Adverse Selection and Income Inequality

Yaping Wu
Institute of Economics and Finance, School of Finance,
Nanjing Audit University

Shenghao Zhu*
School of International Trade and Economics,
University of International Business and Economics

February 4, 2024

Abstract

This study investigates the inequality implications of incentive-feasible contracts, including optimal and nonoptimal contracts, when there is a trade-off between rent extraction and efficiency. We show that information frictions cause inequality. And changes in social norms influence inequality under the optimal contract.

Keywords: contract, monotonicity, income inequality, information friction, social-norm change

JEL classification: D31, D63, D82

*We would like to thank Jess Benhabib, Jie Cai, Zhiyun Li, Thomas Sargent, Yifei Sun, and Yitong Wang for their comments, suggestions and help. We also thank Yunling Kang for his help on mathematics and thank Nanxi Wen for her help on programming.

1 Introduction

In this study, we investigate the inequality implications of incentive-feasible contracts, including optimal and nonoptimal contracts, when a trade-off exists between rent extraction and efficiency. In our model, contracts pin down the agent's payment. Thus, contracts transform their types into incomes. Piketty and Saez (2003) demonstrate that the top income shares in the United States have shown a rapidly increasing trend since the 1970s. Piketty (2020) argues that ideology is an important force determining social inequality.[1] We show that different contracts (i.e., mechanisms) have distinct implications for income inequality.

There is a continuum of agents with measure 1 in the economy. An agent's marginal cost of production is θ. Even though the distribution of Θ is common knowledge, the realization of θ is the agent's private information and is unobservable to the principal. We study the adverse selection problem and show that information rent increases income inequality. We investigate the inequality implications of incentive-feasible contracts. Incentive-feasible contracts satisfy both incentive and participation constraints. We find that information rents increase income inequality for incentive-feasible contracts under certain conditions.

Furthermore, we find that output scheme determines the payment schedule for any feasible contract. The payment includes two parts: production cost and information rent. Information rent can be expressed as a function of output. Therefore, payment is a function of the output level. We can characterize the payment schedule by investigating the properties of the output scheme. In this sense, the output scheme contains sufficient information to study the inequality implications of any feasible contract.

For the two-type distribution of Θ, we find that information rant increases the income inequality. For the continuous type, we find that information rent increases the income inequality if ratio $\frac{U(\theta)}{\theta q(\theta)}$ is decreaseing in θ. We verify that the constant elasticity distribution and the uniform distribution imply the monotonicity of this ratio. We also find a sufficient condition of the output scheme guaranteeing the monotonicity of this ratio.

[1] Ideology refers to "a set of a priori plausible ideas and discourses describing how society should be structured."

Using this sufficient condition, we find more examples of the distribution. One of them has non-differentiable density function.

This study investigates income distribution when agents' payments are not in line with the marginal cost. In the neoclassical theory of distribution, however, actors receive payments according to their marginal product. Information frictions create a wedge between agents' payments and their marginal costs. We find that information frictions cause inequality and situations under which we can rank contracts according to their induced income inequality. We rank the income equality of these contracts using Lorenz ordering. The Lorenz curves can be ranked without intersections.

To investigate the effects of asymmetric information on income inequality, we identify the optimal contract under asymmetric information and under complete information. Each contract induces income inequality in the economy. We then compare inequality under asymmetric information and under complete information. We find that inequality under asymmetric information is less equal than that under complete information. The optimal contract incurs less equal output distribution under incomplete information than under complete information. We invent a decomposition technique, which separates two channels of output distortion and information rent. Through both channels asymmetric information causes income inequality. A less equal output distribution is the first channel, and information rent further exaggerates income inequality.

We also investigate the transformation of the type distribution. Distribution $G(\theta)$ is derived by transformation $G(\theta) = F(\theta)^{\beta}$ for $\beta \geq 1$. If distribution $F(\theta)$ satisfies the condition by which the income distributions under the second-best contract are less equal than those under the first-best contract, then we can show that $G(\theta)$ also satisfies that condition

We study the inequality implications of incentive-feasible contracts. However, the principal's objective function in our model does not reflect the equity concern. Baron and Myerson (1982) use the weighted sum of expected gains for consumers and the expected profit for the firm as the social welfare function. Specifically, they use a parameter to represent the relative weight between consumers and the firm. Based on their study, we introduce the parameter $\alpha \in [0, 1]$, regarding the concern for the agent's utility, into the social welfare function. The change in α represents changes in social norms. We find that inequality, measured by Lorenz ordering in our model, increases if α decreases. Changes in social norms influence inequality under the optimal contract. Piketty and Saez (2003) propose that changes in social norms with regard to inequality partly explain the increase in top wage shares since the 1970s.

1.1 Related literature

This study is also related to that of Lazear and Rosen (1981). Both studies compare income distributions under different contracts. Lazear and Rosen (1981) investigate how different incentive-inducing contracts under moral hazard generate different income distributions. Our model considers income distributions implied by different incentive-inducing contracts under adverse selection. Lazear and Rosen (1981) find that the tournament mechanism can produce skewed income distribution. Similarly, we find that the optimal contract under asymmetric information can generate income distribution that is more dispersed than that under complete information. Lazear and Rosen (1981) investigate aggregate welfare for two cases: that of risk-neutral agents and that of risk-averse agents.[2] We focus on the situation of risk-neutral agents; thus, the principal's objective function does not incorporate the risk-sharing incentive in our model.

Ekeland (2010) views the optimal contract as the transportation map. It transforms the type ditribution into the salary distribution. We investigate the inequality implications of the contract. We compare the income distributions under different contracts and study how the income distribution changes when the type distribuiton changes. Campbell et al. (2021) investigate principal-agent mean field games. The principal-agent mean field game consists of an infinite number of agents and the principal cares about the state distribution among agents. While Campbell et al. (2021) consider dynamic games between the principal and agents, we concentrate on a static situation and investigate the impacts of the contract on the income distribution.

[2] Other studies of tournament-based compensation schemes include Green and Stokey (1983) and Nalebuff and Stiglitz (1983). A tournament is a typical scheme in a moral hazard problem with many agents. Similar to studies in the optimal taxation literature, our model focuses on independent contracts among agents.

Fernandez and Gali (1999) compare markets and tournaments in an economy with borrowing constraints. They focus on the allocation efficiency of these two mechanisms. As a complement to studies that compare the efficiency of different mechanisms, our study compares income distributions under different contracts. Even though in Fernandez and Gali (1999) the initial wealth distribution is exogenous, income distribution depends on the mechanism in their model and is endogenous. In our model, income distribution depends on contracts and is also endogenous.

Dworczak et al. (2021) investigate the role of price regulation redistribution in a market with private information. While Dworczak et al. (2021) study the optimal mechanism design, the policymakers in their model have equity concern. The principal in our model has no equity concern. Thus, we focus on the implications of optimal contracts that only reflect the production (allocation) dimension. We intentionally shut down equity concern in the contract design. In this sense, inequality is an “unintentional” product of the contract's incentive stimulus effects.

Mirrlees (1971) and Saez (2001) investigate optimal income taxation in models with unobservable productivity. The objective function of the principal (the government) is utilitarian; it is the sum of the utility function of all agents in the economy. The social welfare function incorporates equity concern since the agent's utility function is risk-averse and has curvature. Our model differs from this in two aspects. First, the principal's objective function has no equity concern and only reflects the production dimension. Second, the optimal taxation literature usually focuses on the optimal tax scheme whereas we investigate the inequality implications of incentive-feasible contracts, including optimal and nonoptimal contracts.

The rest of this paper is organized as follows. We present the inequality implications of feasible contracts in Section 2. We investigate the effect of information frictions on income distribution in Section 3. We study the type-distribution transformation in Section 4. Section 5 contains an analysis of social-norm change. Section 6 concludes the paper.

2 Feasible contracts and inequality

There is a continuum of agents with measure 1 in the economy. The agent's marginal cost of production is θ. Even though the distribution of Θ is common knowledge, the realization of θ is unobservable to the principal, and it is the agent's private information.

2.1 Characterization of feasible contracts

We assume the following:

Assumption 1: Θ follows a two-type discrete probability distribution,

$$\Theta = \begin{cases} \underline{\theta}, & \text{with probability } v \\ \bar{\theta}, & \text{with probabiity } 1-v \end{cases}.$$

Let $\Delta\theta$ denote the spread of marginal cost,

$$\Delta\theta = \bar{\theta} - \underline{\theta} > 0.$$

There is no heterogeneity among principals. Principals run firms and hire agents. The principal offers the contract $\{(t(\underline{\theta}), q(\underline{\theta})); (t(\bar{\theta}), q(\bar{\theta}))\}$. The agent chooses to claim his type $\tilde{\theta}$. If $\tilde{\theta} = \underline{\theta}$, the agent receives payment $t(\underline{\theta})$ and provides output $q(\underline{\theta})$ to the principal's firm. If $\tilde{\theta} = \bar{\theta}$, the agent receives payment $t(\bar{\theta})$ and provides output $q(\bar{\theta})$ to the principal's firm. We view $t(\underline{\theta})$ and $t(\bar{\theta})$ as agents' incomes. The agents have quasi-linear preferences, $t(\underline{\theta}) - \underline{\theta}q(\underline{\theta})$ and $t(\bar{\theta}) - \bar{\theta}q(\bar{\theta})$. The incentive compatibility constraints are

$$t(\underline{\theta}) - \underline{\theta}q(\underline{\theta}) \geq t(\bar{\theta}) - \underline{\theta}q(\bar{\theta}), \tag{1}$$

and

$$t(\bar{\theta}) - \bar{\theta}q(\bar{\theta}) \geq t(\underline{\theta}) - \bar{\theta}q(\underline{\theta}). \tag{2}$$

Let $\underline{t} = t(\underline{\theta})$, $\underline{q} = q(\underline{\theta})$, $\bar{t} = t(\bar{\theta})$, and $\bar{q} = q(\bar{\theta})$. Thus, any pair of outputs $(\underline{q}, \bar{q})$ that are implementable must satisfy the implementability condition $\underline{q} \geq \bar{q}$. From the efficiency perspective, the principal provides the $\underline{\theta}-$type agents with incentives to induce them to produce more outputs than the $\bar{\theta}-$type agents. Even though the principal's objective function might only care about efficiency or rent extraction, the incentive compatibility conditions themselves have income inequality implications. As in Laffont and Martimort (2002), we define information rents as

$$\underline{U} = \underline{t} - \underline{\theta}\underline{q},$$

for efficient agents and

$$\bar{U} = \bar{t} - \bar{\theta}\bar{q},$$

for inefficient agents. A feasible contract must satisfy incentive compatibility constraints (1) and (2) and participation constraints

$$\underline{U} \geq 0,$$

and

$$\bar{U} \geq 0.$$

The agent's payment scheme consists of

$$\underline{t} = \underline{\theta}\underline{q} + \underline{U},$$

and

$$\bar{t} = \bar{\theta}\bar{q} + \bar{U}.$$

Unless we have $\bar{U} = \underline{U} = 0$, the agents' payments are not in line with the cost.

To maximize surplus, the principal sets $\bar{U} = 0$. Thus, $\bar{t} = \bar{\theta}\bar{q}$. For the $\bar{\theta}-$type agent, payment $\bar{t}$ equals production cost $\bar{\theta}\bar{q}$. For the $\underline{\theta}-$type agent, the incentive compatibility constraints cause a disparity between payments and production costs, $\underline{U} = \Delta\theta\bar{q}$.

Proposition 1 *For the two-type case, we have* $\underline{t}/\bar{t} \geq (\underline{\theta}\underline{q})/(\bar{\theta}\bar{q})$.

Owing to information rent, the income difference between these two types is larger than the difference in production costs. To keep the incentive compatibility constraint, the principals must provide sufficient incentives to the $\underline{\theta}-$type agents. From an efficiency perspective, the principal provides the $\underline{\theta}-$type agent payments greater than than the production costs. However, a feasible contract enlarges the income difference among agents. We use this two-type distribution as a simple motivation case. However, the intuition that incentive compatibility constraints cause income differences among agents holds in more general cases.

For the continuous type, we assume the following:

Assumption 1′: Θ has a cumulative distribution function $F(\theta)$ and a probability density function $f(\theta) > 0$ on $[a, b]$.

The incentive compatibility condition is

$$t(\theta) - \theta q(\theta) \geq t(\tilde{\theta}) - \theta q(\tilde{\theta}),$$

for any $(\theta, \tilde{\theta})$ in $[a, b] \times [a, b]$. As in Myerson (1981), the local incentive compatibility implies that $q(\theta)$ is nonincreasing in θ. Therefore, $q(\theta)$ is almost everywhere differentiable on $[a, b]$.

The information rent function is

$$U(\theta) = t(\theta) - \theta q(\theta),$$

for θ in $[a, b]$. The incentive compatibility implies that $U(\theta)$ is continuous and nonincreasing in θ. Therefore, $U(\theta)$ is almost everywhere differentiable on $[a, b]$,

$$\dot{U}(\theta) = -q(\theta) \quad a.e.$$

for θ in $[a, b]$. Thus, we have

$$U(\tilde{\theta}) = U(\theta) - \int_{\theta}^{\tilde{\theta}} q(\tau) d\tau,$$

for $a \leq \theta \leq \tilde{\theta} \leq b$. To maximize surplus, the principal sets $U(b) = 0$. Thus, we have

$$U(\theta) = \int_{\theta}^{b} q(\tau) d\tau,$$

for θ in $[a, b]$. The relationship between information rent and the output is due to the incentive compatibility condition. Thus, the shape of the output scheme determines information rent.

Let $\rho(\theta)$ denote compensation for the cost of the $\theta-$type agent, $\rho(\theta) = \theta q(\theta)$. We investigate the ratio of information rent to $\rho(\theta)$.

Proposition 2 *For the continuous-type case, if*

$$\frac{\theta_2}{\theta_1} \geq \frac{\int_{\theta_2}^{b} \frac{q(\tau)}{q(\theta_2)} d\tau}{\int_{\theta_1}^{b} \frac{q(\tau)}{q(\theta_1)} d\tau},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$*, then we have*

$$\frac{t(\theta_1)}{\rho(\theta_1)} \geq \frac{t(\theta_2)}{\rho(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$*.*

Proof: For θ in $[a, b]$, we have $\frac{U(\theta)}{\theta q(\theta)} = \frac{\int_\theta^b q(\tau) d\tau}{\theta q(\theta)} = \frac{1}{\theta} \int_\theta^b \frac{q(\tau)}{q(\theta)} d\tau$. Thus, we have

$$\frac{U(\theta_1)}{\theta_1 q(\theta_1)} = \frac{1}{\theta_1} \int_{\theta_1}^{b} \frac{q(\tau)}{q(\theta_1)} d\tau \geq \frac{1}{\theta_2} \int_{\theta_2}^{b} \frac{q(\tau)}{q(\theta_2)} d\tau = \frac{U(\theta_2)}{\theta_2 q(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, since $\frac{\theta_2}{\theta_1} \geq \frac{\int_{\theta_2}^{b} \frac{q(\tau)}{q(\theta_2)} d\tau}{\int_{\theta_1}^{b} \frac{q(\tau)}{q(\theta_1)} d\tau}$ for $a \leq \theta_1 \leq \theta_2 \leq b$.

Therefore, we have

$$\frac{t(\theta_1)}{\rho(\theta_1)} \geq \frac{t(\theta_2)}{\rho(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, Since $\frac{t(\theta)}{\rho(\theta)} = 1 + \frac{U(\theta)}{\rho(\theta)} = 1 + \frac{U(\theta)}{\theta q(\theta)}$ for θ in $[a, b]$. ■

Since $U(b) = 0$, we have $t(b) = \rho(b)$. For θ in $[a, b]$, the payment function is $t(\theta)$ and the cost compensation function is $\rho(\theta)$. If $\rho(\theta)$ is decreasing in θ, payment schedule $t(\theta)$ is decreasing in θ since $t(\theta) = \rho(\theta) + U(\theta)$. Proposition 2 states that $t(\theta)$ decreases relatively faster than $\rho(\theta)$.

We find that information rent increases the income inequality if ratio $\frac{U(\theta)}{\theta q(\theta)}$ is decreaseing in θ. The wedge between the payment schedule and the cost compensation is the information rent, which amplifies the income difference among agents. We use the monotonicity of $\frac{U(\theta)}{\theta q(\theta)}$ to investigate income inequality. For $a \leq \theta_1 \leq \theta_2 \leq b$, $\frac{t(\theta_1)}{\rho(\theta_1)} \geq \frac{t(\theta_2)}{\rho(\theta_2)}$ also implies that information rent accounts for a higher proportion of income for the more efficient agents (the agents with lower θ) since $\frac{U(\theta)}{t(\theta)} = 1 - \frac{\rho(\theta)}{t(\theta)}$.

However, condition

$$\frac{\theta_2}{\theta_1} \geq \frac{\int_{\theta_2}^{b} \frac{q(\tau)}{q(\theta_2)} d\tau}{\int_{\theta_1}^{b} \frac{q(\tau)}{q(\theta_1)} d\tau},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, might not be easily verified. We need to find a sufficient condition for it.

Lemma 1 *For the continuous-type distribution, we have*

$$\frac{U(\theta_1)}{b-\theta_1} \geq \frac{U(\theta_2)}{b-\theta_2},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$.

Proof: Since $q(\theta)$ is nonincreasing in θ, we have $\frac{\int_{\theta_1}^{\theta_2} q(\tau)d\tau}{\theta_2-\theta_1} \geq q(\theta_2)$ and $\frac{\int_{\theta_2}^{b} q(\tau)d\tau}{b-\theta_2} \leq q(\theta_2)$, for $a \leq \theta_1 \leq \theta_2 \leq b$. Thus, we have

$$\frac{\int_{\theta_1}^{\theta_2} q(\tau)d\tau}{\theta_2-\theta_1} \geq \frac{\int_{\theta_2}^{b} q(\tau)d\tau}{b-\theta_2},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. Furthermore, we have

$$\begin{aligned}
\frac{\int_{\theta_1}^{b} q(\tau)d\tau}{b-\theta_1} &= \frac{\theta_2-\theta_1}{b-\theta_1}\frac{\int_{\theta_1}^{\theta_2} q(\tau)d\tau}{\theta_2-\theta_1} + \frac{b-\theta_1}{b-\theta_2}\frac{\int_{\theta_2}^{b} q(\tau)d\tau}{b-\theta_2} \\
&\geq \frac{\theta_2-\theta_1}{b-\theta_1}\frac{\int_{\theta_2}^{b} q(\tau)d\tau}{b-\theta_2} + \frac{b-\theta_1}{b-\theta_2}\frac{\int_{\theta_2}^{b} q(\tau)d\tau}{b-\theta_2} \\
&= \frac{\int_{\theta_2}^{b} q(\tau)d\tau}{b-\theta_2},
\end{aligned}$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. ■

The result of this lemma is intuitive. Ratio $\frac{U(\theta)}{b-\theta}$ represents the average output level in $[\theta, b]$. The average output level is decreasing since $q(\theta)$ is nonincreasing in θ. Using this lemma, we find a sufficient condition for the monotonicity of $\frac{t(\theta)}{\rho(\theta)}$.

Proposition 3 *For the continuous-type case, if*

$$\frac{\theta_1 q(\theta_1)}{b-\theta_1} \leq \frac{\theta_2 q(\theta_2)}{b-\theta_2},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, then we have

$$\frac{t(\theta_1)}{\rho(\theta_1)} \geq \frac{t(\theta_2)}{\rho(\theta_2)}.$$

for $a \leq \theta_1 \leq \theta_2 \leq b$.

Proof: From Lemma 1 we have $\frac{\int_{\theta_1}^b q(\tau)d\tau}{b-\theta_1} \geq \frac{\int_{\theta_2}^b q(\tau)d\tau}{b-\theta_2}$, for $a \leq \theta_1 \leq \theta_2 \leq b$. If $\frac{\theta_1 q(\theta_1)}{b-\theta_1} \leq \frac{\theta_2 q(\theta_2)}{b-\theta_2}$, we have

$$\frac{\int_{\theta_1}^b q(\tau)d\tau}{b-\theta_1}\frac{b-\theta_1}{\theta_1 q(\theta_1)} \geq \frac{\int_{\theta_2}^b q(\tau)d\tau}{b-\theta_2}\frac{b-\theta_2}{\theta_2 q(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. Thus, we have

$$\frac{\int_{\theta_1}^b q(\tau)d\tau}{\theta_1 q(\theta_1)} \geq \frac{\int_{\theta_2}^b q(\tau)d\tau}{\theta_2 q(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. Therefore, we have

$$\frac{t(\theta_1)}{\rho(\theta_1)} \geq \frac{t(\theta_2)}{\rho(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, since $\frac{t(\theta)}{\rho(\theta)} = 1 + \frac{U(\theta)}{\rho(\theta)} = 1 + \frac{U(\theta)}{\theta q(\theta)}$, for θ in $[a, b]$. ■

Whether infromation rent increases the income inequality depends on the shape of output scheme. Specifically, the monotonicity of $\frac{\theta}{b-\theta}q(\theta)$ implies that infromation rent increases the income inequality. To verify the condition in Proposition 3 we do not have to calculate the integral involved in Proposition 2. Note that $\frac{\theta}{b-\theta}$ is increasing in θ, while $q(\theta)$ is nonincreasing in θ. Our sufficient condition requires that $q(\theta)$ does not decrease too fast as θ increases. In Section 3 we see that the uniform distribution satisfies this condition. We concentrate on the impact of information rent on income inequation in this section, and separate two different channels, through which information frictions cause income inequality, in Section 3.

2.2 Lorenz ordering

We can use Lorenz ordering to compare income distributions under different contracts. Let $F_X(\cdot)$ be the distribution function of a nonnegative random variable X with a finite positive mean. Following Gastwirth (1971), we define the Lorenz curve as follows:

Definition 1 *The Lorenz curve of X, $L_X(p)$, is defined as*

$$L_X(p) = \frac{1}{E(X)}\int_0^p F_X^{-1}(r)dr, \forall p \in [0,1],$$

where $F_X^{-1}(r) = \inf\{x \geq 0 : F_X(x) \geq r\}$.

From the definition of the Lorenz curve, we know that for any constant $c > 0$, X and cX share the same Lorenz curve. Thus, multiplying a random variable by a positive constant does not influence its Lorenz curve. We then define Lorenz ordering as follows:

Definition 2 *For two nonnegative random variables X and Y, X Lorenz dominates Y if and only if*

$$L_X(p) \geq L_Y(p),$$

for all p in $[0,1]$, denoted as $X \succeq_L Y$.

$X \succeq_L Y$ implies that Y is less equal than X. Then the Gini coefficient of X is smaller than that of Y. To establish the Lorenz ordering between two nonnegative random variables, we can find the connection between Lorenz ordering and second-order stochastic dominance. Following Ok (2023), we define second-order stochastic dominance as follows:

Definition 3 *Let $F_X(\cdot)$ and $F_Y(\cdot)$ be the distribution functions of random variables X and Y, respectively. X second-order stochastically dominates Y, denoted as $X \succeq_{SSD} Y$, if and only if*

$$\int_{-\infty}^{z} F_X(\tau)d\tau \leq \int_{-\infty}^{z} F_Y(\tau)d\tau,$$

where all $z \in \mathbb{R}$, provided the integrals exist.

Following Shaked and Shanthikumar (2010), we define the convex order of two random variables as follows:

Definition 4 *For two random variables X and Y, X is smaller than Y in the convex order, denoted as $X \preceq_{cx} Y$, if and only if*

$$E[\phi(X)] \leq E[\phi(Y)],$$

where all convex functions $\phi : \mathbb{R} \to \mathbb{R}$, provided the expectations exist.

From Theorem 3.A.1 in Shaked and Shanthikumar (2010), we have the following:

Proposition 4 *Let X and Y be two random variables such that $E(X) = E(Y)$. Then, $X \preceq_{cx} Y$ if and only if $X \succeq_{SSD} Y$; that is,*

$$X \preceq_{cx} Y \Longleftrightarrow X \succeq_{SSD} Y.$$

Theorem 3.A.10 in Shaked and Shanthikumar (2010) states the following:

Proposition 5 *Let X and Y be two nonnegative random variables such that $E(X) = E(Y)$. Then, $X \preceq_{cx} Y$ if and only if $X \succeq_L Y$; that is,*

$$X \preceq_{cx} Y \Longleftrightarrow X \succeq_L Y.$$

Propositions 4 and 5 imply that $X \succeq_L Y$, $X \preceq_{cx} Y$, and $X \succeq_{SSD} Y$ are equivalent if X and Y are two nonnegative random variables with equal means. For two nonnegative random variables X and Y such that $E(X) > 0$, $E(Y) > 0$, and $E(X) \neq E(Y)$, we cannot use Propositions 4 and 5 directly. However, X and $\frac{X}{E(X)}$ have the same Lorenz curve. Thus, $X \succeq_L Y$ is equivalent to $\frac{X}{E(X)} \succeq_L \frac{Y}{E(Y)}$. To compare the Lorenz curves of random variables X and Y, we can investigate $\frac{X}{E(X)}$ and $\frac{Y}{E(Y)}$ since $E\left(\frac{X}{E(X)}\right) = E\left(\frac{Y}{E(Y)}\right) = 1$.

For the two-type distribution, we compare income distribution without information rents and that with information rents. Let

$$X = \begin{cases} \underline{\theta}\underline{q}, & \text{with probability } v \\ \bar{\theta}\bar{q}, & \text{with probabiity } 1-v \end{cases}$$

and

$$Y = \begin{cases} \underline{\theta}\underline{q} + \underline{U}, & \text{with probability } v \\ \bar{\theta}\bar{q}, & \text{with probabiity } 1-v \end{cases}.$$

Theorem 1 *Under Assumption 1, we have $X \succeq_L Y$.*

Proof: see Appendix.

The $\underline{\theta}-$type agents receive information rents under incomplete information. Information rents are transfers beyond the compensation for the production cost. The $\bar{\theta}-$type agents do not receive information rents. We find that information rents increase income inequality for incentive-feasible contracts. Figure 1 shows that distribution X is more dispersed than Y.

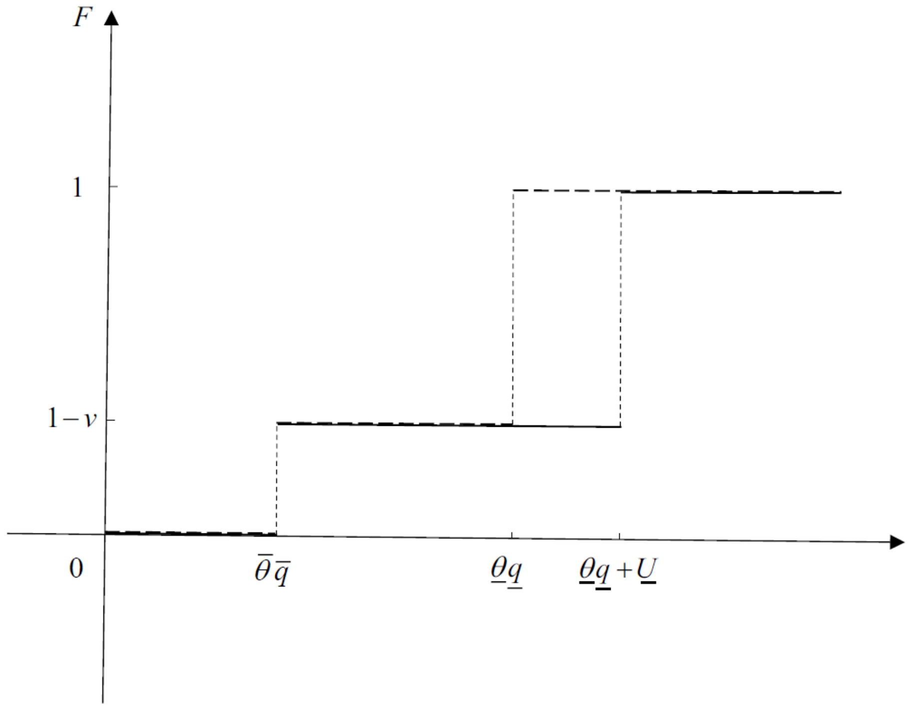

Figure 1: The role of information rent with a discrete distribution.

For two contracts $\{(t(\underline{\theta}), q(\underline{\theta})); (t(\bar{\theta}), q(\bar{\theta}))\}$ and $\{(\tilde{t}(\underline{\theta}), \tilde{q}(\underline{\theta})); (\tilde{t}(\bar{\theta}), \tilde{q}(\bar{\theta}))\}$, we have the following:

Proposition 6 *For the two-type case, $\tilde{q}(\underline{\theta})/\tilde{q}(\bar{\theta}) \geq q(\underline{\theta})/q(\bar{\theta})$ implies that $\tilde{t}(\underline{\theta})/\tilde{t}(\bar{\theta}) \geq t(\underline{\theta})/t(\bar{\theta})$.*

Proof: We have

$$\begin{aligned} t(\underline{\theta})/t(\bar{\theta}) &= [\underline{\theta}q(\underline{\theta}) + \Delta\theta q(\bar{\theta})]/(\bar{\theta}q(\bar{\theta})) \\ &= (\underline{\theta}q(\underline{\theta}))/(\bar{\theta}q(\bar{\theta})) + \Delta\theta/\bar{\theta} \\ &\leq (\underline{\theta}\tilde{q}(\underline{\theta}))/(\bar{\theta}\tilde{q}(\bar{\theta})) + \Delta\theta/\bar{\theta} \\ &= \tilde{t}(\underline{\theta})/\tilde{t}(\bar{\theta}). \end{aligned}$$

■

The role of information rents depend on the output level owing to incentive compatibility constraints. Thus, the output level plays an important role in understanding income equality. The output scheme determines the payment schedule for any feasible contracts. Information rents can be expressed as a function of the output. Therefore, the payment is a function of the output level. We can characterize the payment schedule by investigating the properties of the output scheme. In this sense, the output scheme contains sufficient information to study the inequality implications of any feasible contract. Proposition 6 shows that the contract that induces output levels with higher differences offers incomes with higher differences.

For the continuous type, let

$$X(\theta) = \theta q(\theta),$$

and

$$Y(\theta) = \theta q(\theta) + U(\theta),$$

for θ in $[a, b]$. $X(\Theta)$ represents the cost-compensation distribution, and $Y(\Theta)$ represents income distribution with information rents.

Theorem 2 *Under Assumption 1′, if* $\frac{\theta_2}{\theta_1} \geq \frac{\int_{\theta_2}^{b} \frac{q(\tau)}{q(\theta_2)} d\tau}{\int_{\theta_1}^{b} \frac{q(\tau)}{q(\theta_1)} d\tau}$ *and* $\rho(\theta_1) \geq \rho(\theta_2)$*, for* $a \leq \theta_1 \leq \theta_2 \leq b$*, then we have* $X \succeq_L Y$*.*

Proof: Applying Proposition 2 and Proposition 10 in the Appendix, we obtain the result. ■

Theorem 2 shows that information rents increase income inequality for incentive-feasible contracts under certain conditions. Under the feasible contract, the agents' payments are not in line with the marginal cost. This is different from the neoclassical theory of distribution. With information frictions, the contract creates a wedge between agents' payments and their marginal costs. We find conditions under which we can rank contracts according to their induced income inequality. We rank the income equality of these contracts using Lorenz ordering. The Lorenz curves can be ranked without intersections. Figure 2 shows this result.

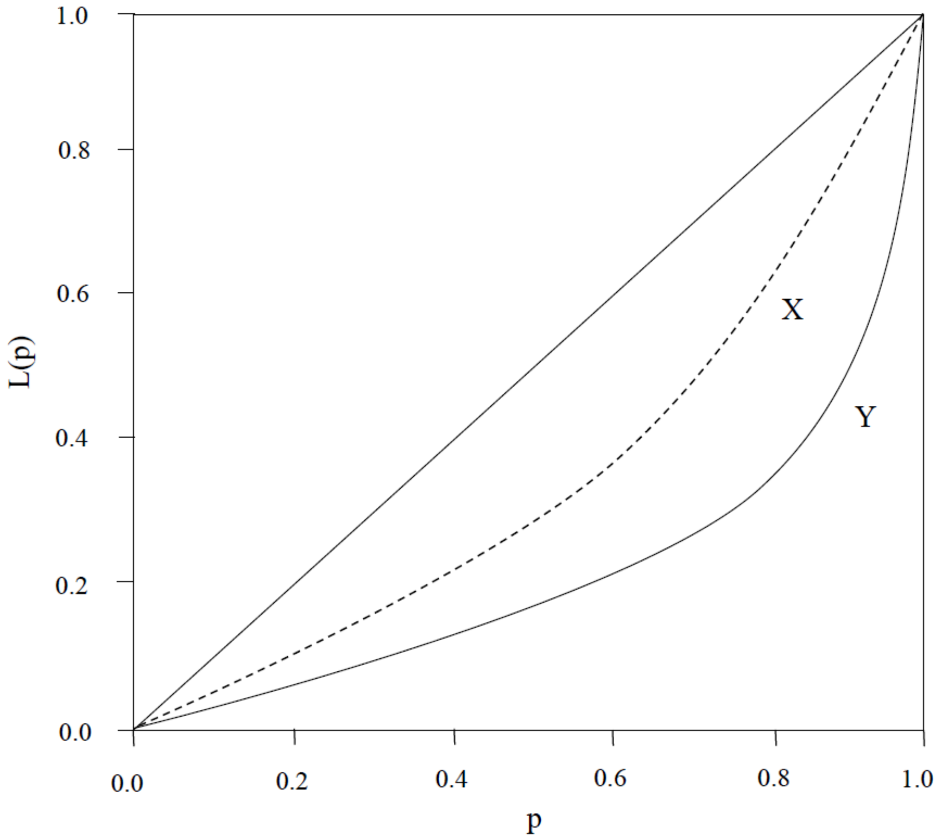

Figure 2: The role of information rent with a continuous distribution.

3 Information frictions

We compare income distribution under incomplete information and under complete information to find the effects of asymmetric information on income distribution.

3.1 Two types

The principal's firm has production function $S(\cdot)$. We have an assumption of $S(\cdot)$.

Assumption 2: The production function satisfies

$$S'(q) > 0,\ S''(q) < 0,\ \lim_{q \to 0} S'(q) = \infty,\ \text{and}\ \lim_{q \to 0} S'(q)q = 0.$$

Assumption 2 implies that the marginal product is infinity at $q = 0$. Thus, there is no shutdown for $\bar{\theta}-$type agents, and q is always greater than 0.

The principal has the objective function

$$\max_{\left\{\left(\underline{t},\underline{q}\right);\left(\bar{t},\bar{q}\right)\right\}} v\left[S\left(\underline{q}\right) - \underline{t}\right] + (1-v)\left[S\left(\bar{q}\right) - \bar{t}\right].$$

With two types, the optimal contract problem is

$$\begin{aligned}
\max_{\left\{\left(\underline{U},\underline{q}\right);\left(\bar{U},\bar{q}\right)\right\}} \quad & v\left[S\left(\underline{q}\right) - \underline{\theta}\underline{q}\right] + (1-v)\left[S\left(\bar{q}\right) - \bar{\theta}\bar{q}\right] - \left[v\underline{U} + (1-v)\bar{U}\right] \\
s.t. \quad & \underline{U} \geq \bar{U} + \Delta\theta\bar{q}, && (3) \\
& \bar{U} \geq \underline{U} - \Delta\theta\underline{q}, && (4) \\
& \underline{U} \geq 0, && (5) \\
& \bar{U} \geq 0. && (6)
\end{aligned}$$

Constraints (3) and (4) are from incentive compatibility constraints (1) and (2). Constraints (5) and (6) are the participation constraints.

We call the optimal contract under incomplete information the second-best contract. Under Assumptions 1 and 2, the optimal contract under asymmetric information satisfies

$$S'(\underline{q}^{SB}) = \underline{\theta}$$

and

$$S'\left(\bar{q}^{SB}\right) = \bar{\theta} + \frac{v}{1-v}\Delta\theta,$$

and the agents' incomes under the second-best contract are

$$\underline{t}^{SB} = \underline{\theta}\underline{q}^{SB} + \Delta\theta\bar{q}^{SB}$$

and

$$\bar{t}^{SB} = \bar{\theta}\bar{q}^{SB}.$$

The optimal contract under asymmetric information is influenced by the information structure and the production function. The output level and transfers are determined by the optimal contract. To extract rents from the agents, principals distort the output level of inefficient agents. To implement $\underline{q}^{SB}$ and $\bar{q}^{SB}$, principals offer transfers $\underline{t}^{SB}$ and $\bar{t}^{SB}$. These transfers determine income distribution in the economy under the second-best contract.

The contract under complete information only has to satisfy the participation constraints and does not have to obey the incentive compatibility constraints. The optimal contract problem under complete information is

$$\begin{aligned} \max_{\{(\bar{U},\bar{q});(\underline{U},\underline{q})\}} \quad & v\left[S\left(\underline{q}\right) - \underline{\theta}\underline{q}\right] + (1-v)\left[S(\bar{q}) - \bar{\theta}\bar{q}\right] - \left[v\underline{U} + (1-v)\bar{U}\right] \\ s.t. \quad & \underline{U} \geq 0, \\ & \bar{U} \geq 0. \end{aligned}$$

We call the optimal contract under complete information the first-best contract. Thus, we have

$$S'\left(\underline{q}^{FB}\right) = \underline{\theta}$$

and

$$S'\left(\bar{q}^{FB}\right) = \bar{\theta},$$

and the agents' incomes under the first-best contract are

$$\underline{t}^{FB} = \underline{\theta}\underline{q}^{FB}$$

and

$$\bar{t}^{FB} = \bar{\theta}\bar{q}^{FB}.$$

To compare income distribution under the firs-best contract and under the second-best contract, we start from an observation.

Lemma 2 *Under Assumptions 1 and 2, we have*

$$\underline{q}^{FB}/\bar{q}^{FB} \leq \underline{q}^{SB}/\bar{q}^{SB}.$$

The second-best contract incurs less equal output distribution than the first-best contract. The $\underline{\theta}-$type agents under the second-best contract have the same output level as those under the first-best contract, $\underline{q}^{SB} = \underline{q}^{FB}$. Owing to the incentive compatibility constraints, the $\bar{\theta}-$type agents under the second-best contract have a lower output level than those under the first-best contract, $\bar{q}^{SB} \leq \bar{q}^{FB}$. Incomplete information causes a downward output distortion for the $\bar{\theta}-$type agents. This distortion induces an efficiency loss and a less equal output distribution. Therefore, we have $\underline{q}^{FB}/\bar{q}^{FB} \leq \underline{q}^{SB}/\bar{q}^{SB}$.

Proposition 7 *Under Assumptions 1 and 2, we have*

$$\underline{t}^{FB}/\bar{t}^{FB} \leq \underline{t}^{SB}/\bar{t}^{SB}.$$

Proof: We have

$$\begin{aligned}
\underline{t}^{FB}/\bar{t}^{FB} &= (\underline{\theta}\underline{q}^{FB})/(\bar{\theta}\bar{q}^{FB}) \\
&\leq (\underline{\theta}\underline{q}^{SB})/(\bar{\theta}\bar{q}^{SB}) \qquad \text{(output distortion)} \\
&\leq (\underline{\theta}\underline{q}^{SB} + \underline{U}^{SB})/(\bar{\theta}\bar{q}^{SB}) \quad \text{(information rent)} \\
&= \underline{t}^{SB}/\bar{t}^{SB},
\end{aligned}$$

where the first inequality is attributable to the output distortion, and the second inequality comes from information rent. ■

The proof of Proposition 7 employs a decomposition technique, through which we separate two channels of output distortion and information rent. Under complete information, agents produce output levels that are determined by the marginal cost of production, and they receive compensation for the production cost. They do not receive information rents. Under incomplete information, the $\underline{\theta}-$type agents produce output levels according to the marginal cost of production while the $\bar{\theta}-$type agents suffer from a downward output distortion. A less equal output distribution is the first channel through which information frictions cause income inequality. Information rents increase income inequality. This is the second channel through which information frictions cause income inequality.

Let W^{SB} be income distribution in the economy under the second-best contract,

$$W^{SB} = \begin{cases} \underline{t}^{SB}, & \text{with probability } v \\ \bar{t}^{SB}, & \text{with probabiity } 1 - v \end{cases}.$$

Let W^{FB} be income distribution in the economy with complete information,

$$W^{FB} = \begin{cases} \underline{t}^{FB}, & \text{with probability } v \\ \bar{t}^{FB}, & \text{with probabiity } 1 - v \end{cases}.$$

The fact that $\underline{t}^{FB}/\bar{t}^{FB} \leq \underline{t}^{SB}/\bar{t}^{SB}$ implies that asymmetric information causes a larger relative difference between the income of the $\underline{\theta}-$type agents and that of the $\bar{\theta}-$type agents.

Theorem 3 *Under Assumptions 1 and 2, we have* $W^{FB} \succeq_L W^{SB}$.

Proof: see Appendix.

Income distribution under incomplete information is less equal than that under complete information. Output distortion owing to asymmetric information causes income inequality, and information rent further exaggerates income inequality.

3.2 Continuous types

With the continuous type, using the rent variable $U(\theta) = t(\theta) - \theta q(\theta)$, the optimization problem of the principal becomes

$$\begin{aligned}
\max_{\{U(\cdot),q(\cdot))\}} \quad & \int_a^b \left[S\left(q\left(\theta\right)\right) - \theta q\left(\theta\right) - U\left(\theta\right) \right] f\left(\theta\right) d\theta \\
s.t. \quad & \dot{U}\left(\theta\right) = -q\left(\theta\right) \qquad a.e. & (7) \\
& \dot{q}\left(\theta\right) \leq 0 \qquad a.e. & (8) \\
& U\left(\theta\right) \geq 0, & (9)
\end{aligned}$$

where (7) is the incentive compatibility constraint, (8) is the implementability condition, and (9) is the participation constraint.

We assume the following:

Assumption 2′: The production function has the form

$$S(q) = \frac{1}{\gamma} q^{\gamma}, \ \gamma \in (0,1).$$

Under Assumptions 1′ and 2′, the optimal contract under asymmetric information satisfies

$$S'\left(q^{SB}(\theta)\right) = \theta + \frac{F(\theta)}{f(\theta)}.$$

Information rent is

$$U^{SB}\left(\theta\right) = \int_\theta^b q^{SB}\left(\tau\right) d\tau.$$

The agent's income under the second-best contract is

$$t^{SB}\left(\theta\right) = \theta q^{SB}\left(\theta\right) + U^{SB}\left(\theta\right).$$

Under complete information, the principal's problem is

$$\begin{aligned}
\max_{\{(U(\cdot),q(\cdot))\}} \quad & \int_a^b [S(q(\theta)) - \theta q(\theta) - U(\theta)] f(\theta) d\theta \\
s.t. \quad & U\left(\theta\right) \geq 0.
\end{aligned}$$

The first-best output is determined by

$$S'\left(q^{FB}(\theta)\right) = \theta.$$

The agent's income under the first-best contract is

$$t^{FB}(\theta) = \theta q^{FB}(\theta).$$

Following Veres-Ferrer and Pavía (2022), we define the elasticity of a random variable as follows:

Definition 5 *The elasticity of a random variable* Θ *is defined as*

$$e(\theta) = \frac{\theta f(\theta)}{F(\theta)},$$

for θ *in* $[a, b]$*, where* $f(\theta)$ *is the probability density function of* Θ*, and* $F(\theta)$ *is its cumulative distribution function.*

Since $F(\theta) = e^{-\int_\theta^b \frac{e(\tau)}{\tau} d\tau}$, for θ in $[a, b]$, we know that the elasticity of a random variable determines its cumulative distribution function. We assume the following:

Assumption 3: $e(\theta)$ is decreasing in θ.

Assumption 3 implies the monotone hazard rate property, $\frac{f(\theta_1)}{F(\theta_1)} \geq \frac{f(\theta_2)}{F(\theta_2)}$, for $a \leq \theta_1 \leq \theta_2 \leq b$.

We have

$$q^{SB}(\theta) = \left(\theta + \frac{F(\theta)}{f(\theta)}\right)^{\frac{1}{\gamma-1}}.$$

for θ in $[a, b]$, and

$$q^{FB}(\theta) = \theta^{\frac{1}{\gamma-1}}.$$

Lemma 3 *Under Assumptions 1′, 2′, and 3, we have*

$$\frac{q^{SB}(\theta_1)}{q^{FB}(\theta_1)} \geq \frac{q^{SB}(\theta_2)}{q^{FB}(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$*.*

Proof: We know from Assumption 3 that $e(\theta_1) \geq e(\theta_2)$, for $a \leq \theta_1 \leq \theta_2 \leq b$. Since

$$\frac{q^{SB}(\theta)}{q^{FB}(\theta)} = \frac{\left(\theta + \frac{F(\theta)}{f(\theta)}\right)^{\frac{1}{\gamma-1}}}{\theta^{\frac{1}{\gamma-1}}} = \left(1 + \frac{1}{e(\theta)}\right)^{\frac{1}{\gamma-1}},$$

for θ in $[a, b]$, and $\frac{1}{\gamma-1} < 0$, we have $\frac{q^{SB}(\theta_1)}{q^{FB}(\theta_1)} \geq \frac{q^{SB}(\theta_2)}{q^{FB}(\theta_2)}$, for $a \leq \theta_1 \leq \theta_2 \leq b$.
■

Lemma 3 implies that

$$q^{SB}(\theta_1)/q^{SB}(\theta_2) \geq q^{FB}(\theta_1)/q^{FB}(\theta_2),$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. The output difference under the second-best contract is larger than that under the first-best contract.

Ratio $\frac{q^{SB}(\theta)}{q^{FB}(\theta)}$ also plays an important role for investigating ratio $\frac{t^{SB}(\theta)}{t^{FB}(\theta)}$. Since $t^{FB}(\theta) = \theta q^{FB}(\theta)$ and $\rho^{SB}(\theta) = \theta q^{SB}(\theta)$, for θ in $[a, b]$, we have $\frac{\rho^{SB}(\theta)}{t^{FB}(\theta)} = \frac{q^{SB}(\theta)}{q^{FB}(\theta)}$. Furthermore, we have

$$\frac{t^{SB}(\theta)}{t^{FB}(\theta)} = \frac{\rho^{SB}(\theta)}{t^{FB}(\theta)} \frac{t^{SB}(\theta)}{\rho^{SB}(\theta)} = \frac{q^{SB}(\theta)}{q^{FB}(\theta)} \frac{t^{SB}(\theta)}{\rho^{SB}(\theta)},$$

for θ in $[a, b]$. Through this relationship we can employ a decomposition technique, which separates two channels of output distortion and information rent. Both channels contribute to the income inequality under incomplete information.

Proposition 8 *Under Assumptions 1′, 2′, and 3, if*

$$\left(1 + \frac{1}{e(\theta)}\right) \theta^{\gamma} (b - \theta)^{1-\gamma}$$

is decreasing in θ, we have

$$\frac{t^{SB}(\theta_1)}{t^{FB}(\theta_1)} \geq \frac{t^{SB}(\theta_2)}{t^{FB}(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$.

Proof: If $\left(1 + \frac{1}{e(\theta)}\right) \theta^{\gamma} (b - \theta)^{1-\gamma}$ is decreasing in θ, we have

$$\left(1 + \frac{1}{e(\theta_1)}\right) \theta_1^{\gamma} (b - \theta_1)^{1-\gamma} \geq \left(1 + \frac{1}{e(\theta_2)}\right) \theta_2^{\gamma} (b - \theta_2)^{1-\gamma},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. Thus, we have

$$\frac{\theta_1\left(1+\frac{1}{e(\theta_1)}\right)}{\theta_2\left(1+\frac{1}{e(\theta_2)}\right)} \geq \left(\frac{\theta_2\,(b-\theta_1)}{\theta_1\,(b-\theta_2)}\right)^{\gamma-1}.$$

Therefore, we have

$$\frac{\theta_1^{\frac{1}{\gamma-1}}\left(1+\frac{1}{e(\theta_1)}\right)^{\frac{1}{\gamma-1}}}{\theta_2^{\frac{1}{\gamma-1}}\left(1+\frac{1}{e(\theta_2)}\right)^{\frac{1}{\gamma-1}}} \leq \frac{\theta_2\,(b-\theta_1)}{\theta_1\,(b-\theta_2)}.$$

Since $q^{SB}\left(\theta\right) = \theta^{\frac{1}{\gamma-1}}\left(1+\frac{F(\theta)}{\theta f(\theta)}\right)^{\frac{1}{\gamma-1}} = \theta^{\frac{1}{\gamma-1}}\left(1+\frac{1}{e(\theta)}\right)^{\frac{1}{\gamma-1}}$, we have

$$\frac{q^{SB}\left(\theta_1\right)}{q^{SB}\left(\theta_2\right)} \leq \frac{\theta_2\,(b-\theta_1)}{\theta_1\,(b-\theta_2)},$$

and

$$\frac{\theta_1 q^{SB}\left(\theta_1\right)}{b-\theta_1} \leq \frac{\theta_2 q^{SB}\left(\theta_2\right)}{b-\theta_2}.$$

From Proposition 3 we know that

$$\frac{t^{SB}(\theta_1)}{\rho^{SB}(\theta_1)} \geq \frac{t^{SB}(\theta_2)}{\rho^{SB}(\theta_2)}.$$

for $a \leq \theta_1 \leq \theta_2 \leq b$.

We know from Lemma 3 that $\frac{q^{SB}(\theta_1)}{q^{FB}(\theta_1)} \geq \frac{q^{SB}(\theta_2)}{q^{FB}(\theta_2)}$, for $a \leq \theta_1 \leq \theta_2 \leq b$. Thus, we know that

$$\frac{t^{SB}(\theta_1)}{t^{FB}(\theta_1)} \geq \frac{t^{SB}(\theta_2)}{t^{FB}(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, since $\frac{t^{SB}(\theta)}{t^{FB}(\theta)} = \frac{\rho^{SB}(\theta)}{t^{FB}(\theta)}\frac{t^{SB}(\theta)}{\rho^{SB}(\theta)} = \frac{q^{SB}(\theta)}{q^{FB}(\theta)}\frac{t^{SB}(\theta)}{\rho^{SB}(\theta)}$, for θ in $[a, b]$. ■

The result of Proposition 8 is equivalent to

$$t^{SB}(\theta_1)/t^{SB}(\theta_2) \geq t^{FB}(\theta_1)/t^{FB}(\theta_2),$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. The income difference under the second-best contract is larger than that under the first-best contract.

Let $t^{SB}(\Theta)$ be income distribution under the second-best contract and $t^{FB}(\Theta)$ be income distribution under the first-best contract.

Theorem 4 *Under Assumptions 1′, 2′, and 3, if* $\left(1+\frac{1}{e(\theta)}\right)\theta^{\gamma}(b-\theta)^{1-\gamma}$ *is decreasing in* θ*, then we have* $t^{FB}(\Theta) \succeq_L t^{SB}(\Theta)$*.*

Proof: Applying Proposition 8 and Proposition 10 in the Appendix, we obtain the result. ■

Theorem 4 implies that income distribution under the second-best contract is less equal than that under the first-best contract.

Table 1: Examples of distributions

Distribution, $F(\theta)$	**Elasticity**, $e(\theta)$	**Output**, $q^{SB}(\theta)$	**Condition used**
Formula of Equation (10)	See Appendix		$\left(1+\frac{1}{e(\theta)}\right)\theta^{\gamma}(b-\theta)^{1-\gamma}\downarrow$
$\theta^{\kappa},\theta\in[0,1]$	κ	$\left(\frac{1+\kappa}{\kappa}\right)^{\frac{1}{\gamma-1}}\theta^{\frac{1}{\gamma-1}}$	$\frac{U(\theta)}{\theta q(\theta)}\downarrow$
$\frac{\theta-a}{b-a},\theta\in[a,b],\frac{b}{a}\leq\frac{3-\gamma}{2}$	$\frac{\theta}{\theta-a}$	$(2\theta-a)^{\frac{1}{\gamma-1}}$	$\frac{\theta q(\theta)}{b-\theta}\uparrow$
$\frac{\log(\theta)-\log(a)}{\log(b)-\log(a)},\theta\in[a,b],\frac{b}{a}\leq\sqrt{2-\gamma}$	$\frac{1}{\log(\theta)-\log(a)}$	$\theta^{\frac{1}{\gamma-1}}(\log(\theta)-\log(a)+1)^{\frac{1}{\gamma-1}}$	$\frac{\theta q(\theta)}{b-\theta}\uparrow$
$\sqrt{\frac{\log(\theta)}{\log(b)}},\theta\in[1,b],b\leq\sqrt{\frac{3-\gamma}{2}}$	$\frac{1}{2\log(\theta)}$	$\theta^{\frac{1}{\gamma-1}}(2\log(\theta)+1)^{\frac{1}{\gamma-1}}$	$\frac{\theta q(\theta)}{b-\theta}\uparrow$
$\frac{\theta\log(\theta)}{b\log(b)},\theta\in[1,b],b\leq\sqrt{2-\gamma}$	$\frac{\log(\theta)+1}{\log(\theta)}$	$\theta^{\frac{1}{\gamma-1}}\left(\frac{2\log(\theta)+1}{\log(\theta)+1}\right)^{\frac{1}{\gamma-1}}$	$\frac{\theta q(\theta)}{b-\theta}\uparrow$

All examples in Table 1 satisfy Assumption 3. For these examples, the income distributions under the second-best contract are less equal than those under the first-best contract. However, with respect to different distributions, we use different conditions to draw result $\frac{t^{SB}(\theta_1)}{\rho^{SB}(\theta_1)} \geq \frac{t^{SB}(\theta_2)}{\rho^{SB}(\theta_2)}$, for $a \leq \theta_1 \leq \theta_2 \leq b$. Verification of these examples is in Section 7.3 of Appendix. Using the procedure in Section 4 we could find more examples by investigating the transformation of the distributions in Table 1.

The uniform distribution on $[0,1]$ is a special case of the distribution with a constant elaticity $\kappa = 1$. We investigate the income distribution under the second-best contract if the type distribution has a constant elasticity κ. The income distribution under the optimal contract has an asymptotical

Pareto tail with exponent $\chi = \kappa\left(\frac{1}{\gamma} - 1\right)$,

$$\lim_{x\to\infty} \frac{\Pr\left(t^{SB}(\Theta) > x\right)}{x^{-\chi}} = \gamma^{-(1-\gamma)\frac{\kappa}{\gamma}} \left(1 + \frac{1}{\kappa}\right)^{-\frac{\kappa}{\gamma}}.$$

The Pareto exponent χ represents the fatness of the tail. The income distribution with a lower the Pareto exponent is more dispersed. The lower κ, the fatter the tail of the income distribution. The higher γ, the fatter the tail of the income distribution.

We investigate income distribution when the agents' payments are not in line with the marginal cost. In the neoclassical theory of distribution, factors receive payments according to their marginal product. Some studies use neoclassical theory to explain the observed income distribution. Sattinger (1975) investigates how comparative advantage connects ability distribution and income distribution in the Roy model. Heckman and Honoré (1990) extend the classical Roy model. Gabaix and Landier (2008) and Terviö (2008, 2009) use the sorting mechanism in assignment models to investigate income distribution. However, they find that the sorting mechanism itself is not enough to generate the fat tail of income distribution. Geerolf (2017) uses an assignment model with complementarities to generate a Pareto tail of income distribution.

We then consider the type distribuion

$$F(\theta) = \begin{cases} h(\theta - a) + \frac{2(1-h(b-a))}{(b-a)^2}(\theta - a)^2, & \theta \in \left[a, \frac{a+b}{2}\right]; \\ 1 - h(b-\theta) - \frac{2(1-h(b-a))}{(b-a)^2}(b-\theta)^2, & \theta \in \left[\frac{a+b}{2}, b\right]. \end{cases} \tag{10}$$

Figure 3 plots the density function of this example. The density function of type distribution is non-differentiable at some θ. Thus, output function $q(\theta)$ is non-differentiable at that point. We find this example to show that our results hold for the case in which output function $q(\theta)$ is non-differentiable at some θ.

With $a = 1$, $b = 1.2$, and $h = 4$, if $\gamma \leq 0.368$, then $\left(1 + \frac{1}{e(\theta)}\right)\theta^{\gamma}(b-\theta)^{1-\gamma}$ is decreasing in θ. We have

$$\frac{t^{SB}(\theta_1)}{t^{FB}(\theta_1)} \geq \frac{t^{SB}(\theta_2)}{t^{FB}(\theta_2)},$$

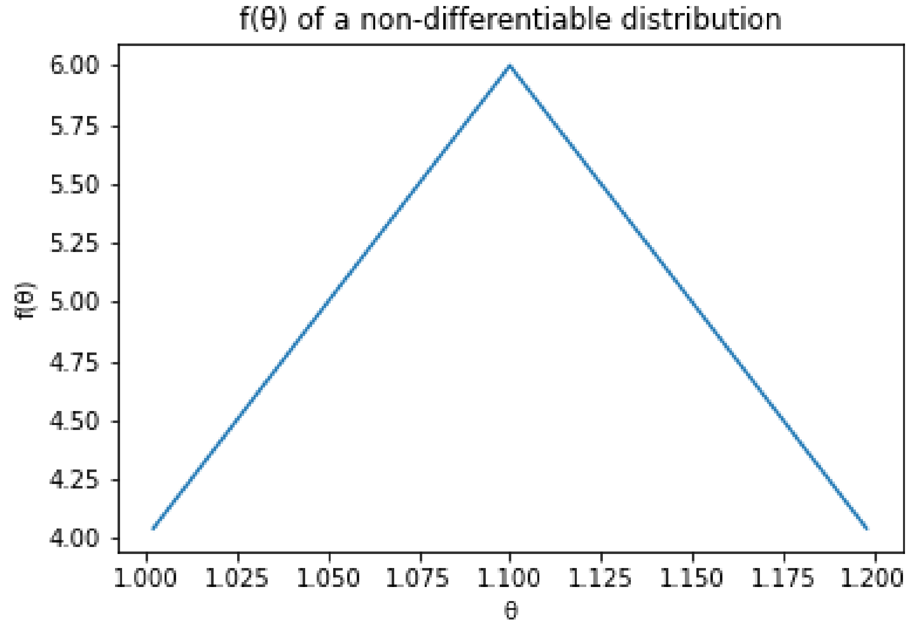

Figure 3: A non-differentiable density function.

for $a \leq \theta_1 \leq \theta_2 \leq b$.

4 Transformation of the type distribution

We have verified examples in Table 1. In this section, we find more examples by the transformation of the type distribution. If the type distribution changes, the optimal contract changes accordingly. However, we show that the condition in Proposition 8 still holds under certain transformation.

Let $\tilde{\theta} = F^{-1}\left(F(\theta)^{\frac{1}{\beta}}\right)$, for $\beta > 0$ be the transformation of the type. We have $G(\theta) = F(\theta)^{\beta}$, for θ in $[a, b]$. We know that $G(\theta)$ satisfies Assumption 3 if $F(\theta)$ satisfies it, since $e_G(\theta) = \beta e_F(\theta)$, for θ in $[a, b]$.

If $F(\theta)$ satisfies the condition in Proposition 8, we have

$$\left(1 + \frac{1}{e_F(\theta_1)}\right)\theta_1^{\gamma}(b - \theta_1)^{1-\gamma} \geq \left(1 + \frac{1}{e_F(\theta_2)}\right)\theta_2^{\gamma}(b - \theta_2)^{1-\gamma}, \qquad (11)$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, which implies that

$$\begin{aligned} & \theta_1^\gamma (b-\theta_1)^{1-\gamma} - \theta_2^\gamma (b-\theta_2)^{1-\gamma} \\ \geq & \tfrac{1}{e_F(\theta_2)} \theta_2^\gamma (b-\theta_2)^{1-\gamma} - \tfrac{1}{e_F(\theta_1)} \theta_1^\gamma (b-\theta_1)^{1-\gamma} \quad . \end{aligned} \tag{12}$$

We know that $\left(1+\frac{1}{e_F(\theta_1)}\right) \leq \left(1+\frac{1}{e_F(\theta_2)}\right)$, since we have $e_F(\theta_1) \geq e_F(\theta_2$ from Assumption 3. Thus, we have

$$\theta_1^\gamma (b-\theta_1)^{1-\gamma} \geq \theta_2^\gamma (b-\theta_2)^{1-\gamma},$$

from relationship (11). Thus, we have

$$\beta \left[\theta_1^\gamma (b-\theta_1)^{1-\gamma} - \theta_2^\gamma (b-\theta_2)^{1-\gamma}\right] \geq \theta_1^\gamma (b-\theta_1)^{1-\gamma} - \theta_2^\gamma (b-\theta_2)^{1-\gamma}, \tag{13}$$

for $\beta \geq 1$.

Combing relationships (12) and (13), we have

$$\left(1+\frac{1}{\beta e_F(\theta_1)}\right) \theta_1^\gamma (b-\theta_1)^{1-\gamma} \geq \left(1+\frac{1}{\beta e_F(\theta_2)}\right) \theta_2^\gamma (b-\theta_2)^{1-\gamma},$$

which implies that

$$\left(1+\frac{1}{e_G(\theta_1)}\right) \theta_1^\gamma (b-\theta_1)^{1-\gamma} \geq \left(1+\frac{1}{e_G(\theta_2)}\right) \theta_2^\gamma (b-\theta_2)^{1-\gamma}, \tag{14}$$

since $e_G(\theta) = \beta e_F(\theta)$, for θ in $[a, b]$.

For $\beta \geq 1$, if $G(\theta) = F(\theta)^\beta$ and $F(\theta)$ satisfies the condition in Proposition 8, then $G(\theta)$ also satisfies the condition in Proposition 8. Thus, each examples in Table 1 represents a class of distributions for which the income distributions under the second-best contract are less equal than those under the first-best contract.

According to the definition in Laffont and Tirole (1993), distribution $F(\theta)$ on $[a, b]$ is more favorable than distribution $G(\theta)$ on the same interval if $G(\theta) \leq F(\theta)$ for all θ and $\frac{f(\theta)}{F(\theta)} \leq \frac{g(\theta)}{G(\theta)}$ for θ in $[a, b]$. Under our transformation, we know that $F(\theta)$ is more favorable than $G(\theta) = F(\theta)^\beta$ for $\beta \geq 1$.

5 Social norms

Piketty and Saez (2003) propose that changing social norms regarding inequality have played an important role in rising inequality in the United States since the 1970s. Thus, we introduce a parameter of norms regarding social welfare into our benchmark model and investigate the effect of social-norm change on income distribution.

The principal maximizes the weighted average of its surplus and of the agent's rent U with $\alpha \in [0,1]$ as the weight for the agent's rent. Here, α represents the social norm:

$$\begin{aligned}
\max_{\{(\underline{U},\underline{q});(\bar{U},\bar{q})\}} \quad & v\left[S\left(\underline{q}\right) - \underline{\theta}\underline{q}\right] + (1-v)\left[S\left(\bar{q}\right) - \bar{\theta}\bar{q}\right] - (1-\alpha)\left[v\underline{U} + (1-v)\bar{U}\right] \\
s.t. \quad & \underline{U} \geq \bar{U} + \Delta\theta\bar{q}, \\
& \bar{U} \geq \underline{U} - \Delta\theta\underline{q}, \\
& \underline{U} \geq 0, \\
& \bar{U} \geq 0.
\end{aligned}$$

The outputs are

$$S'\left(\underline{q}^{\alpha}\right) = \underline{\theta},$$

and

$$S'\left(\bar{q}^{\alpha}\right) = \bar{\theta} + \frac{v}{1-v}(1-\alpha)\Delta\theta.$$

The agents' incomes under the second-best contract are now

$$\underline{t}^{\alpha} = \underline{\theta}\underline{q}^{\alpha} + \Delta\theta\bar{q}^{\alpha},$$

and

$$\bar{t}^{\alpha} = \bar{\theta}\bar{q}^{\alpha}.$$

For the two-type case, if $\alpha_1 \geq \alpha_2$, we have $\bar{q}^{\alpha_2} \leq \bar{q}^{\alpha_1}$. Thus, we know that $\underline{q}^{\alpha_2}/\bar{q}^{\alpha_2} \geq \underline{q}^{\alpha_1}/\bar{q}^{\alpha_1}$. From Proposition 6, we have $\underline{t}^{\alpha_2}/\bar{t}^{\alpha_2} \geq \underline{t}^{\alpha_1}/\bar{t}^{\alpha_1}$. The income difference under α_2 is larger than that under α_1. If the principal places lower weight on the agent's rent, the optimal contract causes less equal income distribution. When the principal increases weight α, both efficiency and equity increase. Then, we find that the optimal contract in our benchmark model, corresponding to $\alpha = 0$, is neither efficient nor equal. The optimal contract under a trade-off between rent extraction and efficiency causes less efficient output levels and less equal income distribution.

For the continuous type, the principal's problem is now

$$\begin{aligned}
\max_{\{(U(\cdot),q(\cdot)\}} \quad & \int_a^b \left[S\left(q\left(\theta\right)\right) - \theta q\left(\theta\right) - (1-\alpha)U(\theta)\right] f(\theta) d\theta, \\
s.t. \quad & \dot{U}\left(\theta\right) = -q\left(\theta\right) \qquad a.e., \\
& \dot{q}(\theta) \leq 0 \qquad a.e., \\
& U\left(\theta\right) \geq 0.
\end{aligned}$$

The output is

$$q^{\alpha}(\theta) = \left(\theta \frac{1-\alpha+e(\theta)}{e(\theta)}\right)^{\frac{1}{\gamma-1}},$$

for θ in $[a, b]$.

Lemma 4 *Under Assumptions 1′, 2′, and 3, we have*

$$\frac{q^{\alpha_2}(\theta_1)}{q^{\alpha_1}(\theta_1)} \geq \frac{q^{\alpha_2}(\theta_2)}{q^{\alpha_1}(\theta_2)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, *if* $\alpha_1 \geq \alpha_2$.

Proof: We have

$$\frac{q^{\alpha_2}(\theta)}{q^{\alpha_1}(\theta)} = \left(\frac{1-\alpha_2+e(\theta)}{1-\alpha_1+e(\theta)}\right)^{\frac{1}{\gamma-1}} = \left(1+\frac{\alpha_1-\alpha_2}{1-\alpha_1+e(\theta)}\right)^{\frac{1}{\gamma-1}},$$

for θ in $[a, b]$. If $\alpha_1 \geq \alpha_2$, we know that $\frac{q^{\alpha_2}(\theta_1)}{q^{\alpha_1}(\theta_1)} \geq \frac{q^{\alpha_2}(\theta_2)}{q^{\alpha_1}(\theta_2)}$, for $a \leq \theta_1 \leq \theta_2 \leq b$, since $e(\theta_1) \geq e(\theta_2)$ from Assumption 3. ■

For uniform distribution on $[a, b]$, we have

$$\frac{t^{\alpha_2}\left(\theta_1\right)}{t^{\alpha_1}\left(\theta_1\right)} \geq \frac{t^{\alpha_2}\left(\theta_2\right)}{t^{\alpha_1}\left(\theta_2\right)}$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, if $\alpha_1 \geq \alpha_2$. Thus, we have $t^{\alpha_2}(\theta_1)/t^{\alpha_2}(\theta_2) \geq t^{\alpha_1}(\theta_1)/t^{\alpha_1}(\theta_2)$ for $a \leq \theta_1 \leq \theta_2 \leq b$. The income difference under α_2 is larger than that under α_1. The derivation is in Section 7.4 of Appendix.

Applying Proposition 10 in the Appendix, we have $t^{\alpha_1}(\Theta) \succeq_L t^{\alpha_2}(\Theta)$, if $\alpha_1 \geq \alpha_2$. Income distribution under α_2 is less equal than that under α_1. The marginal cost of production and the type distribution do not change. The principal offers a contract that displays a less equal wage profile, if $\alpha_1 \geq \alpha_2$. As in Costinot and Vogel (2010), changes in the wage profile reflect changes in the return to skill. If the principal places lower weight on the agent's rent, the optimal contract causes less equal income distribution. The relative income difference between the efficient- and inefficient-type agents becomes larger.

6 Conclusion

This study investigates how information frictions influence income distributions when there is a trade-off between rent extraction and efficiency. We examine the adverse selection problem in this study and show that information rent increases income inequality. We rank the income equality of these contracts using Lorenz ordering. The Lorenz curves can be ranked without intersections.

We find that the output scheme determines the payment schedule for any feasible contracts. In this sense, the output scheme contains sufficient information to study the inequality implications of any feasible contract. For the continuous type, the elasticity of the type distribution determines the payment schedule. We also find that changes in social norms influence inequality under the optimal contract.

References

[1] Baron, David and Roger Myerson (1982): ”Rgulating a monopolist with unknow costs,” *Econometrica*, 50, 911-930.

[2] Campbell, Steven, Yichao Chen, Arvind Shrivates, and Sebastian Jaimungal (2021): ”Deep learning for principal-agent mean field games,” mimeo, University of Toronto.

[3] Costinot, Arnaud and Jonathan Vogel (2010): ”Matching and inequality in the world economy,” *Journal of Political Economy*, 118, 747-786.

[4] Dworczak, Piotr, Scott Kominers, and Mohammad Akbarpour (2021): ”Redistribution through markets,” *Econometrica*, 89, 1665-1698.

[5] Ekeland, Ivar (2010): ”Notes on optimal transportation,” *Economic Theory*, 42, 437-459.

[6] Fellman J. (1976): ”The effct of transformations on Lorenz curves,” *Econometrica*, 44, 823-824.

[7] Fernandez, Raquel and Jordi Gali (1999): ”To each according to...? Markets, tournaments, and the matching problem with borrowing constraints,” *Review of Economic Studies*, 66, 799-824.

[8] Gastwirth, Joseph (1971): ”A general definition of the Lorenz curve,” *Econometrica*, 39, 1037-1039.

[9] Geerolf, Francois (2017): ”A theory of Pareto distributions,” mimeo, Sciences Po.

[10] Green, Jerry and Nancy Stokey (1983): ”A comparison of tournaments and contracts,” *Journal of Political Economy*, 91, 349-364.

[11] Heckman, James and Bo Honoré (1990): ”The empirical content of the Roy model,” *Econometrica*, 58, 1121-1149.

[12] Laffont, Jean-Jacques and David Martimort (2002): *The Theory of Incentives: The Principal-agent Model*, Princeton University Press, Princeton, NJ.

[13] Laffont, Jean-Jacques and Jean Tirole (1993): *The Theory of Incentives in Procurement and Regulation*, The MIT Press, Cambridge, MA.

[14] Lazear, Edward and Sherwin Rosen (1981): "Rank-order tournaments as optimal labor contracts," *Journal of Political Economy*, 89, 841-864.

[15] Mirrlees, James (1971): "An exploration in the theory of optimum income taxation," *Review of Economic Studies*, 38, 175-208.

[16] Myerson, Roger (1981): "Optimal auction design," *Mathematics of Operations Research*, 6, 58-73.

[17] Nalebuff, Barry and Joseph Stiglitz (1983): "Prizes and incentives: Towards a general theory of compensation and competition," *Bell Journal of Economics*, 14, 21-43.

[18] Ok, Efe (2023): *Measure and Probability Theory with Economic Applications*, mimeo, New York University.

[19] Piketty, Thomas (2020): *Capital and Ideology*, Harvard University Press, Cambridge, MA.

[20] Piketty, Thomas and Emmanuel Saez (2003): "Income inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118, 1-39.

[21] Saez, Emmanuel (2001): "Using elasticities to derive optimal income tax rates," *Review of Economic Studies*, 68, 205-229.

[22] Saez, Emmanuel and Gabriel Zucman (2016): "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data," *Quarterly Journal of Economics*, 131, 519-578.

[23] Sattinger, Michael (1975): "Comparative advantage and the distributions of earnings and abilities," *Econometrica*, 43, 455-468.

[24] Shaked, Moshe and George Shanthikumar (2010): *Stochastic Orders*, Springer, New York, NY.

[25] Terviö, Marko (2008): "The difference that CEOs make: An assignment model approach," *American Economic Review*, 98, 642-668.

[26] Terviö, Marko (2009): "Superstars and mediocrities: Market failure in the discovery of talent," *Review of Economic Studies*, 76, 829-850.

[27] Teulings, Coen (1995): "The wage distribution in a model of the assignment of skills to jobs," *Journal of Political Economy*, 103, 280-315.

[28] Veres-Ferrer, Ernesto-Jesús and Jose M. Pavía (2022): "The elasticity of a random variable as a tool for measuring and assessing risks," *Risks*, 10, 68.

[29] Zhu, Shenghao (2013): "Comparisons of stationary distributions of linear models," *Economics Letters*, 119, 221-223.

7 Appendix

7.1 Inequality tools

The coefficient of variation (CV) of a random variable X is defined as

$$CV(X) = \sqrt{\frac{E\left(X - EX\right)^2}{\left(EX\right)^2}}.$$

From Proposition 5, we know that $\frac{X}{E(X)} \succeq_L \frac{Y}{E(Y)}$ implies that $\frac{X}{E(X)} \preceq_{cx} \frac{Y}{E(Y)}$. By the definition of the convex order, we know that

$$E\left(\frac{X}{E(X)} - 1\right)^2 \leq E\left(\frac{Y}{E(Y)} - 1\right)^2,$$

since $\phi(x) = (x-1)^2$ is a convex function. Therefore, $X \succeq_L Y$ implies

$$CV(X) = \sqrt{E\left(\frac{X}{E(X)} - 1\right)^2} \leq \sqrt{E\left(\frac{Y}{E(Y)} - 1\right)^2} = CV(Y).$$

Proposition 9 *For two nonnegative random variables X and Y with $E(X) = E(Y)$,*

$$X = \begin{cases} x, & \text{with probability } v \\ x', & \text{with probabiity } 1-v \end{cases},$$

where $x' \leq x$. Further,

$$Y = \begin{cases} y, & \text{with probability } v \\ y', & \text{with probabiity } 1-v \end{cases},$$

where $y' \leq y$. If $x' \geq y'$, then we have $X \succeq_{SSD} Y$.

Proof: The distribution function of X, $F_X(\tau)$, is[3]

$$F_X(\tau) = (1-v)I_{[x',x)}(\tau) + I_{[x,\infty)}(\tau),\ \tau \in [0,\infty),$$

[3]The indicator function $I_A(\tau)$ is defined as

$$I_A(\tau) = \begin{cases} 1, & \text{if } \tau \in A \\ 0, & \text{if } \tau \notin A \end{cases}.$$

and the distribution function of Y, $F_Y(\tau)$, is

$$F_Y(\tau) = (1-v)I_{[y',y)}(\tau) + I_{[y,\infty)}(\tau),\ \tau \in [0,\infty).$$

Since $E(X) = E(Y)$, $x' \geq y'$ implies $x \leq y$. Thus, for $z \in [0, y')$, we have $\int_0^z [F_X(\tau) - F_Y(\tau)]\, d\tau = 0$, and for $z \in [y', x')$ we have

$$\int_0^z [F_X(\tau) - F_Y(\tau)]\, ds = -\int_{y'}^z (1-v) ds = -(1-v)\,(z-y') \leq 0.$$

For $z \in [x', x)$, we have

$$\int_0^z [F_X(\tau) - F_Y(\tau)]\, d\tau = -\int_{y'}^{x'} (1-v) ds = -(1-v)\,(x'-y') \leq 0.$$

For $z \in [x, y)$, we have

$$\begin{aligned}
\int_0^z [F_X(\tau) - F_Y(\tau)]\, d\tau &= -(1-v)\,(x'-y') + \int_x^z v d\tau \\
&\leq -(1-v)\,(x'-y') + v\,(y-x) \\
&= E\,(Y) - E\,(X) \\
&= 0.
\end{aligned}$$

For $z \in [y, \infty)$, we have

$$\begin{aligned}
\int_0^z [F_X(\tau) - F_Y(\tau)]\, d\tau &= -(1-v)\,(x'-y') + v\,(y-x) \\
&= E\,(Y) - E\,(X) \\
&= 0.
\end{aligned}$$

Thus, we have

$$\int_0^z [F_X(\tau) - F_Y(\tau)]\, d\tau \leq 0,\ \text{for } \forall z \in [0,\infty),$$

which implies $\int_0^z F_X(\tau) d\tau \leq \int_0^z F_Y(\tau) d\tau$ for all $z \in [0,\infty)$. Therefore, we have $X \succeq_{SSD} Y$. ■

Proof of Theorem 1: We have $E(X) = v\underline{\theta}\underline{q} + (1-v)\bar{\theta}\bar{q}$ and $E(Y) =$

$v\left(\underline{\theta}\underline{q}+\underline{U}\right)+(1-v)\bar{\theta}\bar{q}$. Let

$$\hat{X}=\begin{cases}\underline{\theta}\underline{q}/E(X), & \text{with probability } v\\ \bar{\theta}\bar{q}/E(X), & \text{with probabiity } 1-v\end{cases}$$

and

$$\hat{Y}=\begin{cases}\left(\underline{\theta}\underline{q}+\underline{U}\right)/E(Y), & \text{with probability } v\\ \bar{\theta}\bar{q}/E(Y), & \text{with probabiity } 1-v\end{cases}.$$

Since $E\left(\hat{X}\right)=E\left(\hat{Y}\right)=1$ and $\bar{\theta}\bar{q}/E(X)\geq\bar{\theta}\bar{q}/E(Y)$, we know that $\hat{X}\succeq_{SSD}\hat{Y}$ from Proposition 9. Therefore, we have $\hat{X}\succeq_{L}\hat{Y}$ from Propositions 4 and cxL. Thus, we know that $X\succeq_{L}Y$. ■

To compare inequality for the continuous types, we need the following proposition, which extends the results of Fellman (1976):

Proposition 10 *Let X be a nonnegative nondegenerate random variable on $[a,b]$, and let $m(x)$ and $n(x)$ be nonnegative decreasing functions of $x\in[a,b]$ such that $m(x)>0$ and $n(x)>0$ for $x\in[a,b]$. Then, we have*

$$n(X)\succeq_{L}m(X),$$

if $\frac{m(x)}{n(x)}$ is decreasing in $x\in[a,b]$.

Proof: Assume that $X=F_X^{-1}(R)$, where R is a uniform distribution on $(0,1)$. Let $Y=m(X)$ and $Z=n(X)$. We have

$$Y=m(F_X^{-1}(1-R))$$

and

$$Z=n(F_X^{-1}(1-R)).$$

The Lorenz curve of X, $L_X(p)$, is defined as

$$L_X(p)=\frac{1}{E(X)}\int_0^p F_X^{-1}(r)dr,\forall p\in[0,1],$$

where $F_X^{-1}(r)=\inf\{x\geq 0:F_X(x)\geq r\}$.

Thus, we have

$$L_Y(p) = \frac{1}{E(Y)} \int_0^p F_Y^{-1}(r)dr = \frac{1}{E(Y)} \int_0^p m(F_X^{-1}(1-r))dr,$$

where $p \in [0,1]$, and $E(Y) = \int_0^p m(F_X^{-1}(1-r))dr$. Similarly,

$$L_Z(p) = \frac{1}{E(Z)} \int_0^p F_Z^{-1}(r)dr = \frac{1}{E(Z)} \int_0^p n(F_X^{-1}(1-r))dr,$$

where $p \in [0,1]$, and $E(Z) = \int_0^p n(F_X^{-1}(1-r))dr$.

Therefore, for $p \in [0,1]$, we have

$$\begin{aligned} L_Y(p) - L_Z(p) &= \int_0^p \left(\frac{m(F_X^{-1}(1-r))}{E(Y)} - \frac{n(F_X^{-1}(1-r))}{E(Z)} \right) dr \\ &= \int_0^p \frac{1}{n(F_X^{-1}(1-r))} \left(\frac{m(F_X^{-1}(1-r))}{n(F_X^{-1}(1-r))E(Y)} - \frac{1}{E(Z)} \right) dr \leq 0. \end{aligned}$$

Thus, we have $n(X) \succeq_L m(X)$. ■

7.2 Proof of Theorem 3

Proof: Income per capita under the second-best contract is

$$E\left(W^{SB}\right) = v\underline{t}^{SB} + (1-v)\bar{t}^{SB}.$$

Let $\hat{W}^{SB}$ be the normalized income distribution,

$$\hat{W}^{SB} = \frac{W^{SB}}{E\left(W^{SB}\right)} = \begin{cases} \underline{t}^{SB}/E\left(W^{SB}\right), & \text{with probability } v \\ \bar{t}^{SB}/E\left(W^{SB}\right), & \text{with probabiity } (1-v) \end{cases}.$$

Obviously, we have $E\left(\hat{W}^{SB}\right) = 1$. Income per capita under the first-best contract is $E\left(W^{FB}\right) = v\underline{t}^{FB} + (1-v)\bar{t}^{FB}$. Let $\hat{Y}^{FB}$ be the normalized income distribution,

$$\hat{W}^{FB} = \frac{W^{FB}}{E\left(W^{FB}\right)} = \begin{cases} \underline{t}^{FB}/E\left(W^{FB}\right), & \text{with probability } v \\ \bar{t}^{FB}/E\left(W^{FB}\right), & \text{with probabiity } (1-v) \end{cases}.$$

Thus, we have $E\left(\hat{W}^{FB}\right) = 1$.

We have

$$\begin{aligned}
\frac{\bar{t}^{FB}}{E\left(W^{FB}\right)} &= \frac{\bar{t}^{FB}}{v\underline{t}^{FB}+(1-v)\bar{t}^{FB}} \\
&= \frac{1}{v\frac{\underline{t}^{FB}}{\bar{t}^{FB}}+1-v} \\
&\geq \frac{1}{v\frac{\underline{t}^{SB}}{\bar{t}^{SB}}+1-v} \\
&= \frac{\bar{t}^{SB}}{E\left(W^{SB}\right)},
\end{aligned}$$

since we know that $\underline{t}^{FB}/\bar{t}^{FB} \leq \underline{t}^{SB}/\bar{t}^{SB}$ from Proposition 7.

We know that

$$\hat{W}^{FB} \succeq_{SSD} \hat{W}^{SB}$$

from Proposition 9. Thus, by Propositions 4 and 5, we have $\hat{W}^{FB} \succeq_L \hat{W}^{SB}$. Thus, we know that $W^{FB} \succeq_L W^{SB}$. ■

7.3 Examples in Table 1

For the case of the distribution with a constant elaticity $\kappa > 0$, we have $F(\theta) = \theta^\kappa$, for θ in $[a, b]$. Thus, we knnow that $\frac{U^{SB}(\theta)}{\theta q^{SB}(\theta)} = \frac{1-\gamma}{\gamma}\left[1 - \left(\frac{\theta}{b}\right)^{\frac{\gamma}{1-\gamma}}\right]$ is decreasing in θ.

For the case of the uniform distribution, $F(\theta) = \frac{\theta - a}{b-a}$, θ in $[a, b]$. We have $q\left(\theta\right) = (2\theta - a)^{\frac{1}{\gamma-1}}$. We would like to prove that

$$\frac{\theta_1 q(\theta_1)}{b-\theta_1} \leq \frac{\theta_2 q(\theta_2)}{b-\theta_2},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$. That is

$$\frac{q(\theta_2)}{q(\theta_1)} \geq \frac{\theta_1\left(b-\theta_2\right)}{\theta_2\left(b-\theta_1\right)},$$

which is eqivalent to

$$\frac{2\theta_2 - a}{2\theta_1 - a} \leq \left(\frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)}\right)^{\gamma-1}.$$

Let $x = \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)} \leq 1$. Since $x^{\gamma^{-1}} \geq 1 + (1-\gamma)(1-x)$ for x in $(0, 1]$, it is sufficient to show that

$$\frac{2\theta_2 - a}{2\theta_1 - a} - 1 \leq (1-\gamma)\left(1 - \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)}\right).$$

It is sufficient to show that

$$\frac{2(b-\theta_1)}{2\theta_1 - a} \leq (1-\gamma)\frac{b}{\theta_2}.$$

A sufficient condition for this relationship is

$$\frac{b}{a} \leq \frac{3-\gamma}{2}.$$

For the case of distribution $F(\theta) = \frac{\log(\theta)-\log(a)}{\log(b)-\log(a)}$, θ in $[a,b]$, we have $q(\theta) = [\theta(\log(\theta) - \log(a) + 1)]^{\frac{1}{\gamma-1}}$. We would like to prove that

$$\frac{q(\theta_2)}{q(\theta_1)} \geq \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)},$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, which is equivalent to

$$\frac{\theta_2(\log(\theta_2) - \log(a) + 1)}{\theta_1(\log(\theta_1) - \log(a) + 1)} \leq \left(\frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)}\right)^{\gamma-1}.$$

Let $x = \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)} \leq 1$. Since $x^{\gamma-1} \geq 1 + (1-\gamma)(1-x)$ for x in $(0,1]$, it is sufficient to show that

$$(1-\gamma)\frac{(\theta_2-\theta_1)\,b}{\theta_2(b-\theta_1)} \geq \frac{\theta_2(\log(\theta_2) - \log(\theta_1)) + (\theta_2-\theta_1)(\log(\theta_1) - \log(a) + 1)}{\theta_1(\log(\theta_1) - \log(a) + 1)}.$$

Since $log(\theta)$ is a concave function, we have $\log(\theta_2) - \log(\theta_1) \leq \frac{1}{\theta_1}(\theta_2 - \theta_1)$. It is sufficient to prove that

$$(1-\gamma)\,b \geq \frac{\theta_2^2(b-\theta_1)}{\theta_1^2(\log(\theta_1) - \log(a) + 1)} + \frac{\theta_2(b-\theta_1)}{\theta_1}.$$

A sufficient condition for this relationship is

$$\frac{b}{a} \leq \sqrt{2-\gamma}.$$

For the case of distribution $F(\theta) = \sqrt{\frac{\log(\theta)}{\log(b)}}$, $1 \leq \theta \leq b$, we have $q^{SB}(\theta) = [\theta(2\log(\theta)+1)]^{\frac{1}{\gamma-1}}$. We would like to prove that

$$\frac{q(\theta_2)}{q(\theta_1)} \geq \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)},$$

for $1 \leq \theta_1 \leq \theta_2 \leq b$, which is equivalent to

$$\frac{\theta_2(2\log(\theta_2)+1)}{\theta_1(2\log(\theta_1)+1)} \leq \left(\frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)}\right)^{\gamma-1}.$$

Let $x = \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)} \leq 1$. Since $x^{\gamma-1} \geq 1+(1-\gamma)(1-x)$ for x in $(0,1]$, it is sufficient to show that

$$(1-\gamma)\frac{(\theta_2-\theta_1)\,b}{\theta_2(b-\theta_1)} \geq \frac{2\theta_2(\log(\theta_2)-\log(\theta_1))+(\theta_2-\theta_1)(2\log(\theta_1)+1)}{\theta_1(2\log(\theta_1)+1)}.$$

Since $log(\theta)$ is a concave function, we have $\log(\theta_2)-\log(\theta_1) \leq \frac{1}{\theta_1}(\theta_2-\theta_1)$. It is sufficient to prove that

$$(1-\gamma)\,b \geq \frac{2\theta_2^2(b-\theta_1)}{\theta_1^2(2\log(\theta_1)+1)} + \frac{\theta_2(b-\theta_1)}{\theta_1}.$$

A sufficient condition for this relationship is

$$b \leq \sqrt{\frac{3-\gamma}{2}}.$$

For the case of distribution $F(\theta) = \frac{\theta\log(\theta)}{b\log(b)}$, $1 \leq \theta \leq b$, we have $q^{SB}(\theta) = \left[\theta\left(\frac{2\log(\theta)+1}{\log(\theta)+1}\right)\right]^{\frac{1}{\gamma-1}}$. We would like to prove that

$$\frac{q(\theta_2)}{q(\theta_1)} \geq \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)},$$

for $1 \leq \theta_1 \leq \theta_2 \leq b$, which is equivalent to

$$\frac{\theta_2\left(\frac{2\log(\theta_2)+1}{\log(\theta_2)+1}\right)}{\theta_1\left(\frac{2\log(\theta_1)+1}{\log(\theta_1)+1}\right)} \leq \left(\frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)}\right)^{\gamma-1}.$$

Let $x = \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)} \leq 1$. Since $x^{\gamma-1} \geq 1 + (1-\gamma)(1-x)$ for x in $(0,1]$, it is sufficient to show that

$$(1-\gamma)\frac{(\theta_2-\theta_1)\,b}{\theta_2\,(b-\theta_1)} \geq \frac{\theta_2\left(\frac{\log(\theta_2)}{\log(\theta_2)+1} - \frac{\log(\theta_1)}{\log(\theta_1)+1}\right) + (\theta_2-\theta_1)\left(\frac{\log(\theta_1)}{\log(\theta_1)+1}+1\right)}{\theta_1\left(\frac{\log(\theta_1)}{\log(\theta_1)+1}+1\right)}.$$

$1 \leq \theta_1 \leq \theta_2$ implies $\frac{1}{\log(\theta_2)+1} \leq \frac{1}{\log(\theta_1)+1}$. Since $log(\theta)$ is a concave function, we have $\log(\theta_2) - \log(\theta_1) \leq \frac{1}{\theta_1}(\theta_2-\theta_1)$. It is sufficient to prove that

$$(1-\gamma)\,b \geq \frac{\theta_2^2\,(b-\theta_1)}{\theta_1^2\,(2\log(\theta_1)+1)\,(\log(\theta_1)+1)} + \frac{\theta_2\,(b-\theta_1)}{\theta_1}.$$

A sufficient condition for this relationship is

$$b \leq \sqrt{2-\gamma}.$$

Consider the following density function:

$$f(\theta) = \begin{cases} \frac{4(1-h(b-a))}{(b-a)^2}(\theta-a)+h, \; for\,\theta \in \left[a, \frac{a+b}{2}\right]; \\ \frac{4(1-h(b-a))}{(b-a)^2}(b-\theta)+h, \; for\,\theta \in \left[\frac{a+b}{2}, b\right], \end{cases}$$

with cumulative distribution function

$$F(\theta) = \begin{cases} h(\theta-a) + \frac{2(1-h(b-a))}{(b-a)^2}(\theta-a)^2, \; for\,\theta \in \left[a, \frac{a+b}{2}\right]; \\ 1 - h(b-\theta) - \frac{2(1-h(b-a))}{(b-a)^2}(b-\theta)^2, \; for\,\theta \in \left[\frac{a+b}{2}, b\right]. \end{cases}$$

Therefore,

$$e(\theta) = \frac{\theta f(\theta)}{F(\theta)} = \begin{cases} e_1(\theta) = \frac{\theta\left[4(1-h(b-a))(\theta-a)+h(b-a)^2\right]}{h(b-a)^2(\theta-a)+2(1-h(b-a))(\theta-a)^2}, \; for\,\theta \in \left[a, \frac{a+b}{2}\right]; \\ e_2(\theta) = \frac{\theta\left[4(1-h(b-a))(b-\theta)+h(b-a)^2\right]}{(b-a)^2-h(b-a)^2(b-\theta)-2(1-h(b-a))(b-\theta)^2}, \; for\,\theta \in \left[\frac{a+b}{2}, b\right]. \end{cases}$$

We first check that $e(\theta)$ is decreasing in θ. With $a=1$, $b=1.2$, and $h=4$, $e(\theta)$ can be rewritten as

$$e(\theta) = \begin{cases} e_1(\theta) = 1 + \frac{\theta^2-0.6}{\theta^2-1.6\theta+0.6}, \; for\,\theta \in [1, 1.1]; \\ e_2(\theta) = 1 + \frac{0.728-0.4\theta^2}{1.12\theta-0.4\theta^2-0.728}, \; for\,\theta \in [1.1, 1.2]. \end{cases}$$

Since $\theta^2 - 1.6\theta + 0.6 > 0$, for all $(\theta_1, \theta_2) \in [1, 1.1] \times [1, 1.1]$ and $\theta_1 \le \theta_2$,

$$\begin{aligned} &\frac{\theta_2^2 - 0.6}{\theta_2^2 - 1.6\theta_2 + 0.6} - \frac{\theta_1^2 - 0.6}{\theta_1^2 - 1.6\theta_1 + 0.6} \\ &= \frac{(\theta_2 - \theta_1)\left(1.2\left(\theta_2 + \theta_1\right) - 1.6\theta_1\theta_2 - 0.96\right)}{\left(\theta_2^2 - 1.6\theta_2 + 0.6\right)\left(\theta_1^2 - 1.6\theta_1 + 0.6\right)}, \end{aligned}$$

where $1.2\left(\theta_2 + \theta_1\right) - 1.6\theta_1\theta_2 - 0.96 < 0$, because when $\theta_2 = \theta_1$, $1.2\left(\theta_2 + \theta_1\right) - 1.6\theta_1\theta_2 - 0.96 = -1.6\theta_1^2 + 2.4\theta_1 - 0.96 < 0$ and since $1.2 - 1.6\theta_1 < 0$, $1.2\left(\theta_2 + \theta_1\right) - 1.6\theta_1\theta_2 - 0.96 = \theta_2\left(1.2 - 1.6\theta_1\right) + 1.2\theta_1 - 0.96$ is decreasing in θ_2. Therefore, $e_1(\theta)$ is decreasing. Then, $1.12\theta - 0.4\theta^2 - 0.728 > 0$ increases in θ and $0.728 - 0.4\theta^2 > 0$ imples that $e_2(\theta)$ is decreasing. Moreover, $e_1(1.1) = e_2(1.1)$. Hence, $e(\theta)$ is decreasing.

Then, we would like to prove that

$$\left(1 + \frac{1}{e(\theta)}\right)\theta^{\gamma}(b - \theta)^{1-\gamma}$$

is decreasing; that is, for all $(\theta_1, \theta_2) \in [a, b] \times [a, b]$ and $\theta_1 \le \theta_2$,

$$\frac{\theta_2\left(1 + \frac{1}{e(\theta_2)}\right)}{\theta_1\left(1 + \frac{1}{e(\theta_1)}\right)} \le \left(\frac{\theta_1\left(b - \theta_2\right)}{\theta_2\left(b - \theta_1\right)}\right)^{\gamma-1}.$$

Denote $X \equiv \frac{\theta_1(b-\theta_2)}{\theta_2(b-\theta_1)} \le 1$. Since

$$X^{\gamma-1} \ge 1 + (1 - \gamma)(1 - X), \tag{15}$$

we need to prove that

$$(1 - \gamma)\frac{(\theta_2 - \theta_1)\, b}{\theta_2\left(b - \theta_1\right)} \ge \frac{\theta_2\left(1 + \frac{1}{e(\theta_2)}\right)}{\theta_1\left(1 + \frac{1}{e(\theta_1)}\right)} - 1 \tag{16}$$

which is

$$(1 - \gamma)\frac{(\theta_2 - \theta_1)\, b}{\theta_2\left(b - \theta_1\right)} \ge \frac{\theta_2\left(\frac{1}{e(\theta_2)} - \frac{1}{e(\theta_1)}\right) + (\theta_2 - \theta_1)\left(1 + \frac{1}{e(\theta_1)}\right)}{\theta_1\left(1 + \frac{1}{e(\theta_1)}\right)}. \tag{17}$$

We will prove it in the following three cases with $a = 1$, $b = 1.2$, and

$h = 4$:

i. for all $(\theta_1, \theta_2) \in [a, \frac{a+b}{2}] \times [a, \frac{a+b}{2}]$ and $\theta_1 \leq \theta_2$;

ii. for all $(\theta_1, \theta_2) \in [\frac{a+b}{2}, b] \times [\frac{a+b}{2}, b]$ and $\theta_1 \leq \theta_2$;

iii. for all $\theta_1 \in [a, \frac{a+b}{2}]$, $\theta_2 \in [\frac{a+b}{2}, b]$ and $\theta_1 \leq \theta_2$.

For case i, taking $\theta_1 \leq \theta_2$ in $[1, 1.1]$,

$$\begin{aligned}
&\left(\theta_2 - \tfrac{1}{e_1(\theta_2)}\right) - \left(\theta_1 - \tfrac{1}{e_1(\theta_1)}\right) \\
&= \tfrac{\left(\theta_2^3 - 1.3\theta_2^2 + 0.8\theta_2 - 0.3\right)\left(\theta_1^2 - 0.8\theta_1\right) - \left(\theta_1^3 - 1.3\theta_1^2 + 0.8\theta_1 - 0.3\right)\left(\theta_2^2 - 0.8\theta_2\right)}{\left(\theta_2^2 - 0.8\theta_2\right)\left(\theta_1^2 - 0.8\theta_1\right)}, \\
&= \tfrac{(\theta_2 - \theta_1)\left[\left(\theta_1^2 - 0.8\theta_1\right)\theta_2^2 + (0.3 + 0.24\theta_1 - 0.8\theta_1^2)\theta_2 + 0.3\theta_1 - 0.24\right]}{\left(\theta_2^2 - 0.8\theta_2\right)\left(\theta_1^2 - 0.8\theta_1\right)}.
\end{aligned}$$

The axis of symmetry of the polynomial $(\theta_1^2 - 0.8\theta_1)\,\theta_2^2 + (0.3 + 0.24\theta_1 - 0.8\theta_1^2)\theta_2 + 0.3\theta_1 - 0.24$ is $0.4 + \frac{0.1875}{\theta_1} + \frac{0.025}{2\theta_1 - 1.6}$ which decreases in θ_1. Therefore, the maximum of the axis of symmetry is $0.65 < 1$. Since $\theta_1^2 - 0.8\theta_1 > 0$, $(\theta_1^2 - 0.8\theta_1)\,\theta_2^2 + (0.3 + 0.24\theta_1 - 0.8\theta_1^2)\theta_2 + 0.3\theta_1 - 0.24$ increases in θ_2 on $[1, 1.1]$ and the minimum at $\theta_2 = 1$ is $0.2\theta_1^2 - 0.26\theta_1 + 0.06$ which has two roots being 0.3 and 1. Hence, we have $0.2\theta_1^2 - 0.26\theta_1 + 0.06 \geq 0$ on the interval $[1, 1.1]$ and thus the minimum of $(\theta_1^2 - 0.8\theta_1)\,\theta_2^2 + (0.3 + 0.24\theta_1 - 0.8\theta_1^2)\theta_2 + 0.3\theta_1 - 0.24$ is greater or equal to zero. Hence, we have $\theta_2 - \frac{1}{e_1(\theta_2)} \geq \theta_1 - \frac{1}{e_1(\theta_1)}$, thus $\frac{1}{e_1(\theta_2)} - \frac{1}{e_1(\theta_1)} \leq \theta_2 - \theta_1$. Therefore, a sufficient condition for inequality (17) is

$$(1-\gamma)\,\frac{(\theta_2 - \theta_1)\,b}{\theta_2\,(b - \theta_1)} \geq \frac{\theta_2\,(\theta_2 - \theta_1) + (\theta_2 - \theta_1)\left(1 + \frac{1}{e_1(\theta_1)}\right)}{\theta_1\left(1 + \frac{1}{e_1(\theta_1)}\right)},$$

which is

$$(1-\gamma)\,b \geq \frac{\theta_2^2}{\frac{\theta_1}{(b-\theta_1)}\left(1 + \frac{1}{e(\theta_1)}\right)} + \frac{\theta_2}{\frac{\theta_1}{(b-\theta_1)}}.$$

Then, since $e_1(\theta_1)$ is decreasing and $\frac{\theta_1}{(b-\theta_1)}$ is increasing, $\frac{\theta_1}{(b-\theta_1)}\left(1 + \frac{1}{e_1(\theta_1)}\right)$ is increasing, and a sufficient condition on γ is taking $\theta_2 = 1.1$ and $\theta_1 = 1$. Therefore, we obtain $\gamma \leq 0.615$.

For case ii, taking $\theta_1 \leq \theta_2$ in $[1.1, 1.2]$,

$$\begin{aligned}
&\left(2\theta_2 - \tfrac{1}{e_2(\theta_2)}\right) - \left(2\theta_1 - \tfrac{1}{e_2(\theta_1)}\right) \\
&= \tfrac{\left(-4\theta_2^3+6.6\theta_2^2-2.8\theta_2+1.82\right)\left(2.8\theta_1-2\theta_1^2\right)-\left(-4\theta_1^3+6.6\theta_1^2-2.8\theta_1+1.82\right)\left(2.8\theta_2-2\theta_2^2\right)}{\left(2.8\theta_2-2\theta_2^2\right)\left(2.8\theta_1-2\theta_1^2\right)}, \\
&= \tfrac{(\theta_2-\theta_1)\left[\left(8\theta_1^2-11.2\theta_1\right)\theta_2^2+(3.64+12.88\theta_1-11.2\theta_1^2)\theta_2+3.64\theta_1-5.096\right]}{\left(2.8\theta_2-2\theta_2^2\right)\left(2.8\theta_1-2\theta_1^2\right)},
\end{aligned}$$

The axis of symmetry of the polynomial $\left(8\theta_1^2 - 11.2\theta_1\right)\theta_2^2 + (3.64 + 12.88\theta_1 - 11.2\theta_1^2)\theta_2 + 3.64\theta_1 - 5.096$ is $0.7 + \frac{0.05}{\theta_1} - \frac{0.1625}{5.6-4\theta_1}$ which decreases in θ_1. Therefore, the maximum of the axis of symmetry is less than $0.6101 < 1$. Since $8\theta_1^2 - 11.2\theta_1 < 0$, $\left(8\theta_1^2 - 11.2\theta_1\right)\theta_2^2 + (3.64 + 12.88\theta_1 - 11.2\theta_1^2)\theta_2 + 3.64\theta_1 - 5.096$ decreases in θ_2 on $[1.1, 1.2]$ and the minimum at $\theta_2 = 1.2$ is $-1.92\theta_1^2 + 2.968\theta_1 - 0.728$ which has the axis of symmetry less than 1.1 and its minimum at $\theta_1 = 1.2$ is $0.0688 > 0$. Hence, we have $-1.92\theta_1^2 + 2.968\theta_1 - 0.728 > 0$ on the interval $[1.1, 1.2]$ and the minimum of $\left(8\theta_1^2 - 11.2\theta_1\right)\theta_2^2 + (3.64 + 12.88\theta_1 - 11.2\theta_1^2)\theta_2 + 3.64\theta_1 - 5.096$ is positive. Hence, we have $2\theta_2 - \frac{1}{e_2(\theta_2)} \geq 2\theta_1 - \frac{1}{e_2(\theta_1)}$, thus $\frac{1}{e_2(\theta_2)} - \frac{1}{e_2(\theta_1)} \leq 2\left(\theta_2 - \theta_1\right)$. Therefore, a sufficient condition for inequality (17) is

$$(1-\gamma)\, b \geq \frac{2\theta_2^2}{\frac{\theta_1}{(b-\theta_1)}\left(1 + \frac{1}{e(\theta_1)}\right)} + \frac{\theta_2}{\frac{\theta_1}{(b-\theta_1)}},$$

and we obtain a sufficient condition $\gamma \leq 0.706$.

For case iii, $\theta_1 \in [1, 1.1]$ and $\theta_2 \in [1.1, 1.2]$. $1.8\theta_2 - \frac{1}{e_2(\theta_2)}$ is increasing in θ_2 because taking $\theta_{21} \leq \theta_{22} \in [1.1, 1.2]$,

$$\begin{aligned}
&\left(1.8\theta_{22} - \tfrac{1}{e_2(\theta_{22})}\right) - \left(1.8\theta_{21} - \tfrac{1}{e_2(\theta_{21})}\right) \\
&= \tfrac{\left(-3.6\theta_{22}^3+6.04\theta_{22}^2-2.8\theta_{22}+1.82\right)\left(2.8\theta_{21}-2\theta_{21}^2\right)-\left(-3.6\theta_{21}^3+6.04\theta_{21}^2-2.8\theta_{21}+1.82\right)\left(2.8\theta_{22}-2\theta_{22}^2\right)}{\left(2.8\theta_{22}-2\theta_{22}^2\right)\left(2.8\theta_{21}-2\theta_{21}^2\right)}, \\
&= \tfrac{(\theta_{22}-\theta_{21})\left[\left(7.2\theta_{21}^2-10.08\theta_{21}\right)\theta_{22}^2+\left(3.64+11.312\theta_{21}-10.08\theta_{21}^2\right)\theta_{22}+3.64\theta_{21}-5.096\right]}{\left(2.8\theta_{22}-2\theta_{22}^2\right)\left(2.8\theta_{21}-2\theta_{21}^2\right)}.
\end{aligned}$$

The axis of symmetry of the polynomial $\left(7.2\theta_{21}^2 - 10.08\theta_{21}\right)\theta_{22}^2 + \left(3.64 + 11.312\theta_{21} - 10.08\theta_{21}^2\right)\theta_{22} +$ $3.64\theta_{21} - 5.096$ is $0.7 + \frac{(1.82/10.08)}{\theta_{21}} - \frac{0.1}{10.08-7.2\theta_{21}}$ which decreases in θ_{21}. Therefore, the maximum of the axis of symmetry is less than $0.818 < 1.1$. Since $7.2\theta_{21}^2 - 10.08\theta_{21} < 0$, $\left(7.2\theta_{21}^2 - 10.08\theta_{21}\right)\theta_{22}^2 + \left(3.64 + 11.312\theta_{21} - 10.08\theta_{21}^2\right)\theta_{22} +$

$3.64\theta_{21} - 5.096$ decreases in θ_{22} on $[1.1, 1.2]$ and the minimum at $\theta_{22} = 1.2$ is $-1.728\theta_{21}^2 + 2.6992\theta_{21} - 0.728$ which has the axis of symmetry less than 1.1 and its minimum at $\theta_{21} = 1.2$ is $0.0222 > 0$. Hence, we have $-1.728\theta_{21}^2 + 2.6992\theta_{21} - 0.728 > 0$ on the interval $[1.1, 1.2]$ and the minimum of $(7.2\theta_{21}^2 - 10.08\theta_{21})\,\theta_{22}^2 + (3.64 + 11.312\theta_{21} - 10.08\theta_{21}^2)\,\theta_{22} + 3.64\theta_{21} - 5.096$ is positive. Therefore, $1.8\theta_2 - \frac{1}{e_2(\theta_2)}$ is increasing in θ_2.

Moreover, $1.8\theta_1 - \frac{1}{e_1(\theta_1)}$ is increasing in θ_1 because taking $\theta_{11} \leq \theta_{12} \in [1, 1.1]$,

$$\begin{aligned}
&\left(1.8\theta_{12} - \tfrac{1}{e_1(\theta_{12})}\right) - \left(1.8\theta_{11} - \tfrac{1}{e_1(\theta_{11})}\right) \\
&= \tfrac{\left(1.8\theta_{12}^3 - 1.94\theta_{12}^2 + 0.8\theta_{12} - 0.3\right)\left(\theta_{11}^2 - 0.8\theta_{11}\right) - \left(1.8\theta_{11}^3 - 1.94\theta_{11}^2 + 0.8\theta_{11} - 0.3\right)\left(\theta_{12}^2 - 0.8\theta_{12}\right)}{\left(\theta_{12}^2 - 0.8\theta_{12}\right)\left(\theta_{11}^2 - 0.8\theta_{11}\right)}, \\
&= \tfrac{(\theta_{12} - \theta_{11})\left[\left(1.8\theta_{11}^2 - 1.44\theta_{11}\right)\theta_{12}^2 + \left(0.3 + 0.752\theta_{11} - 1.44\theta_{11}^2\right)\theta_{12} + 0.3\theta_{11} - 0.24\right]}{\left(\theta_{12}^2 - 0.8\theta_{12}\right)\left(\theta_{11}^2 - 0.8\theta_{11}\right)}.
\end{aligned}$$

The axis of symmetry of the polynomial $(1.8\theta_{11}^2 - 1.44\theta_{11})\,\theta_{12}^2 + (0.3 + 0.752\theta_{11} - 1.44\theta_{11}^2)\,\theta_{12} + 0.3\theta_{11} - 0.24$ is $0.4 + \frac{(0.15/1.44)}{\theta_{11}} + \frac{0.0125}{1.8\theta_{11} - 1.44}$ which decreases in θ_1. Therefore, the maximum of the axis of symmetry is less than $0.5389 < 1$. Since $1.8\theta_{11}^2 - 1.44\theta_{11} > 0$, $(1.8\theta_{11}^2 - 1.44\theta_{11})\,\theta_{12}^2 + (0.3 + 0.752\theta_{11} - 1.44\theta_{11}^2)\,\theta_{12} + 0.3\theta_{11} - 0.24$ increases in θ_{12} on $[1, 1.1]$ and the minimum at $\theta_{12} = 1$ is $0.36\theta_{11}^2 - 0.388\theta_{11} + 0.06$ which has the axis of symmetry less than 1 and its minimum at $\theta_{11} = 1$ is $0.032 > 0$. Hence, we have $0.36\theta_{11}^2 - 0.388\theta_{11} + 0.06 > 0$ on the interval $[1, 1.1]$ and the minimum of $(1.8\theta_{11}^2 - 1.44\theta_{11})\,\theta_{12}^2 + (0.3 + 0.752\theta_{11} - 1.44\theta_{11}^2)\,\theta_{12} + 0.3\theta_{11} - 0.24$ is positive. Therefore, $1.8\theta_1 - \frac{1}{e_1(\theta_1)}$ is increasing in θ_1.

Furthermore, $\min\left(1.8\theta_2 - \frac{1}{e_2(\theta_2)}\right) = 1.8 \cdot 1.1 - \frac{1}{e_2(1.1)} = \max\left(1.8\theta_1 - \frac{1}{e_1(\theta_1)}\right) = 1.8 \cdot 1.1 - \frac{1}{e_1(1.1)} = 1.98 - \frac{0.02}{0.264}$. Hence, we have $1.8\theta_2 - \frac{1}{e_2(\theta_2)} \geq 1.8\theta_1 - \frac{1}{e_1(\theta_1)}$, thus $\frac{1}{e_2(\theta_2)} - \frac{1}{e_1(\theta_1)} \leq 1.8\,(\theta_2 - \theta_1)$. Therefore, a sufficient condition for inequality (17) is

$$(1-\gamma)\,b \geq \frac{1.8\theta_2^2}{\frac{\theta_1}{(b-\theta_1)}\left(1 + \frac{1}{e_1(\theta_1)}\right)} + \frac{\theta_2}{\frac{\theta_1}{(b-\theta_1)}},$$

and we obtain a sufficient condition $\gamma \leq 0.368$.

In all, a sufficient conditon for $\left(1 + \frac{1}{e(\theta)}\right)\theta^{\gamma}(b-\theta)^{1-\gamma}$ to be decreasing is $\gamma \leq 0.368$.

7.4 Derivation of social norms

For uniform distribution on $[a,b]$, we have $e(\theta)=\frac{\theta}{\theta-a}$, and the second-best

$$
\begin{aligned}
q^{\alpha}(\theta) &= (2-\alpha)^{\frac{1}{\gamma-1}}\left(\theta-\frac{1-\alpha}{2-\alpha}a\right)^{\frac{1}{\gamma-1}}, \\
t^{\alpha}(\theta) &= (2-\alpha)^{\frac{1}{\gamma-1}}\frac{\gamma-1}{\gamma}\left[\left(b-\frac{1-\alpha}{2-\alpha}a\right)^{\frac{\gamma}{\gamma-1}}-\left(\theta-\frac{1-\alpha}{2-\alpha}a\right)^{\frac{\gamma}{\gamma-1}}\right] \\
&\quad +(2-\alpha)^{\frac{1}{\gamma-1}}\left(\theta-\frac{1-\alpha}{2-\alpha}a\right)^{\frac{1}{\gamma-1}}\theta.
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{t^{\alpha_2}(\theta)}{t^{\alpha_1}(\theta)} = \\
&\frac{(2-\alpha_2)^{\frac{1}{\gamma-1}}\frac{\gamma-1}{\gamma}\left[\left(b-\frac{1-\alpha_2}{2-\alpha_2}a\right)^{\frac{\gamma}{\gamma-1}}-\left(\theta-\frac{1-\alpha_2}{2-\alpha_2}a\right)^{\frac{\gamma}{\gamma-1}}\right]+(2-\alpha_2)^{\frac{1}{\gamma-1}}\left(\theta-\frac{1-\alpha_2}{2-\alpha_2}a\right)^{\frac{1}{\gamma-1}}\theta}{(2-\alpha_1)^{\frac{1}{\gamma-1}}\frac{\gamma-1}{\gamma}\left[\left(b-\frac{1-\alpha_1}{2-\alpha_1}a\right)^{\frac{\gamma}{\gamma-1}}-\left(\theta-\frac{1-\alpha_1}{2-\alpha_1}a\right)^{\frac{\gamma}{\gamma-1}}\right]+(2-\alpha_1)^{\frac{1}{\gamma-1}}\left(\theta-\frac{1-\alpha_1}{2-\alpha_1}a\right)^{\frac{1}{\gamma-1}}\theta}.
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
&\frac{\partial}{\partial\theta}\left(\frac{t^{\alpha_2}(\theta)}{t^{\alpha_1}(\theta)}\right) = \\
&\frac{1}{\left((2-\alpha_1)^{\frac{1}{\gamma-1}}\frac{\gamma-1}{\gamma}\left[\left(b-\frac{1-\alpha_1}{2-\alpha_1}a\right)^{\frac{\gamma}{\gamma-1}}-\left(\theta-\frac{1-\alpha_1}{2-\alpha_1}a\right)^{\frac{\gamma}{\gamma-1}}\right]+(2-\alpha_1)^{\frac{1}{\gamma-1}}\left(\theta-\frac{1-\alpha_1}{2-\alpha_1}a\right)^{\frac{1}{\gamma-1}}\theta\right)^2} \\
&\cdot\frac{1}{\gamma}(2-\alpha_2)^{\frac{1}{\gamma-1}}(2-\alpha_1)^{\frac{1}{\gamma-1}}\left(\theta-\frac{1-\alpha_2}{2-\alpha_2}a\right)^{\frac{1}{\gamma-1}-1}\left(\theta-\frac{1-\alpha_1}{2-\alpha_1}a\right)^{\frac{1}{\gamma-1}-1}\theta \\
&\cdot\left\{\left(\theta-\frac{1-\alpha_1}{2-\alpha_1}a\right)\left[\left(\frac{b-\frac{1-\alpha_1}{2-\alpha_1}a}{\theta-\frac{1-\alpha_1}{2-\alpha_1}a}\right)^{\frac{\gamma}{\gamma-1}}-1\right]-\left(\theta-\frac{1-\alpha_2}{2-\alpha_2}a\right)\left[\left(\frac{b-\frac{1-\alpha_2}{2-\alpha_2}a}{\theta-\frac{1-\alpha_2}{2-\alpha_2}a}\right)^{\frac{\gamma}{\gamma-1}}-1\right]\right\}.
\end{aligned}
$$

Denote $g = g(\alpha) \equiv \frac{1-\alpha}{2-\alpha}$, $g'(\alpha) < 0$. In addition,

$$\begin{aligned}
& \frac{\partial}{\partial g}\left[(\theta - ga)\left[\left(\frac{b-ga}{\theta-ga}\right)^{\frac{\gamma}{\gamma-1}} - 1\right]\right] \\
= & \frac{\partial}{\partial g}\left[\frac{(b-ga)^{\frac{\gamma}{\gamma-1}}}{(\theta-ga)^{\frac{1}{\gamma-1}}} - \theta + ga\right] \\
= & \frac{1}{(\theta-ga)^{\frac{2}{\gamma-1}}}\frac{1}{\gamma-1}a(b-ga)^{\frac{\gamma}{\gamma-1}-1}(\theta-ga)^{\frac{1}{\gamma-1}-1}\left[b-ga-\gamma(\theta-ga)\right] + a.
\end{aligned}$$

With $a > 0$, $\gamma - 1 < 0$, $\frac{1}{(\theta-ga)^{\frac{2}{\gamma-1}}}\frac{1}{\gamma-1}a(b-ga)^{\frac{\gamma}{\gamma-1}-1}(\theta-ga)^{\frac{1}{\gamma-1}-1}\left[b-ga-\gamma(\theta-ga)\right] + a \geq 0$ if and only if

$$\frac{(b-ga)-\gamma(\theta-ga)}{(\theta-ga)^{\frac{\gamma}{\gamma-1}}} \leq \frac{(b-ga)-\gamma(b-ga)}{(b-ga)^{\frac{\gamma}{\gamma-1}}}$$

which is true because

$$\begin{aligned}
& \frac{\partial}{\partial \theta}\frac{(b-ga)-\gamma(\theta-ga)}{(\theta-ga)^{\frac{\gamma}{\gamma-1}}} \\
= & \frac{1}{(\theta-ga)^{\frac{2\gamma}{\gamma-1}}}\frac{\gamma}{1-\gamma}(\theta-ga)^{\frac{\gamma}{\gamma-1}-1}\left[(b-ga)-\gamma(\theta-ga)-(1-\gamma)(\theta-ga)\right], \\
= & \frac{1}{(\theta-ga)^{\frac{2\gamma}{\gamma-1}}}\frac{\gamma}{1-\gamma}(\theta-ga)^{\frac{\gamma}{\gamma-1}-1}\left[(b-ga)-(\theta-ga)\right] \geq 0.
\end{aligned}$$

Therefore, $\frac{\partial}{\partial g}\left[(\theta-ga)\left[\left(\frac{b-ga}{\theta-ga}\right)^{\frac{\gamma}{\gamma-1}} - 1\right]\right] \geq 0$ and $g'(\alpha) < 0$ implies that

$$\left(\theta - \frac{1-\alpha_1}{2-\alpha_1}a\right)\left[\left(\frac{b-\frac{1-\alpha_1}{2-\alpha_1}a}{\theta-\frac{1-\alpha_1}{2-\alpha_1}a}\right)^{\frac{\gamma}{\gamma-1}} - 1\right] - \left(\theta - \frac{1-\alpha_2}{2-\alpha_2}a\right)\left[\left(\frac{b-\frac{1-\alpha_2}{2-\alpha_2}a}{\theta-\frac{1-\alpha_2}{2-\alpha_2}a}\right)^{\frac{\gamma}{\gamma-1}} - 1\right] \leq 0,$$

and thus

$$\frac{\partial}{\partial \theta}\left(\frac{t^{\alpha_2}(\theta)}{t^{\alpha_1}(\theta)}\right) \leq 0.$$

Thus, we have

$$\frac{t^{\alpha_2}(\theta_1)}{t^{\alpha_1}(\theta_1)} \geq \frac{t^{\alpha_2}(\theta_2)}{t^{\alpha_1}(\theta_2)}$$

for $a \leq \theta_1 \leq \theta_2 \leq b$, if $\alpha_1 \geq \alpha_2$.

7.5 Taking into account the principal's income

In the main text, we only investigate the income distribution among agents. In this extension, we investigate the income distribution which incoporates the principal's profits.

There is a principal with measure p_1 and a continuum of agents with measure $1-p_1$ in the economy. The principal roughly represents the top 1% of the income distribution, $p_1 = 0.01$. If $S(q) = \frac{1}{\gamma}q^\gamma$, $\gamma \in (0,1)$, the income share of agents in the total income under complete information is γ, which is the same as that under incomplete information.

We introduce a variable Ψ into the model, where $\Psi = a$ represents an agent, and $\Psi = p$ represents a principal. We use the decomposition technique to investigate income distribution.

Let δ^{SB} be the principal's profit in an economy with incomplete information,

$$\begin{aligned}\delta^{SB} = \left(\frac{1}{p_1} - 1\right) \quad \times \quad & \left\{ v\left[S\left(\underline{q}^{SB}\right) - \underline{\theta}\underline{q}^{SB}\right] + (1-v)\left[S\left(q^{\bar{S}B}\right) - \bar{\theta}q^{\bar{S}B}\right] \right.\\ & \left. - \left[v\underline{U}^{SB} + (1-v)\bar{U}^{SB}\right]\right\}.\end{aligned}$$

Let Y^{SB} be income distribution in the economy with incomplete information,

$$Y^{SB} = \begin{cases} \delta^{SB}, & \text{with probability } p_1 \\ \underline{t}^{SB}, & \text{with probability } (1-p_1)\,v \\ \bar{t}^{SB}, & \text{with probabiity } (1-p_1)\,(1-v) \end{cases}.$$

Thus, income per capita is

$$E\left(Y^{SB}\right) = (1-p_1)\left[vS\left(\underline{q}^{SB}\right) + (1-v)S\left(\bar{q}^{SB}\right)\right].$$

Let $\hat{Y}^{SB}$ be the normalized income distribution under the second-best contract,

$$\hat{Y}^{SB} = \frac{Y^{SB}}{E\left(Y^{SB}\right)}.$$

$\hat{Y}^{SB}$ can be viewed as a mixture of distributions,

$$\left[\hat{Y}^{SB}\middle|\,\Psi = a\right] = \frac{W^{SB}}{E\left(Y^{SB}\right)}$$

and

$$\left[\hat{Y}^{SB}\middle|\,\Psi = p\right] = \frac{\delta^{SB}}{E\left(Y^{SB}\right)} = \frac{1-\gamma}{p_1}.$$

Let δ^{FB} be the principal's profit in an economy with incomplete information,

$$\delta^{FB} = \left(\frac{1}{p_1} - 1\right)\left\{v\left[S\left(\underline{q}^{SB}\right) - \underline{\theta}\underline{q}^{SB}\right] + (1-v)\left[S\left(q^{\bar{S}B}\right) - \bar{\theta}q^{\bar{S}B}\right]\right.$$

Let Y^{FB} be income distribution in an economy with complete information,

$$Y^{FB} = \begin{cases} \delta^{FB}, & \text{with probability } p_1 \\ \underline{t}^{FB}, & \text{with probability } (1-p_1)\,v \\ \bar{t}^{FB}, & \text{with probabiity } (1-p_1)\,(1-v) \end{cases}.$$

Thus, income per capita is

$$E\left(Y^{FB}\right) = (1-p_1)\left[vS\left(\underline{q}^{FB}\right) + (1-v)S\left(\bar{q}^{FB}\right)\right].$$

Let $\hat{Y}^{FB}$ be normalized income distribution under the first-best contract,

$$\hat{Y}^{FB} = \frac{Y^{FB}}{E\left(Y^{SB}\right)}.$$

$\hat{Y}^{FB}$ can be viewed as a mixture of distributions,

$$\left[\hat{Y}^{SB}\middle|\,\Psi = a\right] = \frac{W^{FB}}{E\left(Y^{SB}\right)}$$

and

$$\left[\hat{Y}^{SB}\middle|\,\Psi = p\right] = \frac{\delta^{SB}}{E\left(Y^{SB}\right)} = \frac{1-\gamma}{p_1}.$$

Under Assumptions 1 and 2, we have $W^{FB} \succeq_L W^{SB}$.

Theorem 5 *Under Assumptions 1 and 2, we have* $Y^{FB} \succeq_L Y^{SB}$.

Proof: From

$$W=\left\{\begin{array}{ll}\underline{t}^{SB}, & \text{with probability } v \\ \bar{t}^{SB}, & \text{with probabiity } 1-v\end{array}\right. ,$$

we know that $E\left(W\right)=\gamma\left[vS\left(\underline{q}^{SB}\right)+(1-v)S\left(\overline{q}^{SB}\right)\right]$. We also have $E\left(Y^{SB}\right)=(1-p_1)\left[vS\left(\underline{q}^{SB}\right)+(1-v)S\left(\overline{q}^{SB}\right)\right]$. Therefore, we have

$$\frac{E\left(W\right)}{E\left(Y^{SB}\right)}=\frac{\gamma}{1-p_1}.$$

Thus, we know that

$$\left[\hat{Y}^{SB}\middle|\,\Psi=a\right]=\frac{W}{E\left(Y^{SB}\right)}=\frac{E\left(W\right)}{E\left(Y^{SB}\right)}\frac{W}{E\left(W\right)}=\frac{\gamma}{1-p_1}\frac{W}{E\left(W\right)}.$$

Similarly, we have

$$\left[\hat{Y}^{FB}\middle|\,\Psi=a\right]=\frac{W^{FB}}{E\left(Y^{FB}\right)}=\frac{E\left(W^{FB}\right)}{E\left(Y^{FB}\right)}\frac{W^{FB}}{E\left(W^{FB}\right)}=\frac{\gamma}{1-p_1}\frac{W^{FB}}{E\left(W^{FB}\right)}.$$

From Theorem 3 and Proposition 5, we know that

$$\frac{W^{FB}}{E\left(W^{FB}\right)}\preceq_{cx}\frac{W^{SB}}{E\left(W^{SB}\right)}.$$

Thus, based on Corollary 3.A.22 in Shaked and Shanthikumar (2010), we have

$$\left[\hat{Y}^{FB}\middle|\,\Psi=a\right]\preceq_{cx}\left[\hat{Y}^{SB}\middle|\,\Psi=a\right].$$

Since $\left[\hat{Y}^{FB}\middle|\,\Psi=p\right]=\left[\hat{Y}^{SB}\middle|\,\Psi=p\right]=\frac{1-\gamma}{p_1}$, we have

$$\hat{Y}^{FB}\preceq_{cx}\hat{Y}^{SB},$$

from part (b) of Theorem 3.A.12 in Shaked and Shanthikumar (2010).

Since $E\left(\hat{Y}^{FB}\right)=E\left(\hat{Y}^{SB}\right)=1$, we know that

$$Y^{FB}\succeq_{L}Y^{SB},$$

from Proposition 5. ■