On the Optimal Skill Distribution in a Mirrleesian Economy∗

Tin Cheuk Leung† Hakki Yazici‡
Chinese University of Hong Kong Sabanci University

August 29, 2010

Abstract

A fundamental assumption in Mirrlees (1971) and the literature that follows is that society can only provide redistribution ex post, via transferring consumption between agents. This paper introduces a novel channel of redistribution by allowing planner to choose ex ante distribution of skills and asks: what does the optimal distribution of skills look like? To answer this question we consider a static Mirrleesian economy with two types of agents, high and low. We take the fraction of high and low types as exogenously given and analyze the optimal way to distribute a given level of total skills between these types. We find that it is always socially optimal to choose the perfectly unequal distribution, allocating all the skills to high types and leaving all the low types completely unskilled. We suspect the assumption that there are no complementarities between skilled and unskilled labor is the main reason for this extreme result, which may cast some doubt on other normative results in the Mirrleesian optimal income tax literature.

∗We want to thank James A. Mirrlees for his detailed comments and encouragement.
†Contact: Department of Economics, Room 914, Esther Lee Building, Chung Chi Campus, Shatin, Hong Kong
‡Contact: Faculty of Arts and Social Sciences, Office 2093, Orhanli, Tuzla, 34956 Istanbul, Turkey
1 Introduction

People are heterogenous in their skills to turn effort into output. A central question in normative public economics is to redistribute resources from more skilled individuals to less skilled ones in an efficient way. One policy tool to achieve this is to use income taxation, redistributing resources from high-income individuals to low-income ones. As is well-known, however, income taxation is distortionary when individuals’ skills and efforts are private information (See Mirrlees (1971)).

In this paper, we consider an additional policy tool of redistribution. Suppose, for a given level of average skill in the economy, planner can choose the dispersion of skill distribution. In this case, by choosing a less dispersed distribution, a planner can create an economy with more equal earnings capacity among agents. This implies a more equal distribution of consumption for given income taxes. However, changing skill distribution also has an effect on the overall productivity of the economy. The amount of output that can be produced by a given labor force depends on skill distribution chosen.

We ask how a planner should use these two policy tools jointly to redistribute resources in an efficient way. To answer this question we consider a static Mirrleesian economy in which the planner chooses the skill distribution and income taxes. In the model, the planner first chooses the skill distribution, agents then draw their types from the skill distribution privately, the planner chooses the income tax levels, and finally agents work, pay taxes and consume. The main difference between our model and standard ones is the initial stage of skill distribution choice. The planner, taking the average level of skills as given, chooses the dispersion of the skill distribution. We restrict the set of skill distribution available to the planner to discrete distribution with finite number of mass points. The planner then chooses the value of the mass points.

In such a world, at one extreme, planner can choose a skill distribution in which the value of all mass points is equal to the average skill level. In this extreme, after the skill draw, all agents have the same earnings capacity. We call this perfectly equal skill distribution. At another

\footnote{We make two assumptions here. First, the number of mass points is fixed. Second, the planner takes the probability attached to each mass point as given when he chooses the value of the mass points. The latter assumption is not important since our main result holds regardless of the probability assigned to each mass point.}
extreme lies a skill distribution in which the value of all but one mass point is set to zero. In this extreme, after the draw, a fraction of agents have very high earnings capacity while the rest are completely unproductive. We call this perfectly unequal skill distribution. In between, there is a continuum of skill distribution available to the planner, each with a different level of skill inequality. The main result of our paper is striking: it is socially optimal to have either perfectly equal or perfectly unequal skill distribution, an interior level of skill inequality is never optimal.

In the main body of the paper, we assume that the planner faces a linear skill constraint with two mass points. More precisely, the planner chooses the mass points \(w_1, w_2\) subject to the following skill constraint

\[p_1 w_1 + p_2 w_2 = \alpha,\]

where \(p_1, p_2\) are probabilities attached to the mass points and \(\alpha\) is the average skill level in the society. In this case, we show that socially optimal skill distribution is always the perfectly unequal one, i.e., \(w_i = 0\), for some \(i\). The intuition for the result is as follows. Suppose \(w_i \neq 0\), for all \(i\). In this case, it is obvious that labor levels is positive for both types. Then, by moving to a skill distribution in which the type with higher labor level have all the earnings capacity and setting the labor level of the other type to zero, the planner increases total output and decreases total disutility. This shows that increasing skill inequality benefits the society because it increases productive efficiency. Under full information, income taxes are not distortionary, which means the planner can distribute consumption according to its will using income taxes at no cost. This implies that the productive efficiency gain is the only effect of increasing skill inequality on the economy. Thus, under full information, perfectly unequal skill distribution is socially optimal. However, when skill levels are private information, income taxes are distortionary, and increasing skill inequality exacerbates the distortion associated with income taxation, since the benefit of pretending to be low-skilled is higher for the high-skilled agents. Therefore, increasing skill inequality imposes a cost on the society as well. We show that the productive efficiency gain of increasing skill inequality always dominates its cost.

---

\(^2\)The assumption of two mass points is not restrictive. We show all of our results hold in the case with \(N\) types.
One might think that the optimality of perfectly unequal skill distribution is extreme depends on the specifics of our model. To investigate this claim, we analyze several extensions of the baseline model. We find that the optimal skill distribution is either perfectly unequal or perfectly equal one, but never one in between. First, we relax the linear skill constraint assumption by allowing the skill constraint to be convex in skill levels. In particular, the skill constraint becomes the following

\[ p_1 w_1^\beta + p_2 w_2^\beta = \alpha, \]

where \( \beta \geq 1 \) is the convexity parameter. We show that, under full information, the socially optimal skill distribution is either perfectly unequal or perfectly equal one, depending on the convexity of the skill constraint and the convexity of the disutility function. When there is private information about skills, it is hard to provide an analytical solution. Thus, we parameterize the utility and disutility functions and solve the planner’s problem numerically. We show that the socially optimal skill distribution is again either perfectly equal or perfectly unequal one. This time the concavity of the utility function is also one of the determinants of the solution.

Second, we consider a model that is the same as the baseline model, except that there is an exogenous lower bound on the skill levels of agents. The lower bound can be interpreted as skills that agents are born with. In this case, we compute the socially optimal allocation numerically and find very robustly that the optimal skill distribution is again an extreme one. If the lower bound is low (high) relative to average skill level, then the optimal skill distribution is still the perfectly unequal (equal) one. Finally, we extend the set of skill distributions available to the planner to discrete distributions with three mass points. We show, as in the baseline model, that the optimal skill distribution is again the perfectly unequal one, in which the values of two mass points are set to zero and the value of the remaining mass point to maximum.

The paper closest to ours is Cremer, Pestieau, and Racionero (2007) which has a similar model setup. Their paper only compares the two extreme skill distributions, perfectly equal and perfectly unequal, and show that the perfectly unequal one is better under linear skill constraint. Our paper differs from theirs in two respects. First, they only compare the two extreme skill distributions, perfectly equal and perfectly unequal, and show that the perfectly
unequal one is better under linear skill constraint, whereas we show that the perfectly unequal one is the best among all feasible skill distributions, meaning that it is the socially optimal one. Second, the extensions in our paper suggest that the socially optimal skill distribution is an extreme one, but not necessarily the perfectly unequal one as they imply. A few other papers analyze comparative static properties of optimal allocations with respect to individual skill levels and population fractions of types. Instead of providing a comprehensive survey of this literature, we only talk about two examples briefly. Brett and Weymark (2008) investigates the effect of changing an agent’s skill level on the solution of a Mirrlees optimal income tax problem. Hamilton and Pestieau (2005) studies the effect on the individual utilities of changing the fraction of individuals when the social welfare function is either maximin or maximax. The main difference between these papers and ours is an apparent one: while we analyze the optimal skill distribution, they only study comparative static properties.

In our analysis we do not take a stance on any particular interpretation of the skill distribution choice. However, if we think that skills can be partly attained through education, our model may have implications about education policy. In this regard, our paper is related to several papers that consider education policy as a redistribution tool in the presence of income taxation. Hare and Ulph (1979) shows that, when agents’ learning abilities are heterogeneous and skill types are observable, “optimal choice of education policy reinforces redistributive effect of income tax.” Bovenberg and Jacobs (2006) has a model in which agents choose the education level. Government can only influence education, and thus skill distribution, through education subsidy. They show that providing more subsidies to smarter agents (regressive education policy) would make it more incentive compatible to do more redistribution through income taxes.

Another potential interpretation of skill distribution choice is Skill-Biased Technical Change (SBTC), which refers to “a shift in the production technology that favors skilled over unskilled labor by increasing its relative productivity.” Violante (2009) Under this interpretation, the planner chooses the level of SBTC which determines the level of skill inequality between skilled and unskilled labor. While most papers in the SBTC literature are positive studies on the growth and income distribution implications of SBTC, this paper could be interpreted as a normative

\footnote{Other papers which provide such comparative static results are Boadway and Pestieau (2006), Brett and Weymark (2008), and Simula (2007).}
analysis of the optimal level of SBTC. However, this interpretation of our model should be taken with caution since we treat skilled and unskilled labor as perfect substitutes as in almost all the Mirleesian taxation literature (see Naito (1999) and Stiglitz (1982) for exceptions), contrary to the empirically relevant case of production functions with complementarity.

The rest of the paper is structured as follows. In Section 2 we introduce the model formally. Section 3 calculates optimal skill distribution in the benchmark case where each agent’s skill is publicly observable. Section 4 analyzes the optimal skill distribution problem under private information. Section 5 discusses the extensions of the baseline model. Finally, Section ?? concludes.

2 Model

There is a unit measure of agents and they produce output individually according to the production function

\[ y = wl, \]

where \( y \) denotes output, \( w \) denotes skill level, and \( l \) denotes labor effort.

Each agent’s preferences are given by

\[ u(c) - v(l), \]

where \( c \) is consumption and \( u \) and \( v \) satisfy \( u’, -u'', v’ > 0 \) and \( v'' \geq 0 \).

The novelty of our analysis is that we allow society to choose the distribution of skills. For tractability, we assume that the planner has to choose a distribution in which skills can take only two values, \( w_1 \) and \( w_2 \). The probability of drawing \( w_1 \) is \( p_1 \) and the probability of drawing \( w_2 \) is \( p_2 \). We allow the planner to choose \( w_1 \) and \( w_2 \), but \( p_1 \) and \( p_2 \) are exogenously given. We take the average skill level of the economy as given, at \( \alpha \). We assume, in the main body of the paper, that planner faces a linear skill constraint. Therefore, planner chooses \( w_1 \) and \( w_2 \) subject to

\[ p_1 w_1 + p_2 w_2 \leq \alpha \]
and

\[ w_1, w_2 \geq 0. \]

The first inequality says the average skill level of the distribution chosen by planner cannot exceed \( \alpha \), whereas the second simply says skill levels chosen have to be non-negative.

An allocation in this economy is defined to be \( (w_i, c_i, l_i) \) where \( c_i \) and \( l_i \) represent consumption and labor allocation of \( i = 1, 2 \) type, respectively.

An allocation is socially feasible if

\[ p_2 c_2 + p_1 c_1 \leq p_2 w_2 l_2 + p_1 w_1 l_1 \]  

\[ (1) \]

\[ p_1 w_1 + p_2 w_2 \leq \alpha. \]  

\[ (2) \]

\[ w_1, w_2, c_1, c_2, l_1, l_2 \geq 0. \]  

\[ (3) \]

The first inequality here says that total consumption cannot exceed total output. The second inequality makes sure that the average skill level of the distribution chosen by planner cannot exceed \( \alpha \). Finally, the third inequality is just non-negativity of skill, consumption and labor allocations.

The timing of the events is as follows. First, planner chooses the skill distribution. Then, each agent privately draws her skill from this distribution. Finally, planner chooses the consumption and labor allocation, agents announce their types and receive the corresponding allocation. This informational friction requires the allocation to satisfy the following familiar incentive compatibility conditions:

\[ u(c_2) - v(l_2) \geq u(c_1) - v(w_1 l_1 / w_2) \]  

\[ (4) \]

\[ u(c_1) - v(l_1) \geq u(c_2) - v(w_2 l_2 / w_1) \]  

\[ (5) \]

A social planner chooses the level of consumption, labor and the skill distribution to maximize
total welfare subject to social feasibility and incentive compatibility constraints.

An allocation is a social optimum if it solves\(^4\)

\[
\max_{w_1,w_2,c_1,l_1,c_2,l_2} p_2[u(c_2) - v(l_2)] + p_1[u(c_1) - v(l_1)]
\]

s.t. (1), (2), (3), (4), and (5).

We denote the optimal allocation by \((w_1^*, w_2^*, c_1^*, l_1^*, c_2^*, l_2^*)\).

The main issue we are interested in this paper is the socially optimal skill distribution. Therefore, we focus on \(w_1^*\) and \(w_2^*\) in the above problem. To understand the question at hand, it is helpful to consider the set of distributions that are available to the society. On the one extreme, we can set \(w_1 = 0\) and \(w_2 = \alpha p_2\), or \(w_1 = \alpha p_1\) and \(w_2 = 0\). In both of these cases, we are essentially making agents draw their skills from a perfectly unequal skill distribution. On the other extreme, we can set \(w_1 = w_2 = \alpha\) and make everyone in the society identical. In between, there is a whole range of skill distributions in which both \(w_1, w_2 > 0\). In some of these distributions, \(w_1 > w_2\) and in some \(w_1 < w_2\).

From now on, we will call the type that the planner allocates higher skills \(H\) and the other type \(L\), i.e., \(w_i = w_H\) and \(w_j = w_L\), if \(w_i > w_j\). Also, \(p_i = p_H\) and \(p_j = p_L\). So, redefine an allocation as \((w_H, w_L, c_H, l_H, c_L, l_L)\). The main result of the paper, which we prove in section 4, is that the socially optimal skill distribution involves perfect inequality, meaning \(w_L = 0\) and \(w_H = \frac{\alpha}{p_H}\). As an additional point, we also show which perfectly unequal skill distribution is optimal: one in which type 1 is \(H\) or type 2 is \(H\).

### 2.1 Rewriting Planner’s Problem

Let \(\theta = \frac{w_L}{w_H}\). Observe that \(\theta = 0\) is the case in which there is perfect inequality in skill distribution. As we increase \(\theta\) towards 1, inequality in skill distribution decreases and at \(\theta = 1\) there is perfect equality of skills. In the rest of the paper, we will be interested in the value of socially optimal \(\theta\).

\(^4\)We use a utilitarian social welfare function with equal weights on all agents. However, all of our results hold under any social welfare function that values equality beyond laissez-faire market outcome. The reason is that the only feature of this utilitarian social welfare function that we are relying on is that the high-skilled type’s incentive constraint binds, which is true under any social welfare function that values equality.
It is a well-known result that only type $H$ incentive constraint binds under Utilitarian social welfare function with equal weights. Now we can rewrite the problem as:

$$\max_{\theta,c_L,l_L,c_H,l_H} \quad p_H[u(c_H) - v(l_H)] + p_L[u(c_L) - v(l_L)]$$

s.t.

$$p_H c_H + p_L c_L \leq \alpha l_H + \frac{\alpha p_L \theta (l_L - l_H)}{p_H + p_L \theta}$$

$$u(c_H) - v(l_H) \geq u(c_L) - v(\theta l_L)$$

$$c_L, c_H, l_L, l_H \geq 0.$$  
$$\theta \in [0,1].$$

If the planner sets $\theta = 1$, then agents choose their types from the perfectly equal skill distribution where all agents have the skill level $\alpha$. In that case, right hand side of feasibility becomes $p_H \alpha l_H + p_L \alpha l_L$ and incentive compatibility constraint disappears.

3 Benchmark: Full Information Social Optimum

We first analyze the benchmark case with full information. The planner’s problem is the same as above problem except for there is no incentive compatibility constraint:

$$\max_{\theta,c_L,l_L,c_H,l_H} \quad p_H[u(c_H) - v(l_H)] + p_L[u(c_L) - v(l_L)]$$

s.t.

$$p_H c_H + p_L c_L \leq \alpha l_H + \frac{\alpha p_L \theta (l_L - l_H)}{p_H + p_L \theta}$$

$$c_L, c_H, l_L, l_H \geq 0.$$  
$$\theta \in [0,1].$$

It is always optimal for planner to choose the perfectly unequal skill distribution, that is
giving all the skill to one of the two types, in the full information case. It is intuitive to see why. In the absence of the incentive constraint, planner can distribute output the way he wants. Therefore, skill distribution choice has no effect on consumption distribution across agents, which means that the only criterion of optimal skill distribution is productive efficiency. It is easy to see that productive efficiency requires skill distribution to be perfectly unequal. Suppose this is not true, \( \theta^* \in (0, 1] \). From the first order optimality condition between labor of \( H \) and \( L \), we have:

\[
v'(l^*_L) = \theta^* v'(l^*_H)
\]

which implies \( l^*_H \geq l^*_L \) since \( v'' > 0 \) and \( \theta^* \in (0, 1] \). We can find another feasible allocation that strictly improves welfare. Consider a new allocation \( \hat{\theta} = 0, \hat{l}_L = 0 \), and everything else stays the same. The feasibility constraint is relaxed since

\[
\alpha l^*_H + \frac{\alpha p_L (l^*_L - l^*_H)}{p_H + p_L \hat{\theta}} = \alpha l^*_H + \frac{\alpha p_L \theta^* (l^*_L - l^*_H)}{p_H + p_L \theta^*}
\]

and the disutility of \( L \) type decreases since

\[
v(\hat{l}_L) = v(0) < v(l^*_L)
\]

which says that any \( \theta \in (0, 1] \) cannot be optimal. Therefore, we have the following theorem:

**Theorem 1.** In the full information social optimum with linear skill constraint, \( \theta^* = 0 \).

The above theorem does not indicate whether type 1 or type 2 should be type \( H \). However, one can easily note that in the full information optimum with linear skill constraint, the type with lower measure is chosen to be type \( H \), meaning if \( p_i > p_j \), then \( w^*_i = 0, w^*_j = \frac{\alpha}{p_j} \), since a lower measure of agents work to produce the same amount of output.

4 Private Information Social Optimum

With private information, the choice of skill distribution not only affects productive efficiency, but also, through the incentive constraint, restricts the set of consumption distributions available to the planner.
The following lemma characterizes a property of the resource allocation which will be used in Theorem 2.

**Lemma 1.** For any given \( \theta \in [0, 1] \), if \( c_H(\theta), c_L(\theta), l_H(\theta), l_L(\theta) \) solves the planner’s problem above and \( l_H(\theta) \geq l_L(\theta) \), then we have \( u(\frac{\alpha}{p_H}l_H(\theta)) - v(l_H(\theta)) \geq u(0) - v(0) \).

**Proof.** First, we can show that \( \frac{\alpha}{p_H}l_H(\theta) > c_H(\theta) \) since the resource constraint is

\[
(p_H)c_H(\theta) + p_Lc_L(\theta) = \alpha[l_H(\theta) + \frac{p_Lw(l_L(\theta) - l_H(\theta))}{p_H + p_L\theta} \leq 0]
\]

if we give all skill to type \( H \) and set \( \hat{\theta} = 0 \), total output increases to \( \alpha l_H(\theta) \). If we give all output to type \( H \), he gets \( \frac{\alpha}{p_H}l_H(\theta) > c_H(\theta) \).

Second, we can show \( u(\frac{\alpha}{p_H}l_H(\theta)) - v(l_H(\theta)) \geq u(0) - v(0) \) because

\[
\begin{align*}
u(\frac{\alpha}{p_H}l_H(\theta)) - v(l_H(\theta)) &> u(c_H(\theta)) - v(l_H(\theta)) \\
&= u(c_L(\theta)) - v(\theta l_L(\theta)) \\
&\geq u(c_L(\theta)) - v(l_L(\theta)) \\
&> u(0) - v(0)
\end{align*}
\]

The first inequality is because we show \( \frac{\alpha}{p_H}l_H(\theta) > c_H(\theta) \). The second line is the incentive compatibility constraint. The last inequality holds because \( c_H(\theta), l_H(\theta), c_L(\theta), l_L(\theta) = 0 \) would be an optimal allocation otherwise.

As it turns out, in Theorem 2, the optimal skill distribution is still the perfectly unequal one.

**Theorem 2.** In the private information social optimum, \( \theta^* = 0 \).

**Proof.** We proceed with two steps.

**Step 1: \( \theta^* \) cannot be interior.**

Suppose not, \( \theta^* \in (0, 1) \), there are two cases to consider.

**Case 1:** \( l_H^* \geq l_L^* \)
Thanks to Lemma 1, we can set $\hat{\theta} = 0$, $\hat{l}_L = 0$, $\hat{c}_H > c^*_H$ and $\hat{c}_L$ such that the incentive constraint still holds:

$$u(c_H) - v(l_H^*) = u(\hat{c}_L) - v(0)$$

The resource constraint is more relaxed since

$$\alpha l_H^* + \frac{\alpha p_L \hat{\theta}(\hat{l}_L - l_H^*)}{p_H + p_L \hat{\theta}} = \alpha l_H^* + \frac{\alpha p_L \theta^*(l_L^* - l_H^*)}{p_H + p_L \theta^*}$$

The overall utility increases since $\hat{c}_H > c^*_H$ and $\hat{c}_L = 0 < l_L^*$.

Case 2: $l_H^* > l_L^*$

We can set $\hat{\theta} = 1$, then resource constraint is relaxed since

$$\alpha l_H^* + \frac{\alpha p_L \hat{\theta}(l_L^* - l_H^*)}{p_H + p_L \hat{\theta}} = \alpha l_H^* + \frac{\alpha p_L \theta^*(l_L^* - l_H^*)}{p_H + p_L \theta^*}$$

The incentive constraint is also relaxed since

$$u(c_H^*) - v(l_H^*) = u(c_L^*) - v(\theta^* l_L^*) > u(c_L^*) - v(\hat{\theta} l_L^*) = u(c_L^*) - v(\hat{l}_L^*)$$

Thus it is easy to find another allocation that improves welfare.

**Step 2: $\theta^* = 0$**

Suppose not, $\theta^* = 1$, then it is easy to show $c_H^* = c_L^*$ and $l_H^* = l_L^*$, but then we can set $\hat{\theta} = 0$ and $\hat{l}_L = 0$. The resource constraint is unaffected since

$$\alpha l_H^* + \frac{\alpha p_L \hat{\theta}(l_L^* - l_H^*)}{p_H + p_L \hat{\theta}} = \alpha l_H^* + \frac{\alpha p_L \hat{\theta}(\hat{l}_L - l_H^*)}{p_H + p_L \hat{\theta}} = \alpha l_H^*$$

And we can set $\hat{c}_H > c^*_H$ and $\hat{c}_L$ such that the incentive constraint holds due to Lemma 1

$$u(c_H) - v(l_H^*) = u(\hat{c}_L) - v(0)$$

This improves welfare since $\hat{c}_H > c^*_H$ and $\hat{l}_L = 0 < l_L^*$.

In the private information social optimum as well, the planner would distribute all skills to the type which has a lower measure.
5 Extensions

In this section, we argue that the main result of the paper, that socially optimal allocation of income generating skills is either perfectly equal or perfectly unequal, is actually quite robust.

5.1 Convex Skill Constraint

In the previous analysis, we analyze a special case in which the skill constraint is linear. In reality, however, there can be decreasing return to scale when we transfer skill from low to high type. We here assume a convex skill constraint as follow:

\[
p_L w_L^\beta + p_H w_H^\beta \leq \alpha \tag{6}
\]

where \( \beta \geq 1 \) is the scalar for convexity of the skill constraint.

Using this skill constraint, we can write the total output as:

\[
p_L w_L l_L + p_H w_H l_H = w_H(\theta p_L l_L + p_H l_H) = \left( \frac{\alpha}{p_L \theta^\beta + p_H} \right)^{1/\beta}(\theta p_L l_L + p_H l_H)
\]

5.1.1 Full Information

First, we analyze the benchmark case with full information. Using the above expression for total output, we can write the planner problem as:

\[
\max_{\theta, c_H, c_L, l_H, l_L} p_H[u(c_H) - v(l_H)] + p_L[u(c_L) - v(l_L)]
\]

s.t.

\[
p_H c_H + p_L c_L = \left( \frac{\alpha}{p_L \theta^\beta + p_H} \right)^{1/\beta}(\theta p_L l_L + p_H l_H)
\]

Denote \( c_H^*(\theta), c_L^*(\theta), l_H^*(\theta), l_L^*(\theta) \) as the values of \( c_H, c_L, l_H, l_L \) that maximizes the above problem for a given \( \theta \), and denote \( U^* \) as the maximized total utility:
$$U^* \equiv p_H[u(c^*_H(\theta)) - v(l^*_H(\theta))] + p_L[u(c^*_L(\theta)) - v(l^*_L(\theta))]$$

We are interested in the optimal value of \( \theta \) that maximizes \( U^* \). The Envelop Theorem implies that the total derivative of the total utility with respect to \( \theta \) can be expressed as the product of the multiplier to the resource constraint (\( \lambda \)) and the partial derivative of total output with respect to \( \theta \):

$$dU^*/dw_L = \frac{\partial \lambda}{\partial \theta} \left[ \left( \frac{p_H^\alpha}{p_L^\beta + p_H} \right)^{1/\beta} (\theta p_L l^*_L(\theta) + p_H l^*_H(\theta)) - p_H c^*_H(\theta) - p_L c^*_L(\theta) \right]$$

The above expression says that whether more skill equality (higher \( \theta \)) can increase welfare depends on whether it can increase total output. Productive efficiency is the only concern since, under full information, there is no incentive constraint to restrict the planner from equalizing consumption. Using the above expression, whether total output would increase depends on the sign of this expression:

$$[l^*_L(\theta) - \theta^{\beta-1} l^*_H(\theta)] = l^*_L(\theta)[1 - \frac{\theta^{\beta-1}}{l^*_L(\theta)/l^*_H(\theta)}]$$

Whether output would increase with higher skill equality depends on two things. First, the convexity of disutility function would determine the labor ratio, \( \frac{l^*_L}{l^*_H} \). In particular, the less convex the disutility is, the higher the labor ratio would be. Second, \( \beta \) determines the convexity of skill constraint. When skill constraint is more convex, the marginal gain from increasing skill equality would be higher. Here let us assume the disutility has the form of \( v(l) = l^\gamma \), then the above expression becomes

$$[l^*_L(\theta) - \theta^{\beta-1} l^*_H(\theta)] = l^*_L(\theta)[1 - \theta^{\beta-1} + \frac{1}{1-\gamma}]$$

Thus, how welfare changes w.r.t. \( \theta \) depends on the sign of \( \beta - 1 + \frac{1}{1-\gamma} \)
5.1.2 Private Information

Under private information, there is an additional incentive constraint in the planner problem:

\[
\max_{\theta, c_H, c_L, l_H, l_L} p_H[u(c_H) - v(l_H)] + p_L[u(c_L) - v(l_L)]
\]

s.t.

\[
\begin{align*}
p_H c_H + p_L c_L &= (\frac{\alpha}{p_L \theta^\beta + p_H})^{1/\beta}(\theta p_L l_L + p_H l_H) \\
u(c_H) - v(l_H) &= u(c_L) - v(\theta l_L)
\end{align*}
\]

Let \( c_L^*(\theta), l_L^*(\theta), c_H^*(\theta), l_H^*(\theta) \) be the optimal allocations at \( \theta \), and \( L^* \) be the value of the Lagrangian, then the Envelop theorem says that the total derivative of the total utility, \( U^* \), w.r.t. \( \theta \) is the same as the partial derivative of the Lagrangian, \( L^* \), w.r.t. \( \theta \), where

\[
\frac{dU^*}{d\theta} = \frac{\partial L^*}{\partial \theta} = \frac{\alpha^{1/\beta} p_H p_L}{(\theta p_L l_L^*(\theta) + p_H l_H^*(\theta))^{1/\beta + 1}} [l_L^*(\theta) - \theta^{\beta-1} l_H^*(\theta)] + \frac{\phi v'(\theta l_L^*(\theta)) l_L^*(\theta)}{\triangle \text{utility due to incentive constraint (positive)}}
\]

It is difficult to derive analytical solution. We provide numerical solution at this moment. To do this, we have to specify functional forms for utility and disutility.

For utility, we use constant relative risk aversion utility

\[
u(c) = \frac{c^{1-\sigma}}{1-\sigma}
\]

For disutility, we use a simple convex function

\[
v(l) = l^\gamma
\]

In this numerical exercise, we set \( \gamma = 2 \). Since the concavity of the utility function has an impact on the result, we compute the results with \( \sigma = 0.5 \) and \( \sigma = 3 \). Also, to illustrate
the importance of convexity of the skill constraint on the result, we compute the results with 4 values of $\beta$.

Figures 1 and 2 show the separate marginal effects of resource and incentive constraints on total utility with $\sigma = 0.5$ and $\sigma = 3$ respectively. The green curve is the marginal reduction in utility due to reduction in output when $\theta$ is higher; the blue curve is the marginal increase in utility due to the relaxed incentive constraint. Judging from the figures, it is easy to see that there cannot be any interior solution for $\theta$.

As Figures 3 and 4 show, there is no interior solution for $\theta$. Several factors determine which corner the optimal $\theta$ would lie. The first factor is the convexity of the skill constraint. A more linear skill constraint makes unequal skill distribution less costly. As we prove in previous section, when $\beta = 1$, $\theta = 0$ is the optimal solution, which is confirmed by Figures 3 and 4.

Second, in full information case, when $\beta - 1 + \frac{1}{1 - \gamma} \geq 0$, $\theta = 1$ would be optimal. It is still true in private information case since the labor ratio, $\frac{\ell_H}{\ell_L}$, is lower under private information due to the incentive constraint. This is confirmed by the lower panels of Figures 3 and 4.

Third, for $\beta - 1 + \frac{1}{1 - \gamma} < 0$ and $\beta > 1$, the optimal $\theta$ would depend on the concavity of the utility ($\sigma$). In particular, with a more concave utility (higher $\sigma$), the planner would make the labor ratio, $\frac{\ell_H}{\ell_L}$, lower to further equalize the consumptions among the two types. This would decrease the marginal cost of equalizing skills (increasing $\theta$), thus making skill equalization ($\theta = 1$) more appealing. This is confirmed by the right upper sub-figures of Figures 3 and 4.

5.2 Additive innate skills

Consider again the baseline model with linear skill constraint. Suppose everyone has $K$ units of innate skills.

Define $\tilde{\theta} = \frac{w_L + K}{w_H + K}$ and $\tilde{\alpha} = \alpha + K$.

Then, feasibility is:

$$p_H c_H + p_L c_L \leq \tilde{\alpha} l_H + \frac{\tilde{\alpha} p_L \tilde{\theta}(l_L - l_H)}{p_H + p_L \tilde{\theta}}.$$ (7)
Figure 1:

Marginal Cost (Reduction in Output) and Marginal Gain (Relaxed Incentive Constraint) of Higher $\theta$ ($\sigma = 0.5$)
Figure 2:

Marginal Cost (Reduction in Output) and Marginal Gain (Relaxed Incentive Constraint) of Higher $\theta$ ($\sigma = 3$)
Figure 3:

Total Utility \((\sigma = 0.5)\)
Figure 4:

Total Utility ($\sigma = 3$)
And incentive compatibility is:

\[ u(c_H) - v(l_H) \geq u(c_L) - v(\tilde{\theta}l_L). \]

So, this problem is exactly the same as the problem in the baseline model where the average skill in the society is \( \bar{\alpha} \) and the planner is restricted to choose among skill distributions in which each mass point has to be greater equal to \( K \). In terms of \( \theta \), this means the planner is restricted to choose \( \tilde{\theta} \in [\frac{K}{\bar{\alpha}/p_H+K}, 1] \). When \( K \) is low relative to \( \alpha \), that means the lower bound on \( \tilde{\theta} \) is low and Figure 4 tells us that we still have \( \theta^* = 0 \) as socially optimal. On the other hand, when \( K \) is high enough relative to \( \alpha \), the lower bound on \( \tilde{\theta} \) is high and hence Figure 4 tells us that \( \theta^* = 1 \), meaning perfectly equal skill distribution is socially optimal. As a result, even in the extended model with innate skills, we never have an interior distribution to be optimal. In a similar fashion, one can show that the extreme result that either \( \theta^* = 0 \) or \( \theta^* = 1 \) is robust to introducing innate abilities to the convex skill constraint model as well.

5.3 3 types

**Theorem 3.** Under full information, \( \theta_1 = \theta_2 = 0 \) is socially optimal.

*Proof.* Same as two types case. \( \square \)

**Theorem 4.** Under private information, \( \theta_1 = \theta_2 = 0 \) is socially optimal.

*Proof.* Suppose for contradiction that both \( \theta_1 \) and \( \theta_2 \) are both strictly positive (other cases are straightforward implications of two types case).

There are three possibilities:

Case 1. \( l_3^1 \) is the largest labor level.

In this case, set \( \theta_1 = \theta_2 = 0 \) and \( l_1 = l_2 = 0 \). This increases total output while decreasing total disutility. Observe that \( u(c_3^*) - v(l_3^*) < u(c_2^*) - v(0) \) and \( u(c_2^*) - v(0) > u(c_1^*) - v(0) \), so for incentive compatibility we need \( \tilde{c}_2 = \tilde{c}_1 \), and \( u(\tilde{c}_3) - v(l_3^*) \geq u(\tilde{c}_2) - v(0) \). We need a version of Lemma 3 to show that the latter holds, i.e. we need \( u(\alpha l_3^*) - v(l_3^*) \geq u(0) - v(0) \). Suppose not.

21
Then, $u(c_3^*) - v(l_3^*) < u(0) - v(0)$, which implies $u(c_1^*) - v(l_1^*) < u(c_2^*) - v(l_2^*) < u(c_3^*) - v(l_3^*) < u(0) - v(0)$, which is a contradiction since then making all produce an deal $0$ dominates the efficient allocation. So, in the new allocation, we have $u(\hat{c}_3) - v(l_3^*) = u(\hat{c}_2) - v(0) = u(\hat{c}_1) - v(0)$. If $\hat{c}_3 \geq c_3^*$, then that means type 3 is better off which further means type 1 and type 2 are strictly better off. That means new allocation improves over efficient one, which is a contradiction. If $\hat{c}_3 < c_3^*$, then another allocation in which we set $c_3$ equals to $c_3^*$ and divide the rest of output equally between type 1 and type 2 strictly improves over the efficient allocation and is incentive comapitable.

**Case 2.** $l_2^* < l_1^*$.

Set $\hat{w}_3 = w_3 - \epsilon$ and $\hat{w}_1 = w_1 + \epsilon$. This increases total output and relaxes incentive constraint from type 3 to type 2 and type 2 to type 1, creating a contradiction.

**Case 3.** $l_1^* < l_3^* < l_2^*$.

Set $\hat{w}_1 = 0$ and $\frac{w_2}{w_3} = \frac{\hat{w}_2}{\hat{w}_3}$. This increases total output and decreases type 1’s disutility to $0$. Lemma 3 in this case is as follows: We want to show that there exists $\hat{c}_1, \hat{c}_2, \hat{c}_3$ such that $u(\hat{c}_3) - v(l_3^*) \geq u(0) - v(0)$, $u(\hat{c}_2) - v(l_2^*) \geq u(0) - v(0)$, and $u(\hat{c}_3) - v(l_3^*) \geq u(\hat{c}_2) - v(0)$. Use extra output to increase $c_2, c_3$ above $c_2^*, c_3^*$ such that the third holds with equality. Suppose the first does not hold. Then, $u(c_3^*) - v(l_3^*) < u(0) - v(0)$, which gives a contradiction as in case 1. Now suppose for contradiction that $u(\hat{c}_2) - v(l_2^*) < u(0) - v(0)$. Then, consider a new allocation in which we give both type 1 and 2 $(0, 0)$ and type 3 $(c_3^*, l_3^*)$. This is incentive compatible, feasible, and gives strictly higher utility, a contradiction. So, in the perturbed allocation, $u(\hat{c}_3) - v(l_3^*) > u(\hat{c}_2) - v(l_2^*) = u(\hat{c}_1) - v(0)$. Since $u(\hat{c}_3) - v(l_3^*) > u(c_3^*) - v(l_3^*)$ and $u(\hat{c}_2) - v(l_2^*) = u(\hat{c}_1) - v(0) > u(c_2^*) - v(l_2^*) > u(c_1^*) - v(l_1^*)$, perturbed allocation improves over the efficient allocation, yielding the desired contradiction.

\[\Box\]

### 6 Conclusion

To Be Written...
References


