INTERNATIONAL ECONOMIC REVIEW
Vol. 0, No. 0, January 2025 DOI: 10.1111/iere.12756

OPTIMAL REDISTRIBUTION AND EDUCATION SIGNALING*

By Spencer Bastani, Tomer Blumkin, and Luca Micheletto

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

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

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 (IC) considerations justify the introduction of distortions, often in the form of positive marginal tax rates. These distortions facilitate targeted transfers to low-income individuals while providing incentives for high-income individuals to exert labor effort.

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

*Manuscript received August 2022; revised October 2024.

We are grateful to the editor, Dirk Krueger, and three anonymous reviewers for very helpful comments. We are also grateful to Håkan Selin and Anna Sjögren for providing valuable comments on an earlier draft of the article. Spencer Bastani gratefully acknowledges financial support from Jan Wallanders och Tom Hedelius stiftelse (grant P2018-162). Please address correspondence to: Spencer Bastani, Institute for Evaluation of Labour Market and Education Policy (IFAU), Uppsala, Sweden. E-mail: *spencer.bastani@ifau.uu.se*.

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

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

While the current article shares the feature of a second layer of asymmetric information with the two studies mentioned above, it differs from them in at least five ways. First, we focus on worker signaling through investment in education. Despite its central role in economics, its prominence in economics curricula around the world, and its relevance in policy discussions (e.g., Caplan 2018), it is surprising that signaling has been addressed in only a few papers in the vast literature on optimal tax design since the seminal work of Mirrlees (1971).[1] Second, we consider multidimensional signaling in the context of taxation, which allows us to retain the realistic Mirrleesian feature that firms are more informed about workers than the government. Third, we consider tax systems that depend not only on income but also on the signals that the government can observe in the labor market.[2] Fourth, in line with Bastani et al. (2015) but in contrast to Stantcheva (2014), we emphasize the important role of redistribution through the wage (as opposed to the income) channel.[3] Finally, in contrast to Stantcheva (2014), we show that the presence of adverse selection due to asymmetric information between firms and workers does not necessarily lead to a higher level of welfare in the social optimum than that achieved in a Mirrleesian setup where worker types are observable by firms.

The details of our analysis are as follows: Consistent with the prevailing literature on optimal income taxation, we assume that workers differ in their intrinsic productive capabilities, which are unobservable to the government. However, unlike most studies in this area, and in line with the two studies discussed above, we extend this unobservability to potential employers. The distinguishing feature of our analysis is that workers must signal their pro-

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

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

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

ductivity to firms by making costly effort decisions, allowing for information transmission between workers and firms along two dimensions: the quantity (e.g., years of schooling) and the quality (e.g., the difficulty or intensity of a particular educational pathway) of their education. Whereas quantity is observable to both the government and employers, quality is only observable to employers, reflecting an environment in which employers have better information than the government. To make signaling feasible, we assume that workers differ not only in their innate productive abilities but also in their costs of signaling (e.g., the cost of obtaining education), with signals representing components of educational effort that realistically also increase human capital.[4]

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

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

We then characterize the constrained efficient allocation assuming a max–min social objective, called the max-min optimum (MMO), and show that it is given by either an STE or a PTE, depending on which equilibrium configuration produces the highest level of social welfare. We derive necessary and sufficient conditions for the MMO to feature predistribution, emphasizing the role of both differences in agents' innate productivities and differences in the costs of signaling. Note that in our setting, incentive constraints can flow from low- to high types as well as from high- to low types. Low types may have an incentive to invest more in signaling in order to qualify for higher compensation, whereas high types may have an incentive to mimic low types in order to qualify for a more lenient tax treatment.

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

The policy framework required to implement the MMO (whether provided by an STE or a PTE) can take several forms. We propose two simple implementation schemes that combine

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

a nonlinear income tax with income-tested education subsidies or mandates. A nonlinear income tax system—in practice often piecewise linear with multiple brackets—encourages individuals to locate at targeted income levels. Income-tested subsidies and mandates, on the other hand, ensure that higher-ability individuals are not incentivized to deviate from their lower-ability counterparts by choosing lower levels of education (i.e., lower quantity effort) conditional on income level. The design of income-tested education subsidies and mandates is driven by the need to provide the right incentives locally—at a given income level—without distorting the incentives to acquire education at other income levels.[5]

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

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

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

The article is organized as follows: In Section 2, we outline the structure of the game, the equilibrium concept that we use, and the role of government in the economy. We then define the STE and PTE in the presence of a general tax function, and describe the government optimization problem and the concept of MMO. In Section 3, we characterize the optimal wedges

[5] Education subsidies have traditionally been used to correct market failures and redistribute income. In the optimal tax literature, they serve two primary functions: (i) to mitigate the negative effects of income taxation on human capital formation, and (ii) to enhance redistribution. See, for example, Ulph (1977), Tuomala (1986), Boadway and Marchand (1995), Brett and Weymark (2003), Bovenberg and Jacobs (2005), and Maldonado (2008). Some studies, such as Blumkin and Sadka (2008), also explore the possibility of education taxes due to the positive correlation between education and unobserved ability. More recently, Findeisen and Sachs (2016) examine income-contingent student loans and suggest that it may be optimal for very high-income individuals to repay more than the value of their loans, effectively creating an education tax.

associated with the MMO. In Section 4, we discuss how these wedges can be implemented using means-tested education subsidies or mandates. In Section 5, we discuss alternative observational assumptions and some robustness and extensions of the basic setup. Section 6 concludes. Most of our formal derivations and proofs are relegated to the Online Appendix.

2. THE MODEL

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

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

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

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

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

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

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

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

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

where c is consumption and

$$R^i(e_s, e_q) = p^i_s e_s + p^i_q e_q \tag{3}$$

is the cost function for agents of type i, where p^i_s and p^i_q 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:

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

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

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

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

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

$$\left(y^{1*}, e^{1*}_s, e^{1*}_q\right) = \arg\max_{y^1, e^1_s, e^1_q} \left\{y^1 - T\left(y^1, e^1_s, e^1_q\right) - p_s e^1_s - p^1_q e^1_q\right\} \quad \text{subject to}$$
$$y^1 \leq \Theta(e^1_s, e^1_q) \cdot h(e^1_s, e^1_q). \tag{5}$$
$$\left(y^{2*}, e^{2*}_s, e^{2*}_q\right) = \arg\max_{y^2, e^2_s, e^2_q} \left\{y^2 - T\left(y^2, e^2_s, e^2_q\right) - p_s e^2_s - p^2_q e^2_q\right\} \quad \text{subject to}$$
$$y^2 \leq \Theta(e^2_s, e^2_q) \cdot h(e^2_s, e^2_q), \tag{6}$$

and the government's revenue constraint holds:

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

The wage offer function $\Theta(e_s, e_q)$ is updated as firms learn from the observed signals provided by workers. The equilibrium wage function, $\Theta^*(e_s, e_q)$, represents the wage that emerges

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

when firms' beliefs are consistent with the observed equilibrium behavior of workers. An equilibrium is a stable point, in the sense that beliefs are correct given strategies, and strategies are sequentially rational given beliefs.

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

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

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

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

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

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

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

We begin by characterizing the STE.

LEMMA 1 (SEPARATING TAX EQUILIBRIUM, STE). *Suppose that, given the general tax function $T(y, e_s, e_q)$, the allocations*

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

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

The equilibrium allocations satisfy the following conditions:

(a) OPTIMALITY FOR TYPE-1 WORKERS

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

subject to:

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

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

(b) OPTIMALITY FOR TYPE-2 WORKERS

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

subject to:

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

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

(c) NO PROFITABLE DEVIATIONS OFF THE EQUILIBRIUM PATH
No allocation $(y, e_s, e_q) \in \mathbb{R}_+^3$ satisfies:

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

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

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

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

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

We turn next to characterize the pooling refined PBE.

LEMMA 2 (POOLING TAX EQUILIBRIUM, PTE). *Suppose that, given the general tax function $T(y, e_s, e_q)$, the singleton allocation*

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

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

The equilibrium allocation satisfies the following conditions:

(a) *NO PROFITABLE DEVIATIONS FOR TYPE-1 WORKERS*
There is no $(y^1, e_s^1, e_q^1) \in \mathbb{R}_+^3$ satisfying $y^1 \leq \theta^1 h(e_s^1, e_q^1)$ such that:

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

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

(b) *NO PROFITABLE DEVIATIONS FOR TYPE-2 WORKERS*
There is no $(y^2, e_s^2, e_q^2) \in \mathbb{R}_+^3$ satisfying $y^2 \leq \theta^2 h(e_s^2, e_q^2)$ such that:

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

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

(c) *NO JOINT DEVIATIONS TO A NEW POOLING ALLOCATION*
There is no $(\widehat{y}, \widehat{e}_s, \widehat{e}_q) \in \mathbb{R}_+^3 \setminus \{\widehat{y}^*, \widehat{e}_s^*, \widehat{e}_q^*\}$ satisfying $\widehat{y} \leq \bar{\theta} h(\widehat{e}_s, \widehat{e}_q)$ such that, for both $i = 1, 2$,

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

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

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

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

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

DEFINITION 1 (MAX-MIN OPTIMUM, MMO). An MMO is given by a solution to:

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

subject to the government revenue constraint

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

where the wage rate is given by

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

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

and the IC constraints are

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

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

where

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

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

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

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

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

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

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

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

[12] For example, if the revenue constraint is slack (a budget surplus), the government could offer a small lump-sum transfer to both types. Continuity would ensure that the revenue constraint is not violated. IC would be maintained

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

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

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

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

2.3. *When is Predistribution Optimal?* Let us now analyze the social optimality of both STE and PTE configurations. In an STE, the government typically cannot fully eliminate the information rents that arise from productivity differences between workers. In contrast, in a PTE, the government fully eliminates these information rents by enforcing full wage compression, although the PTE generally has less desirable efficiency properties. The equity-efficiency trade-off between pooling and separation depends critically on the magnitude of the produc-

by the linearity of utility in consumption. If the firm makes positive profits, the government could slightly increase the compensation level, y, which would maintain nonnegative profits due to continuity. This would create a fiscal surplus that could be refunded as a lump-sum transfer.

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

tivity differences, and of the heterogeneity in the costs associated with acquiring the quality signal e_q, between the two types of workers.[14] Proposition 1 summarizes the main results.

PROPOSITION 1 (OPTIMALITY OF PREDISTRIBUTION). *Let the ability advantage of type-2 agents be denoted by $\epsilon = \theta^2 - \theta^1 \geq 0$, and the cost disadvantage of type-1 agents by $\delta = p_q^1 - p_q^2 \geq 0$. The MMO can be characterized as follows:*

(a) There is a nonempty set of parameters in the (ϵ, δ)-space for which the MMO is given by a PTE (and thus features predistribution).
(b) For any $\epsilon > 0$, there exists a threshold $\delta^*(\epsilon) \geq 0$ such that the MMO is given by an STE for $\delta > \delta^*$ and a PTE for $\delta < \delta^*$.
(c) There exists some cutoff $\varepsilon^* > 0$ such that $\delta^*(\varepsilon) = 0$ for any $\varepsilon > \varepsilon^*$ (and thus the MMO is an STE for all δ), whereas $\delta^*(\varepsilon) > 0$ for all $\varepsilon < \varepsilon^*$ (so the MMO is either an STE or a PTE, depending on the value of δ).

PROOF. See Online Appendix B. □

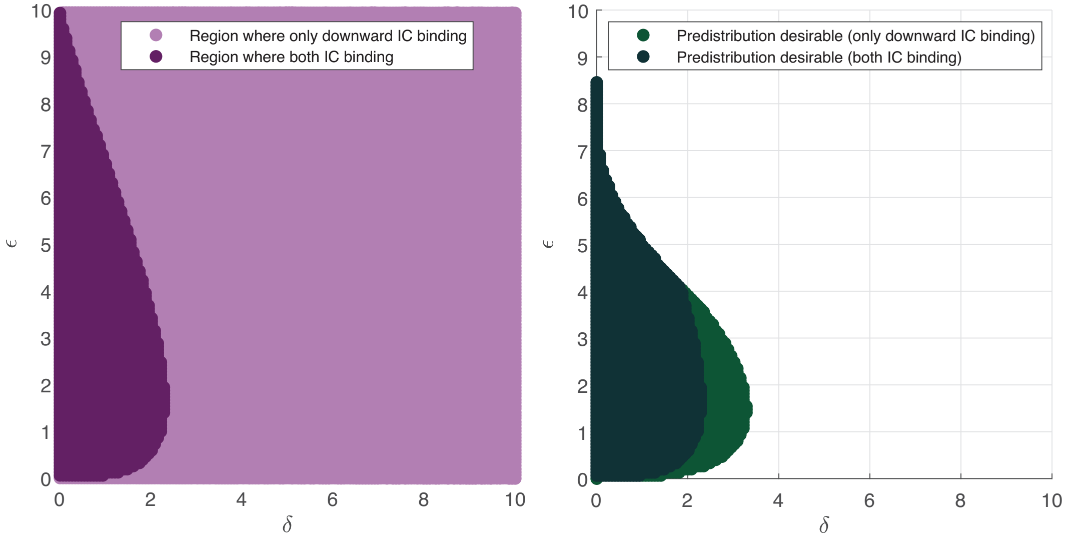
While augmenting the income tax system with taxes or subsidies on education may improve redistribution under separation—by alleviating the binding IC constraints faced by the government—Proposition 1 outlines cases where taxing or subsidizing education, by enabling the implementation of a PTE, increases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

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

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

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

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



NOTES: 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

ILLUSTRATION OF THE CASE FOR PREDISTRIBUTION AND THE PATTERN OF BINDING IC CONSTRAINTS

setting than in a "Mirrleesian" setting where firms observe workers' productivity is shown formally in Proposition 2.

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

PROOF. See Online Appendix C. □

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

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

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

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

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

3. WEDGES IN THE CONSTRAINED EFFICIENT ALLOCATION

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

PROPOSITION 3.

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

$$\widehat{\Omega}^1_{e_s,e_q} \equiv \widehat{MRTS}^1 - \frac{p_s}{p_q^1} = 0 \quad \textit{and} \quad \widehat{\Omega}^2_{e_s,e_q} \equiv \widehat{MRTS}^2 - \frac{p_s}{p_q^2} < 0, \tag{27}$$

$$\widehat{\Omega}^1_{e_s,c} \equiv 1 - \frac{p_s}{\theta^1 h_1(\widehat{e_s}, \widehat{e_q})} = \widehat{\Omega}^1_{e_q,c} = 1 - \frac{p_q^1}{\theta^1 h_2(\widehat{e_s}, \widehat{e_q})} < 0, \tag{28}$$

$$\widehat{\Omega}^2_{e_q,c} \equiv 1 - \frac{p_q^2}{\theta^2 h_2(\widehat{e_s}, \widehat{e_q})} > \widehat{\Omega}^2_{e_s,c} = 1 - \frac{p_s}{\theta^2 h_1(\widehat{e_s}, \widehat{e_q})} > 0, \tag{29}$$

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

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

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

$$\Omega^1_{e_q,c} \equiv 1 - \frac{p_q^1}{\theta^1 h_2\left(e_s^1, e_q^1\right)} = \frac{\lambda^2}{\gamma^1}\left(\frac{p_q^1}{\theta^1 h_2\left(e_s^1, e_q^1\right)} - \frac{p_q^2}{\overline{\theta} h_2\left(e_s^1, \widehat{e_q^2}\right)}\right) > 0, \tag{31}$$

$$\Omega^1_{e_s,c} \equiv 1 - \frac{p_s}{\theta^1 h_1\left(e_s^1, e_q^1\right)} = \frac{\lambda^2}{\gamma^1}\left(\frac{h_1\left(e_s^1, \widehat{e_q^2}\right)}{h_1\left(e_s^1, e_q^1\right)}\frac{1}{\theta^1} - \frac{1}{\overline{\theta}}\right)\frac{p_q^2}{h_2\left(e_s^1, \widehat{e_q^2}\right)}, \tag{32}$$

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

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

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

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

PROOF. See Online Appendix E. □

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

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

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

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

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

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

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

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

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

Finally, looking at (35), we can see that e_s^2 is always undistorted relative to consumption. The reason for this is a combination of two circumstances. First, the marginal cost of acquiring e_s is the same for all agents. Second, when acting as mimickers, type-1 agents replicate the

[17] A necessary but not sufficient condition for $\lambda^1 > 0$ is that the upward IC constraint associated with the low-skilled workers is binding under *laissez-faire*. This is because the redistribution in favor of type-1 agents that occurs through the tax system necessarily reduces the incentive for type-1 agents to mimic type-2 agents.

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 THROUGH MEANS-TESTED EDUCATION SUBSIDIES OR MANDATES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

PROOF. See Online Appendix F. □

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

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

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

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

Regarding how much larger than σ^* the subsidy rate for $e_s \leq e_s^1$ should be, one should consider the various off-equilibrium strategies available to type-2 agents if they decide to behave as mimickers. One possibility is for them to earn y^1 by pooling with their type-1 counterpart on a common effort vector. In this case, type-2 agents would be rewarded according to the average productivity $\overline{\theta}$, and (37) defines the subsidy rate needed to induce them to choose $e_s = e_s^1$. In fact, $\widehat{\sigma}$ is defined to reflect the wedge faced by type-2 agents at the off-equilibrium effort mix $\left(e_s^1, \widehat{e}_q^2\right)$, where $\widehat{e}_q^2$ is implicitly defined as the solution to the equation:[19]

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

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

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

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

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

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

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

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

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

Whether or not an additional third segment is needed depends on the location of the point defined as $\left(e_s^{int}, e_q^{int}\right)$ in Proposition 4, where the two isocost lines (38) and (39) intersect. The first isocost line is for type-1 agents and passes through the point $\left(e_s^1, e_q^1\right)$; the second (flatter than the first because $p_q^1 > p_q^2$) belongs to type-2 agents and passes through the point $\left(e_s^1, \widehat{e}_q^2\right)$. If $\theta^2 h\left(e_s^{int}, e_q^{int}\right) \leq y^1$, there is no need for an additional segment on the subsidy schedule. The reason is that on the isoquant $y^1 = \theta^2 h\left(e_s, e_q\right)$ there is no pair $\left(e_s, e_q\right)$ that is also below the isocost line (39) (i.e., for type-2 agents, it entails an effort cost lower than the one they would incur if they mimicked by pooling, i.e., earning y^1 while being paid according to the average productivity $\overline{\theta}$) and above the isocost line (38) (meaning that type-1 agents would be discouraged from replicating the effort choices of type-2 agents). Instead, if $\theta^2 h\left(e_s^{int}, e_q^{int}\right) > y^1$, type-2 mimickers are better off earning y^1 and achieving segregation than pooling with type-1 agents at the effort mix $\left(e_s^1, \widehat{e}_q^2\right)$. But since the IC constraint (23) is binding in the MMO, it follows that a two-bracket subsidy schedule with $\sigma = \widehat{\sigma}$ for $e_s \leq e_s^1$ and $\sigma = \sigma^*$ for $e_s > e_s^1$ does not ensure implementation. In this case, the schedule $S\left(e_s; y^1\right)$ must be adjusted by introducing an additional segment, for $e_s \leq e_s^{int}$, with an associated subsidy rate of 100%.[20] This full subsidy implies that, on the isoquant $\theta^2 h\left(e_s, e_q\right) = y^1$, any point that allows type-2 agents to achieve separation entails for them a higher cost than the one they would incur by pooling with type-1 agents at the effort mix $\left(e_s^1, \widehat{e}_q^2\right)$.

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

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

Finally, note that, somewhat surprisingly, we obtain the canonical efficiency-at-the-top result for high-ability agents (Sadka, 1976).[22] One might have expected, for example, that the tax function should have been used to distort the effort allocation of type-2 agents toward e_q—the dimension in which they have a comparative advantage—in order to discourage type-1 agents from mimicking them. However, a key insight is that the IC constraint for type-1 agents is already embedded in the laissez-faire equilibrium. As a result, type-2 agents have already internalized this constraint when making their decisions. The labor contract offered to type-2 agents is designed to maximize their utility, subject to the IC constraint of type-1 agents. This is consistent with the government's goal of extracting as much revenue as possible from type-2 agents to facilitate redistribution. It is also worth noting that although the

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

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

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

marginal tax rates for high-skilled agents are zero under the implementing tax function, the income level y^2 in the STE is lower than under laissez-faire when the upward IC constraint for low-skilled workers binds in the laissez-faire scenario. This result arises because redistribution through the tax system increases the utility of type-1 agents relative to their utility under laissez-faire, thereby reducing their incentive to imitate type-2 agents.

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

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

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

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:

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

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

$$T(y) = \begin{cases} \left(\frac{1}{\theta^1} - \frac{1}{\bar{\bar{\theta}}}\right) \frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)} \widehat{y} + \left(\frac{1}{\bar{\bar{\theta}}} - \frac{1}{\theta^1}\right) \frac{p_q^1}{h_2\left(\widehat{e}_s, \widehat{e}_q\right)} y, & \text{for all } y \in [0, \widehat{y}] \\ (y - \widehat{y}) \max\left\{1 - \frac{p_q^2}{\bar{\theta} h_2\left(\widehat{e}_s, \widehat{e}_q\right)}, \frac{\left[\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}\right] - \left[\widehat{y} - p_s \widehat{e}_s - p_q^2 \widehat{e}_q\right]}{\underline{y}^{sep} - \widehat{y}}\right\}, & \text{for all } y > \widehat{y}. \end{cases} \tag{45}$$

PROOF. See Online Appendix G. □

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

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

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

In particular, note that the marginal tax rate given by the second term in the max operator is defined as an expression that depends on both $\underline{y}^{sep}$ and $\theta^2\left(e_s^{2*}, e_q^{2*}\right) - p_s e_s^{2*} - p_q^2 e_q^{2*}$. The former represents the minimum amount of income that type-2 agents would need to earn to achieve separation from their type-1 counterparts in an environment where $T(y) = 0$ for

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

$y > \widehat{y}$. In turn, $\underline{y}^{sep}$ is a function of $e_q^{\min}$, which represents the minimum level of e_q that allows type-2 agents to achieve separation when choosing $e_s = \widehat{e}_s$.

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

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

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

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

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

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

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

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

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

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

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

The threshold and repayment rate vary depending on the borrower's income level. In 2023–24, the income threshold is AUD 51,550 and above this threshold the repayment rate varies from 1% to a maximum of 10% for incomes above AUD 151,201. The income-contingent repayment system is essentially a means-tested progressive tax on graduates, as high-achieving students reach the income threshold earlier and earn higher wages. Similar income-contingent repayment systems exist in the United Kingdom and Sweden, as well as in many other countries.[26]

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

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

5. DISCUSSION

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

5.1. *The Case Where Neither Signal is Observable.* In Online Appendix H, we study the case where the government can only observe income. In this case, due to the weaker policy instruments available, the possibilities for mimickers to deviate are expanded. The main insight from our analysis is that predistribution is not feasible with only an income tax. Thus, the ability to tax the signals transmitted in the labor market is essential to achieve predistribution. In Online Appendix K, we use the case with only an income tax as a benchmark to numerically quantify the welfare gains of taxing the quantity signal. Note that the welfare gains from taxing the education signal arise regardless of whether the MMO features predistribution or not. However, consistent with Proposition 1, the results show that the MMO tends to feature predistribution when the productivity variance between the two categories of workers and the discrepancy in the cost of obtaining the quality signal across types are moderate.

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

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

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

With respect to wedges, the most interesting difference between the STE when the observable signal is e_s and the STE when the observable signal is e_q is that in the latter case it is a priori ambiguous in which direction it is optimal to distort the effort mix of type-1 agents. This contrasts with the result provided by (30) for the case where the observable signal is e_s, namely, that the effort mix chosen by type-1 agents should be distorted toward 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 toward e_q.[29]

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

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

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

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

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 whereas others are separated. Nevertheless, the main qualitative insight that constrained efficient allocations may involve predistribution is not sensitive to the number of types. Several features stand out, however.

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

6. CONCLUDING REMARKS

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

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

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

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

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

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

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

DATA AVAILABILITY STATEMENT Data sharing not applicable to this article as no data sets were generated or analyzed during the current study.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure 2: The welfare gains from predistribution

REFERENCES

Akerlof, G., "The Economics of Caste and of the Rat Race and Other Woeful Tales," *Quarterly Journal of Economics* 90 (1976), 599–617.

Andersson, F., "Income Taxation and Job-Market Signaling," *Journal of Public Economics* 59 (1996), 277–98.

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

Bastani, S., T. Blumkin, and L. Micheletto, "Optimal Wage Redistribution in the Presence of Adverse Selection in the Labor Market," *Journal of Public Economics* 131 (2015), 41–57.

———, ———, and ———, "The Welfare-Enhancing Role of Parental Leave Mandates," *Journal of Law, Economics, and Organization* 35 (11 2019), 77–126.

Blomquist, S., and V. Christiansen, "Taxation and Heterogenous Preferences," *FinanzArchiv* 64 (2008), 218–44.

Blumkin, T., and E. Sadka, "A Case for Taxing Education," *International Tax and Public Finance* 15 (2008), 145–63.

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

Bovenberg, A. L., and B. Jacobs, “Redistribution and Education Subsidies are Siamese Twins,” *Journal of Public Economics* 89 (2005), 2005–35.
Brett, C., and J. A. Weymark, “Financing Education Using Optimal Redistributive Taxation,” *Journal of Public Economics* 87 (2003), 2549–69.
Caplan, B., *The Case Against Education: Why the Education System is a Waste of Time and Money* (NJ: Princeton University Press, 2018).
Cho, I.-K., and D. M. Kreps, “Signaling Games and Stable Equilibria,” *Quarterly Journal of Economics* 102 (1987), 179–221.
Craig, A. C., “Optimal Income Taxation with Spillovers from Employer Learning,” *American Economic Journal: Economic Policy* 15 (May 2023), 82–125.
Ebert, U., “A Reexamination of the Optimal Nonlinear Income Tax,” *Journal of Public Economics* 49 (1992), 47–73.
Findeisen, S., and D. Sachs, “Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans,” *Journal of Public Economics* 138 (2016), 1–21.
Fudenberg, D., and J. Tirole, “Perfect Bayesian Equilibrium and Sequential Equilibrium,” *Journal of Economic Theory* 53 (1991), 236–60.
Grossman, S. J., and M. Perry, “Perfect Sequential Equilibrium,” *Journal of Economic Theory* 39 (1986), 97–119.
Maldonado, D., “Education Policies and Optimal Taxation,” *International Tax and Public Finance* 15 (2008), 131–43.
Manoli, D., “A Signaling Approach to Optimal Income Taxation,” Unpublished Working Paper, 2006.
Mirrlees, J. A., “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies* 38 (1971), 175–208.
Riley, J. G., “Silver Signals: Twenty-Five Years of Screening and Signaling,” *Journal of Economic Literature* 39 (June 2001), 432–78.
Rothschild, C., and F. Scheuer, “Redistributive Taxation in the Roy Model,” *Quarterly Journal of Economics* 128 (2013), 623–68.
Sachs, D., A. Tsyvinski, and N. Werquin, “Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium,” *Econometrica* 88 (2020), 469–93.
Sadka, E., “On Income Distribution, Incentive Effects and Optimal Income Taxation,” *Review of Economic Studies* 43(06) (1976), 261–67.
Spence, M., “Job Market Signaling,” *Quarterly Journal of Economics* 87 (1973), 355–74.
———, “Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution,” *Journal of Economic Theory* 7 (1974), 296–332.
Stantcheva, S., “Optimal Income Taxation with Adverse Selection in the Labor Market,” *Review of Economic Studies* 81 (2014), 1296–329.
———, “Optimal Taxation and Human Capital Policies over the Life Cycle,” *Journal of Political Economy* 125 (2017), 1931–90.
Stiglitz, J. E., “Self-Selection and Pareto-Efficient Taxation,” *Journal of Public Economics* 17 (1982), 213–40.
Sztutman, A., “Dynamic Job Market Signaling and Optimal Taxation,” Working Paper, Tepper School of Business, Carnegie Mellon University, 2024.
Tuomala, M., “On the Optimal Income Taxation and Educational Decisions,” *Journal of Public Economics* 30 (1986), 183–98.
Ulph, D., “On the Optimal Distribution of Income and Educational Expenditure,” *Journal of Public Economics* 8 (1977), 341–56.