Journal of Comparative Economics, March 1992, 16, pp. 27-46

EQUITY, EFFICIENCY, AND INCENTIVES IN A LARGE ECONOMY[1]

Yingyi Qian

Stanford University, Stanford, CA 94305-6072, U.S.A.

Abstract

This paper analyzes the impact of egalitarianism on economic efficiency and individual incentives in an economy under uncertainty with many agents. The socialist state, with an inequality-averse social welfare function based on the agent's *ex post* utilities, is shown facing a tradeoff between equity and efficiency when the agent's actions are not observable. However, it is possible that effort is over-supplied under egalitarianism due to an effect coming from a change in the relative size of the different income groups. The analysis is also related to excessive profit-levelling among state-owned enterprises in the reforming socialist economies.

Journal of Economic Literature Classification Numbers: D63, D82, P20

1. INTRODUCTION

The socialist ideology of egalitarianism is well-known. Concern, not only for economic efficiency, but also for equity in distribution, has a profound impact on people's behavior. Equity, like freedom and efficiency, is often conceived as one of the fundamental values of human society. However, egalitarianism can also be utilized by politicians to rally political support or maintain social stability.

The potential conflict, or tradeoff, between equity and efficiency has been known for a long time.[2] It is only recently, however, that economists have been able to analyze the issue rigorously thanks to the development of information and incentive theory. Indeed, the conflict is possible only if there exists imperfect information and uncertainty at some point, because remuneration for performance serves two roles in such a situation: as an incentive scheme on the one hand and as income distribution on the other.[3]

This paper analyzes the impact of egalitarianism on economic efficiency and individual incentives under moral hazard in an economy with infinitely many agents so that the Law of Large Numbers can be applied. The egalitarian preference of the state is represented by an increasing and concave function G, defined on the individual's *ex post* utility of all agents in the economy.[4] The equity issue addressed here is the *ex post* equity or equity on results. Unlike many models on the subject restricted within the framework of partial equilibrium, our model is general equilibrium in nature.[5]

In order to highlight the difference between *ex ante* and *ex post*, we make an assumption that all agents start from an identical position *a priori*, and we choose to model a problem of moral hazard instead of adverse selection. Our purpose is to make the following point: even if everyone starts from an identical position *ex ante*, an unequal distribution may still arise *ex post*, because of incomplete insurance in the presence of moral hazard.

Following the tradition of incentive theory, we study the strategic interaction between the state and the agents where the state offers incentive schemes first and the agents take an unobservable action afterward. On the one hand, we do not place any advance restrictions on the form of incentive schemes that the state may choose beyond those required by the allocational and informational feasibility conditions. On the other hand, we enlarge the action space of the agent so that there is only one incentive scheme for the state implementing each action.

First, we show that in the model the possibility of conflict between equity and efficiency depends critically on the informational assumption about the agent's actions. There will be no conflict under perfect information, both full efficiency and absolute equity being achieved. Conflict can arise when actions of the agent are not observable by the state. Intuitively, the tradeoff occurs when the constrained efficient allocation calls for an incentive scheme with a large variance in remunerations and hence results in a large variance in welfare distribution, which the egalitarian state tries to avoid. Therefore, the reformers in socialist economies may face the dilemmas of equity versus efficiency when improving efficiency is their goal but the ideology of egalitarianism remains strong.

The model then is used to study the impact of egalitarianism on individual incentives for the supply of effort when effort is not observable. In our model, regardless of the risk attitude of the agent, it is necessary to use an incentive scheme with a wider spread of utility differentiation between good and bad outcomes in order to induce a higher effort. Nevertheless, we find that egalitarianism does not necessarily lower the incentives for the supply of effort from the agent relative to that under utilitarianism. The reason for this seemingly counter intuitive result is as follows. There are two effects on welfare distribution by implementing the higher effort levels from the agent in the large economy. The first is the known effect, which leads to a less equal welfare distribution due to greater rewards for success and heavier punishment for failure in order to satisfy the incentive compatibility condition. However, when the Monotonic Likelihood Ratio Property is satisfied, increased effort also raises the probability of success, and then the Law of Large Numbers implies that the size of the high-income group increases and the size of the low-income group diminishes. This second size effect may lead to a more egalitarian distribution.

Finally, we interpret the agents as state-owned enterprises and relate our model of egalitarianism to an empirical observation of excessive profit-levelling across the enterprises in the reforming socialist economies.

The paper is organized as follows. Section 2 introduces the model. Section 3 derives properties of the optimal incentive schemes. Section 4 examines the trade-off between equity and efficiency. Section 5 analyzes the impact of egalitarianism on the agent's incentive for the supply of effort. Section 6 relates the model to the issue of profit-levelling.

2. THE MODEL

We consider a large economy consisting of an infinite number of homogeneous economic agents. The state is endowed with legal power to promulgate and enforce contracts between the state and all economic agents in the economy. After the state announces the contract, agents take action independently. The state is assumed not be able to observe the agents' actions. The state does, however, observe the outcome of these actions. The outcome could be regarded as output, profit, or revenue. For the sake of simplicity, we use revenue throughout the paper. Following Grossman and Hart (1983), we assume that there are only n possible revenue levels, denoted as a vector $\mathbf{x}=(x_1,...,x_n)$ where $x_1>x_2>...>x_n$. Let A be the set of actions available to the agent. Let $P=\{\mathbf{p}=(p_1,...,p_n)|p_i\geq 0,\ \Sigma_1^n p_i=1\}$ be the unit simplex of dimension n-1. We assume that there is a continuous function $\mathbf{p}$: $A\rightarrow P$, where $\mathbf{p}(\mathbf{a})=(p_1(\mathbf{a}),...,p_n(\mathbf{a}))$ gives the probabilities of the n outcomes $x_1,...,x_n$ if action $\mathbf{a}$ is selected. The randomness is assumed to be independent across agents. The agent knows the probability function $\mathbf{p}$ but not the revenue level that results from his action. Let $Q=\mathbf{p}(A)$, a subset of P. The choice of action $\mathbf{a}$ implies the unique distribution $\mathbf{p}(\mathbf{a})\in Q$. Given that the function $\mathbf{p}(\mathbf{a})$

is common knowledge, a choice of $\mathbf{a}\in A$ is equivalent to a choice of $\mathbf{p}\in Q$. Therefore we can view **p** itself as the agent's action.

For example, only two possible outcomes exist such that x_1 is the revenue for success and x_2 the revenue for failure. Let a be a measure of agents' effort, which is one dimensional. Let $\mathbf{p}(a)=(p(a),1-p(a))$, where $0\leq p\leq 1$, and $p'(a)>0$. Then p represents the probability of success, that is, revenue level x_1. Since the higher the level of agent's effort, the higher the probability of success, we may regard p itself as the effort level.

The agents are all employed by the state, and we assume they have identical von Neumann-Morgenstern utility function $u(y)-v(\mathbf{p})$ which depends both on his income y received from the state and on his action **p**. Here, u is an increasing and concave function, and $v(\mathbf{p}) = v(p_1,...,p_{n-1})$ represents any disutility that the agent suffers by taking action **p**. If the agent is paid y_i on outcome x_i, then the agent chooses action $\mathbf{p}\in Q$ to maximize $\Sigma_1^n\{p_i u(y_i)\}-v(\mathbf{p})$, where $p_n=1-\Sigma_1^{n-1}p_i$.

Since **a** or $\mathbf{p}(\mathbf{a})$ is not observable by the state, it is impossible for the state to pay the agents based on **a** or **p**. Instead, the state will pay the agent according to outcome **x** of his action. An incentive scheme is therefore an n-dimensional vector $\mathbf{s}=(s_1,...,s_n)$, where s_i is the agent's remuneration in the event the agent's revenue is x_i.

The preferences of the state is such that *ex ante* the same weight is given to each agent in the evaluation of social welfare, but *ex post* an equal distribution of utilities among agents is favored. To that end, an increasing and concave function G is introduced so that if a fraction q_i of agents has an *ex post* utility level U_i, where i runs from 1 to m, $q_i\geq 0$, and $\Sigma_1^m q_i=1$, then the state will evaluate social welfare as

$\Sigma_1^m q_i G(U_i)$. In a special case when G is linear, the state has the utilitarian social welfare function, paying no attention to *ex post* equity. In particular, if all agents take the same action **p**, and if the incentive scheme **s** is chosen by the state, exactly fraction p_i of agents generate revenue x_i *ex post*, and therefore receive payment s_i, according to the Law of Large Numbers.

In such a case, social welfare is given by $\Sigma_1^n\{p_iG(u(s_i)-v(\mathbf{p}))\}$. Because of the independence of the randomness, the state faces no uncertainty at all in this large economy.

The constraint we impose on the choice set of the state is the requirement that the state not run a deficit. Therefore, $\mathbf{sp}\le\mathbf{xp}$ must be satisfied. This is indeed a minimum requirement whereas the feasible set of the state is quite large. From the equation $\mathbf{sp}\le\mathbf{xp}$, we see a second role of **s** in addition to that of incentive scheme: the role of income redistribution among all agents in the economy. It is the dual roles of **s**, i.e., the role of an *ex ante* contract between the state and the agents across the state of nature, and the role of the income distribution of *ex post* realization of uncertainty across the agents of the population, that is the cause of the conflict between equity and efficiency.[6]

The problem of the state can be formulated into the following optimization problem:

$$\max_{\mathbf{s},\ \mathbf{p}\in Q} \quad \Sigma_1^n\, p_iG(u(s_i)-v(\mathbf{p})) \tag{2.1}$$

$$\text{s.t.} \quad \text{(i) } \mathbf{sp}\le\mathbf{xp}$$

$$\text{(ii) } \Sigma_1^n\{p_iu(s_i)\}-v(\mathbf{p})\ge\Sigma_1^n\{p_i'u(s_i)\}-v(\mathbf{p}'), \text{ for all } \mathbf{p}'\in Q$$

Constraint (i) is a feasibility condition that requires that the state not run a deficit. Constraint (ii) is an incentive compatibility condition. In comparison, a classical model of the owner-manager relationship (e.g., Grossman and Hart (1983)) is formulated as follows:

$$\max_{\mathbf{s},\ \mathbf{p}\in Q} \quad \Sigma_1^n\{p_i(x_i-s_i)\} \tag{2.2}$$

$$\text{s.t.} \quad \text{(i) } \Sigma_1^n\{p_iu(s_i)\}-v(\mathbf{p})\ge\bar{u}$$

$$\text{(ii) } \Sigma_1^n\{p_iu(s_i)\}-v(\mathbf{p})\ge\Sigma_1^n\{p_i'u(s_i)\}-v(\mathbf{p}'), \text{ for all } \mathbf{p}'\in Q$$

where $\bar{u}$ is the reservation utility of the manager.

There are two major differences between our model (2.1) and the model of the owner-manager relationship (2.2). First, our model is not one of partial equilibrium. Since agents in our model are not allowed to quit, the usual individual rationality constraint arising from the partial equilibrium analysis is not imposed. Instead, the global condition that the total payments to the agents cannot be greater than the total revenues generated in the economy is included. Second, it is assumed that the state is not a profit maximizer. In fact, the social welfare function here does not assign any weight to profits. The state only evaluates *ex post* utilities of the agents and nothing more.

In order to solve the model, it is instructive to decompose the main problem (2.1) into two sub-problems. First, we find schemes to implement an action **p**, called implementation, by solving the following problem for a given $\mathbf{p} \in Q$:

$$\text{(i) } \mathbf{sp} \leq \mathbf{xp} \tag{2.3}$$
$$\text{(ii) } \Sigma_1^n\{p_i u(s_i)\} - v(\mathbf{p}) \geq \Sigma_1^n\{p_i' u(s_i)\} - v(\mathbf{p}'), \text{ for all } \mathbf{p}' \in Q$$

Let the solutions be the correspondence $\mathbf{S}(\mathbf{p})$. In general, **p** may or may not be implementable, therefore, $\mathbf{S}(\mathbf{p})$ may be non-empty or empty. In the case when **p** is implementable, $\mathbf{S}(\mathbf{p})$ can be either single valued or multiple-valued.

In the second step, called optimization, the following problem is solved by searching for the best **p** and $\mathbf{s}(\mathbf{p}) \in \mathbf{S}(\mathbf{p})$:

$$\max_{\mathbf{p} \in Q,\ \mathbf{s}(\mathbf{p}) \in \mathbf{S}(\mathbf{p})} \Sigma_1^n\{p_i G(u(\mathbf{s}_i(\mathbf{p})) - v(\mathbf{p}))\}. \tag{2.4}$$

We note that the state budget constraint $\mathbf{sp} \leq \mathbf{xp}$ is always binding at solutions of (2.1). If this were not true, then by marginally increasing payments to agents for each possible revenue achieved, and differential amounts may be needed to achieve this, the new payment scheme would still be feasible, and the incentive constraint would still be satisfied. As a result, the utilities of the agents would increase. Because G is an increasing function in u, social welfare would also increase. That implies that the original scheme would not be the solution to (2.1). Therefore, in the following discussion, the budget constraint $\mathbf{sp} \le \mathbf{xp}$ is replaced by the balanced budget constraint $\mathbf{sp} = \mathbf{xp}$.

3. PROPERTIES OF THE OPTIMAL INCENTIVE SCHEME

Many useful insights will be obtained by analyzing properties of incentive schemes implementing a given action, which can be achieved by working out the first step, implementation, alone.

Let Q be a k-dimensional manifold in P. For any given action $\mathbf{p} \in Q$ chosen by the agent, in order to select incentive schemes implementing $\mathbf{p}$, the state chooses solution $\mathbf{s}$ to (2.3). Since $\mathbf{s}$ is an element in an n-dimensional space, solutions to (2.3) are generally of dimension (n-k-1), where 1 represents the additional constraint $\mathbf{sp} = \mathbf{xp}$. It is usually assumed that the agent's feasible set of action A is 1-dimensional, e.g., a is effort, or k=1. If n>2, there is typically an (n-2)-dimensional continuum of solutions to (2.3), that is, (n-2)-dimensional continuum incentive schemes to implement the particular action $\mathbf{p} \in Q$. As the dimension of the agent's action space grows from 1, the dimension of the state's options for implementing a given behavior becomes correspondingly more limited. At the extreme, when the agent's action space is of full dimension, that is, k=n-1, all incentive schemes to implement a given action will be of dimension 0.[7]

Increasing dimensionality of the agent's action space to k=n-1 is one way of obtaining a definite form of the incentive scheme selected by the state. In fact, we will show that under certain conditions, every action $\mathbf{p}$ of the agent is not only implementable by the state but also implementable in a unique way:[8]

Lemma 1. Assume (i) Q has a non-empty interior in P; (ii) u(x) is an increasing and concave function defined on $(-\infty,+\infty)$ such that $u(+\infty)=+\infty$ and $u(-\infty)=-\infty$; and (iii) $v(\mathbf{p})$ is continuously differentiable and convex on Q. Then
(1) for any action **p** in the interior of Q, there exists a unique incentive scheme **s** which implements **p**;
(2) the incentive scheme $\mathbf{s}(\mathbf{p})$ is continuous in action **p**; and
(3) monotonicity of the partial marginal disutility $\partial v(\mathbf{p})/\partial p_1 \geq \ldots \geq \partial v(\mathbf{p})/\partial p_{n-1} \geq 0$ implies monotonicity of the incentive scheme: $s_1 \geq s_2 \geq \ldots \geq s_n$.

Proof: See Appendix.

When there are only two outcomes, the incentive scheme implementing action $\mathbf{p}=(p,1-p)$ can be characterized in more detail. Applying part (3) of Lemma 1 to the case of two outcomes, i.e., n=2, the condition $v'(p) \geq 0$ implies $s_1 \geq s_2$.[9] In fact, there exist in such a case other monotonic relationships between the effort level of the agent and the related incentive scheme chosen by the state, as shown in the following propositions.

Proposition 1. Let the incentive scheme $\mathbf{s}=(s_1,s_2)$ implement $p \in (0,1)$ and incentive scheme $\mathbf{s}'=(s_1',s_2')$ implement $p' \in (0,1)$. Then $p \leq p'$ implies

$$u(s_1) - u(s_2) \leq u(s_1') - u(s_2').$$

Proof: See Appendix.

Proposition 1 implies that implementing a higher effort level always requires a greater discrepancy of utilities from incomes for success and for failure, regardless of risk attitude and the shape of the disutility function $v(\mathbf{p})$. Notice that this property is derived exclusively from the incentive compatibility condition. Therefore, a lower effort, which leads to a smaller probability of success, corresponds to a more lenient contract with a smaller spread of remuneration on two outcomes; and a higher effort, which leads to a larger probability of success, calls for a harsher contract with a larger spread.

Proposition 2. Let the incentive scheme $\mathbf{s}=(s_1,s_2)$ implement $p\in(0,1)$ and let $\mathbf{x}=(x_1,x_2)$ implement $p^a\in(0,1)$. p^a is called self reliant action. Then,
(1) implementing any lower effort $p\leq p^a$ requires a smaller reward ($s_1\leq x_1$) and a smaller punishment ($s_2\geq x_2$); and
(2) implementing any higher effort $p\geq p^a$ requires a greater reward ($s_1\geq x_1$) and a greater punishment ($s_2\leq x_2$).

Proof: See Appendix.

Corollary: For any action pair p and p' satisfying $p\leq p^a\leq p'$, implementing the higher effort p' requires both a higher reward for success ($s_1'\geq s_1$), and a greater punishment for failure ($s_2'\leq s_2$).

Here the reference point is p^a, which can be viewed as the hands-off incentive scheme, because no taxes or subsidies will be provided under any circumstance ($\mathbf{s}=\mathbf{x}$). To implement $p\geq p^a$, Proposition 1 tells us that $u(s_1)-u(s_2)\geq u(x_1)-u(x_2)$. Proposition 2 says more: the reward for success s_1 must be actually higher than x_1 and punishment for failure must be greater so that $s_2\leq x_2$. Therefore, whenever the state in this model starts to tax successful agents and to subsidize unsuccessful ones, the state can only implement an effort level lower than p^a. In order for the state to implement an effort from the agent higher than p^a, the state must do the opposite: subsidize the successful agents and tax the unsuccessful ones.

The Corollary characterizes the incentive schemes to implement any two effort levels that are separated by the self reliance action p^a. What about the two effort levels that are both greater or smaller than p^a?

Proposition 3. Let the incentive scheme $\mathbf{s}=(s_1,s_2)$ implement $p\in(0,1)$ and the incentive scheme $\mathbf{s}'=(s_1',s_2')$ implement $p'\in(0,1)$. Then,

(1) if $p\leq p'\leq p^a$, implementing the higher effort p' needs a higher reward for success ($s_1'\geq s_1$); and

(2) if $p^a\leq p\leq p'$, implementing the higher effort p' needs a greater punishment for failure ($s_2'\leq s_2$).

Proof: See Appendix.

Proposition 3 implies that in order to implement a higher effort than the current one from the agents, additional reward is more effective when the current reward is low relative to the reference point x_1; on the other hand, additional punishment is more desirable when current reward is already high relative to the reference point x_1.

When the agent is risk neutral (i.e., u(y)=y), the optimal incentive scheme that implements **p** can be easily solved in a closed form from equations (A.1) and (A.2) in the Appendix. For example, when n=2, $s_1(p)=px_1+(1-p)x_2+(1-p)v'(p)$ and $s_2(p)=px_1+(1-p)x_2-pv'(p)$. This formula can be interpreted as an average income $(px_1+(1-p)x_2)$ plus a bonus (1-p)v'(p) in the event of success (x_1), and an average income with a punishment equals to pv'(p) in the event of failure (x_2). Furthermore, the difference between the bonus and the punishment should be exactly $s_1-s_2 = v'(p)$. That is to say, the marginal disutility of an additional unit of effort should be exactly equal to s_1 minus s_2.

4. EQUITY AND EFFICIENCY

An allocation is called efficient if there exists no other feasible allocation that makes one agent better off without making another agent worse off. Under alternative informational assumptions, efficiency depends on the set of informationally feasible allocations. Under full efficiency, the first best, information is perfect and any physically feasible allocation is informationally feasible. Under constrained efficiency, the second best, a feasible allocation must be both physically feasible and incentive compatible. In our model, a constrained efficient allocation is then defined as action $\mathbf{p}^c$, implemented by $\mathbf{s}^c$, such that no other solution $(\mathbf{p},\mathbf{s})$ to (2.3) gives higher expected utility to the agents.

Lemma 2. Let Q be compact, such that for all $\mathbf{p}\in Q$, $p_i>0$ $1\leq i\leq n$. Then,
(1) a full efficient allocation exists, and
(2) a constrained efficient action $\mathbf{p}^c$ and the associated incentive scheme $\mathbf{s}^c$ also exists.

Proof: See Appendix.

Consider first the case of perfect information. We show that both efficiency and absolute equity can be achieved by the state. Therefore, there is no conflict or tradeoff between equity and efficiency when action is observable:

Proposition 4. Assume G is an increasing and concave function. Then if the agent's action is observable by the state, the state will command a full efficient allocation $\mathbf{p}^f$ with absolute equality of income, $\mathbf{xp}^f$ for everyone.

Proof: See Appendix.

Now consider the case of imperfect information where the agent's action is not observable. First note that in such a case a fully efficient allocation is not sustainable under any system if u is strictly concave. This is because when agents are risk averse, a fully efficient allocation calls for a complete insurance so that the agent's remuneration is independent of outcome, that is, $s=s_1=s_2=...=s_n$. However under such a scheme, the agent's utility function becomes $u(s)-v(\mathbf{p})$, so he will minimize his disutility $v(\mathbf{p})$

which makes him choose an action other than $\mathbf{p}^f$ in general.

But the constrained efficient allocation can be supported by a hypothetical market in our model when action is not observable. Imagine that there are many insurance companies offering insurance policies **s**. Assume that the insurance companies are pure expected profit-maximizers in perfect competition. Then their objective function is E(**x**-**s**)=**xp**-**sp** and each earns zero expected profits in equilibrium. With imperfect information, an insurance company solves the following problem by choosing **p** and **s**:

$$\max_{\mathbf{s},\ \mathbf{p}\in Q} \ (\mathbf{xp}-\mathbf{sp})$$
$$\text{s.t.}\quad \Sigma_1^n\{p_i u(s_i)\}-v(\mathbf{p})\geq\Sigma_1^n\{p_i' u(s_i)\}-v(\mathbf{p}'),\ \text{for all } \mathbf{p}'\in Q$$

Perfect competition implies that **xp**-**sp**=0. It is easy to see that in equilibrium, (**p**,**s**) is a constrained efficient allocation. Therefore, in this model, the hypothesis of pure expected profit maximizing is consistent with the constrained efficiency criterion even when imperfect information is present.[10]

In comparison, the state chooses the optimal $\mathbf{p}\in Q$ to solve (2.4), namely, to carry out the second step of optimization.

Proposition 5. Assume action is not observable. Then,
(1) an optimal solution of the state still exists; and
(2) if u^p denotes the expected utility of the agent under the optimal choice by the state, and u^c represents the expected utility of the agent under conditions of constrained efficient allocation, then $u^p\leq u^c$.[11]

Proof: See Appendix.

When actions of agents are not observable, Proposition 5 shows that the egalitarian state may implement an inefficient action even when the informational constraint is taken into account. That is, the efficiency loss may be additional to that caused by imperfect information in a utilitarian world, and the optimal choice of the state may not be the second best solution as in a comparable model of the owner-manager relationship.

The intuitive reason for the conflict between efficiency and equity under the circumstance of imperfect information can be explained as follows. In the Taylor approximation of function G, the first two terms are the mean and variance of the agents' *ex post* utility distribution. Since G is concave, the variance is entered as a negative term. The tradeoff between the mean and the variance is the main reason for the state to face the conflict between efficiency and equity. When information is perfect, we can obtain the highest mean, without any variance, so that both efficiency and equity are obtained. When action is not observable, an incentive scheme is needed to induce the mean in the constrained efficient allocation, but likely accompanying a higher variance. If the state is utilitarian, i.e., G is linear, the variance does not enter the picture, so only the mean of the *ex post* utility distribution matters. But that coincides with the expected utility of the agent, hence efficiency is maintained. If the state is egalitarian, i.e., G is concave, the high variance of welfare distribution is evaluated negatively. The state may prefer a welfare distribution with a low variance, even when that implies a possible low mean, implying a loss of efficiency.

We have shown that *ex post* egalitarianism leads to an *ex ante* loss for everyone. Therefore, *ex ante* every agent prefers a constrained efficient allocation to an allocation chosen by the egalitarian state. Apparently, the Pareto principle of social choice, i.e., if every one in the society prefers A to B, then the society prefers A to B, is violated in a certain sense.[12] This illustrates a point that, in an uncertain environment, the perspectives of *ex ante* and *ex post* are different. If individuals evaluate their utilities in *ex ante* expected terms while the state or society cares about *ex post* equality in distribution, then the tension between *ex ante* efficiency and *ex post* equity may arise.[13]

5. EGALITARIANISM AND INCENTIVES

The result of the efficiency loss due to egalitarianism is not surprising. However, the direction of the distortion, that is, the effects of egalitarianism on the agent's incentive for the supply of effort has not been analyzed. Intuition seems to suggest that equity consideration should reduce people's incentive to work. But this intuition is not quite right. It is possible that effort is over-supplied under egalitarianism, as we show below.

We assume that n=2 throughout this section, since effort is measured in one dimension. Also for the sake of simplicity, we assume risk-neutrality of the agent, which implies that the first best and the second best coincide, and that efficient allocation p^a is achieved by the hands-off incentive scheme $\mathbf{x}=(x_1,x_2)$, such that $v'(p^a)=x_1-x_2$ or $p^a=1$ if $v'(1)<x_1-x_2$.[14]

Starting with the first best effort level $p^a \in (0,1)$ with no efficiency loss, we increase the effort by a tiny amount $dp>0$, to $p=p^a+dp$. By Proposition 2, the incentive scheme $\mathbf{s}=(s_1,s_2)$ implementing p must satisfy $s_1 \geq x_1$ and $s_2 \leq x_2$. That is, for this slightly higher effort, more reward for success and greater punishment for failure are required. Note that moving away from p^a by dp does not change the mean of the welfare distribution, since p^a maximizes the mean. The intuitive argument goes as follows: Since the new income difference between the two income groups is s_1-s_2, which is larger than x_1-x_2, any egalitarianism would prefer p^a to p and effort will never be increased in the neighborhood of p^a.

The above reasoning is incomplete, because it implicitly takes the size of each income group as fixed and considers only the effect on income discrepancy, which only makes the variance larger. Suppose a fraction p of people receives s_1 and the remaining (1-p) people receive s_2, and $s_1-s_2 \geq x_1-x_2$. Then the variance of income distribution is given by

$$ps_1^2+(1-p)s_2^2-(ps_1+(1-p)s_2)^2$$
$$= p(1-p)(s_1-s_2)^2,$$

which is larger than $p(1-p)(x_1-x_2)^2$, the variance of income for the same fraction p of people who receives x_1 and the remaining (1-p) who receive x_2.

However, as we increase the effort level from p^a to $p=p^a+dp$, by the Law of Large Numbers, we also increase the size of the high income group from p^a to $p=p^a+dp$.[15] Here is the second effect, the effect on the size of income groups comes into the picture. Since more people receive higher income and less people remain in the low income group, it is possible that this effect makes the variance of the income distribution smaller. If the second effect is indeed good on income distribution, and furthermore, if it is strong enough to offset the first bad effect of the increased income discrepancy between the two income groups, the variance of income distribution will be smaller by implementing a higher effort.

To illustrate the point clearly, imagine only two effort levels: High and Low. Assume that Low gives slightly higher mean utility net of effort than High, so Low is more efficient. Assuming further that implementing Low needs an incentive scheme that leaves half of the population in the higher income group and another half in the low income group, and thus there will be an income discrepancy. It is also assumed that implementing High requires a much larger reward for success, as well as a heavy punishment for failure; however, if High is implemented, the project will be successful for sure, i.e., the project is safe. Thus, by the Law of Large Numbers, everyone takes a high reward by exerting effort High, therefore everyone is in the high income group and no one receives heavy punishment.[16] Apparently, society is more egalitarian if High is implemented than if Low is implemented, and the egalitarian state will choose High rather than Low, although the latter is more efficient. Here the second effect of the higher effort is so strong that it virtually merges the two income groups all the way into one, making everyone absolutely equal.[17]

Therefore, the effect on the relative size of income groups can not be neglected in analyzing incentive effects of egalitarianism, despite the fact that the intuitive argument usually goes without such a consideration.[18] Although effort may be over-supplied under egalitarianism, the efficiency loss is still present because of the lower expected utility of individual agents. The higher effort, therefore the harsher contract, is used in order to reduce the variance of remuneration across different outcomes and the variance of welfare distribution by increasing (or decreasing) the size of the high (or low) income group in the population.

In what follows, we analyze two cases of function G: the maximin case and the quadratic case. We show that with the maximin preference of the state, the effort level exerted by the agent is unambiguously reduced relative to p^a. However, with the quadratic preference of the state, both situations exist; one in which effort is reduced, and one in which effort is over-supplied, relative to p^a.

The Maximin Case

The maximin state maximizes[19]

$$pG(s_1-v(p))+(1-p)G(s_2-v(p))$$
$$=\min\{s_1-v(p),s_2-v(p)\}.$$

Substituting the result obtained in Section 3, social welfare becomes

$$W = \min\{s_1-v(p),s_2-v(p)\}$$
$$=s_2-v(p)$$
$$=px_1+(1-p)x_2-pv'(p)-v(p)$$
$$=\{px_1+(1-p)x_2-v(p)\}-pv'(p).$$

Since p^a maximizes $px_1+(1-p)x_2-v(p)$, evaluated at $p=p^a$,

$$dW/dp = -d(pv'(p))/dp < 0.$$

Therefore, with the maximin preference of the state, the effort level exerted by the agent is unambiguously reduced.[20] This is because in such a case only the utility of the worst treated person is evaluated, and the size of income groups is never taken into consideration by the egalitarian state. As higher effort requires greater punishment, the state will avoid higher effort as long as there is some one ($p^a \neq 0$) to be punished more.

The Quadratic Case

With G being a quadratic function, the Taylor approximation of G becomes precise, the state is only sensitive to the mean and the variance of *ex post* utilities of the agent. When the agent is risk neutral, the conflict between the state and the agent arises entirely from the part of the variance of incentive schemes, or the variance of *ex post* utilities among the population.

Let $G(y)=(b/2)y^2+y+b$. Then $G'(y)=by+1$, and $G"(y)=b$. Assume $b<0$, and $by+1>0$ in the relevant region. Then G is increasing and concave. From

$$G(y)=G(Ey)+G'(Ey)(y-Ey)+(1/2)G"(Ey)(y-Ey)^2,$$

where E is the expectation operator, we obtain

$$EG(y)=G(Ey)+(1/2)G"(Ey)E(y-Ey)^2.$$

This implies social welfare

$$\begin{aligned} W &= G(Es-v(p))+(b/2)\bullet Var(s) \\ &= G(ps_1+(1-p)s_2-v(p))+(b/2)\bullet(s_1-s_2)^2 \\ &= G(px_1+(1-p)x_2-v(p))+(b/2)\bullet p(1-p)v'(p)^2. \end{aligned}$$

Therefore,

$$\begin{aligned} dW/dp = {} & G'(px_1+(1-p)x_2-v(p))(x_1-x_2-v'(p))+ \\ & +(b/2)\bullet[(1-2p)v'(p)^2+2p(1-p)v'(p)v"(p)]. \end{aligned}$$

Since p^a maximizes $px_1+(1-p)x_2-v(p)$, the first term evaluated at $p=p^a$ vanishes, the sign of dW/dp at $p=p^a$ is the same as the sign of the second term. Because p^a is determined by $v'(p^a)=x_1-x_2$ alone, p_a may lie anywhere between 0 and 1.

We first show that under certain conditions it is possible for effort to be reduced, relative to p^a due to egalitarian considerations. Consider the case $p^a\leq 1/2$. Then it is clear that $dW/dp<0$ at p^a because $b<0$. This means that effort should always be reduced locally due to egalitarian considerations when only a minority of the population receive the higher income at the first best allocation.[21] This is because by increasing the effort level, the size of the smaller high income group increases, which can only increase the variance of welfare distribution, other things being equal. As the only way to get good effect from increasing effort is blocked, the effort must be reduced even though that means more people will be put into the low income group and be punished.

We next show that under alternative conditions it is also possible for effort to be over-supplied relative to p^a due to egalitarian considerations. Consider the case $p^a>1/2$. If $p^a v''(p^a)/v'(p^a)<(2p^a-1)/(2(1-p^a))$, then $dW/dp>0$ at $p=p^a$. A sufficient condition for the above inequality to be satisfied is that $v''(p)$ is bounded and p^a is close to 1. Under such a condition, only a handful people are in the low income group at the first best, then, by increasing effort, although income discrepancy becomes larger, some people in the tiny group of low income will be transferred to the dominating group of high income, the welfare distribution actually improves.

6. CONCLUDING REMARKS

A generic term agent is used in the paper. If an individual citizen is used for agent, the model is naturally interpreted as that of optimal income taxation. Another option is the manager of state-owned enterprises in socialist economies. A third possible interpretation of agent is the state-owned enterprise itself in socialist economies, which represents workers and managers.[22] With this interpretation, our model may be related to the empirical observation of excessive profit-levelling across state-owned enterprises in the reforming socialist economies.[23]

It is sometimes suggested that egalitarian consideration is one reason for profit-levelling in socialist economies.[24] From the above analysis, we know that the egalitarian preference of the state over the *ex post* distribution of profits retained by the enterprises does not always imply profit-levelling. Only if we assume that the size effect of income groups alone does not reduce very much the variance as the effort increases, does our model, with a certain form of the social welfare function, imply that both reward and punishment will be reduced with the result that profit-levelling will occur. Despite the efficiency loss under egalitarianism, the variance of after-tax profits among state-owned enterprises as well as variance of workers' income and benefits received from their units could be smaller.

APPENDIX: MATHEMATICAL PROOFS

Proof of Lemma 1: (1) Because of the convexity of $v(\mathbf{p})$, for any $\mathbf{p}$ in the interior of Q, $\mathbf{s}$ implements $\mathbf{p}$ if and only if $\mathbf{s}$ solves the first order conditions:

$$\begin{cases} u(s_1) - u(s_n) = \partial v(\mathbf{p})/\partial p_1 \\ \quad \dots \; \dots \\ u(s_{n-1}) - u(s_n) = \partial v(\mathbf{p})/\partial p_{n-1} \end{cases} \qquad (A.1)$$

$$p_1 s_1 + \dots + p_n s_n = p_1 x_1 + \dots + p_n x_n \qquad (A.2)$$

Define $w_i = u(s_i)$, and let $h = u^{-1}$. Then, $s_i = h(w_i)$ and function h is increasing and convex. Substituting $w_i = w_n + (\partial v(\mathbf{p})/\partial p_i)$ into equation (A.2), we have

$$p_1 h(w_n + (\partial v(\mathbf{p})/\partial p_1)) + \dots + p_{n-1} h(w_n + (\partial v(\mathbf{p})/\partial p_{n-1})) + p_n h(w_n) = \mathbf{xp} \qquad (A.3)$$

Now consider the function

$$g(y) = p_1 h(y + (\partial v(\mathbf{p})/\partial p_1)) + \dots + p_{n-1} h(y + (\partial v(\mathbf{p})/\partial p_{n-1})) + p_n h(y).$$

Since $g' > 0$, $g(y)$ is strictly increasing. Moreover, $g(+\infty) = +\infty$ and $g(-\infty) = -\infty$ since $h = u^{-1}$. Therefore, there exits a unique solution w_n to equation (A.3). Because the relevant transformations are one-to-one, there exists a unique solution $\mathbf{s}$ to (A.1) and (A.2).

(2) The continuity of function $\mathbf{s}(\mathbf{p})$ follows from the uniqueness. Since equations in (A.1) are linear in w_i, it is possible to consider only equation (A.3). Suppose there were a discontinuity at some $\mathbf{p}$. Take two sequences $\{p_i\}$ and $\{p_i'\}$ leading to $\mathbf{p}$, but the corresponding limits of w_n are different. By the continuity of the underlying functions, (A.3) has two different solutions for a given $\mathbf{p}$, thus contradicting the result of uniqueness.

(3) The monotonicity of s_i follows from equations in (A.1).■

Proof of Proposition 1: From the incentive compatibility conditions,

$$pu(s_1) + (1-p)u(s_2) - v(p) \geq p'u(s_1) + (1-p')u(s_2) - v(p'),$$

and

$$p'u(s_1') + (1-p')u(s_2') - v(p') \geq pu(s_1') + (1-p)u(s_2') - v(p).$$

Adding up these two equations, we obtain

$$(p'-p)[u(s_1)-u(s_2)]\le(p'-p)[u(s_1')-u(s_2')].$$

Therefore, $p<p'$ implies

$$u(s_1)-u(s_2)\le u(s_1')-u(s_2').\blacksquare \tag{A.4}$$

Proof of Proposition 2: (1) Take any $p\le p^a$, and let $p'=p^a$ in (A.4), we have

$$u(s_1) - u(s_2) \le u(x_1) - u(x_2).$$

Because of the balance budget constraint

$$ps_1+(1-p)s_2 = px_1+(1-p)x_2,$$

it must be the case

$$s_1\le x_1 \text{ and } s_2\ge x_2.$$

A similar proof applies to part (2).■

Proof of Proposition 3: From the balance budget constraints,

$$ps_1+(1-p)s_2 = px_1+(1-p)x_2 \tag{A.5}$$

and

$$p's_1'+(1-p')s_2 = p'x_1+(1-p')x_2. \tag{A.6}$$

Subtracting (A.5) from (A.6), we get

$$p's_1'+(1-p')s_2'-ps_1-(1-p)s_2 = (p'-p)(x_1-x_2).$$

Rearranging terms, we obtain

$$p'(s_1'-s_1)+(1-p')(s_2'-s_2) =(p'-p)[(x_1-x_2)-(s_1-s_2)]. \tag{A.7}$$

Let $p\le p'\le p^a$. Proposition 2 implies $s_1-s_2 \le x_1-x_2$. Therefore, (A.7) implies

$$p'(s_1'-s_1)+(1-p')(s_2'-s_2) \ge 0. \tag{A.8}$$

Suppose $s_1'<s_1$, then (A.8) implies $s_2'>s_2$, contradicting to Proposition 1. Therefore, $s_1'\ge s_1$. A similar proof applies to $p^a\le p\le p'$.■

Proof of Lemma 2: (1) The full efficient action p^f is obtained by solving max $u(\mathbf{xp})-v(\mathbf{p})$ for $\mathbf{p}\in Q$. Since Q is compact, and the objective function is continuous, the existence of $\mathbf{p}^f$ is guaranteed.

(2) Since every $\mathbf{p}\in Q$ is in the interior of P, by Lemma 1, **s** is uniquely determined and is continuous in **p**. The constrained efficient action is

obtained by solving max $\Sigma_1^n\{p_i u(s_i(\mathbf{p}))\}-v(\mathbf{p})$ for $\mathbf{p}\in Q$. Since the objective function is continuous in $\mathbf{p}$, compactness of Q guarantees the existence of a constrained efficient allocation.■

Proof of Proposition 4: For any action $\mathbf{p}\in Q$ and any feasible payment schedule $(y_1,...,y_n)$, the utility of the state is evaluated by

$$\Sigma_1^n\{p_i G(u(y_i)-v(\mathbf{p}))\}\le G(\Sigma_1^n\{p_i u(y_i)\}-v(\mathbf{p}))\le G(u(\Sigma_1^n p_i y_i)-v(\mathbf{p}))\le G(u^f),$$

because of the concavity of u and of G.■

Proof of Proposition 5: The state solves the problem

$$\max\ \Sigma_1^n\{p_i G(u(s_i(\mathbf{p}))-v(\mathbf{p}))\},\ \text{for}\ \mathbf{p}\in Q.$$

Since G and **s** are continuous, compactness of Q implies that an optimal solution does exist. On the other hand, since

$$u^c=\max\ \Sigma\{p_i u(s_i(\mathbf{p}))\}-v(\mathbf{p})\ \text{for}\ \mathbf{p}\in Q,$$

which is maximized on the same set as the state's problem. Therefore the optimal solution of the state $(\mathbf{p}^p,\mathbf{s}^p)$ is also a feasible choice here, which implies $u^p\le u^c$.■

Footnotes

1. I am most grateful to Janos Kornai, Andreu Mas-Colell and Eric Maskin for their advice and encouragement. I also wish to thank Abihijit Banerjee, Oliver Hart, Amartya Sen, Michael Spagat, Martin Weitzman and Chenggang Xu for helpful discussions.

2. Okun's *Equality and Efficiency: The Big Tradeoff* (1975) is a popular book on the subject. Kornai (1986) studies conflicts between economic efficiency and the principles of socialist ethics in socialist economies. At a more philosophic level, Gordon (1980), among many others, provides a pluralist view on values of welfare, justice, and freedom, and illustrates possible complementarity and conflicts between them.

3. For examples, Conn (1982) and Murrell and Miller (1984) investigate the possible conflict using models of a central planner and a manager under the circumstance that the manager possesses some information the center does not.

4. Hammond (1983) argues for a welfare function based on the individual's *ex post* utility in order to obtain a dynamically consistent objective for collective choice under uncertainty. The static model here may be viewed as a reduced form of a dynamic model for the purpose of studying dynamically consistent policies of the state.

5. Our model in many aspects is similar to the models of optimal income taxation (e.g., Mirrlees (1971)). However, those are adverse selection models, where ability is known to the individual, but not to the state.

6. Note that, in our model, agents maximize expected utility where the concavity of u means risk-aversion across different states of nature. Meanwhile, the state maximizes social welfare (under certainty) where the concavity of G represents inequality-aversion across agents of the population.

7. See Holmstrom and Milgrom (1987) on this point.

8. In the remaining part of the paper, we will maintain the assumptions of Lemma 1 and always use the unique incentive scheme $\mathbf{s}(\mathbf{p})$ which implements $\mathbf{p} \in Q$.

9. It is worth noting that the Monotone Likelihood Ratio Property (MLRP) is satisfied automatically in this case, because actions $p'\leq p$ if and only if the likelihood ratios $(p'/p)\leq(1-p')/(1-p)$. We know already that MLRP would imply $s_1\geq s_2$ in a more general model.

10. However, with heterogeneous agents, a competitive equilibrium in general needs not to be constrained efficient.

11. When G is linear, i.e., inequality neutral, efficiency is obtained: $u^p=u^c$.

12. The optimal choice of the state is still *ex post* efficient. That is, after states of nature have been revealed, there is no other feasible allocation that makes one agent better off without making any other agent worse off.

13. One possible way to reconcile the tension is to internalize the externality of the inequality by redefining the individual's utility function. For example, Sen (1966) defines $U_i=\Sigma\gamma_j u_j$, where $\Sigma\gamma_j=1$ and $\gamma_j\geq 0$, in order to incorporate sympathy among people.

14. The second order condition is satisfied because $v(\bullet)$ is convex.

15. What is important here is the Monotone Likelihood Ratio Property (MLRP), which says that a higher effort makes success more likely to happen. Then, by the Law of Large Numbers, more people are successful and end up in the high income group.

16. It should be noted that although implementing a safe project always implies a safe remuneration and therefore implies an equal income distribution, the converse is not true at all: A sure remuneration may implement a very risky project. What is important in the consideration of income distribution is not the variance of the projects that are undertaken, but rather the variance of income distribution, which is equal to the variance of the underlying incentive scheme in our model.

17. Note that in this example, in order to give a sharp comparison, we have used p=1 or 1-p=0. But our argument does not depend on this. Therefore, our

argument here differs from the famous Mirrlees (1974) argument for non-existence of the second best, when the first best can be arbitrarily approached by applying larger and larger punishment with smaller and smaller probability ($p \Rightarrow 0$). In our model in the paper, this non-existence problem is ruled out by requiring Q to be compact and in the interior of P.

18. In fact, any inequality index used in reality is always related to the size of income groups, like the Gini Coefficient and the income shares of the bottom or the top 20 percent population.

19. Consider the narrow power function family: $G(x)=[1/(1-B)]x^{1-B}$, where $B \geq 0$, the relative risk aversion coefficient. When B=0, the linear function $G(x)=x$ is obtained. If B=1, $G(x)=\log(x)$. Let B go to $+\infty$, then G(x) is equivalent to the maximin function.

20. Since $d^2W/dp^2=-3v''-pv'''$, the second order condition is satisfied provided $v'''(p)>0$ for all $0<p<1$.

21. A sufficient condition for effort to be reduced globally is $pv''(p)/v'(p) \geq 1/(2(1-p))$ for all $p \geq 1/2$. This is because $G'(x_1-x_2-v'(p))$ is always negative for $p \geq p^a$, and $(b/2)[(1-2p)v'(p)^2+2p(1-p)v'(p)v''(p)]$ is non-positive for $p \geq p^a$ under the given condition, therefore, $dW/dp \leq 0$ for all $p>p^a$.

22. It is well-known that enterprises in socialist economies are not only work places, but also welfare organizations. To a very large extent, workers' well-being depends on the welfare benefits they receive from the enterprise. The care for distribution of profits of the enterprises is an indirect way of caring for the distribution of workers' income. According to this interpretation, v(**p**) is the aggregated disutility of all workers in the enterprise.

23. In China, there officially exists an adjustment tax with the explicit purpose of equalizing after-tax-profits between enterprises (Naughton (1985)). In Hungary, data shows very little correlation between a firm's pre-tax profitability and its profitability after levelling (see Table 3 on page 233 in Kornai (1983)).

24. Some people are doubtful, and correctly so, that a government in a socialist economy really pursues egalitarianism economywide. Although it is true that in socialist economies the non-monetary income distribution is much less equal than monetary income distribution, that occurs mainly along the vertical lines of the hierarchy. Since our interpretation of agent here is the state-owned enterprise, which is positioned horizontally rather than vertically in the hierarchy, egalitarianism may still be a relevant issue.

REFERENCES

Conn, David, "Effort, Efficiency, and Incentives in Economic Organizations." *Journal of Comparative Economics*, **6**, 3:223-234, September, 1982.

Gordon, Howard Scott, *Welfare, Justice and Freedom*. New York: Columbia University Press, 1980.

Grossman, Sanford, and Oliver Hart, "An Analysis of the Principle-Agent Problem." *Econometrica*, **51**, 1:7-45, January, 1983.

Hammond, Peter, "Ex Post Optimality as a Dynamically Consistent Objective for Collective Choice under Uncertainty." In Prasantak K. Pattanaik and Maurice Salles, Eds., *Social Choice and Welfare*, pp. 175-205. New York: North-Holland, 1983.

Holmstrom, Bengt, and Paul Milgrom, "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, **55**, 2: 303-328, March, 1987.

Kornai, Janos, "Efficiency and the Principles of Socialist Ethics." In Janos Kornai, *Contradictions and Dilemmas*, pp. 124-138. Cambridge: MIT Press, 1986.

Kornai, Janos, "Comments on the Present State and the Prospects of the Hungarian Economic Reforms." *Journal of Comparative Economics*, **7**, 3:225-252, September, 1983.

Mirrlees, James, "An Exploration in the Theory of Optimal Income Taxation." *Review of Economic Studies*, **38**, 175-208, April, 1971.

Mirrlees, James, "Notes on Welfare Economics, Information and Uncertainty." In M. Balch, D. MacFadden and S.Y. Wu, Eds, *Essays on Economic Behavior under Uncertainty*, pp. 243-261. New York: North Holland, 1974.

Murrell, Peter and Jeffrey Miller, "The Applicability of Information-Revealing Schemes in Economic Organizations." *Journal of Comparative Economics*, **8**, 3:277-289, September, 1984.

Naughton, Barry, "False Starts And Second Wind: Financial Reforms In the Chinese Industrial Systems." In Elizabeth Perry and Christine Wong, Eds, *The Political Economy of Reform in Post-Mao China*, pp. 223-252. Cambridge: Harvard University Press, 1985.

Okun, Arthur M, *Equality and Efficiency: The Big Tradeoff*. Washington: The Brookings Institution, 1975.

Sen, Amartya K., "Labor Allocation in a Cooperative Enterprise." *Review of Economic Studies*, **33**, 361-371, October, 1966.