Optimal redistribution and education signaling

Spencer Bastani, Tomer Blumkin, Luca Micheletto

Discussion Paper No. 24-13

March 26, 2024

Monaster Center for
Economic Research
Ben-Gurion University of the Negev
P.O. Box 653
Beer Sheva, Israel

Fax: 972-8-6472941
Tel: 972-8-6472286

Optimal redistribution and education signaling*

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

March 26, 2024

Abstract

This paper studies optimal taxation of income and education when employers cannot observe workers' productivity and workers signal their productivity to firms by choosing both quantity and quality of education. We characterize constrained efficient allocations and derive conditions under which there is *predistribution*, i.e., redistribution through wage compression. Implementation through income and education dependent taxes is discussed, as well as education mandates. A key insight is that achieving predistribution requires complementing the income tax with additional policy instruments that regulate the flow of information in the labor market and prevent high-skilled individuals from separating themselves from their low-skilled counterparts.

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

JEL classification: D82, H21, H52, J31

*We are grateful to the editor, Dirk Krueger, and three anonymous reviewers for insightful comments. We are also grateful to Håkan Selin and Anna Sjögren for providing insightful comments on an earlier draft of the paper. Spencer Bastani gratefully acknowledges financial support from Jan Wallanders och Tom Hedelius stiftelse (grant P2018-162).

[†]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.

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

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

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. In scenarios involving extensive margin choices, such as migration and labor market participation, negative marginal tax rates may be desirable.

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, highlighting how firms screen workers based on their choices about working hours. Stantcheva (2014) examines the implications of adverse selection in the labor market for the design of optimal income taxation. She shows that firms' use of hours and compensation as screening tools can help governments achieve redistributive goals by mitigating the adverse response of high-skilled workers to progressive taxation. Bastani et al. (2015) adopt a similar screening framework, focusing on how governments can use the progressive income tax schedule to influence the wage distribution by implementing bunching or pooling of worker types.

The purpose of the current paper is to propose a framework for evaluating optimal redistributive policies in the presence of educational signaling, recognizing the possibility of redistribution through wage compression, which we refer to as *predistribution*. Despite the central role of signaling theory in economics, its place as a cornerstone of economics curricula around the world, and its presence in policy discussions (as evidenced, for example, by Caplan 2018), it is surprising that signaling has received the attention of only a handful of papers in the vast literature on the optimal design of tax systems that has followed the seminal work of Mirrlees (1971). By studying optimal taxation in the presence of education signaling, our paper seeks to make progress in filling this knowledge gap in the literature.

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 the vast majority of papers in the optimal income tax literature, and in line with the two studies discussed above, we assume that these abilities are also unobservable to potential employers. The distinguishing feature of our analysis is that workers must signal their type to firms by making costly effort choices. In addition, we allow information transmission between workers and firms to occur along more than one dimension, and we analyze constrained efficient allocations that combine taxes on income and taxes on the signaling activity that the government can observe in the labor market. As we show, these aspects are essential for understanding how to design optimal redistribution policies in economies with two levels of asymmetric information. In particular, they determine the feasibility and social desirability of redistribution via the wage channel (predistribution).[1]

While the current paper shares the feature of a second layer of asymmetric information with the studies mentioned above, it differs from them in at least five ways. First, we focus on worker signaling through investment in education.[2] Second, we consider multidimensional signaling in the context of 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.[3] 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. 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 under a Mirrleesian setup in which worker types are observable by firms.

Our analytical vehicle is a model that captures the equity-efficiency tradeoff in a manner similar to the two-type Stiglitz (1982) version of the Mirrlees (1971) optimal income tax model.

[1] Notably, predistribution is a phenomenon which occurs even when the production technology is linear and skill types are perfect substitutes, as in Mirrlees (1971). The mechanism behind wage compression in our signaling context is therefore different from the one emphasized in papers studying optimal income taxation in a general equilibrium context, where redistribution through the wage channel occurs due to sectoral reallocation of labor (see, for example, Stiglitz 1982, Rothschild and Scheuer 2013, and Sachs et al. 2020).

[2] By focusing on signaling, our study is related to recent work by Craig (2023), who 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, and Sztutman (2024), who studies optimal taxation in a dynamic job signaling model where the career profile of labor supply conveys information about worker productivity. Other papers that discuss signaling in the context of taxation are Spence (1974) and Manoli (2006).

[3] 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).

We characterize optimal nonlinear tax policy by focusing on feasible and incentive-compatible (constrained efficient) allocations, invoking the revelation principle, and solving for the optimal direct revelation mechanism. We consider a realistic setting where employers are uninformed about workers' productivity, but have better information than the government. We do so by assuming that workers' signaling has two dimensions: quantity (e.g., years of schooling), which is universally observable (i.e., by both the government and employers), and quality (e.g., the difficulty or intensity of a particular educational track), which is observed only by employers.[4] To make signaling feasible, we assume that workers differ not only in their innate productive ability, but also in their costs of signaling (e.g., the cost of obtaining education). A natural interpretation of the signals in our model is that they represent components of educational effort. Consistent with this interpretation, the signals in our model are not pure waste, but realistically also enhance human capital.[5]

We begin by defining the labor market equilibrium in the presence of a general tax function, recognizing that it can be given by either a *separating tax equilibrium* (STE) or a *pooling tax equilibrium* (PTE). We recognize that the richness of the tax function plays a key role in determining whether a predistributive PTE is attainable. We then characterize the constrained efficient allocation (CEA), assuming a max-min social objective, and show that it is either given by an STE or a PTE, depending on which equilibrium configuration produces the highest level of social welfare. We derive necessary and sufficient conditions for the CEA to feature predistribution, emphasizing the role of both differences in agents' innate productivities and differences in the costs of signaling. We also discuss how the optimality of predistribution depends on the fraction of high-skilled agents in the population, the properties of the human capital production function, and the pattern of binding self-selection constraints faced by the government. Note that in our setting, incentive constraints can flow from low types to high types as well as from high types 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.

We also derive the optimal wedges in the CEA and discuss how these wedges can be implemented using either a tax function that depends jointly on income and the observable component of educational effort, or a tax function that depends only on income but is complemented by educational mandates. Our study highlights a central 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

[4]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.

[5]Our paper is thus related to papers that study the design of optimal income taxation in the presence of human capital investment and learning-by-doing, see, for example, Stantcheva (2017).

flow of information between workers and firms and prevent high-skilled individuals from separating themselves from their low-skilled counterparts.[6]

The policy framework needed to implement the CEA (either given by an STE or by a PTE) can take several forms, including education mandates and means-tested education taxes/subsidies that supplement the income tax system. These policy tools serve to limit the ability of high-achieving individuals to engage in signaling tactics, and thereby separating themselves from their lower-skill counterparts. For example, to discourage signaling, the government could impose penalties on students who complete their education unusually quickly or restrict the ability of students to enroll in multiple programs simultaneously, a practice often used for signaling purposes. Our study also suggests that income-contingent student loans, which are commonly offered on favorable terms in many countries, can be instrumental in achieving redistribution. The main re-distributive feature of these programs is that they tend to subsidize education for individuals with lower skill levels. Our findings argue for improving these programs by focusing subsidies on dimensions of educational attainment in which low-skilled workers have a comparative advantage.[7]

If the CEA is an STE, a (means-tested) subsidy on the observable (quantity) dimension of educational effort in which the low-skilled have a comparative advantage serves to alleviate the binding incentive constraint associated with high-income workers, thereby increasing the amount of redistribution that can be achieved through traditional income redistribution. The logic is similar to what happens in models of optimal hybrid taxation (combining income and commodity taxation) when low-skilled and high-skilled workers have different consumption preferences (see, e.g., Blomquist and Christiansen 2008). In this case, using the income tax to finance subsidies for the good preferred by low-skilled agents allows redistribution to be achieved at a lower efficiency cost than if the income tax were used alone.

If the CEA is given by a PTE, workers earn a common income and exert a common level of (observable) quantity effort. In this case, implementation requires a kink in the tax schedule along both the income and the effort dimensions. In particular, there is no redistribution through the income channel in a PTE, so the main role of the tax schedule in this case is to maintain the pooling equilibrium and thereby support predistribution. The latter is achieved by discouraging high-skilled agents from differentiating themselves from their low-skilled counterparts by choosing an off-equilibrium lower level of quantity effort. This is done by offering a quantity effort subsidy for levels of quantity effort below the equilibrium ("common") effort level, which

[6] Our formal analysis, detailed in Appendix H, shows that, in our setting, predistribution cannot be achieved by an income tax system in isolation.

[7] Education subsidies have traditionally been used to correct market failures and directly redistribute income. In the optimal tax literature, education subsidies are often used for two purposes: i) to mitigate/offset 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 Jacobs (2005), Maldonado (2008). Some studies also suggest the possibility of education taxes. For example, Blumkin and Sadka (2008) justifies education taxes by the positive correlation between observed education and unobserved ability. More recently, Findeisen and Sachs (2016) study income-contingent loans and find that it may be optimal to make very rich individuals pay back more than the value of the loan, implying an education tax.

effectively imposes a marginal tax on downward deviations along the quantity effort dimension. Preventing such deviations by high-skilled mimickers can be achieved more easily by supplementing the non-linear income tax system with a binding education mandate (thus setting a lower bound on quantity effort).

Note that while it is commonly understood that kinks in the income tax schedule result in individuals with different labor productivities being pooled at the same pre-tax income and receiving the same after-tax income (which can sometimes serve redistributive purposes, see, e.g., Ebert 1992), our study highlights that kinks in the education tax/subsidy schedule can serve to bunch people together at the education choice, thereby causing bunching in the wage dimension, achieving redistribution through wage compression.

While our primary focus is on how tax systems that depend jointly on income and education can achieve predistribution in a signaling context, predistribution can also be achieved through other policy instruments in different contexts. For example, anti-discrimination laws can discourage firms from screening or discriminating on the basis of observable characteristics or choices, thereby promoting a more equitable wage distribution. Thus, our policy insight extends beyond education, although we consider it a central tenet. We believe that redistribution through wage compression is an important aspect of real-world economies, and future empirical research would be valuable in quantifying the degree of redistribution associated with different policies that discourage signaling/screening through educational choice or other means in the broader labor market, thereby causing workers with different labor productivities to receive the same compensation.

The paper is organized as follows. In section 2, we present our basic model, define the STE and PTE in the presence of a general tax function, and describe the government optimization problem and the concept of CEA. In section 3, we characterize the optimal wedges associated with the CEA. In section 4, we discuss how these wedges can be implemented using a tax schedule that depends jointly on income and the observable component of educational effort, or using an income tax function that depends only on income supplemented by educational mandates, and discuss the relationship to existing policy instruments. 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 Appendix.

2 The model

Consider an economy with a competitive labor market consisting of two types of workers: a low-skilled worker, denoted by $i = 1$, and a high-skilled worker, denoted by $i = 2$, who differ in their innate ability. The proportion of workers of type i in the population (normalized to a unit measure, without loss of generality) is denoted by $0 < \gamma^i < 1$. The innate ability of a worker is assumed to be private information not available to the firm. Thus, we deviate from the standard Mirrleesian setup by considering two layers of asymmetric information, one between the government and private agents and one between workers and firms.[8]

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 in two dimensions. The first dimension 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 is 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, we will discuss the implications of other observability assumptions later.

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 type i.[9] 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:

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

[8] Other papers that have considered two layers of asymmetric information in the context of optimal policy design are Stantcheva (2014), Bastani et al. (2015, 2019), Craig (2023), and Sztutman (2024).

[9] 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 $\lim_{e_s \to \infty} \frac{\partial h}{\partial e_s} = \lim_{e_q \to \infty} \frac{\partial h}{\partial e_q} = 0$.

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 focus on the following two-stage signaling game. In the first stage, workers choose their level of effort (both quality and quantity components), (e_s^i, e_q^i); $i = 1, 2$. In the second stage, each firm offers a labor contract that specifies the income level as a function of the observed signals, $y(e_s, e_q)$. We characterize the set of Perfect Bayesian Equilibria of this signaling game.[10] Based on the observed signals, firms form their beliefs $\theta(e_s, e_q)$ about the types of workers. In equilibrium, choices are consistent in the sense that firms maximize their expected profits by choosing labor contracts given their beliefs; and workers maximize their utility (by choosing their signal/effort levels) given the labor contracts offered by firms and the applicable tax schedule. Due to the existence of asymmetric information between firms and workers, we need to consider both the possibility of a pooling equilibrium and a separating equilibrium.

The characteristic feature of signaling games is that the informed party moves first. Thus, a critical question is how a less informed agent will react to an unexpected action by the first mover. The beliefs that players have about what it means to observe an unexpected signal are called off-equilibrium path beliefs. Different off-equilibrium path beliefs can support different equilibria. In our analysis, we rely on the commonly used intuitive criterion (Cho and Kreps 1987), which is a so-called equilibrium "refinement" used to narrow the set of plausible equilibria and eliminate equilibria that are considered unreasonable or implausible. In particular, following Grossman and Perry (1986), we consider an extended version of the intuitive criterion that requires the equilibrium to be robust to credible profitable deviations by a subset of types.

The above modeling choices are consistent with Riley (2001), who shows that this refinement condition implies that under a *laissez-faire* regime with no taxes, there is no pooling equilibrium, and a separating equilibrium exists only if the fraction of low-skilled types is sufficiently large. However, these results may change in the presence of taxes. In particular, depending on what is observable by the government, and thus on how large the set of variables on which an individual's tax liability can be conditioned, there may always be a separating equilibrium, and pooling equilibria may also become sustainable.

Below we provide a formal definition of separating and pooling tax equilibria, which refer to the labor market equilibrium obtained under a general tax function. Although the analysis

[10]For a formal exposition of the notion of Perfect Bayesian Equilibrium, see Fudenberg and Tirole (1991).

that follows will mostly focus on the realistic baseline case where, based on the information available to the government, the tax function is conditioned on y and e_s, our definition of tax equilibria is broader since it considers a general (nonlinear) tax function that is allowed to depend on any of the variables y, e_s, and e_q. The alternative tax systems will be analyzed later and compared to the baseline case. We will assume throughout that the government cannot run a deficit, and since we are primarily interested in the use of taxation as a redistributive policy tool, we will assume without loss of generality that the government has no revenue needs. We begin by defining the concept of separating tax equilibrium (STE).

Definition 1 (Separating Tax Equilibrium, STE). *Let $T(y, e_s, e_q)$ denote a general tax function. Define*

$$\left(y^{1*}, e_s^{1^*}, e_q^{1^*}\right) = \underset{y^1, e_s^1, e_q^1}{\operatorname{argmax}} \left\{y^1 - T\left(y^1, e_s^1, e_q^1\right) - p_s e_s^1 - p_q^1 e_q^1\right\} \quad \textit{subject to} \quad y^1 \leq \theta^1 h(e_s^1, e_q^1) \tag{5}$$

$$\left(y^{2*}, e_s^{2^*}, e_q^{2^*}\right) = \underset{y^2, e_s^2, e_q^2}{\operatorname{argmax}} \left\{y^2 - T\left(y^2, e_s^2, e_q^2\right) - p_s e_s^2 - p_q^2 e_q^2\right\} \quad \textit{subject to} \tag{6}$$

$$y^{1*} - T\left(y^{1*}, e_s^{1*}, e_q^{1*}\right) - p_s e_s^{1*} - p_q^1 e_q^{1*} \geq y^2 - T\left(y^2, e_s^2, e_q^2\right) - p_s e_s^2 - p_q^1 e_q^2 \tag{7}$$

$$y^2 \leq \theta^2 h(e_s^2, e_q^2). \tag{8}$$

An allocation given by the quadruplets

$$\begin{aligned}
\left(y^{1*}, c^{1*}, e_s^{1^*}, e_q^{1^*}\right) &= \left(y^{1*}, y^{1*} - T\left(y^{1*}, e_s^{1*}, e_q^{1*}\right), e_s^{1^*}, e_q^{1^*}\right), \\
\left(y^{2*}, c^{2*}, e_s^{2^*}, e_q^{2^*}\right) &= \left(y^{2*}, y^{2*} - T\left(y^{2*}, e_s^{2*}, e_q^{2*}\right), e_s^{2^*}, e_q^{2^*}\right),
\end{aligned}$$

with $\left(e_s^{1^*}, e_q^{1^*}\right) \neq \left(e_s^{2^*}, e_q^{2^*}\right)$, is a Separating Tax Equilibrium (STE) under the tax function $T(y, e_s, e_q)$ if

$$\gamma^1 T\left(y^{1*}, e_s^{1*}, e_q^{1*}\right) + \gamma^2 T\left(y^{2*}, e_s^{2*}, e_q^{2*}\right) \geq 0 \tag{9}$$

and there is no allocation $(y, e_s, e_q) \in \mathbb{R}^{3+}$ jointly satisfying the three inequalities below:

$$y \leq \bar{\theta} h(e_s, e_q) \tag{10}$$

$$y - T(y, e_s, e_q) - p_s e_s - p_q^1 e_q > y^{1*} - T\left(y^{1*}, e_s^{1*}, e_q^{1*}\right) - p_s e_s^{1*} - p_q^1 e_q^{1*} \tag{11}$$

$$y - T(y, e_s, e_q) - p_s e_s - p_q^2 e_q > y^{2*} - T\left(y^{2*}, e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}. \tag{12}$$

Definition 1 captures that the STE allocation is stable since it is immune to strictly profitable deviations both on and off the equilibrium path (applying, in the context of the latter, the extended intuitive criterion previously mentioned). The interpretation of the above set of conditions is as follows. Condition (5) states that the choices of low-skilled agents in equilibrium maximize their utility, provided that the firm makes non-negative profits. Condition (6) states that the equilibrium choices of high-skilled workers maximize their utility subject to the firm making non-negative profits (given by condition (8)) and the incentive compatibility constraint associated with a mimicking low-skilled worker (given by condition (7)). Taken together, conditions (5)–(8) guarantee that no type can strictly benefit from deviating to an allocation that separates it from the other type while allowing the firm to make non-negative profits. Condition (9) is the government's revenue constraint, which states that the government maintains a weakly positive budget surplus. Finally, conditions (11)–(12) state that both types cannot strictly profit from deviating to a pooling allocation while allowing the firm to make non-negative profits (given by condition (10)).

We turn next to define the concept of pooling tax equilibrium (PTE).

Definition 2 (Pooling Tax Equilibrium, PTE)**.** *An allocation given by the quadruplet*

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

is a Pooling Tax Equilibrium (PTE) under the tax function $T\left(y, e_e, e_q\right)$ *if:*

$$\widehat{y}^* \leq \bar{\theta} h\left(\widehat{e}_s^*, \widehat{e}_q^*\right) \tag{13}$$

$$T\left(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*\right) \geq 0, \tag{14}$$

and the following conditions jointly hold:

a) 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:*

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

b) 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:*

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

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

c) 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* $i = 1$ *and* $i = 2$*:*

$$\widehat{y} - T\left(\widehat{y}, \widehat{e}_s, \widehat{e}_q\right) - p_s \widehat{e}_s - p_q^i \widehat{e}_q > \widehat{y}^* - T\left(\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*\right) - p_s \widehat{e}_s^* - p_q^i \widehat{e}_q^*. \tag{18}$$

Definition 2 captures that the PTE allocation is stable since it is immune to strictly profitable deviations off the equilibrium path. The interpretation of the above conditions is as follows. Condition (13) states that the firm makes non-negative profits under the PTE allocation. Condition (14) states that the government runs a weakly positive budget surplus. Condition (15) states that low-skilled workers cannot strictly benefit from deviating to an allocation that separates them from their high-skilled counterparts while allowing the firm to earn non-negative profits.

Conditions (16) and (17) state that high-skilled workers cannot strictly profit from deviating to an allocation that separates them from their low-skilled counterpart while allowing the firm to earn non-negative profits. Finally, condition (18) states that both types cannot strictly benefit from deviating to an alternative pooling allocation while allowing the firm to make non-negative profits.

Before turning to the introduction of the government problem, several observations are in order. First, note that by invoking a general tax function, our formulation of the tax equilibria nests all possible configurations, including laissez-faire, in which the tax is set identically to zero across the board, the income-tax-only regime, in which neither of the two education signals is taxed, and the unrestricted case, in which both income and the two education signals are subject to taxation. Note, however, that the general formulation does not imply that an equilibrium exists for each and every possible configuration. In particular, a pooling equilibrium does not exist under the laissez-faire regime (a standard result in the literature), nor does it exist under the income-tax-only regime (formally proved in Appendix H). Intuitively, to sustain a pooling equilibrium, the set of policy instruments must be sufficiently rich to guarantee that type-2 agents are denied the possibility of exploiting their comparative advantage (in one of the effort dimensions) to achieve separation from their low-ability counterpart. Framing the argument in terms of the conditions stated in Definition 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, a separating equilibrium does not necessarily exist under the laissez-faire regime, since one can find an allocation that satisfies conditions (10)–(12) if the fraction of low-skilled workers in the population is sufficiently small.

Finally, note that in formulating the tax equilibria, we have assumed that firms earn non-negative profits, rather than imposing a zero-profit condition. The latter will necessarily hold (due to the competition among firms for workers and due to the continuity of $h(e_s, e_q)$) when at least one of the three variables y, e_s, and e_q is not subject to taxation. However, for a (potentially discontinuous) tax function that depends on all three variables, a zero-profit condition does not necessarily hold (e.g., setting a confiscatory tax on any allocation other than the pooling equilibrium allocation in which the firm earns positive profits). Of course, for continuous tax functions, the zero-profit condition would hold (e.g., this would be the case for the separating equilibrium under the laissez-faire regime, if it exists). When we solve for the optimal tax equilibrium in what follows, the zero-profit condition will necessarily hold because the government could always change the tax function to extract the positive profits and thereby increase redistribution.

2.2 The government's 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 will 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.[11] We also assume that the social welfare function is of the max-min type. In line 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 3 (Constrained Efficient Allocation, CEA)**.** *A Constrained Efficient Allocation (CEA) is given by the solution to:*

$$\left\{\left(y^1, c^1, e_s^1, e_q^1\right), \left(y^2, c^2, e_s^2, e_q^2\right)\right\} = \underset{y^1,c^1,e_s^1,e_q^1,y^2,c^2,e_s^2,e_q^2}{\arg\max} \; c^1 - R^1\left(e_s^1, e_q^1\right), \tag{19}$$

subject to the government revenue constraint

$$\left(y^1 - c^1\right)\gamma^1 + \left(y^2 - c^2\right)\gamma^2 = 0, \tag{20}$$

the zero-profit conditions which require for $i = 1, 2$ that

$$y^i = \begin{cases} h\left(e_s^i, e_q^i\right)\theta^i, & \textit{for all } \left(e_s^1, e_q^1\right) \neq \left(e_s^2, e_q^2\right) \\ h\left(e_s^i, e_q^i\right)\overline{\theta}, & \textit{for all } \left(e_s^1, e_q^1\right) = \left(e_s^2, e_q^2\right), \end{cases} \tag{21}$$

and the incentive-compatibility (IC) constraints

$$c^2 - R^2\left(e_s^2, e_q^2\right) \geq c^1 - R^2\left(e_s^1, \widehat{e}_q^2\right), \tag{22}$$

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

where

$$\widehat{e}_q^2 = \begin{cases} e_q \textit{ which solves } y^1 = h\left(e_s^1, e_q\right)\overline{\theta}, & \textit{for all } \left(e_s^1, e_q^1\right) \neq \left(e_s^2, e_q^2\right) \\ e_q^1, & \textit{for all } \left(e_s^1, e_q^1\right) = \left(e_s^2, e_q^2\right). \end{cases} \tag{24}$$

We make several observations. First, as shown in Appendix A, the feasible set defined by the constraints in the optimization problem in Definition 3 contains both STE and PTE. Second, the social objective invoked by the government is max-min, which aims to maximize the welfare of type-1 workers. This means that our definition of a CEA refers to a specific point on the second-best Pareto frontier.[12] Finally, note that the above formulation of CEA assumes that the zero-profit condition holds and that the government's revenue constraint is binding.

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

[12] To relax the assumption of a max-min social objective, it would suffice to add an additional constraint to our maximization problem, namely a constraint requiring that the utility achieved by type-2 agents is weakly greater than a given pre-specified target level $\overline{V}$. All points on the second-best Pareto frontier could be obtained by varying $\overline{V}$ and repeatedly solving the government's optimization problem.

A relaxation of either condition would allow the government to change the tax function and increase redistribution.[13]

In the following, we discuss the two possible configurations of CEA separately, leaving the discussion of the conditions under which the social optimum is given by an STE or a PTE to section 2.3.

The CEA is given by an STE In this case, 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 according to their true productivity. Redistribution to the least well-off agents (type-1) occurs through the traditional ex-post tax/transfer channel, with high ability agents (type-2) paying a tax that is used to finance a transfer to their low ability counterparts (type-1).

Even if taxation is used to redistribute to type-1 agents, asymmetric information between firms and workers implies that both the IC constraint associated with type-2 mimicking type-1 and the IC constraint associated with type-1 mimicking type-2 must be taken into account (and indeed both are often binding in the optimal solution). The former IC constraint is relevant because the government wants to redistribute from type-2 agents to their type-1 counterparts; this implies that type-2 agents may have an incentive to mimic type-1 agents in order to qualify for more a lenient tax treatment (i.e., to receive a tax transfer instead of paying a tax). The latter IC constraint is relevant due to the problem of asymmetric information in the labor market; this implies that type-1 agents may have an incentive to mimic type-2 agents in order to be rewarded according to a higher productivity than their real one.

Note that in the standard setup with no 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. The two IC constraints are given by (22)–(23). Their structure reflects the idea that, since the government can condition the tax schedule on both y and e_s, the only potential room for maneuver for a mimicker is in the choice of the quality signal e_q (since this signal is unobservable to the government).

The incentive constraint for type-1 agents (constraint 23) is easy to interpret. In order for a type-1 agent to qualify for a higher wage, the only way is to replicate both effort dimensions of type-2 agents, since the firm observes *both* dimensions of education. Therefore, the constraint (23) requires that type-1 agents must weakly prefer their bundle to replicating the effort mix of their high-productivity (type-2) counterparts.

For type 2 agents, things are a bit more complicated because one has to consider possible off-equilibrium deviations. In order to qualify for the low-skill tax treatment, they must replicate the pre-tax income level and quantity effort e_s of type-1 agents. Type-2 agents may also replicate the

[13] To see this, note that when the revenue constraint is slack (a budget surplus), the government can offer a small lump-sum transfer to both types. Because of continuity, it will not violate the revenue constraint. Incentive compatibility is maintained by the linearity of utility in consumption. If the firm makes positive profits, then one can slightly raise the compensation level, y, which will maintain non-negative profits by continuity. The latter will create a fiscal surplus which, according to the previous argument, can be refunded as a lump sum transfer.

quality effort e_q of type-1 agents, which would make the two types indistinguishable to the firm (which would therefore treat them both as low-skilled types). To prevent such a deviation, the social planner must pay type-2 agents an information rent, because type-2 agents can earn the same income (y^1) as the low-skilled type by incurring lower costs due to the fact that $p_q^2 < p_q^1$.

Although firms do not observe the productivity of workers, and thus type-2 agents cannot identify themselves as high-productivity types while mimicking their type-1 counterparts, there is a possibility of out-of-equilibrium deviation that is even more profitable for type-2 workers than simply replicating the choices of type-1 agents. To see this possibility, note that if the lower-quality effort chosen by type-2 agents were also chosen by type-1 agents, the two agents would become indistinguishable for the firm, leading the firm to pay a wage equal to average productivity (since it rationally expects to hire both types of workers). For type-2 agents this off-equilibrium deviation is more appealing because the level of quality effort required to earn y^1 while being paid the average wage is always less than the level of quality effort required to earn y^1 while being paid the wage θ^1.

Two observations about the incentive constraint (22) are in order. First, violating this constraint leads to violating the extended intuitive criterion (discussed at the beginning of section 2.1 above) —since both types would find it strictly profitable to deviate to the pooling allocation associated with y^1. Second, constraint (22) reflects an information rent, accruing to high-ability workers, associated with the productivity difference between types (a type-2 agent, behaving as a mimicker, is rewarded according to average productivity rather than low productivity, as would be the case if he/she replicated both low-type signals). Note, however, that this information rent is smaller than in the standard Mirrleesian setup, where a type-2 mimicker would be rewarded according to his/her true productivity. Thus, asymmetric information between firms and workers may make it less attractive for high-skilled types to mimic low-skilled types, which may serve to enhance redistribution relative to the standard setup (see also Stantcheva 2014). However, this is not a general result because, as our analysis shows, one must also consider the potentially binding upward incentive constraint relative to the standard Mirrlees setup. We return to this issue in section 2.4.

The CEA is given by a PTE In this case we have $\left(y^1, c^1, e_s^1, e_q^1\right) = \left(y^2, c^2, e_s^2, e_q^2\right) \equiv (\widehat{y}, \widehat{c}, \widehat{e}_s, \widehat{e}_q)$. In a PTE, all agents are induced to choose the same pair (e_s, e_q), and thus each is compensated by firms according to average productivity $\bar{\theta}$ (earning $\widehat{y} = \bar{\theta} h\left(\widehat{e}_s, \widehat{e}_q\right)$). As a consequence of the fact that all agents make the same choices and earn the same income, everyone pays the same tax, which is zero given our assumption that there is no exogenous revenue requirement for the government. Redistribution occurs, but instead of working through the traditional income channel (taxes are paid by high-income earners and used to finance a transfer to low-income earners), it works through the wage channel by compressing (suppressing) wage inequality.

To distinguish between these two channels of redistribution, we will use the term ”predistribution” to refer to the redistribution that occurs through the wage channel. Note also that under a PTE, the incentive compatibility constraints (22)–(23) become de facto irrelevant since they are trivially satisfied.

2.3 When is predistribution optimal?

Let us now turn to an analysis of the social optimality of STE and PTE. In an STE, the government typically cannot fully eliminate the information rents associated with productivity differences. In a PTE, on the other hand, the government fully eliminates these information rents by forcing full wage compression, but the PTE typically has less desirable efficiency properties. The equity-efficiency tradeoff between pooling and separation depends crucially on the differences in productivity and in the cost of acquiring the quality signal e_q between the two types of workers.[14] Proposition 1 below gives the main results.

Proposition 1 (Optimality of Predistribution)**.** *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* $\delta = p_q^1 - p_q^2 > 0$*. The CEA can be characterized as follows:*

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

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

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

Proof. See Appendix B □

While augmenting the income tax system with taxes/subsidies on education can improve redistribution under separation —by alleviating the binding IC constraints faced by the government —Proposition 1 delineates scenarios in which the taxation/subsidization of education, by enabling the implementation of a PTE, allows increasing social welfare beyond what's achievable under an STE.

Part i) of Proposition 1 establishes the case for predistribution by identifying a non-empty set of parameters under which pooling is welfare enhancing relative to separation. Part ii) shows that pooling is socially desirable when the difference in the cost of obtaining the quality signal across types is moderate. In such a scenario, type-1 workers—who tend to invest more effort in the quality dimension to qualify for higher wage rates—are more inclined to engage in mimicking. In contrast to the standard Mirrlees model, in this scenario both IC constraints

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

are binding, and the efficiency gains from separation over pooling are limited. Part iii) shows that pooling is socially preferred when the productivity gap between the two types of workers is moderate, making the efficiency loss from wage compression relatively small.

2.4 Welfare implications of the second layer of asymmetric information

Stantcheva (2014) shows 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. 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 Appendix B) that when the CEA 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 formally shown in Proposition 2.

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

Proof. See Appendix C. □

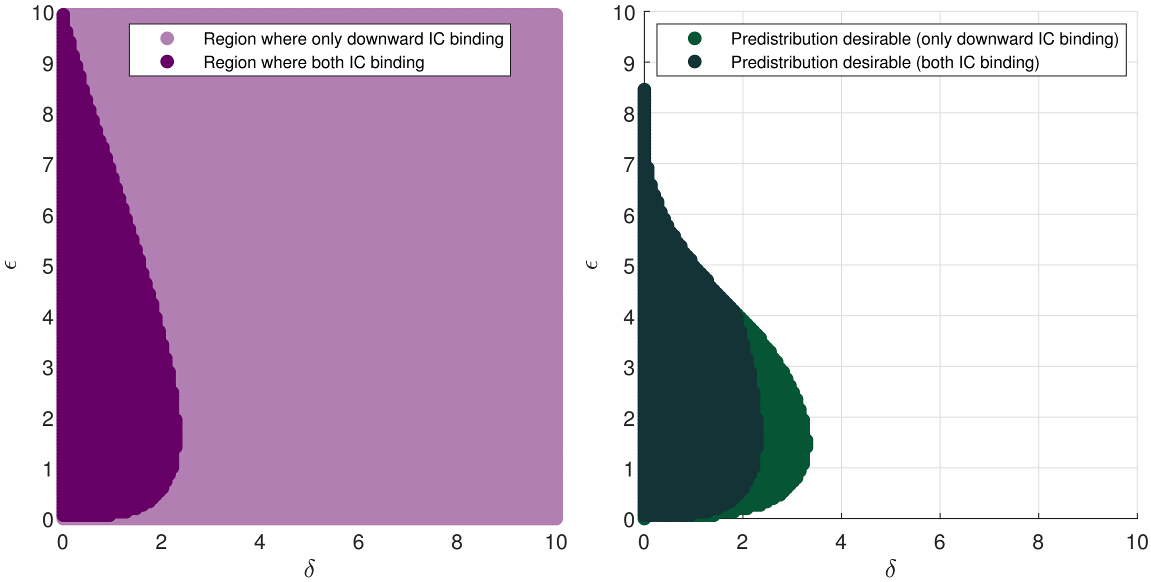
2.5 Further discussion of predistribution optimality

Given a particular functional form of the human capital production function, we can provide a complete analytical representation of the ϵ and δ combinations where predistribution is favorable. To do so, we assume that

$$h(e_s, e_q) = (e_s e_q)^{\beta}, \tag{25}$$

where $0 < \beta < 1/2$ (implying that the production function is strictly concave), and we adopt the following approach. For each combination (ϵ, δ), based on Definition 3, we compute the *optimal* STE that yields the highest welfare for type-1 agents and the *optimal* PTE that yields the highest welfare for type-1 agents, and then compare them to assess which one is superior. We then provide a graphical illustration of the parameter regions in which predistribution (PTE) constitutes the social optimum, highlighting how these regions depend on the set of binding incentive constraints in the optimal STE. The analytical inequalities characterizing the regions are derived in Appendix D and summarized in Appendix D.4. We evaluate the regions in Figure 1 with the parameters $\beta = 0.10$, $\gamma^1 = \gamma^2 = 0.5$, $p_q^1 = 10$, and $\theta^2 = 10$, with δ ranging from 0 to p_q^1 and ϵ ranging 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 region in the right panel) largely overlaps with the region where both IC constraints are binding in the STE (the dark purple region in the left panel). However, for moderate values of ϵ, one can also find cases (see the light green region in the right panel) where the PTE is welfare superior to the STE, even though in the latter only the downward IC constraint is binding.

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.[15] The fact that type-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

[15]The reason is as follows. When $p_q^1 = p_q^2$, the left side of the downward IC constraint (22) coincides with the right side of the upward IC constraint (23) (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 (22) is strictly larger than the left-hand side of the upward IC constraint (23) whenever type-1 agents are required to produce a positive amount of output. Thus, for $p_q^1 = p_q^2$, the IC constraints (22)–(23) can be jointly satisfied only if $y^1 = 0$, implying that type-1 agents remain idle.

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 (22)–(23), 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 $p_q^1 e_q^1 - p_q^2 \hat{e}_q^2$ 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 $y^1 > 0$, $\hat{e}_q^2 < e_q^1$. The latter component reflects the information rent arising from productivity differences and, under the assumption given by (25), 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.

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 (23)) 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.[16]

In Appendix K, we provide an additional analysis that quantifies 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 CEA, 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.

[16]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.

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

$$\widehat{\Omega}^1_{e_s,e_q} \equiv \widehat{MRTS}^1 - \frac{p_s}{p^1_q} = 0 \quad \textit{and} \quad \widehat{\Omega}^2_{e_s,e_q} \equiv \widehat{MRTS}^2 - \frac{p_s}{p^2_q} < 0, \tag{26}$$

$$\widehat{\Omega}^1_{e_s,c} \equiv 1 - \frac{p_s}{\theta^1 h_1\left(\widehat{e}_s, \widehat{e}_q\right)} = \widehat{\Omega}^1_{e_q,c} = 1 - \frac{p^1_q}{\theta^1 h_2\left(\widehat{e}_s, \widehat{e}_q\right)} < 0, \tag{27}$$

$$\widehat{\Omega}^2_{e_q,c} \equiv 1 - \frac{p^2_q}{\theta^2 h_2\left(\widehat{e}_s, \widehat{e}_q\right)} > \widehat{\Omega}^2_{e_s,c} = 1 - \frac{p_s}{\theta^2 h_1\left(\widehat{e}_s, \widehat{e}_q\right)} > 0, \tag{28}$$

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

(ii) If the CEA 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:

$$\Omega^1_{e_s,e_q} \equiv MRTS^1 - \frac{p_s}{p^1_q} = \frac{\lambda^2}{\gamma^1}\left(MRTS^{21}\frac{p^2_q}{p^1_q} - MRTS^1\right) < 0, \tag{29}$$

$$\Omega^1_{e_q,c} \equiv 1 - \frac{p^1_q}{\theta^1 h_2\left(e^1_s, e^1_q\right)} = \frac{\lambda^2}{\gamma^1}\left(\frac{p^1_q}{\theta^1 h_2\left(e^1_s, e^1_q\right)} - \frac{p^2_q}{\overline{\theta} h_2\left(e^1_s, \widehat{e}^2_q\right)}\right) > 0, \tag{30}$$

$$\Omega^1_{e_s,c} \equiv 1 - \frac{p_s}{\theta^1 h_1\left(e^1_s, e^1_q\right)} = \frac{\lambda^2}{\gamma^1}\left(\frac{h_1\left(e^1_s, \widehat{e}^2_q\right)}{h_1\left(e^1_s, e^1_q\right)}\frac{1}{\theta^1} - \frac{1}{\overline{\theta}}\right)\frac{p^2_q}{h_2\left(e^1_s, \widehat{e}^2_q\right)}, \tag{31}$$

$$\Omega^2_{e_s,e_q} \equiv MRTS^2 - \frac{p_s}{p^2_q} = \frac{\lambda^1}{\gamma^2}\left(\frac{p^1_q}{p^2_q} - 1\right)\cdot MRTS^2 \geq 0, \tag{32}$$

$$\Omega^2_{e_q,c} \equiv 1 - \frac{p^2_q}{\theta^2 h_2\left(e^2_s, e^2_q\right)} = \frac{\lambda^1}{\gamma^2}\frac{p^2_q - p^1_q}{\theta^2 h_2\left(e^2_s, e^2_q\right)} \leq 0, \tag{33}$$

$$\Omega^2_{e_s,c} \equiv 1 - \frac{p_s}{\theta^2 h_1\left(e^2_s, e^2_q\right)} = 0, \tag{34}$$

where λ^2 and λ^1 denote the Lagrange multipliers associated with constraint (22) and constraint (23), 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 (24).

Proof. See Appendix E. □

Starting with part i), condition (26) 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. The first result is driven by our assumption that the social objective is to maximize the welfare of type-1 agents. This assumption implies that the effort mix $(\hat{e}_s, \hat{e}_q)$ is chosen so that type-1 agents earn $\hat{y}$ most efficiently. Since type-2 agents also have to exert $(\hat{e}_s, \hat{e}_q)$ in the PTE, and since they have a comparative advantage in the quality dimension, they would prefer more quality effort e_q and correspondingly less quantity effort e_s, and thus face a distortion in their effort mix. Since h represents the acquired human capital, condition (26) also implies that in a PTE the acquired human capital of type-1 agents is distorted upward (equation (27)), while the acquired human capital of type-2 agents is downward distorted (equation (28)).

Now consider part ii) of Proposition 3, which refers to the case of an STE. Condition (29) implies that the effort mix of type-1 agents is distorted towards e_s (i.e., e_s^1 is distorted upward vis-à-vis e_q^1).[17] To provide an intuition for this result, consider the following. For a given isoquant $\theta^1 h(e_s, e_q) = y^1$, assume that type-1 agents are induced to choose the effort mix (e_s^1, e_q^1) that satisfies the no-distortion condition $MRTS^1 = \frac{p_s}{p_q^1}$. Since the government observes e_s, a type-2 mimicker must choose e_s^1, while $\widehat{e}_q^2$ satisfies the equation $\overline{\theta}h(e_s^1, e_q) = y^1$, so $\widehat{e}_q^2 < e_q^1$. It follows that $MRTS^{21} < MRTS^1 = \frac{p_s}{p_q^1} < \frac{p_s}{p_q^2}$. Thus, in this case, type-2 mimickers are induced to choose an effort mix that is distorted toward e_s. Condition (29) implies that instead of letting type-1 agents satisfy the condition $MRTS^1 = \frac{p_s}{p_q^1}$, it is welfare superior to induce type-1 agents to choose an effort mix that is slightly distorted toward e_s. To see this, note that if the distortion is small, it will have only a second-order effect on the total cost $p_s e_s^1 + p_q^1 e_q^1$ incurred by type-1 agents. However, it will have a *negative* first-order welfare effect on type 2 mimickers, increasing their total cost $p_s e_s^1 + p_q^2 \widehat{e}_q^2$ (since it exacerbates the initial distortion in their effort mix).[18]

Eqs. (30)–(31) shed light on the distortion of each given dimension of effort vis-à-vis consumption. According to (30), 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 ($p_q^1 > p_q^2$) compared to type-2 agents, 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 (again, this is due to the fact that $\overline{\theta} > \theta^1$ implies $\widehat{e}_q^2 < e_q^1$ and therefore $h_2\left(e_s^1, e_q^1\right) < h_2\left(e_s^1, \widehat{e}_q^2\right)$). 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 mimickers.

Eq. (31) tells us that one cannot unambiguously determine the direction of the optimal distortion of e_s^1 (relative to consumption). This is due to the fact that one cannot unambiguously say whether the marginal productivity of e_s is higher or lower for a type-1 agent than for 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 distort 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 the cross derivative h_{12} is positive) 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

[17] Recall that our focus on a max-min social objective implies that the downward IC constraint (22) is necessarily binding, i.e., $\lambda^2 > 0$.

[18] Inducing type-1 agents to choose an effort mix that is slightly distorted toward e_q is welfare inferior. To see this, note that if the distortion is small, it will again have only a second-order effect on the total costs $p_s e_s^1 + p_q^1 e_q^1$ incurred by type-1 agents; but will have a first-order *beneficial* welfare effect on type-2 mimickers, reducing their total cost $p_s e_s^1 + p_q^2 \widehat{e}_q^2$ (since it alleviates the initial distortion in their effort mix).

and e_q, one can conclude that e_s^1 is distorted downward relative to consumption.

Now consider the equations (32)–(34), which provide expressions for the wedges characterizing the allocation obtained by type-2 agents. The first thing to note is that λ^1 can be either positive (the upward IC constraint (23) is binding) or zero (the upward IC constraint (23) is slack). [19] When $\lambda^1 = 0$, the equations (32)–(34) tell us that all wedges are zero in the allocation obtained by type-2 agents.

Consider the case where $\lambda^1 > 0$. Eq. (32) tells us that the effort mix of type-2 agents is distorted toward e_q (i.e., e_q^2 is distorted upward vis-à-vis 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 mimicking by type-1 agents less attractive. The intuition behind this result is as follows. 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 the constraint (23) 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 when acting as mimickers, type-1 agents are forced to choose an effort mix that is distorted toward e_q: $MRTS^2 > \frac{p_s}{p_q^1}$. 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 towards e_q. If the distortion is small, it will have only a second-order effect on the total cost $p_s e_s^2 + p_q^2 e_q^2$ incurred by type-2 agents; however, it will have a first-order adverse effect on type-1 mimickers, since the total cost $p_s e_s^2 + p_q^1 e_q^2$ will increase.

According to (33), e_q^2 is unambiguously distorted upward relative to consumption. This happens because compared to type-1 agents, and therefore also compared to type-1 as a mimicker, 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 (34), we can see that e_s^2 is not distorted 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 mimickers.

4 Implementation

We now turn to discuss how the wedges given in Proposition 3 translate into properties of the implementing tax function $T(y, e_s)$. We start with the case where the CEA is given by an STE.

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

4.1 Implementation of the STE

Under a tax schedule that is a function of both y and e_s, type 1 agents solve the following optimization problem:

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

Using subscripts on T to denote partial derivatives, the corresponding first-order conditions are

$$1 - T_1'\left(\theta^1 h\left(e_s^1, e_q^1\right), e_s^1\right) - \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - \frac{T_2'\left(\theta^1 h\left(e_s^1, e_q^1\right), e_s^1\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)} = 0, \tag{35}$$

$$1 - T_1'\left(\theta^1 h\left(e_s^1, e_q^1\right), e_s^1\right) - \frac{p_q^1}{\theta^1 h_2\left(e_s^1, e_q^1\right)} = 0, \tag{36}$$

which leads to the following implicit characterization of marginal tax rates:

$$T_1'\left(\theta^1 h\left(e_s^1, e_q^1\right), e_s^1\right) = 1 - \frac{p_q^1}{\theta^1 h_2\left(e_s^1, e_q^1\right)}, \tag{37}$$

$$T_2'\left(\theta^1 h\left(e_s^1, e_q^1\right), e_s^1\right) = \frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} p_q^1 - p_s. \tag{38}$$

Instead, type 2 agents solve the following optimization problem:

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

subject to:

$$U^1 \geq \theta^2 h\left(e_s^2, e_q^2\right) - p_s e_s^2 - p_q^2 e_q^2 - T\left(\theta^2 h\left(e_s^2, e_q^2\right), e_s^2\right) - \left(p_q^1 - p_q^2\right) e_q^2. \tag{39}$$

Denoting by ϕ the Lagrange multiplier attached to the IC constraint (39), the associated first-order conditions are

$$1 - T_1'\left(\theta^2 h\left(e_s^2, e_q^2\right), e_s^2\right) - \frac{p_s}{\theta^2 h_1\left(e_s^2, e_q^2\right)} - \frac{T_2'\left(\theta^2 h\left(e_s^2, e_q^2\right), e_s^2\right)}{\theta^2 h_1\left(e_s^2, e_q^2\right)} = 0, \tag{40}$$

$$1 - T_1'\left(\theta^2 h\left(e_s^2, e_q^2\right), e_s^2\right) - \frac{p_q^2}{\theta^2 h_2\left(e_s^2, e_q^2\right)} = -\frac{\phi}{1-\phi} \frac{p_q^1 - p_q^2}{\theta^2 h_2\left(e_s^2, e_q^2\right)}, \tag{41}$$

which leads to the following implicit characterization of marginal tax rates:

$$T_1'\left(\theta^2 h\left(e_s^2, e_q^2\right), e_s^2\right) = 1 - \frac{p_q^2}{\theta^2 h_2\left(e_s^2, e_q^2\right)} + \frac{\phi}{1-\phi} \frac{p_q^1 - p_q^2}{\theta^2 h_2\left(e_s^2, e_q^2\right)}, \tag{42}$$

$$T_2'\left(\theta^2 h\left(e_s^2, e_q^2\right), e_s^2\right) = \frac{h_1\left(e_s^2, e_q^2\right)}{h_2\left(e_s^2, e_q^2\right)} p_q^2 - p_s - \frac{\phi}{1-\phi} \frac{h_1\left(e_s^2, e_q^2\right)}{h_2\left(e_s^2, e_q^2\right)} \left(p_q^1 - p_q^2\right). \tag{43}$$

We can then establish the following.

Proposition 4. *The implicit marginal tax rates characterizing the constrained efficient allocation in the case of an STE* $\{(c^1, y^1, e_s^1, e_q^1), (c^2, y^2, e_s^2, e_q^2)\}$ *are as follows:*

$$T_1'\left(y^1, e_s^1\right) = \frac{\lambda^2}{\gamma^1}\left(\frac{p_q^1}{\theta^1 h_2\left(e_s^1, e_q^1\right)} - \frac{p_q^2}{\bar{\theta} h_2\left(e_s^1, \widehat{e}_q^2\right)}\right) > 0, \tag{44}$$

$$T_2'\left(y^1, e_s^1\right) = \frac{p_q^1 \lambda^2}{\gamma^1}\left(MRTS^{21}\frac{p_q^2}{p_q^1} - MRTS^1\right) < 0, \tag{45}$$

$$T_1'\left(y^2, e_s^2\right) = T_2'\left(y^2, e_s^2\right) = 0. \tag{46}$$

Proof. See Appendix F. □

The fact that we get the canonical "efficiency-at-the-top" result for high-ability agents with bounded skill distributions (Sadka 1976) is somewhat surprising. One might have expected, for example, that it would be desirable to use the tax function to distort the effort mix of type-2 agents toward e_q (the component in which type-2 agents have a comparative advantage) in order to discourage mimicking by type-1 agents. A key point, however, is that the IC constraint associated with type-1 agents is already embedded in the laissez-faire equilibrium, and is therefore already internalized by type-2 agents in their decision-making process. The labor contract offered to type-2 agents maximizes their utility subject to the IC constraint of type-1 workers. This maximization is consistent with the government's goal of extracting the maximum amount of taxes from type-2 agents and thereby enhancing redistribution. Note also that although for high-skilled agents $T_1' = T_2' = 0$ under the implementing tax function, y^2 in the STE will be lower than under laissez-faire if the upward IC constraint associated with low-skilled workers is binding under laissez-faire. This is due to the redistribution performed by the tax system, which increases the utility of type-1 agents (relative to their utility under laissez-faire), thereby making them less inclined to imitate their type-2 counterparts.

The above discussion suggests that implementing an STE requires supplementing the income tax system with an education subsidy 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 by taxation). The latter allows the government to extract more taxes from high-skilled workers, thereby increasing redistribution.

Finally, note that it may also be possible to achieve implementation by means of a tax function that depends only on income, supplemented by a mandate on e_s that enforces a lower bound, set at $e_s = e_s^1$, on the value that agents can choose for this variable. However, this alternative implementation scheme is only an option in the case where the inequality $e_s^1 \leq e_s^2$ is satisfied in the CEA.

4.2 Implementation of the PTE

If the CEA 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:*

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

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:

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

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

$$T\left(y\right) = \begin{cases} \left(\frac{1}{\theta^1} - \frac{1}{\overline{\theta}}\right) \frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)} \widehat{y} + \left(\frac{1}{\overline{\theta}} - \frac{1}{\theta^1}\right) \frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)} y, & \textit{for all } y \in [0, \widehat{y}] \\ \left(y - \widehat{y}\right) \max\left\{1 - \frac{p_q^2}{\overline{\theta} h_2\left(\widehat{e}_s, \widehat{e}_q\right)}, \frac{\left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s \widehat{e}_s - p_q^2 \widehat{e}_q\right]}{\underline{y}^{sep} - \widehat{y}}\right\}, & \textit{for all } y > \widehat{y}. \end{cases} \tag{47}$$

Proof. See Appendix G. □

Formula (47) 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 $\left(\widehat{e}_s, \widehat{e}_q\right)$.[20] The (positive) marginal tax rate on the second bracket serves to distort downward the acquired human capital of type-2 agents, and it is designed to be high enough to achieve two goals. One is to ensure that type-2 agents (weakly) prefer pooling at $\widehat{y}$ to pooling at a higher income; the other is 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.

[20] If 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 $e_q < \widehat{e}_q$.

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 the 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 (47), namely $\widetilde{T}(y) = T(y) + p_s\widehat{e}_s$.[21]

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 CEA (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.[22] 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 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

[21] 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 at the PTE an income tax of $p_s\widehat{e}_s$, allowing the government to raise enough revenue to cover the public expenditures associated with public provision.

[22] 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.

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.[23]

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 CEA 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 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 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 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 CEA features predistribution or not. However, consistent with Proposition 1, the results show that the CEA 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 is moderate.

[23]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.

5.2 The case when both signals are taxed

In 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 Appendix I shows, this implies that the CEA 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.[24] 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.[25] 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 (29) 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.[26]

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

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

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

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 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 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.[27] 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 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 Conclusions

We have analyzed optimal redistribution in the presence of signaling, introducing two realistic new features to the standard Mirrleesian framework: (i) the existence of a second layer of asymmetric information between employers and workers regarding the productivity of workers, with the latter having the possibility to engage in signaling to credibly convey this information to potential employers; (ii) a tax system that conditions taxes/transfers on the income earned by workers as well as on the signals that the government can observe in the labor market. The combination of these two new features is shown to preserve the second-best nature of the government optimization problem and the inherent trade-off between conflicting equity and efficiency considerations.

We focused on a model with two-dimensional signals (quality and quantity of education) and

[27] 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.

two types of agents (low- and high-skilled) that differ in the cost of acquiring the signal(s), in which high-skilled agents have a (weak) absolute advantage in signaling through each channel and a comparative advantage in the quality signal.

We have shown that constrained efficient allocations can be given by either a separating or a pooling tax equilibrium, the latter implying that the government engages in *predistribution*, which we have defined as a change in the wage structure that results in cross-subsidization between skill levels.

From a policy perspective, by treating educational attainment as a signaling device used by high-skilled workers to differentiate themselves in the labor market from their low-skilled counterparts, our analysis sheds new light on commonly used policy instruments such as education mandates and (means-tested) education subsidies. These policy instruments are often justified on efficiency grounds to address pervasive market failures (e.g., alleviating credit constraints or internalizing externalities). However, we argue that these instruments can also be viewed as a form of tax on the signals acquired by workers, and thus serve to promote redistributive goals by interfering with the exchange of information between workers and employers.

Our analysis is a first step in exploring redistributive policy in the presence of signaling, and we invoke several restrictive simplifying assumptions to gain tractability. In particular, we have limited the analysis to a setup with two types of agents. However, our main qualitative insights carry over to the general case with many types. In the general case, the social optimum could be given by a hybrid allocation that combines predistribution (bunching) with redistribution, rather than taking one of the two extreme configurations (full separation or pooling) as in the two-type setting.

Finally, taking a broader view of the social desirability of predistribution, this paper emphasizes the signaling role of educational attainment and the potential welfare-enhancing role of educational subsidies and mandates, but there are clearly other contexts in which additional policy instruments can serve to enhance predistribution. A notable example is the widespread use of anti-discrimination legislation to limit the ability of firms to engage in screening or statistical discrimination.

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. ISPE conference on Public economics and redistribution.

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.

A Feasibility of both STE and PTE

If the government has access to a general tax $T(y, e_s)$, both configurations of tax equilibria always exist, regardless of the relative size of the two groups of agents and regardless of how large the difference $p_q^1 - p_q^2$ is.

First, consider the STE. It is always possible to design a tax function that maintains a separating equilibrium where $y^1 = e_s^1 = 0$ (implying $e_q^1 = \widehat{e}_q^2 = 0$). To see that this is indeed the case, invert the function $y = \theta h(e_s, e_q)$ to get $e_q = f\left(\frac{y}{\theta}, e_s\right)$, and 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

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

and

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

It is straightforward to verify that they satisfy the government revenue constraint (20):

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

Furthermore, they also satisfy the incentive constraints (22) and (23), 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 $R^2\left(e_s^1, \widehat{e}_q^2\right) = R^1\left(e_s^1, e_q^1\right) = 0$, constraint (22) simplifies to

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

and constraint (23) simplifies to

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

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 (22) takes the same value as the left-hand side of the constraint (23). Thus, if the incentive constraint (22) that applies to type-2 agents is satisfied, the incentive constraint (23) that applies to type-1 agents is necessarily slack (due to the assumption that $p_q^1 > p_q^2$).

The separating equilibrium represented by the quadruplets (A1)-(A2) can be implemented by a tax function characterized by the following properties:

$$T\left(y^*, e_s^*\right) = \gamma^1 \left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right) \right] > 0, \tag{A6}$$

$$T\left(0, 0\right) = -\gamma^2 \left[y^* - p_s e_s^* - p_q^2 f\left(\frac{y^*}{\theta^2}, e_s^*\right) \right] < 0, \tag{A7}$$

$$T\left(y, e_s\right) = \infty \quad \text{for} \quad \left(y, e_s\right) \notin \left\{ \left(0, 0\right), \left(y^*, e_s^*\right) \right\}. \tag{A8}$$

Next, consider the PTE. Let $y^{1*} = y^{2*} = c^{1*} = c^{2*}, e_s^{1*} = e_s^{2*}, e_q^{1*} = e_q^{2*}$ and $y^{i*} = h\left(e_s^{i*}, e_q^{i*}\right)\overline{\theta}$ denote a pooling allocation. It is easy to check that the zero-profit condition is satisfied and that the tax revenue is zero. A tax function that implements the pooling allocation is given by $T(y^{i*}, e_s^{i*}) = 0$ and $T\left(y, e_s\right) = \infty \quad \text{for} \quad \left(y, e_s\right) \neq \left(y^{i*}, e_s^{i*}\right)$. To see this, note that given the existing tax function, the only possible profitable deviation is given by a downward adjustment along the (untaxed) quality effort dimension e_q. However, such a deviation would not allow type 2 to separate from its less skilled counterpart. The reason is that, if both types deviate (implying that workers are compensated based on average productivity), firms will suffer losses, since $y^{i*} > h\left(e_s^{i*}, e_q^{i}\right)\overline{\theta}$ for $e_q^i < e_q^{i*}$.

B Proof of Proposition 1

Part (i) 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.

Denote the resulting *Rawlsian* 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):

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

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:

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

Using the envelope theorem, it follows that:

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

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

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:

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

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

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

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

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 (i).

Part (ii) 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 (i)), 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

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

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 (ii).

Part (iii) 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:

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

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

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 (iii) of Proposition 1 [see Appendix B, Eqs. (B19) and (B20)], under the above parametric assumptions, the CEA 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 CEA 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 CEA 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.[28]

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 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.

[28]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.

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 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 (25).

D.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

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

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

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

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

Dividing (D3) by (D4) gives

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

Noticing that (D3) can be restated as

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

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

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

Finally, given that from (D3) we have

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

using (D5) we get that

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

which implies

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

and therefore

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

We can then conclude that

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

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

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

where

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

D.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:

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

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

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:

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

subject to the downward IC-constraint

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

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:

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

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

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

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

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

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

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

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

Now rewrite (D16) and (D18) as:

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

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

Dividing (D24) by (D25) gives

$$y^1 = \frac{p_s e_s^1}{\beta}. \tag{D26}$$

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.

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

and therefore

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

Combining (D22) and (D27) gives

$$\begin{aligned} \frac{e_s^1}{e_s^2} &= \frac{\dfrac{\left(\gamma^1\right)^{\frac{\beta}{1-2\beta}}\left(\beta\right)^{\frac{1}{1-2\beta}}\left(p_s\right)^{\frac{\beta-1}{1-2\beta}}}{\left[\frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}}-\frac{\gamma^2 p_q^2}{\left(\bar{\theta}\right)^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}}}{\dfrac{\left(\theta^2\beta\right)^{\frac{1}{1-2\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}}\left(p_q^2\right)^{\frac{\beta}{1-2\beta}}}}=\frac{\left(\gamma^1\right)^{\frac{\beta}{1-2\beta}}\left(\beta\right)^{\frac{1}{1-2\beta}}\left(p_s\right)^{\frac{\beta-1}{1-2\beta}}}{\left[\frac{p_q^1}{\left(\theta^1\right)^{\frac{1}{\beta}}}-\frac{\gamma^2 p_q^2}{\left(\bar{\theta}\right)^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}}\frac{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}}\left(p_q^2\right)^{\frac{\beta}{1-2\beta}}}{\left(\theta^2\beta\right)^{\frac{1}{1-2\beta}}} \\ &= \left[\frac{\gamma^1 p_q^2}{p_q^1\left(\frac{\theta^2}{\theta^1}\right)^{\frac{1}{\beta}}-\gamma^2 p_q^2\left(\frac{\theta^2}{\bar{\theta}}\right)^{\frac{1}{\beta}}}\right]^{\frac{\beta}{1-2\beta}}. \end{aligned} \tag{D29}$$

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

$$\begin{aligned} p_q^1 e_q^1 &= \left(\frac{y^1}{\theta^1}\right)^{\frac{1}{\beta}}\frac{p_q^1}{e_s^1}=p_s e_s^1-\left[p_s-\left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}}\frac{p_q^2}{\left(e_s^1\right)^2}\right]\gamma^2 e_s^1 \\ &= \gamma^1 p_s e_s^1+\gamma^2\left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}}\frac{p_q^2}{e_s^1}. \end{aligned} \tag{D30}$$

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

$$\begin{aligned} y^1 - p_s e_s^1 - p_q^1 e_q^1 &= \frac{p_s e_s^1}{\beta} - p_s e_s^1 - \gamma^1 p_s e_s^1 - \gamma^2 \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^1} \\ &= \frac{1-\beta-\gamma^1\beta}{\beta} p_s e_s^1 - \gamma^2 p_q^2 \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{1}{e_s^1} \\ &= \frac{1-2\beta}{\beta} p_s e_s^1 + \gamma^2 \left[p_s e_s^1 - p_q^2 \left(\frac{y^1}{\bar{\theta}}\right)^{\frac{1}{\beta}} \frac{1}{e_s^1} \right] \\ &= \frac{1-2\beta}{\beta} p_s e_s^1 + \gamma^2 \left[p_s e_s^1 - p_q^2 \left(\frac{p_s}{\beta\bar{\theta}}\right)^{\frac{1}{\beta}} \left(e_s^1\right)^{\frac{1-\beta}{\beta}} \right]. \qquad \text{(D31)} \end{aligned}$$

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

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

and therefore

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

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

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

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

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

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

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

i.e., exploiting (D22),

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

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

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

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:

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

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

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

i.e.,

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

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

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

i.e.,

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

i.e.,

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

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))

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

it follows that

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

and therefore, since

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

we can rewrite condition (D37) as

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

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

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

D.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)

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

and the upward IC-constraint

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

Rewriting (D40) we have that

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

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

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

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

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

or equivalently,

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

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

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

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

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

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

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

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

from which one obtains that

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

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

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

and therefore

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

from which we also obtain that

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

and therefore

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

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

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

i.e.,

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

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.,

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

from which one obtains that

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

or equivalently

$$\begin{aligned} e_s^1 \quad = \quad & \left\{\frac{\gamma^1 \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} + \gamma^2 \left(\theta^2\right)^{\frac{1}{1-\beta}} \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}}{\gamma^1 \left(p_q^1 - p_q^2\right) \left(\overline{\theta}\right)^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1 \left(\overline{\theta}\right)^{\frac{1}{\beta}} - p_q^2 \left(\theta^1\right)^{\frac{1}{\beta}}\right]}\right\}^{\frac{\beta}{1-2\beta}} \\ & \times \frac{\left(\beta\right)^{\frac{1}{1-2\beta}} \left(\theta^1\overline{\theta}\right)^{\frac{1}{1-\beta}} \left(p_q^1 - p_q^2\right)^{\frac{\beta}{1-\beta}}}{\left(p_s\right)^{\frac{1-\beta}{1-2\beta}} \left(p_q^1\right)^{\frac{\beta}{1-2\beta}}}. \end{aligned} \tag{D52}$$

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

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

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:

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

i.e.,

$$\begin{aligned} e_s^2 &= \left[p_q^1(\overline{\theta})^{\frac{1}{\beta}} - p_q^2(\theta^1)^{\frac{1}{\beta}} \right]^{\frac{\beta}{1-\beta}} \frac{(\beta)^{\frac{1}{1-2\beta}} (\theta^2)^{\frac{1}{1-\beta}}}{(p_s)^{\frac{1-\beta}{1-2\beta}} (p_q^1)^{\frac{\beta}{1-2\beta}}} \\ &\quad \times \left\{ \frac{\gamma^1 (p_q^1 - p_q^2)^{\frac{\beta}{1-\beta}} (\theta^1\overline{\theta})^{\frac{1}{1-\beta}} + \gamma^2 (\theta^2)^{\frac{1}{1-\beta}} \left[p_q^1(\overline{\theta})^{\frac{1}{\beta}} - p_q^2(\theta^1)^{\frac{1}{\beta}} \right]^{\frac{\beta}{1-\beta}}}{\gamma^1 (p_q^1 - p_q^2)(\overline{\theta})^{\frac{1}{\beta}} + \gamma^2 \left[p_q^1(\overline{\theta})^{\frac{1}{\beta}} - p_q^2(\theta^1)^{\frac{1}{\beta}} \right]} \right\}^{\frac{\beta}{1-2\beta}} \end{aligned} \tag{D55}$$

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

$$y^2 = \frac{p_s e_s^2}{\beta}. \tag{D56}$$

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

$$\begin{aligned} U^{sep} = &\left(\frac{1-\beta}{\beta}\right) \left(\gamma^2 e_s^2 + \gamma^1 e_s^1\right) p_s + \\ &\gamma^2 p_q^2 \left(\frac{p_s}{\beta\overline{\theta}}\right)^{\frac{1}{\beta}} (e_s^1)^{\frac{1-\beta}{\beta}} - p_q^1 \left(\frac{p_s}{\beta\theta^1}\right)^{\frac{1}{\beta}} (e_s^1)^{\frac{1-\beta}{\beta}} - \gamma^2 p_q^2 \left(\frac{p_s}{\beta\theta^2}\right)^{\frac{1}{\beta}} (e_s^2)^{\frac{1-\beta}{\beta}}. \end{aligned} \tag{D57}$$

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

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

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

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

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

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

i.e.,

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

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

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

i.e.,

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

or equivalently,

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

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}}}{(p_s)^{\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}}}{(p_s)^{\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.4 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

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

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 $(\delta, \varepsilon) \in [0, p_q^1] \times [0, \theta^2]$, as

$$f(\delta,\varepsilon) \equiv \frac{\delta}{(\theta^2)^{\frac{1}{\beta}}} - \left[\frac{p_q^1}{(\theta^2-\varepsilon)^{\frac{1}{\beta}}} - \frac{p_q^1-\delta}{(\theta^2-\gamma^1\varepsilon)^{\frac{1}{\beta}}}\right]\left[\frac{\gamma^1(p_q^1-\delta)}{p_q^1\left(\frac{\theta^2}{\theta^2-\varepsilon}\right)^{\frac{1}{\beta}} - \gamma^2(p_q^1-\delta)\left(\frac{\theta^2}{\theta^2-\gamma^1\varepsilon}\right)^{\frac{1}{\beta}}}\right]^{\frac{1-\beta}{1-2\beta}}. \tag{D61}$$

Our results imply that:

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

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

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

$$(\theta^2-\gamma^1\varepsilon)^{\frac{1}{1-2\beta}} > \frac{\left\{\gamma^2\left[p_q^1-(p_q^1-\delta)\left(\frac{\theta^2-\varepsilon}{\theta^2-\gamma^1\varepsilon}\right)^{\frac{1}{\beta}}\right]^{\frac{\beta}{1-\beta}}(\theta^2)^{\frac{1}{1-\beta}} + \gamma^1(\theta^2-\varepsilon)^{\frac{1}{1-\beta}}(\delta)^{\frac{\beta}{1-\beta}}\right\}^{\frac{1-\beta}{1-2\beta}}}{\left\{\gamma^1\delta+\gamma^2\left[p_q^1-(p_q^1-\delta)\left(\frac{\theta^2-\varepsilon}{\theta^2-\gamma^1\varepsilon}\right)^{\frac{1}{\beta}}\right]\right\}^{\frac{\beta}{1-2\beta}}}. \tag{D63}$$

E Proof of Proposition 3

Part (i) 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 ii) 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

$$-\frac{\partial R^1(y^1,e_s^1)}{\partial y^1}+\lambda^2\frac{\partial \widehat{R}^2(y^1,e_s^1)}{\partial y^1}-\lambda^1\frac{\partial R^1(y^1,e_s^1)}{\partial y^1}+\mu\gamma^1 = 0, \tag{E13}$$

$$-\frac{\partial R^1(y^1,e_s^1)}{\partial e_s^1}+\lambda^2\frac{\partial \widehat{R}^2(y^1,e_s^1)}{\partial e_s^1}-\lambda^1\frac{\partial R^1(y^1,e_s^1)}{\partial e_s^1} = 0, \tag{E14}$$

$$1-\lambda^2+\lambda^1-\mu\gamma^1 = 0, \tag{E15}$$

$$-\lambda^2\frac{\partial R^2(y^2,e_s^2)}{\partial y^2}+\lambda^1\frac{\partial \tilde{R}^1(y^2,e_s^2)}{\partial y^2}+\mu\gamma^2 = 0, \tag{E16}$$

$$-\lambda^2\frac{\partial R^2(y^2,e_s^2)}{\partial e_s^2}+\lambda^1\frac{\partial \tilde{R}^1(y^2,e_s^2)}{\partial e_s^2} = 0, \tag{E17}$$

$$\lambda^2-\lambda^1-\mu\gamma^2 = 0. \tag{E18}$$

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

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

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. (32). We therefore have that

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

Combining (E16) and (E18) gives

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

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. (33). The result stated by eq. (34) can then be easily obtained combining the results provided by (E19) and (E21).

From (E14)-(E15) we have that

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

or equivalently

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

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

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

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

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

from which eq. (29) 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

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

From (E13) and (E15) we have that

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

or equivalently

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

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

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

from which eq. (30) 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. (31) can be easily obtained combining the results provided by (29) and (30).

F Proof of Proposition 4

Condition (37) tells us that the right hand side of (30) can be interpreted as $T_1'\left(y^1, e_s^1\right)$, namely, the marginal income tax rate. Condition (38) tells us that the right hand side of (29) can be interpreted as $T_2'\left(y^1, e_s^1\right)/p_q^1$, namely, the marginal subsidy on e_s faced by low-ability agents (discounted by p_q^1). Next, we show that (42) and (43) imply that the tax function implementing the separating tax equilibrium is characterized by the property that $T_1'\left(y^2, e_s^2\right) = T_2'\left(y^2, e_s^2\right) = 0$. Exploiting (43), from (E19) we obtain that

$$T_2\left(y^2, e_s^2\right) = \left(\frac{\lambda^1}{\gamma^2} - \frac{\phi}{1-\phi}\right)\left(p_q^1 - p_q^2\right)\frac{h_1\left(e_s^2, f\left(y^2, e_s^2, \theta^2\right)\right)}{h_2\left(e_s^2, f\left(y^2, e_s^2, \theta^2\right)\right)}. \tag{F1}$$

Exploiting (42), from (E21) we have that

$$T_1\left(\theta^2 h\left(e_s^2, e_q^2\right), e_s^2\right) = \left(\frac{\phi}{1-\phi} - \frac{\lambda^1}{\gamma^2}\right)\frac{p_q^1 - p_q^2}{\theta^2 h_2\left(e_s^2, f\left(y^2, e_s^2, \theta^2\right)\right)}. \tag{F2}$$

Taking into account that the government's budget constraint requires that $c^2 = y^2 + \frac{\gamma^1}{\gamma^2}\left(y^1 - c^1\right)$, notice that for given values of c^1, y^1 and e_s^1, the socially optimal values of y^2 and e_s^2 are those maximizing

$$y^2 + \frac{\gamma^1}{\gamma^2}\left(y^1 - c^1\right) - R^2\left(y^2, e_s^2\right) \tag{F3}$$

subject to the IC-constraint

$$c^1 - R^1\left(y^1, e_s^1\right) \geq y^2 + \frac{\gamma^1}{\gamma^2}\left(y^1 - c^1\right) - \tilde{R}^1(y^2, e_s^2). \tag{F4}$$

But these are precisely the choices that would be made by type-2 agents if they were subject to a lump-sum tax of amount $\frac{\gamma^1}{\gamma^2}\left(y^1 - c^1\right)$. This means that, at an optimum, $\frac{\lambda^1}{\gamma^2} = \frac{\phi}{1-\phi}$ or, equivalently (taking into account that from the government's first order conditions we have that $\gamma^2 = \lambda^2 - \lambda^1$), $\phi = \lambda^1/\lambda^2$. One can then conclude that, along the tax function implementing the socially optimal separating equilibrium,

$$T_1\left(y^2, e_s^2\right) = T_2\left(y^2, e_s^2\right) = 0. \tag{F5}$$

Exploiting (38), and taking into account that $\frac{\partial R^1\left(y^1, e_s^1\right)}{\partial e_s^1} = p_s - p_q^1 \frac{h_1\left(e_s^1, f\left(y^1, e_s^1, \theta^1\right)\right)}{h_2\left(e_s^1, f\left(y^1, e_s^1, \theta^1\right)\right)}$, it follows from (E23) that

$$T_2\left(y^1, e_s^1\right) \quad = \quad \frac{\lambda^2}{\gamma^1}\left[\frac{h_1\left(e_s^1, f\left(y^1, e_s^1, \bar{\theta}\right)\right)}{h_2\left(e_s^1, f\left(y^1, e_s^1, \bar{\theta}\right)\right)}p_q^2 - \frac{h_1\left(e_s^1, f\left(y^1, e_s^1, \theta^1\right)\right)}{h_2\left(e_s^1, f\left(y^1, e_s^1, \theta^1\right)\right)}p_q^1\right] < 0. \tag{F6}$$

Exploiting (37) , it follows from (E29) that

$$T_1\left(y^1,e_s^1\right)=\frac{\lambda^2}{\gamma^1}\left(\frac{p_q^1}{\theta^1 h_2\left(e_s^1,f\left(y^1,e_s^1,\theta^1\right)\right)}-\frac{p_q^2}{\overline{\theta}h_2\left(e_s^1,f\left(y^1,e_s^1,\overline{\theta}\right)\right)}\right)>0. \tag{F7}$$

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 $e_s \geq \widehat{e}_s$:

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

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 CEA satisfies (E3), it follows that the effort mix $(\widehat{e}_s,\widehat{e}_q)$ satisfies the first order condition (G5). Moreover, given that the CEA 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:

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

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 $\theta^1 h(e_s,e_q)-T\left(\theta^1 h(e_s,e_q)\right)-p_s e_s-p_q^1 e_q$ 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 $y^2 \in (\widehat{y}, \theta^2 h_2(\widehat{e}_s, \widehat{e}_q)]$.[29] A similar argument 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

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

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

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

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

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

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

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

[29] 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:

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

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

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

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

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

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 $\overline{\theta} h(\widehat{e}_s, e_q^{\min})$.[30] Clearly, given that $p_q^2 < p_q^1$ and the income tax

[30]Notice that we can safely disregard the case of pooling at $\overline{\theta} h(\widehat{e}_s, e_q^{\min}) < y < \theta^2 h(\widehat{e}_s, e_q^{\min})$; 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

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 $y \in [\widehat{y}, \theta^2 h(\widehat{e}_s, e_q^{\min}))$. 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 $1 - \frac{p_q^2}{\overline{\theta} h_2(\widehat{e}_s, \widehat{e}_q)}$.[31] Notice, also, that the benefit of pooling at a marginally higher level of income is decreasing in income.[32] 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

$$T(y) = \begin{cases} \left(\frac{1}{\theta^1} - \frac{1}{\overline{\overline{\theta}}}\right) \frac{p_q^1}{h_2(\widehat{e}_s, \widehat{e}_q)} \widehat{y} + \left(\frac{1}{\overline{\overline{\theta}}} - \frac{1}{\theta^1}\right) \frac{p_q^1}{h_2(\widehat{e}_s, \widehat{e}_q)} y, & \text{for all } y \in [0, \widehat{y}] \\ (y - \widehat{y}) \max\left\{1 - \frac{p_q^2}{\overline{\theta} h_2(\widehat{e}_s, \widehat{e}_q)}, \frac{\left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s \widehat{e}_s - p_q^2 \widehat{e}_q\right]}{\underline{y}^{sep} - \widehat{y}}\right\}, & \text{for all } y > \widehat{y}. \end{cases} \tag{G14}$$

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 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

pre-tax income higher than $\overline{\theta} h(\widehat{e}_s, e_q^{\min})$.

[31] 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.

[32] 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 $\widehat{y} + 2\epsilon$.

net income denoted by $\hat{c}$. Lemma 1 below shows that pooling equilibria where all agents choose the same effort vector do not exist.

Lemma 1. *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 $\widehat{y} = \theta^2 h\left(e_s, e_q\right)$ is strictly below the isoquant $\widehat{y} = \bar{\theta}h\left(e_s, e_q\right)$, 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 2 shows that such an equilibrium will never be the social optimum.

Lemma 2. *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.[33] □

[33] 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).

Lemma 1 and Lemma 2 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.

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:

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

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

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

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:

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

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

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

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 3 shows that, in equilibrium, this out-of-equilibrium deviation would never be profitable for type-1 agents.

Lemma 3. *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 1, 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 4. *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:

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

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:

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

subject to the government budget constraint

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

and the downward IC-constraints

$$c^2 - R^2(y^2, c^2, y^1, c^1) \geq c^1 - \widetilde{R}^2(y^1), \tag{H18}$$
$$c^2 - R^2(y^2, c^2, y^1, c^1) \geq c^1 - \hat{R}^2(y^1). \tag{H19}$$

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

$$-\frac{\partial R^1(y^1)}{\partial y^1} - \left(\lambda^{2s} + \lambda^{2p}\right) \frac{\partial R^2(y^2, c^2, y^1, c^1)}{\partial y^1} + \lambda^{2s} \frac{\partial \widetilde{R}^2(y^1)}{\partial y^1} + \lambda^{2p} \frac{\partial \widehat{R}^2(y^1)}{\partial y^1} + \mu\gamma^1 = 0, \tag{H20}$$

$$1 - \left(\lambda^{2s} + \lambda^{2p}\right) - \left(\lambda^{2s} + \lambda^{2p}\right) \frac{\partial R^2(y^2, c^2, y^1, c^1)}{\partial c^1} - \mu\gamma^1 = 0, \tag{H21}$$

$$-\left(\lambda^{2s} + \lambda^{2p}\right) \frac{\partial R^2(y^2)}{\partial y^2} + \mu\gamma^2 = 0, \tag{H22}$$

$$\lambda^{2s} + \lambda^{2p} - \left(\lambda^{2s} + \lambda^{2p}\right) \frac{\partial R^2(y^2, c^2, y^1, c^1)}{\partial c^2} - \mu\gamma^2 = 0. \tag{H23}$$

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

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

But given that $\frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial c^1} + \frac{\partial R^2(y^2,c^2,y^1,c^1)}{\partial c^2} = 0$ (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 $R^2(y^2, c^2, y^1, c^1)$), eq. (H24) implies that $\mu = 1$. Taking this into account we can rewrite (H22) and (H23) as, respectively

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

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

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

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

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 $T(y)$. This can be described as follows:

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

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

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

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

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

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

or equivalently:

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

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 $R^2\left(y^2,c^2,y^1,c^1\right)$. Thus, (H37) can be equivalently restated as

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

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:

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

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

$$V^1(y, y - T(y)) = y - T(y) - R(y), \tag{H40}$$

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

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

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

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

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

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

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

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

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

$$\frac{de_s}{dp_q} = \frac{\theta h_1 h_2}{\Gamma}, \tag{H53}$$

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

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

$$\begin{aligned}
&\frac{\partial(\theta h_1)}{\partial \theta} + \frac{\partial(\theta h_1)}{\partial p_q}\frac{dp_q}{d\theta} \\
&= h_1 + \theta\left[h_{11}\left(-h\frac{p_q h_{12} - p_s h_{22}}{\Gamma} + \frac{\theta h_1 h_2}{\Gamma}\frac{dp_q}{d\theta}\right) + h_{12}\left(h\frac{p_q h_{11} - p_s h_{12}}{\Gamma} - \frac{\theta (h_1)^2}{\Gamma}\frac{dp_q}{d\theta}\right)\right] \\
&= h_1 + \frac{\theta}{\Gamma}\left[-p_q h h_{11} h_{12} + p_s h h_{11} h_{22} + p_q h h_{12} h_{11} - p_s h h_{12} h_{12} + \theta h_1 h_2 h_{11}\frac{dp_q}{d\theta} - \theta (h_1)^2 h_{12}\frac{dp_q}{d\theta}\right] \\
&= h_1 + \frac{\theta}{\Gamma}\left[(h_{11}h_{22} - h_{12}h_{12}) p_s h + (h_2 h_{11} - h_1 h_{12})\theta h_1 \frac{dp_q}{d\theta}\right].
\end{aligned} \tag{H55}$$

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

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

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:

$$p_s \widetilde{e}_s^2 + p_q^1 \widetilde{e}_q^2 = p_s e_s^1 + p_q^1 e_q^1, \tag{H57}$$

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

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:

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

Defining Ψ as

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

we have that

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

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

and therefore

$$\begin{aligned} \frac{\partial \widetilde{R}^2\left(y^1\right)}{\partial y^1} &= p_s \frac{d\widetilde{e}_s^2}{dy^1} + p_q^2 \frac{d\widetilde{e}_q^2}{dy^1} = p_s \frac{\frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right) - p_q^1}{\Psi} + p_q^2 \frac{p_s - \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)}\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\Psi} \\ &= \frac{\left(p_q^2 - p_q^1\right)p_s + \left(p_s\right)^2 \frac{\theta^2 h_2\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)} - p_s p_q^2 \frac{\theta^2 h_1\left(\widetilde{e}_s^2, \widetilde{e}_q^2\right)}{\theta^1 h_1\left(e_s^1, e_q^1\right)}}{\Psi}. \end{aligned} \tag{H63}$$

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

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

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

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

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 $\widehat{R}^2(y^1)$ 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 $\frac{\partial \widehat{R}^2(y^1)}{\partial y^1} = \frac{p_s}{\overline{\theta} h_1(\widehat{e}_s, \widehat{e}_q)} = \frac{p_q^2}{\overline{\theta} h_2(\widehat{e}_s, \widehat{e}_q)}$. It then follows that, when $\lambda^{2p} > 0$, it must necessarily be that

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

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 $\frac{\partial R^1(y^1)}{\partial y^1} - \frac{\partial \widehat{R}^2(y^1)}{\partial y^1} > 0$.

Going back to (H43) and summarizing our results for $T'(y^1)$, we have that $T'(y^1)$ 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

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

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

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

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

$$T'(y^1) = \left[\lambda^{2s}\left(\frac{h_1(\widetilde{e}_s^2, \widetilde{e}_q^2)}{h_1(e_s^1, e_q^1)}\frac{\theta^2}{\theta^1} - 1\right)\frac{p_q^2 - p_q^1}{\Psi} + \lambda^{2p}\left(\frac{1}{\theta^1 h_1(e_s^1, e_q^1)} - \frac{1}{\overline{\theta} h_1(\widehat{e}_s, \widehat{e}_q)}\right)\right]\frac{p_s}{\gamma^1} > 0. \tag{H71}$$

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

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

I The case when both signals are observable

We turn first to formulate the maximization program associated with a CEA 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:

$$\max_{c,e_s,e_q} \left[c - (e_s + e_q p_q^1)\right], \tag{I1}$$

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:

$$\max_{\{(c^i,e_s^i,e_q^i)\}_{i=1,2}} \left[c^1 - (e_s^1 + e_q^1 p_q^1)\right] \tag{I4}$$

subject to:

$$c^1 - (e_s^1 + e_q^1 p_q^1) \geq c^2 - (e_s^2 + e_q^2 p_q^1) \tag{I5}$$
$$c^2 - (e_s^2 + e_q^2 p_q^2) \geq c^1 - (e_s^1 + e_q^1 p_q^2) \tag{I6}$$
$$\gamma^1 \theta^1 h(e_s^1, e_q^1) + \gamma^2 \theta^2 h(e_s^2, e_q^2) \geq \gamma^1 c^1 + \gamma^2 c^2 \tag{I7}$$

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:

$$\Delta Y = [\gamma^2\theta^2\delta - \gamma^1\theta^1\varepsilon][h_1(e_s^*, e_q^*) + h_2(e_s^*, e_q^*)] \tag{I8}$$

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 obviously 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

$$\frac{h_1\left(e_s^1, e_q^1\right)}{h_2\left(e_s^1, e_q^1\right)} - \frac{p_s}{p_q^1} = -\frac{\gamma^2\left(p_q^1 - p_q^2\right)}{\gamma^1 p_q^1 + \gamma^2\left(p_q^1 - p_q^2\right)} p_s < 0. \tag{I12}$$

J The observable signal is e_q instead of e_s

We divide this Appendix in two parts. In part i) 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 ii) 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 i) Consider the right-hand side of the (downward) IC-constraint (22), 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 $\{(y^1, c^1, e_s^1, e_q^1), (y^2, c^2, e_s^2, e_q^2)\}$ 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 ii) Consider the following. For a given isoquant $\theta^1 h(e_s, e_q) = y^1$, assume that type-1 agents are induced to choose the effort mix (e_s^1, e_q^1) that satisfies the no-distortion condition $\frac{h_1(e_s^1,e_q^1)}{h_2(e_s^1,e_q^1)} = \frac{p_s}{p_q^1}$. 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 $\overline{\theta} h(e_s, e_q^1) = y^1$, (which implies that $\widehat{e}_s^2 < e_s^1$). Given that $p_q^2 < p_q^1$, but at the same time $\frac{h_1(\widehat{e}_s^2,\widehat{e}_q^2)}{h_2(\widehat{e}_s^2,\widehat{e}_q^2)} > \frac{h_1(e_s^1,e_q^1)}{h_2(e_s^1,e_q^1)}$, it follows that one cannot a priori establish whether the effort-mix of type-2 mimickers is distorted towards e_s (i.e., $\frac{h_1(\widehat{e}_s^2,\widehat{e}_q^2)}{h_2(\widehat{e}_s^2,\widehat{e}_q^2)} < \frac{p_s}{p_q^2}$) or towards e_q (i.e., $\frac{h_1(\widehat{e}_s^2,\widehat{e}_q^2)}{h_2(\widehat{e}_s^2,\widehat{e}_q^2)} > \frac{p_s}{p_q^2}$).

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 (25) 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:

$$\max_{c^1,c^2,y^1,y^2,e_s^2} c^1 - R^1\left(y^1\right) \tag{K1}$$

subject to the budget constraint

$$\sum_i \gamma^i(y^i - c^i) = 0, \tag{K2}$$

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 (25) to express e_q^2 as a function of y^2 and e_s^2, namely $e_q^2 = (y^2/\theta^2)^{\frac{1}{\beta}}/e_s^2$.

Exploiting the assumption (25) 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:

$$\min_{e_s,e_q} p_s e_s + p_q^1 e_q \quad \text{subject to} \quad (e_s e_q)^\beta \theta^1 = y^1. \tag{K6}$$

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.[34] 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

$$\begin{aligned} &c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \\ &\geq c^1 - \sqrt{\frac{p_s}{p_q^1} (y^1)^{\frac{1}{\beta}} \frac{\left(p_q^1 - p_q^2\right)\left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}} + 2p_q^2\sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}}}\left[\sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} - \left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}}}\right]}{\sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}}} + \sqrt{\left(\frac{1}{\theta^1}\right)^{\frac{1}{\beta}} - \left(\frac{1}{\theta^2}\right)^{\frac{1}{\beta}}}}}. \end{aligned} \tag{K11}$$

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:

$$c^2 - p_s e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{\frac{1}{\beta}} \frac{p_q^2}{e_s^2} \geq c^1 - 2\sqrt{\left(y^1/\overline{\theta}\right)^{1/\beta} p_s p_q^2}. \tag{K12}$$

[34] 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:

$$e_q^{i*} = \left(\frac{y^i}{\theta^i}\right)^{1/\beta} \frac{1}{e_s^i}, (i = 1, 2) \quad \text{and} \quad \hat{e}_q = \left(\frac{y^1}{\bar{\theta}}\right)^{1/\beta} \frac{1}{e_s^1}. \tag{K13}$$

Thus, the IC-constraints (22)–(23) can be written as follows:

$$c^1 - p_s^1 e_s^1 - \left(\frac{y^1}{\theta^1}\right)^{1/\beta} \frac{p_q^1}{e_s^1} \geq c^2 - p_s^1 e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{1/\beta} \frac{p_q^1}{e_s^2}, \tag{K14}$$

$$c^2 - p_s^2 e_s^2 - \left(\frac{y^2}{\theta^2}\right)^{1/\beta} \frac{p_q^2}{e_s^2} \geq c^1 - p_s^2 e_s^1 - \left(\frac{y^1}{\bar{\theta}}\right)^{1/\beta} \frac{p_q^2}{e_s^1}. \tag{K15}$$

When implementing a pooling equilibrium, IC-constraints can be neglected, and the government chooses (y, e_s) to maximize

$$u^1 = y - p_s^1 e_s + p_q^1 \widehat{e}_q(y, e_s), \tag{K16}$$

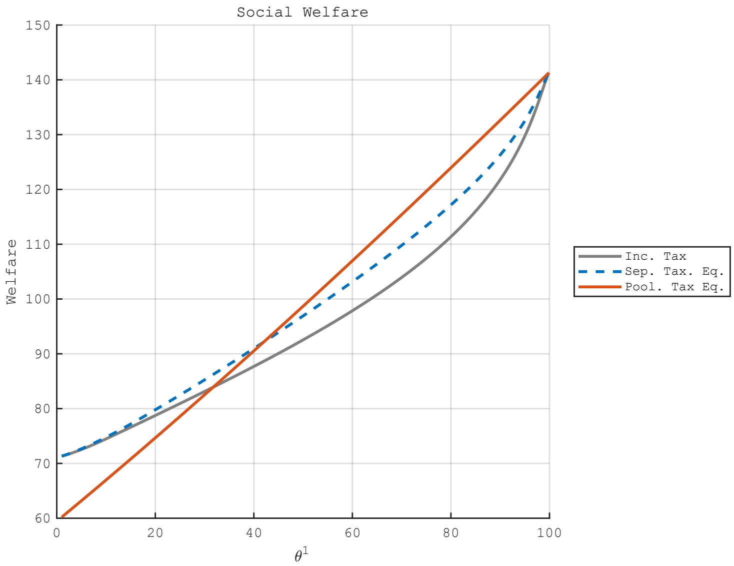
where $\widehat{e}_q(e_s, y)$ is the value of e_q which solves the equation $y = (e_s e_q)^\beta \bar{\theta}$.

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 CEA, 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 CEA 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 CEA (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 CEA (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 CEA 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.