that type-2 agents, when acting as mimickers and earning y^1, might find it optimal to choose a different value for e_s; in particular, given that $\frac{p_s}{p_q^2} > \frac{p_s}{p_q^1}$, they may have an incentive to choose $e_s < e_s^1$.[18] To avoid this possibility, a kinked schedule is needed: for $e_s \leq e_s^1$, the subsidy rate should be large enough to ensure that type-2 agents, if they behave as mimickers and earn y^1, have no incentive to choose a value for e_s that is less than e_s^1; for $e_s > e_s^1$, the subsidy rate should be small enough to ensure that type-1 agents have no incentive to choose a value for e_s that is greater than e_s^1. Thus, one should set $\sigma > \sigma^*$ for $e_s \leq e_s^1$ and $\sigma = \sigma^*$ for $e_s > e_s^1$.

Regarding how much larger than σ^* the subsidy rate for $e_s \leq e_s^1$ should be, one should consider the various off-equilibrium strategies available to type-2 agents if they decide to behave as mimickers. One possibility is for them to earn y^1 by pooling with their type-1 counterpart on a common effort vector. In this case, type-2 agents would be rewarded according to the average productivity $\overline{\theta}$, and (37) defines the subsidy rate needed to induce them to choose $e_s = e_s^1$. In fact, $\widehat{\sigma}$ is defined to reflect the wedge faced by type-2 agents at the off-equilibrium effort mix $\left(e_s^1, \widehat{e}_q^2\right)$, where $\widehat{e}_q^2$ is implicitly defined as the solution to

[18] The reason this poses an implementability problem is that the incentive compatibility constraint (23) is binding in the MMO. The right-hand side of this constraint provides the utility of type-2 agents as mimickers when earning y^1 and pooling with type-1 agents at the effort mix $\left(e_s^1, \widehat{e}_q^2\right)$. Thus, if type-2 mimickers have a better deviation strategy available, implementability breaks down.

the equation:[19]

$$\overline{\theta} h\left(e_s^1, \widehat{e}_q^2\right) = y^1.$$

In other words, the subsidy rate $\widehat{\sigma}$ satisfies the first-order condition:

$$\frac{h_1\left(e_s^1, \widehat{e}_q^2\right)}{h_2\left(e_s^1, \widehat{e}_q^2\right)} = \frac{(1-\widehat{\sigma})\, p_s}{p_q^2} = \frac{p_s}{p_q^2} - \frac{\widehat{\sigma} p_s}{p_q^2}.$$

The other available off-equilibrium strategy is for type-2 agents to earn y^1 by choosing an effort mix that allows them to achieve separation from their type-1 counterpart. Since type-2 agents have a comparative advantage in the e_q dimension, separation would necessarily require them to choose, on the isoquant $\theta^2 h\left(e_s, e_q\right) = y^1$, an effort mix such that $e_s < e_s^1$ and $e_q > e_q^1$.

The problem with letting $\sigma = \widehat{\sigma}$ for all values of $e_s \leq e_s^1$ is that, in general, it does not exclude the possibility that, as mimickers, type-2 agents may be better off earning y^1 and achieving separation than earning y^1 and pooling with type-1 agents at the effort mix $\left(e_s^1, \widehat{e}_q^2\right)$. For this reason, implementation may require the introduction of a third segment on the subsidy schedule $S\left(e_s, y^1\right)$.

Whether or not an additional third segment is needed depends on the location of the point defined as $\left(e_s^{int}, e_q^{int}\right)$ in Proposition 4, where the two isocost lines (38) and (39) intersect. The first isocost line is for type-1 agents and passes through the point $\left(e_s^1, e_q^1\right)$; the second (flatter than the first because $p_q^1 > p_q^2$) belongs to type-2 agents and passes through the point $\left(e_s^1, \widehat{e}_q^2\right)$. If $\theta^2 h\left(e_s^{int}, e_q^{int}\right) \leq y^1$, there is no need for an additional segment on the subsidy schedule. The reason is that on the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$ there is no pair $\left(e_s, e_q\right)$ that is also below the isocost line (39) (i.e., for type-2 agents, it entails an effort cost lower than the one they would incur if they mimicked by pooling,

[19]Notice that, exploiting the fact that $\frac{\lambda^2}{\gamma^1}\left(MRTS^1 - MRTS^{21}\frac{p_q^2}{p_q^1}\right)\frac{p_q^1}{p_s} = 1 - MRTS^1\frac{p_q^1}{p_s}$, we have that

$$\widehat{\sigma} = \left(1 + \frac{\lambda^2}{\gamma^1}\right)\left(MRTS^1 - MRTS^{21}\frac{p_q^2}{p_q^1}\right)\frac{p_q^1}{p_s} = 1 - MRTS^{21}\frac{p_q^2}{p_s},$$

where $MRTS^1 = \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)}$ and $MRTS^{21} = \frac{h_1\left(e_s^1, \widehat{e}_q^2\right)}{h_2\left(e_s^1, \widehat{e}_q^2\right)}$.

i.e., earning y^1 while being paid according to the average productivity $\overline{\theta}$) and above the isocost line (38) (meaning that type-1 agents would be discouraged from replicating the effort choices of type-2 agents). Instead, if $\theta^2 h\left(e_s^{int}, e_q^{int}\right) > y^1$, type-2 mimickers are better off earning y^1 and achieving segregation than pooling with type-1 agents at the effort mix $\left(e_s^1, \widehat{e}_q^2\right)$. But since the incentive compatibility constraint (23) is binding in the MMO, it follows that a two-bracket subsidy schedule with $\sigma = \widehat{\sigma}$ for $e_s \leq e_s^1$ and $\sigma = \sigma^*$ for $e_s > e_s^1$ does not ensure implementation. In this case, the schedule $S\left(e_s; y^1\right)$ must be adjusted by introducing an additional segment, for $e_s \leq e_s^{int}$, with an associated subsidy rate of 100%.[20] This full subsidy implies that, on the isoquant $\theta^2 h\left(e_s, e_q\right) = y^1$, any point that allows type-2 agents to achieve separation entails for them a higher cost than the one they would incur by pooling with type-1 agents at the effort mix $\left(e_s^1, \widehat{e}_q^2\right)$.

Overall, the subsidy schedule incentivizes the choice $e_s = e_s^1$ by all agents earning y^1, regardless of their type.

The income tax levied, provided by (41) and (43), is designed to balance the public budget, with the revenue raised by taxing the income earned by type-2 agents (i.e., $T\left(y^2\right) = y^2 - c^2$) is used to finance the transfer received by type-1 agents. Moreover, since this transfer is at least partly provided by education subsidies, $T\left(y^1\right)$ must be different depending on whether the subsidy schedule $S\left(e_s; y^1\right)$ has two or three segments.[21]

Finally, note that, somewhat surprisingly, we obtain the canonical efficiency-at-the-top result for high-ability agents (Sadka 1976).[22] One might have expected, for example, that the tax function should have been used to distort the effort allocation of type-2

[20] In order to maintain a balanced public budget, the introduction of an additional segment requires a corresponding adjustment of the income tax levied at $y = y^1$.

[21] As we discuss in the Online Appendix, another implementing scheme could be obtained by assuming that, for $y = y^1$, $\sigma\left(e_s\right) = 100\%$ for $e_s \leq e_s^1$ and $\sigma\left(e_s\right) = 0$ for $e_s > e_s^1$, which would be equivalent to adopting a system with an income-based education mandate. In this case, in order to maintain a balanced public budget, one should properly adjust the income tax function by increasing $T\left(y^1\right)$; in particular, one should set $T\left(y^1\right) = y^1 - c^1 + p_s e_s^1$. For the case where $e_s^1 \leq e_s^2$, an income-independent education mandate requiring that $e_s \geq e_s^1$ would suffice.

[22] To see this, note that for any y other than y^1 there is only an income tax and no education subsidy. Moreover, for all y other than y^1, the income tax is constant, implying a zero marginal tax rate.

agents toward e_q—the dimension in which they have a comparative advantage—in order to discourage type-1 agents from mimicking them. However, a key insight is that the IC constraint for type-1 agents is already embedded in the laissez-faire equilibrium. As a result, type-2 agents have already internalized this constraint when making their decisions. The labor contract offered to type-2 agents is designed to maximize their utility, subject to the IC constraint of type-1 agents. This is consistent with the government's goal of extracting as much revenue as possible from type-2 agents to facilitate redistribution. It is also worth noting that although the marginal tax rates for high-skilled agents are zero under the implementing tax function, the income level y^2 in the STE is lower than under laissez-faire when the upward IC constraint for low-skilled workers binds in the laissez-faire scenario. This result arises because redistribution through the tax system increases the utility of type-1 agents relative to their utility under laissez-faire, thereby reducing their incentive to imitate type-2 agents.

4.2 Implementation of the PTE

If the MMO is given by a PTE, the implementation can be achieved by the combined use of a tax that depends only on income and a mandate that enforces a lower bound on e_s. In particular, one can obtain the following result.

Proposition 5. *Let $e_q^{\min}$ be the value of e_q that solves the following problem:*

$$\min_{e_q} \theta^2 h\left(\widehat{e}_s, e_q\right) \quad \text{subject to } \theta^1 h\left(\widehat{e}_s, \widehat{e}_q\right) - T\left(\theta^1 h\left(\widehat{e}_s, \widehat{e}_q\right)\right) - p_q^1 \widehat{e}_q \geq \theta^2 h\left(\widehat{e}_s, e_q\right) - p_q^1 e_q,$$

and define $\underline{y}^{sep}$ as $\underline{y}^{sep} \equiv \theta^2 h\left(\widehat{e}_s, e_q^{\min}\right)$. Furthermore, denote by $\left(e_s^{2*}, e_q^{2*}\right)$ the effort mix that solves the following unconstrained maximization problem:

$$\max_{e_s, e_q} \theta^2 h\left(e_s, e_q\right) - p_s e_s - p_q^2 e_q. \tag{44}$$

Implementation can be achieved by combining a binding mandate on e_s, set to $e_s = \widehat{e}_s$, *with an income tax $T(y)$ such that*

(45)

$$T(y) = \begin{cases} \left(\frac{1}{\theta^1} - \frac{1}{\overline{\overline{\theta}}}\right) \frac{p_q^1}{h_2(\widehat{e}_s,\widehat{e}_q)}\widehat{y} + \left(\frac{1}{\overline{\overline{\theta}}} - \frac{1}{\theta^1}\right) \frac{p_q^1}{h_2(\widehat{e}_s,\widehat{e}_q)}y, & \textit{for all } y \in [0, \widehat{y}] \\ (y - \widehat{y}) \max\left\{1 - \frac{p_q^2}{\overline{\overline{\theta}} h_2(\widehat{e}_s,\widehat{e}_q)}, \frac{\left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s \widehat{e}_s - p_q^2 \widehat{e}_q\right]}{\underline{y}^{sep} - \widehat{y}}\right\}, & \textit{for all } y > \widehat{y}. \end{cases}$$

Proof. See Online Appendix G. □

Formula (45) defines a two-bracket piecewise linear income tax with a kink at $y = \widehat{y}$, a negative marginal tax rate on the first bracket, a positive marginal tax rate on the second bracket, and a U-shaped profile of average tax rates (always positive except at $y = \widehat{y}$, where the average tax rate is zero). The negative marginal tax rate on the first bracket serves to distort the acquired human capital of type-1 agents upward and to incentivize them to choose the effort mix $(\widehat{e}_s, \widehat{e}_q)$.[23]

The (positive) marginal tax rate on the second bracket serves to distort downward the acquired human capital of type-2 agents. It is designed to be high enough to achieve two goals: (i) to ensure that type-2 agents (weakly) prefer pooling at $\widehat{y}$ to pooling at a higher income, and, (ii) to discourage type-2 agents from choosing an effort mix that would allow them to achieve separation from their low-ability counterpart at an income level higher than $\widehat{y}$.

The marginal tax rate on the second bracket achieves both of these goals because it is given by the maximum of two quantities. The first term in the max operator represents the tax rate that guarantees that type-2 agents will not prefer to pool at an income higher than $\widehat{y}$. The second term represents the tax rate that guarantees that type-2 agents will be discouraged from achieving separation from their low-ability counterpart.

In particular, note that the marginal tax rate given by the second term in the max operator is defined as an expression that depends on both $\underline{y}^{sep}$ and $\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}$. The former represents the minimum amount of income that type-2 agents would

[23] Faced with a zero marginal tax rate, type-1 agents would choose $e_s = \widehat{e}_s$ (because of the lower bound on e_s set by the mandate), but $e_q < \widehat{e}_q$.

need to earn to achieve separation from their type-1 counterparts in an environment where $T(y) = 0$ for $y > \widehat{y}$. In turn, $\underline{y}^{sep}$ is a function of $e_q^{\min}$, which represents the minimum level of e_q that allows type-2 agents to achieve separation when choosing $e_s = \widehat{e}_s$.

The quantity $\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}$ represents an upper bound for the utility that could be achieved by type-2 agents under laissez-faire. In particular, given that the effort mix $\left(e_s^{2*}, e_q^{2*}\right)$ is defined as the one that solves the unconstrained maximization problem (44), the quantity $\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}$ represents the utility that would be achieved by type-2 agents in a laissez-faire setting without asymmetric information in the labor market.

In a setting where $T(y) = 0$ for $y > \widehat{y}$, the gain that type-2 agents can achieve by separating from their low-ability counterpart (instead of choosing the effort mix $(\widehat{e}_s, \widehat{e}_q)$ and pooling with them at $\widehat{y}$) cannot exceed the amount:

$$\left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s\widehat{e}_s - p_q^2\widehat{e}_q\right].$$

Note, however, that this is exactly the income tax that would be paid at $y = \underline{y}^{sep}$, based on the definition of the marginal tax rate provided by the second term in the max operator of (47). Thus, such a marginal tax rate prevents type-2 agents from being tempted to achieve separation from their type-1 counterparts.

The binding mandate on e_s serves primarily to ensure the stability of the PTE. The reason is that it prevents type-2 agents from choosing an effort mix that would allow them to earn $\hat{y}$ while being compensated according to their true productivity θ^2 rather than the average productivity $\overline{\theta}$. More generally, the lower bound on e_s helps preserve the PTE because it effectively raises the cost that type-2 agents would have to incur to achieve separation.

Note also that a binding mandate on e_s, set at $e_s = \widehat{e}_s$ is an extreme version of a nonlinear tax on e_s with a large marginal subsidy for values of e_s less than $e_s = \widehat{e}_s$ and a zero marginal tax/subsidy elsewhere. This suggests that the implementation of the PTE could also be achieved by supplementing a piecewise linear tax on income with a piecewise linear tax on e_s with a sufficiently large marginal subsidy on the first bracket.

Finally, note that public provision of education is another way to implement PTE. In particular, suppose that the government publicly provides e_s free of charge up to a maximum amount $\widehat{e}_s$, so that agents only have to bear the marginal cost p_s for those units of e_s that exceed $\widehat{e}_s$. The implementation of the PTE could then be achieved by supplementing this public provision scheme with an income tax $\widetilde{T}(y)$ given by a uniform upward shift, by an amount $p_s\widehat{e}_s$, of the income tax function $T(y)$ provided in (45), namely $\widetilde{T}(y) = T(y) + p_s\widehat{e}_s$.[24]

4.3 Relation to existing policy instruments

In the previous subsection, we have shown how supplementing the income tax system with a means-tested education subsidy or an education mandate serves to implement the MMO (given by either an STE or a PTE).

Means-tested subsidies for education, which play a dual role of correcting market failures and achieving redistributive goals, exist in many countries, either as part of the general tax system or, as has become quite common in recent years, in the form of income-contingent student loans. Student loans are often offered on favorable terms and are used to cover tuition fees and/or living expenses, depending on the country. The size of the subsidy depends on the difference between the tuition charged and the actual cost of providing the education, as well as the extent to which the loans are offered at below-market (subsidized) rates. A notable example is Australia's Higher Education Loan Program (HELP), where students receive loans to finance their education, which are repaid once their income exceeds a certain threshold.[25] The threshold and repayment rate vary depending on the borrower's income level. In 2023–2024 the income threshold is AUD 51,550 and above this threshold the repayment rate varies from 1 percent to a maximum of

[24]The uniform upward shift is necessary to ensure that the government's budget constraint is still satisfied. In particular, under this alternative implementation scheme, each agent will pay an income tax of $p_s\widehat{e}_s$ at the PTE, allowing the government to raise enough revenue to cover the public expenditures associated with public provision.

[25]According to Australian Government Department of Education, Skills and Employment (2020), approximately 2.8 million Australians will owe AUD 68.1 billion in HELP debt in 2020.

10 percent for incomes above AUD 151,201. The income-contingent repayment system is essentially a means-tested progressive tax on graduates, as high-achieving students reach the income threshold earlier and earn higher wages. Similar income-contingent repayment systems exist in the United Kingdom and Sweden, as well as in many other countries.[26]

Education mandates are common in the real world and are often justified on both efficiency and equity grounds. Such mandates typically take the form of minimum compulsory schooling laws, commonly applied in the context of primary/secondary education.

We offer novel normative justifications for the use of both means-tested education subsidies and education mandates (in the context of postsecondary education) to promote redistributive goals by limiting the ability of high-skilled individuals to engage in signaling that serves to separate them from their low-skilled counterparts. Accordingly, a notable feature of our analysis is that both policy instruments should target those components of educational effort in which low-skilled agents have a comparative advantage.

5 Discussion

We next discuss how the case for predistribution in the MMO depends on the observability assumptions (sections 5.1–5.3), the number of signals (section 5.4), and the number of types in the economy (section 5.5).

5.1 The case where neither signal is observable

In Online Appendix H we study the case where the government can only observe income. In this case, due to the weaker policy instruments available, the possibilities for mimickers

[26] In Sweden, student loans have relatively favorable terms compared to many other countries. Repayment usually begins the year after the student graduates, and the repayment period can last up to 25 years. The interest rate on these loans is set by the government and is usually very low. Notably, the repayment amount is based on the borrower's income, making it an income-contingent repayment plan. This means that the amount a graduate pays back each year is a percentage of his or her income above a certain threshold, ensuring that repayments are affordable. If a borrower's income is below that threshold, he or she may be eligible for a repayment waiver for that year.

to deviate are expanded. The main insight from our analysis is that predistribution is not feasible with only an income tax. Thus, the ability to tax the signals transmitted in the labor market is essential to achieve predistribution. In Online Appendix K, we use the case with only an income tax as a benchmark to numerically quantify the welfare gains of taxing the quantity signal. Note that the welfare gains from taxing the education signal arise regardless of whether the MMO features predistribution or not. However, consistent with Proposition 1, the results show that the MMO tends to feature predistribution when the productivity variance between the two categories of workers and the discrepancy in the cost of obtaining the quality signal across types are moderate.

5.2 The case when both signals are taxed

In Online Appendix I we characterize the optimal tax structure under the assumption that the government can tax both quantity and quality signals. In this case, while the government can eliminate the information rent from productivity differences between workers, a residual information rent remains for type-2 workers due to the difference in the cost of acquiring the quality signal. Thus, the first-best allocation remains unattainable. The government's options are the same as when it could only tax the quantity signal: it can implement a pooling or a separating equilibrium. However, there is a difference now: with both signals being observable by the government, a mimicker is always forced to replicate the effort choices of the mimicked type (the mimicker cannot adapt in any other way). As Online Appendix I shows, this implies that the MMO is always an STE. When both signals can be taxed/subsidized, a separating equilibrium is cheaper (more efficient) than a pooling equilibrium in eliminating the information rent arising from productivity differences. A key insight from this analysis is that while the feasibility of predistribution hinges on the ability to tax at least one of the two signals, the desirability of predistribution depends crucially on the government's inability to tax both signals.

5.3 The observable signal is e_q instead of e_s

Our analysis has focused on the case where the signal observable to the government is e_s. It is worth noting that while the PTE does not change depending on which of the two signals is assumed to be observable to the government, the same is not true for the STE.[27] Consequently, the assumption about which signal is observable is not unimportant for comparing the welfare properties of pooling and separating tax equilibria. For the Cobb-Douglas example studied in section 2.5, one can show that the STE achieved when the government observes e_s is always welfare superior to the STE achieved when the government observes e_q.[28] This implies that a PTE becomes relatively more attractive when the signal observed by the government is the one for which type-2 agents have a comparative advantage.

With respect to wedges, the most interesting difference between the STE when the observable signal is e_s and the STE when the observable signal is e_q is that in the latter case it is a priori ambiguous in which direction it is optimal to distort the effort mix of type-1 agents. This contrasts with the result provided by (30) for the case where the observable signal is e_s, namely that the effort mix chosen by type-1 agents should be distorted towards the effort dimension at which they have a comparative advantage (i.e., e_s). When the observable signal is e_q instead of e_s, it may happen that mimicking-deterrence considerations justify distorting the effort mix chosen by type-1 agents towards e_q.[29]

5.4 More than two signals

As noted above, if the government cannot observe and tax/subsidize (both income and) all signals, then it can only reduce (but not eliminate) the information rent from productivity differences. One might therefore think that a pooling equilibrium would be better for

[27] An intuition for this result is provided in the first part of Online Appendix J.

[28] However, this is not a general result, and one can easily construct counterexamples where the opposite result holds.

[29] An intuition for this result is provided in the second part of Online Appendix J.

equity, and that the case for pooling would be stronger, if fewer signals were taxed. The problem with this argument is that it ignores the fact that pooling must be sustainable in order to be socially desirable. In general, where there are n signals, pooling will be sustainable either if the government taxes (at least) $n-1$ signals, or if it taxes $n-j$ signals (with $1<j<n$) and the high-skill types have no comparative advantage in the untaxed signals.

A possible example of adding more signals is when individuals can commit to their hours of work/availability (in addition to the quality and quantity of educational effort). Maintaining our assumptions that $p_s^1 = p_s^2$ and $p_q^1 > p_q^2$, and assuming that work/leisure preferences are the same across types and that labor costs are separable, the result would be that conditioning the tax function on both income and the quantity signal e_s would not be sufficient to make predistribution feasible. The reason is that within the set of untaxed signals (in this case e_q and hours worked), the high-skilled types have a comparative advantage in one dimension (e_q). However, predistribution would be feasible if the observable signal were e_q (instead of e_s). This is because in such a case the tax function could be conditioned on both income and e_q, implying that the high-skilled types have no comparative advantage within the set of untaxed signals (e_s and hours worked).

Of course, endogenizing labor supply in this way hinges on the assumption that the worker pre-commits to his workload, and then the firm uses this information (as well as information about the worker's educational background) to decide on the level of compensation. Alternatively, one could assume that the order is reversed (the firm is the first mover), in which case the model combines signaling (via ex-ante investment in education) with screening (via ex-post choice of hours), making the analysis much more complicated. This latter configuration, while interesting, is beyond the scope of the current analysis.

Before concluding this subsection, a note on the measurement of comparative advantage is in order. For simplicity, our model assumes that agents are free to adjust the signal (quantity and quality efforts are continuous variables). In reality, such adjustment is usually more constrained. For example, schooling may be limited to a high school diploma or a college degree, and working hours (except in the "gig" economy) may be limited to full-time (say, 40 hours per week) or part-time (20 hours per week). This should be taken into account, at least empirically, when assessing comparative advantage. Within the limited set, high-skilled types may not be able to distinguish themselves from their low-skilled counterparts.

5.5 More than two types

To keep our analysis tractable, we have limited our attention to a model with two types. The case with more than two types is more complex because the number of incentive constraints increases significantly. There are also more tax equilibrium configurations to consider, since some types may be pooled while others are separated. Nevertheless, the main qualitative insight that constrained efficient allocations may involve predistribution is not sensitive to the number of types. Several features stand out, however.

First, as in the two-type case, predistribution is not feasible when neither signal is observable, since the high-skill types can always separate from the low-skill types. Second, when only the quantity signal is observable, partial pooling (bunching) becomes feasible and may be superior to full pooling and full separation. Third, when both signals are taxed, while a pooling equilibrium with full wage compression can still be shown to be suboptimal (using a similar argument as in Online Appendix I), partial pooling (bunching of a subset of types) can be shown to be desirable and superior to full separation. The reason is that bunching can serve to mitigate the downward ("adjacent") IC constraints (type j mimicking type $j-1$), so as to reduce the information rent associated with the cost of acquiring the quality signal. This serves to enhance redistribution through the income channel while achieving redistribution through the wage channel.[30] The reason that bunching is desirable is not to eliminate the information rents associated with the difference in productivity between types (the latter is taken care of by the ability to tax

[30]For example, consider the case with three types 1, 2, and 3, where 3 represents the high-skilled type and 1 represents the low-skilled agent. Implementing a hybrid allocation in which types 1 and 2 are bunched together could be superior to a fully separating allocation by allowing a combination of redistribution from type 3 to its low-skilled counterparts and predistribution between types 1 and 2.

both signals), but rather to increase redistribution along the income channel. Pooling, on the other hand, does not achieve redistribution through the income channel and is therefore suboptimal.

6 Concluding remarks

In this paper, we have introduced a new dimension to the traditional Mirrleesian framework by incorporating a second layer of asymmetric information—between workers and employers—and by allowing the tax system to depend on both income and observable signals in the labor market. Our analysis examines how workers engage in multidimensional signaling through both the quantity and quality of education, and how these signals affect optimal tax policies.

Using a mechanism design approach to the analysis of optimal income taxation, we show that allocations that maximize the utility of low-skilled workers, subject to information and resource constraints, can lead to either separating or pooling equilibria. In the case of separating equilibria, incentive constraints operate in both directions: low-skilled workers may attempt to mimic high-skilled workers to obtain higher compensation, while high-skilled workers may mimic low-skilled workers to reduce their tax burden. This dynamic implies that the effect of the second layer of asymmetric information (between workers and firms) on the level of social welfare achievable through optimal tax policy is generally ambiguous.

In pooling equilibria, predistribution occurs through wage compression, with changes in the wage structure creating cross-subsidies between different skill levels. However, such predistribution is only feasible if signaling activities in the labor market are taxed, making such taxes complementary to traditional instruments for achieving redistribution, such as progressive income taxation. From a policy perspective, we suggest that education mandates and means-tested education subsidies, which are traditionally used to address market failures, can also function as redistributive instruments by mitigating the effects of signaling and achieving predistribution.

Although our model is based on a simplified two-type agent framework, the central insights are likely to extend to more complex settings involving multiple types and signals. In these more complicated settings, the social optimum might combine both predistribution and traditional redistribution, rather than feature full separation or full pooling as in the two-type case.

Our findings suggest the potential effectiveness not only of education mandates and subsidies, but also of a wide range of policies that affect incentives to engage in signaling. These could include policies such as penalties for students who complete their education unusually quickly or restrictions on simultaneous enrollment in multiple programs. Refining income-contingent student loan programs to better target subsidies in areas where low-skilled workers have a comparative advantage would further improve redistributive outcomes. In addition, anti-discrimination laws can play an important role in promoting a more equitable wage distribution by reducing the ability of firms to engage in screening or statistical discrimination.

In conclusion, our paper suggests that predistribution through wage compression is an important and underexplored mechanism for redistribution in the real economy. Future empirical research is needed to examine how policies that limit signaling and screening—whether through educational choice or broader labor market interventions—can promote more equitable compensation for workers at different productivity levels. Such work would be crucial for guiding policymakers in designing effective strategies to reduce inequality.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study

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.

Andersson, F. (1996). Income taxation and job-market signaling. *Journal of Public Economics*, 59(2):277–298.

Australian Government Department of Education, Skills and Employment (2020). 2020-21 Higher Education Loan Program (HELP) Annual Report.

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.

Blomquist, S. and Christiansen, V. (2008). Taxation and heterogenous preferences. *FinanzArchiv*, 64:218–244.

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.

Caplan, B. (2018). *The Case Against Education: Why the Education System Is a Waste of Time and Money.* NJ: Princeton University Press.

Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.

Craig, A. C. (2023). Optimal income taxation with spillovers from employer learning. *American Economic Journal: Economic Policy*, 15(2):82–125.

Cremer, H., Pestieau, P., and Racionero, M. (2011). Unequal wages for equal utilities. *International Tax and Public Finance*, 18:383–398.

Ebert, U. (1992). A reexamination of the optimal nonlinear income tax. *Journal of Public Economics*, 49(1):47–73.

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.

Grossman, S. J. and Perry, M. (1986). Perfect sequential equilibrium. *Journal of Economic Theory*, 39(1):97–119.

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

Manoli, D. (2006). A signaling approach to optimal income taxation. *Unpublished Working Paper*.

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

Riley, J. G. (2001). Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature*, 39(2):432–478.

Rothschild, C. and Scheuer, F. (2013). Redistributive taxation in the Roy model. *The Quarterly Journal of Economics*, 128(2):623–668.

Sachs, D., Tsyvinski, A., and Werquin, N. (2020). Nonlinear tax incidence and optimal taxation in general equilibrium. *Econometrica*, 88(2):469–493.

Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *The Review of Economic Studies*, 43(2):261–267.

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

Spence, M. (1974). Competitive and optimal responses to signals: An analysis of efficiency and distribution. *Journal of Economic Theory*, 7(3):296–332.

Stantcheva, S. (2014). Optimal income taxation with adverse selection in the labor market. *Review of Economic Studies*, 81:1296–1329.

Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy*, 125(6):1931–1990.

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

Sztutman, A. (2024). Dynamic job market signaling and optimal taxation. *Working Paper, Tepper School of Business, Carnegie Mellon University.*

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.

Online Appendix

Optimal Redistribution and Education Signaling

By Spencer Bastani, Tomer Blumkin, and Luca Micheletto

January, 2025.

A Existence of MMO

Consider the optimization problem in Definition 1. This is essentially a nested optimization in which the MMO effectively chooses between separating and pooling optimal allocations, depending on which yields the highest level of social welfare. Note that an optimal pooling allocation exists trivially, since it can be expressed as the solution to an unconstrained optimization problem with a strictly concave objective function that reaches a (unique) global maximum due to the Inada conditions. As for the optimal separating allocation, we show below that the feasible set defined by the constraints is nonempty and compact, which together with the continuity of the objective function imply that an optimal separating allocation always exists by the extreme value theorem. Together, this establishes that the optimization problem in Definition 1 is well defined.

Non-emptiness of the set of feasible separating allocations We prove that the feasible set is non empty by constructing an incentive compatible separating allocation in which $y^1 = e_s^1 = 0$ (implying $e_q^1 = \widehat{e}_q^2 = 0$). Invert the function $y = \theta h(e_s, e_q)$ to obtain $e_q = f\left(\frac{y}{\theta}, e_s\right)$, and further denote by y^* and e_s^* the values of y and e_s that maximize $y - p_s e_s - p_q^2 f\left(\frac{y}{\theta^2}, e_s\right)$. Consider the two quadruplets

$$\text{(A1)} \qquad \left(y^1, c^1, e_s^1, e_q^1\right) = \left(0, \gamma^2\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right], 0, 0\right)$$

and

$$\text{(A2)} \qquad \left(y^2, c^2, e_s^2, e_q^2\right) = \left(y^*, y^* - \gamma^1\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right], e_s^*, f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right).$$

It is straightforward to verify that they satisfy the government revenue constraint (21):

$$\text{(A3)} \qquad \left(y^1 - c^1\right)\gamma^1 + \left(y^2 - c^2\right)\gamma^2 = -\gamma^1\gamma^2\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right]$$
$$+ \gamma^2 y^* - \gamma^2\left\{y^* - \gamma^1\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right]\right\} = 0.$$

Furthermore, they also satisfy the incentive constraints (23) and (24), the former as an equality and the latter as a strict inequality. In particular, taking into account that $\left(y^1, c^1, e_s^1, e_q^1\right) = \left(0, \gamma^2\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right], 0, 0\right)$ implies $R^2\left(e_s^1, \hat{e}_q^2\right) = R^1\left(e_s^1, e_q^1\right) = 0$, constraint (23) simplifies to

$$\text{(A4)} \qquad c^2 - R^2\left(e_s^2, e_q^2\right) \geq c^1,$$

and constraint (24) simplifies to

$$\text{(A5)} \qquad c^1 \geq c^2 - R^1\left(e_s^2, e_q^2\right).$$

Substituting $y^* - \gamma^1\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right]$ for c^2, $\gamma^2\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right]$ for c^1, e_s^* for e_s^2 and $f\left(\frac{y^*}{\theta^2}, e_s^*\right)$ for e_q^2, constraints (A4) and (A5) become, respectively:

$$y^* - \gamma^1\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right] - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right) = \gamma^2\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right],$$

$$\gamma^2\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right] > y^* - \gamma^1\left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right)\right] - p_s e_s^* - p_q^1 f\left(\frac{y^*}{\theta^2}, e_s^*\right).$$

The key point to note is that if type-1 agents refrain from investing in education, the information rent enjoyed by type-2 agents can be driven to zero. This implies that the right-hand side of the constraint (23) takes the same value as the left-hand side of the constraint (24). Thus, if the incentive constraint (23) that applies to type-2 agents is satisfied, the incentive constraint (24) that applies to type-1 agents is necessarily slack (due to the assumption that $p_q^1 > p_q^2$). This concludes the proof.

Compactness of the set of feasible separating allocations Let $\mathcal{S}$ denote the set of feasible separating allocations defined by all allocations with $\left(e_s^1, e_q^1\right) \neq \left(e_s^2, e_q^2\right)$ that satisfy the constraints in Definition 1. Since we restrict our analysis to subsets of a Euclidean space, $\mathcal{S}$ is compact if and only if it is closed and bounded. $\mathcal{S}$ is bounded by the Inada conditions. To prove that $\mathcal{S}$ is closed, we need to show that $\mathcal{S}$ contains all its limit points. We will prove that $\mathcal{S}$ is closed by contradiction. Suppose $\mathcal{S}$ is not closed, and hence there exists a sequence of feasible separating allocations that converges to an allocation that is not an element of $\mathcal{S}$. There are two possible limit points to consider, one with $\left(e_s^1, e_q^1\right) \neq \left(e_s^2, e_q^2\right)$ and the other with $\left(e_s^1, e_q^1\right) = \left(e_s^2, e_q^2\right)$. In the first case, due to the continuity of the constraints defining a separating allocation on the set $\mathcal{S}$, the limit allocation is necessarily an element of $\mathcal{S}$. Therefore, we consider the only other possibility, namely a sequence of feasible allocations with $\left(e_s^1, e_q^1\right) \neq \left(e_s^2, e_q^2\right)$, which converges to an allocation where $\left(e_s^1, e_q^1\right) = \left(e_s^2, e_q^2\right)$.

Rewrite the incentive constraint in (23) to get

$$g(e_s^1, e_q^1, e_s^2, e_q^2) = c^1 - c^2 + R^2\left(e_s^2, e_q^2\right) - R^2\left(e_s^1, \hat{e}_q^2\right) \leq 0, \tag{A6}$$

where $R^2\left(e_s^1, \hat{e}_q^2\right) < R^2(e_s^1, e_q^1)$ since $\hat{e}_q^2 < e_q^1$. Note that the incentive constraint in (24) implies that

$$c^1 - c^2 \geq R^1\left(e_s^1, e_q^1\right) - R^1\left(e_s^2, e_q^2\right). \tag{A7}$$

Thus, we have

$$g(e_s^1, e_q^1, e_s^2, e_q^2) \geq R^1\left(e_s^1, e_q^1\right) - R^1\left(e_s^2, e_q^2\right) + R^2\left(e_s^2, e_q^2\right) - R^2\left(e_s^1, \hat{e}_q^2\right) \tag{A8}$$

$$> R^1\left(e_s^1, e_q^1\right) - R^1\left(e_s^2, e_q^2\right) + R^2\left(e_s^2, e_q^2\right) - R^2\left(e_s^1, e_q^1\right). \tag{A9}$$

Since the right side of (A9) converges to zero as $(e_s^1, e_q^1) \to (e_s^2, e_q^2)$, condition (A6) is necessarily violated if $\|(e_s^1, e_q^1) - (e_s^2, e_q^2)\| < \delta$ for some sufficiently small $\delta > 0$. We thus obtained a contradiction to the existence of a limit allocation in which effort levels are pooled (identical between types), since there exists a small neighborhood of the (presumed) limit point that contains no feasible separating allocations. We conclude that $\mathcal{S}$ must be closed.

B Proof of Proposition 1

Part (a) We first prove that there exist some $\varepsilon > 0$ and $\delta > 0$, where $\varepsilon = \theta^2 - \theta^1$ and $\delta = p_q^1 - p_q^2$, such that the PTE is welfare superior to the separating tax equilibrium. We let $p_s = 1$ without loss of generality; fix p_q^1, θ^1, and γ^1, and let $\varepsilon = \delta > 0$ and small. For the set of given parameters, $(\gamma^1, \theta^1, p_q^1, \varepsilon)$, we can solve for the optimal separating and pooling optima (since both exist, see appendix A). Denote the resulting max-min welfare measures by:

$$W^{sep}(\gamma^1, \theta^1, p_q^1, \varepsilon) \tag{B1}$$

$$W^{pool}(\gamma^1, \theta^1, p_q^1, \varepsilon) \tag{B2}$$

Obviously, for $\varepsilon = 0$, $W^{sep} = W^{pool}$. We will show that for $\varepsilon > 0$ and small, $W^{sep} < W^{pool}$. Using a first-order approximation, it suffices to show this:

$$\left.\frac{\partial W^{sep}(\gamma^1, \theta^1, p_q^1, \varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} < \left.\frac{\partial W^{pool}(\gamma^1, \theta^1, p_q^1, \varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0}. \tag{B3}$$

The separating optimum is given by the solution of the following maximization program (formulated in a *Lagrangian* form for convenience):

$$W^{sep}(\varepsilon) = \max\Big\{[c^1 - e_s^1 - e_q^1 p_q^1] + \mu[\gamma^1(\theta^1 h(e_s^1, e_q^1) - c^1) + (1-\gamma^1)((\theta^1+\varepsilon)h(e_s^2, e_q^2) - c^2)]$$
$$+ \lambda[(c^2 - e_s^2 - e_q^2(p_q^1 - \varepsilon)) - (c^1 - e_s^1 - \hat{e}_q^1(p_q^1 - \varepsilon))]$$
$$\text{(B4)} \qquad + \eta \cdot [\theta^1 h(e_s^1, e_q^1) - (\theta^1 + (1-\gamma^1)\varepsilon)h(e_s^1, \hat{e}_q^1)]\Big\},$$

where μ, λ, and η correspond to the *Lagrange* multipliers associated with the revenue constraint, the type-2 IC-constraint, and the condition implicitly defining the off-equilibrium quality signal effort chosen by a type-2 mimicker (the effort is defined by $\hat{e}_q^1$). Note that we implicitly assume that the IC-constraint of the low-skilled (type-1) agent is slack.

The pooling optimum is given by the following maximization program:

$$\text{(B5)} \qquad W^{pool}(\varepsilon) = \max\{(\theta^1 + (1-\gamma^1)\varepsilon)h(e_s, e_q) - e_s - e_q p_q^1\}.$$

Using the envelope theorem, it follows that:

$$\text{(B6)} \qquad \left.\frac{\partial W^{pool}(\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} = (1-\gamma^1)h(e_s^*, e_q^*)$$

$$\text{(B7)} \qquad \left.\frac{\partial W^{sep}(\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} = (\mu^* - \eta^*)(1-\gamma^1)h(e_s^*, e_q^*),$$

where the asterisk ($*$) refers to the optimal allocations under the separation and pooling configurations. Note that for $\varepsilon = 0$, the effort choices coincide.

To prove our claim, it suffices to show that $\mu^* - \eta^* < 1$. Deriving the first-order conditions of the separating optimal allocation with respect to e_q^1, c^1, c^2 and $\hat{e}_q^1$, evaluated at $\varepsilon = 0$, yields the following:

$$\text{(B8)} \qquad -p_q^1 + \mu^*\gamma^1\theta^1 \frac{\partial h}{\partial e_q^1} + \eta^*\theta^1 \frac{\partial h}{\partial e_q^1} = 0$$

$$\text{(B9)} \qquad 1 - \mu^*\gamma^1 - \lambda^* = 0$$

$$\text{(B10)} \qquad -\mu^*(1-\gamma^1) + \lambda^* = 0$$

$$\text{(B11)} \qquad \lambda^* p_q^1 - \eta^*\theta^1 \frac{\partial h}{\partial e_q^1} = 0.$$

After some algebraic manipulations, one can show that:

$$\mu^* = 1 \tag{B12}$$

$$\lambda^* = (1 - \gamma^1) \tag{B13}$$

$$p_q^1 = \theta^1 \frac{\partial h}{\partial e_q^1} \tag{B14}$$

$$\eta^* = \frac{(1 - \gamma^1) p_q^1}{\theta^1 \frac{\partial h}{\partial e_q^1}} < 1. \tag{B15}$$

Thus, $\mu^* - \eta^* < 1$ as needed (note that $p_q^1 = \theta^1 \frac{\partial h}{\partial e_q^1}$ defines the efficiency condition for the quality signal effort choice, which trivially holds for $\varepsilon = 0$). Note that in the maximization program associated with the separating allocation, we have assumed that the incentive constraint associated with the low-skilled (type-1) agents is slack. It clearly follows that, under the parametric assumptions of the proposition, the pooling allocation is welfare superior to the separating allocation when we account for this additional (potentially binding) constraint. This concludes the proof of part (a).

Part (b) Fixing p_q^1, θ^2, and γ^1, we next prove that for any difference in the productivity between the two types of workers, $0 < \varepsilon \leq \theta^2$, there exists $0 \leq \delta^*(\epsilon) \leq p_q^1$, representing the difference in the cost of acquiring the quality signal, such that the separating allocation is welfare superior to the pooling allocation when $\delta > \delta^*(\epsilon)$, whereas the pooling allocation is welfare superior to the separating allocation when $\delta < \delta^*(\epsilon)$. Reformulation of the maximization program associated with the optimal separating allocation (similar to part (a)), but now taking into account the type-1 incentive constraint:

$$\begin{aligned} W^{sep}(\varepsilon, \delta) = \max \Big\{ & [c^1 - e_s^1 - e_q^1 p_q^1] + \mu[\gamma^1((\theta^2 - \varepsilon) h(e_s^1, e_q^1) - c^1) \\ & + (1 - \gamma^1)(\theta^2 h(e_s^2, e_q^2) - c^2)] + \lambda[(c^2 - e_s^2 - e_q^2(p_q^1 - \delta)) \\ & - (c^1 - e_s^1 - \hat{e}_q^1(p_q^1 - \delta))] + \eta[\theta^1 h(e_s^1, e_q^1) - (\theta^2 - \gamma^1 \varepsilon) h(e_s^1, \hat{e}_q^1)] \\ & + \phi[(c^1 - e_s^1 - e_q^1 p_q^1) - (c^2 - e_s^2 - e_q^2 p_q^1)] \Big\}, \end{aligned} \tag{B16}$$

where μ, λ, η, and ϕ correspond to the *Lagrange* multipliers associated with the revenue constraint, the type-2 IC-constraint, the condition implicitly defining the off-equilibrium quality signal effort chosen by a type-2 mimicker (the effort is defined by $\hat{e}_q^1$), and the type-1 IC-constraint. Reformulation of the maximization program associated with the optimal pooling allocation yields:

$$W^{pool}(\varepsilon, \delta) = \max\{(\theta^2 - \gamma^1 \varepsilon) h(e_s, e_q) - e_s - e_q p_q^1\}. \tag{B17}$$

Using the envelope theorem implies that

$$\text{(B18)} \qquad \frac{\partial W^{sep}(\varepsilon,\delta)}{\partial \delta} = \lambda^*[e_q^{2*} - \hat{e}_q^{1*}] > 0,$$

where the asterisk $(*)$ refers to the optimal allocation under the separating equilibrium, where $\lambda^* > 0$ denotes the multiplier associated with the type-2 binding IC constraint due to the max-min social welfare function, and where $e_q^{2*} > \hat{e}_q^{1*}$ denote the quality signal effort levels associated with type-2 and a mimicking type-2, respectively. Note that the strict inequality follows from the construction of the separating equilibrium (note that a separating equilibrium always exists, even if $\delta \to 0$, in which case $\hat{e}_q^1 = e_q^1 = e_s = 0$).

Clearly, $\frac{\partial W^{pool}(\varepsilon,\delta)}{\partial \delta} = 0$, by virtue of the max-min welfare function and as by construction both types choose the same bundle under a pooling allocation. By virtue of the signs of the derivatives, fixing ε, it follows that $W^{sep}(\varepsilon,\delta)$ and $W^{pool}(\varepsilon,\delta)$ intersect at most once. Let $\delta^*(\epsilon)$ denote the implicit solution to $W^{sep}(\varepsilon,\delta) = W^{pool}(\varepsilon,\delta)$ if it exists, and otherwise let $\delta^*(\epsilon) = 0$ if $W^{sep}(\varepsilon,\delta) > W^{pool}(\varepsilon,\delta)$ for all δ and $\delta^*(\epsilon) = p_q^1$ if $W^{sep}(\varepsilon,\delta) < W^{pool}(\varepsilon,\delta)$ for all δ, which completes the proof of part (b).

Part (c) Fixing p_q^1, θ^2, and γ^1, and denoting by ε and δ, as in the previous parts, the difference in productivity and the cost of acquiring the quality signal, respectively, we turn next to prove that there exists some cutoff, $0 < \varepsilon^* < \theta^2$, such that $\delta^*(\varepsilon) = 0$ for any $\varepsilon > \varepsilon^*$, while $\delta^*(\varepsilon) > 0$ for any $\varepsilon < \varepsilon^*$.

Consider the case where $\delta = 0$. Note that in this case the optimal separating equilibrium is given by the two triplets: $(c^1, e_s^1 = e_q^1 = 0)$ and (c^2, e_s^2, e_q^2), which maximize c^1 subject to:

$$\text{(B19)} \qquad c^2 - (e_s^2 + e_q^2 p_q^1) = c^1$$

$$\text{(B20)} \qquad (1-\gamma^1)\theta^2 h(e_s^2, e_q^2) = \gamma^1 c^1 + (1-\gamma^1)c^2$$

where the first equality condition (B19) denotes the binding incentive constraint (for both types!) and the second equality condition (B20) denotes the binding revenue constraint.

Since $\delta = 0$, the two types of agents are observationally equivalent, and thus for the separating allocation to be incentive compatible, the output produced by the low-skilled (type-1) agents must be zero. If output were bounded away from zero, the high-skilled (type-2) agents could mimic by choosing (off-equilibrium) a lower level of the quality signal than the level chosen (on the equilibrium path) by the type-1 agents. Then the two IC-constraints associated with the two types of workers could not hold simultaneously.

To see this formally, assume that the output level associated with the type-1 bundle is positive. The IC constraints associated with type-2 and type-1 would then be:

$$c^2 - (e_s^2 + e_q^2 p_q^1) \geq c^1 - (e_s^1 + \hat{e}_q^1 p_q^1) \tag{B21}$$

$$c^1 - (e_s^1 + e_q^1 p_q^1) \geq c^2 - (e_s^2 + e_q^2 p_q^1) \tag{B22}$$

where $e_q^1 > \hat{e}_q^1$ and $\hat{e}_q^1$ being the implicit solution to $(\theta^2 - \varepsilon)h(e_s^1, e_q^1) = (\theta^2 - \gamma^1\varepsilon)h(e_s^1, \hat{e}_q^1)$.

However, using the two IC-constraints implies that:

$$c^1 - (e_s^1 + e_q^1 p_q^1) \geq c^2 - (e_s^2 + e_q^2 p_q^1) \geq c^1 - (e_s^1 + \hat{e}_q^1 p_q^1) \tag{B23}$$

Hence,

$$c^1 - (e_s^1 + e_q^1 p_q^1) \geq c^1 - (e_s^1 + \hat{e}_q^1 p_q^1) \Leftrightarrow \hat{e}_q^1 p_q^1 \geq e_q^1 p_q^1 \tag{B24}$$

but this clearly contradicts $e_q^1 > \hat{e}_q^1$.

Using the two binding conditions (B19) and (B20) yields that the welfare level associated with the optimal separating equilibrium is given by:

$$W^{sep}(\epsilon, \delta = 0) = \max\{(1 - \gamma^1)[\theta^2 h(e_s^2, e_q^2) - (e_s^2 + e_q^2 p_q^1)]\} \tag{B25}$$

The optimal pooling equilibrium is given by:

$$W^{pool}(\varepsilon, \delta = 0) = \max\{(\theta^2 - \gamma^1\varepsilon)h(e_s, e_q) - (e_s + e_q p_q^1)\} \tag{B26}$$

Let $\Omega(\varepsilon, \delta = 0) \equiv W^{sep}(\epsilon, \delta = 0) - W^{pool}(\varepsilon, \delta = 0)$. It is easy to verify that $\Omega(0, \delta = 0) < 0$, and $\Omega(\theta^2, \delta = 0) > 0$. Thus, by continuity, using the Intermediate Value Theorem, there exists some $0 < \varepsilon^* < \theta^2$ such that $\Omega(\varepsilon^*, \delta = 0) = 0$.

Denoting by $e_s^*(\varepsilon)$ and $e_q^*(\varepsilon)$ the effort levels associated with the quantity and quality signals in the optimal pooling equilibrium when the productivity difference is ε, using the envelope theorem, it follows that

$$\frac{\partial \Omega(\varepsilon, \delta = 0)}{\partial \varepsilon} = \gamma^1 h[e_s^*(\varepsilon), e_q^*(\varepsilon)] > 0. \tag{B27}$$

It follows that for all $\varepsilon < \varepsilon^*$, $W^{sep}(\epsilon, \delta = 0) < W^{pool}(\varepsilon, \delta = 0)$, while for all $\varepsilon > \varepsilon^*$, $W^{sep}(\epsilon, \delta = 0) > W^{pool}(\varepsilon, \delta = 0)$. This completes the proof.

C Proof of Proposition 2

Let $\varepsilon \to \theta^2$ (with $\varepsilon < \theta^2$) and further let $\delta = 0$, hence $p_q^1 = p_q^2 = p_q$. Without loss of generality let $p_s = 1$. Our result will extend by continuity to sufficiently small values of

$\delta > 0$. As shown in the proof of part (c) of Proposition 1 [see Appendix B, Eqs. (B19) and (B20)], under the above parametric assumptions, the MMO is given by an STE in which the effort levels associated with type-1 (low-skilled) workers are given by $e_s^1 = e_q^1 = 0$. Formally, the MMO is given by the solution to the following maximization program:

P1

$$\max_{e_s^2, e_q^2, c^1, c^2} \{c^1\} \quad \text{subject to:} \tag{C1}$$

$$c^2 - e_s^2 - p_q e_q^2 = c^1, \tag{C2}$$

$$\gamma^2 \theta^2 h(e_s^2, e_q^2) = \gamma^1 c^1 + \gamma^2 c^2, \tag{C3}$$

where (C2) and (C3) replicate (B19) and (B20), representing the binding IC-constraint (associated with both type-1 and type-2 workers) and the binding revenue constraint, respectively.

Now consider the MMO associated with an STE under a "Mirrleesian" setup in which firms observe worker types (but the government doesn't):

P2

$$\max_{e_s^1, e_q^1, e_s^2, e_q^2, \hat{e}_q^1, c^1, c^2} \{c^1 - e_s^1 - p_q e_q^1\} \quad \text{subject to:} \tag{C4}$$

$$c^2 - e_s^2 - p_q e_q^2 = c^1 - e_s^1 - p_q \hat{e}_q^1, \tag{C5}$$

$$\gamma^1 \theta^1 h(e_s^1, e_q^1) + \gamma^2 \theta^2 h(e_s^2, e_q^2) = \gamma^1 c^1 + \gamma^2 c^2, \tag{C6}$$

$$\theta^1 h(e_s^1, e_q^1) = \theta^2 h(e_s^1, \hat{e}_q^1), \tag{C7}$$

where condition (C5) is the binding IC-constraint associated with type-2 workers (the IC-constraint associated with type-1 workers is slack due to a single-crossing property and is therefore omitted), and condition (C6) is the binding revenue constraint. The quality effort chosen by the type-2 mimicker is implicitly given by condition (C7), which states that type-2 receives the same compensation as type-1, $y^1 = \theta^1 h(e_s^1, e_q^1)$, and chooses the same quantity effort as type-1, e_s^1. However, type-2 agents choose a lower quality effort level than type-1 agents, $\hat{e}_q^1 \leq e_q^1$, with strict inequality when $e_s^1 > 0$ and $e_q^1 > 0$, and are compensated according to their true productivity, θ^2.[31]

Comparing the maximization programs **P1** and **P2**, one can see that problem **P1** is obtained by setting the effort levels associated with type-1 workers to zero in the

[31] Recall the difference from the case where types are unobservable by firms, in which a mimicking type-2 will choose a lower quality of effort than type-1 but higher than $\hat{e}_q^1$, and be rewarded according to average productivity rather than true productivity.

formulation of problem **P2**. Thus, the optimal solution to problem **P1** is a feasible solution (but not necessarily the optimal one) to problem **P2**. To show that the maximization program **P2** yields a higher level of welfare than the maximization program **P1**, we construct an alternative feasible allocation with strictly positive effort levels for type-1 workers and show that it increases their level of utility.

Formally, let the allocation ($e_s^{1*} = e_q^{1*} = 0$, e_s^{2*}, e_q^{2*}, c^{1*}, c^{2*}) denote the optimal solution for the program **P1**, and consider the following small perturbation of this allocation: $e_s^2 = e_s^{2*}, e_q^2 = e_q^{2*}, e_s^1 = e_q^1 = \sigma > 0$ where σ is small, $c^1 = c^{1*} + \gamma^1\theta^1 h(\sigma,\sigma)$, $c^2 = c^{2*} + \gamma^1\theta^1 h(\sigma,\sigma)$, and $\hat{e}_q^1$ is implicitly given by $\theta^1 h(\sigma,\sigma) = \theta^2 h(\sigma, \hat{e}_q^1)$. In other words, the effort vector of the type-1 worker is raised slightly above zero, and the resulting fiscal surplus is returned to both types of workers as a lump-sum transfer. The proposed perturbation leads to a relaxation of the type-2 worker's IC-constraint and satisfies the revenue constraint. By continuity, the perturbation maintains the slack in the type-1 worker's IC-constraint. Thus, the perturbed allocation is a feasible solution for program **P2**. The change in type-1 utility due to the proposed perturbation is given by

$$\Delta u^1 = \gamma^1\theta^1 h(\sigma,\sigma) - \sigma(1 + p_q). \tag{C8}$$

hence,

$$\Delta u^1 > 0 \Leftrightarrow \frac{h(\sigma,\sigma)}{\sigma} > \frac{(1+p_q)}{\gamma^1\theta^1} \tag{C9}$$

Since $\lim_{\sigma\to 0}\left[\frac{h(\sigma,\sigma)}{\sigma}\right] = \lim_{\sigma\to 0}[h_1(\sigma,\sigma) + h_2(\sigma,\sigma)] = \infty$ by the Inada conditions of the human capital production function, it follows that for σ sufficiently small, $\Delta u^1 > 0$. This concludes the proof.

D Supplementary analysis for Figure 1

D.1 Detailed discussion about the shape of the regions in Figure 1

The vertical axis $\delta = 0$ Along the vertical axis, where $\delta = 0$ (i.e., $p_q^1 = p_q^2$), pooling dominates separation for almost all values of ϵ. To understand this result, note that when $p_q^1 = p_q^2$, it must necessarily be the case that $y^1 = e_s^1 = 0$ in the STE.[32] The fact that type-

[32] The reason is as follows. When $p_q^1 = p_q^2$, the left side of the downward IC constraint (23) coincides with the right side of the upward IC constraint (24) (since $p_q^1 = p_q^2$ implies $R^2\left(e_s^2, e_q^2\right) = R^1\left(e_s^2, e_q^2\right)$). At the same time, however, the right-hand side of the downward IC constraint (23) is strictly larger than the left-hand side of the upward IC constraint (24) whenever type-1 agents are required to produce a

1 agents are idle in the STE is not a major concern when θ^1 is very small (corresponding in our Figure 1 to cases where ϵ is large): when their productivity is very small, the efficiency loss from leaving them idle is also very small. However, as θ^1 increases (i.e., as ϵ decreases in Figure 1), the efficiency loss associated with setting $y^1 = 0$ becomes larger and larger.

Small $\delta > 0$ Now consider the case where δ is small but strictly positive. By replicating the choices of type-2 agents and acting as mimickers, type-1 agents achieve a utility that is strictly lower than the utility of type-2 agents; however, given the assumption that δ is small, the difference between the two utilities is also small. Thus, in order to jointly satisfy the IC constraints (23)–(24), the information rent enjoyed by type-2 agents must also be small. This information rent, which reflects the difference between the utility of type-2 agents behaving as mimickers and the utility of type-1 agents, is given by $p_q^1 e_q^1 - p_q^2 \widehat{e}_q^2$ and can be decomposed into two components. One is due to the fact that $\delta > 0$, and the other is due to the fact that, for $y^1 > 0$, $\widehat{e}_q^2 < e_q^1$. The latter component reflects the information rent arising from productivity differences and, under the assumption given by (26), it is given by $\left[\left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}} - \left(\frac{y^1}{\overline{\theta}}\right)^{\frac{1}{\beta}}\right] \frac{1}{e_s^1}$, and is therefore increasing in y^1 (and increasing in the productivity difference between the two types).

Thus, if δ is small, an STE will be characterized by a value of y^1 that is positive but necessarily small. Whether this leads to a large or small efficiency loss depends on the productivity of type-1 agents. In particular, forcing y^1 to be very small is more costly the higher the productivity of type-1 agents (i.e. the lower is ϵ in Figure 1). This observation seems to suggest that when δ is small, a PTE dominates an STE, provided that the productivity of type-1 agents is sufficiently large (i.e., ϵ is sufficiently low). Looking at Figure 1, we can see that this intuition is only partially confirmed. In particular, we can see that for small values of δ, an STE dominates a PTE both when ϵ is sufficiently high and when it is sufficiently low. The fact that separating dominates when ϵ is sufficiently high is consistent with the intuition that follows from our argument.

What remains to be explained is why STE dominates PTE when ϵ is sufficiently small. When ϵ is small, θ^1 is close to θ^2, and this implies that in an STE, the information rent to type-2 agents (arising from the difference in productivities) is small. But this in turn means that the equity gains from moving from an STE to a PTE are also small. The reason is that these equity gains arise from the elimination of the information rent associated with the difference in productivities enjoyed by type-2 agents in an STE.

positive amount of output. Thus, for $p_q^1 = p_q^2$, the IC constraints (23)–(24) can be jointly satisfied only if $y^1 = 0$, implying that type-1 agents remain idle.

Large δ Figure 1 also shows that for sufficiently large values of δ, an STE is always superior to a PTE. Two things should be noted when interpreting this result. First, if δ is sufficiently large, the upward IC constraint (constraint (24)) is slack regardless of the value of ϵ, and therefore the effort exerted by type-2 agents is first-best optimal in an STE (i.e., it satisfies the equality $h_1\left(e_s^2, e_q^2\right)/h_2\left(e_s^2, e_q^2\right) = p_s/p_q^2$). Second, for a given value of ϵ, the PTE is invariant to changes in δ (in our Figure 1, a variation in δ corresponds to a variation in p_q^2, for given p_q^1); this implies that, as δ increases, the efficiency loss associated with pooling all agents at a common effort mix becomes larger. Thus, for sufficiently large values of δ, the efficiency losses of switching from a separating to a pooling equilibrium outweigh any possible equity gains.

Finally, two remarks are in order regarding the effect of changes in the relative size of the two groups of individuals and the effect of changes in the parameter β on the trade-off between STE and PTE. To save space, we do not provide the figures, but the simulations are available upon request.

The role of γ^1 A PTE tends to become more attractive when the size of the two groups of agents is more similar. This is because when γ^1 is very small, the efficiency cost of leaving type-1 agents idle becomes small, regardless of their productivity. Thus, even if IC considerations require that y^1 be set close to zero in an STE, the associated efficiency cost is negligible. At the other extreme, when γ^1 is very large, the difference between $\overline{\theta}$ and θ^1 also becomes small, implying that in an STE the information rent to type-2 agents, arising from the productivity difference, is quite small. This in turn implies that the equity gains of moving from an STE to a PTE are also small.

The role of β A PTE tends to become more attractive when β is small, i.e., when the degree of decreasing returns to scale characterizing the h function is large. Intuitively, note that the output production function is given by a product of θ, the innate productivity, and h, the acquired human capital, and exhibits overall increasing returns to scale. Consequently, there is an efficiency loss associated with pooling all agents in a common effort mix relative to a separating allocation. As β increases and the h function approaches constant returns to scale, the efficiency loss becomes more pronounced, making pooling less desirable.[33]

[33] The argument is similar to a study by Cremer et al. (2011), which shows that a meritocratic education system (a "separating" allocation with unequal wages) supplemented by a progressive labor income tax system would be preferable to an egalitarian education system (a "pooling" allocation with equal wages) from a redistributive perspective.

D.2 Closed-form solutions for social welfare and derivations for Figure 1

In the first part of this appendix we derive the pooling tax equilibrium and the associated social welfare value. In the second and third parts, we derive the optimal allocation under a separating tax equilibrium and the associated value of social welfare, first for the case when the upward IC constraint is not binding and then for the case when it is binding. Finally, in the last part of the appendix, we use our results to derive the inequalities characterizing the regions described in Figure 1. All results are based on the functional form assumption (26).

D.2.1 Pooling tax equilibrium

When implementing a pooling equilibrium, the government chooses (y, e_s) to maximize

$$u^1 = y - p_s e_s - p_q^1 e_q(y, e_s, \bar{\theta}), \tag{D1}$$

where $e_q(y, e_s, \bar{\theta})$ is the value of e_q which solves the equation $y = (e_s e_q)^\beta \bar{\theta}$, i.e. $\widehat{e}_q = \left(\frac{y}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{1}{e_s}$.

We can then rewrite the government's objective function as

$$U^{pool} = y - p_s e_s - \left(\frac{y}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^1}{e_s}, \tag{D2}$$

which has to be maximized by the optimal choice of y and e_s.

The first order conditions with respect to e_s and y are respectively given by

$$\left(\frac{y}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^1}{(e_s)^2} = p_s, \tag{D3}$$

$$\frac{p_q^1}{e_s} \frac{1}{\beta\bar{\theta}} \left(\frac{y}{\bar{\theta}}\right)^{\frac{1-\beta}{\beta}} = 1. \tag{D4}$$

Dividing (D3) by (D4) gives

$$y = \frac{p_s e_s}{\beta}. \tag{D5}$$

Noticing that (D3) can be restated as

$$\left(\frac{y}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^1}{e_s} = p_s e_s, \tag{D6}$$

it follows that, by using (D5) and (D6), we can re-express U^{pool}, given by (D2), as

$$U^{pool} = y - p_s e_s - \left(\frac{y}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^1}{e_s} = \frac{p_s e_s}{\beta} - p_s e_s - p_s e_s = \frac{1-2\beta}{\beta} p_s e_s. \tag{D7}$$

Finally, given that from (D3) we have

(D8) $$(e_s)^2 = \frac{p_q^1}{p_s}\left(\frac{y}{\bar{\theta}}\right)^{\frac{1}{\beta}},$$

using (D5) we get that

$$(e_s)^2 = \frac{p_q^1}{p_s}\left(\frac{p_s e_s}{\beta\bar{\theta}}\right)^{\frac{1}{\beta}},$$

which implies

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

and therefore

(D9) $$e_s = (p_s)^{\frac{1-\beta}{\beta}\frac{\beta}{2\beta-1}}\left(p_q^1\right)^{\frac{\beta}{2\beta-1}}\left(\frac{1}{\bar{\theta}\beta}\right)^{\frac{\beta}{2\beta-1}\frac{1}{\beta}} = \frac{\left(\bar{\theta}\beta\right)^{\frac{1}{1-2\beta}}}{(p_s)^{\frac{1-\beta}{1-2\beta}}\left(p_q^1\right)^{\frac{\beta}{1-2\beta}}}.$$

We can then conclude that

(D10) $$U^{pool} = \frac{1-2\beta}{\beta}p_s e_s = \frac{1-2\beta}{\beta}p_s\frac{(p_s)^{\frac{1-\beta}{2\beta-1}}}{\left(p_q^1\right)^{\frac{\beta}{1-2\beta}}}\left(\bar{\theta}\beta\right)^{\frac{1}{1-2\beta}} = \frac{1-2\beta}{\beta}\left[\frac{\bar{\theta}\beta}{\left(p_s p_q^1\right)^{\beta}}\right]^{\frac{1}{1-2\beta}}.$$

Since we have that $y = \frac{p_s e_s}{\beta}$, we can equivalently express U^{pool} as

(D11) $$U^{pool} = (1-2\beta)\,y,$$

where

(D12) $$y = \frac{\left(\bar{\theta}\right)^{\frac{1}{1-2\beta}}\beta^{\frac{2\beta}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}}.$$

D.2.2 Separating tax equilibrium when only the downward IC-constraint is binding

Consider now the separating tax equilibrium. The upward and downward IC-constraints are given, respectively, by:

(D13) $$c^1 - p_s e_s^1 - \left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}}\frac{p_q^1}{e_s^1} \geq c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}}\frac{p_q^1}{e_s^2},$$

(D14) $$c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}}\frac{p_q^2}{e_s^2} \geq c^1 - p_s e_s^1 - \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}}\frac{p_q^2}{e_s^1}.$$

We first proceed to solve the government's problem assuming that the upward IC constraint can be neglected. Substituting the resource constraint of the economy into the government's objective function, we can rewrite the government's problem as follows:

$$\max_{y^2,y^1,c^2,e_s^1,e_s^2} \frac{\gamma^2}{\gamma^1}\left(y^2-c^2\right)+y^1-p_se_s^1-p_q^1\left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}}\frac{1}{e_s^1}$$

subject to the downward IC-constraint

$$c^2-\frac{\gamma^2}{\gamma^1}\left(y^2-c^2\right)-y^1-p_se_s^2-\left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}}\frac{p_q^2}{e_s^2}+p_se_s^1+\left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}}\frac{p_q^2}{e_s^1}\geq 0. \tag{D15}$$

Denote by λ the multiplier attached to the IC-constraint (D15). From the first order condition with respect to c^2 we have that $\lambda=\gamma^2$. Taking this into account, the first order conditions with respect to e_s^1, e_s^2, y^1 and y^2 are, respectively:

$$-p_s+\left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}}\frac{p_q^1}{\left(e_s^1\right)^2}+\gamma^2p_s-\gamma^2\left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}}\frac{p_q^2}{\left(e_s^1\right)^2}=0, \tag{D16}$$

$$\left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}}\frac{p_q^2}{\left(e_s^2\right)^2}=p_s, \tag{D17}$$

$$1-\left(y^1\right)^{\frac{1-\beta}{\beta}}\frac{p_q^1}{e_s^1}\frac{1}{\left(\theta^1\right)^{\frac{1}{\beta}}}\frac{1}{\beta}-\gamma^2+\gamma^2\left(y^1\right)^{\frac{1-\beta}{\beta}}\frac{p_q^2}{e_s^1}\frac{1}{\left(\bar{\theta}\right)^{\frac{1}{\beta}}}\frac{1}{\beta}=0, \tag{D18}$$

$$\frac{\gamma^2}{\gamma^1}-\gamma^2\frac{\gamma^2}{\gamma^1}-\gamma^2\left(y^2\right)^{\frac{1-\beta}{\beta}}\frac{p_q^2}{e_s^2}\frac{1}{\left(\theta^2\right)^{\frac{1}{\beta}}}\frac{1}{\beta}=0. \tag{D19}$$

Rewrite (D19) as

$$\left(y^2\right)^{\frac{1-\beta}{\beta}}\frac{p_q^2}{e_s^2}\frac{1}{\left(\theta^2\right)^{\frac{1}{\beta}}}\frac{1}{\beta}=1. \tag{D20}$$

Dividing (D17) by (D20) gives

$$y^2=\frac{p_se_s^2}{\beta}, \tag{D21}$$

from which, substituting in (D17), we obtain

$$e_s^2=\frac{\left(\theta^2\beta\right)^{\frac{1}{1-2\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}}\left(p_q^2\right)^{\frac{\beta}{1-2\beta}}}, \tag{D22}$$

and therefore

$$\text{(D23)} \qquad y^2 = \frac{p_s e_s^2}{\beta} = \frac{p_s}{\beta} \frac{\left(\theta^2 \beta\right)^{\frac{1}{1-2\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}} \left(p_q^2\right)^{\frac{\beta}{1-2\beta}}} = \theta^2 \left[\frac{\theta^2 \beta}{\left(p_s p_q^2\right)^{\frac{1}{2}}} \right]^{\frac{2\beta}{1-2\beta}}.$$

Now rewrite (D16) and (D18) as:

$$\text{(D24)} \qquad \left[\frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}} - \frac{\gamma^2 p_q^2}{\left(\bar{\theta}\right)^{\frac{1}{\beta}}} \right] \left(y^1\right)^{\frac{1}{\beta}} = p_s \left(e_s^1\right)^2 - \gamma^2 p_s \left(e_s^1\right)^2,$$

$$\text{(D25)} \qquad \left[\frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}} - \frac{\gamma^2 p_q^2}{\left(\bar{\theta}\right)^{\frac{1}{\beta}}} \right] \left(y^1\right)^{\frac{1-\beta}{\beta}} = \beta e_s^1 - \gamma^2 \beta e_s^1.$$

Dividing (D24) by (D25) gives

$$\text{(D26)} \qquad y^1 = \frac{p_s e_s^1}{\beta}.$$

Substituting in (D24) the value for y^1 provided by (D26) gives

$$\left[\frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}} - \frac{\gamma^2 p_q^2}{\left(\bar{\theta}\right)^{\frac{1}{\beta}}} \right] \left(\frac{p_s e_s^1}{\beta} \right)^{\frac{1}{\beta}} = \gamma^1 p_s \left(e_s^1\right)^2,$$

and therefore

$$\left(e_s^1\right)^{\frac{1-2\beta}{\beta}} = \frac{\gamma^1 \left(\beta\right)^{\frac{1}{\beta}} \left(p_s\right)^{\frac{\beta-1}{\beta}}}{p_q^1 \left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} - \gamma^2 p_q^2 \left(\frac{1}{\bar{\theta}}\right)^{\frac{1}{\beta}}} = \frac{\left(\beta\right)^{\frac{1}{\beta}} \left(p_s\right)^{\frac{\beta-1}{\beta}}}{p_q^1 \left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} + \frac{\gamma^2}{1-\gamma^2} \left[p_q^1 \left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} - p_q^2 \left(\frac{1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \right]},$$

i.e.

$$\text{(D27)} \qquad e_s^1 = \frac{\left(\gamma^1\right)^{\frac{\beta}{1-2\beta}} \left(\beta\right)^{\frac{1}{1-2\beta}} \left(p_s\right)^{\frac{\beta-1}{1-2\beta}}}{\left[p_q^1 \left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} - \gamma^2 p_q^2 \left(\frac{1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \right]^{\frac{\beta}{1-2\beta}}} = \frac{\left(\theta^1 \beta\right)^{\frac{1}{1-2\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{\beta}{1-2\beta}} \left[\frac{1}{\gamma^1} - \frac{\gamma^2}{\gamma^1} \frac{p_q^2}{p_q^1} \left(\frac{\theta^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \right]^{\frac{\beta}{1-2\beta}}},$$

and therefore

$$y^1 = \frac{p_s e_s^1}{\beta} = \frac{\left(\gamma^1\right)^{\frac{\beta}{1-2\beta}} (\beta)^{\frac{2\beta}{1-2\beta}}}{\left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{\gamma^2 p_q^2}{(\bar{\theta})^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}} \frac{1}{(p_s)^{\frac{\beta}{1-2\beta}}} = \frac{(\beta)^{\frac{2\beta}{1-2\beta}}}{\left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{\gamma^2 p_q^2}{(\bar{\theta})^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}} \left(\frac{\gamma^1}{p_s}\right)^{\frac{\beta}{1-2\beta}}$$

$$= \frac{\left(\theta^1\right)^{\frac{1}{1-2\beta}} (\beta)^{\frac{2\beta}{1-2\beta}}}{\left\{p_s p_q^1 + \frac{\gamma^2}{1-\gamma^2} p_s \left(\theta^1\right)^{\frac{1}{\beta}} \left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{p_q^2}{(\bar{\theta})^{\frac{1}{\beta}}}\right]\right\}^{\frac{\beta}{1-2\beta}}}. \tag{D28}$$

Combining (D22) and (D27) gives

$$\frac{e_s^1}{e_s^2} = \frac{\dfrac{\left(\gamma^1\right)^{\frac{\beta}{1-2\beta}} (\beta)^{\frac{1}{1-2\beta}} (p_s)^{\frac{\beta-1}{1-2\beta}}}{\left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{\gamma^2 p_q^2}{(\bar{\theta})^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}}}{\dfrac{\left(\theta^2 \beta\right)^{\frac{1}{1-2\beta}}}{(p_s)^{\frac{1-\beta}{1-2\beta}} \left(p_q^2\right)^{\frac{\beta}{1-2\beta}}}} = \frac{\left(\gamma^1\right)^{\frac{\beta}{1-2\beta}} (\beta)^{\frac{1}{1-2\beta}} (p_s)^{\frac{\beta-1}{1-2\beta}}}{\left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{\gamma^2 p_q^2}{(\bar{\theta})^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}} \frac{(p_s)^{\frac{1-\beta}{1-2\beta}} \left(p_q^2\right)^{\frac{\beta}{1-2\beta}}}{\left(\theta^2 \beta\right)^{\frac{1}{1-2\beta}}}$$

$$= \left[\frac{\gamma^1 p_q^2}{p_q^1 \left(\frac{\theta^2}{\theta^1}\right)^{\frac{1}{\beta}} - \gamma^2 p_q^2 \left(\frac{\theta^2}{\bar{\theta}}\right)^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}. \tag{D29}$$

Consider now the difference $y^1 - p_s e_s^1 - p_q^1 e_q^1$. Notice that, by using (D24), we can write

$$p_q^1 e_q^1 = \left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}} \frac{p_q^1}{e_s^1} = p_s e_s^1 - \left[p_s - \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^2}{(e_s^1)^2}\right] \gamma^2 e_s^1$$

$$= \gamma^1 p_s e_s^1 + \gamma^2 \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^1}. \tag{D30}$$

Therefore, using (D26) and (D30), we can express $y^1 - p_s e_s^1 - p_q^1 e_q^1$ as

$$y^1 - p_s e_s^1 - p_q^1 e_q^1 = \frac{p_s e_s^1}{\beta} - p_s e_s^1 - \gamma^1 p_s e_s^1 - \gamma^2 \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^1}$$

$$= \frac{1 - \beta - \gamma^1 \beta}{\beta} p_s e_s^1 - \gamma^2 p_q^2 \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{1}{e_s^1}$$

$$= \frac{1 - 2\beta}{\beta} p_s e_s^1 + \gamma^2 \left[p_s e_s^1 - p_q^2 \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{1}{e_s^1}\right]$$

$$= \frac{1 - 2\beta}{\beta} p_s e_s^1 + \gamma^2 \left[p_s e_s^1 - p_q^2 \left(\frac{p_s}{\beta \bar{\theta}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right]. \tag{D31}$$

Given that the downward IC-constraint (D15) must be binding at the separating tax equilibrium (this follows from our assumption that the social welfare function is of the max-min type), we have that

$$\begin{aligned}
\frac{c^2}{\gamma^1} &= y^1 - p_s e_s^1 - \left(\frac{y^1}{\overline{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^1} + \frac{\gamma^2}{\gamma^1} y^2 + p_s e_s^2 + \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \\
&= \frac{p_s e_s^1}{\beta} - p_s e_s^1 + \frac{\gamma^2}{\gamma^1}\frac{p_s e_s^2}{\beta} + p_s e_s^2 + p_q^2 \left[\left(\frac{p_s e_s^2}{\beta\theta^2}\right)^{\frac{1}{\beta}} \frac{1}{e_s^2} - \left(\frac{p_s e_s^1}{\beta\overline{\theta}}\right)^{\frac{1}{\beta}} \frac{1}{e_s^1}\right] \\
&= \frac{p_s e_s^1}{\beta} - p_s e_s^1 + \frac{\gamma^2}{\gamma^1}\frac{p_s e_s^2}{\beta} + p_s e_s^2 + p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right] \\
&= \frac{1-\beta}{\beta} p_s e_s^1 + \frac{\gamma^2}{\gamma^1}\frac{p_s e_s^2}{\beta} + p_s e_s^2 + p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right],
\end{aligned}$$

and therefore

$$\begin{aligned}
c^2 &= \frac{1-\beta}{\beta}\gamma^1 p_s e_s^1 + \gamma^2 \frac{p_s e_s^2}{\beta} + \gamma^1 p_s e_s^2 + \gamma^1 p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right] \\
&= \left[(1-\beta)\gamma^1 e_s^1 + \gamma^2 e_s^2 + \gamma^1 \beta e_s^2\right] \frac{p_s}{\beta} + \gamma^1 p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right] \\
&= \left(\gamma^1 e_s^1 + \gamma^2 e_s^2\right) \frac{p_s}{\beta} + \left(e_s^2 - e_s^1\right)\gamma^1 p_s + \gamma^1 p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right].
\end{aligned}$$

It then follows that the transfer provided to each type-1 agent, $\frac{\gamma^2}{\gamma^1}\left(y^2 - c^2\right)$, is given by

$$\begin{aligned}
& \frac{\gamma^2}{\gamma^1}\left(y^2 - c^2\right) \\
&= \frac{\gamma^2}{\gamma^1}\frac{p_s e_s^2}{\beta} - \frac{\gamma^2}{\gamma^1}\left(\gamma^1 e_s^1 + \gamma^2 e_s^2\right)\frac{p_s}{\beta} - \frac{\gamma^2}{\gamma^1}\left(e_s^2 - e_s^1\right)\gamma^1 p_s \\
&\quad -\frac{\gamma^2}{\gamma^1}\gamma^1 p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right] \\
&= -\gamma^2\left(e_s^1 - e_s^2\right)\frac{p_s}{\beta} - \left(e_s^2 - e_s^1\right)\gamma^2 p_s \\
&\quad -\gamma^2 p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right] \\
&= \frac{1-\beta}{\beta}\left(e_s^2 - e_s^1\right)\gamma^2 p_s \\
&\quad -\gamma^2 p_q^2 \left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}} \left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} - \left(\frac{1}{\overline{\overline{\theta}}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right].
\end{aligned} \tag{D32}$$

We have now all the ingredients to determine the value of the government's objective function, i.e. the utility of type-1 agents, under a separating tax equilibrium. Using

(D31) and (D32), the government's objective function is given by

$$\begin{aligned}
U^{sep} &= \frac{\gamma^2}{\gamma^1}\left(y^2-c^2\right)+y^1-p_s e_s^1-p_q^1 e_q^1 \\
&= \frac{1-\beta}{\beta}\left(e_s^2-e_s^1\right)\gamma^2 p_s-\gamma^2 p_q^2\left(\frac{p_s}{\beta}\right)^{\frac{1}{\beta}}\left[\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}}\left(e_s^2\right)^{\frac{1-\beta}{\beta}}-\left(\frac{1}{\bar{\bar{\theta}}}\right)^{\frac{1}{\beta}}\left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right] \\
&\quad +\frac{1-2\beta}{\beta}p_s e_s^1+\gamma^2\left[p_s e_s^1-p_q^2\left(\frac{p_s}{\beta\bar{\theta}}\right)^{1/\beta}\left(e_s^1\right)^{\frac{1-\beta}{\beta}}\right] \\
&= \frac{1-\beta}{\beta}\left(e_s^2-e_s^1\right)\gamma^2 p_s+\frac{1-2\beta}{\beta}p_s e_s^1+\gamma^2 p_s e_s^1-\gamma^2 p_q^2\left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}}\left(e_s^2\right)^{\frac{1-\beta}{\beta}} \\
&= \frac{1}{\beta}\left[\gamma^2 e_s^2-\gamma^2 e_s^1-\gamma^2\beta e_s^2+\gamma^2\beta e_s^1+\beta\gamma^2 e_s^1+e_s^1-2\beta e_s^1\right]p_s-\gamma^2 p_q^2\left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}}\left(e_s^2\right)^{\frac{1-\beta}{\beta}} \\
&= \frac{1}{\beta}\left[(1-\beta)\gamma^2 e_s^2+(1-2\beta)\gamma^1 e_s^1\right]p_s-\gamma^2 p_q^2\left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}}\left(e_s^2\right)^{\frac{1-\beta}{\beta}}.
\end{aligned}$$

Given that from (D22) we have

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

it follows that

$$\gamma^2 p_q^2\left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}}\left(e_s^2\right)^{\frac{1-\beta}{\beta}}=\gamma^2 p_q^2\left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}}\frac{\left(\theta^2\beta\right)^{\frac{1}{1-2\beta}\frac{1-\beta}{\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}\frac{1-\beta}{\beta}}\left(p_q^2\right)^{\frac{1-\beta}{1-2\beta}}}=\gamma^2\left(p_s p_q^2\right)^{\frac{\beta}{2\beta-1}}\left(\beta\theta^2\right)^{\frac{1}{1-2\beta}}.$$

Therefore, we can rewrite U^{sep} as

$$U^{sep}=\frac{1}{\beta}\left[(1-\beta)\gamma^2 e_s^2+(1-2\beta)\gamma^1 e_s^1\right]p_s-\gamma^2\left(p_s p_q^2\right)^{\frac{\beta}{2\beta-1}}\left(\beta\theta^2\right)^{\frac{1}{1-2\beta}},$$

i.e., exploiting (D22),

(D33) $$U^{sep}=\frac{1}{\beta}\left[(1-\beta)\gamma^2 e_s^2+(1-2\beta)\gamma^1 e_s^1\right]p_s-\gamma^2 p_s e_s^2=\frac{1-2\beta}{\beta}\left(\gamma^1 e_s^1+\gamma^2 e_s^2\right)p_s,$$

or, equivalently, since $y^1=\frac{p_s e_s^1}{\beta}$ and $y^2=\frac{p_s e_s^2}{\beta}$,

(D34) $$U^{sep}=(1-2\beta)\left(\gamma^1 y^1+\gamma^2 y^2\right),$$

where y^1 is provided by (D28) and y^2 is provided by (D23).

The above analysis was performed under the assumption that the upward IC-constraint could be neglected. We can now proceed to verify under which condition this assumption is justified. The upward IC-constraint is satisfied if the following condition holds:

$$c^1 - p_s e_s^1 - p_q^1 \left(y^1\right)^{\frac{1}{\beta}} \frac{1}{(\theta^1)^{\frac{1}{\beta}}} \frac{1}{e_s^1} \geq c^2 - p_s e_s^2 - p_q^1 \left(y^2\right)^{\frac{1}{\beta}} \frac{1}{(\theta^2)^{\frac{1}{\beta}}} \frac{1}{e_s^2}, \tag{D35}$$

or equivalently, exploiting the fact that we know the downward IC-constraint is necessarily binding,

$$\left(p_q^1 - p_q^2\right) e_q^2 \geq p_q^1 e_q^1 - p_q^2 \widehat{e}_q^2,$$

where the RHS of the inequality represents the difference between the utility of a type-1 as a non-mimicker and the utility of a type-2 behaving as a mimicker, and the LHS represents the difference between the utility of a type-2 as a non-mimicker and the utility of a type-1 as a mimicker.

According to the binding version of the downward IC-constraint, we have that

$$c^2 - p_s e_s^2 - p_q^2 \left(y^2\right)^{\frac{1}{\beta}} \frac{1}{(\theta^2)^{\frac{1}{\beta}}} \frac{1}{e_s^2} = c^1 - p_s e_s^1 - p_q^2 \left(y^1\right)^{\frac{1}{\beta}} \frac{1}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} \frac{1}{e_s^1},$$

i.e.,

$$c^2 - p_s e_s^2 - p_q^1 \left(y^2\right)^{\frac{1}{\beta}} \frac{1}{(\theta^2)^{\frac{1}{\beta}}} \frac{1}{e_s^2} = c^1 - p_s e_s^1 - p_q^2 \left(y^1\right)^{\frac{1}{\beta}} \frac{1}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} \frac{1}{e_s^1} - \left(p_q^1 - p_q^2\right) \left(y^2\right)^{\frac{1}{\beta}} \frac{1}{(\theta^2)^{\frac{1}{\beta}}} \frac{1}{e_s^2}. \tag{D36}$$

Replace in the RHS of (D35) the RHS of (D36). We get:

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

i.e.,

$$-p_q^1 \left(y^1\right)^{\frac{1}{\beta}} \frac{1}{(\theta^1)^{\frac{1}{\beta}}} \frac{1}{e_s^1} \geq -p_q^2 \left(y^1\right)^{\frac{1}{\beta}} \frac{1}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} \frac{1}{e_s^1} - \left(p_q^1 - p_q^2\right) \left(y^2\right)^{\frac{1}{\beta}} \frac{1}{(\theta^2)^{\frac{1}{\beta}}} \frac{1}{e_s^2},$$

i.e.,

$$\left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{p_q^2}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}\right] \frac{(y^1)^{\frac{1}{\beta}}}{e_s^1} \leq \frac{p_q^1 - p_q^2}{(\theta^2)^{\frac{1}{\beta}}} \frac{(y^2)^{\frac{1}{\beta}}}{e_s^2}. \tag{D37}$$

The LHS of (D37) captures the information rent that has to be paid to type-2 agents to deter them from mimicking. The RHS captures the difference between the utility of a type-2 agent and that of a type-1 agent behaving as a mimicker. Intuitively, one can safely disregard the upward IC-constraint if enough redistribution can be performed towards type-1 agents, i.e. if the information rent that accrues to type-2 agents is sufficiently small.

Noticing that (from (D21) and (D26))

$$\begin{aligned}\frac{(y^1)^{\frac{1}{\beta}}}{e_s^1} &= \frac{(p_s)^{\frac{1}{\beta}}(e_s^1)^{\frac{1}{\beta}}}{\beta^{\frac{1}{\beta}}}\frac{1}{e_s^1} = \frac{(p_s)^{\frac{1}{\beta}}(e_s^1)^{\frac{1-\beta}{\beta}}}{\beta^{\frac{1}{\beta}}} = \frac{p_s\,(y^1)^{\frac{1-\beta}{\beta}}}{\beta},\\ \frac{(y^2)^{\frac{1}{\beta}}}{e_s^2} &= \frac{(p_s)^{\frac{1}{\beta}}(e_s^2)^{\frac{1}{\beta}}}{\beta^{\frac{1}{\beta}}}\frac{1}{e_s^2} = \frac{(p_s)^{\frac{1}{\beta}}(e_s^2)^{\frac{1-\beta}{\beta}}}{\beta^{\frac{1}{\beta}}} = \frac{p_s\,(y^2)^{\frac{1-\beta}{\beta}}}{\beta},\end{aligned}$$

it follows that

$$\frac{\frac{(y^1)^{\frac{1}{\beta}}}{e_s^1}}{\frac{(y^2)^{\frac{1}{\beta}}}{e_s^2}} = \frac{(y^1)^{\frac{1-\beta}{\beta}}}{(y^2)^{\frac{1-\beta}{\beta}}},$$

and therefore, since

$$\begin{aligned}\frac{(y^1)^{\frac{1}{\beta}}}{e_s^1} &= \frac{(p_s)^{\frac{1}{\beta}}(e_s^1)^{\frac{1-\beta}{\beta}}}{\beta^{\frac{1}{\beta}}},\\ \frac{(y^2)^{\frac{1}{\beta}}}{e_s^2} &= \frac{(p_s)^{\frac{1}{\beta}}(e_s^2)^{\frac{1-\beta}{\beta}}}{\beta^{\frac{1}{\beta}}},\end{aligned}$$

we can rewrite condition (D37) as

(D38) $$\left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{p_q^2}{(\overline{\theta})^{\frac{1}{\beta}}}\right](e_s^1)^{\frac{1-\beta}{\beta}} \le \frac{p_q^1 - p_q^2}{(\theta^2)^{\frac{1}{\beta}}}(e_s^2)^{\frac{1-\beta}{\beta}}.$$

Finally, exploiting (D29) we can rewrite (D38) as

(D39) $$\frac{p_q^1 - p_q^2}{(\theta^2)^{\frac{1}{\beta}}} \ge \left[\frac{\gamma^1 p_q^2}{p_q^1\left(\frac{\theta^2}{\theta^1}\right)^{\frac{1}{\beta}} - \gamma^2 p_q^2\left(\frac{\theta^2}{\overline{\theta}}\right)^{\frac{1}{\beta}}}\right]^{\frac{1-\beta}{1-2\beta}}\left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{p_q^2}{(\overline{\theta})^{\frac{1}{\beta}}}\right].$$

D.2.3 Separating tax equilibrium when both IC-constraints are binding

Suppose now that (D39) is violated so that both IC-constraints are binding at a separating tax equilibrium. In this case the government maximizes

$$\max_{y^2,y^1,c^2,e_s^1,e_s^2} \frac{\gamma^2}{\gamma^1}\left(y^2 - c^2\right) + y^1 - p_s e_s^1 - p_q^1\left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}}\frac{1}{e_s^1}$$

subject to the binding version of the downward IC-constraint (D15)

(D40) $$c^2 - \frac{\gamma^2}{\gamma^1}\left(y^2 - c^2\right) - y^1 - p_s e_s^2 - p_q^2\left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}}\frac{1}{e_s^2} + p_s e_s^1 + p_q^2\left(\frac{y^1}{\overline{\theta}}\right)^{\frac{1}{\beta}}\frac{1}{e_s^1} = 0$$

and the upward IC-constraint

$$\text{(D41)} \qquad \frac{p_q^1 - p_q^2}{(\theta^2)^{\frac{1}{\beta}}} \frac{(y^2)^{\frac{1}{\beta}}}{e_s^2} = \left[\frac{p_q^1}{(\theta^1)^{\frac{1}{\beta}}} - \frac{p_q^2}{(\overline{\theta})^{\frac{1}{\beta}}} \right] \frac{(y^1)^{\frac{1}{\beta}}}{e_s^1}.$$

Rewriting (D40) we have that

$$\text{(D42)} \quad -\gamma^2 \frac{\gamma^2}{\gamma^1} y^2 - \gamma^2 y^1 - \gamma^2 p_s e_s^2 - \gamma^2 p_q^2 \left(\frac{y^2}{\theta^2} \right)^{\frac{1}{\beta}} \frac{1}{e_s^2} + \gamma^2 p_s e_s^1 + \gamma^2 p_q^2 \left(\frac{y^1}{\overline{\theta}} \right)^{\frac{1}{\beta}} \frac{1}{e_s^1} = -\frac{\gamma^2}{\gamma^1} c^2$$

Therefore, we can restate the government's problem by eliminating c^2 from the objective function (exploiting (D42)) and write

$$\max_{y^2, y^1, e_s^1, e_s^2} \frac{\gamma^2}{\gamma^1} y^2 + y^1 - p_s e_s^1 - p_q^1 \left(\frac{y^1}{\theta^1} \right)^{\frac{1}{\beta}} \frac{1}{e_s^1} - \gamma^2 \frac{\gamma^2}{\gamma^1} y^2 - \gamma^2 y^1 - \gamma^2 p_s e_s^2$$
$$-\gamma^2 p_q^2 \left(\frac{y^2}{\theta^2} \right)^{\frac{1}{\beta}} \frac{1}{e_s^2} + \gamma^2 p_s e_s^1 + \gamma^2 p_q^2 \left(\frac{y^1}{\overline{\theta}} \right)^{\frac{1}{\beta}} \frac{1}{e_s^1}$$

subject to the upward IC-constraint (D41).

Collecting terms in the objective function we can reformulate the government's problem as

$$\max_{y^2, y^1, e_s^1, e_s^2} \gamma^2 y^2 + \gamma^1 y^1 - \gamma^1 p_s e_s^1 - \gamma^2 p_s e_s^2 - p_q^1 \left(\frac{y^1}{\theta^1} \right)^{\frac{1}{\beta}} \frac{1}{e_s^1} - \gamma^2 p_q^2 \left(\frac{y^2}{\theta^2} \right)^{\frac{1}{\beta}} \frac{1}{e_s^2} + \gamma^2 p_q^2 \left(\frac{y^1}{\overline{\theta}} \right)^{\frac{1}{\beta}} \frac{1}{e_s^1}$$

subject to upward IC constraint (D41).

Finally, noticing that the constraint (D41) can be equivalently restated as

$$\text{(D43)} \qquad \frac{1}{e_s^2} = \frac{(\theta^2)^{\frac{1}{\beta}} (y^1)^{\frac{1}{\beta}} \left[p_q^1 (\overline{\theta})^{\frac{1}{\beta}} - p_q^2 (\theta^1)^{\frac{1}{\beta}} \right]}{\left(p_q^1 - p_q^2 \right) (y^2)^{\frac{1}{\beta}} (\theta^1)^{\frac{1}{\beta}} (\overline{\theta})^{\frac{1}{\beta}} e_s^1},$$

we can exploit (D43) to eliminate e_s^2 from the variables entering the objective function. The government's problem can then be restated as

$$\max_{y^2, y^1, e_s^1} \gamma^2 y^2 + \gamma^1 y^1 - \gamma^1 p_s e_s^1 - \gamma^2 p_s \frac{p_q^1 - p_q^2}{(\theta^2)^{\frac{1}{\beta}}} \frac{(y^2)^{\frac{1}{\beta}}}{(y^1)^{\frac{1}{\beta}}} \frac{(\theta^1 \overline{\theta})^{\frac{1}{\beta}}}{p_q^1 (\overline{\theta})^{\frac{1}{\beta}} - p_q^2 (\theta^1)^{\frac{1}{\beta}}} e_s^1$$
$$-p_q^1 \left(\frac{y^1}{\theta^1} \right)^{\frac{1}{\beta}} \frac{1}{e_s^1} - \gamma^2 p_q^2 \left(\frac{y^2}{\theta^2} \right)^{\frac{1}{\beta}} \frac{(y^1 \theta^2)^{\frac{1}{\beta}} \left[p_q^1 (\overline{\theta})^{\frac{1}{\beta}} - p_q^2 (\theta^1)^{\frac{1}{\beta}} \right]}{\left(p_q^1 - p_q^2 \right) \left(y^2 \theta^1 \overline{\theta} \right)^{\frac{1}{\beta}} e_s^1} + \gamma^2 p_q^2 \left(\frac{y^1}{\overline{\theta}} \right)^{\frac{1}{\beta}} \frac{1}{e_s^1},$$

or equivalently,

$$\max_{y^2,y^1,e_s^1} \gamma^2 y^2+\gamma^1 y^1-\gamma^1 p_s e_s^1-\gamma^2\frac{p_q^1-p_q^2}{(\theta^2)^{\frac{1}{\beta}}}\frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}\frac{(y^2)^{\frac{1}{\beta}}}{(y^1)^{\frac{1}{\beta}}}p_s e_s^1$$
$$+\left[\frac{\gamma^2 p_q^1 p_q^2}{p_q^1-p_q^2}\frac{(\theta^1)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}-p_q^1\right]\frac{(y^1)^{\frac{1}{\beta}}}{(\theta^1)^{\frac{1}{\beta}}e_s^1}.$$

The first order condition with respect to y^2, y^1 and e_s^1 are respectively given by:

(D44)
$$\left(y^1\right)^{\frac{1}{\beta}}=\frac{1}{\beta}\frac{p_q^1-p_q^2}{(\theta^2)^{\frac{1}{\beta}}}\frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}\left(y^2\right)^{\frac{1-\beta}{\beta}}p_s e_s^1,$$

(D45)
$$\gamma^1+\frac{1}{\beta}\gamma^2\frac{p_q^1-p_q^2}{(\theta^2)^{\frac{1}{\beta}}}\frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}\frac{(y^2)^{\frac{1}{\beta}}}{(y^1)^{\frac{1+\beta}{\beta}}}p_s e_s^1+\frac{1}{\beta}\left[\frac{\gamma^2 p_q^1 p_q^2}{p_q^1-p_q^2}\frac{(\theta^1)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}-p_q^1\right]\frac{(y^1)^{\frac{1-\beta}{\beta}}}{(\theta^1)^{\frac{1}{\beta}}e_s^1}=0,$$

(D46)
$$-\gamma^1 p_s-\gamma^2\frac{p_q^1-p_q^2}{(\theta^2)^{\frac{1}{\beta}}}\frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}\frac{(y^2)^{\frac{1}{\beta}}}{(y^1)^{\frac{1}{\beta}}}p_s=\left[\frac{\gamma^2 p_q^1 p_q^2}{p_q^1-p_q^2}\frac{(\theta^1)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}-p_q^1\right]\frac{(y^1)^{\frac{1}{\beta}}}{(\theta^1)^{\frac{1}{\beta}}(e_s^1)^2}.$$

We can rewrite (D45)-(D46), respectively, as

$$\left\{\left[\frac{\gamma^2 p_q^1 p_q^2}{p_q^1-p_q^2}\frac{(\theta^1)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}-p_q^1\right]\frac{1}{(\theta^1)^{\frac{1}{\beta}}}+\frac{p_q^1-p_q^2}{(\theta^2)^{\frac{1}{\beta}}}\frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}(e_s^1)^2\gamma^2 p_s}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}p_q^1-(\theta^1)^{\frac{1}{\beta}}p_q^2}\frac{(y^2)^{\frac{1}{\beta}}}{(y^1)^{\frac{2}{\beta}}}\right\}\frac{(y^1)^{\frac{1-\beta}{\beta}}}{e_s^1}=-\beta\gamma^1,$$
(D47)
$$\left\{\left[\frac{\gamma^2 p_q^1 p_q^2}{p_q^1-p_q^2}\frac{(\theta^1)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}-p_q^1\right]\frac{1}{(\theta^1)^{\frac{1}{\beta}}}+\frac{p_q^1-p_q^2}{(\theta^2)^{\frac{1}{\beta}}}\frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}(e_s^1)^2\gamma^2 p_s}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}p_q^1-(\theta^1)^{\frac{1}{\beta}}p_q^2}\frac{(y^2)^{\frac{1}{\beta}}}{(y^1)^{\frac{2}{\beta}}}\right\}\frac{(y^1)^{\frac{1}{\beta}}}{(e_s^1)^2}=-p_s\gamma^1.$$
(D48)

from which one obtains that

(D49)
$$y^1=\frac{p_s e_s^1}{\beta}.$$

Using (D49), from the first order condition (D44) we get

$$\left(\frac{p_s e_s^1}{\beta}\right)^{\frac{1}{\beta}}=\frac{1}{\beta}\frac{p_q^1-p_q^2}{(\theta^2)^{\frac{1}{\beta}}}\frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}\left(y^2\right)^{\frac{1-\beta}{\beta}}p_s e_s^1,$$

and therefore

$$\text{(D50)}\qquad y^2 = \frac{p_s e_s^1}{\beta} \frac{\left(\theta^2\right)^{\frac{1}{1-\beta}}}{\left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}}} \frac{\left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\left(\theta^1 \overline{\theta}\right)^{\frac{1}{1-\beta}}},$$

from which we also obtain that

$$\frac{y^2}{y^1} = \frac{\left(\theta^2\right)^{\frac{1}{1-\beta}}}{\left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}}} \frac{\left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\left(\theta^1 \overline{\theta}\right)^{\frac{1}{1-\beta}}},$$

and therefore

$$\text{(D51)}\qquad \left(\frac{y^2}{y^1}\right)^{\frac{1}{\beta}} = \frac{\left(\theta^2\right)^{\frac{1}{1-\beta}\frac{1}{\beta}}}{\left(p_q^1 - p_q^2\right)^{\frac{1}{1-\beta}}} \frac{\left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{1}{1-\beta}}}{\left(\theta^1 \overline{\theta}\right)^{\frac{1}{1-\beta}\frac{1}{\beta}}}$$

Using (D51) to substitute for $\frac{\left(y^2\right)^{\frac{1}{\beta}}}{\left(y^1\right)^{\frac{1}{\beta}}}$ on the LHS of (D46) gives

$$\begin{aligned} &-\gamma^1 p_s - \gamma^2 \frac{p_q^1 - p_q^2}{\left(\theta^2\right)^{\frac{1}{\beta}}} \frac{\left(\theta^1 \overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}} \frac{\left(\theta^2\right)^{\frac{1}{1-\beta}\frac{1}{\beta}}}{\left(p_q^1 - p_q^2\right)^{\frac{1}{1-\beta}}} \frac{\left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{1}{1-\beta}}}{\left(\theta^1 \overline{\theta}\right)^{\frac{1}{1-\beta}\frac{1}{\beta}}} p_s \\ &= \left[\frac{\gamma^2 p_q^1 p_q^2}{p_q^1 - p_q^2} \frac{\left(\theta^1\right)^{\frac{1}{\beta}} - \left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} - p_q^1\right] \frac{\left(y^1\right)^{\frac{1}{\beta}}}{\left(\theta^1\right)^{\frac{1}{\beta}} \left(e_s^1\right)^2}, \end{aligned}$$

i.e.,

$$\gamma^1 p_s + \gamma^2 p_s \frac{\left(\theta^2\right)^{\frac{1}{1-\beta}}}{\left(\theta^1 \overline{\theta}\right)^{\frac{1}{1-\beta}}} \frac{\left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}}} = \left[p_q^1 - \frac{\gamma^2 p_q^1 p_q^2}{p_q^1 - p_q^2} \frac{\left(\theta^1\right)^{\frac{1}{\beta}} - \left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}\right] \frac{\left(y^1\right)^{\frac{1}{\beta}}}{\left(\theta^1\right)^{\frac{1}{\beta}} \left(e_s^1\right)^2},$$

i.e. (exploiting the fact that $y^1 = \frac{p_s e_s^1}{\beta}$),

$$\gamma^1 + \gamma^2 \left(\frac{\theta^2}{\theta^1 \overline{\theta}}\right)^{\frac{1}{1-\beta}} \left[\frac{p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}}{p_q^1 - p_q^2}\right]^{\frac{\beta}{1-\beta}} = \left[p_q^1 - \frac{\gamma^2 p_q^1 p_q^2}{p_q^1 - p_q^2} \frac{\left(\theta^1\right)^{\frac{1}{\beta}} - \left(\overline{\theta}\right)^{\frac{1}{\beta}}}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}}\right] \frac{\left(e_s^1\right)^{\frac{1-2\beta}{\beta}} \left(p_s\right)^{\frac{1-\beta}{\beta}}}{\left(\beta \theta^1\right)^{\frac{1}{\beta}}},$$

i.e.,

$$\gamma^1+\gamma^2\left(\frac{\theta^2}{\theta^1\overline{\theta}}\right)^{\frac{1}{1-\beta}}\left[\frac{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}{p_q^1-p_q^2}\right]^{\frac{\beta}{1-\beta}}$$

$$=\left\{\frac{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}{p_q^1-p_q^2}+\frac{\gamma^1p_q^2\left[\left(\theta^1\right)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}\right]}{p_q^1-p_q^2}\right\}p_q^1\frac{\left(p_s\right)^{\frac{1-\beta}{\beta}}}{\left(\beta\right)^{\frac{1}{\beta}}}\frac{\left(e_s^1\right)^{\frac{1-2\beta}{\beta}}}{\left(\overline{\theta}\theta^1\right)^{\frac{1}{\beta}}},$$

from which one obtains that

$$e_s^1=\frac{\left(\beta\overline{\theta}\theta^1\right)^{\frac{1}{1-2\beta}}\left\{\gamma^1+\gamma^2\left(\frac{\theta^2}{\theta^1\overline{\theta}}\right)^{\frac{1}{1-\beta}}\left[\frac{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}{p_q^1-p_q^2}\right]^{\frac{\beta}{1-\beta}}\right\}^{\frac{\beta}{1-2\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}}\left(p_q^1\right)^{\frac{\beta}{1-2\beta}}\left\{\frac{\gamma^1p_q^2\left[\left(\theta^1\right)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}\right]}{p_q^1-p_q^2}+\frac{p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}{p_q^1-p_q^2}\right\}^{\frac{\beta}{1-2\beta}}},$$

or equivalently

$$e_s^1=\left\{\frac{\gamma^1\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}+\gamma^2\left(\theta^2\right)^{\frac{1}{1-\beta}}\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]}\right\}^{\frac{\beta}{1-2\beta}}$$

$$\times\frac{\left(\beta\right)^{\frac{1}{1-2\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}}\left(p_q^1\right)^{\frac{\beta}{1-2\beta}}}. \tag{D52}$$

Having found an expression for e_s^1, we have that

$$y^1=\frac{p_se_s^1}{\beta}=\beta^{\frac{2\beta}{1-2\beta}}\frac{\left(\overline{\theta}\theta^1\right)^{\frac{1}{1-2\beta}}\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}}{\left(p_sp_q^1\right)^{\frac{\beta}{1-2\beta}}}$$

$$\times\left\{\frac{\gamma^1\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}+\gamma^2\left(\frac{\theta^2}{\theta^1\overline{\theta}}\right)^{\frac{1}{1-\beta}}\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1p_q^2\left[\left(\theta^1\right)^{\frac{1}{\beta}}-\left(\overline{\theta}\right)^{\frac{1}{\beta}}\right]+p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}}\right\}^{\frac{\beta}{1-2\beta}},$$

or equivalently,

$$y^1 = \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{\beta}{1-2\beta}}$$

$$\times \frac{\beta^{\frac{2\beta}{1-2\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}}, \tag{D53}$$

and using (D50) we get that

$$y^2 = \beta^{\frac{2\beta}{1-2\beta}} \frac{\left(\overline{\theta}\theta^1\right)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}} \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} + \gamma^2 \left(\frac{\theta^2}{\theta^1\overline{\theta}}\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 p_q^2 \left[\left(\theta^1\right)^{\frac{1}{\beta}} - \left(\overline{\theta}\right)^{\frac{1}{\beta}}\right] + p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}} \right\}^{\frac{\beta}{1-2\beta}}$$

$$\times \left(\frac{\theta^2}{\theta^1\overline{\theta}}\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}},$$

or equivalently,

$$y^2 = \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{\beta}{1-2\beta}}$$

$$\times \frac{\left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}} \left(\theta^2\right)^{\frac{1}{1-\beta}} \beta^{\frac{2\beta}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}}. \tag{D54}$$

Having found expressions for y^1, y^2 and e_s^1, we can use (D43) to get the following expression for e_s^2:

$$\begin{aligned} e_s^2 &= \frac{p_q^1 - p_q^2}{\left(\theta^2\right)^{\frac{1}{\beta}}} \frac{\left(y^2\right)^{\frac{1}{\beta}}}{\left(y^1\right)^{\frac{1}{\beta}}} \frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}} e_s^1 \\ &= \frac{p_q^1 - p_q^2}{\left(\theta^2\right)^{\frac{1}{\beta}}} \left[\frac{p_q^1 - p_q^2}{\left(\theta^2\right)^{\frac{1}{\beta}}} \frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}} \right]^{\frac{1}{\beta-1}} \frac{\left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}}}{p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}} \\ &\quad \times \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{\beta}{1-2\beta}} \\ &\quad \times \frac{\left(\beta\right)^{\frac{1}{1-2\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{\beta}{1-2\beta}}}, \end{aligned}$$

i.e.,

$$e_s^2 = \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}} \frac{(\beta)^{\frac{1}{1-2\beta}} \left(\theta^2\right)^{\frac{1}{1-\beta}}}{(p_s)^{\frac{1-\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{\beta}{1-2\beta}}}$$

$$\text{(D55)} \qquad \times \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{\beta}{1-2\beta}}$$

From (D54) and (D55) we also see that

$$\text{(D56)} \qquad y^2 = \frac{p_s e_s^2}{\beta}.$$

We can now calculate the value of the government's objective function

$$U^{sep} = \gamma^2 y^2 + \gamma^1 y^1 - \gamma^1 p_s e_s^1 - \gamma^2 p_s e_s^2 - p_q^1 \left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}} \frac{1}{e_s^1} - \gamma^2 p_q^2 \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{1}{e_s^2} + \gamma^2 p_q^2 \left(\frac{y^1}{\overline{\theta}}\right)^{\frac{1}{\beta}} \frac{1}{e_s^1},$$

which can also be rewritten (using (D49) and (D56)) as

$$\text{(D57)} \quad U^{sep} = \left(\frac{1-\beta}{\beta}\right) \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s +$$
$$\gamma^2 p_q^2 \left(\frac{p_s}{\beta\overline{\theta}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}} - p_q^1 \left(\frac{p_s}{\beta\theta^1}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}} - \gamma^2 p_q^2 \left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}}.$$

Let's first compute $\gamma^2 p_q^2 \left(\frac{p_s}{\beta\overline{\theta}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}} - p_q^1 \left(\frac{p_s}{\beta\theta^1}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}$. Using (D52) we have

$$\gamma^2 p_q^2 \left(\frac{p_s}{\beta\overline{\theta}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}} - p_q^1 \left(\frac{p_s}{\beta\theta^1}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}}$$
$$= \left[\frac{\gamma^2 p_q^2}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} - \frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}}\right] \frac{(\beta)^{\frac{1}{1-2\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right)}{(p_s)^{\frac{\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{1-\beta}{1-2\beta}}}$$
$$\times \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{1-\beta}{1-2\beta}}.$$

Let's now compute $-\gamma^2 p_q^2 \left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}}$. Using (D55) we have

$$\begin{aligned}
& -\gamma^2 p_q^2 \left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}} \left(e_s^2\right)^{\frac{1-\beta}{\beta}} \\
= & -\gamma^2 p_q^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right] \frac{(\beta)^{\frac{1}{1-2\beta}}}{(p_s)^{\frac{\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{1-\beta}{1-2\beta}}} \\
& \times \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{1-\beta}{1-2\beta}} .
\end{aligned}$$

The government's objective function (D57) can then be re-expressed as

$$\begin{aligned}
U^{sep} = & \left(\frac{1-\beta}{\beta}\right) \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s \\
& + \left[\frac{\gamma^2 p_q^2}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} - \frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}}\right] \frac{(\beta)^{\frac{1}{1-2\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right)}{(p_s)^{\frac{\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{1-\beta}{1-2\beta}}} \\
& \times \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{1-\beta}{1-2\beta}} \\
& - \gamma^2 p_q^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right] \frac{(\beta)^{\frac{1}{1-2\beta}}}{(p_s)^{\frac{\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{1-\beta}{1-2\beta}}} \\
& \times \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{1-\beta}{1-2\beta}} ,
\end{aligned}$$

i.e.,

$$\begin{aligned}
U^{sep} = & \left(\frac{1-\beta}{\beta}\right) \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s \\
& + \left\{ \left[\frac{\gamma^2 p_q^2}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} - \frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}}\right] \left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right) - \gamma^2 p_q^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right] \right\} \frac{(\beta)^{\frac{1}{1-2\beta}}}{(p_s)^{\frac{\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{1-\beta}{1-2\beta}}} \\
& \times \left\{ \frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]} \right\}^{\frac{1-\beta}{1-2\beta}} .
\end{aligned}$$

Since we have that

$$\begin{aligned}
& \left[\frac{\gamma^2 p_q^2}{\left(\overline{\theta}\right)^{\frac{1}{\beta}}} - \frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}}\right] \left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right) - \gamma^2 p_q^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right] \\
= \; & \gamma^2 p_q^2 \left(\frac{1}{\overline{\theta}}\right)^{1/\beta} \left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right) - p_q^1 \left(\frac{1}{\theta^1}\right)^{1/\beta} \left(\theta^1\overline{\theta}\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right) - \gamma^2 p_q^2 p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 p_q^2 p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}} \\
= \; & \gamma^2 p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right) - p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} \left(p_q^1 - p_q^2\right) - \gamma^2 p_q^2 p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 p_q^2 p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}} \\
= \; & \left[\gamma^2 p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}} - p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^1 p_q^2 \left(\overline{\theta}\right)^{\frac{1}{\beta}}\right] p_q^1 \\
= \; & -\left\{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\} p_q^1,
\end{aligned}$$

the government's objective function can be re-expressed as

$$\begin{aligned}
U^{sep} = & \left(\frac{1-\beta}{\beta}\right) \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s \\
& - \left\{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\} \frac{\left(\beta\right)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}} \\
& \times \left\{\frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]}\right\}^{\frac{1-\beta}{1-2\beta}},
\end{aligned}$$

i.e.,

$$\begin{aligned}
U^{sep} = & \left(\frac{1-\beta}{\beta}\right) \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s \\
& - \frac{\left(\beta\right)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}} \frac{\left\{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}},
\end{aligned}$$

or equivalently,

$$\begin{aligned}
U^{sep} = & \left(\frac{1-2\beta}{\beta}\right) \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s \\
& - \frac{\left(\beta\right)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}} \frac{\left\{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}} \\
& + \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s.
\end{aligned}$$

Noticing that

$$
\begin{aligned}
& \left(\gamma^2 e_s^2+\gamma^1 e_s^1\right) p_s \\
= \; & \gamma^2 e_s^2 p_s+\gamma^1 e_s^1 p_s \\
= \; & \gamma^2 p_s\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}} \frac{(\beta)^{\frac{1}{1-2\beta}}\left(\theta^2\right)^{\frac{1}{1-\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}}\left(p_q^1\right)^{\frac{\beta}{1-2\beta}}} \\
& \times\left\{\frac{\gamma^1\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}+\gamma^2\left(\theta^2\right)^{\frac{1}{1-\beta}}\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]}\right\}^{\frac{\beta}{1-2\beta}} \\
& +\gamma^1 p_s\left\{\frac{\gamma^1\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}+\gamma^2\left(\theta^2\right)^{\frac{1}{1-\beta}}\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]}\right\}^{\frac{\beta}{1-2\beta}} \\
& \times\frac{(\beta)^{\frac{1}{1-2\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}}\left(p_q^1\right)^{\frac{\beta}{1-2\beta}}} \\
= \; & \left\{\frac{\gamma^1\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}+\gamma^2\left(\theta^2\right)^{\frac{1}{1-\beta}}\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]}\right\}^{\frac{\beta}{1-2\beta}} \frac{(\beta)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}} \\
& \times\left\{\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\left(\theta^2\right)^{\frac{1}{1-\beta}}+\gamma^1\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\right\} \\
= \; & \frac{(\beta)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}} \frac{\left\{\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\left(\theta^2\right)^{\frac{1}{1-\beta}}+\gamma^1\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}},
\end{aligned}
$$

we have that

$$
\begin{aligned}
&\left(\frac{1-2\beta}{\beta}\right)\left(\gamma^2 e_s^2+\gamma^1 e_s^1\right)p_s \\
&-\frac{(\beta)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}}\frac{\left\{\gamma^1\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}+\gamma^2\left(\theta^2\right)^{\frac{1}{1-\beta}}\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}} \\
&+\left(\gamma^2 e_s^2+\gamma^1 e_s^1\right)p_s \\
=&\left(\frac{1-2\beta}{\beta}\right)\left(\gamma^2 e_s^2+\gamma^1 e_s^1\right)p_s \\
&-\frac{(\beta)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}}\frac{\left\{\gamma^1\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}+\gamma^2\left(\theta^2\right)^{\frac{1}{1-\beta}}\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}} \\
&+\frac{(\beta)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}}\frac{\left\{\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\left(\theta^2\right)^{\frac{1}{1-\beta}}+\gamma^1\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}} \\
=&\left(\frac{1-2\beta}{\beta}\right)\left(\gamma^2 e_s^2+\gamma^1 e_s^1\right)p_s.
\end{aligned}
$$

Thus, the government's objective function can be re-expressed as $\left(\frac{1-2\beta}{\beta}\right)\left(\gamma^2 e_s^2+\gamma^1 e_s^1\right)p_s$. Substituting the optimal values for e_s^1 and e_s^2 gives

$$
\begin{aligned}
U^{sep} &= \left(\frac{1-2\beta}{\beta}\right)\left(\gamma^2 e_s^2+\gamma^1 e_s^1\right)p_s \\
&= \left(\frac{1-2\beta}{\beta}\right)\frac{(\beta)^{\frac{1}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}}\frac{\left\{\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}\left(\theta^2\right)^{\frac{1}{1-\beta}}+\gamma^1\left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}}\left(p_q^1-p_q^2\right)^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1\left(p_q^1-p_q^2\right)\left(\overline{\theta}\right)^{\frac{1}{\beta}}+\gamma^2\left[p_q^1\left(\overline{\theta}\right)^{\frac{1}{\beta}}-p_q^2\left(\theta^1\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}}.
\end{aligned}
\tag{D58}
$$

Finally, notice that (D49) and (D56) allow restating U^{sep} as

$$U^{sep} = (1-2\beta)\left(\gamma^2 y^2+\gamma^1 y^1\right). \tag{D59}$$

D.3 The regions used to generate Figure 1

From (D11), (D12), (D34) and (D59), we have that a pooling tax equilibrium dominates a separating tax equilibrium if and only if

$$\text{(D60)} \qquad \frac{\left(\bar{\theta}\right)^{\frac{1}{1-2\beta}} \beta^{\frac{2\beta}{1-2\beta}}}{\left(p_s p_q^1\right)^{\frac{\beta}{1-2\beta}}} > \gamma^1 y^1 + \gamma^2 y^2.$$

Under a separating tax equilibrium the upward IC-constraint can be safely disregarded when (D39) holds. In this case, the values of y^1 and y^2 are given by (D28) and (D23).

When instead condition (D39) is violated, both IC-constraints will be binding at a separating tax equilibrium; in this case the values of y^1 and y^2 are given by (D53) and (D54).

For given values of p_q^1 and θ^2, let the ability advantage of type-2 agents be denoted by $\epsilon = \theta^2 - \theta^1 > 0$ and the cost disadvantage of type-1 agents denoted by $\delta = p_q^1 - p_q^2 > 0$. Moreover, define the function $f(\delta, \varepsilon)$, for $(\delta, \varepsilon) \in \left[0, p_q^1\right] \times [0, \theta^2]$, as

(D61)

$$f(\delta, \varepsilon) \equiv \frac{\delta}{(\theta^2)^{\frac{1}{\beta}}} - \left[\frac{p_q^1}{(\theta^2 - \varepsilon)^{\frac{1}{\beta}}} - \frac{p_q^1 - \delta}{(\theta^2 - \gamma^1 \varepsilon)^{\frac{1}{\beta}}}\right] \left[\frac{\gamma^1 \left(p_q^1 - \delta\right)}{p_q^1 \left(\frac{\theta^2}{\theta^2 - \varepsilon}\right)^{\frac{1}{\beta}} - \gamma^2 \left(p_q^1 - \delta\right) \left(\frac{\theta^2}{\theta^2 - \gamma^1 \varepsilon}\right)^{\frac{1}{\beta}}}\right]^{\frac{1-\beta}{1-2\beta}}.$$

Our results imply that:

i) For (δ, ε)-pairs such that $f(\delta, \varepsilon) \geq 0$, predistribution is desirable if and only if

$$\text{(D62)} \qquad \left[\frac{(\theta^2 - \gamma^1 \varepsilon)^{\frac{1}{\beta}}}{p_q^1}\right]^{\frac{\beta}{1-2\beta}} > \gamma^1 \left[\frac{\gamma^1}{\frac{p_q^1}{(\theta^2 - \varepsilon)^{\frac{1}{\beta}}} - \frac{\gamma^2 (p_q^1 - \delta)}{(\theta^2 - \gamma^1 \varepsilon)^{\frac{1}{\beta}}}}\right]^{\frac{\beta}{1-2\beta}} + \gamma^2 \left[\frac{(\theta^2)^{\frac{1}{\beta}}}{p_q^1 - \delta}\right]^{\frac{\beta}{1-2\beta}}.$$

ii) For (δ, ε)-pairs such that $f(\delta, \varepsilon) < 0$, predistribution is desirable if and only if

(D63)

$$\left(\theta^2 - \gamma^1 \varepsilon\right)^{\frac{1}{1-2\beta}} > \frac{\left\{\gamma^2 \left[p_q^1 - \left(p_q^1 - \delta\right) \left(\frac{\theta^2 - \varepsilon}{\theta^2 - \gamma^1 \varepsilon}\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}} (\theta^2)^{\frac{1}{1-\beta}} + \gamma^1 (\theta^2 - \varepsilon)^{\frac{1}{1-\beta}} (\delta)^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1 \delta + \gamma^2 \left[p_q^1 - \left(p_q^1 - \delta\right) \left(\frac{\theta^2 - \varepsilon}{\theta^2 - \gamma^1 \varepsilon}\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}}.$$

E Proof of Proposition 3

Part (a) The government's problem under pooling can be equivalently stated as

$$\max_{e_s, e_q} \overline{\theta} h\left(e_s, e_q\right) - p_s e_s - p_q^1 e_q, \tag{E1}$$

Denoting by a hat symbol the optimal values of e_s and e_q, and using subscripts on h to denote partial derivatives, the associated first order conditions are

$$1 - \frac{p_s}{\overline{\theta} h_1\left(\widehat{e}_s, \widehat{e}_q\right)} = 0, \tag{E2}$$

$$1 - \frac{p_q^1}{\overline{\theta} h_2\left(\widehat{e}_s, \widehat{e}_q\right)} = 0, \tag{E3}$$

from which it also follows that

$$\frac{p_s}{p_q^1} = \frac{h_1\left(\widehat{e}_s, \widehat{e}_q\right)}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)} < \frac{p_s}{p_q^2}. \tag{E4}$$

Part (b) From the relationship $\theta h\left(e_s, e_q\right) = y$, one can derive a function, denoted by f, that expresses e_q as a function of y, e_s and θ: $e_q = f\left(y, e_s, \theta\right)$. Relying on such a function, define $R^1\left(y^1, e_s^1\right)$, $R^2\left(y^2, e_s^2\right)$, $\widehat{R}^2\left(y^1, e_s^1\right)$ and $\tilde{R}^1(y^2, e_s^2)$ as follows:

$$R^1\left(y^1, e_s^1\right) \equiv p_s e_s^1 + p_q^1 f\left(y^1, e_s^1, \theta^1\right), \tag{E5}$$

$$R^2\left(y^2, e_s^2\right) \equiv p_s e_s^2 + p_q^2 f\left(y^2, e_s^2, \theta^2\right), \tag{E6}$$

$$\widehat{R}^2\left(y^1, e_s^1\right) \equiv p_s e_s^1 + p_q^2 f\left(y^1, e_s^1, \overline{\theta}\right), \tag{E7}$$

$$\tilde{R}^1(y^2, e_s^2) \equiv p_s e_s^2 + p_q^1 f\left(y^2, e_s^2, \theta^2\right). \tag{E8}$$

Based on (E5)-(E8) we can then equivalently reformulate the government's optimal tax problem as

$$\max_{y^1, e_s^1, c^1, y^2, e_s^2, c^2} c^1 - R^1\left(y^1, e_s^1\right) \tag{E9}$$

subject to

$$c^2 - R^2\left(y^2, e_s^2\right) \geq c^1 - \widehat{R}^2\left(y^1, e_s^1\right) \tag{E10}$$

$$c^1 - R^1\left(y^1, e_s^1\right) \geq c^2 - \tilde{R}^1(y^2, e_s^2) \tag{E11}$$

$$\gamma^1\left(y^1 - c^1\right) + \gamma^2\left(y^2 - c^2\right) \geq 0. \tag{E12}$$

Denote by λ^2, λ^1 and μ the Lagrange multipliers of the government's problem. The first order conditions with respect to, respectively $y^1, e_s^1, c^1, y^2, e_s^2, c^2$, are

$$
\begin{aligned}
&\text{(E13)} & -\frac{\partial R^1(y^1,e_s^1)}{\partial y^1} + \lambda^2\frac{\partial \widehat{R}^2(y^1,e_s^1)}{\partial y^1} - \lambda^1\frac{\partial R^1(y^1,e_s^1)}{\partial y^1} + \mu\gamma^1 &= 0,\\
&\text{(E14)} & -\frac{\partial R^1(y^1,e_s^1)}{\partial e_s^1} + \lambda^2\frac{\partial \widehat{R}^2(y^1,e_s^1)}{\partial e_s^1} - \lambda^1\frac{\partial R^1(y^1,e_s^1)}{\partial e_s^1} &= 0,\\
&\text{(E15)} & 1 - \lambda^2 + \lambda^1 - \mu\gamma^1 &= 0,\\
&\text{(E16)} & -\lambda^2\frac{\partial R^2(y^2,e_s^2)}{\partial y^2} + \lambda^1\frac{\partial \tilde{R}^1(y^2,e_s^2)}{\partial y^2} + \mu\gamma^2 &= 0,\\
&\text{(E17)} & -\lambda^2\frac{\partial R^2(y^2,e_s^2)}{\partial e_s^2} + \lambda^1\frac{\partial \tilde{R}^1(y^2,e_s^2)}{\partial e_s^2} &= 0,\\
&\text{(E18)} & \lambda^2 - \lambda^1 - \mu\gamma^2 &= 0.
\end{aligned}
$$

From (E15) and (E18) we get that $\mu = 1$ and $\lambda^2 - \lambda^1 = 1 - \gamma^1 = \gamma^2$. Taking this into account, from (E17)-(E18) we get

$$
\begin{aligned}
-\frac{\partial R^2(y^2,e_s^2)}{\partial e_s^2} &= \frac{\lambda^1}{\gamma^2}\left(\frac{\partial R^2(y^2,e_s^2)}{\partial e_s^2} - \frac{\partial \tilde{R}^1(y^2,e_s^2)}{\partial e_s^2}\right)\\
&= \frac{\lambda^1}{\gamma^2}\left[\left(p_s - \frac{h_1(e_s^2, f(y^2,e_s^2,\theta^2))}{h_2(e_s^2, f(y^2,e_s^2,\theta^2))}p_q^2\right) - \left(p_s - \frac{h_1(e_s^2, f(y^2,e_s^2,\theta^2))}{h_2(e_s^2, f(y^2,e_s^2,\theta^2))}p_q^1\right)\right]\\
\text{(E19)} \qquad &= \frac{\lambda^1}{\gamma^2}\left(p_q^1 - p_q^2\right)\frac{h_1(e_s^2, f(y^2,e_s^2,\theta^2))}{h_2(e_s^2, f(y^2,e_s^2,\theta^2))}.
\end{aligned}
$$

Noticing that $\frac{\partial R^2(y^2,e_s^2)}{\partial e_s^2} = p_s - p_q^2\frac{h_1(e_s^2,f(y^2,e_s^2,\theta^2))}{h_2(e_s^2,f(y^2,e_s^2,\theta^2))}$, eq. (E19) implies eq. (33). We therefore have that

$$
\text{(E20)} \qquad \frac{h_1(e_s^2, f(y^2,e_s^2,\theta^2))}{h_2(e_s^2, f(y^2,e_s^2,\theta^2))} = \frac{h_1(e_s^2,e_q^2)}{h_2(e_s^2,e_q^2)} \geq \frac{p_s}{p_q^2}.
$$

Combining (E16) and (E18) gives

$$
\text{(E21)} \qquad 1 - \frac{\partial R^2(y^2,e_s^2)}{\partial y^2} = \frac{\lambda^1}{\gamma^2}\left(p_q^2 - p_q^1\right)\left(\frac{de_q^2}{dy^2}\right)_{de_s^2=0} = \frac{\lambda^1}{\gamma^2}\frac{p_q^2 - p_q^1}{\theta^2 h_2(e_s^2, f(y^2,e_s^2,\theta^2))} \leq 0.
$$

Noticing that $\frac{\partial R^2(y^2,e_s^2)}{\partial y^2} = \frac{p_q^2}{\theta^2 h_2(e_s^2,f(y^2,e_s^2,\theta^2))}$, eq. (E21) implies eq. (34). The result stated by eq. (35) can then be easily obtained combining the results provided by (E19) and (E21).

From (E14)-(E15) we have that

$$
\text{(E22)} \qquad -\frac{\partial R^1(y^1,e_s^1)}{\partial e_s^1}\left(\lambda^2 + \gamma^1\right) = -\lambda^2\frac{\partial \widehat{R}^2(y^1,e_s^1)}{\partial e_s^1},
$$

or equivalently

$$\text{(E23)} \qquad -\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial e_s^1}=\frac{\lambda^2}{\gamma^1}\left(\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial e_s^1}-\frac{\partial \widehat{R}^2\left(y^1,e_s^1\right)}{\partial e_s^1}\right).$$

Noticing that $\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial e_s^1}=p_s-p_q^1\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}$, eq. (E23) can be restated as

$$\text{(E24)} \qquad -p_s+p_q^1\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}$$

$$\text{(E25)} \qquad =\frac{\lambda^2}{\gamma^1}\left[\left(p_s-\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}p_q^1\right)-\left(p_s-\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}p_q^2\right)\right]$$

$$\text{(E26)} \qquad =\frac{\lambda^2}{\gamma^1}\left[\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}p_q^2-\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}p_q^1\right],$$

from which eq. (30) is obtained. Notice that the right hand side of the equation above has a negative sign given that $p_q^2<p_q^1$ and $f\left(y^1,e_s^1,\overline{\theta}\right)<f\left(y^1,e_s^1,\theta^1\right)$ (which implies that $\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}<\frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}$). It therefore follows that

$$\text{(E27)} \qquad \frac{h_1\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}{h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}=\frac{h_1\left(e_s^1,e_q^1\right)}{h_2\left(e_s^1,e_q^1\right)}<\frac{p_s}{p_q^1}.$$

From (E13) and (E15) we have that

$$\text{(E28)} \qquad -\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial y^1}\left(\lambda^2+\gamma^1\right)=-\lambda^2\frac{\partial \widehat{R}^2\left(y^1,e_s^1\right)}{\partial y^1}-\gamma^1,$$

or equivalently

$$\text{(E29)} \qquad 1-\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial y^1}=\frac{\lambda^2}{\gamma^1}\left(\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial y^1}-\frac{\partial \widehat{R}^2\left(y^1,e_s^1\right)}{\partial y^1}\right).$$

Noticing that $\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial y^1}=\frac{p_q^1}{\theta^1 h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}$, eq. (E29) can be restated as

$$\text{(E30)} \qquad 1-\frac{\partial R^1\left(y^1,e_s^1\right)}{\partial y^1}=\frac{\lambda^2}{\gamma^1}\left(\frac{p_q^1}{\theta^1 h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}-\frac{p_q^2}{\overline{\theta}h_2\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}\right),$$

from which eq. (31) is obtained. Notice that the right hand side of the equation above has a positive sign given that $p_q^1>p_q^2$ and $f\left(y^1,e_s^1,\theta^1\right)>f\left(y^1,e_s^1,\overline{\theta}\right)$ (which implies that $h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)<h_2\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)$). Finally, the result stated by eq. (32) can be easily obtained combining the results provided by (30) and (31).

F Proof of Proposition 4

Assume first that the nonlinear income tax is confiscatory at all levels of income other than y^1 and y^2. The subsidy schedule for $y = y^2$ is flat and has $\sigma(e_s) = 0$ $\forall e_s$. Suppose that, conditional on earning a gross income equal to y^1, the subsidy on e_s is proportional at rate σ. Along the isoquant $y^1 = \theta^1 h(e_s, e_q)$, type-1 agents will choose an effort mix that satisfies the condition $MRTS^1 = (1-\sigma)\, p_s/p_q^1$. Thus, setting $\sigma = 1 - \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} \frac{p_q^1}{p_s}$ will induce type-1 agents to adopt the constrained-efficient effort mix $\left(e_s^1, e_q^1\right)$. From (30) we also have that

$$\sigma = \frac{\lambda^2}{\gamma^1}\left(MRTS^1 - MRTS^{21}\frac{p_q^2}{p_q^1}\right)\frac{p_q^1}{p_s} > 0. \tag{F1}$$

Denote by σ^* the subsidy rate provided by (F1). For later purposes, notice that type-1 agents would still be induced to adopt the constrained-efficient effort mix $\left(e_s^1, e_q^1\right)$ if one were to replace the proportional subsidy with a piecewise-linear structure such that $\sigma > \sigma^*$ for $e_s \leq e_s^1$ and $\sigma = \sigma^*$ for $e_s > e_s^1$. The problem with a proportional subsidy set at rate σ^* is that it does not guarantee that, for a type-2 agent behaving as a mimicker (and therefore earning a gross income equal to y^1), choosing $e_s = e_s^1$ is indeed optimal. To consider the incentives of a type-2 agent earning y^1 , consider first the case of a type-2 agent being remunerated according to the average productivity $\overline{\theta}$. Along the isoquant $y^1 = \overline{\theta} h(e_s, e_q)$, a type-2 mimicker would find it optimal to select $e_s = e_s^1$ provided that the subsidy on e_s is set at a rate such that

$$\frac{h_1\left(e_s^1, \widehat{e}_q^2\right)}{h_2\left(e_s^1, \widehat{e}_q^2\right)} = \frac{(1-\sigma)\, p_s}{p_q^2}. \tag{F2}$$

Denote by $\widehat{\sigma}$ the subsidy that satisfies the condition (F2). Since $p_q^2 < p_q^1$ and $\widehat{e}_q^2 < e_q^1$ (which implies that $\frac{h_1\left(e_s^1, \widehat{e}_q^2\right)}{h_2\left(e_s^1, \widehat{e}_q^2\right)} < \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)}$), it follows that $\widehat{\sigma} > \sigma^*$.

$$\widehat{\sigma} = 1 - \frac{h_1\left(e_s^1, \widehat{e}_q^2\right)}{h_2\left(e_s^1, \widehat{e}_q^2\right)} \frac{p_q^2}{p_s}.$$

Moreover, since

$$\begin{aligned}
\widehat{\sigma} &= 1 - \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} \frac{p_q^1}{p_s} + 1 - \frac{h_1\left(e_s^1, \widehat{e}_q^2\right)}{h_2\left(e_s^1, \widehat{e}_q^2\right)} \frac{p_q^2}{p_s} - 1 + \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} \frac{p_q^1}{p_s} \\
&= 1 - \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} \frac{p_q^1}{p_s} + \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} \frac{p_q^1}{p_s} - \frac{h_1\left(e_s^1, \widehat{e}_q^2\right)}{h_2\left(e_s^1, \widehat{e}_q^2\right)} \frac{p_q^2}{p_s},
\end{aligned}$$

exploiting (F1), we can also re-express $\widehat{\sigma}$ as

$$\widehat{\sigma} = \left(1 + \frac{\lambda^2}{\gamma^1}\right)\left(MRTS^1 - MRTS^{21}\frac{p_q^2}{p_q^1}\right)\frac{p_q^1}{p_s} > \sigma^*.$$

Thus, a piecewise-linear subsidy schedule such that $\sigma = \widehat{\sigma}$ for $e_s \leq e_s^1$ and $\sigma = \sigma^*$ for $e_s > e_s^1$ would ensure that $e_s = e_s^1$ is the optimal choice both for a type-1 agent being remunerated according to the productivity θ^1 and for a type-2 mimicker being remunerated according to the average productivity $\overline{\theta}$.

The last thing that one needs to check is whether the proposed piecewise-linear subsidy schedule is sufficient to deter type-2 agents from earning y^1 while being remunerated according to their true productivity θ^2. Put differently, one should check whether, under the proposed piecewise-linear subsidy schedule, it is indeed the case that type-2 mimickers have no incentive to choose a value $e_s \neq e_s^1$ that allows them to achieve separation at an income equal to y^1. For this purpose, consider in the (e_s, e_q)-plane the point $\left(e_s^1, e_q^1\right)$ (on the isoquant $y^1 = \theta^1 h\left(e_s, e_q\right)$) and the point $\left(e_s^1, \widehat{e}_q^2\right)$ (on the isoquant $y^1 = \overline{\theta} h\left(e_s, e_q\right)$), and consider the two isocost lines $\left(1 - \widehat{\sigma}\right) p_s e_s + p_q^1 e_q = \left(1 - \widehat{\sigma}\right) p_s e_s^1 + p_q^1 e_q^1$ and $\left(1 - \widehat{\sigma}\right) p_s e_s + p_q^2 e_q = \left(1 - \widehat{\sigma}\right) p_s e_s^1 + p_q^2 \widehat{e}_q^2$, the first pertaining to type-1 agents and passing through the point $\left(e_s^1, e_q^1\right)$, and the second pertaining to type-2 agents and passing through the point $\left(e_s^1, \widehat{e}_q^2\right)$ (with the first isocost line being flatter than the second due to the fact that $p_q^1 > p_q^2$). Type-2 mimickers will have no incentive to choose a value $e_s \neq e_s^1$ that allows them to achieve separation at $y = y^1$ if and only if, on the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$, there is no pair $\left(e_s, e_q\right)$ that is at the same time below the isocost line $\left(1 - \widehat{\sigma}\right) p_s e_s + p_q^2 e_q = \left(1 - \widehat{\sigma}\right) p_s e_s^1 + p_q^2 \widehat{e}_q^2$ (meaning that it entails for type-2 agents an effort cost that is lower than the one they would sustain if mimicking by pooling, i.e. earning y^1 while being remunerated according to the average productivity $\overline{\theta}$) and above the isocost line $\left(1 - \widehat{\sigma}\right) p_s e_s + p_q^1 e_q = \left(1 - \widehat{\sigma}\right) p_s e_s^1 + p_q^1 e_q^1$ (meaning that type-1 agents would be deterred from replicating the effort choices of type-2 agents). Denote by $\left(e_s^{int}, e_q^{int}\right)$ the point of intersection of the two isocost lines. As the isocost line associated with type-2 agents is steeper than the isocost line associated with their type-1 counterparts, two observations follow. First, notice that $e_s^{int} < e_s^1$ and $e_q^{int} > e_q^1$. Further notice that an allocation, which is both cost-saving for type-2 agents and above the isocost line associated with type-1 agents, necessarily lies to the left of the intersection point. Two cases, hence, must be distinguished.

Case i) The point $\left(e_s^{int}, e_q^{int}\right)$ is to the left of the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$ (or at a point on this isoquant), i.e. $\theta^2 h\left(e_s^{int}, e_q^{int}\right) \leq y^1$. In this case, type-2 mimickers find it unprofitable to separate themselves from their type-2 counterparts. They are (weakly) better off earning y^1 while being remunerated according to the average productivity $\overline{\theta}$ than earning y^1 while being remunerated according to the true productivity θ^2. It then follows that the two-bracket piecewise-linear subsidy schedule is enough for implementation purposes.

Case ii) The point $\left(e_s^{int}, e_q^{int}\right)$ is to the right of the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$, i.e. $\theta^2 h\left(e_s^{int}, e_q^{int}\right) > y^1$. In this case the proposed two-bracket piecewise-linear subsidy schedule is not enough for implementation purposes. The reason is that type-2 mimickers can earn y^1 while achieving separation (choosing $e_s < e_s^{int}$ and $e_q > e_q^{int}$) and sustaining a cost $(1-\widehat{\sigma})\, p_s e_s + p_q^2 e_q$ which is strictly lower than $(1-\widehat{\sigma})\, p_s e_s^1 + p_q^2 \widehat{e}_q^2$. In particular, the fact that the point $\left(e_s^{int}, e_q^{int}\right)$ is to the right of the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$ means that along the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$ there is a non-empty set of points $\left(e_s, e_q\right)$ that are at the same time below the isocost line $(1-\widehat{\sigma})\, p_s e_s + p_q^2 e_q = (1-\widehat{\sigma})\, p_s e_s^1 + p_q^2 \widehat{e}_q^2$ and above the isocost line $(1-\widehat{\sigma})\, p_s e_s + p_q^1 e_q = (1-\widehat{\sigma})\, p_s e_s^1 + p_q^1 e_q^1$. To guard against the possibility that type-2 mimickers find it preferable to earn y^1 while achieving separation (rather than pooling with type-1 agents at the effort mix $\left(e_s^1, \widehat{e}_q^2\right)$), a third bracket should be included in the subsidy schedule. In particular, let $\sigma = 100\%$ for $e_s \leq e_s^{int}$, $\sigma = \widehat{\sigma}$ for $e_s^{int} < e_s \leq e_s^1$ and $\sigma = \sigma^*$ for $e_s > e_s^1$. This change implies that lowering e_s below e_s^{int} does not entail any cost-saving for an agent; in particular, any combination $\left(e_s, e_q\right)$ (on the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$) such that $e_s < e_s^{int}$ and $e_q > e_q^{int}$ (needed to achieve separation) is more costly for type-2 agents than $p_q^2 e_q^{int}$, and therefore more costly than $(1-\widehat{\sigma})\left(e_s^1 - e_s^{int}\right) p_s + p_q^2 \widehat{e}_q^2$. This implies that for type-2 mimickers it is strictly better to earn y^1 while being remunerated according to the average productivity $\overline{\theta}$ rather than by achieving separation.

Under case i), type-1 agents receive a transfer equal to $\widehat{\sigma} p_s e_s^1$ through the education subsidy schedule. To satisfy the government's budget constraint it must be that $T\left(y^1\right) = y^1 - c^1 + \widehat{\sigma} p_s e_s^1$. Under case ii), type-1 agents receive a transfer equal to $\left[e_s^{int} + \left(e_s^1 - e_s^{int}\right)\widehat{\sigma}\right] p_s$ through the education subsidy schedule. To satisfy the government's budget constraint it must be that $T\left(y^1\right) = y^1 - c^1 + \left[e_s^{int} + \left(e_s^1 - e_s^{int}\right)\widehat{\sigma}\right] p_s$ and $T\left(y^2\right) = y^2 - c^2$.

So far, in our discussion we have been assuming that the nonlinear income tax was set at confiscatory levels for all levels of income other than y^1 and y^2. However, this assumption can be safely replaced with the alternative assumption that $T\left(y \neq y^1\right) = y^2 - c^2$. The reason is that such a change does not affect the incentives of type-2 agents. We already know that, by construction, type-2 agents are weakly better off earning y^2 rather than earning y^1. Now suppose that, for $y \neq \{y^1, y^2\}$, we replace the confiscatory income tax with a lump-sum tax set at $y^2 - c^2$, and suppose that, after this change is implemented, type-2 agents prefer to earn $y \neq y^2$. Given that type-2 agents choose y and e_s to maximize their utility subject to the constraint that type-1 agents have no incentive to replicate their choices, it then follows that the government could have offered to type-2 agents an allocation different than $\left(y^2, c^2, e_s^2, e_q^2\right)$ that jointly satisfies the following three conditions: i) it allows the government to keep unchanged the amount of taxes raised from type-2 agents; ii) it allows type-2 agents to achieve a higher utility; iii) it does not violate the constraint requiring type-1 agents not to be tempted to replicate the choices of type-2 agents. But if this had been possible, the new allocation would have allowed the government to relax the downward IC constraint (23), which has to be binding given our assumption of a max-min social objective, contradicting the assumption that $\left(y^2, c^2, e_s^2, e_q^2\right)$ belongs to the MMO. Thus, setting $T\left(y \neq y^1\right) = y^2 - c^2$ for $y \neq \{y^1, y^2\}$ implies that type-2 agents are better off achieving separation by earning y^2 and choosing $e_s = e_s^2$. At the same time, it is also true that type-2 agents cannot achieve a higher utility by pooling with type-1 agents, and being remunerated according to the average productivity $\overline{\theta}$, at some income level $y \neq y^1$. In fact, suppose that this were indeed possible for them. In order to pool with type-1 agents at $y \neq y^1$, it must be the case that type-1 agents are weakly better off earning y, while being remunerated according to the average productivity, than getting the bundle $\left(y^1, c^1, e_s^1, e_q^1\right)$ intended for them. However, this would contradict the assumption that the MMO was an STE. The reason is that there would exist a pooling allocation, $(y, y^2 - c^2, e_s, e_q)$, such that the utility of type-1 agents is not lower than at the presumed optimal bundle $\left(y^1, c^1, e_s^1, e_q^1\right)$, while at the same time allowing the government to raise a higher net tax revenue (running a budget surplus rather than a budget-balance).[34] Notice that type-1 agents, clearly, have no incentives to separate themselves from their type-2 counterparts by choosing a level of income $y \neq \{y^1, y^2\}$, which is strictly dominated by pooling with type-2 agents at this level of income.

Finally, notice that, for $y = y^1$, the proposed subsidy-schedule follows a declining scale ($\sigma'(e_s) \leq 0$). However, all the incentives would be left unaffected if one were to replace this schedule with a simpler schedule featuring only two brackets: $\sigma(e_s) = 100\%$ for $e_s \leq e_s^1$ and $\sigma(e_s) = 0$ for $e_s > e_s^1$. This would imply extending over the entire interval $(0, e_s^1)$ the most generous subsidy rate. To maintain public budget balance, in this case one should adjust the income tax function by properly raising $T(y^1)$; in particular one should set $T(y^1) = y^1 - c^1 + p_s e_s^1$. This alternative subsidy schedule would be equivalent to adopting a system with an income-dependent mandate. The only difference is that under a mandate the public expenditure for education is nil, since individuals pay the full price but are required to get a minimum amount. This implied that, under an income-dependent mandate, public budget balance is guaranteed by setting $T(y^1) = y^1 - c^1$. Furthermore, if it is the case that $e_s^1 \leq e_s^2$ at the MMO, an even simpler implementing scheme would suffice because in such a case there would be no need to let the mandate be income-dependent.

[34] At the pooling allocation $\left(y, y^2 - c^2, e_s, e_q\right)$, education subsidies are zero and everybody pays an income tax $y^2 - c^2 > 0$.

G Proof of Proposition 5

Consider the optimization problem solved by an agent of type-1 under a tax system featuring a proportional tax/subsidy on income supplemented by a mandate on e_s prescribing that $e_s \geq \widehat{e}_s$:

$$\max_{e_s^1, e_q^1} \ (1-t)\,\theta^1 h\left(e_s^1, e_q^1\right) - p_s e_s^1 - p_q^1 e_q^1 \quad \text{subject to } e_s^1 \geq \widehat{e}_s. \tag{G1}$$

The associated first order conditions would be

$$(1-t)\,\theta^1 h_1\left(e_s^1, e_q^1\right) \leq p_s, \tag{G2}$$

$$(1-t)\,\theta^1 h_2\left(e_s^1, e_q^1\right) = p_q^1. \tag{G3}$$

Suppose that $t = \left(\frac{1}{\overline{\overline{\theta}}} - \frac{1}{\theta^1}\right)\frac{p_q^1}{h_2(\widehat{e}_s, \widehat{e}_q)} < 0$. The first order conditions (G2)-(G3) would become

$$1 + \left(\frac{1}{\theta^1} - \frac{1}{\overline{\overline{\theta}}}\right)\frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)} \leq \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)}, \tag{G4}$$

$$1 + \left(\frac{1}{\theta^1} - \frac{1}{\overline{\overline{\theta}}}\right)\frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)} = \frac{p_q^1}{\theta^1 h_2\left(e_s^1, e_q^1\right)}. \tag{G5}$$

Given that the MMO satisfies (E3), it follows that the effort mix $(\widehat{e}_s, \widehat{e}_q)$ satisfies the first order condition (G5). Moreover, given that the MMO also satisfies (E4), it also follows that the effort mix $(\widehat{e}_s, \widehat{e}_q)$ satisfies (as an equality) the first order condition (G4).

Exploiting the above result, let the mandate on e_s be set at $\widehat{e}_s$, and assume that, for $y \in [0, \widehat{y}]$, the income tax chosen by the government takes the following linear form:

$$T(y) = \underbrace{\left(\frac{1}{\theta^1} - \frac{1}{\overline{\overline{\theta}}}\right)\frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)}}_{>0}\widehat{y} + \underbrace{\left(\frac{1}{\overline{\overline{\theta}}} - \frac{1}{\theta^1}\right)\frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)}}_{<0}y. \tag{G6}$$

Notice that the tax function (G6) features a constant marginal subsidy ($T' = \left(\frac{1}{\overline{\overline{\theta}}} - \frac{1}{\theta^1}\right)\frac{p_q^1}{h_2(\widehat{e}_s, \widehat{e}_q)} < 0$), a decreasing average tax rate, and also that, by construction, it satisfies the condition $T(\widehat{y}) = 0$. Assume initially that $T(y) = 0$ also for $y > \widehat{y}$ (we will later revise this assumption). It follows that $(\widehat{e}_s, \widehat{e}_q)$ represents the effort mix (e_s, e_q) that maximizes $\theta^1 h(e_s, e_q) - T(\theta^1 h(e_s, e_q)) - p_s e_s - p_q^1 e_q$ subject to the constraint $e_s \geq \widehat{e}_s$.

Consider now the various options available to type-2 agents.

i) Suppose that they try to achieve separation from their low-skilled counterpart. Under separation, type-1 agents get a utility equal to $y^1 - T(y^1) - p_s\widehat{e}_s - p_q^1\widehat{e}_q$. Notice also that, given our assumptions about $T(y)$, type-2 agents cannot achieve separation at a level of income $y \in [y^1, \theta^2 h_2(\widehat{e}_s, \widehat{e}_q)]$ (where one should notice that $\theta^2 h_2(\widehat{e}_s, \widehat{e}_q) > \widehat{y} > y^1$)

and would not find attractive to achieve separation at a level of income lower than y^1. To show this, notice first that, to verify whether or not it is possible or attractive for type-2 agents to achieve separation at $y^2 \leq \theta^2 h_2(\widehat{e}_s, \widehat{e}_q)$, it suffices to check whether separation is achievable or attractive when choosing $e_s^2 = \widehat{e}_s$. This is because type-2 agents have a comparative advantage in the e_q-dimension, and the mandate on e_s prevents them from choosing a value of e_s smaller than $\widehat{e}_s$. Consider first what would happen if type-2 agents try to achieve separation at an equilibrium where they earn $\widehat{y}$; they would choose the effort mix $(\widehat{e}_s, e_q(\widehat{y}, \widehat{e}_s, \theta^2))$. However, given that $y^1 = \theta^1 h(\widehat{e}_s, \widehat{e}_q) < \widehat{y}$, $T(y^1) > T(\widehat{y}) = 0$ and $e_q(\widehat{y}, \widehat{e}_s, \theta^2) < \widehat{e}_q$, if type-2 agents were to choose the effort mix $(\widehat{e}_s, e_q(\widehat{y}, \widehat{e}_s, \theta^2))$ they would not succeed in achieving separation (because type-1 agents would be strictly better off by replicating the effort mix of type-2 agents than by choosing $(\widehat{e}_s, \widehat{e}_q)$: $y^1 - T(y^1) - p_s\widehat{e}_s - p_q^1\widehat{e}_q < \widehat{y} - p_s\widehat{e}_s - p_q^1 e_q(\widehat{y}, \widehat{e}_s, \theta^2)$). A similar argument can be invoked to show that type-2 agents could never achieve separation at a level of income y^2 such that $y^2 \in (\widehat{y}, \theta^2 h_2(\widehat{e}_s, \widehat{e}_q)]$.[35] It can also be used to show that type-2 agents could never achieve separation at a level of income y^2 such that $y^2 \in [y^1, \widehat{y})$. In particular, for type-2 agents to be able to achieve separation it must be that

$$y^1 - T(y^1) - p_q^1\widehat{e}_q > y^2 - T(y^2) - p_q^1 e_q(y^2, \widehat{e}_s, \theta^2). \tag{G7}$$

Given the assumptions made about $T(y)$, for $y^2 \in [y^1, \widehat{y})$ we have that $[y^2 - T(y^2)] - [y^1 - T(y^1)] \geq 0$. But given that $e_q(y^2, \widehat{e}_s, \theta^2) - \widehat{e}_q < 0$, the inequality $[y^2 - T(y^2)] - [y^1 - T(y^1)] < p_q^1 [e_q(y^2, \widehat{e}_s, \theta^2) - \widehat{e}_q]$ is violated. Now consider the case where $y^2 < y^1$. In this case we have that $[y^2 - T(y^2)] - [y^1 - T(y^1)] < 0$ and therefore one cannot rule out the possibility that (G7) is satisfied and therefore separation is achievable. However, even if type-2 agents could achieve separation at some value of income smaller than y^1 , they would not have an incentive to do that. In fact, for separation to be attractive for them it must be that

$$y^2 - T(y^2) - p_q^2 e_q(y^2, \widehat{e}_s, \theta^2) > y^1 - T(y^1) - p_q^2 e_q(y^1, \widehat{e}_s, \overline{\theta}). \tag{G8}$$

Together, the two inequalities (G7)-(G8) require that

$$p_q^2 \left[e_q(y^2, \widehat{e}_s, \theta^2) - e_q(y^1, \widehat{e}_s, \overline{\theta})\right] < \left[y^2 - T(y^2)\right] - \left[y^1 - T(y^1)\right] < p_q^1 \left[e_q(y^2, \widehat{e}_s, \theta^2) - \widehat{e}_q\right]. \tag{G9}$$

But since $e_q(y^2, \widehat{e}_s, \theta^2) - \widehat{e}_q < e_q(y^2, \widehat{e}_s, \theta^2) - e_q(y^1, \widehat{e}_s, \overline{\theta}) < 0$, and $p_q^2 < p_q^1$, we have that

$$p_q^2 \left[e_q(y^2, \widehat{e}_s, \theta^2) - e_q(y^1, \widehat{e}_s, \overline{\theta})\right] > p_q^1 \left[e_q(y^2, \widehat{e}_s, \theta^2) - \widehat{e}_q\right], \tag{G10}$$

[35] In this case the argument also takes into account that we have assumed that $T(y) = 0$ for $y > \widehat{y}$.

i.e. (G7)-(G8) cannot be jointly satisfied.

We can then conclude that, for type-2 agents, the only feasible and attractive way to achieve separation requires them to choose $e_q > \widehat{e}_q$ and earn a pre-tax income that is strictly larger than $\theta^2 h(\widehat{e}_s, \widehat{e}_q)$. Denote by $e_q^{\min}$ the minimum level of e_q that allows type-2 agents to achieve separation when choosing $e_s = \widehat{e}_s$. Formally, $e_q^{\min}$ is defined as the solution to the following problem:

$$\text{(G11)} \qquad \min_{e_q} \theta^2 h(\widehat{e}_s, e_q) \quad \text{subject to } y^1 - T\left(y^1\right) - p_q^1 \widehat{e}_q \geq \theta^2 h(\widehat{e}_s, e_q) - p_q^1 e_q.$$

Furthermore, denote by $\left(e_s^{2*}, e_q^{2*}\right)$ the effort mix that solves the following unconstrained maximization problem:

$$\text{(G12)} \qquad \max_{e_s, e_q} \theta^2 h(e_s, e_q) - p_s e_s - p_q^2 e_q.$$

Notice that, since $(\widehat{e}_s, \widehat{e}_q)$ is the effort mix that maximizes $\overline{\theta} h(e_s, e_q) - p_s e_s - p_q^1 e_q$, it must necessarily be that $e_s^{2*} > \widehat{e}_s$ and $e_q^{2*} > \widehat{e}_q$.

Finally, define $\underline{y}^{sep}$ as $\underline{y}^{sep} \equiv \theta^2 h\left(\widehat{e}_s, e_q^{\min}\right)$. Notice that, under the assumption that $T(y) = 0$ for $y \geq \widehat{y}$, the gain that type-2 agents can obtain by separating from their low-ability counterpart (instead of choosing the effort mix $(\widehat{e}_s, \widehat{e}_q)$ and pooling with them at $\widehat{y}$) cannot exceed the amount $\left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s \widehat{e}_s - p_q^2 \widehat{e}_q\right]$. Thus, to ensure that type-2 agents never find attractive to separate from their low-ability counterpart, it would suffice to modify our initial assumption that $T(y) = 0$ for $y \geq \widehat{y}$ and let $T(y)$, for $y \geq \underline{y}^{sep}$, be given by

$$\text{(G13)} \qquad T(y) = \left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s \widehat{e}_s - p_q^2 \widehat{e}_q\right] > 0.$$

ii) What we have established so far is that a pooling equilibrium at $\widehat{y}$, where both agents choose $(\widehat{e}_s, \widehat{e}_q)$, is weakly better for type-2 agents than any separating equilibrium that they can achieve. Moreover, for type-1 agents, a pooling equilibrium at $\widehat{y}$ is strictly better than a separating equilibrium; this is because under separation they achieve a utility equal to $y^1 - T(y^1) - p_s \widehat{e}_s - p_q^1 \widehat{e}_q$, which is lower than $\widehat{y} - p_s \widehat{e}_s - p_q^1 \widehat{e}_q$ (remember that $y^1 < \widehat{y}$ and $T(y^1) > 0$). Notice also that from the perspective of type-1 agents, the best pooling allocation is the one where all agents earn $\widehat{y}$ (the pooling allocation $(\widehat{y}, \widehat{c}, \widehat{e}_s, \widehat{e}_q)$ was obtained as the outcome of the government's problem where the utility of type-1 agents was maximized within the set of pooling allocation, and therefore pooling at $\widehat{y}$ would be the preferred choice of type-1 agents even in the absence of taxes; the conclusion is strengthened by the fact that we have defined an income tax function such that $T(y) \geq 0$ for all values of y). Thus, the only thing that is left to check, in order to establish that our function $T(y)$ implements the optimal pooling allocation, is to verify whether, for type-2 agents, pooling at $\widehat{y}$ is not dominated by pooling at some other level of income weakly smaller than $\overline{\theta}h\left(\widehat{e}_s, e_q^{\min}\right)$.[36] Clearly, given that $p_q^2 < p_q^1$ and the income tax function is regressive in the interval $[0, \widehat{y}]$, pooling at $\widehat{y}$ is strictly preferred by type-2 agents to pooling at a level of income smaller than $\widehat{y}$. Notice, however, that pooling at a value of y slightly higher than $\widehat{y}$ would be strictly better for type-2 agents than pooling at $\widehat{y}$ if, as we have assumed so far, $T(y) = 0$ for $y \in [\widehat{y}, \theta^2 h\left(\widehat{e}_s, e_q^{\min}\right))$. Moreover, pooling at a value of y slightly larger than $\widehat{y}$ would still allow type-1 agents to achieve a utility that is higher than the one achieved at a separating equilibrium. This represents a threat to the implementability of the pooling allocation intended by the government. To eliminate this threat we need to properly adjust the tax schedule $T(y)$. For this purpose, notice that, switching from pooling at $\widehat{y}$ to pooling at a marginally higher level of income entails for type-2 agents a maximum gain that is given by $1 - \frac{p_q^2}{\overline{\theta}h_2(\widehat{e}_s,\widehat{e}_q)}$.[37] Notice, also, that the benefit of pooling at a marginally higher level of income is decreasing in income.[38] Therefore, to make sure that type-2 agents weakly prefer pooling at $\widehat{y}$ to pooling at levels of income higher than $\widehat{y}$, it would suffice to assume that $T'(y) = 1 - \frac{p_q^2}{\overline{\theta}h_2(\widehat{e}_s,\widehat{e}_q)}$ for $y \geq \widehat{y}$.

Combining the insights obtained in i) and ii), it follows that one way to implement the allocation $(\widehat{y}, \widehat{c}, \widehat{e}_s, \widehat{e}_q)$ as a pooling tax equilibrium is to enforce a lower bound on e_s, set at $\widehat{e}_s$, supplemented by a two-bracket piecewise-linear income tax $T(y)$ such that

(G14)

$$T(y) = \begin{cases} \left(\frac{1}{\theta^1} - \frac{1}{\overline{\theta}}\right) \frac{p_q^1}{h_2(\widehat{e}_s,\widehat{e}_q)}\widehat{y} + \left(\frac{1}{\overline{\theta}} - \frac{1}{\theta^1}\right) \frac{p_q^1}{h_2(\widehat{e}_s,\widehat{e}_q)} y, & \text{for all } y \in [0, \widehat{y}] \\ (y - \widehat{y}) \max\left\{1 - \frac{p_q^2}{\overline{\theta}h_2(\widehat{e}_s,\widehat{e}_q)}, \frac{\left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s\widehat{e}_s - p_q^2\widehat{e}_q\right]}{\underline{y}^{sep} - \widehat{y}}\right\}, & \text{for all } y > \widehat{y}. \end{cases}$$

H The case when neither signal is observable

Here we consider the special case where an individual's tax liability is only a function of his or her labor income. 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 $t^i \equiv y^i - c^i$. Recall that the wage rate earned by a given individual is defined as the ratio of his or her pre-tax income y and the value of the h-function evaluated at the effort

[36]Notice that we can safely disregard the case of pooling at $\overline{\theta}h\left(\widehat{e}_s, e_q^{\min}\right) < y < \theta^2 h\left(\widehat{e}_s, e_q^{\min}\right)$; the reason is that type-2 agents achieve separation when choosing $e_q \geq e_q^{\min}$, and therefore there can be no pooling at levels of pre-tax income higher than $\overline{\theta}h\left(\widehat{e}_s, e_q^{\min}\right)$.

[37]Since $\frac{h_1(\widehat{e}_s,\widehat{e}_q)}{h_2(\widehat{e}_s,\widehat{e}_q)} < \frac{p_s}{p_q^2}$, the maximum gain can be calculated assuming that the additional output is produced by only relying on an upward variation in e_q.

[38]Switching from pooling at $\widehat{y}$ to pooling at $\widehat{y} + \epsilon$ raises the utility of type-2 agents by a larger amount than switching from pooling at $\widehat{y} + \epsilon$ to pooling at $\widehat{y} + 2\epsilon$.

vector chosen by the individual.

H.1 A pooling tax equilibrium when neither signal is observable

Since there is no exogenous public revenue requirement, in a pooling equilibrium the income tax system offers the same pre-tax income $\hat{y}$ to both types of agents, which is also equal to the net income denoted by $\hat{c}$. Lemma 3 below shows that pooling equilibria where all agents choose the same effort vector do not exist.

Lemma 3. *With only an income tax in place, a pooling tax equilibrium where both workers choose the same effort vector does not exist.*

Proof. Consider a candidate pooling allocation $(\hat{y}, \hat{c})$ where both workers choose the same effort mix, given by the pair $(\hat{e}_s, \hat{e}_q)$. 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 $\hat{c} - (p_s e_s^i + p_q^i e_q^i) = \hat{u}^i$, for $i = 1, 2$, will describe the indifference curves, in the (e_s, e_q) plane, passing through the point $(\hat{e}_s, \hat{e}_q)$. Since the indifference curve for agents of type i (with $i = 1, 2$) has a slope of $-p_s/p_q^i$, it follows that 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 forked region northwest of point $(\hat{e}_s, \hat{e}_q)$. Now suppose that instead of choosing $(\widehat{e}_s, \widehat{e}_q)$, type-2 agents deviate to the effort mix $\left(\widehat{e}_s - \epsilon, \widehat{e}_q + \frac{p_s}{p_q^2 + \nu}\epsilon\right)$, where $\epsilon > 0$ and $0 < \nu < p_q^1 - p_q^2$. By construction, the effort mixture $\left(\widehat{e}_s - \epsilon, \widehat{e}_q + \frac{p_s}{p_q^2 + \nu}\epsilon\right)$ is inside the above forked region, which implies that it has a lower cost for type-2 agents than the effort mix $(\widehat{e}_s, \widehat{e}_q)$, while it has a higher cost for type-1 agents. Therefore, by deviating to the effort mix $\left(\widehat{e}_s - \epsilon, \widehat{e}_q + \frac{p_s}{p_q^2 + \nu}\epsilon\right)$, type-2 workers can credibly reveal their productivity. Moreover, since in (e_s, e_q)-space the isoquant $\widehat{y} = \theta^2 h(e_s, e_q)$ is strictly below the isoquant $\widehat{y} = \overline{\theta} h(e_s, e_q)$, it follows by continuity that for sufficiently small ϵ, the total output produced by a deviating type-2 worker would strictly exceed $\widehat{y}$. Thus, it would also be the case that firms find it profitable to hire the deviating type-2 worker. □

Note that due to the two dimensions of signaling, it is possible to have pooling in income without pooling in the effort vectors chosen by the two agents. Lemma 4 shows that such an equilibrium will never be the social optimum.

Lemma 4. *With only an income tax in place, pooling on income without pooling on the effort signals observed by firms is socially suboptimal.*

Proof. When there is pooling of income without pooling of effort signals, type 1 individuals are: (a) paid their true productivity before taxes, and (b) pay zero net taxes. Both (a) and (b) are true under laissez-faire. In addition, they have an undistorted effort mix under laissez-faire. So a pooling allocation can't possibly be better than laissez-faire in terms of type 1's welfare. However, the laissez-faire equilibrium is clearly dominated by the socially optimal separating allocation implemented by the income tax, which entails some redistribution and thus yields a higher level of utility for type 1 workers than under laissez-faire. We conclude that pooling of income without pooling of effort signals is suboptimal.[39] □

Lemma 3 and Lemma 4 together imply that under a pure income tax system the optimal solution is given by a separating equilibrium in which types 1 and 2 earn different levels of income.

Proposition 6. *With only an income tax in place, pre-distribution cannot be achieved and the social optimum is always given by a separating tax equilibrium.*

H.2 A separating tax equilibrium when neither signal is observable

In a separating tax equilibrium, agents are paid by the firms according to their true productivity (a type i agent is paid a wage rate of θ^i). The problem of choosing the tax schedule $T(y)$ can be equivalently formulated as the problem of properly selecting two pairs of pre-tax and after-tax incomes (y^i, c^i), where $c^i = y^i - T(y^i)$, $y^1 - c^1 < 0$ and $y^2 - c^2 > 0$. Besides satisfying the government budget constraint, the two bundles must be chosen in such a way that they are incentive-compatible: agents of type i, for $i = 1, 2$ must be weakly better off at the bundle intended for them, i.e. the bundle (y^i, c^i), than at the bundle intended for agents of type $j \neq i$, i.e. the bundle (y^j, c^j). The main difference from the standard Mirrleesian (1971) setup is the presence of a second layer of asymmetric information between workers and employers. The latter implies that to render the allocation incentive compatible, one should not only consider (as stated above) mimicking by replication (that is, choosing the bundle intended for the other type), but also off-equilibrium path mimicking options. We turn next to explore this in detail.

Consider first the bundle associated with type-1 workers. As we will formally prove below, in the socially optimal separating equilibrium, type-2 workers will never resort to mimicking by replication (they will hence strictly prefer their bundle to choosing the bundle intended for type-1 workers). In the standard model, mimicking by replication is the only option available to type-2 workers and hence the associated IC constraint will be binding in the optimal solution. In our setup, in contrast, there will be superior alternatives for type 2, due to the presence of asymmetric information between workers and employers.

[39] The argument is similar to the standard argument why bunching with two types is never optimal in a standard Mirrleesian setting without asymmetric information between firms and workers, see, e.g., Stiglitz (1982).

As in equilibrium, type-2 workers will never mimic by replication, if type-1 agents choose the bundle (y^1, c^1) intended for them, they would select an efficient mix of e_s and e_q, denoted by $(e_s^1(y^1), e_q^1(y^1))$, with an associated cost given by:

$$\text{(H1)} \qquad R^1(y^1) = \min_{e_s, e_q} R^1(e_s, e_q) \quad \text{subject to} \quad h(e_s, e_q)\theta^1 = y^1.$$

The only case in which the effort mix is being distorted in the optimal solution, is when an IC constraint associated with mimicking by replication is binding in the optimal solution. In such a case, distorting the effort mix would serve to mitigate the constraint.

Notice that efficiency in the choice of the effort mix means that $e_s^1(y^1)$ and $e_q^1(y^1)$ satisfy the condition

$$\text{(H2)} \qquad \frac{\partial h\left(e_s^1(y^1), e_q^1(y^1)\right)/\partial e_s^1}{\partial h\left(e_s^1(y^1), e_q^1(y^1)\right)/\partial e_q^1} = \frac{p_s}{p_q^1},$$

which equates the marginal rate of technical substitution (MRTS) to the marginal cost ratio.

In contrast to type-1 workers, the effort mix of type-2 workers may well be distorted in the optimal solution. This is because, as we will formally show below, mimicking by replication would be desirable for type-1 agents. If the associated IC constraint would bind, distorting the effort mix would serve to alleviate the constraint. If they were to choose the bundle (y^2, c^2) intended for them, type-2 agents would select a mix of e_s and e_q, denoted by $(e_s^2(y^2), e_q^2(y^2))$, with an associated cost given by:

$$\text{(H3)} \qquad R^2(y^2) = \min_{e_s, e_q} R^2(e_s, e_q) \quad \text{subject to}$$

$$\text{(H4)} \qquad h(e_s, e_q)\theta^2 = y^2$$

$$\text{(H5)} \qquad c^1 - R^1\left(y^1\right) \geq c^2 - R^2\left(y^2\right) - \left(p_q^1 - p_q^2\right) e_q^2.$$

The second constraint captures the fact that type-2 agents take also into account that the effort mix that they choose must not be attractive for type-1 agents. Therefore, the effort mix chosen by type-2 agents will depend on whether this constraint is binding or slack. If it is slack, the effort mix $(e_s^2(y^2), e_q^2(y^2))$ will satisfy the efficiency condition $\frac{\partial h\left(e_s^2, e_q^2\right)/e_s^2}{\partial h\left(e_s^2, e_q^2\right)/e_q^2} = \frac{p_s}{p_q^2}$; if the constraint is binding, the effort mix will satisfy the inequality $\frac{\partial h\left(e_s^2, e_q^2\right)/e_s^2}{\partial h\left(e_s^2, e_q^2\right)/e_q^2} > \frac{p_s}{p_q^2}$ (i.e., it will be distorted towards e_q, the effort dimension on which type-2 agents have a comparative advantage).

Let's now consider the incentive-compatibility constraints that should be accounted for by the government in the choice of the two bundles (y^i, c^i). To implement a given separating equilibrium, the government must guard against various deviating strategies available to agents, i.e., the government must ensure that no agent has an incentive to deviate from the expected behavior. In principle, there are three deviating strategies that an agent of type i can choose to earn the income y^j intended for the other type. Agents of type i can choose an effort vector that allow them to be compensated according to (i) the productivity of the other type, (ii) the average productivity, or (iii) their true productivity. We consider these three deviating strategies in more detail below. Since a deviating agent is someone who earns an amount of income that is intended for some other type of agent, we will use 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. By behaving in this way, a type-i mimicker would be paid a wage rate θ^j (i.e., according to the productivity of the type being mimicked) and would incur the following costs:

$$\breve{R}^i(y^j) = p_s^i \breve{e}_s^i(y^j) + p_q^i \breve{e}_q^i(y^j), \tag{H6}$$

where $(\breve{e}_s^i(y^j), \breve{e}_q^i(y^j)) = (e_s^j(y^j), e_q^j(y^j))$ denotes the effort mix of a mimicker of type i, which is identical to the effort mix chosen in equilibrium by agents of type j.

In addition to the deviating strategy described above, in which a mimicker of type i chooses the effort vector chosen in equilibrium by agents of type $j \neq i$, there are also deviating strategies in which a mimicker chooses a 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 that is at once: i) different from the one chosen in equilibrium by type-j agents, ii) attractive also to type-j agents, and iii) sufficient to allow firms to make non-negative profits when paying agents according to the average productivity $\bar{\theta}$. For a type-i mimicker, the most attractive of such strategies is the one with associated costs given by:

$$\hat{R}^i(y^j) = \min_{(e_s,e_q)\neq(e_s^j(y^j),e_q^j(y^j))} R^i(e_s, e_q) \tag{H7}$$

subject to:

$$R^j(e_s, e_q) \leq R^j(y^j), \tag{H8}$$

$$y^j \leq h(e_s, e_q)\bar{\theta}. \tag{H9}$$

The constraint (H8) captures the fact that the deviating strategy is feasible in the sense that it also induces type j agents to change their effort vectors. The constraint (H9) ensures that the effort vector is sufficient to provide a non-negative profit for the hiring firm in a pooling equilibrium where both agents are paid according to the average productivity $\bar{\theta}$. Lemma 5 shows that, in equilibrium, this out-of-equilibrium deviation would never be profitable for type-1 agents.

Lemma 5. *The IC constraint associated with type-1 agents mimicking by pooling with their, type-2, high-skilled counterparts will be slack in the optimal solution.*

Proof. Under the proposed 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 $h(e_s, e_q)\bar{\theta} \geq y^2$. To sustain the deviation to the pooling allocation, in equilibrium, type-2 agents should be indifferent between the pooling allocation and the bundle intended for them. The latter follows from a combination of two weak inequalities: the intended bundle should be weakly preferred to the pooling allocation (by construction of the equilibrium) and at the same time the pooling allocation should be weakly preferred to the intended bundle (to make the deviation to the pooling allocation feasible). However, by Lemma 3, we can find an alternative bundle to the presumably optimal bundle offered to type-2 agents in equilibrium, that will separate them from type-1 agents and deliver them a strictly higher level of utility. This yields the desired contradiction, as, by offering the new bundle to type-2 agents, the government can create a slack in the IC-constraint of type-2 agents and thereby enhance redistribution towards type-1 agents. □

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 to the mimicking strategy of replicating the effort vector chosen in equilibrium by type-1 agents.

Lemma 6. *For a type-2 agent, it is always more attractive to earn y^1 while being rewarded according to average productivity $\bar{\theta}$ than to earn y^1 while being rewarded according to low productivity $\theta^1 < \bar{\theta}$. In other words, $\hat{R}^2(y^1) < \breve{R}^2(y^1)$.*

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. □

The other deviating strategy involving the choice of an off-equilibrium effort vector is the one in which type-i agents mimic the earned income y^j of type-j agents, but invest in the signals in such a way as to differentiate themselves from type-j agents and thereby succeed in being compensated by firms according to their true productivity θ^i. For a type-i mimicker, the most attractive of such strategies is the one with associated costs given by:

$$\tilde{R}^i(y^j) = \min_{(e_s,e_q)\neq(e_s^j(y^j),e_q^j(y^j))} R^i(e_s, e_q) \tag{H10}$$

subject to:

$$R^j(e_s, e_q) \geq R^j(y^j), \tag{H11}$$

$$y^j \leq h(e_s, e_q)\theta^i. \tag{H12}$$

In the above problem, the constraint (H11) ensures that the effort vector chosen by type-i mimickers is not attractive to type-j agents, thereby allowing type-i mimickers to separate from their type-j counterparts. The constraint (H12) instead ensures 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 $p_s^1 = p_s^2 \equiv p_s$ and $p_q^1 > p_q^2$, it necessarily follows that $\tilde{R}^2(y^1) < R^1(y^1)$ and $\tilde{R}^1(y^2) > R^2(y^2)$. Notice also that there are two possible scenarios in which type-i agents succeed in separating from type-j agents at income level y^j: one in which the constraint (H11) is binding, and another in which it is slack. In the former case, the agent behaving as a mimicker will use a distorted effort mix (i.e, an effort mix that violates 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.

Consider now which of the deviating strategies described above are really relevant, from the point of view of the government, when choosing the bundles (y^i, c^i). As we have previously pointed out, of the three deviating strategies that are potentially available to type-2 agents, the deviating strategy with associated cost $\hat{R}^2(y^1)$ is necessarily more attractive than the one with associated cost $\breve{R}^2(y^1)$. The government can then safely neglect the latter. Thus, a first incentive-compatibility constraint that is relevant for the government is that

$$c^2 - R^2(y^2) \geq c^1 - \min\left\{\widetilde{R}^2(y^1), \hat{R}^2(y^1)\right\}, \tag{H13}$$

Regarding type-1 agents we know, given the content of Lemma 1, that the only two available strategies are the one with associated cost $\breve{R}^1(y^2)$ and the one with associated cost $\tilde{R}^1(y^2)$. Thus, it would appear that a second IC-constraint that is relevant for the government is

$$c^1 - R^1(y^1) \geq c^2 - \min\left\{\breve{R}^1(y^2), \widetilde{R}^1(y^2)\right\}. \tag{H14}$$

Suppose however that the social optimum is a separating equilibrium and that the constraint (H14) is binding with $\min\left\{\breve{R}^1(y^2), \widetilde{R}^1(y^2)\right\} = \widetilde{R}^1(y^2)$. Given that $y^1 - c^1 < 0$ and $y^2 - c^2 > 0$, the government could then do better by removing (y^1, c^1) from the menu of bundles available on the income tax schedule and letting type-1 agents bear the cost of

$\widetilde{R}^1(y^2)$ and pool with type-2 agents at (y^2, c^2). type-1 agents would not suffer, since we have, by assumption, that $c^1 - R^1(y^1)) = c^2 - \widetilde{R}^1(y^2)$. At the same time, since $y^1 - c^1 < 0$ and $y^2 - c^2 > 0$ in the supposedly optimal separating equilibrium, the government would experience an increase in revenue. Thus, if $\min\left\{\breve{R}^1(y^2), \widetilde{R}^1(y^2)\right\} = \widetilde{R}^1(y^2)$, then the constraint (H14) is necessarily slack. Put differently, the only relevant IC-constraint pertaining to the behavior of type-1 agents is $c^1 - R^1(y^1) \geq c^2 - \breve{R}^1(y^2)$. However, notice that this constraint is already embedded in the optimization problem solved by type-2 agents. This implies that it is possible to formulate the government's problem in a way that does not include this IC-constraint as a separate constraint. In order to do this, the only requirement is that one rewrites the IC-constraint pertaining to type-2 agents as follows:

$$\text{(H15)} \qquad c^2 - R^2(y^2, c^2, y^1, c^1) \geq c^1 - \min\left\{\widetilde{R}^2(y^1), \hat{R}^2(y^1)\right\}.$$

The constraint (H15) embeds implicitly also the IC-constraint pertaining to type-1 agents ($c^1 - R^1(y^1) \geq c^2 - \breve{R}^1(y^2)$) because it highlights that the minimum cost sustained by type-2 agents in order to produce y^2 is not only a function of y^2 but also of the variables c^2, y^1 and c^1, which all affect the incentives for type-1 agents to behave as mimickers. This observation allows restating the government's optimal tax problem in a simplified way as follows:

$$\text{(H16)} \qquad \max_{\{y^i, c^i\}_{i=1,2}} c^1 - R^1(y^1)$$

subject to the government budget constraint

$$\text{(H17)} \qquad \sum_i \gamma^i (y^i - c^i) = 0,$$

and the downward IC-constraints

$$\text{(H18)} \qquad c^2 - R^2(y^2, c^2, y^1, c^1) \geq c^1 - \widetilde{R}^2(y^1),$$

$$\text{(H19)} \qquad c^2 - R^2(y^2, c^2, y^1, c^1) \geq c^1 - \hat{R}^2(y^1).$$

Denote respectively by μ, λ^{2s} and λ^{2p} the Lagrange multipliers attached to constraint (H17), (H18) and (H19). The first order conditions with respect to, respectively y^1, c^1,

y^2 and c^2, are

(H20)

$$-\frac{\partial R^1\left(y^1\right)}{\partial y^1}-\left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial y^1}+\lambda^{2s}\frac{\partial \widetilde{R}^2\left(y^1\right)}{\partial y^1}+\lambda^{2p}\frac{\partial \widehat{R}^2\left(y^1\right)}{\partial y^1}+\mu\gamma^1=0,$$

(H21)

$$1-\left(\lambda^{2s}+\lambda^{2p}\right)-\left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^1}-\mu\gamma^1=0,$$

(H22)

$$-\left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2\left(y^2\right)}{\partial y^2}+\mu\gamma^2=0,$$

(H23)

$$\lambda^{2s}+\lambda^{2p}-\left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^2}-\mu\gamma^2=0.$$

Adding up (H21) and (H23), and simplifying terms, gives

$$1-\left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^1}-\left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^2}=\mu. \tag{H24}$$

But given that $\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^1}+\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^2}=0$ (a joint marginal increase in c^1 and c^2 has no impact on the upward IC-constraint that enters the optimization problem solved by type-2 agents, and therefore has no impact on $R^2\left(y^2,c^2,y^1,c^1\right)$), eq. (H24) implies that $\mu=1$. Taking this into account we can rewrite (H22) and (H23) as, respectively

$$-\left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2\left(y^2\right)}{\partial y^2} = -\gamma^2, \tag{H25}$$

$$\left(\lambda^{2s}+\lambda^{2p}\right)\left(1-\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^2}\right) = \gamma^2. \tag{H26}$$

Dividing (H25) by (H26) and rearranging terms gives

$$1-\frac{\frac{\partial R^2\left(y^2\right)}{\partial y^2}}{1-\frac{\partial R^2\left(y^2,c^2,y^1,c^1\right)}{\partial c^2}}=0. \tag{H27}$$

To interpret (H27) in terms of the properties of the implementing tax function, consider the individual optimization problem for type-2 agents under a nonlinear income tax $T\left(y\right)$. This can be described as follows:

$$\max_{e_s^2,e_q^2}\quad \theta^2 h\left(e_s^2,e_q^2\right)-T\left(\theta^2 h\left(e_s^2,e_q^2\right)\right)-p_s e_s^2-p_q^2 e_q^2 \tag{H28}$$

subject to the IC-constraint

$$U^1 \geq \theta^2 h\left(e_s^2, e_q^2\right) - T\left(\theta^2 h\left(e_s^2, e_q^2\right)\right) - p_s e_s^2 - \left(p_q^1 - p_q^2\right) e_q^2. \tag{H29}$$

Equivalently, the optimization problem of type-2 agents can be reformulated as a two-stage problem. In the first stage, for a given amount of production y, consumption $y-T(y)$, and for a given level of utility U^1 achieved by type-1 agents when not behaving as mimickers, type-2 agents choose the effort mix that minimizes production costs subject to the IC-constraint prescribing that type-1 agents have no incentive to replicate the choices of type-2 agents. This gives a conditional indirect utility function

$$V^2\left(y, y-T(y), U^1\right) = y - T(y) - R\left(y, y-T(y), U^1\right), \tag{H30}$$

where $R\left(y, y-T(y), U^1\right) = p_s e_s^2\left(y, y-T(y), U^1\right) + p_q^2 e_q^2\left(y - T(y), U^1\right)$. Notice that type-2 agents will not necessarily choose an effort mix that satisfies the efficiency condition $\frac{h_1\left(e_s^2, e_q^2\right)}{h_2\left(e_s^2, e_q^2\right)} = \frac{p_s}{p_q^2}$ (where $h_1\left(e_s^2, e_q^2\right) \equiv \frac{\partial h\left(e_s^2, e_q^2\right)}{\partial e_s^2}$ and $h_2\left(e_s^2, e_q^2\right) \equiv \frac{\partial h\left(e_s^2, e_q^2\right)}{\partial e_q^2}$). Given that the optimization problem solved by type-2 agents is subject to an IC-constraint that is aimed at deterring type-1 agents from replicating their effort choices, the efficiency condition $\frac{h_1\left(e_s^2, e_q^2\right)}{h_2\left(e_s^2, e_q^2\right)} = \frac{p_s}{p_q^2}$ will be satisfied only if the IC-constraint is not binding. If instead the IC-constraint is binding, the effort mix chosen by type-2 agents will satisfy the inequality $\frac{h_1\left(e_s^2, e_q^2\right)}{h_2\left(e_s^2, e_q^2\right)} > \frac{p_s}{p_q^2}$ (i.e. it will be distorted towards e_q, which reflects the circumstance that type-2 agents have a comparative advantage in the quality dimension of effort).

At the second stage y is optimally chosen subject to the link between pre-tax earnings and post-tax earnings determined by the tax schedule $T(y)$. The first order condition for this problem is given by

$$1 - T'(y) - \frac{\partial R\left(y, y-T(y), U^1\right)}{\partial y} - \frac{\partial R\left(y, y-T(y), U^1\right)}{\partial c}\left(1 - T'(y)\right) = 0, \tag{H31}$$

from which we can derive the following implicit characterization of the marginal income tax rate faced by type-2 agents:

$$T'(y) = \frac{1 - \frac{\partial R\left(y, y-T(y), U^1\right)}{\partial y} - \frac{\partial R\left(y, y-T(y), U^1\right)}{\partial c}}{1 - \frac{\partial R\left(y, y-T(y), U^1\right)}{\partial c}} = 1 - \frac{\frac{\partial R\left(y, y-T(y), U^1\right)}{\partial y}}{1 - \frac{\partial R\left(y, y-T(y), U^1\right)}{\partial c}}. \tag{H32}$$

Thus, combining (H27) and (H32), we can conclude that

$$T'\left(y^2\right) = 0, \tag{H33}$$

namely that the constrained social optimum can be implemented letting type-2 agents face a zero marginal income tax rate. It is important to emphasize that this result does not imply that y^2 is going to be equal to its first-best efficient level. If the upward IC-constraint that enters the optimization problem of type-2 agents is binding at the constrained social optimum, y^2 is going to exceed its first-best efficient level. What (H27) tells us is that, even if the socially (constrained) optimal value of y^2 is above its first-best efficient level, there is no reason to use the marginal income tax rate faced by type-2 agents to affect the incentives underlying their decision process. The reason is that these incentives are already aligned with those underlying the social decision problem: given that the upward IC-constraint is already part, even in the absence of taxation, of the optimization problem solved by type- 2 agents, there is no need to use the policy instruments to let type-2 agents internalize this constraint.

Consider now the first order conditions (H20)-(H21) and rewrite them, respectively, as

$$\text{(H34)}\quad -\frac{\partial R^1(y^1)}{\partial y^1} = \left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial y^1} - \lambda^{2s}\frac{\partial \widetilde{R}^2(y^1)}{\partial y^1} - \lambda^{2p}\frac{\partial \widehat{R}^2(y^1)}{\partial y^1} - \gamma^1,$$

$$\text{(H35)}\quad 1 = \left(\lambda^{2s}+\lambda^{2p}\right)\left(1+\frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial c^1}\right)+\gamma^1.$$

Dividing (H34) by (H35) and multiplying by the right hand side of (H35) gives

$$\begin{aligned} &-\frac{\partial R^1(y^1)}{\partial y^1}\left[\left(\lambda^{2s}+\lambda^{2p}\right)\left(1+\frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial c^1}\right)+\gamma^1\right] \\ \text{(H36)}\qquad &= \left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial y^1} - \lambda^{2s}\frac{\partial \widetilde{R}^2(y^1)}{\partial y^1} - \lambda^{2p}\frac{\partial \widehat{R}^2(y^1)}{\partial y^1} - \gamma^1, \end{aligned}$$

or equivalently:

$$\begin{aligned} \text{(H37)}\quad \gamma^1\left(1-\frac{\partial R^1(y^1)}{\partial y^1}\right) &= \left(\lambda^{2s}+\lambda^{2p}\right)\frac{\partial R^1(y^1)}{\partial y^1} - \lambda^{2s}\frac{\partial \widetilde{R}^2(y^1)}{\partial y^1} - \lambda^{2p}\frac{\partial \widehat{R}^2(y^1)}{\partial y^1} \\ &+ \left(\lambda^{2s}+\lambda^{2p}\right)\left(\frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial y^1}+\frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial c^1}\frac{\partial R^1(y^1)}{\partial y^1}\right). \end{aligned}$$

Notice however that eq. (H37) can be further simplified by realizing that a marginal increase in y^1, coupled with an upward adjustment in c^1 by $\frac{\partial R^1(y^1)}{\partial y^1}$, leaves unaffected the utility of type-1 agents and therefore has no impact on $R^2(y^2,c^2,y^1,c^1)$. Thus, (H37) can be equivalently restated as

$$\text{(H38)}\qquad 1-\frac{\partial R^1(y^1)}{\partial y^1} = \frac{\lambda^{2s}}{\gamma^1}\left(\frac{\partial R^1(y^1)}{\partial y^1} - \frac{\partial \widetilde{R}^2(y^1)}{\partial y^1}\right) + \frac{\lambda^{2p}}{\gamma^1}\left(\frac{\partial R^1(y^1)}{\partial y^1} - \frac{\partial \widehat{R}^2(y^1)}{\partial y^1}\right).$$

To interpret (H38) in terms of the properties of the implementing tax function, consider the individual optimization problem for type-1 agents under a nonlinear income tax $T(y)$.

This can be described as follows:

(H39) $$\max_{e_s^1, e_q^1} \quad \theta^1 h\left(e_s^1, e_q^1\right) - T\left(\theta^1 h\left(e_s^1, e_q^1\right)\right) - p_s e_s^1 - p_q^1 e_q^1.$$

Once again, this optimization problem can be equivalently reformulated as a two-stage problem. In the first stage, for a given amount of production y and consumption $y - T(y)$, type-1 agents choose the effort mix that minimizes production costs, i.e. the effort mix that satisfies the condition $\frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} = \frac{p_s}{p_q^1}$ (where $h_1\left(e_s^1, e_q^1\right) \equiv \frac{\partial h\left(e_s^1, e_q^1\right)}{\partial e_s^1}$ and $h_2\left(e_s^1, e_q^1\right) \equiv \frac{\partial h\left(e_s^1, e_q^1\right)}{\partial e_q^1}$). This gives a conditional indirect utility function

(H40) $$V^1\left(y, y - T(y)\right) = y - T(y) - R(y),$$

where $R(y) = p_s e_s^1(y) - p_q^1 e_q^1(y)$. At the second stage y is optimally chosen subject to the link between pre-tax earnings and post-tax earnings determined by the tax schedule $T(y)$. The first order condition for this problem is given by

(H41) $$1 - T'(y) - \frac{\partial R(y)}{\partial y} = 0,$$

from which we can derive the following implicit characterization of the marginal income tax rate faced by type-1 agents:

(H42) $$T'(y) = 1 - \frac{\partial R(y)}{\partial y}.$$

Thus, combining (H38) and (H42), we can conclude that

(H43) $$T'\left(y^1\right) = \frac{\lambda^{2s}}{\gamma^1}\left(\frac{\partial R^1\left(y^1\right)}{\partial y^1} - \frac{\partial \widetilde{R}^2\left(y^1\right)}{\partial y^1}\right) + \frac{\lambda^{2p}}{\gamma^1}\left(\frac{\partial R^1\left(y^1\right)}{\partial y^1} - \frac{\partial \widehat{R}^2\left(y^1\right)}{\partial y^1}\right).$$

To shed light on the sign of $T'\left(y^1\right)$, consider first $\frac{\partial R^1\left(y^1\right)}{\partial y^1}$. When the tax liability is only a function of earned income we know that, when not behaving as mimickers, type-1 agents choose an undistorted (efficient) effort mix. This means that

(H44) $$\frac{\partial R^1\left(y^1\right)}{\partial y^1} = \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)} = \frac{p_q^1}{\theta^1 h_2\left(e_s^1, e_q^1\right)},$$

where subscripts on h are again used to denote partial derivatives.

Consider now $\frac{\partial \widetilde{R}^2\left(y^1\right)}{\partial y^1}$. Two possibilities should separately be considered: i) in order to earn y^1 while being remunerated according to their true productivity θ^2, type-2 agents are not forced to choose a distorted effort mix; ii) in order to earn y^1 while being remunerated according to their true productivity θ^2, type-2 agents are forced to choose a distorted effort mix.

Under case i) the effort mix chosen by a type-2 mimicker, denoted by $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$, satisfies the condition

$$h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)/h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) = p_s/p_q^2 \tag{H45}$$

and therefore

$$\frac{\partial \widetilde{R}^2\left(y^1\right)}{\partial y^1} = \frac{p_s}{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)} = \frac{p_q^2}{\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}, \tag{H46}$$

implying that

$$\frac{\partial R^1\left(y^1\right)}{\partial y^1} - \frac{\partial \widetilde{R}^2\left(y^1\right)}{\partial y^1} = \left(\frac{1}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - \frac{1}{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}\right)p_s. \tag{H47}$$

Notice that the sign of (H47) is opposite to the sign of $\frac{\partial(\theta h_1)}{\partial \theta} + \frac{\partial(\theta h_1)}{\partial p_q}\frac{dp_q}{d\theta}$. Furthermore, we have that

$$\frac{\partial\left(\theta h_1\right)}{\partial \theta} + \frac{\partial\left(\theta h_1\right)}{\partial p_q}\frac{dp_q}{d\theta} = h_1 + \theta\left[h_{11}\left(\frac{de_s}{d\theta} + \frac{de_s}{dp_q}\frac{dp_q}{d\theta}\right) + h_{12}\left(\frac{de_q}{d\theta} + \frac{de_q}{dp_q}\frac{dp_q}{d\theta}\right)\right]. \tag{H48}$$

For a given y, an undistorted effort mix solves the system of equations:

$$\theta h\left(e_s, e_q\right) = y, \tag{H49}$$

$$p_q h_1\left(e_s, e_q\right) - p_s h_2\left(e_s, e_q\right) = 0. \tag{H50}$$

Differentiating (H49)-(H50) with respect to e_s, e_q and θ gives, in matrix form

$$\begin{bmatrix} \theta h_1\left(e_s, e_q\right) & \theta h_2\left(e_s, e_q\right) \\ p_q h_{11}\left(e_s, e_q\right) - p_s h_{12}\left(e_s, e_q\right) & p_q h_{12}\left(e_s, e_q\right) - p_s h_{22}\left(e_s, e_q\right)\end{bmatrix}\begin{bmatrix} de_s/d\theta \\ de_q/d\theta\end{bmatrix} = \begin{bmatrix} -h\left(e_s, e_q\right) \\ 0\end{bmatrix},$$

from which one obtains

$$\frac{de_s}{d\theta} = -h\frac{p_q h_{12} - p_s h_{22}}{\Gamma}, \tag{H51}$$

$$\frac{de_q}{d\theta} = h\frac{p_q h_{11} - p_s h_{12}}{\Gamma}, \tag{H52}$$

where $\Gamma \equiv \theta\left\{h_1\left[p_q h_{12} - p_s h_{22}\right] - h_2\left[p_q h_{11} - p_s h_{12}\right]\right\} > 0$.
Differentiating (H49)-(H50) with respect to e_s, e_q and p_q gives, in matrix form

$$\begin{bmatrix} \theta h_1\left(e_s, e_q\right) & \theta h_2\left(e_s, e_q\right) \\ p_q h_{11}\left(e_s, e_q\right) - p_s h_{12}\left(e_s, e_q\right) & p_q h_{12}\left(e_s, e_q\right) - p_s h_{22}\left(e_s, e_q\right)\end{bmatrix}\begin{bmatrix} de_s/dp_q \\ de_q/dp_q\end{bmatrix} = \begin{bmatrix} 0 \\ -h_1\left(e_s, e_q\right)\end{bmatrix},$$

from which one obtains

$$\text{(H53)} \qquad \frac{de_s}{dp_q} = \frac{\theta h_1 h_2}{\Gamma},$$

$$\text{(H54)} \qquad \frac{de_q}{dp_q} = -\frac{\theta (h_1)^2}{\Gamma}.$$

Plugging (H51)-(H54) into (H48) gives

$$\frac{\partial (\theta h_1)}{\partial \theta} + \frac{\partial (\theta h_1)}{\partial p_q} \frac{dp_q}{d\theta}$$

$$= h_1 + \theta \left[h_{11} \left(-h \frac{p_q h_{12} - p_s h_{22}}{\Gamma} + \frac{\theta h_1 h_2}{\Gamma} \frac{dp_q}{d\theta} \right) + h_{12} \left(h \frac{p_q h_{11} - p_s h_{12}}{\Gamma} - \frac{\theta (h_1)^2}{\Gamma} \frac{dp_q}{d\theta} \right) \right]$$

$$= h_1 + \frac{\theta}{\Gamma} \left[-p_q h h_{11} h_{12} + p_s h h_{11} h_{22} + p_q h h_{12} h_{11} - p_s h h_{12} h_{12} + \theta h_1 h_2 h_{11} \frac{dp_q}{d\theta} - \theta (h_1)^2 h_{12} \frac{dp_q}{d\theta} \right]$$

(H55)

$$= h_1 + \frac{\theta}{\Gamma} \left[(h_{11} h_{22} - h_{12} h_{12}) p_s h + (h_2 h_{11} - h_1 h_{12}) \theta h_1 \frac{dp_q}{d\theta} \right].$$

Given that $h_{11} h_{22} - h_{12} h_{12} > 0$ (by concavity of the h-function) and $\frac{dp_q}{d\theta} < 0$, we can conclude that

$$\text{(H56)} \qquad \frac{\partial (\theta h_1)}{\partial \theta} + \frac{\partial (\theta h_1)}{\partial p_q} \frac{dp_q}{d\theta} > 0,$$

which in turn implies that $\frac{1}{\theta^1 h_1 \left(e_s^1, e_q^1\right)} > \frac{1}{\theta^2 h_1 \left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}$ in (H47).

Under case ii) the effort mix $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$ chosen by a type-2 mimicker is by assumption distorted (i.e., it violates the condition $h_1 \left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) / h_2 \left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) = p_s / p_q^2$) and can be obtained as a solution to the following system of equations:

$$\text{(H57)} \qquad p_s \widetilde{e}_s^2 + p_q^1 \widetilde{e}_q^2 = p_s e_s^1 + p_q^1 e_q^1,$$

$$\text{(H58)} \qquad \theta^2 h \left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) = y^1,$$

where $p_s e_s^1 + p_q^1 e_q^1$ represents the total costs incurred by type-1 agents to earn y^1 when abiding by the efficiency condition $h_1 / h_2 = p_s / p_q^1$ and being remunerated according to their true productivity θ^1. For a concave h-function there will be two pairs $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$ that solve the system (H57)-(H58): one pair that lies north-west of $\left(e_s^1, e_q^1\right)$ and one pair that lies south-east of $\left(e_s^1, e_q^1\right)$. Given that $p_q^2 < p_q^1$ and that both pairs $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$ lie on the same iso-cost line, pertaining to type-1 agents, with slope $-p_s / p_q^1$, it follows that the least costly pair for a type-2 mimicker will be the one lying north-west of $\left(e_s^1, e_q^1\right)$. Thus, $\widetilde{e}_s^2 < e_s^1$ and $\widetilde{e}_q^2 > e_q^1$; moreover, at the relevant $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$-pair, the effort mix will be distorted towards e_q, i.e. we will have that $h_1 / h_2 > p_s / p_q^2$.

Taking into account that $d\left(p_s e_s^1 + p_q^1 e_q^1\right)/dy^1 = dR^1(y^1)/dy^1 = p_s/\left[\theta^1 h_1\left(e_s^1, e_q^1\right)\right]$, differentiating (H57)-(H58) with respect to $\widetilde{e}_s^2$, $\widetilde{e}_q^2$ and y^1 gives, in matrix form:

$$\text{(H59)} \qquad \begin{bmatrix} p_s & p_q^1 \\ \theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) & \theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) \end{bmatrix} \begin{bmatrix} d\widetilde{e}_s^2/dy^1 \\ d\widetilde{e}_q^2/dy^1 \end{bmatrix} = \begin{bmatrix} \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)} \\ 1 \end{bmatrix}.$$

Defining Ψ as

$$\text{(H60)} \qquad \Psi \equiv \left[p_s h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) - p_q^1 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)\right]\theta^2,$$

we have that

$$\text{(H61)} \qquad \frac{d\widetilde{e}_s^2}{dy^1} = \frac{\frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) - p_q^1}{\Psi},$$

$$\text{(H62)} \qquad \frac{d\widetilde{e}_q^2}{dy^1} = \frac{p_s - \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\Psi},$$

and therefore

$$\frac{\partial \widetilde{R}^2\left(y^1\right)}{\partial y^1} = p_s \frac{d\widetilde{e}_s^2}{dy^1} + p_q^2 \frac{d\widetilde{e}_q^2}{dy^1} = p_s \frac{\frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) - p_q^1}{\Psi} + p_q^2 \frac{p_s - \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\Psi}$$

$$\text{(H63)} \qquad = \frac{\left(p_q^2 - p_q^1\right)p_s + \left(p_s\right)^2 \frac{\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - p_s p_q^2 \frac{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)}}{\Psi}.$$

Notice that, since $h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)/h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) > p_s/p_q^2$, we have that $\Psi < 0$.

The difference $\frac{\partial R^1(y^1)}{\partial y^1} - \frac{\partial \widetilde{R}^2(y^1)}{\partial y^1}$ is then given by

$$\begin{aligned}
\frac{\partial R^1(y^1)}{\partial y^1} - \frac{\partial \widetilde{R}^2(y^1)}{\partial y^1} &= \frac{p_s}{\theta^1 h_1(e_s^1, e_q^1)} - \frac{\left(p_q^2 - p_q^1\right) p_s + (p_s)^2 \frac{\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - p_s p_q^2 \frac{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)}}{\Psi} \\
&= \frac{\Psi - \left(p_q^2 - p_q^1\right)\theta^1 h_1\left(e_s^1, e_q^1\right) - \left[p_s \frac{\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - p_q^2 \frac{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\right]\theta^1 h_1\left(e_s^1, e_q^1\right)}{\Psi \theta^1 h_1\left(e_s^1, e_q^1\right)} p_s \\
&= \frac{\left[p_s h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) - p_q^1 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)\right]\theta^2 - \left(p_q^2 - p_q^1\right)\theta^1 h_1\left(e_s^1, e_q^1\right)}{\Psi \theta^1 h_1\left(e_s^1, e_q^1\right)} p_s \\
&\quad - \frac{\left[p_s \frac{\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - p_q^2 \frac{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\right]\theta^1 h_1\left(e_s^1, e_q^1\right)}{\Psi \theta^1 h_1\left(e_s^1, e_q^1\right)} p_s \\
&= \frac{p_q^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)\theta^2 - p_q^1 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)\theta^2 - \left(p_q^2 - p_q^1\right)\theta^1 h_1\left(e_s^1, e_q^1\right)}{\Psi \theta^1 h_1\left(e_s^1, e_q^1\right)} p_s \\
\text{(H64)} \qquad &= \left[\frac{h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{h_1\left(e_s^1, e_q^1\right)} \frac{\theta^2}{\theta^1} - 1\right]\left(p_q^2 - p_q^1\right)\frac{p_s}{\Psi}.
\end{aligned}$$

Since $p_q^2 - p_q^1 < 0$ and $\Psi < 0$, we have that

$$\text{(H65)} \qquad sign\left\{\frac{\partial R^1(y^1)}{\partial y^1} - \frac{\partial \widetilde{R}^2(y^1)}{\partial y^1}\right\} = sign\left\{\frac{h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{h_1\left(e_s^1, e_q^1\right)} \frac{\theta^2}{\theta^1} - 1\right\}.$$

Given that, as we have previously noticed, $\widetilde{e}_s^2 < e_s^1$ and $\widetilde{e}_q^2 > e_q^1$, it follows that $h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) > h_1\left(e_s^1, e_q^1\right)$, guaranteeing that $\frac{h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{h_1\left(e_s^1, e_q^1\right)} \frac{\theta^2}{\theta^1} > 1$. Thus, as we had shown for case i), we can once again establish that $\frac{\partial R^1(y^1)}{\partial y^1} - \frac{\partial \widetilde{R}^2(y^1)}{\partial y^1} > 0$.

Now consider the term $\frac{\partial \widehat{R}^2(y^1)}{\partial y^1}$ appearing in (H43). Even for this term one should in principle distinguish two scenarios. Under the first, in order to earn y^1 while being remunerated according to the average productivity $\overline{\theta}$, a type-2 agent is not forced to choose a distorted effort mix. Under the second scenario, in order to earn y^1 while being remunerated according to the average productivity $\overline{\theta}$, a type-2 agent is forced to choose a distorted effort mix. However, it is easy to show that the second scenario can be safely neglected for the purposes of our analysis. The reason is that, if it is indeed the case that, in order to earn y^1 while being remunerated according to the average productivity $\overline{\theta}$, type-2 agents are forced to choose a distorted effort mix, it necessarily follows that there is another, more attractive, deviating strategy available to them. To understand this point, consider first the system of equations that determine $\widehat{e}_s$ and $\widehat{e}_q$ under the assumption that, in order to earn y^1 while being remunerated according to the average productivity $\overline{\theta}$, a type-2 agent is forced to choose a distorted effort mix:

$$p_s\widehat{e}_s + p_q^1\widehat{e}_q = p_s e_s^1 + p_q^1 e_q^1 \tag{H66}$$

$$\overline{\theta}h\left(\widehat{e}_s, \widehat{e}_q\right) = y^1, \tag{H67}$$

where $p_s e_s^1 + p_q^1 e_q^1$ represents the total costs incurred by type-1 agents to earn y^1 when abiding by the efficiency condition $h_1/h_2 = p_s/p_q^1$ and being remunerated according to their true productivity θ^1. For a concave h-function there will be two pairs $(\widehat{e}_s, \widehat{e}_q)$ that solve the system (H66)-(H67): one pair that lies north-west of $\left(e_s^1, e_q^1\right)$ and one pair that lies south-east of $\left(e_s^1, e_q^1\right)$. Given that $p_q^2 < p_q^1$ and that both pairs $(\widehat{e}_s, \widehat{e}_q)$ lie on the same iso-cost line, pertaining to type-1 agents, with slope $-p_s/p_q^1$, it follows that the least costly pair for a type-2 mimicker will be the one lying north-west of $\left(e_s^1, e_q^1\right)$.

Now consider again eqs. (H57)-(H58), i.e. the system of equations that determine $\widetilde{e}_s^2$ and $\widetilde{e}_q^2$ under the assumption that, in order to earn y^1 while being remunerated according to their true productivity θ^2, type-2 agents are forced to choose a distorted effort mix. Given that $\theta^2 > \overline{\theta}$, the isoquant described by (H58) lies strictly below the isoquant described by (H67). Therefore, since both $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$ and $(\widehat{e}_s, \widehat{e}_q)$ lie on the same iso-cost line, pertaining to type-1 agents, with slope $-p_s/p_q^1$, it must be that $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$ lies north-west of $(\widehat{e}_s, \widehat{e}_q)$: $\widetilde{e}_s^2 < \widehat{e}_s$, $\widetilde{e}_q^2 > \widehat{e}_q$. But then, the fact that $p_q^2 < p_q^1$ (implying that the iso-cost lines pertaining to type-2 agents have slope $-p_s/p_q^2 < -p_s/p_q^1$) implies that, necessarily, choosing $\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)$ represents for a type-2 mimicker a more attractive deviating strategy than choosing $(\widehat{e}_s, \widehat{e}_q)$.

As a consequence of the above discussion, a necessary condition for the deviating strategy with associated cost $\widehat{R}^2\left(y^1\right)$ to be the least costly mimicking strategy for a type-2 agent is that, in order to earn y^1 while being remunerated according to the average productivity $\overline{\theta}$, type-2 agents are not forced to choose a distorted effort mix. Put differently, a necessary condition for $\lambda^{2p} > 0$ is that $\frac{\partial \widehat{R}^2\left(y^1\right)}{\partial y^1} = \frac{p_s}{\overline{\theta}h_1(\widehat{e}_s, \widehat{e}_q)} = \frac{p_q^2}{\overline{\theta}h_2(\widehat{e}_s, \widehat{e}_q)}$. It then follows that, when $\lambda^{2p} > 0$, it must necessarily be that

$$\frac{\partial R^1\left(y^1\right)}{\partial y^1} - \frac{\partial \widehat{R}^2\left(y^1\right)}{\partial y^1} = \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - \frac{p_s}{\overline{\theta}h_1\left(\widehat{e}_s, \widehat{e}_q\right)} = \left(\frac{1}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - \frac{1}{\overline{\theta}h_1\left(\widehat{e}_s, \widehat{e}_q\right)}\right) p_s. \tag{H68}$$

Noticing that the sign of (H68) is opposite to the sign of $\frac{\partial(\theta h_1)}{\partial \theta} + \frac{\partial(\theta h_1)}{\partial p_q}\frac{dp_q}{d\theta}$, we can again rely on (H56) to conclude that $\frac{\partial R^1\left(y^1\right)}{\partial y^1} - \frac{\partial \widehat{R}^2\left(y^1\right)}{\partial y^1} > 0$.

Going back to (H43) and summarizing our results for $T'\left(y^1\right)$, we have that $T'\left(y^1\right)$ is necessarily positive (given that our max-min social welfare function implies that at least one of the two downward IC-constraints, with associated multipliers λ^{2s} and λ^{2p}, will be binding). Notice also that, if in order to earn y^1 while being remunerated according to their true productivity θ^2, type-2 agents are not forced to choose a distorted effort mix, it must necessarily be that $\lambda^{2s} > 0$ and $\lambda^{2p} = 0$ (earning y^1 while being remunerated according to the average productivity $\overline{\theta}$ is necessarily a dominated deviating strategy for type-2 agents). Thus, four different scenarios are conceivable. Under the first, $\lambda^{2s} > 0$, $\lambda^{2p} = 0$ and

$$T'\left(y^1\right) = \frac{\lambda^{2s}}{\gamma^1}\left(\frac{1}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - \frac{1}{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}\right) p_s > 0. \tag{H69}$$

Under the second scenario we still have that $\lambda^{2s} > 0$, $\lambda^{2p} = 0$ but this time we have that

$$T'\left(y^1\right) = \frac{\lambda^{2s}}{\gamma^1}\left(\frac{h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)\theta^2}{h_1\left(e_s^1, e_q^1\right)\theta^1} - 1\right)\left(p_q^2 - p_q^1\right)\frac{p_s}{\Psi} > 0. \tag{H70}$$

Under the third scenario $\lambda^{2s} > 0$ and $\lambda^{2p} > 0$; in this case we have that
(H71)

$$T'\left(y^1\right) = \left[\lambda^{2s}\left(\frac{h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)\theta^2}{h_1\left(e_s^1, e_q^1\right)\theta^1} - 1\right)\frac{p_q^2 - p_q^1}{\Psi} + \lambda^{2p}\left(\frac{1}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - \frac{1}{\overline{\theta} h_1\left(\widehat{e}_s, \widehat{e}_q\right)}\right)\right]\frac{p_s}{\gamma^1} > 0.$$

Under the last scenario $\lambda^{2s} = 0$ and $\lambda^{2p} > 0$; in this case we have that

$$T'\left(y^1\right) = \frac{\lambda^{2p}}{\gamma^1}\left(\frac{1}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - \frac{1}{\overline{\theta} h_1\left(\widehat{e}_s, \widehat{e}_q\right)}\right) p_s > 0. \tag{H72}$$

I The case when both signals are observable

We turn first to formulate the maximization program associated with a MMO given by a pooling tax equilibrium. Without loss of generality, we will assume that $p_s = 1$. The pooling tax equilibrium is given by the triplet (c, e_s, e_q) which solves the following maximization problem:

$$\max_{c, e_s, e_q}\left[c - \left(e_s + e_q p_q^1\right)\right], \tag{I1}$$

where

$$c = \bar{\theta} h(e_s, e_q) \tag{I2}$$

$$\bar{\theta} = \gamma^1\theta^1 + \gamma^2\theta^2. \tag{I3}$$

That is, type-1 utility is maximized by choosing effort levels (quantity and quality) subject to the constraints that workers are compensated based on average productivity and zero tax revenues are being collected.

We turn next to formulate the maximization program associated with a constrained efficient allocation given by a separating tax equilibrium. The separating tax equilibrium is given by the two triplets (c^1, e_s^1, e_q^1) and (c^2, e_s^2, e_q^2), which solve the following constrained maximization problem:

(I4) $$\max_{\{(c^i, e_s^i, e_q^i)\}_{i=1,2}} \left[c^1 - (e_s^1 + e_q^1 p_q^1)\right]$$

subject to:

(I5) $$c^1 - (e_s^1 + e_q^1 p_q^1) \geq c^2 - (e_s^2 + e_q^2 p_q^1)$$

(I6) $$c^2 - (e_s^2 + e_q^2 p_q^2) \geq c^1 - (e_s^1 + e_q^1 p_q^2)$$

(I7) $$\gamma^1 \theta^1 h(e_s^1, e_q^1) + \gamma^2 \theta^2 h(e_s^2, e_q^2) \geq \gamma^1 c^1 + \gamma^2 c^2$$

In the separating regime, the utility of type 1 is maximized by offering two different bundles, where types are compensated according to their productivity, and the redistribution is limited to the income channel. Note that the IC-constraints actually only consider mimicking by replication, not (infeasible) off-equilibrium deviations. As $p_q^1 > p_q^2 > 0$, the single-crossing property holds. Thus, as we are considering a Rawlsian welfare function, the only binding IC-constraint is the one associated with the high-skilled (type-2) individual.

Next, we show that pooling is suboptimal, i.e., the socially optimal configuration will be a separating allocation.

Proposition 7. *Assuming that both signals are taxed, the pooling equilibrium is suboptimal and thus predistribution is socially undesirable. Moreover, social welfare is strictly higher compared to the case where only one signal is taxed.*

Proof. Let the triplet (e_s^*, e_q^*, c^*) denote the (presumably) socially optimal pooling allocation, and consider the following alternative separating allocation, obtained as a small perturbation of the pooling allocation and given by the two triplets: (c^1, e_s^1, e_q^1) and (c^2, e_s^2, e_q^2) where $e_s^1 = e_s^* - \varepsilon, e_q^1 = e_q^* - \varepsilon, c^1 = c^* - \varepsilon(1 + p_q^1)$, where $\varepsilon > 0$ and small; and where $e_s^2 = e_s^* + \delta, e_q^2 = e_q^* + \delta, c^2 = c^* + \delta(1 + p_q^2)$, with $\delta > 0$ and small. It is easy to check that since $p_q^1 > p_q^2$, the perturbed allocation is incentive compatible, and that it preserves the utility level of both types as in the pooling allocation.

Invoking a first-order approximation and following some algebraic manipulations, the total effect of the perturbation on the aggregate output (ΔY) is given by:

(I8) $$\Delta Y = [\gamma^2 \theta^2 \delta - \gamma^1 \theta^1 \varepsilon][h_1(e_s^*, e_q^*) + h_2(e_s^*, e_q^*)]$$

where $h_j, j = 1, 2$, denote the partial derivatives with respect to the first and second arguments of $h(\cdot)$.

The corresponding total effect of the perturbation on the aggregate consumption (ΔC)

is given by:

$$\Delta C = \gamma^2 \delta(1 + p_q^2) - \gamma^1 \varepsilon(1 + p_q^1) \tag{I9}$$

Thus,

$$\Delta Y - \Delta C = \gamma^2 \delta \left[\theta^2 [h_1(e_s^*, e_q^*) + h_2(e_s^*, e_q^*)] - (1 + p_q^2)\right] - \gamma^1 \varepsilon \left[\theta^1 [h_1(e_s^*, e_q^*) + h_2(e_s^*, e_q^*)] - (1 + p_q^1)\right] \tag{I10}$$

Suppose now that $\kappa \equiv \gamma^2 \delta = \gamma^1 \varepsilon$. It follows that:

$$\Delta Y - \Delta C = \kappa \left[(\theta^2 - \theta^1)[h_1(e_s^*, e_q^*) + h_2(e_s^*, e_q^*)] + (p_q^1 - p_q^2)\right] > 0, \tag{I11}$$

where the inequality sign follows as $h_1 > 0, h_2 > 0, \kappa > 0, \theta^2 > \theta^1$, and $p_q^1 > p_q^2$.

The resulting tax surplus can be refunded as a lump sum transfer, which increases the utility of both types relative to the pooling allocation without violating the IC-constraints. We have thus obtained a contradiction to the presumed optimality of the pooling allocation as needed.

The fact that social welfare is strictly higher relative to the case where only one signal is taxed follows from the following three observations: (i) the social optimum when both signals are taxed is (always) given by a separating allocation, as just shown, (ii) the downward IC-constraint is tightened (only replication is allowed) relative to the case where only the quantity signal is taxed, and, (iii) the upward IC-constraint never binds (it may bind when only the quantity signal is taxed). This concludes the proof. □

As anticipated, the ability to tax both signals serves to enhance redistribution. However, the interesting insight is that predistribution becomes suboptimal in contrast to the case where only the quantity signal is subject to taxation. If the government has the full capacity to tax both signals, then the elimination of the information rent associated with the difference in productivity between types can be achieved through the separating allocation and does not require the implementation of a pooling allocation. This is a more efficient way to achieve this goal and improve redistribution. The feasibility of predistribution depends on the ability to tax signals, but the social desirability of its use depends critically on the limited ability to tax all signals. Predistribution, which involves large inefficiencies, compensates for the inability to tax the quality signal directly.

Note that unlike the standard (ABC) optimal tax formulas, which usually depend on the skill distribution, the optimal marginal tax rates for the type-1 bundle do not depend on the productivity difference between types, since both signals can be taxed directly to eliminate the information rent of the type-2 bundle (which is undistorted, since the type-1 IC constraint is not binding in the optimal solution). In particular, it can be shown (details available upon request) that the optimal wedge on the effort mix chosen by type-1 agents is given by

$$\text{(I12)} \qquad \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} - \frac{p_s}{p_q^1} = -\frac{\gamma^2\left(p_q^1 - p_q^2\right)}{\gamma^1 p_q^1 + \gamma^2\left(p_q^1 - p_q^2\right)} p_s < 0.$$

J The observable signal is e_q instead of e_s

We divide this Appendix in two parts. In part (a) we provide an intuition for the result that the socially optimal separating tax equilibrium is not invariant to the assumption about which of the two signals is observable by the government. In part (b) we provide an intuition for the result that it is a priori ambiguous in which direction it is optimal to distort the effort-mix of type-1 agents under the socially optimal separating tax equilibrium.

Part (a) Consider the right-hand side of the (downward) IC-constraint (23), which gives the utility attainable by type-2 agents if they behave as mimickers. The constraint shows that, as mimickers, type-2 agents do not need to replicate all the effort choices of type-1 agents; they only need to replicate e_s^1, which is the choice of type-1 agents along the effort-dimension that is observable by the government. The choice of type-2 mimickers along the other effort-dimension, $\widehat{e}_q^2$, is given by the value of e_q that satisfies the equation $\overline{\theta} h\left(e_s^1, e_q\right) = y^1$ (see (24)), which implies that $\widehat{e}_q^2 < e_q^1$ (given that the isoquant $\overline{\theta} h\left(e_s^1, e_q\right) = y^1$ is strictly below the isoquant $\theta^1 h\left(e_s^1, e_q\right) = y^1$). In a setting where the signal observed by the government is e_q instead of e_s, type-2 mimickers would have instead to replicate e_q^1, while $\widehat{e}_s^2$ would be given by the value of e_s that solves the equation $\overline{\theta} h\left(e_s, e_q^1\right) = y^1$ (implying that $\widehat{e}_s^2 < e_s^1$).

Thus, for a given quadruplet $\left(y^1, c^1, e_s^1, e_q^1\right)$ intended for type-1 agents, the utility achievable by type-2 agents will in general differ depending on whether the signal observable by the government is e_s or e_q. This implies the following two possibilities. i) The allocation $\left\{\left(y^1, c^1, e_s^1, e_q^1\right), \left(y^2, c^2, e_s^2, e_q^2\right)\right\}$, which is the socially optimal separating tax equilibrium in a setting where the signal observed by the government is e_s, is not feasible, because it violates the downward IC-constraint in a setting where the observed signal is e_q; ii) the allocation $\left\{\left(y^1, c^1, e_s^1, e_q^1\right), \left(y^2, c^2, e_s^2, e_q^2\right)\right\}$ which represents the socially optimal separating tax equilibrium in a setting where the signal observed by the government is e_s is also feasible when the observed signal is e_q, but does not represent the socially optimal separating tax equilibrium in the latter setting (because the downward IC-constraint is slack).

Part (b) Consider the following. For a given isoquant $\theta^1 h\left(e_s, e_q\right) = y^1$, assume that type-1 agents are induced to choose the effort mix $\left(e_s^1, e_q^1\right)$ that satisfies the no-distortion condition $\frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} = \frac{p_s}{p_q^1}$. In a setting where the observed signal is e_q, a type-2 mimicker must choose $\widehat{e}_q^2 = e_q^1$, while $\widehat{e}_s^2$ is set to satisfy the equation $\overline{\theta} h\left(e_s, e_q^1\right) = y^1$, (which implies that $\widehat{e}_s^2 < e_s^1$). Given that $p_q^2 < p_q^1$, but at the same time $\frac{h_1\left(\widehat{e}_s^2, \widehat{e}_q^2\right)}{h_2\left(\widehat{e}_s^2, \widehat{e}_q^2\right)} > \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)}$, it follows that one cannot a priori establish whether the effort-mix of type-2 mimickers is distorted towards e_s (i.e., $\frac{h_1\left(\widehat{e}_s^2, \widehat{e}_q^2\right)}{h_2\left(\widehat{e}_s^2, \widehat{e}_q^2\right)} < \frac{p_s}{p_q^2}$) or towards e_q (i.e., $\frac{h_1\left(\widehat{e}_s^2, \widehat{e}_q^2\right)}{h_2\left(\widehat{e}_s^2, \widehat{e}_q^2\right)} > \frac{p_s}{p_q^2}$).

This ambiguity is essentially the reason why it is not possible to determine once and for all in which direction it is desirable to distort the effort mix chosen by type-1 agents. Suppose for instance that it is indeed the case that when type-1 agents choose an undistorted effort mix, the effort mix chosen by type-2 mimickers is distorted in the direction of e_q. Then it will be welfare enhancing to induce type-1 agents to choose an effort mix that is slightly distorted towards e_q. If the distortion is small, it will have only a second-order effect on the total costs $(p_s e_s^1 + p_q^1 e_q^1)$ borne by type-1 agents; but it will have a first-order negative effect on type-2 mimickers, increasing the total cost $p_s \widehat{e}_s^1 + p_q^2 e_q^1$ (because the initial distortion in their effort mix is exacerbated).

Finally, note that distorting the effort mix chosen by type-1 agents in the direction of e_q is more likely to be desirable when the difference $p_q^1 - p_q^2$ is relatively small and the ratio h_1/h_2 increases rapidly when lowering e_s (for given e_q).

K The welfare gains from predistribution

This section uses the functional form in equation (26) to illustrate the welfare gains from predistribution. We do this by setting up the constrained nonlinear optimization problem faced by the government using AMPL and solving it using the state-of-the-art nonlinear optimization package, KNITRO.

K.1 The income tax regime

Given that, as we explained in Appendix H.2, the incentives underlying the decision problem of type-2 agents when they are not acting as mimickers are aligned with the incentives underlying the social decision problem, the government's problem can be equivalently reformulated as follows:

(K1)
$$\max_{c^1, c^2, y^1, y^2, e_s^2} c^1 - R^1\left(y^1\right)$$

subject to the budget constraint

(K2)
$$\sum_i \gamma^i (y^i - c^i) = 0,$$

the upward IC-constraint

$$c^1 - R^1\left(y^1\right) \geq c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^1}{e_s^2}, \tag{K3}$$

and the downward IC-constraints

$$c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \geq c^1 - \widetilde{R}^2(y^1), \tag{K4}$$

$$c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \geq c^1 - \hat{R}^2(y^1). \tag{K5}$$

In the reformulated version of the government problem, we have included e_s^2 as an artificial control variable for the government; for this reason, we have also explicitly included the upward IC-constraint in the government problem. Note also that we used assumption (26) to express e_q^2 as a function of y^2 and e_s^2, namely $e_q^2 = \left(y^2/\theta^2\right)^{\frac{1}{\beta}}/e_s^2$.

Exploiting the assumption (26) also allows us to obtain closed-form expressions for the government's objective function and the right hand side of the incentive constraints (K4)–(K5). To achieve this goal, we begin by deriving the effort costs incurred by agents who choose the point on the income tax schedule intended for them.

Choices of a truthfully reporting agent of type 1 Consider agents of type 1 who earn the income level y^1 that the government intends for them. They will choose an efficient mix of e_s and e_q and solve:

$$\min_{e_s, e_q} \; p_s e_s + p_q^1 e_q \quad \text{subject to} \quad \left(e_s e_q\right)^{\beta} \theta^1 = y^1. \tag{K6}$$

The optimal effort choices are given by

$$e_s^1\left(y^1\right) = \sqrt{\left(\frac{y^1}{\theta^1}\right)^{1/\beta} \frac{p_q^1}{p_s}} \quad \text{and} \quad e_q^1\left(y^1\right) = \sqrt{\left(\frac{y^1}{\theta^1}\right)^{1/\beta} \frac{p_s}{p_q^1}}. \tag{K7}$$

Inserting (K7) into the cost function yields

$$R^1(y^1) = p_s \sqrt{\left(\frac{y^1}{\theta^1}\right)^{1/\beta} \frac{p_q^1}{p_s}} + p_q^1 \sqrt{\left(\frac{y^1}{\theta^1}\right)^{1/\beta} \frac{p_s}{p_q^1}} = 2\sqrt{\left(\frac{y^1}{\theta^1}\right)^{1/\beta} p_s p_q^1}. \tag{K8}$$

Optimal deviating strategies for agents of type 2 Now consider the different strategies available to type-2 agents. There are three cases to consider, depending on which of the two constraints (K4)–(K5) is relevant.[40] These cases can be distinguished using conditions that depend on the ratio θ^2/θ^1, the relative size of the two groups (γ^1 and γ^2), and a constant defined as:

$$\Omega \equiv \left[\left(p_q^2 + p_q^1\right) / \left(2\sqrt{p_q^2 p_q^1}\right)\right]^{2\beta}. \tag{K9}$$

Case 1: $\theta^2/\theta^1 \leq \Omega$ In this case we have that $\min\left\{\widetilde{R}^2(y^1), \hat{R}^2(y^1)\right\} = \widetilde{R}^2(y^1)$, and therefore only constraint (K4) is relevant. The effort mix chosen by a type-2 mimicker under its optimal deviation strategy is undistorted (satisfies $e_q/e_s = p_s/p_q^2$) and $\widetilde{R}^2(y^1) = 2\sqrt{(y^1/\theta^2)^{\frac{1}{\beta}}\, p_q^2 p_s}$. Thus, the relevant downward IC-constraint can be expressed as:

$$c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \geq c^1 - 2\sqrt{(y^1/\theta^2)^{\frac{1}{\beta}}\, p_q^2 p_s}. \tag{K10}$$

Case 2: $\overline{\theta}/\theta^1 < \Omega < \theta^2/\theta^1$ In this case, we again have $\min\left\{\widetilde{R}^2(y^1), \hat{R}^2(y^1)\right\} = \widetilde{R}^2(y^1)$. This time, however, type-2 mimickers must choose a distorted effort mix ($e_q/e_s \neq p_s/p_q^2$) in order to achieve separation and be paid according to their true productivity. Thus, the relevant downward IC-constraint can be formulated as

$$c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \geq c^1 - \sqrt{\frac{p_s}{p_q^1}\,(y^1)^{\frac{1}{\beta}}\, \frac{\left(p_q^1 - p_q^2\right)\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} + 2p_q^2\sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}}}\left[\sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} - \left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}}}\right]}{\sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} - \left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}}}}}. \tag{K11}$$

Case 3: $\overline{\theta}/\theta^1 \geq \Omega$ In this case, it is not possible to determine unambiguously whether $\widetilde{R}^2(y^1) < \hat{R}^2(y^1)$, $\widetilde{R}^2(y^1) > \hat{R}^2(y^1)$, or $\widetilde{R}^2(y^1) = \hat{R}^2(y^1)$. What can be established is that the mimicking strategy with associated cost $\widetilde{R}^2(y^1)$ necessarily requires that a type 2 mimicker chooses a distorted effort mix. Thus, there are two relevant downward IC-constraints, one given by (K11) (the one associated with the cost $\widetilde{R}^2(y^1)$), and the other (associated with the cost $\hat{R}^2(y^1)$) given by:

$$c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \geq c^1 - 2\sqrt{\left(y^1/\overline{\theta}\right)^{1/\beta} p_s p_q^2}. \tag{K12}$$

[40] In our numerical example, we will vary the parameters so that all three cases are considered. The derivations needed to distinguish between the different cases are available on request.

K.2 Government problem, constrained efficient allocation

We start with the separating equilibrium. In this case, denoting by an asterisk symbol the effort choices of workers in equilibrium, and by a hat symbol the quality effort choice of a mimicking type-2 worker, it follows:

$$\text{(K13)} \qquad e_q^{i*} = \left(\frac{y^i}{\theta^i}\right)^{1/\beta} \frac{1}{e_s^i}, (i = 1, 2) \quad \text{and} \quad \hat{e}_q = \left(\frac{y^1}{\bar{\theta}}\right)^{1/\beta} \frac{1}{e_s^1}.$$

Thus, the IC-constraints (23)–(24) can be written as follows:

$$\text{(K14)} \qquad c^1 - p_s^1 e_s^1 - \left(\frac{y^1}{\theta^1}\right)^{1/\beta} \frac{p_q^1}{e_s^1} \geq c^2 - p_s^1 e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{1/\beta} \frac{p_q^1}{e_s^2},$$

$$\text{(K15)} \qquad c^2 - p_s^2 e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{1/\beta} \frac{p_q^2}{e_s^2} \geq c^1 - p_s^2 e_s^1 - \left(\frac{y^1}{\bar{\theta}}\right)^{1/\beta} \frac{p_q^2}{e_s^1}.$$

When implementing a pooling equilibrium, IC-constraints can be neglected, and the government chooses (y, e_s) to maximize

$$\text{(K16)} \qquad u^1 = y - p_s^1 e_s + p_q^1 \widehat{e}_q(y, e_s),$$

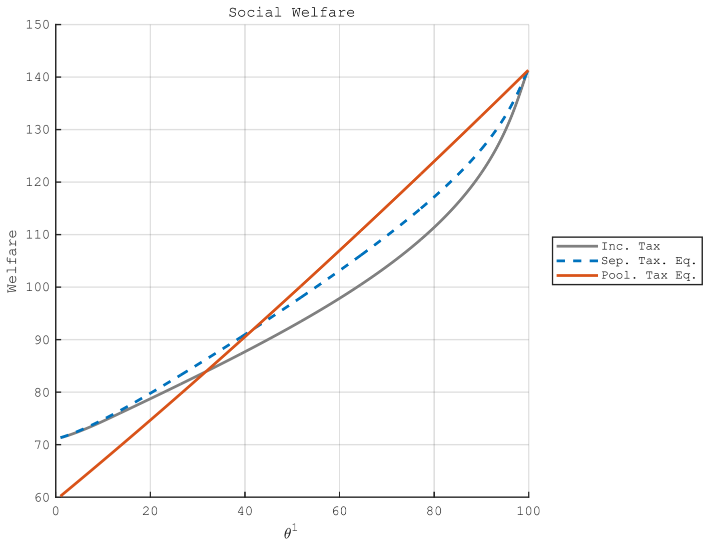
where $\widehat{e}_q(e_s, y)$ is the value of e_q which solves the equation $y = (e_s e_q)^\beta \bar{\theta}$.

K.3 Welfare gains

We fix type 2 productivity at $\theta^2 = 100$ and compute the social welfare level of the case with only an income tax and compare it to the social welfare level in the MMO, while letting θ^1 vary between 1 and 100. In this way, we consider a wide range of values for the ratio θ^1/θ^2. We keep the normalization $p_s^1 = p_s^2 = p_q^2 = 1$ and set $\beta = 0.10$, $\gamma^1 = \gamma^2 = 0.5$, $p_q^1 = 1.05$. We then compute the maximum achievable welfare gain from predistribution, which is achieved at the value of θ^1 at which the difference between the social welfare level in the MMO and the social welfare level in the income tax system is greatest. We express this maximum achievable welfare gain in equivalent-variation terms by first computing the minimum amount of resources that must be injected into the income-tax-only case in order to achieve the social welfare level of the MMO (by repeatedly solving the government's optimization program), and then dividing this number by the total output of the income-tax-only case to obtain a measure of the welfare gain expressed as a fraction of output.

Figure 2 shows the results. As expected, we see that the MMO (given by either an STE or a PTE, depending on which results in the highest social welfare) always dominates the case with only an income tax. We see that it is optimal to implement a separating allocation when θ^1 takes low and intermediate values, while the pooling allocation dominates when θ^1 is relatively close to θ^2. Notice that when θ^1 is very close to 100, the separating allocation dominates, although it is not visible in the figure. This is a knife edge case of no practical relevance. The maximum welfare gain from the MMO relative to the pure income tax regime is obtained at $\theta^1 = 75.1$, amounts to 12.44% of total output, and is associated with the implementation of a pooling allocation.

Figure 2: The welfare gains from predistribution



Note: Inc. Tax = case with only one income tax. Sep. Tax. Eq. = separating tax equilibrium, Pool. Tax. Eq. = pooling tax equilibrium. The constrained efficient allocation is given by the upper envelope of the dashed blue line and the solid orange line. The maximum welfare gain from predistribution occurs at $\theta^1 = 75.1$ and amounts to 12.44% of total output.