The effect of Better Information on Income Inequality

Itzhak Zilcha and Bernard Eckwert

7-07

June, 2007

The Foerder Institute for Economic Research
and
The Sackler Institute of Economic Studies

The effect of Better Information on Income Inequality

Itzhak Zilcha and Bernard Eckwert

Abstract

We consider an OLG economy with endogenous investment in human capital. Heterogeneity in individual human capital levels in modelled by a distribution of innate ability across agents. This distribution is common knowledge but, at young age, no agent knows his/hers ability. The production of human capital depends on each individual's investment in education. This investment decision is taken only after observing a signal which is correlated to his/her true ability, and which is used for updating beliefs. Thus, a better information system affects the distribution of human capital in each generation. Assuming separable and identical preferences for all individuals, we derive the following results in equilibrium: (a) if the relative measure of risk aversion is less (more) than 1 then more information raises (reduces) income inequality. (b) When a risk sharing market is available better information results in higher inequality regardless of the measure risk aversion.

Keywords w: information system, income inequality, risk sharing markets

JEL D80 J24 J30

The Effect of Better Information on Income Inequality

by

Bernhard Eckwert and Itzhak Zilcha*

2nd revision March 2005

Bernhard Eckwert
Department of Economics
University of Bielefeld
Germany

Itzhak Zilcha
The Eitan Berglas School of Economics
and CESifo
Tel Aviv University
Israel

*We are grateful to Alex Cukierman, Zvi Eckstein, Yona Rubinstein and Daniel Tsiddon. An anonymous referee made extremely useful comments and suggestions. Financial support from the German-Israeli Foundation for Scientific Research and Development (GIF) is gratefully acknowledged.

Abstract

We consider an OLG economy with endogenous investment in human capital. Heterogeneity in individual human capital levels is modelled by a distribution of innate ability across agents. This distribution is common knowledge but, at young age, no agent knows his/her ability. The production of human capital depends on each individual's investment in education. This investment decision is taken only after observing a signal which is correlated to his/her true ability, and which is used for updating beliefs. Thus, a better information system affects the distribution of human capital in each generation. Assuming separable and identical preferences for all individuals, we derive the following results in equilibrium: (a) If the relative measure of risk aversion is less (more) than 1 then more information raises (reduces) income inequality. (b) When a risk sharing market is available better information results in higher inequality regardless of the measure risk aversion.

Keywords: Information system, income inequality, risk sharing markets.

JEL classification numbers: D80, J24, J30.

1 Introduction

In recent years we have witnessed growing interest of economists in the determinants of income inequality and its evolution in dynamic models. Many papers in this field focus on the role of education systems for the distribution of income. These papers have produced mixed results, both on the theoretical and empirical level, suggesting that better public education systems do not necessarily lead to less income inequality [e.g., Glomm and Ravikumar (1992), Sylwester (2002a,2002b)]. Another main issue relates to the effects of various institutional settings and macroeconomic policies on the distribution of income in equilibrium [see, for example, Loury (1981), Galor and Zeira (1993), Benabou (1996), Orazem and Tesfatsion (1997), Aghion (2002)].

The endogenous growth literature has investigated the causes of inequality in income distribution, concentrating on four main transmission channels. Firstly, differences in unobservable individual talent may generate income inequality [e.g., Juhn et al. (1993)]. Secondly, based on subjective assessments of individual talent agents may choose different levels of investment in education [e.g., Galor and Tsiddon (1997), Viaene and Zilcha (2002)]. Thirdly, the stock of human capital of parents may affect their children's learning. If this linkage is specific to the household it will contribute to income inequality [e.g., Hassler and Mora (2000)]. Fourthly, the rate of technological change may affect the return to investment in education. Rubinstein and Tsiddon (2002) show that this mechanism accounts for a significant part of the inequality between different education groups. Our paper focuses on the first two channels. We develop an approach which embeds the social mechanism of selecting individual education levels in an endogenous growth model where agents differ with respect to ability. When young, each agent is screened by an information system which provides him with imperfect information about his talent. The central issue of our study is how better information (i.e., more efficient screening in the education period) affects the intragenerational income distribution.

Over the last decades the literature in the field of the economics of information has put much emphasis on the issue how welfare effects of information are related to the market structure of an economy. Many studies have examined the value of information in various partial equilibrium models [e.g.,Blackwell (1951,1953), Green

(1981)] and in general equilibrium frameworks [e.g., Hirshleifer (1971,1975), Orosel (1996), Schlee (2001), Eckwert and Zilcha (2001)]. However, to the best of our knowledge, this paper contains the first attempt to study the impact of information systems on the distribution of income in an endogenous growth model.

Our analytical framework is an OLG economy where investment in education is done under uncertainty. Individuals in the same generation differ in their (random) innate abilities. When ability is still unknown each individual decides how much 'effort' to invest in his/her education and training. At this time, the return to this investment, in term of wages during the working period, is perceived as random since it depends on the realization of the ability. The effort decision is made after observing a signal which contains information about the agent's random ability. We analyze the effect of better information system, i.e., more efficient screening, upon the distribution of income in each generation.[1]

Our analysis concentrates on the intragenerational distribution of average income across groups of individuals with a given, but unknown, ability. We demonstrate that income inequality may either increase or decrease with a better information system. More precisely, assuming constant relative risk aversion utility functions, we show that better information increases (decreases) income inequality if the relative measure of risk aversion is smaller (larger) than 1. Risk aversion plays such a crucial role because it affects the behavior of the optimal effort level: the effort level increases (decreases) as the agent receives a more favorable information signal if relative risk aversion is below 1 (above 1). Thus, better information may either increase or decrease the dispersion in the distribution of investments in education – a fact which critically contributes to our result about the consequences of better information for income inequality.

We also study the role of risk sharing arrangements for the link between information and the distribution of income. Assuming that the signals convey information related to an insurable part of the random ability (or, the rate of return to invest-

[1]Of course, better information has an impact on the accumulation of human capital as well and, hence, on economic growth. We have studied this aspect in a separate paper [see, Eckwert and Zilcha (2004)]. This paper uses a different information concept and it does not address the issue of income distribution studied here.

ment in human capital) we show that better information always enhances income inequality.

The paper is organized as follows. In section 2 we describe the OLG economy and define a concept of informativeness which is called 'reliability'. In section 3 we study the effect of more reliable signals on income inequality. Section 4 deals with the same issue in the presence of an insurance market. To facilitate reading all proofs are relegated to the Appendix.

2 The Model

Consider an overlapping generations economy with a single commodity which is traded each period $t = 0, 1, \cdots$. The commodity can either be consumed or used as an input (physical capital) in a production process (see, e.g., Azariadis and Drazen (1990)) . The generations reproduce identically over time. Each generation consists of a continuum of individuals $i \in [0, 1]$ who live for three periods. In their first period ('youth') agents obtain education while they are still supported by their parents. In their second period ('middle-age') they work and spend part of the labor income for consumption; and in their third period ('retirement') they consume their savings. We denote by $G_t, t = 0, 1, \cdots$ the generation of all agents born at date $t - 1$.

Each agent earns labor income which depends on his human capital. Human capital, h, of an agent in G_t is determined by his innate ability, $A \in \mathbb{R}_+$, the effort $e \in \mathbb{R}_+$ invested in education by this individual, and the 'environment' when education takes place represented by the average human capital of agents in the previous generation:

$$h = Ag(H_{t-1}, e) \tag{1}$$

H_{t-1} denotes the average human capital of G_{t-1} (to be defined below), and $g : \mathbb{R}_+^2 \to \mathbb{R}_+$ is a twice differentiable function which is strictly increasing in both arguments.

Agents differ only with respect to the abilities assigned to them by nature. Random individual ability $\tilde{A}$ realizes at the beginning of an agent's middle-age period according to the probability distribution ν. $\tilde{A}$ is a non-degenerate random variable. Hence, there is risk at the individual level. We assume that the individual risk about ability is identical across agents and that there is no risk in the aggregate, i.e., the ex post distribution of ability is exactly ν. This distribution ν is common knowledge. In view of a result by Feldman and Gilles (1985, p. 29, Proposition 2) such a probabilistic setting exists. In this setting the individual risks are identical but not independent (see also the discussion in Khan and Sun (1999)).

Before an agent chooses optimal effort in the youth period nature assigns to him a deterministic signal $y \in Y \subset \mathbb{R}$ which contains information about his unknown ability. The signals assigned to agents with ability A are distributed according to the density $f(\cdot|A)$. Invoking, once again, the result by Feldman and Gilles (1985, p. 29, Proposition 2) we may assume that $f(\cdot|A)$ is exactly the ex post distribution of signals across agents with ability A. The conditional densities $f(\cdot|A)$ are known to all agents. Thus, by construction, the distributions of signals and of abilities across agents are correlated. Since an agent in his youth period is ignorant about what ability nature has assigned to him, he forms expectations on the basis of the signal. Therefore, an agent's decision about how much effort, e, to invest in education, will be based on the conditional distribution of A given his signal y. Hence, as agents in the same generation differ only with respect to (unknown) ability, any two individuals who receive the same signal choose the same effort level.

The distribution of signals received by agents in the same generation has the density

$$\mu(y) = \int_{\mathbb{R}_+} f(y|A)\nu(A)\,\mathrm{d}A. \tag{2}$$

Denoting by $\nu_y(\cdot)$ the density of the conditional distribution of A given the signal y, average ability of all agents who have received the signal y is

$$\bar{A}(\nu_y) := \int_{\mathbb{R}_+} A\nu_y(A)\,\mathrm{d}A. \tag{3}$$

The agents are expected utility maximizers with von-Neumann Morgenstern lifetime utility function

$$U(e, c_1, c_2) = v(e) + u_1(c_1) + u_2(c_2). \tag{4}$$

Individuals derive negative utility from 'effort' while they are young and positive utility from consumption in the working period, c_1, and from consumption in the retirement period, c_2.

Assumption 1 *The utility functions v and u_j, $j = 1, 2$, have the following properties:*

(i) $v : \mathbb{R}_+ \to \mathbb{R}$ *is decreasing and strictly concave,*

(ii) $u_j : \mathbb{R}_+ \to \mathbb{R}$ *is increasing and strictly concave, $j = 1, 2$.*

In each period, competitive firms transform physical capital K and human capital H into a consumption/investment good. The transformation process can be described by an aggregate production function $F(K, H)$ which exhibits constant returns to scale. If individual i supplies l^i units of labor in his 'working period', his supply of human capital equals $l^i h^i$. We assume inelastic labor supply, i.e., that l^i is a constant and it is equal to 1 for all i.

Assumption 2 *$F(K, H)$ is concave, homogeneous of degree 1, and satisfies $F_K > 0$, $F_H > 0, F_{KK} < 0, F_{HH} < 0$.*

We also assume throughout this paper full international capital mobility, while human capital is assumed to be immobile. As a consequence, at each date t the interest rate $\bar{r}_t$ is exogenously given and marginal productivity of aggregate physical capital K_t is equal to $1 + \bar{r}_t$ (assuming full depreciation of capital in each period). Thus, given the aggregate stock of human capital at date t, H_t, the stock K_t must adjust such that

$$1 + \bar{r}_t = F_K(K_t, H_t) \qquad t = 1, 2, 3, \cdots \tag{5}$$

holds. Equation (5) and Assumption 2 imply that $\frac{K_t}{H_t}$ is determined by the international rate of interest $\bar{r}_t$. Hence the wage rate w_t (price of one unit of human capital), is given in equilibrium by the marginal product of aggregate human capital, is also determined once $\bar{r}_t$ is given. Thus we may write

$$w_t = F_L\Big(\frac{K_t}{H_t}, 1\Big) =: \zeta(\bar{r}_t) \qquad t = 1, 2, 3, \cdots . \tag{6}$$

Now let us consider the optimization problem of an agent in G_t, given $\bar{r}_t, w_t$, and H_{t-1}. At date $t - 1$, when 'young', this individual chooses the optimal level of effort employed in obtaining education. This decision is made after the individual has observed his signal y and, hence, is based on the conditional distribution $\nu_y(\cdot)$

of ability. The decision about saving, s, to be used for consumption when 'old' is taken in the second period, after the agent has learned his ability, A, and, hence, his human capital, h. Thus s will depend on h via the wage earnings, $w_t h$.

For given levels of h, w_t and $\bar{r}_t$, the optimal saving decision of the agent is determined by

$$\begin{aligned} \max_s \ & u_1(c_1) + u_2(c_2) \\ \text{s.t. } & c_1 = w_t h - s \\ & c_2 = (1+\bar{r}_t)s \end{aligned} \tag{7}$$

and satisfies the necessary and sufficient first order condition

$$-u_1'(w_t h - s) + (1+\bar{r}_t)u_2'((1+\bar{r}_t)s) = 0 \tag{8}$$

for all h. From equation (8) we find optimal saving as a function of each realized h, i.e., $s = s_t(h)$. The optimal level of effort invested in education, e, is determined by

$$\begin{aligned} \max_e \ & E[v(e) + u_1(c_1) + u_2(c_2)|y] \\ \text{s.t. } & c_1 = w_t h - s \\ & c_2 = (1+\bar{r}_t)s, \end{aligned} \tag{9}$$

where h is given by equation (1) and s satisfies equation (8). Note that from the perspective of the agent his (unknown) ability A is a random variable which is distributed according to $\nu_y(\cdot)$. Similarly, h is also perceived to be random as human capital depends on ability. Due to the Envelope theorem and the strict concavity of the utility functions, problem (9) has a unique solution determined by the first order condition

$$v'(e) + w_t g_2(H_{t-1}, e)E[Au_1'(w_t h - s)|y] = 0. \tag{10}$$

Since u_1' is a decreasing function we also conclude from (8) that $s_t(h)$ and $w_t h - s_t(h)$ are both increasing in h. This implies, in particular, that the LHS in (10) is strictly decreasing in e. Similarly, from equation (10) we obtain the optimal level of effort as a function of the conditional distribution ν_y, i.e., $e = e_t(\nu_y)$. Note that any two agents in generation t who receive the same individual signal will choose the same effort level.

Using (2) and (3) the aggregate stock of human capital at date t can be expressed as[2]

$$H_t = E_y[\bar{h}_t(\nu_y)] = \int_Y \bar{h}_t(\nu_y)\mu(y)\mathrm{d}y, \tag{11}$$

where

$$\bar{h}_t(\nu_y) := \bar{A}(\nu_y)g(H_{t-1}, e_t(\nu_y)) \tag{12}$$

is the average human capital of agents in G_t who have received the signal y.

Definition 1 *Given the international interest rates* $(\bar{r}_t)$ *and the initial stock of human capital* H_0*, a competitive equilibrium consists of a sequence* $\{(e^i, s^i)_{i\in G_t}\}_{t=1}^{\infty}$*, and a sequence of wages* $(w_t)_{t=1}^{\infty}$*, such that:*

(i) *At each date t, given* $\bar{r}_t$*,* H_{t-1}*, and* w_t*, the optimum for each* $i \in G_t$ *in problems (9) and (7) is given by* (e^i, s^i)*.*

(ii) *The aggregate stocks of human capital,* $H_t, t = 1, 2, \cdots$*, satisfy (11).*

(iii) *Wage rates* $w_t, t = 1, 2, \cdots$*, are determined by (6).*

2.1 Information Systems

When young, agents are ignorant about their abilities. Hence, a young individual perceives his own ability as being randomly distributed according to ν. Yet, he correctly understands that the distribution of signals and of abilities across the individuals in his generation are correlated. Therefore, he uses his signal, y, to update the perceived prior probability distribution, ν, of his ability. The updated (or, posterior) distribution has density

$$\nu_y(A) = f(y|A)\nu(A)/\mu(y). \tag{13}$$

[2] H_t can be rewritten as

$$H_t = \int_{\mathbb{R}_+}\left[\int_Y g(H_{t-1}, e_t(\nu_y))f(y|A)\,\mathrm{d}y\right]A\nu(A)\,\mathrm{d}A.$$

Therefore the aggregate human capital stock is deterministic.

In particular, ex ante the perceived posterior distributions of ability are the same for all agents who have received the same signal.

An information system, which will be represented by $f : Y \times \mathbb{R}_+ \to \mathbb{R}_+$ throughout the paper, specifies for each level of ability $A \in \mathbb{R}_+$ a conditional density function over the set of signals. The positive real number $f(y|A)$ is the conditional density of all agents with ability A to whom nature has assigned the signal y.

Assumption 3 *The densities $\{f(\cdot|A), A \in \mathbb{R}_+\}$ have the monotone likelihood ratio property (MLRP): given any prior distribution for A, $y' > y$ implies that the posterior distribution of A conditional on y' dominates the posterior distribution of A conditional on y in the sense of first-order stochastic dominance.*

MLRP implies that higher signal is 'good news' (see Milgrom (1981)). As a consequence, $\int_{\mathbb{R}_+} \varphi(A)\nu_{y'}(A)\mathrm{d}A \geq \int_{\mathbb{R}_+} \varphi(A)\nu_y(A)\mathrm{d}A$ holds for any strictly increasing function φ.

Blackwell (1953) proposed a criterion that compares different information systems by their informational contents. Suppose $\bar{f}$ and $\hat{f}$ are two information systems with associated density functions $\bar{\nu}_y$, $\hat{\nu}_y$, $\bar{\mu}$, $\hat{\mu}$. Blackwell defined the informativeness of an information system as follows:

Definition 2 *Let $\bar{f}$ and $\hat{f}$ be two information systems. $\bar{f}$ is said to be more informative in the Blackwell sense than $\hat{f}$(expressed by $\bar{f} \succ_{\text{inf}} \hat{f}$), if there exists an integrable function $\lambda : Y^2 \to \mathbb{R}_+$ such that*

$$\int_Y \lambda(y', y)\mathrm{d}y' = 1 \tag{14}$$

holds for all y, and

$$\hat{f}(y'|A) = \int_Y \bar{f}(y|A)\lambda(y', y)\mathrm{d}y \tag{15}$$

holds for all $A \in \mathbb{R}_+$.

According to this criterion, $\bar{f} \succ_{\text{inf}} \hat{f}$ holds if $\hat{f}$ can be obtained from $\bar{f}$ through a process of randomization, i.e., by adding some random noise.

Given an information system f and an ability level A, define

$$L_f^A(z) := \int_{\frac{f_A}{f}(y|A)\le z} f(y|A)\,\mathrm{d}y$$

where f_A denotes the partial derivative. $L_f^A(z)$ is called the *likelihood ratio distribution* of an agent with ability A under information system f.

Lemma 1 $\frac{f_A}{f}(y|A)$ *is monotone increasing in* y *and, hence, the likelihood ratio distribution function can be written as*

$$L_f^A(z) = F\left(\left(\frac{f_A}{f}\right)^{-1}(z)\Big|A\right), \tag{16}$$

where $F(y|A)$ *is the c.d.f. for the distribution of signals across agents with ability* A.

Kim (1995) has shown that the likelihood ratio distribution under $\bar{f}$, $L_{\bar{f}}^A(z)$, is a mean preserving spread of that under $\hat{f}$, $L_{\hat{f}}^A(z)$, if $\bar{f}$ is more informative (in the Blackwell sense) than $\hat{f}$:

Lemma 2 *Let* $\bar{f}$ *and* $\hat{f}$ *be two information systems such that* $\bar{f} \succ_{\text{inf}} \hat{f}$. *For any* $A \in \mathbb{R}_+$, $L_{\bar{f}}^A(z)$ *is a mean preserving spread (MPS) of* $L_{\hat{f}}^A(z)$. *That is, both distribution functions have the same mean and*

$$\int^{z'} \bar{F}\left(\left(\frac{\bar{f}_A}{\bar{f}}\right)^{-1}(z)\Big|A\right)\mathrm{d}z \ge \int^{z'} \hat{F}\left(\left(\frac{\hat{f}_A}{\hat{f}}\right)^{-1}(z)\Big|A\right)\mathrm{d}z \quad \forall z' \in \mathbb{R} \tag{17}$$

with the strict inequality holding for some z'.

Proof: see Kim (1995).

Inequality (17) can be transformed into an integral condition that will turn out to be a useful tool for the analysis in this paper.

Lemma 3 *Inequality (17) is satisfied for all* $z' \in \mathbb{R}$ *if and only if the following integral condition holds for all* $\vartheta \in [0,1]$*:*

$$S(\vartheta|A) := \int_0^{\vartheta}\left[\frac{\bar{f}_A}{\bar{f}}\left(\bar{F}^{-1}(s|A)\Big|A\right) - \frac{\hat{f}_A}{\hat{f}}\left(\hat{F}^{-1}(s|A)\Big|A\right)\right]\mathrm{d}s \le 0. \tag{18}$$

Proof: This lemma is a straightforward modification of Proposition 3 in Demougin and Fluet (2001). The proof is therefore omitted. □

Being ignorant about what level of ability nature has assigned to him, a young agent perceives his ability as random. For clarity, in the sequel we shall mark variables which are perceived to be random by a $\tilde{}$. Our analysis is based on a special concept of informativeness which we call 'reliability':

Definition 3 (reliability) *Let $\bar{f}$ and $\hat{f}$ be two information systems. $\bar{f}$ is more reliable than $\hat{f}$ (expressed by $\bar{f} \succ_{\text{rel}} \hat{f}$), if:*

(i) *[Adding Noise Reduces Reliability] For any $A \in \mathbb{R}_+$, the likelihood ratio distribution under $\bar{f}$, $L^A_{\bar{f}}(z)$, is a MPS of that under $\hat{f}$, $L^A_{\hat{f}}(z)$.*

(ii) *[Good News Become Better News Under More Reliability] For any $A \in \mathbb{R}_+$ and any increasing function $\pi : \mathbb{R}_+ \to \mathbb{R}_{++}$,*

$$\frac{E^{\bar{f}}[\pi(\tilde{A})|\bar{F}^{-1}(s''|A)]}{E^{\bar{f}}[\pi(\tilde{A})|\bar{F}^{-1}(s'|A)]} \geq \frac{E^{\hat{f}}[\pi(\tilde{A})|\hat{F}^{-1}(s''|A)]}{E^{\hat{f}}[\pi(\tilde{A})|\hat{F}^{-1}(s'|A)]} \tag{19}$$

holds for any two signals $s'' \geq s'$. If π is a decreasing function, then the inequality in (19) reverses.

According to Lemma 2, condition (i) is weaker than the Blackwell criterion, i.e., condition (i) is implied by $\bar{f} \succ_{\text{inf}} \hat{f}$. Condition (ii) postulates that under a more reliable information system the conditional expectation of any increasing transformation of the state variable (ability) reacts more sensitively to changes in the (transformed) signal s. In this sense an increase in the transformed signal s constitutes better news if the signal is more reliable. Note that

$$\frac{E^{\bar{f}}[\pi(\tilde{A})|\bar{F}^{-1}(s|A)]}{E^{\hat{f}}[\pi(\tilde{A})|\hat{F}^{-1}(s|A)]} \tag{20}$$

is monotone increasing in the signal s, if and only if condition (ii) holds.

Our definition of *reliability* contains two conditions: condition (i), regarding the likelihood ratio distribution, has been used by Kim (1995), who demonstrated that it is implied by the Blackwell sufficiency condition. Our condition (ii) bears similarity with Lehmann's (1988) concept of *accuracy* (discussed and characterized by Persico (2000)). Intuitively speaking, a signal is more accurate, if it is more closely related (in a statistical sense) to the state variable. Our reliability concept admits a similar interpretation. Ericson (1969) derived a sufficient condition which guarantees that property (ii), or (19), holds. He also showed that this condition is satisfied for many classes of prior/posterior pairs. As a result, restricting our distributions to such a class renders our reliability concept to be weaker than Blackwell's sufficiency condition (using Kim's (1995) result). However, in general, there are few open questions here: Under what conditions does Blackwell's sufficiency condition imply reliability? Under what conditions does reliability imply accuracy, and vice versa?

Observe that any two information systems which can be ordered according to Blackwell's criterion, $\bar{f} \succ_{\text{inf}} \hat{f}$, satisfy the inequality in (19) for $s' = 0$ and $s'' = 1$. In particular, for increasing π, the term in (20) is less than 1 for $s = 0$ and larger than 1 for $s = 1$. This observation follows from the following assessment (for arbitrary $\hat{A} \in \mathbb{R}_+$):[3]

$$\begin{aligned} E^{\bar{f}}[\pi(\tilde{A})|\bar{F}^{-1}(0|\hat{A})] &= E^{\bar{f}}[\pi(\tilde{A})|\underline{y}] = \int_{\mathcal{A}} \pi(A)\bar{\nu}_{\underline{y}}(A)\,\mathrm{d}A \\ &\leq \frac{1}{\hat{\mu}(\underline{y})} \int_Y \bar{\mu}(y')\lambda(\underline{y}, y') \int_{\mathcal{A}} \pi(A)\bar{\nu}_{y'}(A)\,\mathrm{d}A\mathrm{d}y' \\ &= \int_{\mathcal{A}} \pi(A)\hat{\nu}_{\underline{y}}(A)\,\mathrm{d}A = E^{\hat{f}}[\pi(\tilde{A})|\hat{F}^{-1}(0|\hat{A})]. \end{aligned} \tag{21}$$

By a similar argument we get

$$E^{\bar{f}}[\pi(\tilde{A})|\bar{F}^{-1}(1|\hat{A})] \geq E^{\hat{f}}[\pi(\tilde{A})|\hat{F}^{-1}(1|\hat{A})], \quad \forall \hat{A} \in \mathcal{A}. \tag{22}$$

In view of (21) and (22), the conditional expectations $E^{\bar{f}}[\pi(\tilde{A})|\bar{F}^{-1}(s|\hat{A})]$ and $E^{\hat{f}}[\pi(\tilde{A})|\hat{F}^{-1}(s|\hat{A})]$ cross at least once.

Thus the Blackwell criterion has important implications for the curvatures of con-

[3] The third equality makes use of Bayes' rule which implies

$$\hat{\nu}_y(A) = \frac{1}{\hat{\mu}(y)} \int_Y \bar{\mu}(y')\bar{\nu}_{y'}(A)\lambda(y, y')\,\mathrm{d}y'.$$

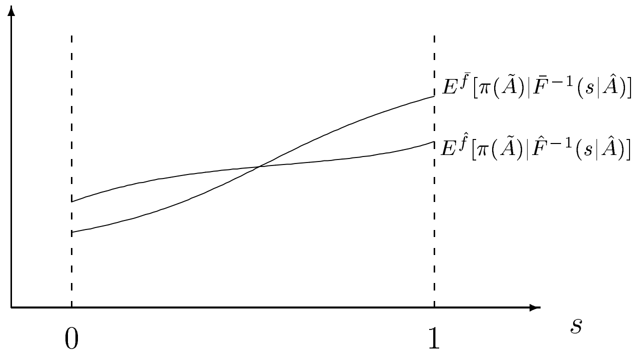

Figure 1: Conditional expectations under information systems f and $\hat{f}$.

ditional expectations under the information systems $\bar{f}$ and $\hat{f}$. These implications are strengthened by the condition in (19), which implies that the conditional expectations in Figure 1 have the single crossing property.[4]

3 Information and Inequality Without Risk Sharing

To facilitate the comparison of income distributions under different information systems we restrict the utility functions $u_1(\cdot)$, $u_2(\cdot)$, and $v(\cdot)$ to be in the family of CRRA:

$$u_1(c_1) = \frac{c_1^{1-\gamma_u}}{1-\gamma_u}; \quad u_2(c_2) = \beta\frac{c_2^{1-\gamma_u}}{1-\gamma_u}; \quad v(e) = -\frac{e^{\gamma_v+1}}{\gamma_v+1}. \tag{23}$$

[4]In many economic applications signals are generated according to $\tilde{s} = A + \delta\tilde{\varepsilon}$. For a variety of distributions, which include the cases where $\tilde{A}$ and $\tilde{s}$ are jointly normal or where the log of their densities is quadratic (see Ericson (1969)), this specification implies that conditional expectations can be represented as $E[\tilde{A}|s] = E\tilde{A} + \vartheta(s - E\tilde{s})$ with $\vartheta = \sigma_A^2/(\sigma_A^2 + \delta^2\sigma_\varepsilon^2)$. Typically, in these applications ϑ is used as a measure of reliability. In such a framework condition (19) is always satisfied: Let $\bar{f}$ and $\hat{f}$ be information systems corresponding to $\bar{\vartheta}$ and $\hat{\vartheta}$, $\bar{\vartheta} > \hat{\vartheta}$. We conclude

$$\frac{E^{\bar{f}}[\pi(\tilde{A})|s]}{E^{\hat{f}}[\pi(\tilde{A})|s]} = \frac{E[\pi(\tilde{A})] + \bar{\vartheta}(s - E\tilde{s})}{E[\pi(\tilde{A})] + \hat{\vartheta}(s - E\tilde{s})},$$

which is increasing in s.

γ_u and γ_v are strictly positive constants. γ_v parametrizes the curvature of the utility function in the youth period, v; and γ_u parametrizes the curvature of the utility functions in the middle age period and retirement period, u_i, $i = 1, 2$.

We also assume that the function g in (1) has the form

$$g(H, e) = \hat{g}(H)e^{\alpha}, \tag{24}$$

where $\hat{g}$ is strictly increasing in H, and $\alpha \in (0, 1)$.

Using the functional forms of u_j, $j = 1, 2$, in (23), it follows from equation (8) that, given $\bar{r}_t$ and w_t, the saving s is proportional to the human capital level h. In other words, for each t and for any agent in G_t we have:

$$s = m_t h, \quad 0 < m_t < w_t, \quad t = 1, 2, \cdots \tag{25}$$

The specifications in (23), (24) and (25) allow us to solve equation (10) for the optimal effort level as a function of the conditional distribution ν_y:

$$e_t(\nu_y) = \delta_t \Big(E[\tilde{A}^{1-\gamma_u} | y] \Big)^{\rho/\alpha} \tag{26}$$

where

$$\delta_t := \left[\frac{\alpha w_t (\hat{g}(H_{t-1}))^{1-\gamma_u}}{(w_t - m_t)^{\gamma_u}} \right]^{\rho/\alpha} ; \quad \rho = \frac{\alpha}{\gamma_v + \alpha(\gamma_u - 1) + 1}.$$

We will discuss the role of information for the distribution of expected individual incomes across agents of different types *before* ability and signals are revealed. For that purpose we focus on the average income of all agents with given ability A. The distribution of average income within this class of agents across different ability levels will serve as a measure of inequality.

Let

$$\begin{aligned} I_t^f(A) &:= \int_Y w_t h f(y|A)\, \mathrm{d}y \\ &= w_t \hat{g}(H_{t-1}) \delta_t^{\alpha} A \int_Y \Big(E^f[\tilde{A}^{1-\gamma_u} | y] \Big)^{\rho} f(y|A)\, \mathrm{d}y \end{aligned} \tag{27}$$

be the expected income, as of date $t-1$ of an agent in generation t with ability A. We measure ex ante income inequality by the elasticity of expected income with respect to ability.

Definition 4 (income inequality) *Income distribution under information system $\bar{f}$ is said to be more unequal than income distribution under information system $\hat{f}$, if*

$$\varepsilon[I_t^{\bar{f}}, A] := \frac{\partial I_t^{\bar{f}}(A)}{\partial A} \frac{A}{I_t^{\bar{f}}(A)} \geq \frac{\partial I_t^{\hat{f}}(A)}{\partial A} \frac{A}{I_t^{\hat{f}}(A)} =: \varepsilon[I_t^{\hat{f}}, A] \tag{28}$$

holds for all $t \geq 0$ and all A.

We will show below (cf. part (i) in Theorem 1) that for $\gamma_u \leq 1$ and given A the elasticity $\varepsilon[I_t^f, A]$ is bounded from below by $\varepsilon[I_t^{f^0}, A]$, where f^0 denotes the uninformative system. From (27) it is immediate that $\varepsilon[I_t^{f^0}, A] = 1$ for all A. Similarly, for $\gamma_u \geq 1$ the elasticity $\varepsilon[I_t^f, A]$ is bounded from below by $\varepsilon[I_t^{f^1}, A]$ where f^1 is the fully reliable system. Since under f^1 the signal reveals an agent's talent, (27) reduces to

$$[w_t \hat{g}(H_{t-1})\delta_t^\alpha]^{-1} I_t^{f^1}(A) = A^{1+(1-\gamma_u)\rho}$$

and, hence, $\varepsilon[I_t^{f^1}, A] = 1 + (1-\gamma_u)\rho = (1+\gamma_v)/[1+\gamma_v+\alpha(\gamma_u - 1)] > 0$. Thus the elasticities in (28) are strictly positive.

Remark: The above definition of income inequality implies the definition given by Atkinson (1970) in the sense that the normalized income under $\hat{f}$, $||I_t^{\hat{f}}(A)|| := I_t^{\hat{f}}(A)/E[I_t^{\hat{f}}(A)]$, dominates in second degree stochastic dominance the normalized income under $\bar{f}$, $||I_t^{\bar{f}}(A)|| := I_t^{\bar{f}}(A)/E[I_t^{\bar{f}}(A)]$. In other words, $||I_t^{\bar{f}}(A)||$ and $||I_t^{\hat{f}}(A)||$ differ by a MPS and, hence, the Lorenz curve for $||I_t^{\hat{f}}(A)||$ lies strictly above that for $||I_t^{\bar{f}}(A)||$.

Observe that

$$\operatorname{sign}\big(\varepsilon[I_t^{\bar{f}}, A] - \varepsilon[I_t^{\hat{f}}, A]\big) = \operatorname{sign}\big(\varepsilon[\bar{I}^{\bar{f}}, A] - \varepsilon[\bar{I}^{\hat{f}}, A]\big), \tag{29}$$

where

$$\bar{I}^f(A) := I_t^f(A)/w_t \hat{g}(H_{t-1})\delta_t^\alpha A, \quad f = \bar{f}, \hat{f}.$$

$\bar{I}^f(A)$ can be rewritten as

$$\begin{aligned} \bar{I}^f(A) &= \int_0^1 \Big(E^f\big[\tilde{A}^{1-\gamma_u} \big| F^{-1}(s|A)\big]\Big)^\rho f\big(F^{-1}(s|A)\big|A\big)\big(F^{-1}\big)'(s)\,\mathrm{d}s \\ &= \int_0^1 \Big(E^f\big[\tilde{A}^{1-\gamma_u} \big| F^{-1}(s|A)\big]\Big)^\rho \mathrm{d}s, \quad (f,F) = (\bar{f},\bar{F}), (\hat{f},\hat{F}) \end{aligned} \tag{30}$$

Differentiating (30) we obtain

$$\frac{\partial \bar{I}^f(A)}{\partial A}\frac{1}{\bar{I}^f(A)} = \frac{1}{\bar{I}^f(A)}\int_Y \left(E^f[\tilde{A}^{1-\gamma_u}|y]\right)^\rho f_A(y|A)\,\mathrm{d}y$$

$$= \int_0^1 \frac{\left(E^f\left[\tilde{A}^{1-\gamma_u}\middle|F^{-1}(s|A)\right]\right)^\rho}{\int_0^1 \left(E^f\left[\tilde{A}^{1-\gamma_u}\middle|F^{-1}(\hat{s}|A)\right]\right)^\rho \mathrm{d}\hat{s}}\frac{f_A}{f}\left(F^{-1}(s|A)\middle|A\right)\mathrm{d}s. \quad (31)$$

Combining (29) and (31) yields a useful characterization of ex ante inequality. Define

$$\Gamma^A(f|s) := \frac{\left(E^f\left[\tilde{A}^{1-\gamma_u}\middle|F^{-1}(s|A)\right]\right)^\rho}{\int_0^1 \left(E^f\left[\tilde{A}^{1-\gamma_u}\middle|F^{-1}(\hat{s}|A)\right]\right)^\rho \mathrm{d}\hat{s}}; \qquad (f,F) = (\bar{f},\bar{F}),(\hat{f},\hat{F}) \quad (32)$$

Proposition 1 *Income inequality in the sense of Definition 4 is higher under information system $\bar{f}$ than under information system $\hat{f}$, if and only if*

$$\int_0^1 \left[\Gamma^A(\bar{f}|s)\frac{\bar{f}_A}{\bar{f}}\left(\bar{F}^{-1}(s|A)\middle|A\right) - \Gamma^A(\hat{f}|s)\frac{\hat{f}_A}{\hat{f}}\left(\hat{F}^{-1}(s|A)\middle|A\right)\right]\mathrm{d}s \geq 0, \quad (33)$$

holds for all A.

If relative risk aversion, γ_u, is equal to 1, expected individual income is linear in ability and, hence, income inequality is not affected by more reliable information. To verify this claim we first observe that

$$\int_0^1 \frac{f_A}{f}\left(F^{-1}(s|A)\middle|A\right)\mathrm{d}s = \int_Y f_A(y|A)\,\mathrm{d}y = 0. \quad (34)$$

Proposition 2 *If $\gamma_u = 1$, income inequality does not depend on the chosen information system.*

Proof: $\gamma_u = 1$ and equation (34) together imply that (33) is satisfied with equality for all A. Thus inequality of the income distribution is the same for any two information systems $\bar{f}$ and $\hat{f}$. □

For $\gamma_u = 1$ the individual effort level is not responsive to the information revealed by a signal (cf. (26)). As a consequence, the distribution of expected incomes across agents does not depend on the information system either. Yet, when γ_u differs from 1 a more reliable information system may give rise to more or less income inequality.

Theorem 1 *Let $\bar{f}$ and $\hat{f}$ be two information systems with $\bar{f} \succ_{\text{rel}} \hat{f}$.*

(i) *If $\gamma_u \leq 1$, the income distribution under $\bar{f}$ is more unequal than the income distribution under $\hat{f}$.*

(ii) *If $\gamma_u \geq 1$, the income distribution under $\hat{f}$ is more unequal than the income distribution under $\bar{f}$.*

Information affects the distribution of efforts at each date. From (26) we know that effort is increasing in the signal if $\gamma_u \leq 1$, and decreasing if $\gamma_u \geq 1$. Consider the case $\gamma_u \leq 1$. Under a more reliable information system signals more accurately reflect the agents' talents. As a consequence, due to more favorable signals, agents with high talents tend to choose higher effort levels; and agents with low talents tend to choose lower effort levels. This mechanism produces more income inequality. Similarly, if $\gamma_u \geq 1$ the mechanism works in reverse and reduces income inequality.

Remark: More reliable information leads to higher income inequality whenever effort is increasing in the signal. If the utility functions u_1 and u_2 satisfy the restriction in (23), then assuming $\gamma_u \leq 1$ generates this monotonicity. Otherwise, additional assumptions are needed. If, e.g., u_1 and u_2 are restricted to be in the family of CRRA, then effort is increasing in the signal if $\gamma_{u_2} \leq \gamma_{u_1} \leq 1$ holds; and effort is decreasing in the signal if $\gamma_{u_2} \geq \gamma_{u_1} \geq 1$ is satisfied. This claim can be verified through inspection of equations (8) and (10).

4 Information and Inequality with Risk Sharing

This section proceeds on the assumption that part of the perceived uncertainty of an agent's ability is insurable. Let $A = A_1 \cdot A_2$ with $(A_1, A_2) \in \mathcal{A} := \mathbb{R}_+^2$. We assume that the distributions of A_1 and of A_2 across agents in H_t are stochastically independent. Before agents make decisions about effort they can insure the perceived risk which is associated with the A_1- component of their (unknown) ability. Since there is no aggregate risk in the economy the insurance market for the A_1-component of ability will be unbiased, i.e., the agents can share part of the perceived risk on fair terms. In Section 3.1 the signals affected only uninsurable risks. In this section we assume that the signals contain only information about the insurable risk factor A_1.[5]

In order to introduce the risk sharing market we need to assume that the A_1-component of individual ability is verifiable by the insurers. The future income of each individual, perceived as random at young age, will then have an insurable component as well as an uninsurable component.[6] Denote by $\bar{A}_1(\nu_y)$ the expected value of $\tilde{A}_1$ if the signal y has been observed,

$$\bar{A}_1(\nu_y) := \int_{\mathcal{A}} A_1 \nu_y(A)\,\mathrm{d}A. \tag{35}$$

Since the insurance market is unbiased, all agents find it optimal to completely eliminate the (perceived) A_1- risk from income in their second period of life. Thus the optimal saving and effort decisions of an agent in G_t satisfy the following first order conditions

$$(1+\bar{r}_t)u_2'\big((1+\bar{r}_t)s\big) - u_1'\big(w_t\bar{A}_1(\nu_y)A_2 g(H_{t-1},e) - s\big) = 0 \quad (A_2 \in \mathbb{R}_+) \tag{36}$$

$$v'(e) + w_t g_2(H_{t-1},e)E\big[\tilde{\bar{A}}(\nu_y)u_1'\big(w_t\tilde{\bar{A}}(\nu_y)g(H_{t-1},e) - s\big)|y\big] = 0 \quad (y \in Y), \tag{37}$$

where

$$\tilde{\bar{A}}(\nu_y) := \bar{A}_1(\nu_y)\cdot\tilde{A}_2. \tag{38}$$

It is our aim to analyze the impact of information on income inequality if agents are able to share part of the uncertainty about their random ability. Our next theorem deals with this issue. It investigates whether more reliable information about the insurable risk increases or decreases income inequality.

Using the functional specifications (23)-(26) the average income of an agent with ability $A = A_1 \cdot A_2$ is

$$\begin{aligned} I_t^f(A_1,A_2) &= w_t\delta_t^\alpha \hat{g}(H_{t-1})\Big(E\big[\tilde{A}_2^{1-\gamma_u}\big]\Big)^\rho A_2 \int_Y \Big(E^f\big[\tilde{A}_1|y\big]\Big)^\tau f(y|A_1)\,\mathrm{d}y \\ &= w_t\delta_t^\alpha \hat{g}(H_{t-1})\Big(E\big[\tilde{A}_2^{1-\gamma_u}\big]\Big)^\rho A_2 \int_0^1 \Big(E^f\big[\tilde{A}_1|F^{-1}(s|A_1)\big]\Big)^\tau \mathrm{d}s, \end{aligned} \tag{39}$$

[5]In fact, the analysis in Section 3 can be understood as being conducted in the same framework as here with the signals containing only information about the uninsurable (perceived) risk factor A_2.

[6]Again, we shall mark with a ˜ those variables which are perceived as random by the agent.

where

$$\tau := 1 + \rho(1-\gamma_u) = \frac{1+\gamma_v}{\gamma_v + \alpha\gamma_u + (1-\alpha)} > 0. \tag{40}$$

Since $I_t^f(\cdot)$ is linear in A_2 regardless of the information system, the elasticity of expected income with respect to A_1 is now the relevant measure of inequality: the income distribution is more unequal under $\bar{f}$ than under $\hat{f}$ if $\varepsilon[I_t^{\bar{f}}, A_1] \geq \varepsilon[I_t^{\hat{f}}, A_1]$ holds for all $t \geq 0$ and $A_1 \in \mathbb{R}_+$.

Define

$$\Delta^{A_1}(f|s) := \frac{\left(E^f\left[\tilde{A}_1 \middle| F^{-1}(s|A_1)\right]\right)^\tau}{\int_0^1 \left(E^f\left[\tilde{A}_1 \middle| F^{-1}(\hat{s}|A_1)\right]\right)^\tau \mathrm{d}\hat{s}}. \tag{41}$$

The same procedure as in Section 3 yields

Proposition 3 *Income inequality is higher under information system $\bar{f}$ than under information system $\hat{f}$, if and only if*

$$\int_0^1 \left[\Delta^{A_1}(\bar{f}|s)\frac{\bar{f}_{A_1}}{\bar{f}}\left(\bar{F}^{-1}(s|A_1)\middle|A_1\right) - \Delta^{A_1}(\hat{f}|s)\frac{\hat{f}_{A_1}}{\hat{f}}\left(\hat{F}^{-1}(s|A_1)\middle|A_1\right)\right] \mathrm{d}s \geq 0, \tag{42}$$

holds for all A_1.

In the absence of risk sharing opportunities the impact of more reliable information on income inequality was critically dependent on the risk aversion parameter γ_u. By contrast, if those risks which are affected by the signals can be insured on fair terms more reliable information will always increase income inequality:

Theorem 2 *Assume that an unbiased insurance market for the A_1-risk is available. Let $\bar{f}$ and $\hat{f}$ be two information systems with $\bar{f} \succ_{\text{rel}} \hat{f}$. The income distribution under $\bar{f}$ is more unequal than the income distribution under $\hat{f}$.*

If the signals convey information about *insurable* risks more efficient screening (i.e., a more reliable information system) results in more inequality regardless of the measure of relative risk aversion. In Section 3 we found that the effect of screening information about *non-insurable* risks on income inequality depends on the agents' attitudes towards risk. Why is risk sharing so critical for the link between income inequality and information?

In the absence of risk sharing arrangements agents use their effort decisions as a hedging instrument against talent risks: strongly risk-averse agents choose higher effort in response to a less favorable signal. This mechanism produces less income inequality under a more reliable system. By contrast, moderately risk-averse agents choose higher effort when they receive a more favorable signal, which results in more income inequality.

When an unbiased risk sharing market exists, the implications of screening information on income inequality are different. The opportunity to obtain insurance against talent risks on fair terms obviously eliminates any incentives to use effort as a hedging device. However, under unbiased risk sharing arrangements, agents with favorable signals (i.e., high expected ability) obtain insurance on better terms than agents with unfavorable signals (i.e., low expected ability). This mechanism produces more income inequality under a more reliable information system, regardless of the agents' attitudes towards risk.

Unlike income inequality, economic growth is more intricately related to screening information. Even when risk sharing is possible, attitudes towards risk critically affect the link between growth and information. Eckwert and Zilcha (2004) found that in a framework similar to the one in this paper more reliable information enhances growth when γ_u is less than 1; and more reliable information reduces growth when γ_u exceeds 1. Thus, in economies with low risk aversion ($\gamma_u < 1$) information induced growth always comes at the cost of higher income inequality. For strongly risk-averse economies ($\gamma_u > 1$), by contrast, our analysis suggests an inverse linkage between growth and inequality.

6 Concluding Remarks

Education systems in all modern societies are based on screening mechanisms which generate information about the (unobservable) talent of young individuals. This paper is an attempt to analyze the implications of such mechanisms for the distribution of intragenerational income. A screening mechanism is modeled as an information system that allows the interpretation (in a Bayesian way) of signals which are specific to the agents.

Our analysis has shown that the implementation of a more efficient screening mechanism for individual talent will have effects on the inequality of the income distribution. These effects critically depend on the market structure of the economy. If the screening mechanism is applied to uninsurable risks, the income distribution will become more (less) unequal with more efficient screening if the economy is moderately (strongly) risk averse. Yet, if the screening mechanism applies to insurable risks, better screening always produces more income inequality.

In our framework individuals make their decision about investment in education after they observe a signal correlated to their ability, but prior to knowing the rate of return on this investment, i.e., before the realization of their 'type'. As a result 'income' can be defined at three points of time: (i) ex-ante income, i.e., expected income for each given level of ability A, (ii) expected income given the signal each individual observes, and (iii) ex-post income, i.e., the realized income when state of nature is known. We have considered here the ex-ante income distribution, but we are aware of the fact that our results may significantly change if we choose a different notion of income. Each type of inequality has a different conceptual meaning and it is not the place here to compare them due to the deep normative aspects it entails. We intend, however, to examine the issues discussed in this paper for the other notions of income as well.

To illustrate the significance of the underlying income concept for the properties of the equilibrium income distribution let us consider the subset D of all individuals with the same ability A. Under the ex ante specification (i) all individuals in this subset have the same expected income. Now consider the distribution of income after each person has observed an individual signal (specification (ii)). Obviously, agents in the subset D will almost surely have different incomes unless the information system is either fully informative (where the signal reveals the state) or uninformative (where the signals convey no information). A similar argument applies to the income specification (iii). Clearly we cannot expect that the results of our paper generalize to income concepts which correspond to the specifications (ii) or (iii).

Appendix

Let us prove some preliminary results before we proceed with the proofs of the theorems.

Proof of Lemma 1: Define $h(y|A,\hat{A}) := f(y|A)/f(y|\hat{A})$. Obviously, $h_y(y|A,\hat{A}) = 0$, $\forall y$, if $A = \hat{A}$. Also, by MLRP, $h_y(y|A,\hat{A}) > 0$, $\forall y$, if $A > \hat{A}$. From this observation we conclude

$$0 \leq \frac{\partial}{\partial A}\Big(h_y(y|A,\hat{A})\Big)\Big|_{A=\hat{A}} = \frac{\partial}{\partial y}\left(\frac{\partial h(y|A,\hat{A})}{\partial A}\Big|_{A=\hat{A}}\right) = \frac{\partial}{\partial y}\left(\frac{f_A(y|A)}{f(y|A)}\right),$$

which proves the claim. □

Lemma 4 *Let $\theta : X := [\underline{x}, \bar{x}] \to \mathbb{R}$ be an integrable function satisfying*

$$\int_{\underline{x}}^{\hat{x}} \theta(x)\,\mathrm{d}x \leq 0; \qquad \int_X \theta(x)\,\mathrm{d}x = 0$$

for all $\hat{x} \in X$. Then

$$\int_X h(x)\theta(x)\,\mathrm{d}x \overset{(\leq)}{\geq} 0$$

holds for any differentiable monotone increasing (decreasing) function $h : X \to \mathbb{R}$.

Proof: Integration by parts gives

$$\begin{aligned} \int_X h(x)\theta(x)\,\mathrm{d}x &= h(x)\int_{\underline{x}}^{x} \theta(\hat{x})\,\mathrm{d}\hat{x}\Big|_{\underline{x}}^{\bar{x}} - \int_X h'(x)\int_{\underline{x}}^{x} \theta(\hat{x})\,\mathrm{d}\hat{x}\,\mathrm{d}x \\ &= -\int_X h'(x)\int_{\underline{x}}^{x} \theta(\hat{x})\,\mathrm{d}\hat{x}\,\mathrm{d}x. \end{aligned}$$

The last term is non-negative if $h(x)$ is increasing; and it is non-positive if $h(x)$ is decreasing. □

Lemma 5 *Let $\bar{f}$ and $\hat{f}$ be two information systems with $\bar{f} \succ_{\text{rel}} \hat{f}$.*

(i)

$$\frac{\Gamma^A(\bar{f}|s'')}{\Gamma^A(\bar{f}|s')} \geq \frac{\Gamma^A(\hat{f}|s'')}{\Gamma^A(\hat{f}|s')}, \qquad s'' \geq s',\ A \in \mathbb{R}_+ \tag{43}$$

holds for $\gamma_u \leq 1$,

(ii)

$$\frac{\Gamma^A(\bar{f}|s'')}{\Gamma^A(\bar{f}|s')} \leq \frac{\Gamma^A(\hat{f}|s'')}{\Gamma^A(\hat{f}|s')}, \qquad s'' \geq s',\ A \in \mathbb{R}_+ \tag{44}$$

holds for $\gamma_u \geq 1$.

Proof: The claims in (i) and (ii) follow from part (ii) in Definition 3, if we set $\pi(\tilde{A}) = \tilde{A}^{1-\gamma_u}$.

□

Corollary 1 *Let* $\bar{f}$ *and* $\hat{f}$ *be two information systems with* $\bar{f} \succ_{\text{rel}} \hat{f}$. *Given any* $A \in \mathbb{R}_+$ *and any increasing function* $V : [0,1] \to \mathbb{R}$,

$$\int_0^1 V(s)\Gamma^A(\bar{f}|s)\,\mathrm{d}s \overset{(\leq)}{\geq} \int_0^1 V(s)\Gamma^A(\hat{f}|s)\,\mathrm{d}s$$

holds for $\gamma_u \overset{(\geq)}{\leq} 1$.

Proof: By Lemma 5, the densities $\Gamma^A(f,\cdot)$, $f = \bar{f}, \hat{f}$, satisfy a variant of the MLRP in the sense of (43) for $\gamma_u \leq 1$ and (44) for $\gamma_u \geq 1$. Therefore, the same reasoning as in the proof of Proposition 1 in Milgrom (1981) yields the result claimed in the corollary.

□

Proof of Theorem 1: We apply the characterization in Proposition 1.

(i) According to Lemma 1, $\frac{f_A}{f}\left(F^{-1}(s|A)\middle|A\right)$ is monotone increasing in s for $(f, F) = (\bar{f}, \bar{F}), (\hat{f}, \hat{F})$. The validity of the condition in (33) is then immediate from the following assessment:

$$\begin{aligned}\int_0^1 \frac{\bar{f}_A}{\bar{f}}\left(\bar{F}^{-1}(s|A)\middle|A\right)\Gamma^A(\bar{f}|s)\,\mathrm{d}s \;&\geq\; \int_0^1 \frac{\bar{f}_A}{\bar{f}}\left(\bar{F}^{-1}(s|A)\middle|A\right)\Gamma^A(\hat{f}|s)\,\mathrm{d}s\\ &\geq\; \int_0^1 \frac{\hat{f}_A}{\hat{f}}\left(\hat{F}^{-1}(s|A)\middle|A\right)\Gamma^A(\hat{f}|s)\,\mathrm{d}s\end{aligned}$$

In the first of the above inequalities we have used Corollary 1. The second inequality makes use of (34), Lemma 4, and the integral condition (18).

(ii) If $\gamma_u \geq 1$, the last two inequalities are reversed. Proposition 1 then implies the claim.

□

Lemma 6 *Let $\bar{f}$ and $\hat{f}$ be two information systems with $\bar{f} \succ_{\mathrm{rel}} \hat{f}$.*

$$\frac{\Delta^{A_1}(\bar{f}|s'')}{\Delta^{A_1}(\bar{f}|s')} \geq \frac{\Delta^{A_1}(\hat{f}|s'')}{\Delta^{A_1}(\hat{f}|s')} \tag{45}$$

holds for all A_1, $s'' \geq s'$.

Proof: Since the information systems $\bar{f}$ and $\hat{f}$ satisfy MLRP, with respect to the state variable A_1, (45) follows from part (ii) in Definition 3. □

Corollary 2 *Let $\bar{f}$ and $\hat{f}$ be two information systems with $\bar{f} \succ_{\mathrm{rel}} \hat{f}$. Given any increasing function $V : [0,1] \to \mathbb{R}$,*

$$\int_0^1 V(s)\Delta^{A_1}(\bar{f}|s)\,\mathrm{d}s \geq \int_0^1 V(s)\Delta^{A_1}(\hat{f}|s)\,\mathrm{d}s$$

holds for all A_1.

Proof: By Lemma 6, the densities $\Delta^{A_1}(f,\cdot)$, $f = \bar{f}, \hat{f}$, satisfy a variant of the MLRP in the sense of (45). The result follows by the same reasoning as in the proof of Proposition 1 in Milgrom (1981). □

Proof of Theorem 2: We apply the characterization in Proposition 3. Using Lemma 6 and Corollary 2, the proof proceeds as in part (i) of Theorem 1. □

References

1. Aghion, P., 2002, *Schumpeterian Growth Theory and the Dynamics of Income Inequality*, Econometrica 70(3), 855-882.

2. Azariadis, C. and Drazen, A., 1990, *Threshold Externalities in Economic Development*, Quarterly Journal of Economics 105, 501-526.

3. Becker, G., 1964, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, Columbia University Press, New York.

4. Benabou, R., 1996, *Equity and Efficiency in Human Capital Investment: The Local Connection*, Review of Economic Studies 63, 237-264.

5. Blackwell, D., 1951, *Comparison of Experiments*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics, 93-102.

6. Blackwell, D., 1953, *Equivalent Comparison of Experiments*, Annals of Mathematical Statistics 24, 265-272.

7. Diamond, P., 1965, *National Debt in a Neoclassical Growth Model*, American Economic Review 55, 1126-1150.

8. Demougin, D. and Fluet, C., 2001, *Ranking of Information Systems in Agency Models: An Integral Condition*, Economic Theory 17, 489-496.

9. Eckwert, B. and Zilcha, I., 2001, *The Value of Information in Production Economies*, Journal of Economic Theory 100, 172-186.

10. Eckwert, B. and Zilcha, I., 2004, *Economic Implications of Better Information in a Dynamic Framework*, Economic Theory 24, 561-581.

11. Ericson, W. A., 1969, *A Note on the Posterior Mean of a Population Mean*, Journal of the Royal Statistical Society. Series B, 31(2), 332-334.

12. Feldman, M. and Gilles, C., 1985, *An Expository Note on Individual Risk without Aggregate Uncertainty*, Journal of Economic Theory 35, 26-32.

13. Galor, O. and Tsiddon, D., 1997, *The Distribution of Human Capital and Economic Growth*, Journal of Economic Growth 2, 93-124.

14. Galor, O. and Zeira, J., 1993, *Income Distribution and Macroeconomics*, Review of Economic Studies 60, 35-52.

15. Glomm, G. and Ravikumar, B., 1992, *Private Investment in Human Capital: Endogenous Growth and Income Inequality*, Journal of Political Economy 100, 818-834.

16. Green, J., 1981, *The Value of Information with Sequential Futures Markets*, Econometrica 49, 335-358.

17. Hassler, J. and Mora, V. R., 2000, *Intelligence, Social Mobility, and Growth*, American Economic Review 90(4), 888-908.

18. Hirshleifer, J., 1971, *The Private and Social Value of Information and the Reward to Incentive Activity*, American Economic Review 61, 561-574.

19. Hirshleifer, J., 1975, *Speculation and Equilibrium: Information, Risk and Markets*, Quarterly Journal of Economics 89, 519-542.

20. Juhn, C., Murphy, K. M. and Brooks, P., 1993, *Wage Inequality and the Rise in Return to Skill*, Journal of Political Economy 101(3), 410-442.

21. Khan, M. A. and Sun, Y., 1999, *Weak Measurability and Characterizations of Risk*, Economic Theory 13, 541-560.

22. Kim, S. K., 1995 *Efficiency of an Information System in an Agency Model*, Econometrica 63(1), 89-102.

23. Lehmann, E., 1988, *Comparing Location Experiments*, The Annals of Statistics 16, 521-533.

24. Loury, G., 1981, *Intergenerational Transfers and the Distribution of Earnings*, Econometrica 49(4), 843-867.

25. Milgrom, P. R., 1981, *Good News and Bad News: Representation Theorems and Applications*, Bell Journal of Economics 12, 380-391.

26. Orazem, P. and Tesfatsion, L., 1997, *Macrodynamic Implications of Income-Transfer Policies for Human Capital Investment and School Effort*, Journal of Economic Growth 2, 305-329.

27. Orosel, G. O., 1996, *Informational Efficiency and Welfare in the Stock Market*, European Economic Review 40, 1379-1411.

28. Persico, N., 2000, *Information Acquisition in Auctions*, Econometrica 68, 135-148.

29. Rubinstein, Y. and Tsiddon, D., 2002, *Coping With Technological Changes: The Role of Ability in Making Inequality so Persistent*, Working Paper, The Foerder Institute for Economic Research, Tel Aviv University.

30. Schlee, E., 2001, *The Value of Information in Efficient Risk Sharing Arrangements*, American Economic Review 91(3), 509-524.

31. Stiglitz, J., 1975, *The Theory of 'Screening', Education, and the Distribution of Income*, American Economic Review 65(3), 283-300.

32. Sylwester, K., 2002a, *A Model of Public Education and Income Inequality with a Subsistence Constraint*, Southern Economic Journal 69(1), 144-158.

33. Sylwester, K., 2002b, *Can Education Expenditures Reduce Income Inequality?*, Economics of Education Review 21, 43-52.

34. Viaene, J. M. and Zilcha, I., 2002 *Capital Markets Integration, Growth, and Income Distribution*, European Economic Review 46, 301-327.