Meritocracy and the Inheritance of Advantage

José V. Rodríguez Mora*
University of Edinburgh & CEPR

Michael J. Watts†
University of Edinburgh

January 2015
Draft Version

Abstract

This paper examines the investment decisions of parents in their children and the intergenerational income relationship. Each is a function of the other, and together they determine the equilibrium in our economy. The model provides a new explanation of the Great Gatsby curve based on firms' efficient use of information. We use it to generate comparative statics results from an exogenous increase in either the degree of meritocracy (the quality of information on talent) or the degree of inherited advantage (the quality of information on parental income). The main result is that meritocracy is very akin to inherited advantage. Both have the same effect on mobility and inequality. This results in a perverse effect of meritocracy – it can *lower* mobility. Moreover, in a sufficiently unmeritocratic society, better information on talent can cause firms to use information on parental income *more*. Finally, greater inherited advantages might actually stimulate investment in human capital despite firms rewarding information on human capital less.

*School of Economics, University of Edinburgh, 31 Buccleuch Place, Edinburgh EH8 9JT, United Kingdom. Email: sevimora@gmail.com

†School of Economics, University of Edinburgh, 31 Buccleuch Place, Edinburgh EH8 9JT, United Kingdom. Email: michael.watts@ed.ac.uk

1 Introduction

The word *meritocracy* was coined by the sociologist Michael Young in his book "The Rise of Meritocracy" (Young (1958)). The book explains an imaginary history of the UK up to 2034 (the book is supposedly narrated in 2033). It presents the workings of a society where positions are allocated based solely on merit, and not on birth. This might have been thought of as a superb idea, but turns into a catastrophe in the book. The reason for this is that the distribution of talent is endogenous to the workings of society. The high reward for talent that meritocracy entails fosters inequality in the distribution of income. The children of the rich receive, in such a case, more and better education, and the degree of intergenerational mobility decreases. Meritocracy fosters both inequality and persistence in income across generations. The book finishes with the death of the supposed author at the hands of an anti-meritocratic revolution, when those left behind by history rebel against what they see as unjust inequalities that meritocracy has created.

Meritocracy, thus, was not at its inception a positive word. It had an ironic tone, and it was coined as a warning. It was recognized that it would give rise to a feeling of justice and would foster short-run efficiency, but it was doomed to increase inequality to the point of being unbearable for society.

It has been argued that a large part of the increase in inequality in recent decades could be attributed to a phenomenon akin to the one that the book captures. Our paper models the workings of meritocracy in a context where society would be inclined to treat the children of the rich in a more benevolent manner than it treats the children of the poor. Inherited advantages are meaningful and rational in the model that we develop. No money is left on the table. The market, and competition, ensure that no static inefficiencies take place. Firms pay workers based on their abilities (net of misallocation costs) to the greatest extent possible, given the available information. That information encompasses two things: a signal on the background of the worker, specifically on her parents income; and a signal on her talent or human capital. We say that the weight given by firms to the first signal creates inherited advantages, while the latter signal fosters meritocracy. The onset of meritocracy, understood as an increase in the accuracy of information on the quality of individuals, and not their background, ends up increasing the degree of inequality and the persistence of income across generations.

One of our main insights is that meritocracy is very akin to advantages. They have the same effect on mobility and inequality. Moreover, more meritocracy may, in general, even increase the value of coming from a good background.

The correlation between intergenerational income mobility and inequality was recently considered by Alan Krueger in a presentation he gave to the Center for American Progress (Krueger (2012)). He discussed the Great Gatsby Curve, the literary namesake of which provides an example, of sorts, of upward mobility. It shows a negative correlation between income inequality and mobility. There exist several explanations for this negative correlation in the literature. One such explanation is given by the distance effect in Hassler et al. (2007). In their model, greater mobility leads to a long-run equilibrium in which there are fewer unskilled workers relative to skilled. This reduces the distance between their two incomes, which is their measure of income inequality, and increases the ability of unskilled parents to pay for the education of their children. This feeds back into increased mobility. Another explanation is provided by Solon (2004). In Solons model, both intergenerational income elasticity and inequality are a function of the same factors, including the inheritance of income generating traits and more policy-related factors such as the progressivity of public human capital investment. This would again give rise to the observed negative correlation between income inequality and mobility. Our paper offers an alternative explanation the inheritance of advantage.

To be clear, we are not talking, as Solon did, about the greater opportunity of the children of the rich to accumulate talent. We assume that this is true, and model it through a capital market imperfection allowing the rich to invest more in their childrens development. What we are talking about is the greater efficacy with which the children of the rich inherit the ability to look talented. This relates our paper to the literature on statistical discrimination (such as Coate & Loury (1993)). Our approach differ from these papers' in that our firms discriminate based on endogenously determined variables and do so within a dynamic model. This naturally produces the observed negative correlation between mobility and inequality. As income inequality increases, people differ more, and it becomes easier to identify talented individuals within society. As firms become more certain about who is talented and who is not, this feeds back into income dispersion. In addition, the better firms get at identifying talent, the lower is mobility, because the talented tend to be from rich backgrounds. Additionally, since one of the ways which firms identify talent is through information on family background, this information may be used more in a more meritocratic society, further reducing mobility.

The paper proceeds in the following way. First, we model the investment of parents in the human capital development of their children. This decision is subject to any investment coming from their own income, and to the technology for identifying and rewarding talent used by firms. There is an equilibrium concept here: parents make their investment decisions taking as given the income determination function; and the income determination function which they use must be correct given their investment choices and the behaviour of firms. The bulk of the paper then investigates the behaviour of firms when choosing how much to pay a worker based on the limited information provided by the signals on background and human capital. Once the behavior of parents and firms is established, we solve for the equilibrium investment rule and income determination function.

Our main results are derived from a comparative statics analysis of the steady state economy. We exogenously vary the quality of the background signal (advantage) and the human capital signal (meritocracy) to see what effect this has.

First, there exists a multiplier effect whereby increased inequality improves the information available to firms when selecting workers, increasing discrimination, reducing mobility and encouraging further inequality in the incomes that they pay.

Second, the model implies a perverse effect of meritocracy: if a meritocratic society is one in which firms can more readily identify the quality of workers and pay them accordingly, this will decrease mobility. Contrary to what could be thought, meritocracy does not increase the degree of intergenerational mobility. More information on people's ability is bound to decrease intergenerational mobility because ability and background are correlated and, by increasing income dispersion, meritocracy increases the value of any existing information on people's ability. Firms discriminate more amongst workers of more diverse ability, providing advantages to the talented, but the talented predominantly come from richer backgrounds so this lowers intergenerational mobility.

An implication of this result is that societies with meritocratic institutions (those in which it is easy to signal talent) should tend to have more inequality and less mobility than those without. This is what we observe in the likes of the US and UK, with a hierarchy of higher education institutions from the Ivy League and Oxbridge downwards, compared to the Scandinavian countries with their more egalitarian educational institutions.

In addition to the effect which this has on the value assigned to merit, it also affects the value assigned to inherited advantages, which is itself interesting. One could imagine that in a society where talent is more readily identifiable, background would be unimportant. However, it is quite possible that the opposite is true. In fact we show that when society starts off sufficiently unmeritocratic, increases in meritocracy actually increase the weight which firms give to signals on parental income. Improvements in meritocracy increase the prevalence of inherited advantages.

Lastly, we look at the comparative static effects on investment in human capital and median income (we use median income as income is modelled as log normal and so the median, unlike the mean, is free of any effects coming from changes in the variance of log income).

Greater meritocracy leads to a greater proportion of income being invested in human capital development. Greater inherited advantages may actually do the same because, although human capital is rewarded less by firms, to the extent that it is rewarded at all parents are better able to provide these advantages to their grandchildren and future generations of their family by investing more heavily in their children.

We get the result we would expect on median income, provided misallocation of workers has a sufficiently large negative effect on output. More uncertainty, in the form of noisier signals, reduces median income.

2 Accumulation of Human Capital in an Extended Becker-Tomes Framework

We are going to develop a model that while departing only slightly from the standard Becker-Tomes paradigm, produces interesting and novel outcomes. Those outcomes will be explained in depth in sections 3 and 4. Before that, we consider the investment problem which faces parents. As in Becker & Tomes (1979, 1986), agents decide how much to invest in the human capital of their children. We assume that when they take such a decision they are uncertain about what that level of human capital will ultimately be. The problem they confront is:

$$W\left(Y_t^i\right) = \max_{X_t^i} \left\{ \ln C_t^i + \frac{1}{1+\delta} EW\left(Y_{t+1}^i | Y_t^i, X_t^i\right) \right\} \tag{1}$$

$s.t.$

$$C_t^i = Y_t^i - X_t^i; \quad X_t^i \geq 0 \tag{2}$$

$$Y_{t+1}^i \sim F\left(H_{t+1}^i, Y_t^i\right) \tag{3}$$

$$H_{t+1}^i \sim G\left(X_t^i\right) \tag{4}$$

where Y_t^i is the (lifetime) income of family i at generation t, C_t^i is their consumption, X_t^i is their investment in the human capital of their offspring, and H_{t+1}^i is the human capital that the offspring actually achieves. Equation 2 is the budget constraint that parents face, while equation 4 states that human capital is a stochastic function of the investment that parents make, without a role for other forms of inheritance (genetic or otherwise). We assume for simplicity that each generation of each family has exactly one member.

The novelty of our theory refers to equation 3, which states that children's income is a stochastic function of their human capital and the income that their parents have. In the next sections we will develop a theory such that in equilibrium the income of an individual is a stochastic function of her human capital and her parental income, which has effects *even controlling for the human capital of the child* . Moreover, the relative effect of both human capital and parental income are a function of the degree of inequality of the distribution of income, which is itself an endogenous object of the model.

Equations 3 and 4 together imply a relationship between the income of a child and the income and investment made by their parent. We will begin, in this section, by taking as given the following functional form of this relationship:

$$Y_{t+1}^i = e^{\gamma_0} \left(Y_t^i\right)^{\gamma_1} \left(X_t^i\right)^{\gamma_2} e^{\varepsilon_{t+1}^i} \tag{5}$$

where γ_0, γ_1 and γ_2 are endogenous parameters of this *income determination function* and ε_{t+1}^i is a stochastic error term with a certain mean $\bar{\varepsilon}$ and variance V_ε. At this stage we are guessing that this is the correct functional form, but in subsequent sections we will show that this is, in fact, the correct form of the income determination function and solve for the equilibrium values of γ_0, γ_1 and γ_2 and the functional form of the random variable ε_{t+1}^i.

Note that for equilibrium we must: define parent's investment behaviour where they take as given the income determination function in equation 5; show that, as a consequence of parent's investment behaviour, equation 5 is the correct income determination function. This section will concentrate on the former; section 4 will concentrate on the latter.

The last ingredient which we need to solve the parent's maximisation problem and find the equation for the accumulation of human capital is a functional form for equation 4. We assume that it is:

$$H_{t+1}^i = Z \left(X_t^i\right)^\alpha e^{\tilde{\omega}_{t+1}^i}; \quad \tilde{\omega}_{t+1}^i \sim N\left(-\frac{V_\omega}{2}, V_\omega\right) \tag{6}$$

where Z is a constant akin to total factor productivity, and $\tilde{\omega}_t^i$ is an iid shock normally distributed with a mean such that $E\left(e^{\tilde{\omega}_t^i}\right) = 1$. We assume that $\alpha \in (0, 1)$.

The following result characterizes the optimal investment decisions of dynasties:

Result 1. *The solution of the maximization problem in equation 1 requires that investment in education is a fixed proportion of the individual's income:* $X_t^i = \lambda Y_t^i$, *with:*

$$\lambda = \frac{\gamma_2}{1+\delta-\gamma_1} \tag{7}$$

The value function of agents is $W(Y) = A + B \ln Y$*, with*

$$A = \frac{\bar{\varepsilon} + \gamma_0 + \ln\left[(1+\delta) - (\gamma_1+\gamma_2)\right]^{[(1+\delta)-(\gamma_1+\gamma_2)]} + \ln \frac{\gamma_2^{\gamma_2}}{[(1+\delta)-\gamma_1]^{[(1+\delta)-\gamma_1]}}}{\left[(1+\delta) - (\gamma_1+\gamma_2)\right]\frac{\delta}{1+\delta}} \tag{8}$$

$$B = \frac{(1+\delta)}{(1+\delta) - (\gamma_1+\gamma_2)} \tag{9}$$

Notice that given that investment is a fixed proportion of income (result 1) and that human capital has a constant elasticity to investment (equation 6), it follows that:

Result 2. *Log human capital is a linear function of log parental income. The elasticity is an exogenous parameter* α*, while the constant depends on the investment rate* λ *:*

$$h_{t+1}^i = \ln Z + \alpha \ln \lambda - \frac{V_\omega}{2} + \alpha y_t^i + \omega_{t+1}^i \tag{10}$$

where $\omega_{t+1}^i \sim N(0, V_\omega)$ *is iid noise and* h_{t+1}^i *and* y_t^i *represent the log of the child's human capital and the log of parental income respectively.*

This is the equation for the accumulation of human capital. Human capital depends positively on parental income and the fraction of income which parents invest in their children. This equation is known to firms but the level of human capital of a particular worker is not.

3 The Market Value of Human Capital and Parental Income

We now know how parents invest in their children, and how human capital is accumulated, taking the income determination function 5 as given. The income determination function will itself be determined by how the market values the available information on human capital and parental income. The goal of this section is the explain how this process occurs. For this, we need to consider production in the child's generation.

Firms are assumed to have a production function which is increasing in the human capital of a worker, and decreasing in the square of the deviation of the human capital of the worker from the firm's belief about their human capital. That is, there is a cost to misallocation. If the firm incorrectly perceives the human capital of a worker, the worker is less productive. This is summarised in the following:

$$Y_{t+1}^i = E\left[\exp\left\{h_{t+1}^i - \frac{\theta}{2}\left(h_{t+1}^i - E(h_{t+1}^i|\Omega_{t+1}^i)\right)^2\right\}\right] \tag{11}$$
$$\theta \in R^+$$

Agent i receives an income equal to their expected human capital adjusted for the cost of misallocation. The parameter θ controls the magnitude of the misallocation cost and Ω_{t+1}^i is the set of available information. The information set consists of: (1) a public signal on the parental income of agent i: a_{t+1}^i; (2) a public signal on the human capital of agent i: m_{t+1}^i; (3) the distribution of income in the parent's generation; (4) the equation for the accumulation of human capital given in equation 10. We assume that log income in generation t is distributed normally, $y_t^i \sim N\left(\mu_{yt}, V_{yt}\right)$.

In what follows we will be interested in the steady state distribution of income. The equilibrium and comparative statics which we describe will be in steady state and so, for notational simplicity, we will drop the time subscripts on μ_y and V_y and assume that they are at the steady state values which we solve for below (we have already implicitly assumed that γ_0, γ_1 and γ_2 are at their steady state values). Then the information set available to firms to determine the payment to agent i is:

$$\Omega_{t+1}^i = \left\{a_{t+1}^i, m_{t+1}^i, \mu_y, V_y\right\} \tag{12}$$

An important part of our story is that human capital is not perfectly observed. Instead the market observes the signals a_{t+1}^i and m_{t+1}^i and infers from them what they can about the human capital of the child. The first signal, a_{t+1}^i, is a publicly available signal on parental income such that:

$$a_{t+1}^i = y_t^i + \varepsilon_{t+1}^{ia} \tag{13}$$

where $\varepsilon_{t+1}^{ia} \sim N(0, V_a)$ is iid noise. The second signal, m_{t+1}^i, is a publicly available signal on human capital such that:

$$m_{t+1}^i = h_{t+1}^i + \varepsilon_{t+1}^{im} \tag{14}$$

where $\varepsilon_{t+1}^{im} \sim N(0, V_m)$ is iid noise.

The market pricing of these signals is key to the workings of our model as it feeds into income variance, intergenerational mobility and the investment behaviour of parents.

Result 3. *Given the process of human capital accumulation in equation 10 and the information set given by equation 12, the posterior of the log of human capital is given by:*

$$h_{t+1}^i | \Omega_{t+1}^i \ \sim \ N\left(\mu_{h_{t+1}^i|\Omega_{t+1}^i}, V_{h_{t+1}^i|\Omega_{t+1}^i}\right)$$

with

$$\begin{aligned}
\mu_{h_{t+1}^i|\Omega_{t+1}^i} &= \beta_m m_{t+1}^i + (1-\beta_m)\left[\ln Z + \alpha \ln \lambda - \frac{V_\omega}{2} + \alpha\beta_a a_{t+1}^i + \alpha\left(1-\beta_a\right)\mu_y\right] \\
V_{h_{t+1}^i|\Omega_{t+1}^i} &= \beta_m V_m \\
\beta_a &= \frac{V_y}{V_y + V_a} \\
\beta_m &= \frac{\alpha^2\beta_a V_a + V_\omega}{\alpha^2\beta_a V_a + V_\omega + V_m}
\end{aligned}$$

The posterior belief about the human capital of individual i at time $t+1$ is a stochastic function of their observable signals, a_{t+1}^i and m_{t+1}^i. The variance in those beliefs is actually constant across individuals and (in steady state) across time.

Result 4. *Given the posterior belief about the log of human capital follows a normal distribu-*

tion

$$h^i_{t+1}|\Omega^i_{t+1} \ \sim \ N\left(\mu_{h^i_{t+1}|\Omega^i_{t+1}}, V_{h^i_{t+1}|\Omega^i_{t+1}}\right)$$

and income is given by equation 11, it follows that:

$$Y^i_{t+1} = \frac{1}{\sqrt{1+\theta V_{h^i_{t+1}|\Omega^i_{t+1}}}} \exp\left\{\frac{1}{2}\left(\frac{V_{h^i_{t+1}|\Omega^i_{t+1}}}{1+\theta V_{h^i_{t+1}|\Omega^i_{t+1}}}\right)\right\} \exp\left\{\mu_{h^i_{t+1}|\Omega^i_{t+1}}\right\} \tag{15}$$

By taking logs and substituting from result 3 we can find the log income of individual i with signals a^i_{t+1} and m^i_{t+1}:

$$\begin{aligned} y^i_{t+1} = {} & (1-\beta_m)\left[\ln Z + \alpha \ln \lambda - \frac{V_\omega}{2} + \alpha\left(1-\beta_a\right)\mu_y\right] - \frac{1}{2}\left[\ln\left(1+\theta\beta_m V_m\right) - \frac{\beta_m V_m}{1+\theta\beta_m V_m}\right] \\ & + \alpha\beta_a\left(1-\beta_m\right)a^i_{t+1} + \beta_m m^i_{t+1} \end{aligned} \tag{16}$$

Notice that β_m is the weight given to the signal of an agent's human capital when determining her income, and $\hat{\beta}_a = \alpha\beta_a\left(1-\beta_m\right)$ is the weight given to the signal on parent's income. Thus, if $\hat{\beta}_a$ is high, society is rewarding individuals who appear to have "good" backgrounds. In that case, we will say that advantages are prevalent in society. Likewise, a high value of β_m indicates that society is rewarding "merit", the available information on talent. Notice that there exists a certain trade-off between merit and advantages, as a larger value of β_m is in and by itself bound to decrease in $\hat{\beta}_a$, but it is by no means obvious that this will be the overall effect on the use of the parental income signal, as both β_a and β_m are endogenous. We will see that there are parameter changes that imply an increase in the prevalence of both meritocracy and advantages, while others affect them in opposite directions.

Given that a^i_{t+1} and m^i_{t+1} are both stochastic functions of y^i_t we can write the law of motion of the log of income:

$$y_{t+1}^{i} = \ln Z + \alpha \ln \lambda - \frac{V_{\omega}}{2} + \alpha\left(1-\beta_{a}\right)\left(1-\beta_{m}\right)\mu_{y} - \frac{1}{2}\left[\ln\left(1+\theta\beta_{m}V_{m}\right) - \frac{\beta_{m}V_{m}}{1+\theta\beta_{m}V_{m}}\right]$$
$$+ \alpha\left[\beta_{a}\left(1-\beta_{m}\right)+\beta_{m}\right] y_{t}^{i} + \alpha\beta_{a}\left(1-\beta_{m}\right)\varepsilon_{t+1}^{ia} + \beta_{m}\left(\varepsilon_{t+1}^{im} + \omega_{t+1}^{i}\right) \quad (17)$$

The intergenerational income elasticity is $\rho = \alpha\left[\beta_{a}\left(1-\beta_{m}\right)+\beta_{m}\right]$, and (after some manipulation) the law of motion of the variance of log income is:

$$V_{y_{t+1}} = \alpha^{2}\left[\beta_{a}\left(1-\beta_{m}\right)+\beta_{m}\right] V_{y_{t}} + \beta_{m}V_{\omega} \quad (18)$$

The larger the value of β_a or β_m the larger the dispersion of income next period. This interacts with the equations determining β_a and β_m (how much firms discriminate according to background and perceived merit). The more income variance there is, the larger β_a and β_m. Hence we have a feed back mechanism from discrimination, through income variance, to further discrimination. This is going to be a key feature of our model: any change is parameters which leads to more (or less) discrimination by firms will see this effect multiplied through its effect on income variance. We discuss this *feed back effect* in the context of some special cases of the model below.

We characterize the steady state of our economy in the following result:

Result 5. *Steady State. Given the process of human capital accumulation in equation 10, income given by equation 11, and the information set given by equation 12, there exists a unique steady state which is globally stable. In the steady state log income is normally distributed with* *variance being characterized by the (unique) solution of the following system of equations:*

$$V_y = \frac{\beta_m V_\omega}{1 - \alpha^2 \left[\beta_a \left(1 - \beta_m\right) + \beta_m\right]} \tag{19}$$

$$\beta_a = \frac{V_y}{V_y + V_a} \tag{20}$$

$$\beta_m = \frac{\alpha^2 \ \beta_a \ V_a + V_\omega}{\alpha^2 \ \beta_a \ V_a + V_\omega + V_m} \tag{21}$$

The steady state mean of log income and intergenerational correlation of income are given by:

$$\mu_y = \frac{\ln Z + \alpha \ln \lambda - \frac{V_\omega}{2} - \frac{1}{2}\left[\ln\left(1 + \theta\beta_m V_m\right) - \frac{\beta_m V_m}{1+\theta\beta_m V_m}\right]}{1 - \alpha} \tag{22}$$

$$\rho = \alpha \left[\beta_a \left(1 - \beta_m\right) + \beta_m\right] \tag{23}$$

3.1 An Exogenous Increase in Inherited Advantages

In this section we want to look at the effect of exogenously changing V_a. A reduction in V_a, all else equal, improves the quality of the signal on parental income, providing greater advantages to those from more affluent backgrounds. We will begin by assuming $V_m \to \infty$. This exogenously shuts off the meritocracy signal, since it is useless and completely disregarded by firms, and allows us to consider the comparative statics of V_a on the parental income signal in isolation. In particular, we are interested in the comparative statics effects on income variance, discrimination, and intergenerational mobility. We will then allow for positive but finite values of V_m to look at the comparative static effects of a change in inherited advantages where the two signals can interact.

The first thing to note is that when $V_m \to \infty$, the unique steady state value of V_y is zero for all V_a. Therefore there are no comparative statics results to be gained from V_a. In order to allow us to consider the parental income signal is isolation, we modify the model slightly.

Specifically, we change the production function to:

$$Y_{t+1}^{i}=E\left[\exp\left\{h_{t+1}^{i}-\frac{\theta}{2}\left(h_{t+1}^{i}-E\left(h_{t+1}^{i}|\Omega_{t+1}^{i}\right)\right)^{2}\right\}\right]\exp\left\{u_{t+1}^{i}\right\}, \quad \theta\in R^{+}$$

with $u_{t+1}^{i}\sim N\left(0,V_{u}\right)$ and $V_{u}>0$. The only alteration is the addition of this stochastic noise term which generates some income variance, even in the absence of any discrimination by firms amongst workers, and so prevents firms from ignoring the different backgrounds of their workers.

When we do this, we find that $\alpha\beta_{a}$ is the weight given to the parental income signal in the income determination equation, and so measures the extent to which firms discriminate based on background. The law of motion of log income is given by the following, which is analogous to equation 17 in the general model.

$$\begin{aligned}y_{t+1}^{i}&=\ln Z+\alpha\ln\lambda-\frac{V_{\omega}}{2}-\frac{1}{2}\left[\ln\left(1+\theta\left[\alpha^{2}\ \beta_{a}V_{a}+V_{\omega}\right]\right)-\frac{\alpha^{2}\beta_{a}V_{a}+V_{\omega}}{1+\theta\left[\alpha^{2}\beta_{a}V_{a}+V_{\omega}\right]}\right]\\&\quad+\alpha\left(1-\beta_{a}\right)\mu_{y}+\alpha\beta_{a}y_{t}^{i}+\alpha\beta_{a}\varepsilon_{t+1}^{ai}+u_{t+1}^{i}\end{aligned}\tag{24}$$

The steady state income variance and level of discrimination are found by the (unique) solution to the following two equations:

$$V_{y}=\frac{V_{u}}{1-\alpha^{2}\beta_{a}}\tag{25}$$

$$\beta_{a}=\frac{V_{y}}{V_{y}+V_{a}}\tag{26}$$

Intergenerational mobility is then given by $\rho=\alpha\beta_{a}$.

Equations 25 and 26 completely define the steady state. Together they also show the feed back mechanism which we want to stress. Equation 25 indicates that the larger the role of background in the determination of income (the larger β_{a} is for a given α), the larger the degree of inequality. That is, the greater the advantage given to people from good backgrounds, the more inequality there is. Equation 26, poses the feed back: the more inequality there is, the more that the market will rationally give extra advantages to people from good backgrounds.

A consequence of the feed back mechanism is a multiplier effect that we characterize in the following result[1]:

Result 6. *The full effect of an exogenous change in the parameters α, V_a or V_u is greater than the partial effect due to the feed back from income inequality to discrimination*

$$\left|\frac{dV_y}{d\chi}\right| > \left|\frac{\partial V_y}{\partial \chi}\right| \;\; ; \;\; \left|\frac{d\beta_a}{d\chi}\right| > \left|\frac{\partial \beta_a}{\partial \chi}\right| \quad \textit{where} \quad \chi = \{\alpha, V_a, V_u\}$$

Any exogenous change in parameters that has a direct effect on either the degree of inequality or the use of information on background, will have its effect amplified via the other variable. This is, if there is an increase of the variance of income shocks, V_u, its effect on the variance of income is larger than the direct one on 25, because the increase in inequality produces an increase in the degree of discrimination, which feeds back into inequality, which feeds back into discrimination, and so on.

More important, perhaps, for our discussion is that a decrease in income variance *conditional on the signal* , if as a consequence of a reduction of the informational noise, results in an increase of the *unconditional* income variance, discrimination and intergenerational correlation of incomes. We summarize this in the following result:

Result 7. *For $V_m \to \infty$, the steady state V_y, ρ and β_a are all decreasing in V_a.*

If there is an increase in the precision of the signal, this has an immediate effect of increasing β_a (see equation 26 for a fixed V_y), but its long run effect on discrimination will be larger, as inequality will increase as a consequence of the increased discrimination.

Notice that the result is not obvious. Let's look again at equation 24 and consider an exogenous decrease in the amount of informational noise (a smaller value of V_a). For a given β_a the variance of income will fall in the subsequent generation. However, once we allow β_a to

[1] A second consequence is that if $\alpha \geq 1$ there are two stable steady states. One is analogous to the one we report here. The second has infinite variance of log income and, consequently $\beta_a = 1$. We abstract from this (interesting) possibility by assuming that $\alpha < 1$.

adjust endogenously we find just the opposite: the less noise we feed into the economy in the form of informational noise, the *larger* the variance of income in the economy. The reason is precisely *because* the increase in precision translates into more discrimination (less noise, larger β_a), which itself translates into a larger intergenerational correlation ($\alpha\beta_a$). Thus, the income process becomes more persistent, and this is bound to increase the *unconditional* variance of income, *even if what we did in the first place was to decrease the conditional variance of income.*

We can now add the human capital signal back into the model by assuming V_m is positive but finite. This leads to the more general result:

Result 8. *An increase in the accuracy of the signal on background (a decrease of V_a) results, in steady state, in more inequality, greater persistence of income across generations, more discrimination based on perceptions of the background of an agent, and a smaller elasticity of income to the signal on ability:*

$$\frac{dV_y}{dV_a} < 0; \quad \frac{d\rho}{dV_a} < 0; \quad \frac{d\beta_a}{dV_a} < 0; \quad \frac{d\hat{\beta}_a}{dV_a} < 0; \quad \frac{d\beta_m}{dV_a} > 0$$

More accurate information on the background of an individual is going to increase the attention that firms give to on this signal, increasing the persistence of income across generations and its variance across individuals. The reasons are exactly the same as what we saw in the discussion of result 7. More information implies more connection with the past, which increases income dispersion and makes information on the past even more valuable.

Notice that the effect on meritocracy is the opposite. Better information on background results in *less* elasticity of income on the ability of individuals. This might look surprising, given that we say that income dispersion increases and thus the problem of firms determining abilities seems more severe the more accurate the information of background. The reason is that albeit the unconditional variance of income increases, the variance of income *conditional on the signal a_t^i decreases.* Thus, the dispersion of human capital *conditional on a_t^i* decreases.

More accurate information on people's background does not only increase inequality, the monetary value of appearing from a good background and the degree of persistence in incomes. It also decreases the value of direct information on the human capital of individuals. There is a crowding-out effect on meritocracy, its space in the determination of an individual's income being taken over by inherited advantages. There is less intergenerational mobility, and the talent that is not derived from parental income (ω_t^i) has less value in the determination of income.

3.2 An Exogenous Increase in Meritocracy

Next we want to consider the effect of exogenously reducing V_m. A reduction in V_m, all else equal, improves the quality of the human capital signal, providing greater advantages to those of greater talent. Similar to section 3.1, we will first look at the human capital signal in isolation, by assuming $V_a \to \infty$ and shutting off the parental income signal, then we will look at the effect of an increase in meritocracy where the two signals can interact. Unlike in that section however, we use the production function given in equation 11 (and will continue to use it for the rest of the paper. This is equivalent to $V_u = 0$ from the previous section).

Taking $V_a \to \infty$ changes the steady state defined by equations 19 to 21 in the following way:

$$V_y = \frac{\beta_m V_\omega}{1 - \alpha^2 \beta_m}$$

$$\beta_a = 0$$

$$\beta_m = \frac{\alpha^2 V_y + V_\omega}{\alpha^2 V_y + V_\omega + V_m} = \frac{V_h}{V_h + V_m}$$

The intergenerational correlation of incomes is given by $\rho = \alpha\beta_m$.

Firms no longer discriminate based on parental income since they effectively have no information on which to do so. The same feedback mechanism as in section 3.1 is still at play. The more that firms use the existing information about the human capital of an individual (the larger β_m is), the larger the dispersion of income between individuals. The inheritance process determined by equation 10 translates this increase of income dispersion into an increase of the dispersion of human capital in the following generation. But this increased dispersion of the priors on ability makes the existing information about the human capital of any specific agents more valuable, increasing β_m even further.

With minimum manipulation we can show the following interesting comparative statics:

Result 9. *For $V_a \to \infty$, the steady state values of V_h, V_y, ρ and β_m are all decreasing in V_m.*

Notice its implication. The more meritocratic a society is, the greater the information that is available to firms on the abilities of individuals, the more persistent income shocks will be, and the more inequality there will be.

Contrary to what could be thought, meritocracy does not increase the degree of intergenerational mobility. It decreases it. More information on people's ability is bound to decrease intergenerational mobility because ability and background are correlated and, by increasing income dispersion, meritocracy *increases the value of any existing information on people's ability*. Firms discriminate more amongst workers of more diverse ability, providing advantages to the talented, but the talented predominantly come from richer backgrounds so this lowers intergenerational mobility.

Meritocracy and the advantages generated by the perceptions on one's background work in a very similar manner. As a consequence of more inequality, firms face more uncertainty on the abilities of any specific worker. Consequently, they place more value on any information correlated with her ability, be this direct information on her ability (which would produce meritocracy) or information on her background (which would produce advantages). Firms will use this information more, which itself suffices to increase inequality even further, leading to even greater discrimination as the feed back process continues.

We now consider the effect of an exogenous improvement in the quality of the human capital signal when both signals are available and useful to the firm (i.e. V_a is positive but finite).

Result 10. *An increase in the accuracy of the signal on ability (a decrease of V_m) results, in steady state, in more inequality, greater persistence of income across generations, a larger* *elasticity of income to the signal on ability and more weight given to the signal on background when evaluating an agent's parental income (which is what β_a measures):*

$$\frac{dV_y}{dV_m} < 0; \quad \frac{d\rho}{dV_m} < 0; \quad \frac{d\beta_m}{dV_m} < 0; \quad \frac{d\beta_a}{dV_m} < 0$$

Moreover, given a set of values for $\alpha \in (0,1)$ and $V_a \in R^+$ $(V_a < \infty)$, there exists a variance of the signal on ability $\hat{V}_m$ such that $\left(0 < \hat{V}_m < \infty\right)$

$$\textit{If } V_m < \hat{V}_m, \textit{ then } \frac{d\hat{\beta}_a}{dV_m} > 0$$
$$\textit{If } V_m > \hat{V}_m, \textit{ then } \frac{d\hat{\beta}_a}{dV_m} < 0$$

The value of $\hat{\beta}_a$, the weight given to the signal on background when evaluating an agent's human capital, is maximal if $V_m = \hat{V}_m$.

If society is better endowed to judge its members merit, it is doomed to increase the dispersion of their incomes (paying more to those judged to be better). This increased dispersion has effects on both the value assigned to merit, β_m, and the value assigned to "advantages", $\hat{\beta}_a$.

First it increases the dispersion of the abilities of the children, thus feeding back into increased underlying uncertainty and value of the signal on merit in the following period. Thus, not surprisingly, better information on talent results into a more meritocratic society.

The effects on the weight given to background when determining income ($\hat{\beta}_a$) are more complicated. First of all, there is a "crowding-out effect" of opposite sign to that in result 8. Better information on talent makes you place less weight on background, as the ratio of the variance of abilities *conditional on* a_t^i to the variance of the signals of talent increases: more weight on merit implies less weight on a_t^i, as it is clear from the fact that β_m enters negatively in $\hat{\beta}_a = \alpha\beta_a (1 - \beta_m)$. However, there is an effect on the opposite direction too: as income variance increases, the signal on background becomes more valuable in judging parental income. The net effect on $\hat{\beta}_a$ depends on the relative size of the effects on β_a and β_m.

We can understand the net effect by doing the following mental exercise. Imagine that V_m were very low (and thus, firms know ability quite well). In that case β_m would be very large (close to one), and the effect of the increase of β_a would be very small ($\frac{\partial \hat{\beta}_a}{\partial \beta_a} = \alpha (1 - \beta_m)$). In this case, the net effect of an decrease in V_m (an increase in the amount of information on ability) would be a decrease in the weight of advantages and an increase in the weight of meritocracy. There would be a trade-off between meritocracy and advantages. Imagine the polar opposite where V_m is very large. In such a case β_m would be close to 0 and the value of background information would be dominant. Any increase in the quality of information on ability would increase both meritocracy (β_m) and advantages ($\hat{\beta}_a$).

In any case, notice that the degree of intergenerational mobility *always* decreases whenever the society becomes more meritocratic as a consequence of a decrease of V_m. This occurs both when there is a trade-off (and advantages become less important) or when there is not. This is a consequence of inheritance. The talented become richer, and thus incomes are bound to get more dispersed, and this increased dispersion of incomes is going to be translated into a further dispersion of abilities. Abilities then, being better evaluated, translate into more income for the children of the rich *even if it is perfectly possible that firms care less about the background of agents.*

This is one of the main insights of our paper. Meritocracy in and of itself is not going to increase intergenerational mobility or decrease the prevalence of inheritance. And this is bound to happen even if an increase in meritocracy does produce a decrease in the advantages associated with being from a good background, which is by no means a foregone conclusion.

This is not to say that meritocracy is a bad thing. In the next section, for instance, we will show that in any reasonable context it increases investment in human capital accumulation. The significance of our result is to notice that there are two roads that lead to countries having low intergenerational mobility and high inequality. One is the preeminence of advantages, but a very different road leading to the same mobility and inequality is the meritocratic one. If the aspects of reality that one focuses on are limited to the degree of mobility and inequality, meritocracy and inherited advantages are almost equivalent. They produce the same effects.

Both of them produce a negative correlation between inequality and mobility, reproducing the result depicted in the Great Gatsby curve.

4 Equilibrium

We now need to solve for the equilibrium in our model. Recall that in section 2 we defined an equilibrium to consist of the following: a rule for parent's investment behaviour, *taking the income determination function in equation 5 as given*; and an income determination function of the form given in equation 5 *which results from parent's investment behaviour*. The former is given by equation 7 for some (equilibrium) values of γ_1 and γ_2. Our aim in this section is to describe the income determination function which results from the investment rule in 7, and to find the equilibrium values of γ_0, γ_1 and γ_2 and the function ε_t^i. They are provided in the following result:

Result 11. *Equilibrium. The equilibrium stochastic process of income as a function of parental income and investment is* $Y_{t+1}^i = e^{\gamma_0} \left(Y_t^i\right)^{\gamma_1} \left(X_t^i\right)^{\gamma_2} e^{\varepsilon_{t+1}^i}$ *with:*

$$\begin{aligned}
\gamma_0 &= \ln Z + \alpha\left(1-\beta_m\right)\left[\left(1-\beta_a\right)\mu_y + \ln\lambda\right] - \frac{1}{2}\left[\ln\left(1+\theta\beta_m V_m\right) - \frac{\beta_m V_m}{1+\theta\beta_m V_m} + V_\omega\right] \\
\gamma_1 &= \alpha\beta_a\left(1-\beta_m\right) \\
\gamma_2 &= \alpha\beta_m \\
\varepsilon_{t+1}^i &= \alpha\beta_a\left(1-\beta_m\right)\varepsilon_{t+1}^{ai} + \beta_m\left(\omega_{t+1}^i + \varepsilon_{t+1}^{mi}\right)
\end{aligned}$$

and, consequently, the equilibrium share of income invested in children's education is:

$$\lambda = \frac{\alpha\beta_m}{1+\delta-\alpha\beta_a\left(1-\beta_m\right)} \tag{27}$$

Results 5 and 11 define our equilibrium. The earlier result on steady state is important because the extent to which firms discriminate based on the available signals, and the resulting mean (log) income level, play a role in the equilibrium income determination function.

The final thing which remains for us to examine is the effect of a change in the accuracy of the signals on investment and median income.

Result 12. *An increase in the accuracy of the human capital signal (a decrease of V_m) results, in steady state, in an increase in the proportion of income invested in education. An increase in the accuracy of the signal on background (a decrease of V_a) may increase or decrease investment. A sufficient condition for $\frac{d\lambda}{dV_a} < 0$ is:*

$$\frac{V_m}{V_\omega}\left[\frac{1-\alpha^2\left[\beta_a\left(1-\beta_m\right)+\beta_m\right]}{\left[(1+\delta)-\alpha\beta_a\right]}\right] > 1$$

The first effect is fairly intuitive. Parent's respond to the greater rewards arising from the human capital signal by investing more heavily in the human capital of their child. In addition, to the extent that human capital is rewarded, parents can create greater advantages for their more distant descendants by investing in their children and providing a more favourable background for their grandchildren.

The effect of a decrease in V_a is somewhat different. Better quality information about parental income causes firms to give more weight to the background signal but *less* to the ability signal. To the extent that ability is still rewarded, by investing more you can create advantages for your grandchildren and future generations of your family which will be more readily rewarded. This encourages greater investment. On the other hand, the value of private investment in your children decreases as β_m decreases. This discourages investment.

Result 13. *When misallocation is sufficiently costly (θ is above a threshold level), an increase in the accuracy of either signal will raise median income. The threshold value above which this occurs for the ability signal, $\bar{\theta}$, solves:*

$$\frac{2\alpha}{\lambda}\frac{\frac{d\lambda}{dV_m}}{\left[\beta_m+V_m\frac{d\beta_m}{dV_m}\right]}=\frac{\left[\bar{\theta}\left(1+\bar{\theta}\beta_m V_m\right)-1\right]}{\left(1+\bar{\theta}\beta_m V_m\right)^2} \tag{28}$$

The threshold value above which this occurs for the background signal, $\hat{\theta}$, solves:

$$\frac{2a}{\lambda}\frac{\frac{d\lambda}{dV_a}}{V_m\frac{d\beta_m}{dV_a}} = \frac{\left[\hat{\theta}\left(1+\hat{\theta}\beta_m V_m\right)-1\right]}{\left(1+\hat{\theta}\beta_m V_m\right)^2} \tag{29}$$

It is worth noting that the left-hand side of 28 is negative. As a result, $\theta \geq 1$ is a sufficient condition for a fall in V_m to raise median income.

Provided misallocation of workers has a sufficiently large negative effect on output, we get the result we would expect. More uncertainty, in the form of noisier signals, reduces median income.

It might seem strange that an increase in the accuracy of the human capital signal could decrease median income, particularly since we have shown in result 12 that it increases investment in human capital. To examine this, let us imagine there are no misallocation costs ($\theta = 0$). The log of median income is then given by:

$$\mu_y = \frac{\ln Z - \frac{V_\omega}{2} + \alpha \ln \lambda + \frac{\beta_m V_m}{2}}{1-\alpha}$$

The last two terms in the numerator are endogenous. The λ term is increasing in the accuracy of the human capital signal; the $\beta_m V_m$ term is decreasing. It is the presence of $\beta_m V_m$ which can cause median income to fall. This term is the conditional variance of the log of human capital. It enters the income determination function because the conditional mean of human capital (upon which firms base their wage decisions) depends on the conditional variance of the log of human capital (due to human capital having a log normal distribution). Although a reduction in the noise in the human capital signal has a mean preserving effect on the log of human capital, it does not have a mean preserving effect on the *level* of human capital. It reduces it. As a result, a reduction in noise has a negative effect on median income which may or may not be overcome by increased investment.

The addition of a misallocation cost, which acts in the opposite direction to this conditional variance effect, ensures that, for a sufficiently large cost of misallocation, median income is higher in the presence of less uncertainty. In this case, more meritocratic societies have higher median income.

5 Conclusion

This paper has examined the investment decisions of parents in their children and the intergenerational income relationship (or Becker-Tomes equation). Each was a function of the other, and together they determine the equilibrium in our economy. The model provided a new explanation of the Great Gatsby curve based on firms' efficient use of information. We used it to generate comparative statics results from an exogenous increase in either the degree of meritocracy or the degree of inherited advantage.

One of the most interest results was that, while on the face of it these might seem like very different exercises, the effects of a change in either were very similar. In fact, they had exactly the same qualitative effects on inequality and intergenerational mobility. Whether talent was becoming more easily recognisable, or a wealthy upbringing was, both had the effect of increasing the degree to which firms could discriminate amongst workers. This not only increased inequality, but started a process whereby feedback effects through the degree of discrimination further increased inequality. Mobility fell because the information which firms were using was either directly related to parental income or positively correlated with it.

A further important, and seemingly perverse result, was that an increase in the degree of meritocracy may actually increase the degree of inherited advantage. That is, as talent becomes increasingly easy to identify, the increase in inequality it creates improves the quality of the signal on parental income. As a result, firms use it more. We might naturally assume that better quality information on a worker's talent would cause a substitution away from information on parental income and towards that on human capital. It does, for a fixed income variance. But income variance is not fixed and endogenously responds upwards, increasing the use of both signals. The net effect of these substitution and variance effects is an increase in the prevalence of inherited advantages when society starts off sufficiently unmeritocratic.

A final result to emphasise comes from the effect of increasing advantages on investment. One might presume that if information on the circumstances of birth determine income to a larger extent, and information on talent less so, that this would certainly decrease investment in human capital. The reason it may not, is that while the talent you can generate for your children is less well rewarded, to the extent that it is rewarded at all it can create advantages which are inherited by your grandchildren and great-grandchildren.

A Proof of Result 1

Proof. Solving the program:

$$W\left(Y_t^i\right) = \max_{X_t^i} \left\{ \ln\left[Y_t^i - X_t^i\right] + \frac{1}{1+\delta} EW\left(e^{\gamma_0}\left(Y_t^i\right)^{\gamma_1}\left(X_t^i\right)^{\gamma_2} e^{\varepsilon_t^i}\right)\right\}$$

First we prove that parents invest a fixed percentage of their income in their children:

$$X_t^i = \lambda Y_t^i$$

Guess

$$W\left(Y_t^i\right) = A + B \ln Y_t^i$$

The Euler equation is:

$$\frac{1}{Y_t^i - X_t^i} = \frac{1}{1+\delta} \frac{\partial EW\left(e^{\gamma_0}\left(Y_t^i\right)^{\gamma_1}\left(X_t^i\right)^{\gamma_2} e^{\varepsilon_t^i}\right)}{\partial X_t^i}$$

$$EW\left(e^{\gamma_0}\left(Y_t^i\right)^{\gamma_1}\left(X_t^i\right)^{\gamma_2} e^{\varepsilon_t^i}\right) = A + B\left[\bar{\varepsilon} + \gamma_0 + \gamma_1 \ln Y_t^i + \gamma_2 \ln X_t^i\right]$$

So, the Euler becomes:

$$\frac{1}{Y_t^i - X_t^i} = \frac{1}{1+\delta} B\gamma_2 \frac{1}{X_t^i}$$

Implying:

$$X_t^i = \frac{B\gamma_2}{(1+\delta) + B\gamma_2} Y_t^i$$

and

$$C_t^i = \frac{(1+\delta)}{(1+\delta) + B\gamma_2} Y_t^i$$

Now, substituting X_t^i into the expectation:

$$\begin{aligned} EW\left(Y_{t+1}^i | Y_t^i, X_t^i\right) &= EW\left(e^{\gamma_0}\left(Y_t^i\right)^{\gamma_1}\left(X_t^i\right)^{\gamma_2} e^{\varepsilon_t^i}\right) \\ &= A + B\left[\bar{\varepsilon} + \gamma_0 + (\gamma_1 + \gamma_2) \ln Y_t^i + \gamma_2 \ln \frac{B\gamma_2}{(1+\delta) + B\gamma_2}\right] \end{aligned}$$

And the value function should be:

$$W\left(Y_t^i\right) = \ln C_t^i + \frac{1}{(1+\delta)} EW\left(Y_{t+1}^i | Y_t^i, X_t^i\right)$$

$$\begin{aligned} W\left(Y_t^i\right) = & \ln \frac{(1+\delta)}{(1+\delta) + B\gamma_2} + \ln Y_t^i + \frac{A}{(1+\delta)} \\ & + \frac{B}{(1+\delta)} \left[\bar{\varepsilon} + \gamma_0 + (\gamma_1 + \gamma_2) \ln Y_t^i + \gamma_2 \ln \frac{B\gamma_2}{(1+\delta) + B\gamma_2}\right] \end{aligned}$$

So, if the guess is right:

$$A = \ln \frac{(1+\delta)}{(1+\delta) + B\gamma_2} + \frac{A}{(1+\delta)} + \frac{B}{(1+\delta)} \left[\bar{\varepsilon} + \gamma_0 + \gamma_2 \ln \frac{B\gamma_2}{(1+\delta) + B\gamma_2}\right]$$

and

$$B = \frac{B}{(1+\delta)} (\gamma_1 + \gamma_2) + 1$$

Solving for B

$$B = \frac{(1+\delta)}{(1+\delta) - (\gamma_1 + \gamma_2)}$$

which should be positive. We will show that the equilibrium values of γ_1 and γ_2 are always such that this happens. Finally, solving for A

$$\delta A = (1+\delta) \ln \frac{(1+\delta)}{(1+\delta) + B\gamma_2} + \frac{(1+\delta)}{(1+\delta) - (\gamma_1 + \gamma_2)} \left[\bar{\varepsilon} + \gamma_0 + \gamma_2 \ln \frac{B\gamma_2}{(1+\delta) + B\gamma_2}\right]$$

It is useful to notice that

$$1 - \lambda = \frac{(1+\delta)}{(1+\delta) + B\gamma_2} = \frac{(1+\delta) - (\gamma_1 + \gamma_2)}{(1+\delta) - \gamma_1} = 1 - \frac{\gamma_2}{[(1+\delta) - \gamma_1]}$$

so:

$$\lambda = \frac{\gamma_2}{[(1+\delta) - \gamma_1]}$$

$$\delta A = (1+\delta)\ln\left[1 - \frac{\gamma_2}{[(1+\delta)-\gamma_1]}\right] + \frac{(1+\delta)}{(1+\delta)-(\gamma_1+\gamma_2)}\left[\bar{\varepsilon} + \gamma_0 + \gamma_2 \ln\frac{\gamma_2}{[(1+\delta)-\gamma_1]}\right]$$

$$A = \frac{\bar{\varepsilon} + \gamma_0 + \ln\left[(1+\delta)-(\gamma_1+\gamma_2)\right]^{[(1+\delta)-(\gamma_1+\gamma_2)]} + \ln\frac{\gamma_2^{\gamma_2}}{[(1+\delta)-\gamma_1]^{[(1+\delta)-\gamma_1]}}}{[(1+\delta)-(\gamma_1+\gamma_2)]\,\frac{\delta}{1+\delta}}$$

□

B Proof of Result 2

Proof. This follows directly from the human capital accumulation equation 6 and the investment rule, $X_t^i = \lambda Y_t^i$. □

C Proof of Result 3

Proof. Since $(h_{t+1}^i, a_{t+1}^i, m_{t+1}^i)$ has a multivariate normal distribution, we can appeal to the conditional normal p.d.f. to find the distribution of $h_{t+1}^i|\Omega_{t+1}^i$.

Let $\mathbf{X}$ be a partitioned multivariate normal random vector with $\mathbf{X}^T = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{bmatrix}$, where $\mathbf{X}_1 = \begin{bmatrix} h_{t+1}^i \end{bmatrix}$ and $\mathbf{X}_2^T = \begin{bmatrix} a_{t+1}^i & m_{t+1}^i \end{bmatrix}$. The mean of $\mathbf{X}$ at time t is given by:

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

where $\mu_1 = [\mu_h]$ and $\mu_2^T = [\mu_y \;\; \mu_h]$.

The variance-covariance matrix of $\mathbf{X}$ at time t is given by:

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

where:

$$\mathbf{\Sigma}_{11} = \begin{bmatrix} V_h \end{bmatrix}$$

$$\mathbf{\Sigma}_{12} = \begin{bmatrix} \alpha V_y & V_h \end{bmatrix} = \mathbf{\Sigma}_{21}^T$$

$$\mathbf{\Sigma}_{22} = \begin{bmatrix} V_y + V_a & \alpha V_y \\ \alpha V_y & V_h + V_m \end{bmatrix}$$

Then the distribution of h_{t+1}^i conditional on Ω_{t+1}^i is univariate normal with the following mean and variance:

$$\begin{aligned}
&E\left(h_{t+1}^i | \Omega_{t+1}^i\right) \\
&= \mu_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\left(\mathbf{x}_2 - \mu_2\right) \\
&= \mu_h + \frac{1}{\left(V_h + V_m\right)\left(V_y + V_a\right) - \alpha^2 V_y^2} \begin{bmatrix} \alpha V_y & V_h \end{bmatrix} \begin{bmatrix} V_h + V_m & -\alpha V_y \\ -\alpha V_y & V_y + V_a \end{bmatrix} \begin{bmatrix} a_{t+1}^i - \mu_y \\ m_{t+1}^i - \mu_h \end{bmatrix} \\
&= \mu_h + \alpha\beta_a\left(1 - \beta_m\right)\left[a_{t+1}^i - \mu_y\right] + \beta_m\left[m_{t+1}^i - \mu_h\right] \\
&= \beta_m m_{t+1}^i + \left(1 - \beta_m\right)\left[\ln Z + \alpha \ln \lambda - \frac{V_\omega}{2} + \alpha\beta_a a_{t+1}^i + \alpha\left(1 - \beta_a\right)\mu_y\right]
\end{aligned}$$

and,

$$\begin{aligned}
&Var\left(h_{t+1}^i | \Omega_{t+1}^i\right) \\
&= \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21} \\
&= V_h - \frac{1}{\left(V_h + V_m\right)\left(V_y + V_a\right) - \alpha^2 V_y^2} \begin{bmatrix} \alpha V_y & V_h \end{bmatrix} \begin{bmatrix} V_h + V_m & -\alpha V_y \\ -\alpha V_y & V_y + V_a \end{bmatrix} \begin{bmatrix} \alpha V_y \\ V_h \end{bmatrix} \\
&= V_h - \alpha^2 V_y \beta_a\left(1 - \beta_m\right) - \beta_m V_h \\
&= \alpha^2 V_y\left(1 - \beta_m\right)\left(1 - \beta_a\right) + \left(1 - \beta_m\right)V_\omega \\
&= \beta_m V_m
\end{aligned}$$

□

D Proof of Result 4

Proof. We need to calculate

$$
\begin{aligned}
Y_{t+1}^i &= E\left[\exp\left\{h_t^i - \frac{\theta}{2}\left(h_{t+1}^i - \mu_{h_{t+1}^i|\Omega_{t+1}^i}\right)^2\right\}\right] \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{V_{h_{t+1}^i|\Omega_{t+1}^i}}} \exp\left\{h_t^i - \frac{\theta}{2}\left(h_t^i - \mu_{h_{t+1}^i|\Omega_{t+1}^i}\right)^2 - \frac{1}{2}\frac{\left(h_t^i - \mu_{h_{t+1}^i|\Omega_{t+1}^i}\right)^2}{V_{h_{t+1}^i|\Omega_{t+1}^i}}\right\} dh_t^i
\end{aligned}
$$

After some manipulation, this becomes:

$$
\begin{aligned}
Y_{t+1}^i &= \frac{1}{\sqrt{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}} \exp\left\{\frac{1}{2}\left(\frac{V_{h_{t+1}^i|\Omega_{t+1}^i}}{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}\right) + \mu_{h_{t+1}^i|\Omega_{t+1}^i}\right\} \\
&\cdot \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{\frac{V_{h_{t+1}^i|\Omega_{t+1}^i}}{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}}} \exp\left\{-\frac{1}{2}\frac{\left(h_t^i - \left[\mu_{h_{t+1}^i|\Omega_{t+1}^i} + \left(\frac{V_{h_{t+1}^i|\Omega_{t+1}^i}}{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}\right)\right]\right)^2}{\left(\frac{V_{h_{t+1}^i|\Omega_{t+1}^i}}{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}\right)}\right\} dh_t^i \\
&= \frac{1}{\sqrt{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}} \exp\left\{\frac{1}{2}\left(\frac{V_{h_{t+1}^i|\Omega_{t+1}^i}}{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}\right) + \mu_{h_{t+1}^i|\Omega_{t+1}^i}\right\}
\end{aligned}
$$

Therefore, income of individual i is given by:

$$
Y_{t+1}^i = \frac{1}{\sqrt{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}} \exp\left\{\frac{1}{2}\left(\frac{V_{h_{t+1}^i|\Omega_{t+1}^i}}{1+\theta V_{h_{t+1}^i|\Omega_{t+1}^i}}\right)\right\} \exp\left\{\mu_{h_{t+1}^i|\Omega_{t+1}^i}\right\}
$$

□

E Proof of Result 5

Proof. Imposing the steady state condition, $V_{y_t} = V_{y_{t-1}} = V_y$ on the law of motion of the variance of log income given in equation 18, gives:

$$V_y = \alpha^2 \left[\beta_a(1-\beta_m) + \beta_m\right] V_y + \beta_m V_\omega$$

We will call the right-hand side of this equation $\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)$. It follows that:

$$\begin{aligned}
&\frac{d\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)}{dV_y} \\
&= \frac{\partial\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)}{\partial V_y} + \frac{\partial\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)}{\partial \beta_a}\frac{d\beta_a}{dV_y} + \frac{\partial\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)}{\partial \beta_m}\frac{d\beta_m}{dV_y} \\
&= \alpha^2 \left[\beta_a(1-\beta_m) + \beta_m\right] + \alpha^2(1-\beta_m)V_y\frac{d\beta_a}{dV_y} + \left[\alpha^2\beta_a V_a + V_\omega\right]\frac{d\beta_m}{dV_y}
\end{aligned}$$

Since,

$$\frac{d\beta_a}{dV_y} = \frac{1-\beta_a}{V_y + V_a} > 0$$

$$\frac{d\beta_m}{dV_y} = \alpha^2\frac{(1-\beta_m)(1-\beta_a)^2}{\alpha^2\beta_a V_a + V_\omega + V_m} > 0$$

it follows that,

$$\begin{aligned}
\frac{d\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)}{dV_y} &= \alpha^2 \left[\beta_a(1-\beta_m) + \beta_m\right]\left\{1 + (1-\beta_m)(1-\beta_a)\right\} \\
&= \alpha\rho\ \left[1 + (1-\beta_m)(1-\beta_a)\right] > 0
\end{aligned}$$

while the second derivative is:

$$\begin{aligned}
\frac{d^2\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)}{dV_y^2} &= \frac{\partial\left[\frac{d\Psi(V_y,\beta_a,\beta_m,V_\omega,\alpha)}{dV_y}\right]}{\partial\beta_a}\frac{d\beta_a}{dV_y} + \frac{\partial\left[\frac{d\Psi(V_y,\beta_a,\beta_m,V_\omega,\alpha)}{dV_y}\right]}{\partial\beta_m}\frac{d\beta_m}{dV_y} \\
&= 2\alpha^2(1-\beta_a)(1-\beta_m)\left[(1-\beta_m)\frac{d\beta_a}{dV_y} + (1-\beta_a)\frac{d\beta_m}{dV_y}\right] > 0
\end{aligned}$$

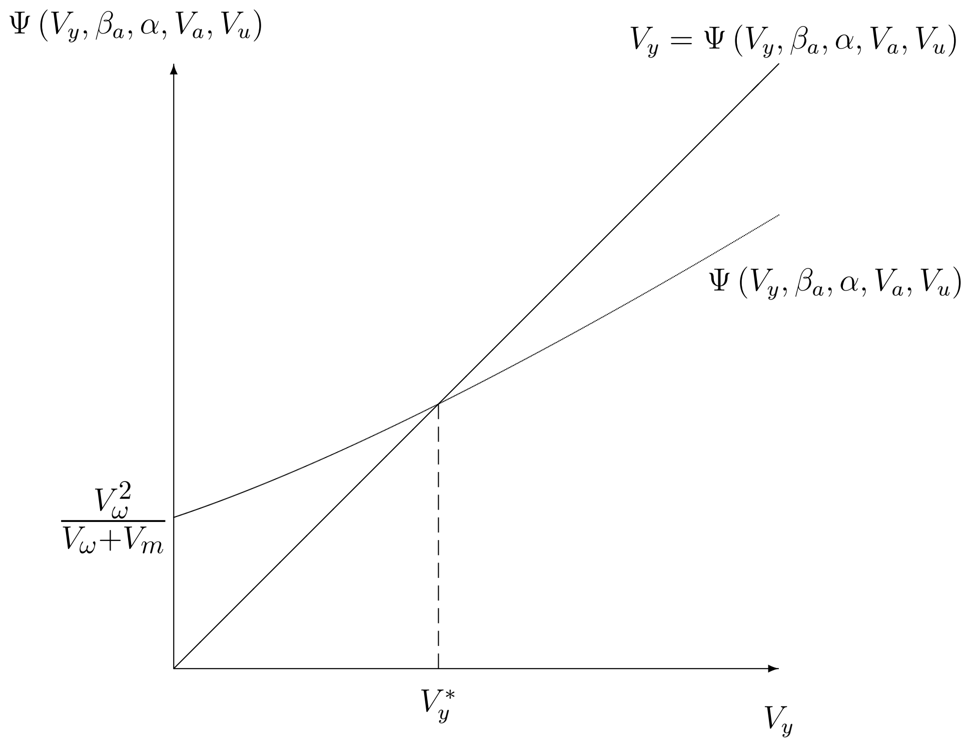

Figure 1: Law of motion of V_y

We also know that:

$$\Psi(0, \beta_a, \beta_m, V_\omega, \alpha) = \frac{V_\omega^2}{V_\omega + V_m}$$

$$\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha) \to \infty \quad \text{as } V_y \to \infty$$

$$\frac{d\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)}{dV_y} \to \alpha^2 \quad \text{as} \quad V_y \to \infty$$

Since $\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)$ starts above the 45 degree line, is upward sloping and convex in V_y, and has a maximum slope of $\alpha^2 < 1$, it must cut the 45 degree line once from above. This gives the unique, stable steady state value of V_y.

Note that the shape of the curve implies there will be a multiplier effect from changes in the parameters. Any parameter change which causes a shift in $\Psi(V_y, \beta_a, \beta_m, V_\omega, \alpha)$ will lead to a larger change in V_y. This multiplier effect will be discussed further when we consider the comparative statics of the model. $\square$

F Proof of Result 6

Proof. Let:

$$\Psi_a(V_y, \alpha, V_a, V_u) = \alpha^2 \frac{V_y^2}{V_y + V_a} + V_u$$

where steady state is defined by $V_y = \Psi_a(V_y, \alpha, V_a, V_u)$. All else equal, $\Psi_a(V_y, \alpha, V_a, V_u)$ is increasing in V_y. As a result, we generate multiplier effects from a change in any of the parameters (α, V_a or V_u) which shift $\Psi_a(V_y, \alpha, V_a, V_u)$. If they cause a fall in $\Psi_a(V_y, \alpha, V_a, V_u)$, V_y will have to fall to bring the steady state equation back into equality, which will further reduce $\Psi_a(V_y, \alpha, V_a, V_u)$ and so on. Similarly, any exogenous change in the parameters which increases $\Psi_a(V_y, \alpha, V_a, V_u)$ will increase V_y, further increasing $\Psi_a(V_y, \alpha, V_a, V_u)$.

For example, a change in V_u has the partial effect:

$$\frac{\partial \Psi_a(V_y, \alpha, V_a, V_u)}{\partial V_u} = 1$$

But the overall effect of the change on steady state V_y is:

$$\begin{aligned}
\frac{d\Psi_a(V_y, \alpha, V_a, V_u)}{dV_u} &= 1 + \frac{\partial \Psi_a(V_y, \alpha, V_a, V_u)}{\partial V_y} \frac{dV_y}{dV_u} \\
\frac{dV_y}{dV_u} &= 1 + \alpha^2 \beta_a \left(1 + (1 - \beta_a)\right) \frac{dV_y}{dV_u} \\
\frac{dV_y}{dV_u} &= \frac{1}{1 - \alpha^2 \left[1 - (1 - \beta_a)^2\right]} > \frac{\partial \Psi_a(V_y, \alpha, V_a, V_u)}{\partial V_u}
\end{aligned}$$

In a similar manner, if we call $B_a(\beta_a, \alpha, V_a, V_u) = \frac{V_u}{V_u + V_a(1-\alpha^2\beta_a)}$, we can show that:

$$\frac{d\beta_a}{dV_u} = \frac{V_a\left(1 - \alpha^2 \beta_a\right)}{\left[V_u + V_a\left(1 - \alpha^2 \beta_a\right)\right]^2 - \alpha^2 V_a V_u} > \frac{\partial B_a(\beta_a, \alpha, V_a, V_u)}{\partial V_u}$$

It is easy to show the partial and total effect from a change in α or V_a in exactly the same way, proving that that the total effect on β_a or V_y from a change in the parameters is greater in absolute terms than the partial effect.

□

G Proof of Result 7

Proof. Comparative statics from a change in the precision of the signal a:

$$\frac{d\Psi_a(V_y,\alpha,V_a,V_u)}{dV_a} = \frac{\partial\Psi_a(V_y,\alpha,V_a,V_u)}{\partial V_a} + \frac{\partial\Psi_a(V_y,\beta_a,\alpha,V_u)}{\partial V_y}\frac{dV_y}{dV_a}$$

$$\frac{dV_y}{dV_a} = -\alpha^2\beta_a^2 + \alpha^2\beta_a\left(1+(1-\beta_a)\right)\frac{dV_y}{dV_a}$$

$$\frac{dV_y}{dV_a} = \frac{-\alpha^2\beta_a^2}{1-\alpha^2\left[1-(1-\beta_a)^2\right]} < 0$$

$$\frac{d\beta_a}{dV_a} = \frac{\partial\beta_a}{\partial V_a} + \frac{\partial\beta_a}{\partial V_y}\frac{dV_y}{dV_a}$$

$$= \frac{-\beta_a(1-\beta_a)}{V_a}\left[\frac{1-\alpha^2\beta_a}{1-\alpha^2\left[1-(1-\beta_a)^2\right]}\right] < 0$$

Since $\rho = \alpha\beta_a$ it follows that:

$$\frac{d\rho}{dV_a} = \frac{-\alpha\beta_a(1-\beta_a)}{V_a}\left[\frac{1-\alpha^2\beta_a}{1-\alpha^2\left[1-(1-\beta_a)^2\right]}\right] < 0$$

□

H Proof of Result 8

Proof. Notice:

$$\frac{d\beta_a}{dV_a} = \frac{\partial\beta_a}{\partial V_y}\frac{dV_y}{dV_a} + \frac{\partial\beta_a}{\partial V_a} = \frac{1}{V_y+V_a}\left[(1-\beta_a)\frac{dV_y}{dV_a} - \beta_a\right]$$

and

$$\frac{d\beta_m}{dV_a} = \frac{\partial\beta_m}{\partial\beta_a}\frac{d\beta_a}{dV_a} + \frac{\partial\beta_m}{\partial V_a} = \frac{\alpha^2(1-\beta_m)}{\alpha^2\beta_aV_a+V_\omega+V_m}\left[(1-\beta_a)^2\frac{dV_y}{dV_a} + \beta_a^2\right]$$

So, after some algebra

$$\frac{dV_y}{dV_a} = \frac{-\alpha^2\beta_a^2(1-\beta_m)^2}{1-\alpha\rho\ \left[1+(1-\beta_m)(1-\beta_a)\right]}$$

Note that:

$$1-\alpha\rho\left[1+(1-\beta_m)(1-\beta_a)\right]=1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]>0$$

Therefore, the effect of a change in V_a on V_y is:

$$\frac{dV_y}{dV_a}=\frac{-\alpha^2\beta_a^2(1-\beta_m)^2}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]}<0$$

The effects of a change in V_a on β_a and β_m are:

$$\begin{aligned}\frac{d\beta_a}{dV_a}&=-\frac{\beta_a(1-\beta_a)}{V_a}\left[1+\frac{\alpha^2\beta_a(1-\beta_a)(1-\beta_m)^2}{(1-\alpha^2)+\alpha^2(1-\beta_a)^2(1-\beta_m)^2}\right]<0\\&=-\frac{\beta_a(1-\beta_a)}{V_a}\left[\frac{1-\alpha^2\left[1-(1-\beta_a)(1-\beta_m)^2\right]}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]}\right]\end{aligned}$$

$$\frac{d\beta_m}{dV_a}=\frac{\alpha^2\beta_a^2(1-\beta_m)^2}{V_m}\left[\frac{1-\alpha^2}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]}\right]>0$$

The extent to which firms value and use the signal on background is not measured by β_a but by $\hat{\beta}_a=\alpha\beta_a(1-\beta_m)$. A change in the precision of the advantage signal has the following effect on $\hat{\beta}_a$:

$$\begin{aligned}\frac{d\hat{\beta}_a}{dV_a}&=\alpha(1-\beta_m)\frac{d\beta_a}{dV_a}-\alpha\beta_a\frac{d\beta_m}{dV_a}<0\\&=-\frac{\alpha\beta_a(1-\beta_m)}{V_aV_m}\left[V_m+\frac{\beta_a(1-\alpha^2)\left[\alpha^2\beta_a(1-\beta_m)V_a-V_m\right]}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]}\right]\end{aligned}$$

The effect of a change in V_a on ρ is:

$$\begin{aligned}\frac{d\rho}{dV_a} &= \alpha\left(1-\beta_m\right)\frac{d\beta_a}{dV_a} + \alpha\left(1-\beta_a\right)\frac{d\beta_m}{dV_a}\\ &= -\frac{\alpha\beta_a^2\left(1-\beta_m\right)}{1-\alpha^2\left[1-\left(1-\beta_m\right)^2\left(1-\beta_a\right)^2\right]}\left\{\frac{\left(1-\alpha^2\right)\left[V_\omega+V_m\right]}{V_y\left[\alpha^2\beta_a V_a+V_\omega+V_m\right]}+\frac{\alpha^2\left(1-\beta_m\right)^2\left(1-\beta_a\right)}{V_y}\right\}\\ &= -\frac{\alpha\beta_a\left(1-\beta_a\right)\left(1-\beta_m\right)^2}{V_aV_m\left[1-\alpha^2\left[1-\left(1-\beta_m\right)^2\left(1-\beta_a\right)^2\right]\right]}\\ &\qquad .\left\{\left(1-\alpha^2\right)\left(V_\omega+V_m\right)+\alpha^2(1-\beta_a)\left(1-\beta_m\right)V_m\right\}<0\end{aligned}$$

□

I Proof of Result 9

Proof. Let $\Psi_m(V_y,\alpha,V_m,V_\omega) = (\alpha^2V_y+V_\omega)\,\beta_m$. Then $V_y = \Psi_m(V_h,\alpha,V_m,V_\omega)$ defines our steady state. Note that this is just equation 18 written in steady state for $V_a\to\infty$.

The comparative statics from a change in the precision of the signal m are then given by:

$$\begin{aligned}\frac{d\Psi_m(V_y,\alpha,V_m,V_\omega)}{dV_m} &= \frac{\partial\Psi_m(V_h,\alpha,V_m,V_\omega)}{\partial V_y}\frac{dV_y}{dV_m}+\frac{\partial\Psi_m(V_h,\alpha,V_m,V_\omega)}{\partial\beta_m}\frac{d\beta_m}{dV_m}\\ \frac{dV_y}{dV_m} &= \alpha^2\beta_m\frac{dV_y}{dV_m}+\left(\alpha^2V_y+V_\omega\right)\frac{d\beta_m}{dV_m}\end{aligned}$$

and

$$\begin{aligned}\frac{d\beta_m}{dV_m} &= \frac{\partial\beta_m}{\partial V_m}+\frac{\partial\beta_m}{\partial V_y}\frac{dV_y}{dV_m}\\ &= -\frac{\beta_m}{\alpha^2V_y+V_\omega+V_m}+\left\{\frac{\alpha^2\left(1-\beta_m\right)}{\alpha^2V_y+V_\omega+V_m}\right\}\frac{dV_y}{dV_m}\end{aligned}$$

Solving these two equations simultaneously gives:

$$\frac{dV_y}{dV_m} = -\frac{\beta_m^2}{1-\alpha^2\left[1-\left(1-\beta_m\right)^2\right]}<0$$

$$\frac{d\beta_m}{dV_m} = -\frac{\beta_m (1-\beta_m)}{V_m}\left[\frac{1-\alpha^2\beta_m}{1-\alpha^2\left[1-(1-\beta_m)^2\right]}\right] < 0$$

Since $\rho = \alpha\beta_m$ and $V_h = \alpha^2 V_y + V_\omega$ it follows that:

$$\frac{d\rho}{dV_m} = -\frac{\alpha\beta_m (1-\beta_m)}{V_m}\left[\frac{1-\alpha^2\beta_m}{1-\alpha^2\left[1-(1-\beta_m)^2\right]}\right] < 0$$

$$\frac{dV_h}{dV_m} = -\frac{\alpha^2\beta_m^2}{1-\alpha^2\left[1-(1-\beta_m)^2\right]} < 0$$

□

J Proof of Result 10

Proof. Notice that:

$$\frac{d\beta_a}{dV_m} = \frac{\partial\beta_a}{\partial V_y}\frac{dV_y}{dV_m} = \frac{(1-\beta_a)}{V_y+V_a}\frac{dV_y}{dV_m}$$

and

$$\frac{d\beta_m}{dV_m} = \frac{\partial\beta_m}{\partial\beta_a}\frac{d\beta_a}{dV_m} + \frac{\partial\beta_m}{\partial V_m}$$
$$= \frac{\alpha^2(1-\beta_a)^2(1-\beta_m)}{\alpha^2\beta_a V_a + V_\omega + V_m}\frac{dV_y}{dV_m} - \frac{\beta_m}{\alpha^2\beta_a V_a + V_\omega + V_m}$$

Then,

$$\frac{dV_y}{dV_m} = \frac{-\beta_m^2}{1-\alpha\rho\ \left[1+(1-\beta_m)(1-\beta_a)\right]}$$

As above:

$$1-\alpha\rho\ \left[1+(1-\beta_m)(1-\beta_a)\right] = 1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]$$

Therefore, the effect of a change in V_m on V_y is:

$$\frac{dV_y}{dV_m} = -\frac{\beta_m^2}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]} < 0$$

The effect of a change in V_m on β_a and β_m is:

$$\frac{d\beta_a}{dV_m} = -\frac{(1-\beta_a)^2}{V_a}\left[\frac{\beta_m^2}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]}\right] < 0$$

$$\frac{d\beta_m}{dV_m} = -\frac{\beta_m(1-\beta_m)}{V_m}\left[1+\frac{\alpha^2\beta_m(1-\beta_m)(1-\beta_a)^2}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]}\right] < 0$$

The effect of a change in V_m on ρ is:

$$\begin{aligned}\frac{d\rho}{dV_m} &= \alpha(1-\beta_m)\frac{d\beta_a}{dV_m} + \alpha(1-\beta_a)\frac{d\beta_m}{dV_m} \\ &= -\frac{\alpha\beta_m(1-\beta_a)(1-\beta_m)}{V_aV_m}\left\{V_a + \frac{(1-\beta_a)\beta_m V_m}{\left[1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]\right]}\right\} < 0\end{aligned}$$

Now turning to the effect on $\hat{\beta}_a$. A change in the precision of the ability signal has the following effect on $\hat{\beta}_a$:

$$\begin{aligned}\frac{d\hat{\beta}_a}{dV_m} &= \alpha(1-\beta_m)\frac{d\beta_a}{dV_m} - \alpha\beta_a\frac{d\beta_m}{dV_m} \\ &= \frac{\alpha\beta_m(1-\beta_m)}{V_aV_m\left[1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]\right]}\left\{\beta_aV_a\left(1-\alpha^2\right) - V_\omega(1-\beta_m)(1-\beta_a)^2\right\}\end{aligned}$$

This is $\lesseqgtr$ zero if:

$$\beta_aV_a\left(1-\alpha^2\right) \lesseqgtr V_\omega(1-\beta_m)(1-\beta_a)^2$$

There is a turning point at $\hat{V}_m$ where $\hat{V}_m$ gives values of β_a and β_m which solve the following equation:

$$\beta_aV_a\left(1-\alpha^2\right) = V_\omega(1-\beta_m)(1-\beta_a)^2$$

The left-hand side of the above equation is increasing in V_m while the right-hand side is decreasing. Therefore $\hat{V}_m$ is a unique turning point and a maximum:

$$\frac{d\left[\beta_a V_a\left(1-\alpha^2\right)\right]}{dV_m}=V_a\left(1-\alpha^2\right)\frac{d\beta_a}{dV_m}<0$$
$$\frac{d\left[V_\omega\left(1-\beta_m\right)\left(1-\beta_a\right)^2\right]}{dV_m}=-V_\omega(1-\beta_a)\left\{2\left(1-\beta_m\right)\frac{d\beta_a}{dV_m}+(1-\beta_a)\frac{d\beta_m}{dV_m}\right\}>0$$

For values of $V_m<\hat{V}_m$, $\beta_a V_a\left(1-\alpha^2\right)>V_\omega\left(1-\beta_m\right)\left(1-\beta_a\right)^2$ and $\frac{d\hat{\beta}_a}{dV_m}>0$.For values of $V_m>\hat{V}_m$, $\beta_a V_a\left(1-\alpha^2\right)<V_\omega\left(1-\beta_m\right)\left(1-\beta_a\right)^2$ and $\frac{d\hat{\beta}_a}{dV_m}<0$. □

K Proof of Result 11

Proof. From equation 16 we can see that:

$$y_{t+1}^i=\left(1-\beta_m\right)\left[\ln Z+\alpha\ln\lambda-\frac{V_\omega}{2}+\alpha\left(1-\beta_a\right)\mu_y\right]-\frac{1}{2}\left[\ln\left(1+\theta\beta_m V_m\right)-\frac{\beta_m V_m}{1+\theta\beta_m V_m}\right]$$
$$+\alpha\beta_a\left(1-\beta_m\right)a_{t+1}^i+\beta_m m_{t+1}^i$$

Substituting for a_t^i and m_t^i:

$$y_{t+1}^i=\left(1-\beta_m\right)\left[\ln Z+\alpha\ln\lambda-\frac{V_\omega}{2}+\alpha\left(1-\beta_a\right)\mu_y\right]-\frac{1}{2}\left[\ln\left(1+\theta\beta_m V_m\right)-\frac{\beta_m V_m}{1+\theta\beta_m V_m}\right]$$
$$+\alpha\beta_a\left(1-\beta_m\right)\left(y_t^i+\varepsilon_{t+1}^{ia}\right)+\beta_m\left(h_{t+1}^i+\varepsilon_{t+1}^{im}\right)$$

Equation 6 then allows us to substitute for log human capital:

$$y_{t+1}^i=\left(1-\beta_m\right)\left[\ln Z+\alpha\ln\lambda-\frac{V_\omega}{2}+\alpha\left(1-\beta_a\right)\mu_y\right]-\frac{1}{2}\left[\ln\left(1+\theta\beta_m V_m\right)-\frac{\beta_m V_m}{1+\theta\beta_m V_m}\right]$$
$$+\alpha\beta_a\left(1-\beta_m\right)\left(y_t^i+\varepsilon_{t+1}^{ia}\right)+\beta_m\left(\ln Z+\alpha\ln X_{t-1}^i-\frac{V_\omega}{2}+\omega_{t+1}^i+\varepsilon_{t+1}^{im}\right)$$

It is then just a matter of rearranging to find γ_0, γ_1, γ_2 and ε

$$y_t^i = \left\{ \ln Z + \alpha\left(1-\beta_m\right)\left[\left(1-\beta_a\right)\mu_y + \ln\lambda\right] - \frac{1}{2}\left[\ln\left(1+\theta\beta_m V_m\right) - \frac{\beta_m V_m}{1+\theta\beta_m V_m} + V_\omega\right]\right\}$$
$$+\alpha\beta_a\left(1-\beta_m\right)y_t^i + \alpha\beta_m \ln X_t^i + \alpha\beta_a\left(1-\beta_m\right)\varepsilon_{t+1}^{ia} + \beta_m\left(\omega_{t+1}^i + \varepsilon_{t+1}^{im}\right)$$

λ is found by substitution of γ_1 and γ_2 into equation 7. Note that $\gamma_1 + \gamma_2 = \rho < 1+\delta$ as required to ensure B is positive.

□

L Proof of Result 12

Proof.

$$\lambda = \frac{\beta_m \alpha}{1+\delta-\left(1-\beta_m\right)\ \alpha\ \beta_a}$$

The partial derivatives of λ with respect to β_m and β_a are therefore:

$$\frac{\partial\lambda}{\partial\beta_m} = \frac{\alpha\left[(1+\delta)-\alpha\beta_a\right]}{\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2} > 0$$
$$\frac{\partial\lambda}{\partial\beta_a} = \frac{\alpha^2\beta_m\left(1-\beta_m\right)}{\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2} > 0$$

So:

$$\frac{d\lambda}{dV_m} = \frac{\alpha\left[(1+\delta)-\alpha\beta_a\right]}{\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2}\frac{d\beta_m}{dV_m} + \frac{\alpha^2\beta_m\left(1-\beta_m\right)}{\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2}\frac{d\beta_a}{dV_m} < 0$$

With respect to V_a:

$$\begin{aligned}\frac{d\lambda}{dV_a} &= \frac{\alpha\left[(1+\delta)-\alpha\beta_a\right]}{\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2}\frac{d\beta_m}{dV_a} - \frac{\alpha^2\beta_m\left(1-\beta_m\right)}{\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2}\frac{d\beta_a}{dV_a} \\ &= \frac{\alpha^2\beta_a\left(1-\beta_m\right)}{V_a V_m\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2}\left\{\frac{\alpha(1-\alpha^2)\beta_a V_a(1-\beta_m)\left[(1+\delta)-\alpha\beta_a\right]}{1-\alpha^2\left[1-\left(1-\beta_m\right)^2\left(1-\beta_a\right)^2\right]}\right\} \\ &- \frac{\alpha^2\beta_a\left(1-\beta_m\right)}{V_a V_m\left[1+\delta-\left(1-\beta_m\right)\alpha\beta_a\right]^2}\left\{\frac{(1-\beta_a)\beta_m V_m\left[1-\alpha^2\left[1-\left(1-\beta_a\right)\left(1-\beta_m\right)^2\right]\right]}{1-\alpha^2\left[1-\left(1-\beta_m\right)^2\left(1-\beta_a\right)^2\right]}\right\}\end{aligned}$$

This is positive if:

$$\alpha\left[\frac{1-\alpha^2}{1-\alpha^2\left[1-(1-\beta_m)^2(1-\beta_a)^2\right]}\right](1-\beta_m)>\frac{V_m}{V_\omega}\left[\frac{1-\alpha^2\left[\beta_a(1-\beta_m)+\beta_m\right]}{\left[(1+\delta)-\alpha\beta_a\right]}\right]$$

Since the left-hand side of this equation is less than 1, it follows that a sufficient (though not necessary) condition for $\frac{d\lambda}{dV_a}<0$ is:

$$\frac{V_m}{V_\omega}\left[\frac{1-\alpha^2\left[\beta_a(1-\beta_m)+\beta_m\right]}{\left[(1+\delta)-\alpha\beta_a\right]}\right]>1$$

□

M Proof of Result 13

Proof. Median income is e^{μ_y} so the log of median income is μ_y.

$$\mu_y=\frac{\ln Z+\alpha\ln\lambda-\frac{1}{2}\left[\ln(1+\theta\beta_m V_m)-\frac{\beta_m V_m}{1+\theta\beta_m V_m}+V_\omega\right]}{(1-\alpha)}$$

The derivative with respect to V_m is given by:

$$\begin{aligned}\frac{d\mu_y}{dV_m}&=\frac{\partial\mu_y}{\partial V_m}+\frac{\partial\mu_y}{\partial\lambda}\frac{d\lambda}{dV_m}+\frac{\partial\mu_y}{\partial\beta_m}\frac{d\beta_m}{dV_m}\\&=-\frac{\left[\theta(1+\theta\beta_m V_m)-1\right]}{(1+\theta\beta_m V_m)^2}\frac{\beta_m}{2(1-\alpha)}+\frac{\alpha}{(1-\alpha)\lambda}\frac{d\lambda}{dV_m}-\frac{\left[\theta(1+\theta\beta_m V_m)-1\right]}{(1+\theta\beta_m V_m)^2}\frac{V_m}{2(1-\alpha)}\frac{d\beta_m}{dV_m}\\&=-\left[\frac{\beta_m+V_m\frac{d\beta_m}{dV_m}}{2(1-\alpha)}\right]\left\{\frac{\theta(1+\theta\beta_m V_m)-1}{(1+\theta\beta_m V_m)^2}-\frac{2\alpha\frac{d\lambda}{dV_m}}{\lambda\left[\beta_m+V_m\frac{d\beta_m}{dV_m}\right]}\right\}\end{aligned}$$

Note that

$$\beta_m+V_m\frac{d\beta_m}{dV_m}=\frac{(1-\alpha^2)\beta_m^2}{(1-\alpha^2)+\alpha^2(1-\beta_m)^2(1-\beta_a)^2}>0$$

The derivative $\frac{d\mu_y}{dV_m}$ is negative if:

$$\frac{2\alpha \frac{d\lambda}{dV_m}}{\lambda \left[\beta_m + V_m \frac{d\beta_m}{dV_m}\right]} < \frac{\theta \left(1 + \theta\beta_m V_m\right) - 1}{\left(1 + \theta\beta_m V_m\right)^2} \tag{30}$$

The left-hand side is independent of θ. The right-hand side is increasing in θ:

$$\frac{d}{d\theta}\left[\frac{\theta \left(1 + \theta\beta_m V_m\right) - 1}{\left(1 + \theta\beta_m V_m\right)^2}\right] = \frac{1 + \beta_m V_m \left(\theta + 2\right)}{\left(1 + \theta\beta_m V_m\right)^3} > 0$$

So if there is a sufficient cost of misallocation, then an increase in the accuracy of the signal on ability leads to higher median income. As the left-hand side of equation 30 is negative, a sufficient condition for $\frac{d\mu_y}{dV_m}$ to be negative is:

$$\theta \left(1 + \theta\beta_m V_m\right) - 1 > 0$$

which holds if $\theta \geq 1$. The derivative of μ_y with respect to V_a is given by:

$$\begin{aligned}
\frac{d\mu_y}{dV_a} &= \frac{\partial\mu_y}{\partial\lambda}\frac{d\lambda}{dV_a} + \frac{\partial\mu_y}{\partial\beta_m}\frac{d\beta_m}{dV_a} \\
&= \frac{a}{\left(1-\alpha\right)\lambda}\frac{d\lambda}{dV_a} - \frac{\left[\theta\left(1 + \theta\beta_m V_m\right) - 1\right]}{\left(1 + \theta\beta_m V_m\right)^2}\frac{V_m}{2\left(1-\alpha\right)}\frac{d\beta_m}{dV_a} \\
&= \frac{V_m}{2\left(1-\alpha\right)}\frac{d\beta_m}{dV_a}\left[\frac{2a}{\lambda V_m}\frac{\frac{d\lambda}{dV_a}}{\frac{d\beta_m}{dV_a}} - \frac{\left[\theta\left(1 + \theta\beta_m V_m\right) - 1\right]}{\left(1 + \theta\beta_m V_m\right)^2}\right]
\end{aligned}$$

This is negative if:

$$\frac{2a}{\lambda V_m}\frac{\frac{d\lambda}{dV_a}}{\frac{d\beta_m}{dV_a}} < \frac{\left[\theta\left(1 + \theta\beta_m V_m\right) - 1\right]}{\left(1 + \theta\beta_m V_m\right)^2} \tag{31}$$

The left-hand side of this equation is also independent of θ. Therefore, a sufficiently large θ will again ensure that the inequality holds and that median income is increasing in the accuracy of the background signal. □

References

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, *87*(6), 1153–1189.

Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, *4*(3), S1–S39.

Coate, S., & Loury, G. (1993). Will affirmative action policies eliminate negative stereotypes? *American Economic Review*, *83*(5), 1220–1240.

Hassler, J., Rodríguez Mora, J. V., & Zeira, J. (2007). Inequality and mobility. *Journal of Economic Growth*, *12*, 235–259.

Krueger, A. B. (2012). The rise and consequences of inequality in the united states. Speech to the Center for American Progress.

Solon, G. (2004). A model of intergenerational mobility variation over time and place. In M. Corak (Ed.) *Generational Income Mobility in North America and Europe*, chap. 2, (pp. 38–47). Cambridge University Press.

Young, M. (1958). *The Rise of Meritocracy, 1870 - 2033*. London: Thames and Hudson.