

Bastani, Spencer; Blumkin, Tomer; Micheletto, Luca

Working Paper

Optimal Wage Redistribution in the Presence of Adverse Selection in the Labor Market

CESifo Working Paper, No. 4938

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Bastani, Spencer; Blumkin, Tomer; Micheletto, Luca (2014) : Optimal Wage Redistribution in the Presence of Adverse Selection in the Labor Market, CESifo Working Paper, No. 4938, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/102095>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Papers

www.cesifo.org/wp

Optimal Wage Redistribution in the Presence of Adverse Selection in the Labor Market

Spencer Bastani
Tomer Blumkin
Luca Micheletto

CESIFO WORKING PAPER NO. 4938

CATEGORY 1: PUBLIC FINANCE

AUGUST 2014

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

CESifo

Center for Economic Studies & Ifo Institute

Optimal Wage Redistribution in the Presence of Adverse Selection in the Labor Market

Abstract

In this paper we allude to a novel role played by the non-linear income tax system in the presence of adverse selection in the labor market due to asymmetric information between workers and firms. We show that an appropriate choice of the tax schedule enables the government to affect the wage distribution by controlling the transmission of information in the labor market. This represents an additional channel through which the government can foster the pursuit of its redistributive goals.

JEL-Code: D820, H210, J310.

Keywords: adverse selection, labor market, optimal taxation, pooling, redistribution.

Spencer Bastani
Department of Economics
Uppsala University
Uppsala / Sweden
spencer.bastani@nek.uu.se

Tomer Blumkin
Department of Economics
Ben Gurion University of the Negev
Israel - 84105 Beer-Sheva
tomerblu@bgu.ac.il

Luca Micheletto
Department of Law
University of Milan
Milan / Italy
luca.micheletto@unibocconi.it

July 23, 2014

Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

1 Introduction

The modern approach to taxation emphasizes information as the fundamental constraint on public policy. The key assumption in the standard Mirrlees (1971) framework is that the government is unable to observe individual productivities (earning capacities) and hence has to redistribute based on observed levels of income. This might invite high-skilled workers to engage in “mimicking”, that is, to reproduce the earned income of a low-skilled worker, in order to benefit from a more lenient tax treatment and thereby derive a higher utility. This means that the income tax must be designed in a way which renders such mimicking unattractive; namely, the income tax must be incentive-compatible.

A standard assumption in the optimal tax literature is that there is symmetric information between workers and firms. In a recent paper, Stantcheva (forthcoming) relaxes this assumption by assuming that firms cannot observe the productivities of workers. Assuming in addition that higher-skilled workers have a weaker taste for leisure, firms have the possibility to screen between high- and low-skilled workers by offering an increased compensation conditional on a higher labor effort. This gives rise to adverse selection where high-skilled agents work more than the efficient amount. Stantcheva shows that when the government is sufficiently egalitarian, social welfare would be higher in the presence of adverse selection than under the Mirrleesian benchmark with symmetric information. The reason for this is that under adverse selection, as labor contracts cannot be conditioned on (unobserved) labor productivity, high-skilled mimickers are not fully remunerated for their higher earning capacity. That is, they have to work longer hours than under a symmetric information regime in order to reproduce the income of the low-skilled workers. This makes less tempting for the high-skilled workers to mimic their low-skilled counterparts and thereby enhances redistribution.

In principle, the government can promote redistributive goals through two different channels: (i) by changing the income distribution, and/or (ii) by affecting the underlying wage distribution. In the standard Mirrlees (1971) setting, the production technology is assumed to be linear, which implies that the wage distribution is exogenous, thereby leaving no scope for the government to further equity goals through the wage channel. By relaxing the assumption of linearity, the subsequent literature has introduced a role for the income tax to affect the wage distribution. Stiglitz (1982)

demonstrates that, when skill types are complements in the production technology, it is socially optimal to marginally subsidize the labor supply of high-skilled workers in order to reduce wage dispersion. This in turn renders the optimal taxes less progressive than under the standard Mirrlees setup with a linear production technology. More recently, Rothschild and Scheuer (2013) have extended the discrete Stiglitz (1982) framework to a continuum of types that differ along a multidimensional skill vector and have allowed for endogenous occupational choices. They show that the redistributive wage channel emphasized by Stiglitz carries over to the more general setting. However, the additional features associated with the occupational choice margin mitigate the general equilibrium effects and make the optimal taxes more progressive (but still less progressive than under the standard Mirrleesian setting).¹

In this paper we connect the analysis of Stantcheva (forthcoming) with the above-mentioned strand of the literature, which emphasizes the wage channel for redistribution. Stantcheva considers a standard linear production technology and restricts attention to separating allocations, in which each type of worker is offered a distinct consumption-labor bundle. In a separating allocation each worker is remunerated according to his/her marginal productivity and wage rates are fixed (as in the symmetric information setting). This leaves no scope for the government to redistribute through the wage channel.

Employing a similar framework, we show that the government can in fact affect the wage distribution. By choosing an appropriate tax system the government can block the possibility for firms to engage in screening and implement a pooling allocation with full wage equalization. When designing the optimal redistributive policy the government has to balance the efficiency gains from screening, associated with implementing a separating allocation, and the equity gains from wage pooling. The pooling

¹The above mentioned literature has limited attention to the role of income taxation in affecting the wage distribution by relying on the complementarity between production factors. Cremer et al. (2011) consider a setting with a linear production technology (that is, no complementarities) but where the government can supplement the nonlinear income tax with education policy that affects the wage distribution. They find, surprisingly, that the most unequal distribution of wages is desirable from the standpoint of social welfare maximisation when the permissible degree of wage differentiation is large. When the permissible degree of wage differentiation is small, they demonstrate that an equal-wage outcome (which obviates the redistributive role of income taxation) may be socially desirable.

allocation turns out to be socially superior when the preferences for redistribution are sufficiently strong and the differences in productivities are not too large.

The general message of our analysis is that one can highlight a novel role played by the non-linear income tax system in the presence of adverse selection in the labor market due to asymmetric information between workers and firms. Under symmetric information, firms observe workers' productivities and therefore remunerate each worker according to his/her marginal productivity in a competitive labor market. Under asymmetric information, however, the translation of differences in productivities into differences in wage rates crucially hinges on the mechanism by which workers and firms exchange information. In this case, a nonlinear income tax can also be used by the government as a device to control the transmission of information in the labor market. Under certain circumstances, this might prove to be for the government an important channel to foster the pursuit of its redistributive goals.

2 The Model

We use the simplest possible model with just the key ingredients necessary to demonstrate our point. Consider an economy with low- and high-skilled workers (indexed by l and h , respectively) that produce a single consumption good (the price of which is normalized to unity) using a production technology exhibiting constant returns to scale and perfect substitutability between the two skill levels. We normalize the workers' population to a unit measure and let the measures of low- and high-skilled workers be given, respectively, by m^l and m^h .

Let the earning capacity (which is equal to the hourly wage rate under a perfectly competitive labor market) of a low- and a high-skilled worker be denoted by w^l and w^h respectively, where $w^h > w^l \geq 0$.

The two types of workers differ in their labor-leisure preferences. The utility of the high-skilled workers is given by $u^h \equiv c^h - g(n^h)$ where c represents consumption, n represents working hours, and where, $g(0) = 0$, $g' > 0$, $g'' > 0$ and $\lim_{n \rightarrow 0} g'(n) = 0$. The utility of the low-skilled workers is given by $u^l \equiv c^l - kg(n^l)$, where $k > 1$. That is, low-skilled workers incur a higher disutility (both total and marginal) from work

relative to their high-skilled counterparts for the same working hours supplied.²

2.1 Labor Market Equilibrium under Asymmetric Information

We deviate from the standard Mirrlees (1971) framework and assume that firms cannot observe the types of their workers when signing a labor contract. An alternative interpretation of the setting would be that firms do observe the types but are not allowed to offer separate contracts due to anti-discrimination legislation. As is well known since the seminal contribution of Rothschild and Stiglitz (1976), adverse selection may arise in such contexts. Before turning to present the optimal tax problem we briefly characterize the *laissez-faire* equilibrium (adopting the Rothschild and Stiglitz (RS) equilibrium concept) and demonstrate the resulting market failure.

2.2 The RS Equilibrium

A typical labor contract specifies the number of working hours, n , and the corresponding total compensation, c . Crucially, a labor contract cannot be made conditional on the type of worker, which is assumed to be private information of the worker and hence unobservable by the hiring firm. The RS equilibrium is defined by a set of labor contracts satisfying two properties: (i) firms make non-negative profits on each contract; and, (ii) there is no other potential contract that would yield non-negative profits if offered (in addition to the equilibrium set of contracts).

Having defined the equilibrium, we turn next to show that the *laissez-faire* allocation under symmetric information may become non incentive-compatible in the presence of asymmetric information.

In a competitive labor market with symmetric information each worker would be remunerated according to his/her earning capacity. Formally, a competitive equilibrium allocation is given by the two consumption-labor bundles (c^{i*}, n^{i*}) ; $i = l, h$, which

²The quasi-linear specification, which is common in the literature [see Diamond (1998) and Salanié (2011) amongst others], is invoked for tractability. Our qualitative results remain robust to incorporation of income effects on labor supply.

satisfy:

$$c^{i*} = w^i n^{i*}; i = l, h, \quad (1)$$

$$w^i = k^i g'(n^{i*}); \quad i = l, h; \quad k^l = k \text{ and } k^h = 1. \quad (2)$$

The first condition is the individual budget constraint driven by the zero-profit free entry requirement, whereas the second condition states that workers optimally choose their labor supply by equating the marginal disutility from work with their hourly wage rate.

In addition, to ensure a separating equilibrium, the allocation has to satisfy two incentive compatibility constraints. These constraints ensure that workers have no incentives to mimic each other. Formally,

$$c^{l*} - kg(n^{l*}) \geq c^{h*} - kg(n^{h*}), \quad (3)$$

$$c^{h*} - g(n^{h*}) \geq c^{l*} - g(n^{l*}). \quad (4)$$

Notice that a high-skilled worker has no incentive to mimic his/her low-skilled counterpart due to the higher hourly wage rate reflected in his/her symmetric information *laissez-faire* contract. Thus, the incentive compatibility constraint (4) for the high-skilled worker is slack.³ However, the incentive compatibility constraint associated with the low-skilled worker may be violated. To see this, reformulate the incentive constraint associated with the low-skilled worker by substituting for w^i and c^{i*} from (1) and (2) into (3) to obtain:

$$kg'(n^{l*}) n^{l*} - kg(n^{l*}) \geq g'(n^{h*}) n^{h*} - kg(n^{h*}). \quad (5)$$

Consider now the limiting case where k converges to 1. Re-formulating (5) by taking the limit yields:

$$g'(n^{l*}) n^{l*} - g(n^{l*}) \geq g'(n^{h*}) n^{h*} - g(n^{h*}) \iff H(n^{l*}) \geq H(n^{h*}), \quad (6)$$

where $H(n) \equiv g'(n)n - g(n)$.

³Formally, $c^{h*} - g(n^{h*}) > w^h n^{l*} - g(n^{l*}) = \frac{w^h}{w^l} c^{l*} - g(n^{l*}) > c^{l*} - g(n^{l*})$, where the first inequality follows by virtue of the strict convexity of g , which implies that n^{h*} is the (unique) optimal labor supply choice of type- h workers under the symmetric information regime, the equality follows from the budget constraint in (1) and the latter inequality follows as $w^h > w^l$.

Differentiation with respect to n yields, $H'(n) = g''(n)n > 0$ where the inequality follows by the strict convexity of g . Thus, by virtue of (5), for the incentive constraint associated with the low-skilled worker to hold it is necessary that $n^{l*} \geq n^{h*}$. However, by virtue of (2), the strict convexity of g and the fact that $w^h > w^l$ it follows that $n^{h*} > n^{l*}$. Thus, by continuity considerations, for k sufficiently close to unity, the incentive constraint associated with the low-skilled workers is violated and, hence, the symmetric information *laissez-faire* allocation is not incentive compatible.⁴

When condition (5) is violated, the *laissez-faire* RS equilibrium allocation is given by the two consumption-labor bundles $(c^{i**}, n^{i**}); i = l, h$, which satisfy:

$$c^{i**} = w^i n^{i**}; i = l, h, \quad (7)$$

$$w^l = kg'(n^{l**}), \quad (8)$$

$$c^{l**} - kg(n^{l**}) = c^{h**} - kg(n^{h**}). \quad (9)$$

Comparing the equilibrium allocations under symmetric and asymmetric information [given, respectively, by conditions (1)-(2) and (7)-(9)] reveals that the labor supply condition for type- h workers under the symmetric information regime is being replaced by the binding incentive constraint of type- l workers under the asymmetric information regime, which implicitly defines the labor contract offered to type- h workers in the asymmetric information equilibrium. Under asymmetric information low-skilled workers are still offered their efficient (symmetric information) allocation, $(n^{l**} = n^{l*})$, whereas high-skilled workers' labor supply choice is distorted, as they work more hours than under their efficient allocation $(n^{h**} > n^{h*})$. This enables the firms to reduce the information-rent associated with type- l workers and render the allocation incentive compatible.

Two final remarks are in order. First, notice that in the RS setting the separating equilibrium characterized by conditions (7)-(9) exists when the fraction of low-skilled workers is sufficiently high [see RS (1976)]. A pooling equilibrium does not exist, as firms can engage in "cream-skimming", by offering a contract that would attract only type- h workers and yield positive profits. Notice further that our assumption that workers differ not only in their earning capacity [as in Mirrlees (1971)] but also in

⁴Notice that when k is sufficiently large; namely, when the disutility from work entailed by the low-skilled workers is sufficiently high, the symmetric information *laissez-faire* allocation would be incentive compatible (and hence first-best efficient).

their labor-leisure preferences is essential for the existence of a separating equilibrium which relies on the ability of firms to screen between workers based on their differences in preferences (higher-skilled workers exhibit weaker taste for leisure). In the absence of such screening capacity the only equilibrium that would sustain under asymmetric information would be one where all workers would be pooled together and each paid an hourly wage rate equal to the average productivity.

3 The Government's Problem

The government is seeking to design a non-linear tax-and-transfer system, which maximizes a welfare function given by a weighted average of the utilities of the two types of workers. Formally,

$$W = \sum_i \beta^i u^i; i = l, h, \quad \text{where} \quad \sum_i \beta^i = 1 \quad \text{and} \quad m^l < \beta^l \leq 1. \quad (10)$$

The fact that the weight assigned to type- l workers strictly exceeds their share in the population reflects the strictly egalitarian preferences of the government with respect to redistribution.

We follow Stantcheva (forthcoming) by considering the regime referred to as “adverse selection with unobservable private contracts” in which neither the firm nor the government observes workers’ types and, in addition, the government has no control over labor contracts.

In the previous subsection we have argued that a pooling equilibrium cannot exist in the RS setting under *laissez faire*, as it will invite “cream-skimming”. However, in the presence of government intervention the government may block the possibility for such “cream-skimming” by an appropriate choice of the tax system. In the absence of screening, all workers will be pooled together and receive the same wage rate equal to the average productivity and the same income level. When designing the optimal redistributive policy, therefore, the government has to account for the trade-off between the efficiency gains from screening associated with implementing a separating equilibrium, which induces high-skilled workers to work more hours than their low-skilled counterparts, and the equity gains from wage pooling associated with implementing a pooling equilibrium.

3.1 The Separating Equilibrium Regime

Invoking the self-selection approach common in the optimal tax literature, a non-linear income tax schedule is given by the tuple $\{y^i, T^i\}; i = l, h$, where y denotes gross income and T denotes the associated tax (possibly negative) which satisfies the balanced budget constraint: $\sum_i m^i T^i = 0$, where we assume with no loss in generality that the government has no exogenous revenue needs. To abbreviate notation, letting $T^l \equiv -T$, it follows, using the balanced budget constraint, that $T^h = \frac{m^l}{m^h} \cdot T$.

We turn next to characterize the separating *RS* equilibrium given the tax schedule in place. Notice that the only margin of maneuver of firms is to set the working hours demanded, denoted by n^i (with $i = l, h$), for each level of gross income. As firms do not observe workers' types, the number of working hours demanded for each level of gross income will be independent of the worker's type. A separating equilibrium has to satisfy the following set of conditions.

First of all, the resulting allocation has to be incentive-compatible; namely, it has to satisfy the following two incentive constraints associated with type-*l* and type-*h*, respectively:

$$y^l + T - kg(n^l) \geq y^h - \frac{m^l}{m^h} \cdot T - kg(n^h), \quad (\text{IC}^l)$$

$$y^h - \frac{m^l}{m^h} \cdot T - g(n^h) \geq y^l + T - g\left(\frac{y^l}{\sum_i m^i w^i}\right). \quad (\text{IC}^h)$$

The incentive constraint associated with type-*l* (IC^l) is rather standard, the only difference from the Mirrlees (1971) setting being that the mimicking type-*l* agent is working the same number of hours as his/her type-*h* counterpart.

The incentive constraint associated with type-*h* is instead non-standard [the argument is similar to Stantcheva (forthcoming)]. Notice that in the second term on the right-hand side of the inequality in (IC^h), we replaced n^l with the term $\frac{y^l}{\sum_i m^i w^i}$. Due to the requirement that firms earn non-negative profits in a separating equilibrium, n^l is necessarily bounded from below by the term $\frac{y^l}{w^l}$. Offering any lower level of n^l [assuming (IC^h) is not violated] would yield negative profits. However, by offering a sufficiently low level of n^l , the firm can attract also type-*h* workers, who are more productive than their type-*l* counterparts. Thus, although the firm suffers losses on type-*l* workers, it is compensated, by gaining on their type-*h* counterparts. The term $\frac{y^l}{\sum_i m^i w^i}$

defines the level of n that would yield the firm zero profits in a pooling equilibrium associated with the income level y^l . That is, the term defines a lower bound on n that can be offered by the firm under such pooling equilibrium. Offering such a pooling equilibrium contract would be more attractive for both types of workers than the separating contract associated with y^l , as $\frac{y^l}{\sum_i m^i w^i} < \frac{y^l}{w^l} \leq n^l$. Incentive compatibility then requires that a type- h worker weakly prefers his/her contract to mimicking type- l and getting the pooling contract that yields zero profits to the firm. Notice that any alternative pooling contract that would yield positive profits would require longer working hours and would hence be clearly dominated by the type- h separating equilibrium contract.⁵

In addition to being incentive compatible, the resulting allocation has to satisfy two zero-profit conditions associated with the contracts offered to type- l and type- h workers, respectively:

$$y^l = w^l n^l, \quad (\text{ZP}^l)$$

$$y^h = w^h n^h. \quad (\text{ZP}^h)$$

Condition (ZP^l) requires that a contract offered to a type- l worker would yield zero profits. Recall that according to the definition of the RS equilibrium any contract has to yield non-negative profits (but not necessarily zero profits). However, if the condition is violated and holds as a strict inequality, a firm can offer a contract that slightly reduces n^l . Clearly, this new contract would attract type- l workers and would yield positive profits, by continuity considerations; therefore, it would not be an equilibrium. In case type- h workers find this contract attractive as well, the firm's profits will further increase (as type- h workers are more productive than their type- l counterparts).

We turn next to condition (ZP^h) . Notice that this condition may hold as a strict inequality. In such a case an additional (complementary-slackness) condition has to be

⁵We would like to make the following technical remark. The binding incentive constraint associated with type- h workers implies that a type- h worker is indifferent between his/her separating contract and the pooling contract associated with y^l that yields zero profits. In principle, one might consider the possibility of firms offering this pooling contract as being a violation of condition (ii) in the definition of the RS equilibrium on page 5. However, notice that the pooling contract is strictly preferred by *all* low-skilled workers to their separating contract. Thus, in order for the pooling contract to yield zero profits, *all* high skilled workers would have to choose the pooling contract even though they are indifferent between the pooling and their separating contracts. We rule out this implausible possibility.

satisfied, namely, requiring that the incentive constraint associated with type- l workers is binding. To see this, notice that when condition (IC^l) is slack and the zero profit condition for type- h workers does not hold, a firm can offer a new contract that slightly decreases n^h , attracting only type- h workers and yielding positive profits, by continuity considerations. When (IC^l) is binding, however, such a decrease in n^h will also attract type- l workers. The resulting pooling allocation has to be unprofitable to sustain the equilibrium. Thus, we need to add the condition $n^h \sum_i m^i w^i \leq y^h$ which implies that for any $n < n^h$, the pooling contract would yield negative profits. However, any profits earned by a firm hiring type- h workers can be taxed away at a confiscatory 100 percent tax rate and paid back to the workers in an incentive compatible manner that renders both types of workers strictly better off. The modified profit cum income tax system is equivalent to an income tax system where $y^h = w^h n^h$, namely, the zero profit condition is satisfied, and the income tax paid by type- h workers is augmented by the amount paid by the firm as profit taxes. Consequently, we will henceforth assume that condition (ZP^h) holds with no loss in generality.

Reformulating the utilities of the two types of workers given the tax schedule in place, employing the two zero-profit conditions, and substituting into the welfare function in (10) yields:

$$W = \beta^l \left[y^l + T - k g \left(\frac{y^l}{w^l} \right) \right] + \beta^h \left[y^h - \frac{m^l}{m^h} \cdot T - g \left(\frac{y^h}{w^h} \right) \right]. \quad (11)$$

Substituting from the zero profit conditions in (ZP^l) and (ZP^h) into the incentive compatibility conditions (IC^l) and (IC^h) yields:

$$y^l + T - k g \left(\frac{y^l}{w^l} \right) \geq y^h - \frac{m^l}{m^h} \cdot T - k g \left(\frac{y^h}{w^h} \right), \quad (12)$$

$$y^h - \frac{m^l}{m^h} \cdot T - g \left(\frac{y^h}{w^h} \right) \geq y^l + T - g \left(\frac{y^l}{\sum_i m^i w^i} \right). \quad (13)$$

Under the separating equilibrium regime the government is choosing the tax parameters y^l , y^h and T so as to maximize the welfare in (11) subject to the two incentive compatibility constraints in (12) and (13).

3.2 The Pooling Equilibrium Regime

Under a pooling equilibrium regime, by choosing an appropriate tax schedule, the government can determine the common gross level of income \bar{y} .⁶ Being unable to distinguish between the two types of workers, firms will pay all workers the same wage rate equaling the average productivity:

$$\bar{w} \equiv \sum_i m^i w^i.$$

Moreover, all agents will work the same number of hours, \bar{n} , given, due to the zero profit condition, by:

$$\bar{n} = \bar{y} / \bar{w}. \quad (14)$$

By virtue of our assumption that the government has no exogenous revenue needs, with no loss in generality, there will be no tax levied at the income level chosen by both types of workers.

Substituting from (14) into the welfare function in (10) yields:

$$W = \beta^l [\bar{y} - k g(\bar{y} / \bar{w})] + \beta^h [\bar{y} - g(\bar{y} / \bar{w})]. \quad (15)$$

Under the pooling equilibrium regime the government is choosing the common gross level of income, \bar{y} , so as to maximize the welfare in (15).

We turn next to compare the two regimes.

3.3 Comparison between the Separating and the Pooling Equilibria

The following proposition characterizes a sufficient condition for the pooling allocation to be the socially desirable equilibrium configuration.

Proposition 1. *When the weight assigned to type-l workers in the welfare function is sufficiently large and the difference in productivities is sufficiently small, the second best optimum is given by a pooling allocation.*

⁶One simple tax schedule that implements the pooling allocation is the one in which the government levies a confiscatory 100 percent tax rate at any level of income other than \bar{y} . Notice that such a schedule prevents firms from engaging in “cream-skimming”. The reason is that they become unable to attract only the high-skilled workers by offering them a higher compensation in exchange for longer working hours.

Proof Consider the max-min (Rawlsian) case where the government assigns a zero weight to the utility of type- h workers, that is it maximizes the utility of type- l workers. The result will extend by continuity to the case where the weight assigned to type- l workers is large enough.

Let us formulate the optimal solution for the government problem under the two alternative regimes: (i) a separating equilibrium; and, (ii) a pooling equilibrium.

The *Lagrangian* associated with the constrained optimization problem under a separating equilibrium is given by:

$$W^s(w^l, w^h) \equiv \max_{y^l, y^h, T, \lambda} \left[\left(y^l + T - kg \left(\frac{y^l}{w^l} \right) \right) + \lambda \left(y^h - \frac{m^l}{m^h} \cdot T - g \left(\frac{y^h}{w^h} \right) - y^l - T + g \left(\frac{y^l}{\sum_i m^i w^i} \right) \right) \right]. \quad (16)$$

Notice that the relevant binding constraint is (IC^h) . We will assume that (IC^l) is slack. Our result will hold a-fortiori when (IC^l) is binding.⁷

The constrained optimization problem under a pooling equilibrium is given by:

$$W^p(w^l, w^h) \equiv \max_y \left(y - kg \left(\frac{y}{\sum_i m^i w^i} \right) \right). \quad (17)$$

We prove the proposition by showing that $W^p(w^l, w^h) > W^s(w^l, w^h)$ when $w^l \rightarrow w^h$.

Notice that when $w^l = w^h$ both regimes coincide, hence $W^p = W^s$. By invoking a first order approximation it suffices then to prove the following:

$$\lim_{w^l \rightarrow w^h} \frac{\partial W^p(w^l, w^h)}{\partial w^l} < \lim_{w^l \rightarrow w^h} \frac{\partial W^s(w^l, w^h)}{\partial w^l}.$$

Differentiation of the expressions in (16) and (17) with respect to w^l employing the envelope condition, yields:

$$\frac{\partial W^s(w^l, w^h)}{\partial w^l} = kg' \left(\frac{y^l}{w^l} \right) \frac{y^l}{w^{l2}} - \lambda g' \left(\frac{y^l}{\sum_i m^i w^i} \right) \frac{m^l y^l}{(\sum_i m^i w^i)^2}, \quad (18)$$

$$\frac{\partial W^p(w^l, w^h)}{\partial w^l} = kg' \left(\frac{y}{\sum_i m^i w^i} \right) \frac{m^l y}{(\sum_i m^i w^i)^2}. \quad (19)$$

⁷Notice that the standard single crossing property doesn't apply which means that if (IC^h) is binding in a separating allocation, this does not imply that (IC^l) is slack. If the pooling equilibrium dominates the separating equilibrium when (IC^l) does not bind, then this is clearly the case also when allowing for the possibility that (IC^l) binds (since that would imply that the welfare associated with the separating equilibrium would be even lower).

By differentiating the *Lagrangian* in (16) with respect to T and equating to zero it follows that $\lambda = m^h$. Substituting for λ into (18) and taking the limit of the expressions in (18) and (19) when $w^l \rightarrow w^h$, employing the fact that when $w^l \rightarrow w^h$, $y^l \rightarrow y^h$ and $y \rightarrow y^h$, yields upon rearrangement:

$$\lim_{w^l \rightarrow w^h} \frac{\partial W^s(w^l, w^h)}{\partial w^l} = (k - m^h m^l) g' \left(\frac{y^h}{w^h} \right) \frac{y^h}{w^{h2}} \quad (20)$$

and

$$\lim_{w^l \rightarrow w^h} \frac{\partial W^p(w^l, w^h)}{\partial w^l} = k m^l g' \left(\frac{y^h}{w^h} \right) \frac{y^h}{w^{h2}}. \quad (21)$$

Thus,

$$\begin{aligned} \lim_{w^l \rightarrow w^h} \frac{\partial W^p(w^l, w^h)}{\partial w^l} &< \lim_{w^l \rightarrow w^h} \frac{\partial W^s(w^l, w^h)}{\partial w^l} \\ &\iff \\ k m^l &< k - m^h m^l \\ &\iff \\ m^l &< k \end{aligned}$$

where the latter equivalence follows from the fact that $m^h = 1 - m^l$ and the last inequality is a consequence of $k > 1 > m^l$.

This completes the proof. \square

The rationale underlying the proposition is as follows. When the difference in productivities is small enough and the government is sufficiently egalitarian, the equity gains associated with wage pooling outweigh the efficiency gains from screening and the optimal redistributive policy is to implement a pooling allocation by choosing a tax schedule that prevents firms from engaging in screening.

It is important to notice the difference between our prediction that implementing a pooling allocation would be desirable and the classic result in Stiglitz (1982) who demonstrates that in a two-type setting pooling will never be the welfare maximizing allocation.⁸ In the Stiglitz (1982) setting, firms are able to observe the earning abilities of their workers. Hence, when the government implements a pooling allocation, it

⁸Notice that in a setting with more than two types bunching may be desirable. See Ebert (1992).

implies that income levels are pooled but not wage rates. In contrast, in our case, as firms cannot distinguish between their workers, a pooling allocation implies that the wage rates are pooled.

To gain further insights we consider a simple numerical example. We make the following parametric assumptions: $g(n) = \frac{n^2}{2}$, $k = 1.05$, $w^h = 100$, $m^h = 0.6$. Our qualitative results remain robust to the parametric specification chosen. The figure below compares the max-min (Rawlsian) welfare levels associated with the pooling and the separating allocations for different values of w^l , the wage rate of the low-skilled workers.

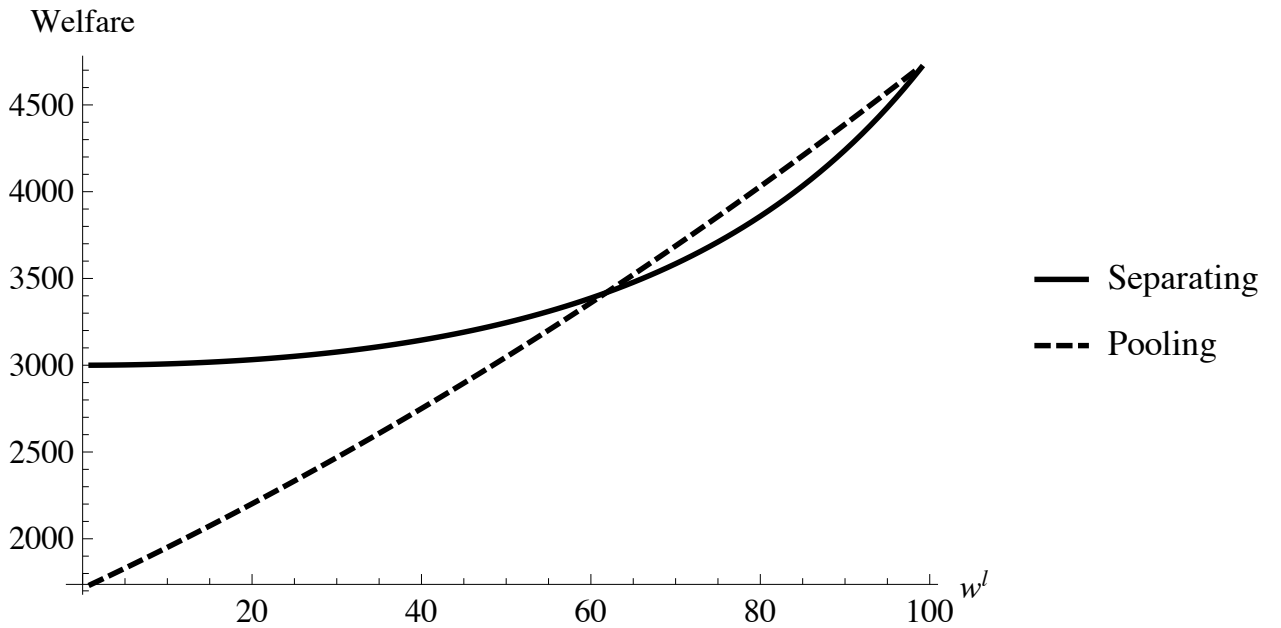


Figure 1: Welfare comparison between the Separating and Pooling Allocation.

Two insights emerge from the figure. First, there exists a threshold level of the wage rate of the low-skilled workers ($w^l \simeq 62$), above which the pooling equilibrium dominates, and below which the separating allocation prevails. Thus, having moderate gains from screening (reflected by a relatively small difference in productivities between the two types of workers) is not only sufficient (as suggested by the proposition) but also necessary for the pooling allocation to be socially desirable. Second, and perhaps most importantly, the superiority of the pooling allocation is not confined to knife-edged cases. Pooling turns out to be socially desirable over a large range of parameters and a shift from a separating to a pooling allocation may yield a substantial welfare gain (up to 4.5 percent increase in the utility of the low-skilled worker when

$w^l \simeq 82$, which is equivalent to an increase of 3.2 percent in the low-skilled consumption level relative to his/her consumption under the separating regime).

4 Conclusion

There are two different channels via which concerns about inequity could be addressed by income taxation: one is by affecting the post-tax income distribution and the other is by affecting the underlying wage distribution. In the standard Mirrlees (1971) setting, labor markets are competitive and wage rates are exogenously given, as skills are perfect substitutes and perfectly observable by the firms. This leaves no scope for redistribution through the wage channel. Stiglitz (1982) and the subsequent literature challenged this prediction focusing on the role of complementarities across different skill types in the production technology.

In this paper we have employed a setting that maintains the Mirrlees (1971) assumption of perfect substitutability across skill types but allows for asymmetric information between firms and workers. We have demonstrated that in such a context, the government can affect the underlying wage distribution by choosing an appropriate tax system that blocks the possibility for firms to engage in screening, thereby implementing a pooling allocation with full wage equalization. We have further shown that the pooling allocation is socially optimal when the preferences for redistribution are sufficiently strong and the differences in productivities are not too large.

The message conveyed by our analysis is fairly general and relates to the role of income taxation as an instrument to attain redistribution through the wage channel by limiting the transmission of information between workers and firms. In this paper we have confined attention to one particular mechanism of information transmission; namely, screening by firms, but other channels, such as signaling by workers, may be considered as well.

References

Cremer, H., Pestieau, P., and Racionero, M. (2011). Unequal wages for equal utilities. *International Tax and Public Finance*, 18(4):383–398.

- Diamond, P. A. (1998). Optimal income taxation: An example with a u-shaped pattern of optimal marginal tax rates. *The American Economic Review*, 88:83–95.
- Ebert, U. (1992). A reexamination of the optimal nonlinear income tax. *Journal of Public Economics*, 49(1):47 – 73.
- Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.
- Rothschild, C. and Scheuer, F. (2013). Redistributive taxation in the Roy model. *The Quarterly Journal of Economics*, 128(2):623–668.
- Rothschild, M. and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649.
- Salanié, B. (2011). *The Economics of Taxation*. The MIT Press.
- Stantcheva, S. (2014). Optimal income taxation with adverse selection in the labor market. *Review of Economic Studies* (forthcoming).
- Stiglitz, J. E. (1982). Self-selection and Pareto-efficient taxation. *Journal of Public Economics*, 17:213–240.