Optimal Skill Distribution under Convex Skill Costs*

Tin Cheuk Leung[†]
Wake Forest University

Hakki Yazici[‡]
Sabanci University

November 17, 2017

Abstract

This paper studies optimal distribution of skills in an optimal income tax framework with convex skill constraints. The problem is cast as a social planning problem where a redistributive planner chooses how to distribute a given amount of aggregate skills across people. We find that optimal skill distribution is either perfectly equal or perfectly unequal, but an interior level of skill inequality is never optimal.

Keywords: Skill Distribution, convex skill costs, optimal taxation

JEL Classification: H2

*We would like to thank James Mirrlees for his detailed comments and encouragement. We also thank Piero Gottardi, Pierre Pestieau and seminar participants at PET 10 for their comments.

[†]Department of Economics, Kirby Hall, Wake Forest University, Box 7505, Winston - Salem NC 27109, USA. Email: leungtc@wfu.edu

[‡]Contact: Faculty of Arts and Social Sciences, Office 2093, Orhanli, Tuzla, 34956 Istanbul, Turkey. Email: hakkiyazici@sabanciuniv.edu

1 Introduction

How should income taxes be designed in the face of economic inequality that stems from differences in worker skills? In a seminal analysis, Mirrlees (1971) analyzes this issue and shows that whenever workers' skills are private information, income taxation is distortionary, and optimal income taxation is shaped by a trade-off between equality and efficiency. In this paper, we extend the analysis in Mirrlees (1971) by allowing the government to choose the distribution of skills in the economy in addition to income taxes.

Specifically, we consider a static Mirrleesian economy in which the planner chooses the skill distribution and income taxes. The timing of events is the model is as follows. First, the government chooses a skill distribution taking the average skill level as given. Second, the government chooses the income tax system. Third, agents draw their types from the skill distribution privately. Finally, given their skills, agents work, pay taxes and consume. The main difference between our model and that of Mirrlees (1971) is the first stage of where, taking the average level of skills as given, the government chooses the dispersion of the skill distribution. Traditional models of Mirrleesian optimal taxation take the distribution of skills in the economy as given, and hence, do not have this initial stage.

We restrict the set of skill distributions available to the government to discrete distributions with two mass points. More precisely, the government chooses mass points w_1, w_2 subject to the following skill constraint,

$$p_1 w_1^\beta + p_2 w_2^\beta = \alpha,$$

where p_1, p_2 are exogenous probabilities attached to the mass points and α is the average skill level in society. The convexity parameter $\beta \geq 1$ controls the technology of skill conversion across agents. When $\beta = 1$, in order to increase type 2 agents' skills by one unit, the government needs to decrease type 1 agents' skills by $\frac{p_2}{p_1}$ units independent of the level of the skill levels. When $\beta > 1$, however, the cost of increasing one type's skills is increasing in that type's skill level. In other words, there is diminishing returns to investing in skills.

In this economy, there are two extremes regarding the dispersion of skills. At one extreme, there is a skill distribution in which $w_1 = w_2$, meaning, all agents have the same earnings capacity. We call this the *perfectly equal* skill distribution. If the government chooses this skill distribution, there is no redistributive purpose for income taxation: redistribution is carried out solely via skill distribution choice. At the other extreme, the government can choose a skill distribution in which only one type has positive skills, i.e., $w_1 = 0$ or $w_2 = 0$. Here, a fraction of agents have very high earnings capacity while the rest are completely unproductive. We call this the *perfectly unequal* skill distribution. Redistribution needs to be carried out ex-post in this economy via income taxes. In between the two extreme distributions, there is a continuum of skill distributions with different levels of skill dispersion.

In a closely related paper, Leung and Yazici (2017), we analyze the optimal skill distribution problem in a similar framework under the assumption that $\beta = 1$, i.e., the planner faces a linear skill constraint. There, we prove that, whenever $\beta = 1$, the socially optimal skill distribution is always perfectly unequal, i.e., $w_i = 0$, for some i. The main novelty of the current paper over Leung and Yazici (2017) is that here we allow for $\beta > 1$, meaning we allow for convex skill distributions. This is an important generalization. As discussed earlier, when $\beta > 1$, the cost of increasing one type's skills is increasing in that type's skill level. In other words, $\beta > 1$ case is akin to the assumption of diminishing returns to investing in skills, and there is a large body of empirical evidence that supports the notion that human capital investment features diminishing returns.[1]

We show that under full information, the socially optimal skill distribution is either perfectly unequal or perfectly equal, depending on the convexity of the skill constraint and the convexity of the disutility function. When there is private information about skills, we provide a sufficient condition for the optimality of perfectly equal skill distribution that depends on the convexity of the skill constraint and the convexity of the disutility function. When this condition does not hold, it is hard to provide an analytical solution. Instead, we parameterize the utility and disutility functions, and solve the optimal skill distribution problem numerically. We find that the socially optimal skill distribution is again either perfectly equal or perfectly unequal. In this case, we observe that, in addition to the convexity of the skill constraint and the convexity of the disutility function, the level of concavity of the utility function also matters for whether perfectly equal or unequal distribution is optimal.

[1]See Mincer (1974), Psacharopoulos (1985), Psacharopoulos (1994) and Harmon and Walker (1999), among others.

This paper is also closely related to Cremer, Pestieau, and Racionero (2010). Like Leung and Yazici (2017), Cremer, Pestieau, and Racionero (2010) assume a linear skill constraint and show that the perfectly unequal skill distribution provides strictly higher social welfare than perfectly equal skill distribution. In the current paper, we go beyond the linear skill constraint assumption and provide an analysis of optimal skill distribution under diminishing returns to skill acquision. Boadway and Pestieau (2006), Simula (2007), and Hamilton and Pestieau (2005) analyze comparative static properties of optimal allocations with respect to certain parameters of the skill distribution.

The rest of this paper is structured as follows. In Section 2, we introduce the model formally. In Section 3, we characterize optimal skill distribution both in the cases in which skills are public and private information. Section 4 provides concluding remarks.

2 Model

There is a unit measure of agents. 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 preference is given by

$$u(c) - v(l),$$

where c is consumption and u and v satisfy $u', -u'', v' > 0$ and $v'' > 0$.

The novelty of our analysis is that we allow the government to choose the distribution of skills. For simplicity, it is assumed that 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 government to choose w_1 and w_2 subject to the given total skill level α, but p_1 and p_2 are exogenously given. In other words, the government chooses w_1 and w_2 subject to the following skill constraint:

$$p_1 w_1^\beta + p_2 w_2^\beta \leq \alpha.$$

Allocation. An *allocation* in this economy is defined as $(w_i, c_i, l_i)_{i=1,2}$, where c_i and l_i represent consumption and labor allocation of type i.

Feasibility. An allocation is *feasible* if

$$p_2c_2 + p_1c_1 \leq p_2w_2l_2 + p_1w_1l_1, \tag{1}$$

$$p_1w_1^\beta + p_2w_2^\beta \leq \alpha, \tag{2}$$

$$w_1, w_2, c_1, c_2, l_1, l_2 \geq 0. \tag{3}$$

The first inequality means that total consumption cannot exceed total output. The second inequality ensures that the average skill level of the distribution chosen by the government does not exceed α. Finally, the third inequality is just the non-negativity of skill, consumption and labor allocations.

The timing of the events is as follows. First, the government chooses the skill distribution. Then, the government chooses a tax function $T : \mathcal{R}_+ \to \mathcal{R}$, where $T(y)$ is the income tax that an agent with income y pays. Then, each agent privately draws her skill from the chosen skill distribution. Finally, each agent chooses his optimal consumption and labor allocation given the tax system. Taking w_i as given, agent i solves the following consumption-labor problem.

$$\begin{gathered}\max\nolimits_{c_i,l_i} u(c_i) - v(l_i) \\ s.t. \\ c_i \leq w_il_i - T(w_il_i).\end{gathered} \tag{4}$$

Government's Optimal Tax Problem. The government chooses the distribution of skills and the tax function to maximize the total welfare in the economy subject to the fact that the resulting allocation solves each agent's problem.

$$\begin{gathered}\max\nolimits_{w_1\geq 0, w_2 \geq 0, T(\cdot)} p_2[u(c_2) - v(l_2)] + p_1[u(c_1) - v(l_1)] \\ \text{s.t. } (2) \text{ and} \\ \text{for each i, } (c_i, l_i) \text{ solves } (4).\end{gathered} \tag{5}$$

Social Planning Problem. By the Revelation Principle, the government's optimal tax problem given by (5) is equivalent to a social planning problem in which a planner chooses the skill distribution and the consumption and labor allocations directly as functions of agents' types. The assumption that taxes only depend on income and not on agents' types in the optimal tax problem introduces the restriction that agents' types are private information in the planning problem. Since it is significantly easier to solve the planning problem and the planning problem and the optimal tax problem give identical solution regarding thhe optimal allocation and the skill distribution, in the rest of the paper we focus on the solution to the social planning problem.

The timing of events in the planning problem is similar to that in the optimal tax problem. The planner first chooses the skill distribution. Then, the planner chooses the consumption and labor allocations as functions of agents' types.Then, each agent privately draws her skill from the chosen skill distribution. Finally, agents announce their types and receive the corresponding allocation. This informational friction requires the allocation to satisfy the following familiar *incentive compatibility* conditions:

Incentive compatibility. An allocation is *incentive compatible* if

$$u(c_2) - v(l_2) \geq u(c_1) - v(w_1 l_1 / w_2) \tag{6}$$

$$u(c_1) - v(l_1) \geq u(c_2) - v(w_2 l_2 / w_1) \tag{7}$$

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.

Social Optimum. An allocation is a *social optimum* if it solves[2]

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

s.t. (1), (2), (3), (6), and (7).

[2]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 the laissez-faire market outcome. The only feature of this utilitarian social welfare function on which we rely is that the high-skilled type's incentive constraint binds, which is true under any social welfare function that values equality.

We denote the optimal allocation by $(w_1^*, w_2^*, c_1^*, l_1^*, c_2^*, l_2^*)$. As we are interested in the socially optimal skill distribution, we focus on w_1^* and w_2^* in the above problem. The range of distributions that are available to society are as follows. At one extreme, the planner can set $w_1 = 0$ and $w_2 = (\frac{\alpha}{p_2})^{1/\beta}$, or $w_1 = (\frac{\alpha}{p_1})^{1/\beta}$ and $w_2 = 0$. In both of these cases, a fraction of agents have very high earnings capacity while the rest are completely unproductive. We call these perfectly unequal skill distributions. On the other extreme, we can set $w_1 = w_2 = \alpha^{1/\beta}$ and make everyone in the society identical. We call this the perfectly equal skill distribution. 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 denote by H the type that the planner allocates higher skills and by L the other type, i.e., $w_i = w_H$ and $w_j = w_L$, if $w_i > w_j$. In addition, let $p_i = p_H$ and $p_j = p_L$. Hence, we redefine an allocation as $(w_H, w_L, c_H, l_H, c_L, l_L)$.

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 θ 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 θ^*.

Using $\theta = \frac{w_L}{w_H}$ transformation, one can rewrite the right-hand side of the resource constraint as follows:

$$\begin{aligned} p_L w_L l_L + p_H w_H l_H &= w_H(\theta p_L l_L + p_H l_H) \\ &= (\frac{\alpha}{p_L \theta^\beta + p_H})^{1/\beta}(\theta p_L l_L + p_H l_H). \end{aligned}$$

It is a well-known result that only the type H incentive constraint binds under a Utilitarian social welfare function with equal weights. Using $\theta = \frac{w_L}{w_H}$ transformation, we can write the H type incentive constraint as:

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

Therefore, the planning problem becomes:

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

s.t.

$$p_H c_H + p_L c_L \leq (\frac{\alpha}{p_L \theta^\beta + p_H})^{1/\beta}(\theta p_L l_L + p_H l_H)$$
$$u(c_H) - v(l_H) \geq u(c_L) - v(\theta l_L)$$
$$c_L, c_H, l_L, l_H \geq 0$$
$$\theta \in [0, 1].$$

If the planner sets $\theta = 1$, then agents choose their types from the perfectly equal skill distribution where all agents have the skill level α. In this case, the incentive compatibility constraint disappears.

3 Characterizing the Optimal Skill Distribution

In this section, we characterize the optimal distribution of skills for the economy outlined in Section 2.

3.1 Full Information

First, we analyze the benchmark case with full information.[3] Using the above expression for total output, we can write the planner problem as:

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

[3]In terms of the optimal tax problem, the full information planning problem corresponds to a tax problem that is identical to (5) except the government chooses type-dependent lump-sum taxes (T_L, T_H) instead of an income tax function.

s.t.

$$p_H c_H + p_L c_L = (\tfrac{\alpha}{p_L \theta^\beta + p_H})^{1/\beta} (\theta p_L l_L + p_H l_H)$$
$$c_H, c_L, l_H, l_L \geq 0, \theta \in [0,1].$$

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 maximize the above problem for a given θ, and $\lambda^*(\theta)$ the optimal multiplier associated with that problem. Denote $U^*(\theta)$ as the maximized total utility:

$$U^*(\theta) \equiv p_H[u(c_H^*(\theta)) - v(l_H^*(\theta))] + p_L[u(c_L^*(\theta)) - v(l_L^*(\theta))].$$

Then, the planner's problem of choosing optimal θ becomes:

$$\max_\theta U^*(\theta) - \lambda^*(\theta)\Big[(\frac{\alpha}{p_L \theta^\beta + p_H})^{1/\beta}(\theta p_L l_L^*(\theta) + p_H l_H^*(\theta)) - p_H c_H^*(\theta) - p_L c_L^*(\theta)\Big]. \tag{8}$$

The solution to this problem is characterized by the first-order condition with respect to θ. Differentiating (8) with respect to θ and equating that to zero reveals that the derivative of $U^*(\theta)$ with respect to θ can be expressed as the product of the multiplier to the resource constraint and the partial derivative of total output with respect to θ:

$$\begin{aligned}\frac{dU^*(\theta)}{d\theta} &= \frac{\partial \lambda^*(\theta)\Big[(\frac{\alpha}{p_L\theta^\beta + p_H})^{1/\beta}(\theta p_L l_L^*(\theta) + p_H l_H^*(\theta)) - p_H c_H^*(\theta) - p_L c_L^*(\theta)\Big]}{\partial \theta} \\ &= \lambda^*(\theta) \frac{\alpha^{1/\beta} p_H p_L}{(\theta p_L l_L^*(\theta) + p_H l_H^*(\theta))^{1/\beta+1}} [l_L^*(\theta) - \theta^{\beta-1} l_H^*(\theta)]. \end{aligned} \tag{9}$$

Expression (9) above shows that whether a higher degree of skill equality (higher θ) can increase welfare depends on whether it can increase total output. Productive efficiency is the only concern because, under full information, no incentive constraint restricts the planner from equalizing consumption. A careful examination of (9) reveals that the sign of the right-hand side of (9) depends on the sign of

$$[l_L^*(\theta) - \theta^{\beta-1} l_H^*(\theta)] = l_L^*(\theta)[1 - \frac{\theta^{\beta-1}}{l_L^*(\theta)/l_H^*(\theta)}], \tag{10}$$

since the remaining part of the right-hand side of (9) has to be positive (notice that the multiplier, $\lambda^*(\theta)$, has to be positive). This implies that whether total output (and hence total welfare) increases with θ or not depends on the sign of (10).

To simplify the analysis, we assume a particular form for the disutility function, $v(l) = l^\gamma$, where $\gamma > 1$. Then, (10) becomes

$$[l_L^*(\theta) - \theta^{\beta-1} l_H^*(\theta)] = l_L^*(\theta)[1 - \theta^{\beta-1+\frac{1}{1-\gamma}}].$$

Recalling that $\theta \in [0, 1]$, this implies that output and hence welfare is strictly increasing with θ at all $\theta \in [0, 1)$ if and only if $\beta - 1 + \frac{1}{1-\gamma} > 0$. Similarly, this implies that output and hence welfare is strictly decreasing with θ at all $\theta \in [0, 1)$ if and only if $\beta - 1 + \frac{1}{1-\gamma} < 0$. Whenever $\beta - 1 + \frac{1}{1-\gamma} = 0$, output is independent of θ, and thus any $\theta \in [0, 1]$ is optimal. The following theorem summarizes these results.

Theorem 1. *In the full information social optimum,*

$$\theta^* = \begin{cases} 0 & \text{if } \beta - 1 + \frac{1}{1-\gamma} < 0; \\ [0,1] & \text{if } \beta - 1 + \frac{1}{1-\gamma} = 0; \\ 1 & \text{if } \beta - 1 + \frac{1}{1-\gamma} > 0. \end{cases}$$

The theorem states that optimal skill distribution is either perfectly equal or unequal, depending on the degree of convexity of the skill constraint, β, and the degree of the convexity of the disutility function, γ. In particular, for a given γ, the higher β is the more likely it becomes that perfect equality in skill distribution ($\theta^* = 1$) is optimal. This is intuitive. As β increases, skill constraint becomes more convex and hence it becomes more and more costly to transfer all skills to one type of workers. Thus, perfect equality is more likely to be optimal. Moreover, for a given β, the higher γ is the more likely it becomes that perfect equality in skill distribution is optimal. This is also intuitive. As γ increases, disutility of effort becomes more convex, which implies that it is optimal for the planner to bring labor supplies of high and low skilled individuals closer to each other in the optimal allocation. As the two types work more similar hours, the benefit of setting a more unequal skill distribution toward high skilled agents decrease. As a result, it becomes more likely that $\theta^* = 1$ is optimal.

Notice that when $\beta = 1$, Theorem 1 implies that, independent of γ (as long as γ is greater than 1 which amounts to disutility being strictly convex), $\theta^* = 0$. Intuitively, $\beta = 1$ implies that there is no cost of giving all the skills to one type of agent since the skill constraint is linear. Thus, under this assumption, as long as $l_H^* > l_L^*$ in the optimal allocation (which holds for any $\gamma > 1$ that is finite), it is optimal to give all the skills to the high type.

3.2 Private Information

Under private information, the planning problem is the same as the full information planning problem, except that there is an additional constraint– the usual incentive compatibility constraint.

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

We know from the full information analysis of the previous subsection that productive efficiency calls for the optimality of the perfectly equal skill distribution whenever $\beta - 1 + \frac{1}{1-\gamma} \geq 0$. As the perfectly equal skill distribution also brings perfect consumption equality without any further need for distortionary income taxation, it is also socially optimal under private information whenever this condition is satisfied. The theorem below follows.

Theorem 2. *If $\beta - 1 + \frac{1}{1-\gamma} \geq 0$, then we have $\theta^* = 1$ in the private information social optimum.*

Proof. Consider a relaxed version of the private information planning problem in which we drop the incentive constraint. That relaxed problem is equivalent to the full information planning problem and we know that in the solution to that problem we have $\theta = 1$, $c_H = c_L$, and $l_H = l_L$. Clearly, this allocation satisfies the incentive constraint and hence is in the constraint set of the original planning problem under private information, which means that it solves the problem. □

It is difficult to provide an analytical solution to this problem when $\beta - 1 + \frac{1}{1-\gamma} < 0$. Therefore, in what follows we provide numerical solutions. We assume utility function is of the constant relative risk aversion utility functional form:

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}.$$

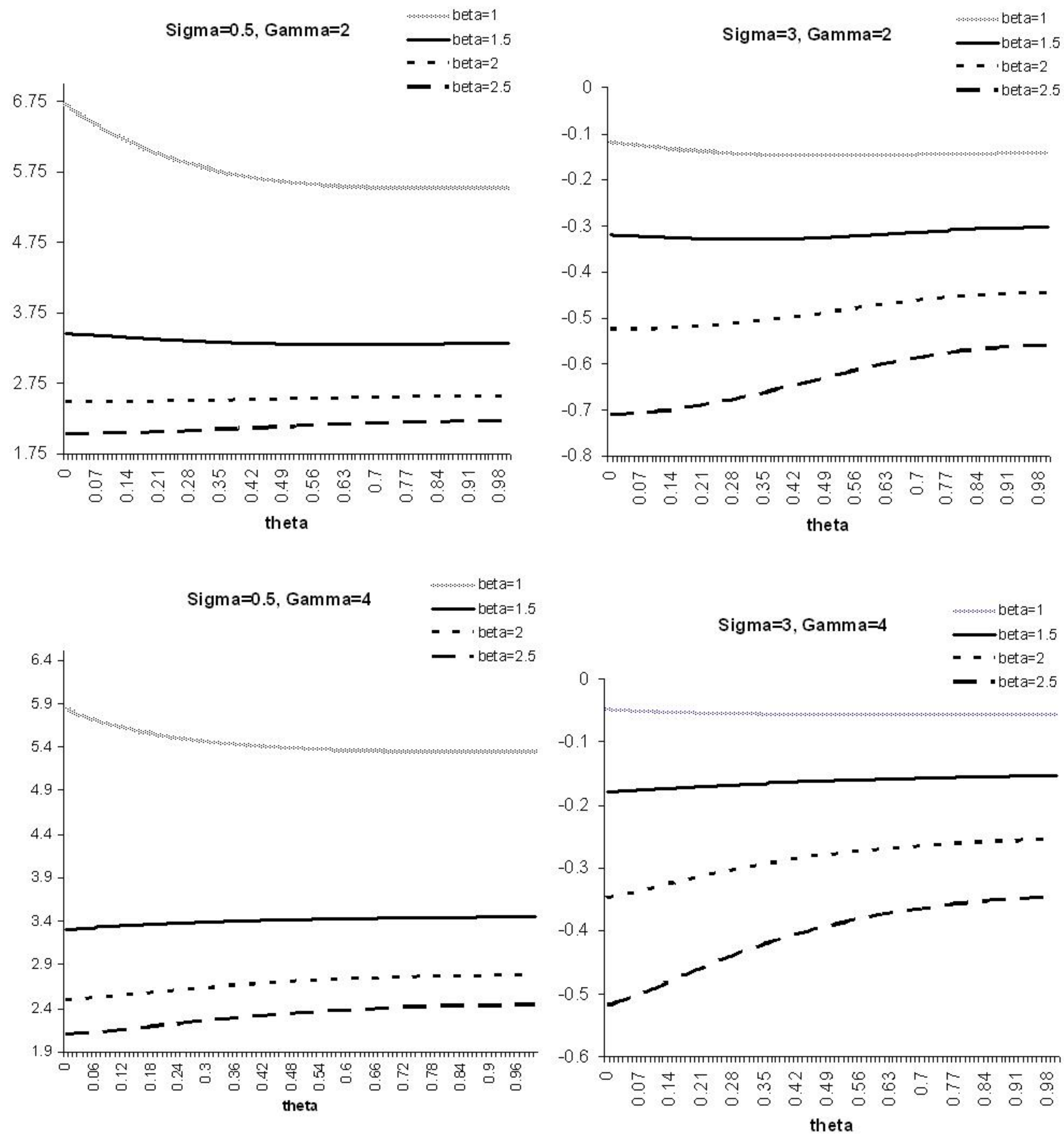
We compute the value of the planning problem for value of θ that is $[0, 1]$, and plot these value functions in Figure 1 under different parameter configurations. As the concavity of the utility function and convexity of the disutility function affect the result, we compute the value function with two values of σ ($\sigma = 0.5$ and $\sigma = 3$) and two values of γ ($\gamma = 2$ and $\gamma = 4$). The four sub-figures of Figure 1 each correspond to one of the four combinations of σ and γ. To illustrate the importance of the convexity of the skill constraint on the result, we compute the value function for 4 different values of β. Within each sub-figure, each plot corresponds to a value function for a different value of β.

The main message of Figure 1 is that there is no combination of parameters (β, γ, σ) for which optimal θ is interior. Several factors determine the corner in which the optimal θ will lie. The first factor is the convexity of the skill constraint. As the intuition in the full information case suggests, a more linear skill constraint (lower β) makes unequal skill distribution less costly and hence more likely to be socially optimal. Figure 1 confirms this: fixing γ and σ (i.e. looking within each sub-figure), θ^* changes from 0 to 1 as β increases. Also, observe that as it is shown in Leung and Yazici (2017), when $\beta = 1$, $\theta^* = 0$, which is true in all four sub-figures.

Second, again similar to the full information case, a more convex disutility (higher γ) implies that $\frac{l_L^*}{l_H^*}$ has to be higher. Thus, increasing skill inequality (lower θ) increases output less, and hence is less likely to be optimal. In Figure 1, holding σ and β constant, a higher γ (from 2 to 4) makes $\theta^* = 1$ more likely. For instance, when $\sigma = 0.5$ and $\beta = 1.5$, θ^* changes from 0 to 1 when γ goes up from 2 to 4.

The third parameter that matters for the optimal θ is σ, the concavity parameter of the utility function. Under private information, for any σ, if the planner wants the high type to produce more output, he must provide the high type with incentives to do that, meaning $c_H > c_L$. Now, keeping everything else constant, if we look at an economy with a higher σ, the planner would like to set consumption levels of the two types closer to each other. To do this, however, the planner has to close the gap between l_H and l_L, which makes skill equality more appealing. In Figure 1, holding β and γ constant, a higher σ (from 0.5 to 3) makes $\theta^* = 1$ more likely. For instance, when $\gamma = 2$ and $\beta = 1.5$, θ^* changes from 0 to 1 when σ goes up from 0.5 to 3.

Figure 1:



Value Functions

4 Conclusion

This paper studies the optimal distribution of skills in a Mirrleesian economy with convex skill constraints. Under full information, the socially optimal skill distribution is either perfectly unequal or perfectly equal, depending on the convexity of the skill constraint and the convexity of the disutility function. For the case in which individual skills are private information, we provide a sufficient condition for the optimality of perfectly equal skill distribution that again depends on the convexity of the skill constraint and the convexity of the disutility function. When this condition does not hold, it is hard to provide an analytical solution. Instead, we provide a numerical analysis of the optimal skill distribution problem. We find that the socially optimal skill distribution is again either perfectly equal or perfectly unequal. In this case, in addition to the convexity of the skill constraint and the convexity of the disutility function, the level of concavity of the utility function also matters for results.

References

Boadway, R. W., and P. Pestieau (2006): "Tagging and Redistributive Taxation," *Annales d'Économie et de Statistique*, (83-84), 123–147.

Cremer, H., P. Pestieau, and M. Racionero (2010): "Unequal Wages for Equal Utilities," *International Tax and Public Finance*, forthcoming.

Hamilton, J., and P. Pestieau (2005): "Optimal Income Taxation and the Ability Distribution: Implications for Migration Equilibria," *International Tax and Public Finance*, 12(1), 29–45.

Harmon, C., and I. Walker (1999): "The marginal and average returns to schooling in the UK," *European Economic Review*, 43(4-6), 879–887.

Leung, T., and H. Yazici (2017): "On Optimal Skill Distribution in a Mirrleesian Economy," *Working Paper*.

Mincer, J. (1974): *Schooling, Experience, and Earnings*. National Bureau of Economic Research, Inc.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38(114), 175–208.

Psacharopoulos, G. (1985): "Returns to Education: A Further International Update and Implications," *Journal of Human Resources*, 20(4), 583–604.

——— (1994): "Returns to investment in education: A global update," *World Development*, 22(9), 1325–1343.

Simula, L. (2007): "Optimality Conditions and Comparative Static Properties of Non-linear Income Taxes Revisited," *Paris School of Economics Working Paper*, (2007-15).