



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

The Value of a Privileged Background

Michael J. Watts

Doctor of Philosophy
The University of Edinburgh
2013

Declaration

I declare that this thesis was composed by myself and that the work contained is my own. No work by any other author has been used without due acknowledgment. This work has not been submitted for any other degree or professional qualification

Michael J. Watts

Abstract

This thesis considers how informational imperfections may give rise to advantages for those born to relatively rich parents. The first chapter focuses on the separation of some societies into different classes. Within the model, classes provide greater advantages to those from privileged backgrounds and, even in the absence of legal barriers preventing the lower classes from accessing skilled jobs, the skilled amongst them are still de facto denied access to high paying jobs through statistical discrimination. This chapter shows that there can be a net benefit from class discrimination, versus a classless state, when it creates information relating to the abilities of the upper class.

This theme is expanded on in chapter two where a signalling model more explicitly describes the statistical discrimination suffered by some members of society. The advantage conferred on those from privileged backgrounds generates income dispersion, which in turn reinforces the advantages of the rich. If this feedback is strong enough, the model may exhibit multiplicity of steady states. This multiplicity of steady states is backward looking: the income dispersion today depends on the extent to which firms use the information available to them, which in turn depends on the income dispersion in the previous generation. The model of chapter two also demonstrates why societies with more “meritocratic” institutions may exhibit less intergenerational income mobility: the income dispersion that meritocracy creates increases the value of a privileged upbringing.

The final chapter adds parental investment to the model. In doing so it brings the model more squarely in line with the statistical discrimination literature, although the

model does not exhibit a multiplicity of equilibria. There is a unique optimal investment rule for parents. Exogenous shocks to meritocracy are again examined. Meritocracy increases income variance and hence, from behind the veil of ignorance, creates greater uncertainty over the income an individual will receive. The model describes how a risk averse person might prefer to be born into an economy where they expect to be poorer but avoid this increased uncertainty, and so despite raising incomes, meritocracy may make agents, on average, more unhappy.

Acknowledgments

I would like to thank my supervisors, Sevi Mora and Maia Güell, for their advice and guidance over the course of my PhD. I also owe them many, many cups of coffee. The ESRC and School of Economics at Edinburgh University have supported me in my research, teaching and travel over the last four years. I have learnt a lot from all three. I am grateful to my examiners, Joseph Zeira, Liam Delaney and Jan Grobovšek, for taking the time to read my thesis and conduct my viva. Finally, thank you to my parents, family, and especially Mairi, for their encouragement and support. If nothing else, I have learnt that if you are going to finish your PhD thesis six weeks after you get married, you need a very understanding wife.

Contents

Declaration	i
Abstract	ii
Acknowledgments	iv
1 Class Divided Societies	1
1.1 Literature Review	2
1.2 A model of class divisions	4
1.3 The steady state without classes	7
1.3.1 Steady states of the economy when $\gamma b < 1$	8
1.3.2 Steady states of the economy when $\gamma b > 1$	10
1.4 The steady state with classes	11
1.4.1 Steady states when $\gamma b < 1$	13
1.4.2 Steady states when $\gamma b > 1$	19
1.5 A numerical example	20

1.6	Extending the model	24
1.7	Conclusion	25
2	The Inheritance of Advantage	27
2.1	Labour market sorting where firms receive a signal on parental income .	30
2.1.1	A firm's beliefs about talent	30
2.1.2	Steady State Beliefs	32
2.1.3	The feedback mechanism	34
2.2	Labour market sorting where firms receive a signal on talent	35
2.2.1	Steady state solutions where $\sigma_{\epsilon\tau}^2 \geq \sigma_{\epsilon s_2}^2$	37
2.2.2	Steady state solutions where $\sigma_{\epsilon\tau}^2 < \sigma_{\epsilon s_2}^2$	38
2.2.3	The curse of meritocracy I	39
2.3	Labour market sorting where firms receive two signals	40
2.3.1	The firm's choice of β_1 when the value of β_2 is fixed	42
2.3.2	The firm's choice of β_2 when the value of β_1 is fixed	43
2.3.3	A numerical example	43
2.3.4	The curse of meritocracy II	45
2.4	A discrete model with finite income variance	49
2.4.1	Steady state conditions	52
2.4.2	The steady state income distribution	55
2.4.3	An Example	56

2.5	An extension	60
2.6	Conclusions	61
3	Inheritance and Investment	63
3.1	An optimal investment rule for children's education	65
3.1.1	The pay decisions of firms	66
3.1.2	The investment decisions of parents	69
3.1.3	Solving the equilibrium payment and investment rules	71
3.2	The investment rule	75
3.3	Steady states of the economy	77
3.3.1	Case 1: The steady state of a weak inheritance economy	78
3.3.2	Case 2: The low steady state in a strong inheritance economy	81
3.3.3	Case 3: The high steady state in a strong inheritance economy	82
3.4	The weak inheritance economy	83
3.4.1	The effects of an improvement in meritocracy in a weak inheritance economy	84
3.4.2	The effects of an increase in inherited advantages in a weak inheritance economy	88
3.4.3	A discussion of the weak inheritance economy	91
3.5	The strong inheritance economy	92
3.5.1	The effects of an improvement in meritocracy in a strong inheritance economy	93

3.5.2	The effects of an increase in inherited advantages in a strong inheritance economy	95
3.6	Conclusions	97
A	The parameters of the value function	98
	Summary	102
	Bibliography	106

Chapter 1

Class Divided Societies

Class divisions have been a long-standing talking point in the UK and our continuing research into social mobility, or the lack thereof, tells of the uneasiness with which we greet this situation. A recent report on social mobility chaired by Alan Milburn MP found that the UK was far from a classless society and that if anything family background has become more important for access to the professions. In the US, a trend towards greater income segregation, in addition to persistent racial inequality, led one liberal American publication to write, “a growing percentage of white workers are now experiencing the same kinds of economic challenges that most African Americans have been facing since the 1970s” (Rodgers (2008)). This suggests that the traditional racial divisions, rather than being eradicated, may be being transformed into a British style class system.

This first chapter of my thesis examines under which circumstances such systems can be sustained. Their emergence is modelled as a consequence of statistical discrimination. In the absence of classes, firms lack the information they require to determine the talent of an individual. They have little to base their pay decisions on and do not know which workers to assign to a relatively complex task. The class signal provides them, perhaps imperfectly, with that information. The cost is that the talented group of people created is smaller than the proportion of workers they would like to assign to

the complex task. Within the model, where the informational gains exist, they dominate the allocation costs. However, there are many steady state divisions of society which go beyond the point at which these gains are exhausted.

As we will see below, there are class divisions of society which can be sustained in steady state for which both classes of society are better off than if no such class division existed. This should not mask the fact, however, that the lower classes are being discriminated against. A talented lower class worker will still be paid less than an untalented upper class one, despite talent being the only difference in productivity between workers. The circumstances of birth deny access to a high paying job for a worker who is perfectly capable of carrying it out, yet at the same time he is paid more than if a class system did not exist. The reason is the firms' lack of information which will be a common feature of much of this thesis.

1.1 Literature Review

The origins of this chapter, and this thesis in general, lie primarily in the discrimination literature. As such, it will be useful to discuss it as background before moving on to the model.

Papers on discrimination generally have their origins in one of two sources: preference-based discrimination (Becker (1957)); and statistical discrimination (Phelps (1972); Arrow (1973)). Becker-style discrimination involves utility maximising agents who suffer a utility cost from employing, working with, or being served by individuals from particular groups. It is, in essence, a model of racism and sexism. The origins of this chapter's model are instead in statistical discrimination. In this line of literature individuals give a signal of their ability but that signal is imperfect. Employers combine a worker's signal with the other information that they have about them, say their race. In some equilibria, beliefs about both groups are the same and so race contains no useful information on which to discriminate. However, there may also be equilibria where the two groups are treated differently. One group is believed to be less talented and

so their investments in education open fewer doors for them. As a result they invest less in education and so are, in fact, less talented. It is a story which has been used in other areas of economics (for its application to fiscal policy and inequality see Piketty (1995, 1998)).

The Phelps article cited above was based on the difference in variance of the signals of the two groups, an assumption which was carried through into Lundberg & Startz (1983) and Lazear (1995). This chapter is more in keeping with the Arrow tradition which can be followed through to the seminal contribution of Coate and Louri on statistical discrimination and affirmative action (Coate & Louri (1993)). That article provided a model which outlined the statistical discrimination mechanism, but there were two important omissions. One was that the productive sector was not specified. The other was that the wage for the more complex task was constant and exogenously given. This meant that the two groups (in their case black and white) were entirely separate, the wage of one unaffected by what was happening to the other.

Moro and Norman, in a number of papers, (Moro (2003); Moro & Norman (2004); Norman (2003)) addressed these issues and extended the Coate and Louri model to include a production function with complementarity between the two tasks, simple and complex, and endogenous determination of both wages. The complex task could only be done by talented individuals and, as such, the wage for performing the complex task was a function of the individual's probability of being talented. This chapter builds on top of those features a dynamic framework where workers are discriminated based on an endogenous class signal. The focus on an endogenous, intrinsically useful signal is distinct from the discrimination of (for the most part ex ante identical) individuals on the basis of race, ethnicity, or gender discussed in much of the literature. The model of section 1.2 will lay the foundations for considering the class divisions which may exist in steady state.

1.2 A model of class divisions

Assume that there is a population of mass one and that each individual has one child. Every individual may be talented or not. If they are talented they have a value of τ equal to one. If they are not, they have a value of τ equal to zero. There are up to two “classes” of individual depending on parental income: if an individual’s parent is rich he is from the “upper” class; if his parent is poor he is from the “lower” class. In what follows, we assume that class is observable but talent is not. In the next chapter we will consider a case where both are imperfectly observable.

The production function is Cobb-Douglas. There are two tasks in society, and our inputs to production are the effective labour input into each task. One task is “complex”. Only talented individuals are able to perform it and any individual trying to perform the complex task who is not talented will produce nothing. The other task is “simple”. Everyone is capable of performing the simple task.

$$y = \tilde{C}^\gamma S^{1-\gamma} \tag{1.1}$$

\tilde{C} is the effective labour input into the complex task. These are the individuals who are doing the complex task *and* are capable of doing it. Firms will not be able to observe whether an individual has the skills to perform the task and so will have to form an expectation based on the observable about an individual: his class. S is the labour input into the simple task.

Every individual in society will be offered a wage for performing each task equal to his or her expected marginal product. Workers have no preference over tasks but will always choose the task for which they are offered the higher wage. If a worker is indifferent, a firm can allocate them to whichever task it prefers.

The effective labour input into the complex task is equal to the number of people

doing the complex task multiplied by the proportion of them which actually have the skills to carry it out: $\tilde{C} = Prob(\tau = 1)K$, where K is the proportion of the population choosing to undertake the complex task.¹ As a result:

$$w^c = \frac{\partial y}{\partial K} = \frac{\partial y}{\partial \tilde{C}} \frac{\partial \tilde{C}}{\partial K} = \frac{\partial y}{\partial \tilde{C}} Prob(\tau = 1)$$

The wage which a firm will offer to an individual choosing the complex task, w^c , is equal to the wage that they would offer a talented person performing the complex task multiplied by the probability that that person is talented. This is because talent is a necessary prerequisite for being able to carry out the complex task.

We will assume that the probability that an individual is talented is an increasing function of their parent's income, w_{-1} . Individuals from a richer background are likely to have had greater investments made in their schooling and childhood development, putting them in a better position to take on skilled tasks when they enter the labour force. This underlying investment process will be considered in detail in chapter 3 but, for now, we will assume it to be true. This probability reaches a maximum of one for those with sufficiently rich parents.

$$Prob(\tau = 1) = \begin{cases} P(w_{-1}) & \text{if } w_{-1} < P^{-1}(1) \\ 1 & \text{if } w_{-1} \geq P^{-1}(1) \end{cases}$$

where $P(\cdot)$ is an increasing function of w_{-1} . We will assume below that $P(w_{-1}) = aw_{-1}^b$.

¹We will see below that all individuals performing the complex task have the same probability of being able to do so in steady state and so the proportion able to perform the task is equal to the probability that any one of them can do so. I use $Prob(\tau = 1)$ to represent the probability of being talented *of someone performing the complex task*. These are the only people whose talent plays a role in output.

The wage that firms will offer to anyone carrying out the simple task is equal to,

$$w^s = \frac{\partial y}{\partial S} \quad (1.2)$$

This is the same for all workers since there is no minimum skill level required to be able to perform the simple task.

Since class is observable, and talent is not, the wage that a worker in the complex sector is offered will be dependent on her class. All individuals will receive one of two wage offers for performing the complex task. If their parent is rich, and hence they are upper class, they will receive a wage of w^c .

$$w^c = \frac{\partial y}{\partial \tilde{C}} \text{Prob}(\tau = 1 | w_{-1} = w_{-1}^c) \quad (1.3)$$

If their parent is poor, and hence they are from the lower class, they will receive a wage $w^{c|s}$.

$$w^{c|s} = \frac{\partial y}{\partial \tilde{C}} \text{Prob}(\tau = 1 | w_{-1} = w_{-1}^s) \quad (1.4)$$

Note that in equations 1.3 and 1.4 we are considering only two options for parent's income in a class divided society, w_{-1}^c and w_{-1}^s . This is because in steady state noone will receive the income $w^{c|s}$. Imagine they did. This would imply that $w^{c|s} > w^s$: the lower class children would choose to undertake the complex task. But, if the lower class would choose the complex task so would the upper class, since firms will believe them to be at least as likely to be able to perform it. Class division then, in a world where workers are free to choose their task, requires that the lower class is sufficiently disadvantaged as to remain lower class which precludes the acceptance of $w^{c|s}$ in steady state.

1.3 The steady state without classes

A steady state of this model is a situation in which the distribution of income is constant across time. The easiest steady states to calculate are those in which all agents are paid the same wage for either task since, instead of an unchanged distribution across time, this amounts to a unique constant wage (although we allow the steady state economy to exhibit a class divide, it does not have to). When this is the case all individuals have the same probability of being talented.

We will begin with a situation in which all individuals within a generation receive the same wage. This is given by w in the current generation and w_{-1} in the previous one.

$$w = \frac{\partial y}{\partial \tilde{C}} \text{Prob}(\tau = 1 | w_{-1} = w_{-1}) = \frac{\partial y}{\partial S}$$

Everyone within a given generation is indifferent between the two tasks since they earn a common wage for both. The effective labour input into each task will be $\tilde{C} = P(w_{-1})K$ and $S = 1 - K$, the latter of these coming from the fact that all workers who are not assigned to the complex task are able to perform the simple one.²

We will use $P(w_{-1}) = aw_{-1}^b$ to describe the relationship between parental income and talent where parental income is sufficiently low that the probability of being talented is less than one. This then leaves us with,

$$K = \gamma$$

$$w = \begin{cases} \gamma^\gamma (1 - \gamma)^{1-\gamma} a^\gamma w_{-1}^{\gamma b} & \text{if } w_{-1} < \left(\frac{1}{a}\right)^{\frac{1}{b}} \\ \gamma^\gamma (1 - \gamma)^{1-\gamma} & \text{if } w_{-1} \geq \left(\frac{1}{a}\right)^{\frac{1}{b}} \end{cases}$$

²Workers are all assigned to tasks by firms since they will be paid the same in either and so will be indifferent as to which task they perform.

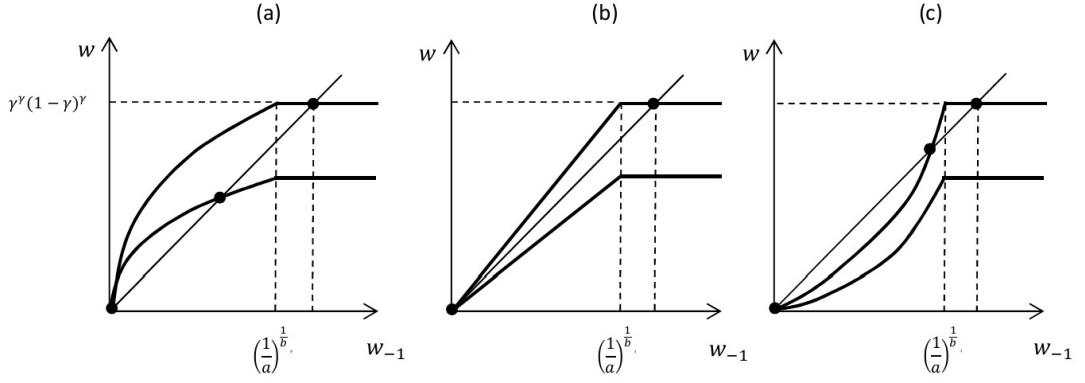


Figure 1.1: Possible steady state wages in the economy with no classes: (a) $\gamma b < 1$; (b) $\gamma b = 1$; and (c) $\gamma b > 1$.

The first of these conditions sets the wages to be equal across tasks within a generation. The second describes the relationship between wages across generations. It is illustrated in figure 1.1 for values of γb less than, greater than and equal to one.³

Since $\gamma b = 1$ is a special case, and an intermediate case of the others, it will not be considered in the discussion below. It should be clear from the diagram which steady states exist in this case. When the function is above or below the $w = w_{-1}$ line, the wage is rising or falling across generations respectively. The steady states are shown at the points where $w = w_{-1}$, the steady state condition. We will now turn to each in turn.

1.3.1 Steady states of the economy when $\gamma b < 1$

These values of the parameters indicate that either complex labour's share of output is low or talent is relatively inelastic in parental income. It means that the wage is a concave function of the parental wage. At relatively low incomes there are large gains in the income of subsequent generations from raising current incomes, large enough to be sustained. This concavity allows the steady state wage to reach a positive value at which it is sufficiently unresponsive to the parental wage to be stable. This section examines each of the steady state values of the wage drawn in panel (a) of figure 1.1.

³Whether γb is greater than or less than 1 describes the convexity or concavity of this function in w_{-1} . For γb equal to one it is linear.

The first possible steady state solution is,

$$w^* = 0 \tag{1.5}$$

This steady state is possible for all values of γ, a and b but, as we can see from the diagram, is not stable when $\gamma b < 1$. Any positive income shock will lead to income growth towards a higher steady state as it leads to sufficient development of talent in subsequent generations.

The other steady state values of the wage drawn in figure 1.1(a) are really two cases of the same one. The lower one only exists for values of γ, a and b which satisfy $a < \frac{1}{\gamma^\gamma b (1-\gamma)^{(1-\gamma)b}}$. If that condition is met, there exists a steady state wage given by,

$$w^* = \left[a^\gamma \gamma^\gamma (1-\gamma)^{1-\gamma} \right]^{\frac{1}{1-\gamma b}} \tag{1.6}$$

In this steady state workers have a positive probability of being talented but the condition on a ensures that they are not talented with certainty. They earn a positive wage in either task and it is stable.

The highest steady state wage drawn is one for which the common wage is sufficiently high that everyone is able to do the complex task. It is the case when $a \geq \frac{1}{\gamma^\gamma b (1-\gamma)^{(1-\gamma)b}}$.

$$w^* = \gamma^\gamma (1-\gamma)^{(1-\gamma)} \tag{1.7}$$

The two conditions on a are mutually exclusive so, as can be seen in figure 1.1,

equations 1.6 and 1.7 essentially describe the same steady state for different values of a . There is one stable steady state level of income. For a given value of γb it may, for a sufficiently high a , lead to all workers being talented. Otherwise only a fraction of them will be talented. The parameter a measures the probability that an individual is talented for a given value of his parent's income. It can be thought of as the technology of talent formation and analogous to the efficiency of the public education system. Here, the steady state wage is increasing in the efficiency of the education system up to the point where all workers are talented. This is very intuitive.

1.3.2 Steady states of the economy when $\gamma b > 1$

In this economy the complex task represents a larger share of production or a worker's talent is relatively elastic with respect to her parent's income. The result is that the wage is a convex function of the parental wage as illustrated in panel (c) of figure 1.1. Changes in parents' wages lead to more than proportional changes in children's wages. Thus if income is high and increasing, it will tend to keep increasing, but if it is low, it will tend to be falling. This leads to an "all-or-nothing" society which tends to stabilise with either everyone, or noone, being talented.

The first possible steady state solution is again $w^* = 0$ only in this case it is stable. An income shock will not affect the steady state wage in the long run. Following the shock the economy would not be able to support a skilled upper class of workers capable of carrying out the complex task because, for low levels of income, the talent gains from positive income shocks are too small. The wage would fall back to zero and so there exists a poverty trap.

The second and third possible steady states have the same level of income as in equations 1.6 and 1.7. The difference when $\gamma b > 1$ is that they either both exist or both do not. If they do, the one given in equation 1.6 is unstable while the one given in 1.7, where all workers are talented, is stable.

The condition for existence of these "high" steady states is $a \geq \frac{1}{\gamma \gamma^b (1-\gamma)^{(1-\gamma)b}}$. If the

economy's output is heavily dependent on the complex sector, or the development of a skilled workforce requires a high income in the previous generation, then an efficient public education system is required to prevent the economy collapsing into poverty. If the education system is of low enough quality, eventually people will not be guaranteed to be talented and there will be too few talented workers to support the complex sector. The economy will go into decline.

As an example, this would suggest that, at least in a classless society when private spending on education is complementary to public spending, economies with relatively large high tech sectors (high γ) or efficient privately funded university systems (high b) may still go into decline if they do not maintain a good enough public education system (low a). Although γ is not endogenous to the model, it also shows that rich countries in which the public education system is going into decline might be better to shift away from skilled labour in their production (reduce γ). This would allow them to maintain a stable steady state level of income below the maximum rather than experiencing a long decline into poverty.

1.4 The steady state with classes

In this section we will consider how the situation changes when firms observe an individual's class when making wage decisions. They now have something to distinguish between workers and so will not necessarily offer the same wages to all. The equilibrium wages paid w^s , w^c , and w^{cls} are given in equations 1.2 to 1.4. These wages are based on the task that the worker is performing and which class they hail from. Depending on how the wages relate to one another, a proportion K of workers will be choosing the complex task and $(1 - K)$ the simple task.

It must be that $w^c > w^{cls}$ in any class-based society. This simply states that those from a richer background are offered a higher wage to perform the complex task since they are more likely to be able to perform that task. As a result K can only take one of three values: $K = 0$; $K = 1$; and $K = K_{-1}$.

If $w^s > w^c > w^{c|s}$ then everyone would want to do the simple task ($K = 0$). Similarly, if $w^c > w^{c|s} > w^s$ then everyone would want to do the complex task ($K = 1$). In either case the economy, which is reliant on both tasks for producing output, would collapse to a classless society with no productive output. Wages would be equal at zero. These can therefore not be the structure of equilibrium wage offers in a class divided society.

If $w^c > w^s > w^{c|s}$ then $K = K_{-1}$ as the upper class individuals perform the complex task and the lower class individuals perform the simple task. This is the only K for which a class division exists in society. Since it is constant through all time we will call it K_0 .⁴ From equations 1.2 to 1.4 we can see the condition for this class division.

$$\frac{\partial y}{\partial \tilde{C}} \text{Prob}(\tau = 1 | w_{-1} = w_{-1}^c) > \frac{\partial y}{\partial S} > \frac{\partial y}{\partial \tilde{C}} \text{Prob}(\tau = 1 | w_{-1} = w_{-1}^s) \quad (1.8)$$

All of the wages in the economy will be determined by two things: the distribution of workers across the two tasks; and the productivity of workers in the complex task. The former we know is given by K_0 in all periods. The latter will be determined by the proportion of complex workers who are talented, itself a function of the steady state wage in the complex sector. We will now focus on determining that, again assuming that $P(w_{-1}) = aw_{-1}^b$.

From equation 1.3 we can see that the wage in the complex sector must satisfy,

$$w^c = \begin{cases} \gamma \left[\frac{1-K_0}{K_0} \right]^{1-\gamma} a^\gamma (w_{-1}^c)^{\gamma b} & \text{if } w_{-1}^c < \left(\frac{1}{a}\right)^{\frac{1}{b}} \\ \gamma \left[\frac{1-K_0}{K_0} \right]^{1-\gamma} & \text{if } w_{-1}^c \geq \left(\frac{1}{a}\right)^{\frac{1}{b}} \end{cases}$$

This describes the relationship of wages across time in the complex sector. Since

⁴Technically $w^c \geq w^s > w^{c|s}$ or $w^c > w^s \geq w^{c|s}$ could be enough since presumably the firm would choose to assign the upper class workers to the complex task in the first case, and the lower class workers to the simple task in the second, in order to prevent output collapsing to zero.

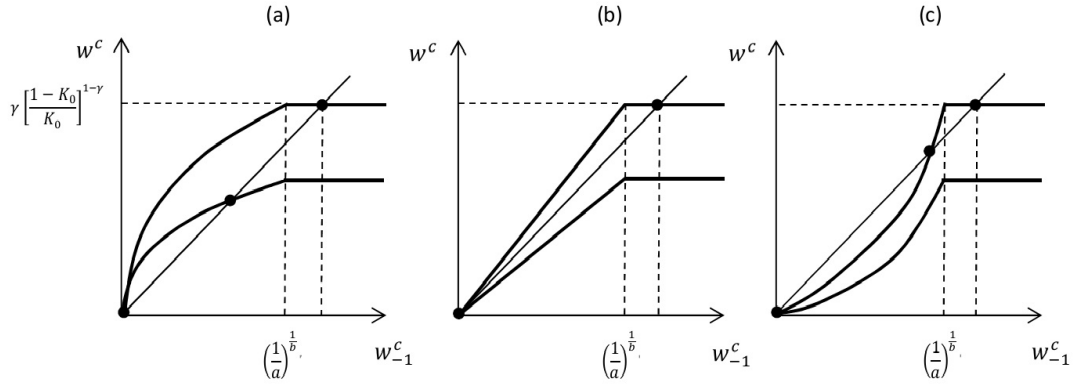


Figure 1.2: Possible steady state wages in the economy with class division: (a) $\gamma b < 1$; (b) $\gamma b = 1$; and (c) $\gamma b > 1$.

the children of complex workers also perform the complex task, it also describes the upper class wage process. It is illustrated in figure 1.2 for values of γb less than, greater than and equal to one. Since a steady state requires that the distribution of income is unchanged over time, and K_0 is constant, the equilibrium wages must be constant across periods. Our steady state condition is therefore $w^c = w_{-1}^c$.

1.4.1 Steady states when $\gamma b < 1$

The first possible steady state is again one where wages are equal to zero. However, clearly this would not satisfy our condition for a class divided equilibrium given in equation 1.8. It therefore cannot lead to a steady state distribution of income with distinct classes. As we can see from the diagram in panel (a) of figure 1.2, there are then two separate cases for which there is a stable steady state value of the upper class wage: in one, not all of the upper class workers are talented; in the other, they are. The two are mutually exclusive. If the upper class wage is stable, so will be the lower class one.

A steady state where not all complex workers are talented

The upper class wage in the situation where they are not all talented is given by,

$$w^{c*} = \left\{ \gamma a^\gamma \left[\frac{1 - K_0}{K_0} \right]^{1-\gamma} \right\}^{\frac{1}{1-\gamma b}} \quad (1.9)$$

The condition which ensures that they are not all talented is given by $w^{c*} < \left(\frac{1}{a}\right)^{\frac{1}{b}}$. Substituting from equation 1.9 this can be shown to give,

$$K_0 > \frac{(a\gamma^b)^{\frac{1}{b(1-\gamma)}}}{1 + (a\gamma^b)^{\frac{1}{b(1-\gamma)}}} \quad (1.10)$$

In order to ensure that not all of the complex workers are talented, there must be a large enough number of them. The marginal productivity of complex workers tends to infinity as their numbers fall to zero, and so eventually they would have to have sufficiently high income to all be talented.

In addition to having a large enough mass on complex workers to ensure that they are not all talented, we also need the distribution of workers to be such that equation 1.8 is satisfied. The upper classes have to choose the complex task and the lower classes the simple one. To ensure this, we need to know the steady state simple wage and the wage offered to lower class workers to perform the complex task. From equations 1.1, 1.2 and 1.9 we can find the steady state value of the simple wage.

$$w^{s*} = (1 - \gamma) \left\{ \gamma^\gamma a^\gamma \left[\frac{1 - K_0}{K_0} \right]^{\gamma(b-1)} \right\}^{\frac{1}{1-\gamma b}} \quad (1.11)$$

Interestingly, the simple wage can be increasing or decreasing in K_0 . As K_0 increases, the number of simple workers decreases, the number of complex workers increases, and the proportion of *talented* complex workers decreases. The first two are standard ways in which the simple wage should increase. The last effect is born out

of the fall in the complex wage. If it is large enough, it can push the effective labour input into the complex task down, despite there being more workers performing it. The scarcity of workers who are able to perform the complex task may in turn lead to a fall in the marginal productivity of simple workers.

In a similar manner to that of w^{s*} , we can find the steady state wage offered to a lower class worker for performing the complex task from equation 1.4.

$$w^{c|s*} = (1 - \gamma)^b \left\{ \gamma^{1-b(1-\gamma b)} a \gamma \left[\frac{1 - K_0}{K_0} \right]^{1-\gamma-b(1-\gamma b)} \right\} \quad (1.12)$$

Again, this may be increasing or decreasing in K_0 depending on whether the simple wage and the effective labour input into the complex task are increasing or decreasing.

Based on the values of the wages, we can simplify equation 1.8 to be, for this steady state,

$$\frac{\gamma(1 - K_0)}{(1 - \gamma)K_0} > 1 > \left[\frac{\gamma(1 - K_0)}{(1 - \gamma)K_0} \right]^{1-b}, \quad K_0 > 0 \quad (1.13)$$

The two conditions that make this inequality hold are $0 < K_0 < \gamma$ and $b > 1$.⁵ It is not surprising that a high value of b is important for the existence of a class divided society. The sensitivity of talent to parental income was always likely to play an important role since it magnifies the advantages of the upper class over the lower.

To summarise, if $1 < b < \frac{1}{\gamma}$ and,

$$\frac{(a\gamma^b)^{\frac{1}{b(1-\gamma)}}}{1 + (a\gamma^b)^{\frac{1}{b(1-\gamma)}}} < K_0 < \gamma \quad (1.14)$$

⁵Since $b > 1$ implies convexity in the relationship between the parental wage and the probability of being talented, it is worth noting that this convexity may not exist across the whole range of $w_{-1}^c \leq (\frac{1}{a})^{\frac{1}{b}}$. However, it is plausible that we would observe this sort of convexity at low values of w_{-1}^c . An example might be where initial parental investment develops basic skills required in order for further investments to raise the probability of being able to perform a complex task.

there exists a steady state distribution of workers in the economy such that a constant fraction K_0 are upper class and performing the complex task in any period. Only upper class workers perform the complex task and not all of them are qualified to do so. Only lower class workers perform the simple task. The wages paid to workers are given by equations 1.9 and 1.11. These wages are stable.

It is worth noting at this point that this steady state will be an important one. It was mentioned that the steady state simple wage could be higher if the lower class contained a larger fraction of the population. When $b > 1$, it will be. This is due to spillovers from the upper classes when they are richer and better able to carry out the complex task (it requires that they are not already perfectly able to carry it out). Thus both classes of worker would be better off in a class based society than a classless one. This will be the only steady state for which this can happen.

A steady state where all complex workers are talented

In this section we will examine the steady state which may arise when $\gamma b < 1$ in which all upper class workers are capable of carrying out the complex task. Their wage is then given by,

$$w^{c*} = \gamma \left[\frac{1 - K_0}{K_0} \right]^{1-\gamma} \quad (1.15)$$

For this to be a steady state wage with class division we again need to satisfy a condition ensuring all complex workers are talented, and a condition ensuring workers sort into jobs in the required manner. The former is given by $w^{c*} \geq \left(\frac{1}{a}\right)^{\frac{1}{b}}$. Substituting from equation 1.15 this can be shown to give:

$$K_0 \leq \frac{(a\gamma^b)^{\frac{1}{b(1-\gamma)}}}{1 + (a\gamma^b)^{\frac{1}{b(1-\gamma)}}} \quad (1.16)$$

In order for there to be a steady state where all complex workers are talented, there must be few enough of them to keep their wage sufficiently high. Comparing equations 1.10 and 1.16 we can see that they are mutually exclusive. Only one of these steady states values of the complex wage can exist at a time. There is no multiplicity in incomes, at least for stable values of income and situations in which there is a class division in society. This is also clear from figure 1.2(a).

The condition on the distribution of workers which we need to satisfy in order for this steady state to be viable is the familiar one from equation 1.8. In this case the steady state wages offered to the lower classes are given by $w^{s*} = (1 - \gamma) \left[\frac{K_0}{1 - K_0} \right]^\gamma$ and $w^{c,s*} = \gamma a (1 - \gamma)^b \left[\frac{1 - K_0}{K_0} \right]^{1 - \gamma(1 + b)}$. Here the wage for the simple task is decreasing as more people carry it out, as we would normally expect. The gains from a higher proportion of talented complex workers have been exhausted.

$$\frac{(a\gamma(1-\gamma)^{b-1})^{\frac{1}{1-\gamma b}}}{1 + (a\gamma(1-\gamma)^{b-1})^{\frac{1}{1-\gamma b}}} < K_0 < \min \left\{ \gamma, \frac{1}{1 + (a(1-\gamma)^b)^{\frac{1}{\gamma b}}} \right\}$$

If γb is less than one, and equation 1.16 is satisfied, it means that reductions in K_0 , channelling people into the simple job, bring down earnings in the simple job to a greater extent than it brings down the earnings the lower class would earn from performing the complex task. The scarcity of workers performing the complex task becomes more of a concern to firms than the relatively low probability that the lower classes can fill in. At low values of K_0 the lower classes would find it optimal to switch to the complex task. Equally, high values of K_0 cannot be sustained in this steady state, either because the upper classes would wish to switch to the simple task or because the

lower classes would earn a sufficiently high wage that they were all qualified to do the complex task. Neither is compatible with an upper and lower class.

$$\frac{\left(a\gamma(1-\gamma)^{b-1}\right)^{\frac{1}{1-\gamma b}}}{1+\left(a\gamma(1-\gamma)^{b-1}\right)^{\frac{1}{1-\gamma b}}} < K_0 < \min \left\{ \gamma, \frac{(a\gamma^b)^{\frac{1}{b(1-\gamma)}}}{1+(a\gamma^b)^{\frac{1}{b(1-\gamma)}}}, \frac{1}{1+\left(a(1-\gamma)^b\right)^{\frac{1}{\gamma b}}} \right\} \quad (1.17)$$

To summarise, if $\gamma b < 1$ and K_0 satisfies equation 1.17, there exists a steady state where there are two classes of worker. A proportion K_0 are upper class and performing the complex task at any time. All upper class workers are qualified to do the complex task but not all lower class workers are, and the wages paid to the upper and lower class workers are stable.

Before we move on to consider the case of $\gamma b > 1$, it may be useful to summarise what we have found for $\gamma b < 1$. If either 1.14 or 1.17 hold, there is a range of values for the size of the upper class and lower classes which can be maintained over time at a constant level. At this value of K_0 there are unique values of the wages paid to the upper and lower classes. There may be values of K_0 at the upper end of this range, given by equation 1.10, for which not all of the upper class workers are talented. In that case, a value of b greater than 1 would be required to accentuate the differences in ability between the two classes. It also opens up the possibility that, were there a choice of K_0 , creating a smaller, higher paid, exclusively talented upper class may improve efficiency, and the incomes of both classes, in the economy. Despite the smaller number of complex workers, effective input into that sector would be higher since the proportion of them which are capable of doing the task increases. The more productive complex sector has spillover effects into the simple sector, raising the wages of lower class workers. We will illustrate these potential gains from class divisions in section 1.5.

1.4.2 Steady states when $\gamma b > 1$

The situation when γb is greater than one is illustrated in figure 1.2(c). As in the previous section, the steady state with complex wage equal to zero is clearly at odds with the notion of an upper class and does not satisfy equation 1.8. The other two steady state levels of the complex wage, as in section 1.3.2, either both exist or both do not. The condition for their existence is given by equation 1.16. This would open up the possibility of multiplicity of income distributions, however, the lower valued one, where upper class workers are not all talented, is an unstable steady state. As a result, we will not discuss it much further, other than to say that the wages are as in section 1.4.1 where all workers were not talented.

The only stable steady state complex wage which leads to class division is given by equation 1.15. It is high enough that all upper class individuals are talented. The lower class wage is also as set out in section 1.4.1 and the condition for existence of a steady state income distribution is,

$$0 < K_0 < \min \left\{ \gamma, \frac{(a\gamma(1-\gamma)^{b-1})^{\frac{1}{1-\gamma b}}}{1 + (a\gamma(1-\gamma)^{b-1})^{\frac{1}{1-\gamma b}}}, \frac{(a\gamma^b)^{\frac{1}{b(1-\gamma)}}}{1 + (a\gamma^b)^{\frac{1}{b(1-\gamma)}}}, \frac{1}{1 + (a(1-\gamma)^b)^{\frac{1}{\gamma b}}} \right\} \quad (1.18)$$

In the case of γb less than one, as K_0 became very small the lower classes would find it optimal to switch to the complex task. Here they do not. They are severely disadvantaged by their lower class status. When K_0 is smaller, their low wage makes it very unlikely that they are capable of performing the complex task and so, despite the scarcity of complex workers, they are unwilling to switch. Firms will not offer them a high enough wage to induce them to do so. Societies where skilled labour is an important input into production or where skill development is very elastic in parental income can maintain a very large lower class.

On the other hand, K_0 must be sufficiently small for four things to happen. The first condition prevents the upper class workers from switching to the simple task. The second prevents the lower class workers from switching to the complex task. Since an increase in K_0 lowers the wages of the upper class and raises those of the lower class, either is possible for sufficiently high values of K_0 . The third condition ensures that all workers performing the complex task are capable of doing so, while the last condition ensures that all workers performing the simple task are not. This is essential for class division.

1.5 A numerical example

To make these ideas clearer, let us consider a numerical example for which $\gamma b < 1$. Let $\gamma = 0.5$, $b = 1.5$ and $a = 1$. From equations 1.14 and 1.17 we can calculate that an upper class containing between 1.6 per cent and 50 per cent of the population could be maintained in steady state. If there is less than 20 per cent of the population in the upper class, they will all be talented. At a point below which 1.6 per cent of the population are performing the complex task, the incentive to switch is sufficient for the lower classes that they do so, and there is no longer an upper and lower class within society. This is illustrated in figure 1.3, panel (a).

One of the interesting features of the figure is that if K_0 is below 0.5, output is higher. When not all complex workers are talented ($K_0 > 0.2$),

$$y^* = \left[a^\gamma \gamma^{\gamma b} K_0^{\gamma(1-b)} [1 - K_0]^{1-\gamma} \right]^{\frac{1}{1-\gamma b}}$$

Differentiating with respect to K_0 and using the knowledge that $b > 1$ and $\gamma b < 1$, we find y^* is decreasing in K_0 . Not only that, both the lower and upper classes gain from a rich upper class minority and inequality. This is because the misallocation of workers is reduced once firms have a marker, class, on which to base their wage decisions. Workers self select based on these wages in such a way that the effective labour input

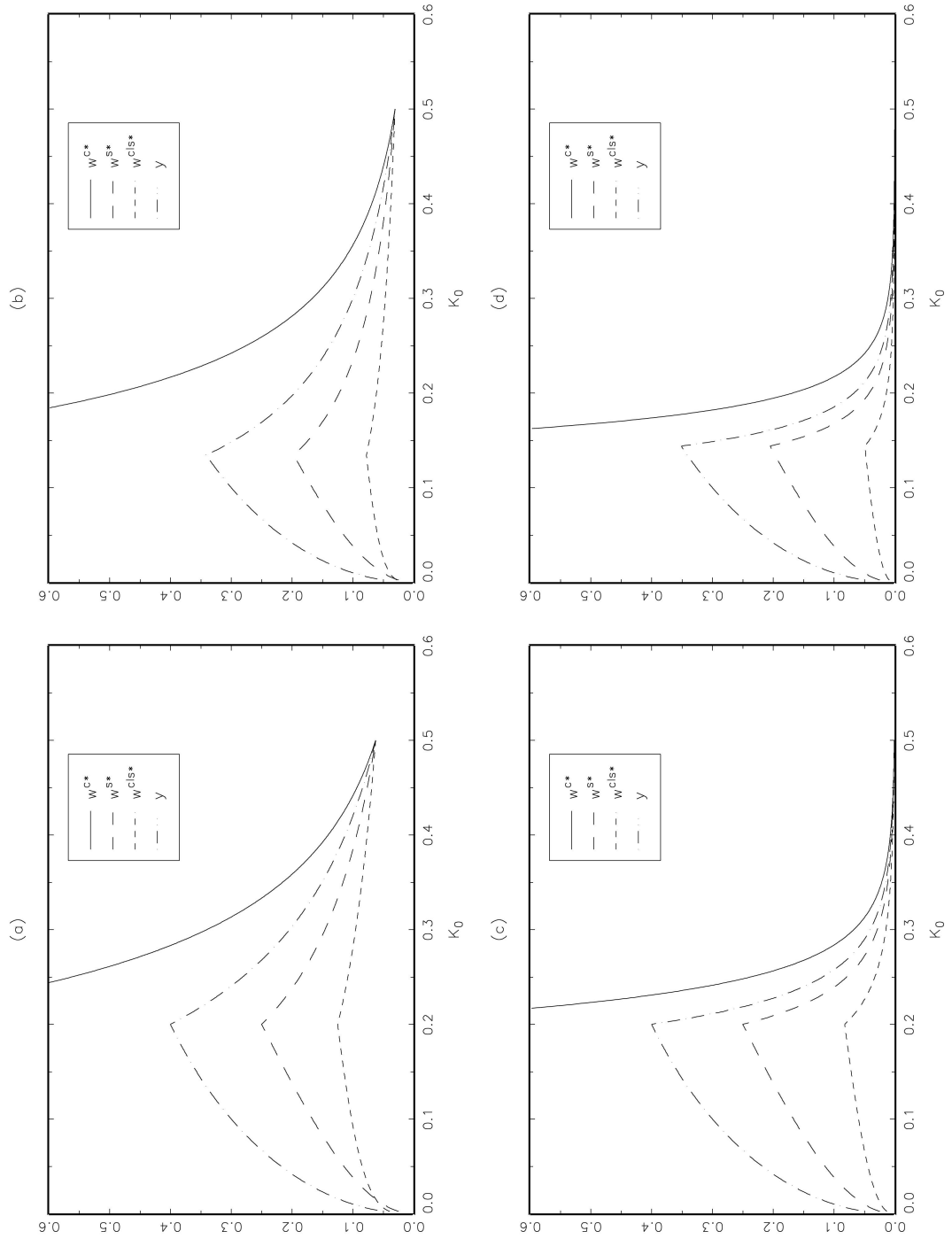


Figure 1.3: Steady state wages and output for different steady state sizes of the upper class when $\gamma b < 1$: (a) $\gamma = 0.5; a = 1; b = 1.5$; (b) $\gamma = 0.5; a = 0.7; b = 1.5$; (c) $\gamma = 0.5; a = 1; b = 1.8$; (d) $\gamma = 0.5; a = 0.7; b = 1.8$

into the complex task is increasing, even as fewer people are doing it. Their increased productivity spills over into the simple sector. These gains are exhausted at K_0 equal to 0.2 because at this point all the complex workers are talented. Steady state output is then given by,

$$y^{**} = K_0^\gamma [1 - K_0]^{1-\gamma}$$

which is increasing in K_0 . Thus the optimum steady state in terms of total output is at $K = 0.2$ and not $K_0 = 0.5$.

Panel (b) decreases a to 0.7. Thus for a given parental income, the probability that a child is talented is lower. This could be one effect of a relatively weak education system. The gains from having fewer complex workers extend to lower values of K_0 since, for a given K_0 the upper class wage will be lower when a is lower, and for a given upper class wage the education system struggles to produce exclusively talented workers. In fact output will, in general, tend to be relatively low in such a society and the gains from class division are weaker. The level of talent in the classless steady state is lowered and the peak output level falls from 0.4 to below 0.35.

An effect of weakness in the public education system is likely to be a relatively high value of b . Private investments are likely to play a greater role in talent development when public education is performing poorly. An increase in b to 1.8 is illustrated in panel (c). Its main effect is the increase in convexity of the wage and output curves. When private investment in children's development plays a large role, the gains from having an identifiable class of skilled workers come at relatively low values of K_0 . There are large benefits to having a small upper class since it generates a sufficiently high probability that the upper classes are able to perform their task. ⁶

Panel (d) of figure 1.3 illustrates both a fall in a and an increase in b relative to panel (a). It combines the changes which took place in panels (b) and (c). If this roughly corresponds to a fall in the efficiency of the public education system, both through lowering the probability of being talented for a given parental income and increasing

⁶Considering figure 1.2(a), the multiplier effect from a fall in K_0 is greater when b is larger. Thus there is more to gain in this case from squeezing K_0 and creating a very small upper class, as it can raise their steady state wage and probability of being talented greatly.

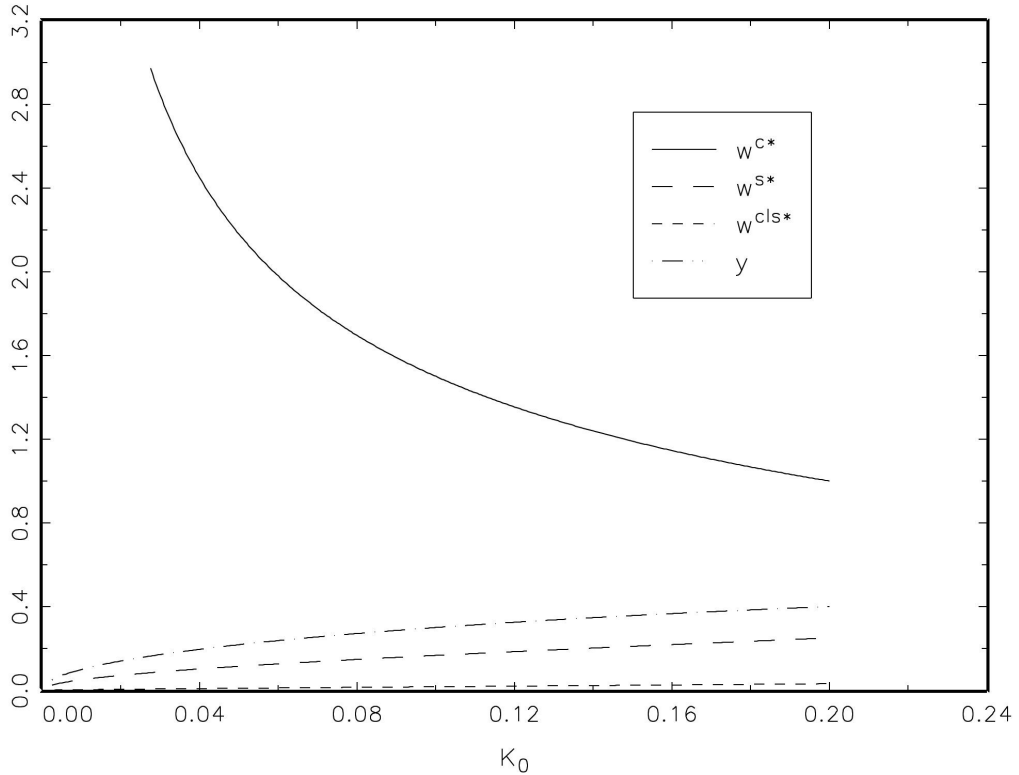


Figure 1.4: Steady state wages and output for different steady state sizes of the upper class when $\gamma b > 1$: $\gamma = 0.5$; $a = 1$; $b = 2.5$

the role of private investments, a small, rich upper class will maximise output and provide a Pareto improvement over the classless steady state. A small number of workers in the complex sector is required to create a class that is significantly more talented than the other.

In all four of these illustrations the overall message is the same: by creating an identifiable class of workers who are more likely to be talented, the economy, and everyone in it, can do better.

Figure 1.4 shows the situation when $\gamma = 0.5$, $b = 2.5$ and $a = 1$. In this case $\gamma b > 1$ and so the only stable steady state complex wage is one where all of the upper class are talented. In order for this to happen, they must number at most 20 per cent of the population. Any reduction in their size below this causes a larger imbalance between the complex and simple sector workforces (each has a 50 per cent share of output) with no gain in terms of the information available to firms.

1.6 Extending the model

Section 1.4 found unique steady state distributions of workers and income where a proportion K_0 of individuals were performing the complex task and receiving w^{c*} and a proportion $1 - K_0$ were carrying out the simple task and receiving w^{s*} . These could be thought of as classes, since $w^{c*} > w^{s*}$ and generations of families received the same wage.

This section shows how the same principles could be applied to a richer model of the economy where there are a continuum of tasks to choose from, and hence a continuum of wages and classes. A subscript i will represent the task, where i is on a scale 0 to 1 with the higher valued i 's representing more demanding tasks (those for which it is more important to have talented people). We now use a CES production function of the form,

$$y = \left[\int_0^1 \tilde{\chi}_i^\sigma di \right]^{\frac{1}{\sigma}}$$

where $\tilde{\chi}_i$ is the effective labour input into task i . Let $P_i(w)$ be the probability that an individual with parental wage w can perform task i and for a given wage $P_a(w) > P_b(w)$ for any $a < b$. The steady state wage then becomes,

$$w_i = \frac{\partial y}{\partial \tilde{\chi}_i} P_i(w_i)$$

This is the equivalent of equation 1.3 in steady state. There must also be a similar condition on K to 1.8, which in this case will be,

$$\frac{\partial y}{\partial \tilde{\chi}_i} P_i(w_i) \geq \frac{\partial y}{\partial \tilde{\chi}_j} P_j(w_i) \quad \forall i, j \neq i$$

It states that in steady state K_i must be such that the best any individual can do,

given that he is from class i , is perform task i . This keeps the proportions of people allocated to each task constant across time and maintains the separation between the classes. Individuals would be impeded or advantaged by birth in exactly the same way as in the two sector model described above.

1.7 Conclusion

The chapter lays the foundations upon which I will build in the remainder of the thesis. It investigated the advantages of birth and the extent to which your background, or class, may be used as a signal of your abilities. Equally it describes the impediments which the skilled members of the lower class face. They freely choose the job for which the steady state wage is lower despite being capable of performing the one for which it is higher. Their class ensures that, from their point of view, this is the optimal thing to do.

Convexity has been seen to be important. Without convexity in the probability of being talented with respect to parental income, it is difficult to sustain a steady state with class division. Convexity implies that the elasticity must be greater than one. This perhaps is not surprising. Convexity tends to pull people apart. However, classes may only be of benefit if this does not occur to too great an extent, and if the deviation in the size of classes is not too great compared to the classless steady state.

The main result of this chapter can be seen in figure 1.3. For γb less than one, steady state output is higher in an economy with a class divide than the classless society. Instead of everyone having the same probability of being talented, society specialises. The upper class specialise in the complex task and achieve a high probability of talent and high marginal productivity. The spillover to the lower classes also raises their marginal productivity.

This idea has been captured in other papers, notably those of Norman and Moro (Moro (2003); Moro & Norman (2004); Norman (2003)) where allocative efficiency

could be enhanced by specialisation. They do not derive conditions under which such specialisation is possible in their model nor is the signal upon which division is possible endogenous. This chapter is a step in that direction.

We require $1 < b < 1/\gamma$. The probability of being talented must be sufficiently elastic in parental income that, for lower values of K_0 , the lower classes will not find it optimal to switch to the complex task. They need to be discriminated against to a large enough extent. At the same time, to generate the stable intermediate wages required, it must be that b is not too large. The only class based economies with $b > 1/\gamma$ are ones for which output would be at least as high if the complex sector were larger.

The gains from specialisation extend over a wider range of steady state distributions of workers if a is smaller. If, given parental income, the probability of being talented is lower, perhaps through deficiencies in the education system, the economy can gain from relatively large deviations in the size of the upper class from the classless state. These gains, however, tend to be more limited since deviations from the allocation of workers in the classless steady state come at a cost. Similarly, large deviations from the classless steady state lead to gains in output and wages if b is large. This corresponds to the increasing importance of private investments, a feature of a weak public education system.

A final feature which the class divided society of this model captures is that all groups can gain from specialisation, including the lower classes. As output is increased, the lower classes capture some of the increase. There is a trickle-down effect. So although the lower classes are *relatively* disadvantaged in our class based society, they may be better off than they would be in the classless alternative. However, raising the incomes of the upper classes when $\gamma b > 1$, or when the size of the lower class is large, will produce no such gains amongst the lower classes. They exist only up to a point. Equally this should not mask the discrimination that the skilled lower class workers, and advantages that the unskilled upper class workers, are experiencing. The following two chapters will consider these issues in more detail.

Chapter 2

The Inheritance of Advantage¹

In chapter 1 we saw how class, or parental income, allowed workers born to particular families to inherit advantages. There were upper class agents who were untalented but paid more than lower class agents who were talented. The reason was that statistically, when talent was unobservable, they were a better bet to be talented. It was not whether they were talented that mattered. What mattered was whether their background made them *look* talented.

In this chapter we expand on the same story. There will be three main differences: firstly, parental income (class) will not be observed perfectly but only with some noise; talent will also be observable with noise; and both income and talent will follow a normal distribution. We consider to what extent the dispersion of income, and its correlation across generations, is related to the inheritance of advantages. In so doing it attempts to go beyond a story of capital market imperfections and consider how firms use the information available to them.

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. Krueger’s illustration

¹This chapter is based upon work I have undertaken for a paper of the same title which I am coauthoring with José V. Rodríguez Mora at Edinburgh University

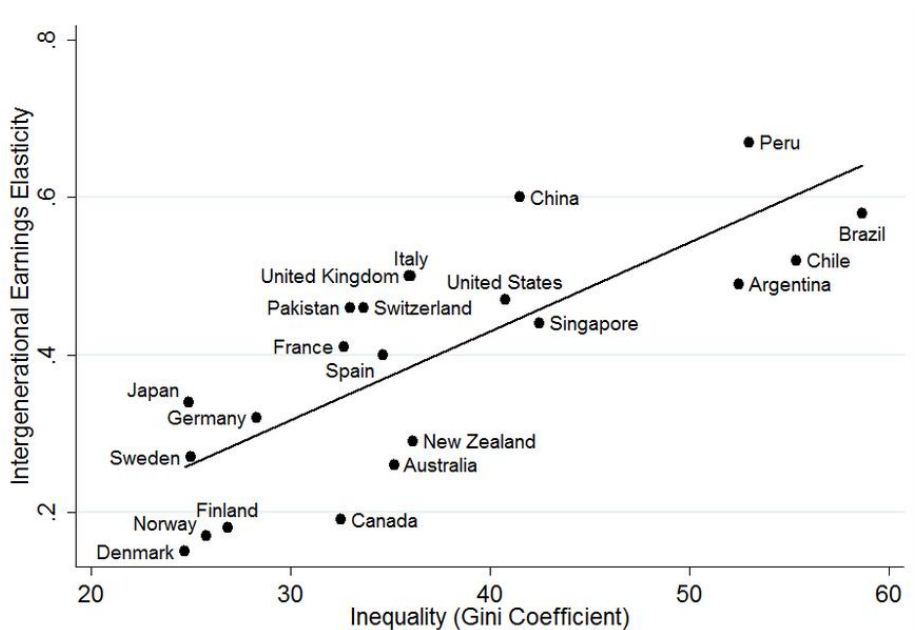


Figure 2.1: Corak’s Great Gatsby Curve

drew on data provided by Miles Corak and Figure 2.1 reproduces Corak’s version of the Great Gatsby Curve (Corak (2012)). Income inequality and the intergenerational earnings elasticity are on the x- and y-axes respectively. We can see that more unequal countries have a greater correlation between the income of fathers and sons or, put another way, that there is 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 in Solon (2004). In Solon’s 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 upward sloping line illustrated in figure 2.1. This paper will offer 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, perhaps through some sort of capital market imperfection allowing the rich to invest more in their children's development. We will look specifically at the underlying investment process in chapter 3. 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) and Moro & Norman (2004)). We 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. This feedback mechanism may lead to multiplicity of steady states. In addition, the better firms get at identifying talent, the lower is mobility, both because the talented tend to be from rich backgrounds and because one of the ways which firms identify talent is through information on family background.

These are the two key points which I would wish to highlight in this chapter. First, that there exists a multiplier effect whereby increased inequality improves the information available to firms when selecting workers, reducing mobility and encouraging further inequality in the incomes that they pay. When there is strong inheritability, in the sense that talent is closely related to family background, this can lead to multiplicity. Second, that it 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 could *decrease* mobility. As a result, 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. I will elaborate further on these points in the models

which follow.

2.1 Labour market sorting where firms receive a signal on parental income

In this section we will consider how firms react to a signal on the parental income of a worker. This is going to be relevant to the firm because parental income and talent will be related in the following way,

$$\tau = \alpha (y_{-1} - \bar{y}_{-1}) + \epsilon_{\tau} \quad (2.1)$$

where τ is a worker's talent, y_{-1} is his parent's income, \bar{y}_{-1} is the mean of parental income, and $\epsilon_{\tau} \sim N(0, \sigma_{\epsilon_{\tau}}^2)$. The talent process is a function of parental income and luck, where the extent of the role played by parental income is governed by the parameter $\alpha > 0$. Talent is the post-education, pre-labour force level of skill of a worker, and its development has a tournament component which will keep average talent centered on zero. We do not consider the factors underlying this relationship in this chapter but you could imagine it being the result of capital market imperfections.

2.1.1 A firm's beliefs about talent

There are firms which pay wages according to their beliefs about the talent an individual has. Let θ_1 be the information available to a firm. θ_1 can be broken down into two elements: μ , information on the distribution of income in the parents' generation; and s_1 , a signal on y_{-1} given by,

$$s_1 = y_{-1} + \epsilon_{s_1} \quad (2.2)$$

where $\epsilon_{s_1} \sim N(0, \sigma_{\epsilon_{s_1}}^2)$. This signal can be thought of as an observation on an individual's background: where they grew up, how they speak, how they dress, and so on. Firms do not, at this stage, receive a direct signal on talent. We assume that the firm knows the distributions of ϵ_{s_1} and ϵ_τ .

Using this information, the firm takes expectations of talent, conditional on the information available, in the following way,

$$E(\tau|\theta_1) = \alpha E(y_{-1} - \bar{y}_{-1}|\theta_1) + E(\epsilon_\tau|\theta_1) \quad (2.3)$$

Since both μ and s_1 are independent of ϵ_τ , the firm's expectation of ϵ_τ conditional on θ_1 will be equal to zero and the firm can concentrate on forming beliefs on parental income. To do this they initially form their prior, conditional on μ , and then update it with the information contained in an individual's signal using Bayesian inference. This leads to an expectation of parental income (as a deviation from its mean) given by,

$$E(y_{-1} - \bar{y}_{-1}|\theta_1) = \frac{\sigma_{y_{-1}}^2}{\sigma_{y_{-1}}^2 + \sigma_{\epsilon_{s_1}}^2} (s_1 - \bar{s}_1)$$

Substituting this back into equation 2.3 gives a posterior belief on talent of,

$$E(\tau|\theta_1) = \frac{\alpha \sigma_{y_{-1}}^2}{\sigma_{y_{-1}}^2 + \sigma_{\epsilon_{s_1}}^2} (s_1 - \bar{s}_1) \quad (2.4)$$

$\frac{\alpha \sigma_{y_{-1}}^2}{\sigma_{y_{-1}}^2 + \sigma_{\epsilon_{s_1}}^2}$ is equal to $\frac{\sigma_{\tau, s_1}}{\sigma_{s_1}^2}$ which is the coefficient from an OLS regression of τ on s_1 .² Thus the firm adjusts away from their prior belief (zero) and towards the signal to the extent that a change in the signal implies a change in talent. When parental income variance is high, the prior gives little information as people are very different. The signal is more heavily used in determining the posterior and hence talent. When parental income variance is zero, the prior gives perfect information and the signal is

²Since τ and s_1 are bivariate normal it can also be read directly from the PDF of the conditional distribution of τ .

disregarded. The point, as we will see in the next section, is that income variance is endogenous.

2.1.2 Steady State Beliefs

Suppose that firms set income equal to their belief about an individual's talent having observed their signal. They have no further opportunity to learn about the talent of their worker. We will write this income as a linear function of the deviation of the signal from its mean value.

$$y = E(\tau|\theta_1) = \beta_1 (s_1 - \bar{s}_1)$$

where β_1 can be read from equation 2.4. Since mean income is then zero in every generation, this implies that the mean signal is zero. We can then write this equation as

$$y = E(\tau|\theta_1) = \beta_1 s_1 \tag{2.5}$$

β_1 measures the optimal reaction of the firm to the signal. The more the firms react to the signals, the more income variance there is. But, as we saw above, the more income variance there is, the more value there is in the signal relative to the prior and so the more the firms react to the signal. β_1 both causes and is a reaction to income variance.

By substitution from equation 2.2 we can calculate the variance of income as a function of the variance in parental income and, by setting $\sigma_y^2 = \sigma_{y-1}^2$, the steady state income variance. This is given by,

$$\sigma_y^2(\beta_1) = \begin{cases} \frac{\beta_1^2 \sigma_{\epsilon s_1}^2}{1 - \beta_1^2} & \text{if } \beta_1 < 1 \\ \infty & \text{if } \beta_1 \geq 1 \end{cases}$$

We can see that, at least up to a point, σ_y^2 is a function of β_1 . We also know from equation 2.4 that $\beta_1 = \frac{\alpha \sigma_y^2}{\sigma_y^2 + \sigma_{\epsilon s_1}^2}$ or, more generally, β_1 is a function of σ_y^2 (which we shall call $F(\cdot)$). Finding a steady state value of β_1 is thus a matter of finding a fixed point of the equation $\beta_1 = F(\sigma_y^2(\beta_1))$ which is given by,

$$\beta_1 = F(\sigma_y^2(\beta_1)) = \frac{\alpha \sigma_y^2(\beta_1)}{\sigma_y^2(\beta_1) + \sigma_{\epsilon s_1}^2} = \begin{cases} \alpha \beta_1^2 & \text{if } \beta_1 < 1 \\ \alpha & \text{if } \beta_1 \geq 1 \end{cases}$$

There are a maximum of three steady state values of β_1 : $\beta_1 = 0$ will always be a steady state; $\beta_1 = 1/\alpha$ is a steady state when α is greater than one; and $\beta_1 = \alpha$ is a steady state when α is greater than one.

The simplest case is where α is less than one. In this instance there is only one steady state value of β_1 equal to zero, corresponding to a steady state income variance of zero. This is very intuitive. If the firm pays every worker the same income then by equation 2.1 their children's talent will be distributed entirely by luck. By definition the firm can infer nothing about luck through the signal, and so they ignore it and pay all the children the same income. The case where α is less than one is shown in figure 2.2 where $F(\beta_1)$ is shorthand for $F(\sigma_y^2(\beta_1))$. The economy will always tend to this steady state.

When α is greater than or equal to one there still exists a steady state with no income variance and no reaction by the firm to an individual's signal. This is stable in the sense that for small perturbations around zero the economy will return to this steady state. There also exist two other steady states: there is an unstable steady state where β_1 is equal to $\frac{1}{\alpha}$; and there is a stable steady state with β_1 equal to α . We are more interested in the latter. The situation where α is greater than one is shown in

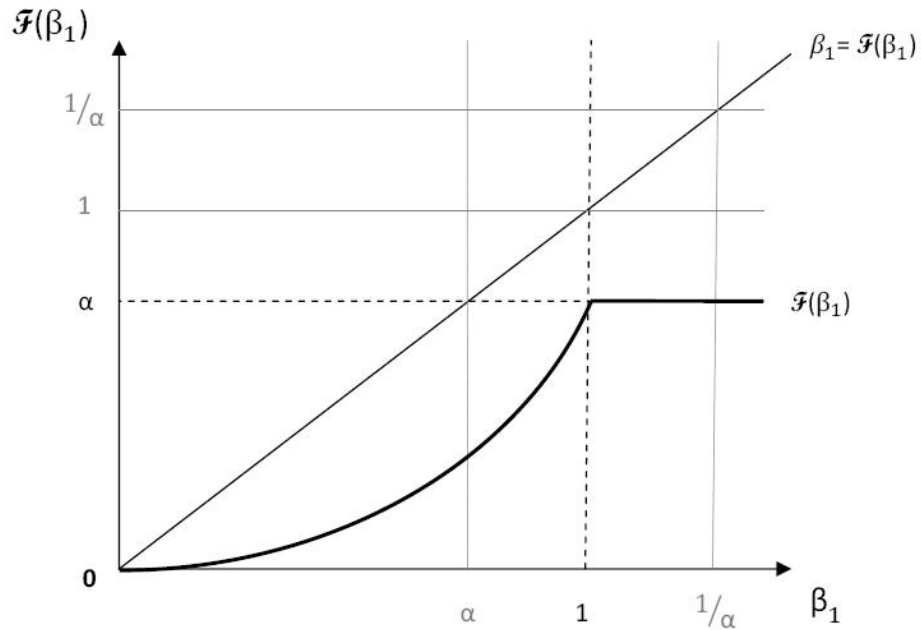


Figure 2.2: Steady State for $\alpha < 1$

figure 2.3.

The stable steady state with β_1 equal to α has income variance growing over time towards infinity. As income variance in the economy becomes very high, incomes will be so different that it will be perfectly evident who is the child of whom. As a result the firm fully uses the signal to the extent that parental income tells them about talent (α).

2.1.3 The feedback mechanism

The fact that two stable steady states may emerge is one of the main findings that we will consider in the rest of this paper. When income variance is decreasing, firms care less about the signal because people are increasingly similar and the signal is known to be uninformative. They use it less and income variance continues to fall. When income variance is increasing, firms care more about the signal because it is becoming more informative and people more different. Increasingly there is an extra meaningful dimension in which people differ. As firms use the signal more, this feeds back into

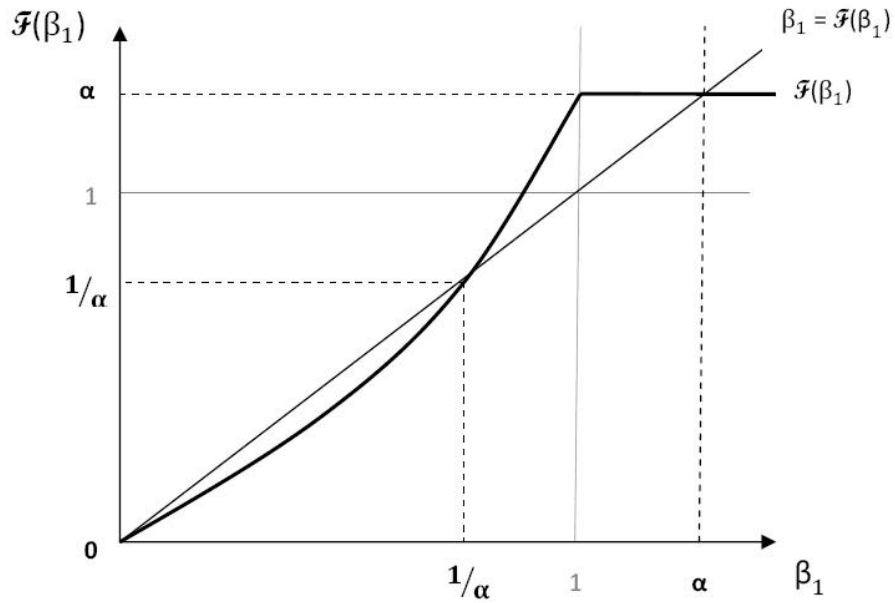


Figure 2.3: Steady States for $\alpha > 1$

greater income inequality. As a result of this latter process, there emerges a portion of the population who are rich and give advantages to their children that go beyond their talents - the ability to display their privileged upbringing and be paid for it.

2.2 Labour market sorting where firms receive a signal on talent

In the model of section 2.1 the information set of the firm consisted of two parts: μ , prior information about the distribution of income in the parent's generation; and s_1 , a signal on parental income. In this section we consider the actions of a firm when faced with an alternative information set θ_2 . They receive the same prior information but an alternative signal, s_2 , on an individual's talent given by,

$$s_2 = \tau + \epsilon_{s_2} \quad (2.6)$$

They do not receive signal s_1 . Talent is still given by equation 2.1 and income by

$E(\tau|\theta_2)$. Later, we will model an improvement in meritocracy as a fall in $\sigma_{\epsilon s_2}^2$ since this improves the precision of the signal and allows talent to be better identified and rewarded.

We calculate the posterior beliefs of the firm, given this new information, which gives it beliefs about talent of,

$$E(\tau|\theta_2) = \frac{\alpha^2 \sigma_{y_{-1}}^2 + \sigma_{\epsilon\tau}^2}{\alpha^2 \sigma_{y_{-1}}^2 + \sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2} s_2 \quad (2.7)$$

The income that the firm pays a worker is a linear function of the new signal,

$$y = E(\tau|\theta_2) = \beta_2 s_2 \quad (2.8)$$

where the value of β_2 can be read from equation 2.7. From this we can quite easily calculate the steady state income variance as a function of β_2 , remembering that $\sigma_y^2 = \sigma_{y_{-1}}^2$ in steady state,

$$\sigma_y^2(\beta_2) = \begin{cases} \frac{\beta_2^2 (\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2)}{1 - (\alpha\beta_2)^2} & \text{if } \beta_2 < \frac{1}{\alpha} \\ \infty & \text{if } \beta_2 \geq \frac{1}{\alpha} \end{cases}$$

We are now in a position to write the fixed point equation $\beta_2 = G(\sigma_y^2(\beta_2))$ where $G(\cdot)$ is the function of parental income variance given in equation 2.7. This gives,

$$\beta_2 = G(\sigma_y^2(\beta_2)) = \frac{\alpha^2 \sigma_y^2(\beta_2) + \sigma_{\epsilon\tau}^2}{\alpha^2 \sigma_y^2(\beta_2) + \sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2} = \begin{cases} \frac{(\alpha\beta_2)^2 \sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2}{\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2} & \text{if } \beta_2 < \frac{1}{\alpha} \\ 1 & \text{if } \beta_2 \geq \frac{1}{\alpha} \end{cases}$$

This is a similar result to what we found in section 2.1. There are up to two steady state values of β_2 less than $\frac{1}{\alpha}$ given by,

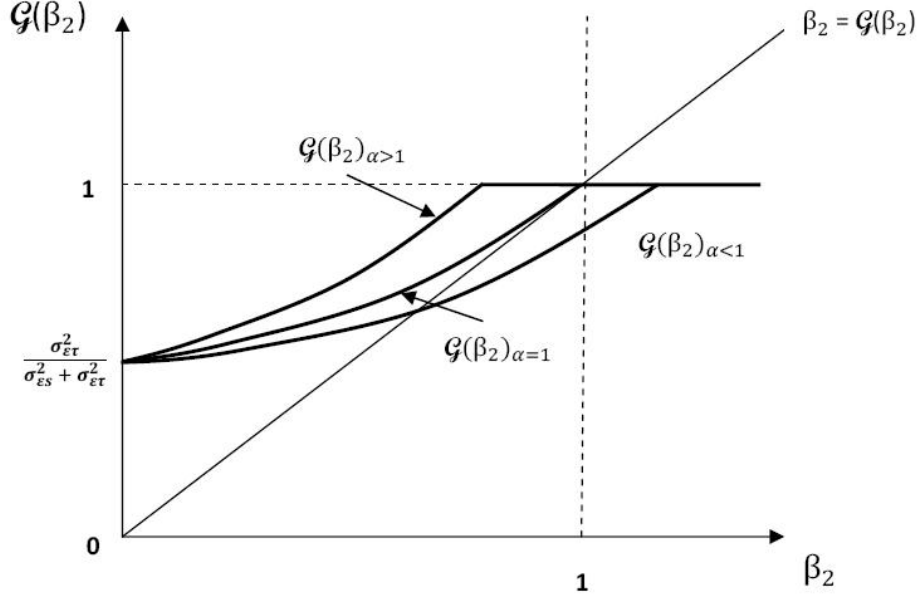


Figure 2.4: Only one steady state exists for a given α and $\sigma_{\epsilon\tau}^2 \geq \sigma_{\epsilon s_2}^2$

$$\beta_2 = \frac{\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2 \pm \sqrt{(\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2)^2 - 4\alpha^2 \sigma_{\epsilon s_2}^2 \sigma_{\epsilon\tau}^2}}{2\alpha^2 \sigma_{\epsilon s_2}^2}$$

We will call the lower of these $\underline{\beta}_2$ and the higher $\bar{\beta}_2$. There is a third possible steady state: $\beta_2 = 1$ for values of β_2 greater than $\frac{1}{\alpha}$. We will consider two possible cases: in the first, $\sigma_{\epsilon\tau}^2 \geq \sigma_{\epsilon s_2}^2$; in the second $\sigma_{\epsilon\tau}^2 < \sigma_{\epsilon s_2}^2$.

2.2.1 Steady state solutions where $\sigma_{\epsilon\tau}^2 \geq \sigma_{\epsilon s_2}^2$

The simplest case is where $\sigma_{\epsilon\tau}^2 \geq \sigma_{\epsilon s_2}^2$. When this condition holds there is only ever one steady state value of β_2 .³ When the variance of talent is largely exogenous ($\sigma_{\epsilon\tau}^2$ is

³Begin by considering $\alpha = 1$. The three possible steady states are then: $\beta_2 = \underline{\beta}_2 = 1$ for $\beta_2 < 1$; $\beta_2 = \bar{\beta}_2 = \sigma_{\epsilon\tau}^2 / \sigma_{\epsilon s_2}^2 > 1$ for $\beta_2 < 1$; and $\beta_2 = 1$ for $\beta_2 \geq 1$. The first two of these are contradictions, leaving only one steady state with $\beta_2 = 1$ and $\sigma_y^2 \rightarrow \infty$. An increase in α causes the $G(\beta_2)$ curve to pivot upwards leaving the only steady state as $\beta_2 = 1$. A decrease in α will cause $\underline{\beta}_2$ to fall below one, confirming it as a valid steady state value. $\bar{\beta}_2$ is ruled out once we consider that any steady state solution must lie on the $G(\beta_2)$ curve which increases monotonically to a maximum value of one at $\beta_2 = 1/\alpha$. Therefore any solution to $\beta_2 = G(\beta_2)$ must have a solution with a value less than or equal to one. A decrease in α below one would cause $\bar{\beta}_2$ to increase, and since it was already above one, this rules it out as a steady state. $\beta_2 = 1$ when $\beta_2 \geq 1/\alpha > 1$ is also ruled out by contradiction. There is therefore never multiplicity when $\sigma_{\epsilon\tau}^2 \geq \sigma_{\epsilon s_2}^2$ for any value of α .

high), and the signal is precise ($\sigma_{\epsilon s_2}^2$ is low), a firm's decision over how much to use the signal is largely independent of what other firms are doing. Since $G(\beta_2)$ measures the firm's response to the prevailing β_2 in society it has little slope in this case. The steady state solutions for α in various ranges are shown in figure 2.4.

2.2.2 Steady state solutions where $\sigma_{\epsilon\tau}^2 < \sigma_{\epsilon s_2}^2$

Where $\sigma_{\epsilon\tau}^2 < \sigma_{\epsilon s_2}^2$ there is a possibility that multiple steady state values of β_2 exist. Intuitively when $\sigma_{\epsilon\tau}^2$ is low and $\sigma_{\epsilon s_2}^2$ is high, the variance in talent and quality of the signal are largely driven by income variance. Since this is endogenously determined by the prevailing β_2 in society, any firm's response to changes in the prevailing β_2 is greater than in the previous section. This is why $G(\beta_2)$ has a steeper slope and why we may reach more than one steady state value of β_2 . This is illustrated in figure 2.5.

Qualitatively the situation is very similar to what we saw for β_1 in section 2.1. There are two stable steady states: one at $\beta_2 = \underline{\beta}_2$ with finite income variance; and one at $\beta_2 = 1$ with income variance growing over time (towards infinity). We will refer to these as the "low" and "high" (stable) steady states.

In this case there is an upper and lower bound on α , equal to one and $\alpha^* = \frac{\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2}{2\sqrt{\sigma_{\epsilon s_2}^2 \sigma_{\epsilon\tau}^2}}$ respectively, for which this multiplicity exists. For values of α less than one there is only one steady state value of β_2 less than $\sigma_{\epsilon\tau}^2 / \sigma_{\epsilon s_2}^2$. There is insufficient inheritance (parental income is not an important enough driver of talent) for heavy use of the signal to translate into high enough quality that this heavy use be sustained. For values of α greater than α^* there is only one steady state value of β_2 equal to one. Here the opposite is true. The additional quality of the talent signal generated by even a small increase in the amount of income variance would be sufficient for the talent signal to be used more and more.

The feedback mechanism works in a similar way to what we saw in section 2.1.3. More income inequality leads to greater dispersion of talent. This in turn leads to greater dispersion of the signals and more value being given to the signals by firms, the

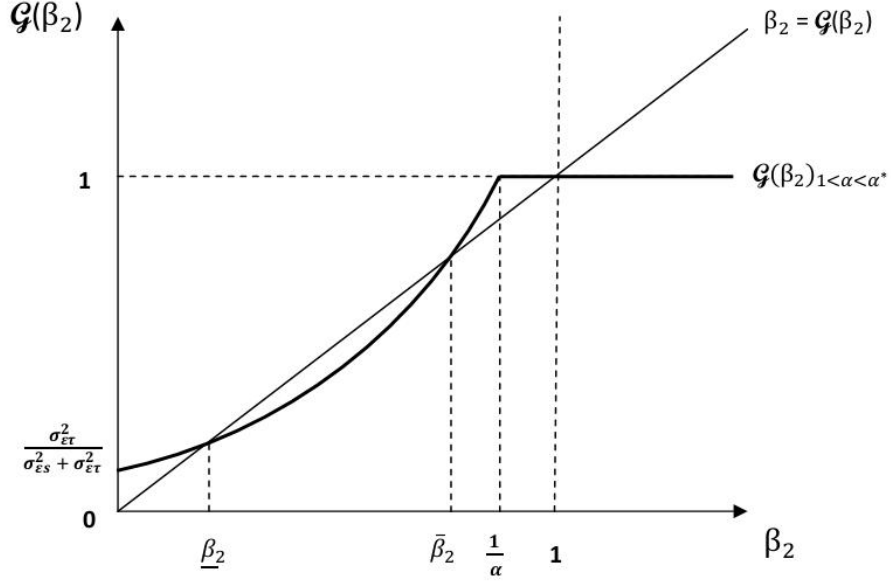


Figure 2.5: When $1 < \alpha < \alpha^*$ and $\sigma_{\epsilon\tau}^2 < \sigma_{\epsilon s_2}^2$ there are 3 steady states, 2 of which are stable

combined effect of which is greater income inequality.

2.2.3 The curse of meritocracy I

We model an increase in meritocracy as a fall in $\sigma_{\epsilon s_2}^2$. This improves the quality of the information which firms have about talent and makes them better able to pay workers according to their talent. From figure 2.5 we can see that a fall in $\sigma_{\epsilon s_2}^2$ causes the intercept of $G(\beta_2)$ to increase and, since it must reach the same point at $\beta_2 = 1/\alpha$, the slope to fall.⁴ $\underline{\beta}_2$ increases. The steady state intergenerational correlation of incomes is given by,

$$\rho_{y,y-1} = \begin{cases} \alpha\beta_2 & \text{if } \beta_2 < 1/\alpha \\ 1 & \text{if } \beta_2 \geq 1/\alpha \end{cases}$$

⁴Formally, the slope of the $G(\beta_2)$ function is increasing in $\sigma_{\epsilon s_2}^2$ at any given value of $\beta_2 < \frac{1}{\alpha}$,

$$\frac{\partial}{\partial \sigma_{\epsilon s_2}^2} \left[\frac{\partial G(\beta_2)}{\partial \beta_2} \right] \Big|_{\beta_2 = \hat{\beta}_2 < \frac{1}{\alpha}} = \frac{2\alpha^2 \hat{\beta}_2 \sigma_{\epsilon\tau}^2}{(\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2)} > 0$$

and so increases in the low steady state as $\sigma_{\epsilon s_2}^2$ falls. Improvements in meritocracy *reduce* mobility.

If we exogenously improve the quality of the signal, firms use it more, increasing the variance of income and talent and further improving the quality of the signal (the feedback mechanism). As this happens, firms get better at identifying talented individuals and paying them accordingly but these individuals tend to be from rich backgrounds, reducing mobility. We produce the observed negative correlation between mobility and inequality in the low steady state. The high steady state is unaffected.

2.3 Labour market sorting where firms receive two signals

In the previous sections we have considered the actions of a firm trying to work out the talent of individual workers with the aid of two alternative information sets, θ_1 and θ_2 . We are now going to combine these so that the firm's information set, which we will now call θ' , has three parts; the prior, μ , based on information on the distribution of income in the parents' generation; signal s_1 on parental income given by equation 2.2; and signal s_2 on talent given by equation 2.6. By reintroducing the signal on parental income, we will show that this makes the curse of meritocracy worse. In addition to firms being better able to pick out the talented, who happen to be predominantly from rich backgrounds, they may also use the signal on parental income more.

We assume that talent is still given by equation 2.1. As before, the distributions of the errors ϵ_{s_1} , ϵ_{s_2} and ϵ_τ are known to the firm and all are independently distributed.

Using Bayesian inference we can calculate the posterior beliefs of a firm conditional on the new information set θ' . This gives a steady state posterior belief of,⁵

⁵The firm's belief, given in equation 2.9 is a weighted average of three things: a prior on talent, $E(\tau|\mu)$, equal to zero; a belief about talent based solely on signal 1, $E(\tau|s_1)$, equal to αs_1 ; and a belief about talent based solely on signal 2, $E(\tau|s_2)$, equal to s_2 . As σ_y^2 goes to infinity the prior is useless and the weights given to the other parts sum to one. Similarly, if $\sigma_{\epsilon s_1}^2$ equals zero then the prior is useless since the first signal perfectly informs the firm about parental income. Again it is thrown away. The weights given to other parts then reflect the extent to which parental income and the second signal inform about talent and sum to one. When $\sigma_{\epsilon s_2}^2$ is zero, only the second signal is used since it perfectly informs about talent so its weight is equal to one.

$$E(\tau|\theta') = \frac{\alpha\sigma_{\epsilon s_2}^2\sigma_y^2}{\alpha^2\sigma_y^2\sigma_{\epsilon s_1}^2 + (\sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2)(\sigma_y^2 + \sigma_{\epsilon s_1}^2)}s_1 + \frac{\alpha^2\sigma_y^2\sigma_{\epsilon s_1}^2 + \sigma_{\epsilon\tau}^2(\sigma_y^2 + \sigma_{\epsilon s_1}^2)}{\alpha^2\sigma_y^2\sigma_{\epsilon s_1}^2 + (\sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2)(\sigma_y^2 + \sigma_{\epsilon s_1}^2)}s_2 \quad (2.9)$$

Since firm's pay wages according to expected talent we can now write the equation for income as,

$$y = E(\tau|\theta') = \beta_1 s_1 + \beta_2 s_2 \quad (2.10)$$

where β_1 and β_2 can be read from equation 2.9. They are the coefficients of a multivariate regression of talent on s_1 and s_2 . This gives steady state income variance of,⁶

$$\sigma_y^2 = \begin{cases} \frac{\beta_1^2\sigma_{\epsilon s_1}^2 + \beta_2^2(\sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2)}{1 - (\beta_1 + \alpha\beta_2)^2} & \text{if } \beta_1 + \alpha\beta_2 < 1 \\ \infty & \text{if } \beta_1 + \alpha\beta_2 \geq 1 \end{cases}$$

We now have three equations in three unknowns: one each for steady state values of β_1 and β_2 as function of the income variance which can be read from equation 2.9; and one for the steady state income variance as a function of β_1 and β_2 . Unfortunately finding analytical solutions to these three unknowns for finite income variance is unmanageable. As such, we will proceed in a slightly different manner to before.

⁶This is derived from,

$$\begin{aligned} \sigma_y^2 &= \beta_1^2(\sigma_{y-1}^2 + \sigma_{\epsilon s_1}^2) + \beta_2^2(\alpha^2\sigma_{y-1}^2 + \sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2) + 2\beta_1\beta_2(\alpha\sigma_{y-1}^2) \\ &= (\beta_1 + \alpha\beta_2)^2\sigma_{y-1}^2 + \beta_1^2\sigma_{\epsilon s_1}^2 + \beta_2^2(\sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2) \end{aligned}$$

Solving for $\sigma_y^2 = \sigma_{y-1}^2$ gives steady state income variance.

2.3.1 The firm's choice of β_1 when the value of β_2 is fixed

Since it is very difficult to find an analytical solution for β_1 when β_2 and σ_y^2 are both being endogenously determined, what we do instead is to exogenously impose a value of β_2 and ask the question: If this were the value of β_2 which firms faced, what value(s) of β_1 would they tend towards? This method allows us to draw a reaction correspondence for β_1 as a function of different exogenously imposed values of β_2 .

To carry out this method, first we need to know the firms' choice of β_1 when β_2 is exogenous. In section 2.1 we noted that the posterior belief in equation 2.4 gave a value of β_1 equal to the coefficient from a regression of τ on s . This worked because the OLS regressor was the value of β_1 which minimised the variance of expected talent around its true value, which also happens to be how a rational agent using Bayes rule behaves. We use this result again here. The value of β_1 which minimises the variance of $E(\tau|\theta')$ around τ is found by solving the minimisation,

$$\min_{\beta_1} E(\tau - \beta_1 s_1 - \beta_2 s_2)^2 \quad (2.11)$$

which gives,

$$\beta_1 = \alpha(1 - \beta_2) \frac{\sigma_y^2}{\sigma_y^2 + \sigma_{\epsilon s_1}^2} \quad (2.12)$$

Fixing the value of β_2 at zero shuts off the talent signal and returns us to the model of section 2.1. Also, if $\beta_1 + \alpha\beta_2 \geq 1$, we can see from above that income variance will tend to infinity and $\beta_1 = \alpha(1 - \beta_2)$. Since rearranging this gives $\beta_1 + \alpha\beta_2 = \alpha$ the requirement for the existence of a steady state with growing variance is again $\alpha \geq 1$.

By substituting the steady state income variance and solving for β_1 , equation 2.12 gives the $\hat{\beta}_1$ reaction correspondence, where the "hat" indicates that this is the value

of β_1 that firms choose *given an exogenously imposed value of β_2* .

2.3.2 The firm's choice of β_2 when the value of β_1 is fixed

We proceed in exactly the same manner as in the previous section. First, we want to solve the minimisation in equation 2.11 for β_2 to give us the firm's choice of β_2 *given β_1* . Solving the minimisation gives,

$$\beta_2 = \frac{\alpha(\alpha - \beta_1)\sigma_y^2 + \sigma_{\epsilon\tau}^2}{\alpha^2\sigma_y^2 + \sigma_{\epsilon\tau}^2 + \sigma_{\epsilon s_2}^2} \quad (2.13)$$

Substituting in the steady state income variance and solving for β_2 gives us a correspondence $\widehat{\beta}_2$ which describes the firm's choice of β_2 in reaction to a particular exogenously imposed value of β_1 . Setting β_1 equal to zero cuts off the signal on parental income and returns us to the model of section 2.2 where the first signal plays no role. Also, when $\beta_1 + \alpha\beta_2 \geq 1$ we again find that there is a steady state at which income variance tends towards infinity and $\beta_1 + \alpha\beta_2 = \alpha$.

2.3.3 A numerical example

The response correspondences in sections 2.3.1 and 2.3.2 do not easily lend themselves to analytical solutions so a numerical analysis was carried out. The correspondences were drawn for the parameter values: $\alpha = 1.1$; $\sigma_{\epsilon s_1}^2 = 5$; and $\sigma_{\epsilon s_2}^2 = 4$; $\sigma_{\epsilon\tau}^2 = 1$. These were chosen based on what had been learned from the one signal models – that to have multiplicity we require: α to be larger than one but below some upper bound α^* ; and the variance in the signal errors ($\sigma_{\epsilon s_1}^2$ and $\sigma_{\epsilon s_2}^2$) to be large compared to the variance in the error on talent ($\sigma_{\epsilon\tau}^2$). Figure 2.6 shows how the reaction correspondences look for these parameter values. We have focussed only on non-negative, non-complex values of β_1 and β_2 .

Along the x- and y-axis, β_1 and β_2 are respectively fixed at zero. Thus the solutions

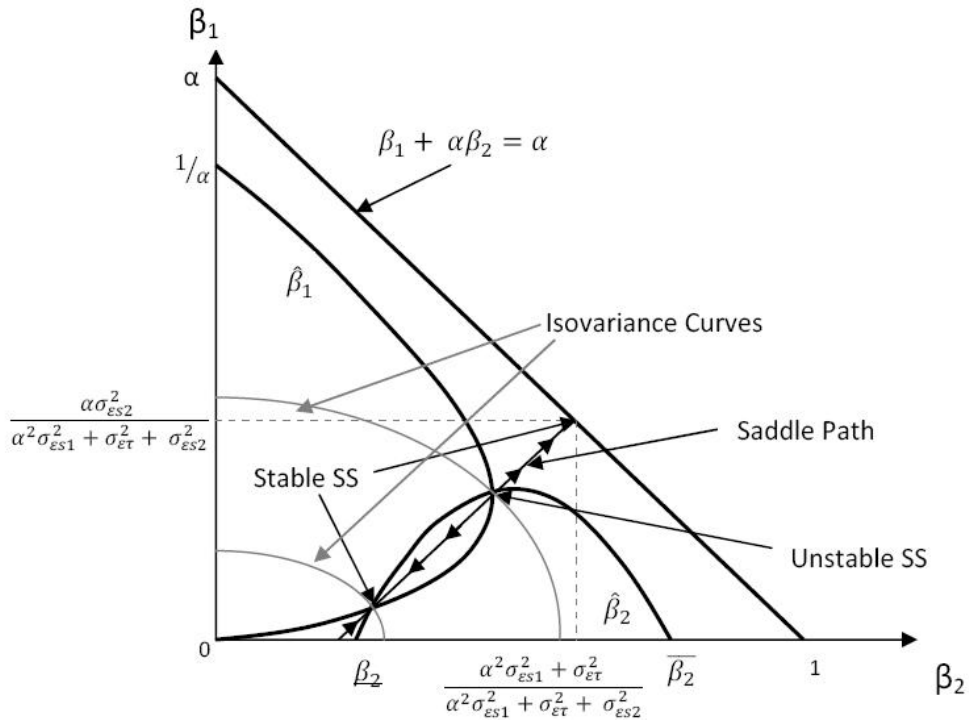


Figure 2.6: An example of multiplicity in the two signal model

along the axes are as in sections 2.2 and 2.1 respectively. The highest of these solutions are connected by the line $\beta_1 + \alpha\beta_2 = \alpha$. From sections 2.3.2 and 2.3.3 we can see that the infinite variance solutions to both $\hat{\beta}_1$ and $\hat{\beta}_2$ lie on this line.

Where the two correspondences cross we find a finite variance steady state solution for β_1 and β_2 with the firm choosing β_1 as a best response to β_2 and β_2 as a best response to β_1 . We find that there are two of these, one with relatively low values of β_1 and β_2 and one with relatively higher ones. The lower one is stable while the higher one is unstable.

The dynamics of the economy work as follows. Suppose β_1 and β_2 are fixed at certain values. This will imply a certain income variance. Then imagine that at some point in time firms are allowed to freely choose β_1 and β_2 . They will do so taking the current income variance as given (since it is our state variable). As a result they act as if they were on the saddle path with the same variance as was created by the initial (β_1, β_2) combination. This is given by the point at which the “isovariance” curve

through this (β_1, β_2) cuts the saddle path. They will jump to a new point on the saddle path, after which they continue to choose values of β_1 and β_2 which follow the saddle path towards steady state. If the initial (β_1, β_2) was inside the higher of the two isovariance curves illustrated, the economy will converge towards the lower stable steady state (with finite variance). If it is outside this isovariance curve the economy will converge towards the high stable steady state with growing income variance. The equation of the saddle path is given by,

$$\beta_1 = -\frac{\sigma_{\epsilon\tau}^2}{\alpha\sigma_{\epsilon s_1}^2} + \frac{\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2}{\alpha\sigma_{\epsilon s_1}^2}\beta_2 \quad (2.14)$$

It is worth noting that, where two stable steady states exist, it is always the case that one has a lower level of both β_1 and β_2 and so there never exists two steady states which substitute one signal for the other.

2.3.4 The curse of meritocracy II

In the two signal model, the intergenerational income correlation is given by,

$$\rho_{y,y-1} = \begin{cases} \beta_1 + \alpha\beta_2 & \text{if } \beta_1 + \alpha\beta_2 < 1 \\ 1 & \text{if } \beta_1 + \alpha\beta_2 \geq 1 \end{cases}$$

Income variance in the finite variance steady state is given by,

$$\sigma_y^2 = \frac{\beta_1^2\sigma_{\epsilon s_1}^2 + \beta_2^2(\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon\tau}^2)}{1 - (\beta_1 + \alpha\beta_2)^2}$$

Since the analytical solutions for β_1 and β_2 are unmanageable, and intergenerational mobility and income inequality are functions of β_1 and β_2 , it is also very difficult to solve analytically the effects of a change in our exogenous parameters on them. We

Parameter	Values the parameter may take
α	0.1, 0.2, ..., 1.5
$\sigma_{\epsilon s_1}^2$	1, 2, ..., 12
$\sigma_{\epsilon s_2}^2$	1, 2, ..., 12
$\sigma_{\epsilon \tau}^2$	0.5, 1.0, ..., 6.0

Table 2.1: Parameter values for the numerical exercise investigating the curse of meritocracy

proceeded with a numerical analysis. First we will examine the effects of a fall in $\sigma_{\epsilon s_2}^2$. This is again how we define an improvement in meritocracy. There are two forces at work as $\sigma_{\epsilon s_2}^2$ falls: *given β_1 and β_2* a fall in the noise will decrease income variance which will lead to a reduction in β_1 and β_2 as firms respond; but *for a given income variance* a fall in $\sigma_{\epsilon s_2}^2$ necessarily increases the use of the talent signal, which in turn will increase income variance and the use of both signals.

We first define our parameter sets to be any combination of parameters from table 2.1. α may take any one of fifteen values, $\sigma_{\epsilon s_1}^2$ any one of twelve, $\sigma_{\epsilon s_2}^2$ one of twelve, and $\sigma_{\epsilon \tau}^2$ one of twelve. This gives close to 26,000 possible combinations of parameters. However for some of those with α greater than one, the finite variance steady state will not exist. We eliminated these cases and were left with 17,918 parameter combinations which produce a finite variance steady state.

For each of these steady states we decreased $\sigma_{\epsilon s_2}^2$ by 0.001. We interpret this as an exogenous marginal increase in meritocracy. In all 17,918 cases for which a finite variance steady state existed, income variance and the intergenerational correlation of incomes both increased. We graph the effect of increasing meritocracy on inequality and mobility for a small number of different parameter sets in figure 2.7. Meritocracy was measured as $\frac{1}{\sigma_{\epsilon s_2}^2}$. The graph illustrates the increase in inequality and fall in mobility which accompanies an increase in meritocracy. There are two possible reasons for the fall in mobility: in all cases, the quality of the second signal has increased and so β_2 increases. This was the effect described in section 2.2.3 where firms get better at picking out the talented but they tend to be from rich backgrounds; additionally, in some cases (all cases for which $\alpha > 1$) the first signal is also used more, reducing mobility since it directly relates to family background.

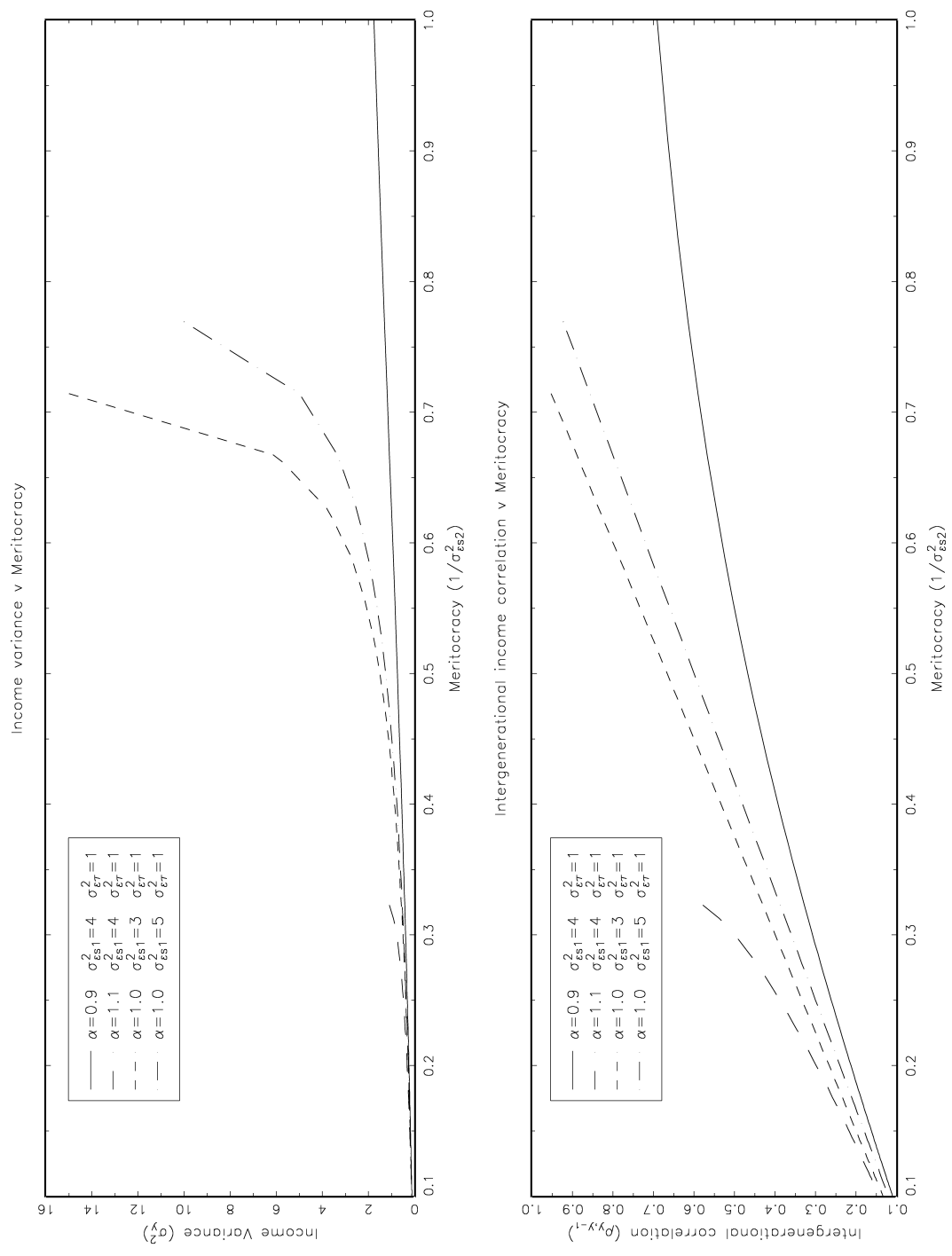


Figure 2.7: The effect of increasing meritocracy on steady state income variance and mobility

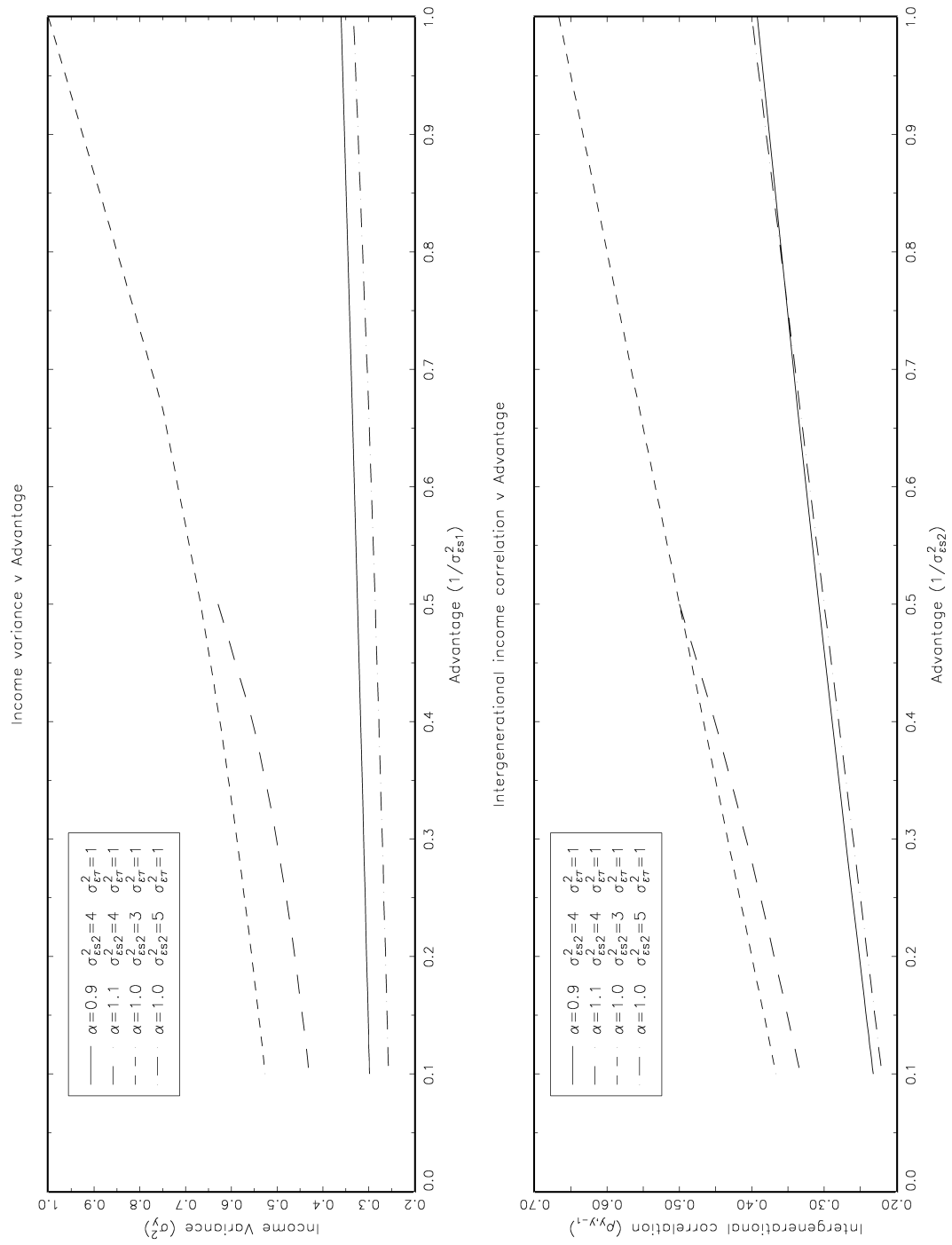


Figure 2.8: The effect of increasing advantages on steady state income variance and mobility

We also investigated whether the correlation of an individual's talent and income increased with a fall in $\sigma_{\epsilon s_2}^2$. If it did not, a fall in $\sigma_{\epsilon s_2}^2$ would not be a useful way of exogenously shocking meritocracy. It occurred in all cases.

In addition to investigating the effects of meritocracy we consider the effect of inherited advantages. An exogenous increase in inherited advantages was defined as a fall in $\sigma_{\epsilon s_1}^2$. With a fall in $\sigma_{\epsilon s_1}^2$, it is easier for firms to identify parental income and, since parental income is correlated with talent, to pay higher incomes to those from richer backgrounds. Thus the advantages of a rich background (and disadvantages of a poor background) are greater. We decreased $\sigma_{\epsilon s_1}^2$ by 0.001 and examined the effects for the parameter sets given in table 2.1. In all 17,981 cases for which a finite variance steady state existed, income variance, the intergenerational correlation of incomes, and the correlation of talent and income increased. Figure 2.8 illustrates how increases in advantages related to background (defined as $\frac{1}{\sigma_{\epsilon s_1}^2}$) lead to greater income inequality and lower intergenerational mobility.

In summary, an increase in the precision of either signal (an increase in meritocracy or advantages related to background) will lead to greater income inequality and further increases in both meritocracy and advantages related to background. This is because the endogenous response of firms to the more precise information is to use one or both of the signals more. We might have expected that when firms are better informed about an individual's talent, they would care less about his background. We show that more often than not (and in all cases where there is sufficient inheritance) this is not true and that even when it is true, mobility still falls due to increased use of the talent signal.

2.4 A discrete model with finite income variance

Until now this chapter has used a model where income and talent were continuous and normally distributed. It has been shown, using three variations on the normal model, how multiplicity may exist, producing two stable steady state outcomes. In the low steady state there was finite income variance, while the high steady state tended

towards infinite income variance. Although the analysis concentrated on the low steady state, the existence of the infinite variance high steady state seems less realistic. We now turn to a model with discrete levels of income, which will provide a simple example of some of the features of the normal model, without the high variance steady state exhibiting infinite steady state income variance.

As in section 2.3, firm's are still unable to observe talent and receive two signals about each worker: the first, which we will call s_r , is a signal on parental income (r represents a rich background); the second, s_t , is a signal on talent. In a departure from the previous model, these signals can only take values of one or zero and their quality will be determined exogenously. Since there are two signals, each of which can take one of two values, there are four "types" of individual. Y_{rt} gives the income of each type, for example Y_{11} is the income of a type 11 worker (one for whom $s_r = 1$ and $s_t = 1$).

Talent, τ , is discrete and only able to take values one or zero. If an individual's parent is sufficiently rich, it is more likely to take a value of one. The probability of being talented is given by,

$$P(\tau = 1) = \begin{cases} \alpha_s + \mu_s(1 - \alpha_s) & \text{if } \tilde{Y}_{rt} > \tilde{K} \\ \mu_s & \text{if } \tilde{Y}_{rt} = \tilde{K} \\ \mu_s(1 - \alpha_s) & \text{if } \tilde{Y}_{rt} < \tilde{K} \end{cases} \quad 0 \leq \alpha_s \leq 1$$

$$S_{rt} = P(\tau = 1 | \tilde{Y}_{rt})$$

\tilde{K} is average talent in the parent's generation which, as we will see below, is equal in steady state to average income in the parent's generation. Those whose parent has above average income are more likely to be talented than those whose parent has below average income. α_s is a parameter controlling inheritance. When α_s is zero, everyone has the same probability of being talented, μ_s , irrespective of who their parent is. When α_s is one, for everyone other than those on the threshold, talent is entirely determined by parental "type". S_{rt} is defined to be the probability of being talented conditional on *parental* income \tilde{Y}_{rt} .

As before, the first signal is on parental income. We will call an individual who looks like she is from a rich background ($s_r = 1$) “advantaged”, while one who looks like she is from a poor background ($s_r = 0$) will be termed “disadvantaged”. An individual’s likelihood of being advantaged is increased if her parent is sufficiently rich. The probability of being advantaged is given by,

$$P(s_r = 1) = \begin{cases} \alpha_r + \mu_r(1 - \alpha_r) & \text{if } \tilde{Y}_{rt} > (1 + \lambda)\tilde{K} \\ \mu_r & \text{if } \tilde{Y}_{rt} = (1 + \lambda)\tilde{K} \\ \mu_r(1 - \alpha_r) & \text{if } \tilde{Y}_{rt} < (1 + \lambda)\tilde{K} \end{cases} \quad 0 \leq \alpha_r \leq 1, 0 \leq \lambda$$

$$R_{rt} = P(s_r = 1 | \tilde{Y}_{rt})$$

\tilde{Y}_{rt} represents parental income so the first line states that an individual’s chances of being advantaged are larger if her parent has an income above $(1 + \lambda)\tilde{K}$. λ is a parameter controlling how rich your parent needs to be, relative to the average, in order to increase your chances of being advantaged. α_r is a parameter which controls the quality of the signal. If α_r is zero then the signal is randomly distributed in the population and hence not correlated with, or informative about, parental income. If α_r is one, the signal tells exactly whose parents have income above and below the threshold $(1 + \lambda)\tilde{K}$. R_{rt} is defined to be the probability that an individual is advantaged given her parent’s income is Y_{rt} .

Finally, the second signal, s_t , is a signal on talent. We will refer to an individual who looks talented ($s_t = 1$) as “promising”, since he is an attractive prospect to a firm, while one which looks untalented ($s_t = 0$) will be referred to to as “unpromising”. The probability that an individual is promising is given by,

$$P(s_t = 1) = \begin{cases} \alpha_t + \mu_t(1 - \alpha_t) & \text{if } \tau = 1 \\ \mu_t(1 - \alpha_t) & \text{if } \tau = 0 \end{cases} \quad 0 \leq \alpha_t \leq 1$$

This states that an individual is more likely be promising if he is actually talented and unpromising if he is not. α_t measures the quality of the signal, where zero is completely uninformative and one completely informative. It is going to be our measure of meritocracy, which will increase as α_t goes from zero to one.

In what follows we are going to make the simplifying assumptions that $\mu_s = \mu_r = \mu_t = \mu$.

2.4.1 Steady state conditions

Let π_{rt} represent the fraction of the population which are of that type. In steady state π_{rt} is constant across generations. We will begin with π_{11} , the proportion of individuals who are both advantaged and promising. The probability that an individual is of type 11 depends on what type their parent is and what their probability of being 11 is conditional on their parents type. This gives,⁷

$$\begin{aligned} P(s_r = 1, s_t = 1) &= \pi_{11} \\ &= \mu(1 - \alpha_t) [R_{11}\tilde{\pi}_{11} + R_{10}\tilde{\pi}_{10} + R_{01}\tilde{\pi}_{01} + R_{00}\tilde{\pi}_{00}] \\ &\quad + \alpha_t [R_{11}S_{11}\tilde{\pi}_{11} + R_{10}S_{10}\tilde{\pi}_{10} + R_{01}S_{01}\tilde{\pi}_{01} + R_{00}S_{00}\tilde{\pi}_{00}] \end{aligned}$$

⁷The proportion of the population which is advantaged and promising is given by,

$$\begin{aligned} P(s_r = 1, s_t = 1) &= \pi_{11} \\ &= P(s_r = 1|\tilde{Y}_{11}) P(s_t = 1|\tilde{Y}_{11}) P(\tilde{Y}_{11}) + P(s_r = 1|\tilde{Y}_{10}) P(s_t = 1|\tilde{Y}_{10}) P(\tilde{Y}_{10}) \\ &\quad + P(s_r = 1|\tilde{Y}_{01}) P(s_t = 1|\tilde{Y}_{01}) P(\tilde{Y}_{01}) + P(s_r = 1|\tilde{Y}_{00}) P(s_t = 1|\tilde{Y}_{00}) P(\tilde{Y}_{00}) \\ &= R_{11}[\alpha_t S_{11} + \mu(1 - \alpha_t)] \tilde{\pi}_{11} + R_{10}[\alpha_t S_{10} + \mu(1 - \alpha_t)] \tilde{\pi}_{10} \\ &\quad + R_{01}[\alpha_t S_{01} + \mu(1 - \alpha_t)] \tilde{\pi}_{01} + R_{00}[\alpha_t S_{00} + \mu(1 - \alpha_t)] \tilde{\pi}_{00} \\ &= \mu(1 - \alpha_t) [R_{11}\tilde{\pi}_{11} + R_{10}\tilde{\pi}_{10} + R_{01}\tilde{\pi}_{01} + R_{00}\tilde{\pi}_{00}] \\ &\quad + \alpha_t [R_{11}S_{11}\tilde{\pi}_{11} + R_{10}S_{10}\tilde{\pi}_{10} + R_{01}S_{01}\tilde{\pi}_{01} + R_{00}S_{00}\tilde{\pi}_{00}] \end{aligned}$$

The third line uses the fact that $P(s_t = 1|\tilde{Y}_{rt}) = P(s_t = 1|\tau = 1)P(\tau = 1|\tilde{Y}_{rt}) + P(s_t = 1|\tau = 0)P(\tau = 0|\tilde{Y}_{rt})$ and that $P(\tilde{Y}_{rt}) = \tilde{\pi}_{rt}$ since all parents of type rt are paid an income Y_{rt} .

We define a steady state to be one in which the proportion of each type is constant over time (i.e. $\pi_{rt} = \tilde{\pi}_{rt}$). We will also define a variable A , which is the proportion of the population which is both advantaged and talented, and is given by,

$$A = R_{11}S_{11}\pi_{11} + R_{10}S_{10}\pi_{10} + R_{01}S_{01}\pi_{01} + R_{00}S_{00}\pi_{00}$$

Then our first steady state condition is,

$$\pi_{11} = \alpha_t A + \mu(1 - \alpha_t) [R_{11}\pi_{11} + R_{10}\pi_{10} + R_{01}\pi_{01} + R_{00}\pi_{00}]$$

Next we want to find the steady state solution for π_{10} , the proportion of the population that is advantaged but unpromising. To do so we will actually find the proportion that are advantaged which equals the unconditional probability that $s_r = 1$. This gives,⁸

$$\begin{aligned} P(s_r = 1) &= \pi_{11} + \pi_{10} \\ &= R_{11}\tilde{\pi}_{11} + R_{10}\tilde{\pi}_{10} + R_{01}\tilde{\pi}_{01} + R_{00}\tilde{\pi}_{00} \end{aligned}$$

By letting $\pi_{rt} = \tilde{\pi}_{rt}$ and substituting our second steady state equation into our first, we are left with the following two steady state conditions:

$$\pi_{11} = \alpha_t A + \mu(1 - \alpha_t) [\pi_{11} + \pi_{10}] \tag{2.15}$$

⁸The proportion of the population which is advantaged is given by,

$$\begin{aligned} P(s_r = 1) &= \pi_{11} + \pi_{10} \\ &= P(s_r = 1|\tilde{Y}_{11})P(\tilde{Y}_{11}) + P(s_r = 1|\tilde{Y}_{10})P(\tilde{Y}_{10}) + P(s_r = 1|\tilde{Y}_{01})P(\tilde{Y}_{01}) + \\ &\quad + P(s_r = 1|\tilde{Y}_{00})P(\tilde{Y}_{00}) \\ &= R_{11}\tilde{\pi}_{11} + R_{10}\tilde{\pi}_{10} + R_{01}\tilde{\pi}_{01} + R_{00}\tilde{\pi}_{00} \end{aligned}$$

$$\pi_{11} + \pi_{10} = R_{11}\pi_{11} + R_{10}\pi_{10} + R_{01}\pi_{01} + R_{00}\pi_{00} \quad (2.16)$$

We derive our next steady state equation by focussing on the proportion of the population which is promising. First, K is the proportion of the population which is talented and so, in steady state, is given by,

$$K = P(\tau = 1) = S_{11}\pi_{11} + S_{10}\pi_{10} + S_{01}\pi_{01} + S_{00}\pi_{00}$$

Using K , we find the following equation for the steady state proportion of promising individuals,⁹

$$\pi_{11} + \pi_{01} = \alpha_t K + \mu(1 - \alpha_t) \quad (2.17)$$

Our last steady state equation draws on the fact that the π_{rt} terms for each generation must sum to one and so,

$$\pi_{11} + \pi_{10} + \pi_{01} + \pi_{00} = 1 \quad (2.18)$$

Equations 2.15 through 2.18 can be solved for the four unknowns (π_{11} , π_{10} , π_{01} and π_{00}) to give the steady state proportions of each type in the population.

⁹The proportion of the population which is promising is given by,

$$\begin{aligned} P(s_t = 1) &= \pi_{11} + \pi_{01} \\ &= P(s_t = 1|\tau = 1)P(\tau = 1) + P(s_t = 1|\tau = 0)P(\tau = 0) \\ &= [\alpha_t + \mu(1 - \alpha_t)]K + \mu(1 - \alpha_t)(1 - K) \\ &= \alpha_t K + \mu(1 - \alpha_t) \end{aligned}$$

2.4.2 The steady state income distribution

Once we have established the steady state distribution of types we can easily define the steady state income which each type receives. We assume that incomes are paid to each type according to the probability that they are talented conditional on their observables. So the income of a type 11 individual will be given by,¹⁰

$$Y_{11} = P(\tau = 1 | s_r = 1, s_t = 1) = \frac{[\alpha_t + \mu(1 - \alpha_t)] A}{\pi_{11}} \quad (2.19)$$

By the same logic,

$$Y_{10} = P(\tau = 1 | s_r = 1, s_t = 0) = \frac{(1 - \alpha_t)(1 - \mu) A}{\pi_{10}} \quad (2.20)$$

$$Y_{01} = P(\tau = 1 | s_r = 0, s_t = 1) = \frac{[\alpha_t + \mu(1 - \alpha_t)] (K - A)}{\pi_{01}} \quad (2.21)$$

$$Y_{00} = P(\tau = 1 | s_r = 0, s_t = 0) = \frac{(1 - \alpha_t)(1 - \mu) (K - A)}{\pi_{00}} \quad (2.22)$$

Any solution to equations 2.15 through 2.22 gives a full characterisation of a steady state distribution of income $(\pi_{11}, \pi_{10}, \pi_{01}, \pi_{00}, Y_{11}, Y_{10}, Y_{01}, Y_{00})$. Equations 2.19 through 2.22 can also be used to show that average income is equal to K ¹¹.

¹⁰The probability that a type 11 individual is talented is given by,

$$\begin{aligned} P(\tau = 1 | s_r = 1, s_t = 1) &= \frac{P(s_r = 1, s_t = 1 | \tau = 1) P(\tau = 1)}{P(s_r = 1, s_t = 1)} \\ &= \frac{P(s_r = 1 | \tau = 1) P(s_t = 1 | \tau = 1) K}{\pi_{11}} \\ &= \frac{[\sum_{rt} P(s_r = 1 | \tilde{Y}_{rt}) P(\tilde{Y}_{rt} | \tau = 1)] [\alpha_t + \mu(1 - \alpha_t)] K}{\pi_{11}} \\ &= \frac{[\sum_{rt} R_{rt} S_{rt} \pi_{rt}] [\alpha_t + \mu(1 - \alpha_t)]}{\pi_{11}} \\ &= \frac{[\alpha_t + \mu(1 - \alpha_t)] A}{\pi_{11}} \end{aligned}$$

¹¹Average income, \bar{Y} , is given by $Y_{11}\pi_{11} + Y_{10}\pi_{10} + Y_{01}\pi_{01} + Y_{00}\pi_{00} = A + (K - A) = K$

2.4.3 An Example

The goals of this chapter, and the focus of the earlier model, were two-fold: to demonstrate the multiplier effect whereby raising income inequality improved information on workers' quality and caused firms to pay more varied wages. Where there was sufficient inheritance this led to multiplicity; and secondly, to demonstrate that a more meritocratic society, defined as one in which firms are better informed of worker quality, displays lower mobility. When both of these points are combined, we see that a society in which firms are better informed will lead to both greater inequality and less mobility in a natural way. We now plan to demonstrate these points within our discrete model, without having infinite income variance in the high steady state.

The feedback mechanism

We begin by setting α_t equal to zero (with $\lambda > 0$). This means that the signal on talent is randomly distributed amongst society and hence uncorrelated with actual talent. It is disregarded by the firm. This is the situation that was previously described in section 2.1. I will use an X subscript to indicate that a signal is irrelevant so, for example, Y_{1X} is the income of an individual who is advantaged when the talent signal is not being used by the firm. In this situation, there can be at most two income levels where firms discriminate amongst workers based on whether or not they are advantaged.

When α_t is equal to zero, the simplest steady state occurs where both income levels lie below the income threshold $(1 + \lambda)K$. This implies that all individuals have the same probability of being advantaged, and so being advantaged or not is not a useful signal to the firm. They disregard it, along with the talent signal, and pay everyone the same income, K . This common income level can be shown to equal μ since,

$$K = P(\tau = 1) = P(\tau = 1 | \tilde{Y} = K) P(\tilde{Y} = K) = \mu$$

This steady state is illustrated in figure 2.9(a) and essentially describes the same

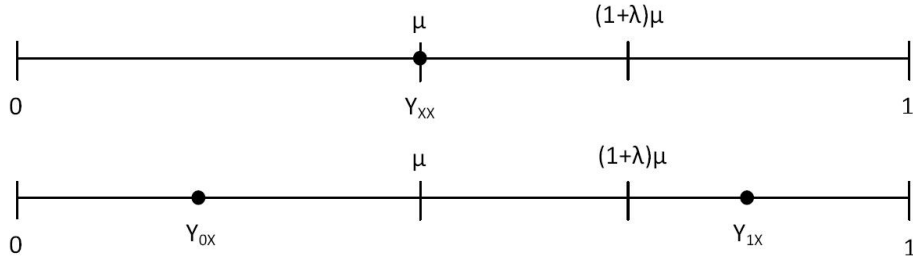


Figure 2.9: The steady states of the economy when α_t equals zero. There are two possibilities: (a) everyone has the same probability of being advantaged and so everyone receives the same income; (b) some people are more likely to be advantaged than others and the advantaged are paid more.

steady state as the zero variance case in section 2.1. This steady state will always exist.

There is an alternative steady state which can be reached where there are two groups of individuals with different probabilities of being advantaged. For this to be a steady state, it must be that the two incomes are sufficiently different so that one group is more likely to be both advantaged and talented, implying advantage is a useful signal of talent. This implies we require that $Y_{1X} > (1 + \lambda)K$ and that $Y_{0X} < K$. It is relatively easy to show that should such a steady state exist there would be a proportion μ of the population who are advantaged and a proportion μ who are talented. The level of incomes would be given by,

$$Y_{1X} = \alpha_r \alpha_s + \mu (1 - \alpha_r \alpha_s)$$

$$Y_{0X} = \mu (1 - \alpha_r \alpha_s)$$

This steady state is illustrated in figure 2.9(b). When the advantaged group is more likely to be talented, they are paid more, implying they are more likely to be talented. It is similar to the high steady state in section 2.1. In particular, this steady state will not always exist (the advantaged group need to be paid *sufficiently* more) but is more likely the greater is the inheritability of talent through parental income.¹²

¹²This steady state exists if $\lambda < \frac{\alpha_r \alpha_s (1 - \mu)}{\mu}$, so a higher value of α_s will improve the likelihood of multiplicity as in section 2.1.

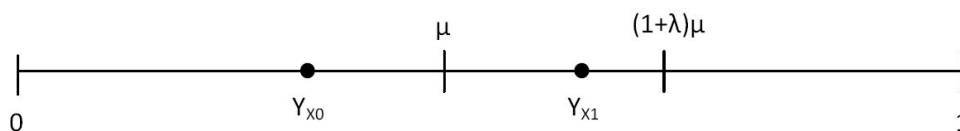


Figure 2.10: The steady states of the economy when α_t is larger than zero but less than $\frac{\lambda\mu}{1-\mu}$

The curse of meritocracy I

We now turn our attention to demonstrating how improvements in meritocracy may lead to increased inequality and falling mobility. To do so, we will focus on the low steady state illustrated in figure 2.9(a). We will increase meritocracy by increasing α_t , the parameter which governs the information which firms have on an individual's talent. The resulting steady state is illustrated in figure 2.10.

As we increase α_t above zero, we produce an economy in which there are two groups of individuals, one of which is more likely to be talented and hence more likely to be promising. The probability of being advantaged is the same for everyone. We can see that the proportion in the promising group in steady state is given by,

$$P(s_t = 1) = \pi_{X1} = \alpha_t K + \mu(1 - \alpha_t)$$

Similarly we can show that the proportion of talented individuals is given by,

$$P(\tau = 1) = K = \alpha_s \pi_{X1} + \mu(1 - \alpha_s)$$

Solving these two equations we find that a proportion μ of the population is promising and that a proportion μ of the population is talented. The correlation between the membership of these two groups will determine the extent to which the signal is useful to the firm. This is determined by α_t . It is easy to show that the incomes which the

firms will pay are given by,

$$Y_{X1} = P(\tau = 1 | s_t = 1) = \alpha_t + \mu(1 - \alpha_t)$$

$$Y_{X0} = P(\tau = 1 | s_t = 0) = \mu(1 - \alpha_t)$$

The difference in the income paid to the two groups depends on the quality of the information available to the firm, α_t . As we increase α_t , Y_{X1} increases and Y_{X0} decreases. In addition, as a higher income is paid to those who are promising (those who are more likely to be talented), and those who are promising tend to come from richer backgrounds, this reduces mobility. The increase in income dispersion and fall in mobility due to increased use of the talent signal are the same effects as we saw in section 2.2.3. This steady state distribution of incomes exists for all α_t less than $\frac{\lambda\mu}{1-\mu}$.¹³

The curse of meritocracy II

Finally, we want to consider what happens when α_t reaches $\frac{\lambda\mu}{1-\mu}$. At this point, Y_{X1} reaches the threshold $(1 + \lambda)\mu$ and the probability of being advantaged jumps from $\mu(1 - \alpha_F)$ to μ for the children of the rich. The fact that some of those from rich backgrounds, who are more likely to be talented, are also now more likely to be advantaged, means that both signals are now useful to the firm. There are four income groups formed by each of our previous groups, Y_{X1} and Y_{X0} , splitting in two. The advantaged members of each group move to a higher steady state income, while the disadvantaged in each group move to a lower steady state income. At this point there are several steady states which the economy could jump to. I will focus attention on one.¹⁴ This is the one where Y_{01} and Y_{10} do not jump far from Y_{X1} and Y_{X0} respectively so that $Y_{10} < K < Y_{01} < (1 + \lambda)K$.¹⁵ This is illustrated in figure 2.11. K is again equal to μ .

¹³At this point Y_{X1} reaches the threshold at which the parental income signal starts to have some value. Although we did not restrict λ , we will assume that this threshold is less than one otherwise we would never leave this steady state.

¹⁴The goal of this section is not to fully characterise all the possible steady state income distributions but instead to provide simple examples of the effects described earlier in the chapter.

¹⁵For this to be the steady state that is reached at $\alpha_t = \frac{\lambda\mu}{1-\mu}$ we require that $\alpha_s < \min \left\{ \frac{\lambda(1-\mu)}{\mu\alpha_r[\lambda^2 + (1+\lambda)(1-\alpha_r)[1-\mu(1+\lambda)]]}, \frac{\lambda\mu}{\alpha_r[\lambda\mu + [1-\mu(1+\lambda)]^2]} \right\}$.

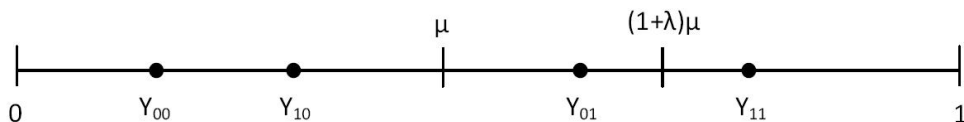


Figure 2.11: One steady state of the economy when α_t is equal to $\frac{\lambda\mu}{1-\mu}$

As could happen in in the normal version of the model, we observe complementarity between the two signals. As we improve the quality of the signal on talent, there will come a point where the firm begins to use the signal on parental income. This is shown in figure 2.11. At the point where firms begin using a person’s advantage to place them in a job there will be a discrete fall in mobility. This is brought about by the increased inequality associated with improved meritocracy. The idea that meritocracy can simultaneously increase inequality and lower mobility was one of the central themes in the previous sections of this chapter.

2.5 An extension

Throughout this chapter we have assumed that firms pay workers a wage equal to their expected talent at the point of entrance into the workforce. There is no further opportunity for them to learn about their worker’s talent. There is a strand of literature which investigates whether, if firms statistically discriminate when sorting new workers, the effects of this discrimination fade as firms observe their worker’s ability on-the-job (Farber & Gibbons (1996); Altonji & Pierret (2001)). Most of this literature has investigated whether a schooling signal is used by firms in the absence of better information on talent, although a paper by Arcidiacono et al. (2008) also looks at race.

A simple way to include such considerations in our framework would be to have the (lifetime) income of an individual be a weighted average of his talent and expected talent.

$$y = \Psi\tau + (1 - \Psi) E(\tau|s_1, s_2)$$

The role played by expected talent reflects the times when the worker's ability is unobservable to his firm and so it relies on observable signals. The relative weight given to expected talent, $(1 - \Psi)$, would indicate the amount of time that the worker spends in this situation during his lifetime, and hence things such as the speed of learning by employers (the subject of a paper by Lange (2007)). This change would reduce, but not eliminate, the role of signalling. This chapter has investigated the situation where $\Psi = 0$. It would be interesting to see for which other values of Ψ the above conclusions hold.

2.6 Conclusions

In this model, inequality leads to discrimination because when people are more different, the signals they provide are more informative. This discrimination feeds back into inequality. This feedback mechanism may lead to multiplicity of steady states, one with a relatively small amount of discrimination and inequality and one with a relatively large amount of discrimination and ever increasing inequality.

There exists a negative correlation between inequality and mobility, which was observed in Corak's "Great Gatsby Curve". We believe this may occur for two reasons: the increased discrimination which accompanies greater inequality results in firms being better able to identify talent, which is correlated with parental income; and it may result in them using family background information more. This move towards greater inequality and lower mobility may be started by an exogenous *increase* in the degree of meritocracy in society, which improves the quality of information available to firms, feeding into discrimination and inequality.

This model can also be used to investigate the effects of a particular policy: early childhood intervention. We could model the early intervention as a way of relieving the credit constraints on poorer families through an extension of public education. This would weaken the relationship between talent development and parental income, which in this model refers to a fall in α . This pre-school investment would lead to a weakening

of the relationship between talent and parental income, making the parental income signal less useful, lowering discrimination and in turn lowering inequality and raising mobility.

The intention here is not to advocate such a policy or to presuppose the outcome of a welfare analysis, but instead to consider the effects of this intervention. Early childhood interventions have been considered in the literature, particularly by James Heckman and co-authors (such as Cunha & Heckman (2007) and Heckman & Masterov (2007)), who concluded that complementarity and self-productivity of skills developed at different points during childhood make early interventions effective at reducing inequality. We reach the same conclusion, but inequality is reduced by early interventions limiting the availability of information upon which discrimination is based. It is a very different mechanism. The following chapter will consider the welfare implications of a similar model, but it is worth noting what this model implies about reductions in α : by weakening the link between parental income and talent development, it increases mobility and reduces income inequality.

Chapter 3

Inheritance and Investment

In chapter 2 we considered the effects of statistical discrimination in labour markets. Firms had a limited information set available to them, and they pieced together a belief about an individual's talent by weighting the information that they had according to its precision. This statistical discrimination fed through to income inequality, but also fed off past inequality. In order to keep that model simple, and to show in a stark way how that feedback mechanism worked, it was limited to two key variables, income and talent, and three key concepts, inequality, intergenerational mobility, and meritocracy. How these three aggregate measures of the state of our economy interacted was the focus of that chapter.

The aim of this chapter is to build on that framework a mechanism which permits forward-looking behaviour. If we look through the more recent statistical discrimination literature we see that it is this forward-looking mechanism which determines equilibrium behaviour. For example, consider the discriminating equilibrium in Coate & Loury (1993). It is the belief of black workers that firms will use their race to sort them that prevents them from investing in the skills that would prevent firms from doing so. Up until now we have not considered the forward-looking investment behaviour of optimising agents. This will be the focus of chapter 3. We will build on top of the model from chapter 2, with its potential for multiplicity of steady states, the Coate

and Loury mechanism whereby agents' expectations of the return to investment in skills affect their decision over investment. While this creates the potential for multiple self-fulfilling equilibria, we will see that in this model, that is not observed.

In order to build in investment decisions, I am going to alter the model in one key way. The distribution of income is going to be modelled as lognormal rather than normal. Parents will be the optimising agents, each choosing how much to invest in their child's talent development (education). They are going to be credit constrained. As a result their investment will be a function of their income, and in order to keep investments confined to the set of non-negative real numbers we want incomes to be non-negative. Hence we use the lognormal income distribution. It has the additional feature of adding more realism to the model. Section 3.1 will derive the optimal investment rule for parents. It is to invest a fraction of their income. The rule will be discussed in section 3.2.

This version of the model will also allow for growth in mean income. The model in the second chapter was primarily concerned with the variance and correlation of incomes and so we abstracted from changes in mean income. Again, this will add some additional realism to the model.

This chapter will proceed in the following way: first, we will derive an optimal investment rule which governs parents' investment in their children's education; second we will show that this is unique and why this uniqueness comes about; finally, given the investment rule, and the income process which it implies, we will look at some features of the economy and how they change following an exogenous improvement in meritocracy or inherited advantages. A key finding is that, for certain sets of parameters, a risk averse agent may have lower utility on average in a country with higher mean income and greater meritocracy, if that meritocracy brings with it greater variance, and hence uncertainty, over income.

3.1 An optimal investment rule for children's education

We will begin with an equation describing the formation of talent. In this case we are going to consider the log of talent, where talent is lognormally distributed. As in the previous chapters, an individual's level of talent describes their skill level after education and before entering the labour force, and so is dependent on their human capital formation as a child. Our assumption is that this depends on the income level of their parent when they are growing up and the level of investment which their parent makes in them (we assume that each parent has one child). There is an underlying credit constraint which states that it is impossible to borrow in order to invest in your child. Hence, coming from a rich background is beneficial not only through the environment of your childhood but because it better enables your parent to invest in your development. The equation for the log of talent is given by,

$$\ln\tau = \phi + \alpha_y \ln y_{-1} + \alpha_x \ln x + \epsilon_\tau \quad (3.1)$$

where τ is an individual's talent, y_{-1} is his parent's income level, x is the investment which his parent chose to make in him, and ϵ_τ is a mean zero normally distributed error with variance $\sigma_{\epsilon_\tau}^2$. ϕ is a constant term, while α_x and α_y are the parameters which govern the advantages which come from being part of a rich family. In sections 3.4 and 3.5 we will refer to an economy with "low" values of α_x and α_y ($\alpha_x + \alpha_y < 1$) as a weak inheritance economy and one with "high" values ($\alpha_x + \alpha_y \geq 1$) as a strong inheritance economy.

We will call $\hat{x}(y_{-1})$ the equilibrium investment rule but we will not assume any structure on it. Instead we will allow for any investment rule and solve the parent's investment choice problem to find the structure that it has in equilibrium. The equilibrium investment rule is known to all.

3.1.1 The pay decisions of firms

As outlined in the introduction, we are going to assume in this chapter that income follows a log-normal distribution,

$$\ln y_{-1} \sim N\left(\mu_{y_{-1}}, \sigma_{y_{-1}}^2\right)$$

If there were no additional information available to the firm, they would believe everyone to be of average log income and the precision in their beliefs would be $1/\sigma_{y_{-1}}^2$.

Suppose there are two signals available to the firm which provide some information on a worker's talent. The first is a noisy signal of her parent's income given by $s_1 = y_{-1}e^{\epsilon_{s_1}}$ where ϵ_{s_1} is a mean zero normally distributed error with variance $\sigma_{\epsilon_{s_1}}^2$. Neither this signal nor the prior provide the firm with any information about the error in talent, ϵ_τ but, as in the chapter 2 model, they do allow the firm to form a belief about parental income which is itself a signal of talent. Firms, as before, form a posterior belief about log parental income conditional on the signal s_1 . The distribution of these posterior beliefs is given by

$$\ln y_{-1}|s_1 \sim N\left(\frac{\sigma_{\epsilon_{s_1}}^2}{\sigma_{y_{-1}}^2 + \sigma_{\epsilon_{s_1}}^2}\mu_{y_{-1}} + \frac{\sigma_{y_{-1}}^2}{\sigma_{y_{-1}}^2 + \sigma_{\epsilon_{s_1}}^2}\ln s_1, \frac{\sigma_{y_{-1}}^2\sigma_{\epsilon_{s_1}}^2}{\sigma_{y_{-1}}^2 + \sigma_{\epsilon_{s_1}}^2}\right)$$

We will refer to the mean of this distribution as $\mu_{y_{-1}|s_1}$. It is the firm's belief about the parental income of an individual with signal s_1 . We will refer to the variance as $\sigma_{y_{-1}|s_1}^2$. The precision in the firm's belief is therefore $1/\sigma_{y_{-1}|s_1}^2$.

Given the human capital development function in (3.1), firms can form a belief about talent based on their posterior beliefs about parental income. They must consider how much parental investment the individual received and in doing so take the

equilibrium investment rule as given. Let $\mu_{x|\hat{x}}$ be the firms' belief about an individual's $\ln x$, given the investment rule, and $\sigma_{x|\hat{x}}^2$ be the variance in those beliefs. Although the equilibrium investment rule will be known to the firm, parental income is only known with some error which leads to the variance in their beliefs about investment. $\mu_{x|\hat{x}}$ and $\sigma_{x|\hat{x}}^2$ are known to all.

This gives the following conditional distribution of talent,

$$\ln \tau | s_1 \sim N \left(\phi + \alpha_y \mu_{y-1|s_1} + \alpha_x \mu_{x|\hat{x}} + \alpha_y^2 \sigma_{y-1|s_1}^2 + \alpha_x^2 \sigma_{x|\hat{x}}^2 + 2\alpha_y \alpha_x \sigma_{y-1|s_1, x|\hat{x}} + \sigma_{\epsilon\tau}^2 \right) \quad (3.2)$$

Here $\sigma_{y-1|s_1, x|\hat{x}}$ is the covariance of the firm's belief about a parent's income, conditional on the parental income signal, and their belief about the investment that the parent has made, given the equilibrium investment rule. There is a statistical advantage to looking like you are from a privileged background, both directly and through the investment you will be believed to have received.

The second signal is a noisy signal on the individual's talent given by $s_2 = \tau e^{\epsilon_{s_2}}$ where ϵ_{s_2} is a mean zero normally distributed error with variance $\sigma_{\epsilon_{s_2}}^2$. Each firm is now in a position to choose what weights to give to their beliefs based on the the information they have about an individual's parental income (μ_{y-1}, s_1) and the information they have about her talent (s_2). This gives the following result:

$$\ln \tau | s_1, s_2 \sim N \left([1 - \beta_2] \left[\phi + \alpha_y \mu_{y-1|s_1} + \alpha_x \mu_{x|\hat{x}} \right] + \beta_2 \ln s_2, \beta_2 \sigma_{\epsilon_{s_2}}^2 \right) \quad (3.3)$$

where

$$\beta_2 = \frac{\alpha_y^2 \sigma_{y-1|s_1}^2 + \alpha_x^2 \sigma_{x|\hat{x}}^2 + 2\alpha_y \alpha_x \sigma_{y-1|s_1, x|\hat{x}} + \sigma_{\epsilon\tau}^2}{\alpha_y^2 \sigma_{y-1|s_1}^2 + \alpha_x^2 \sigma_{x|\hat{x}}^2 + 2\alpha_y \alpha_x \sigma_{y-1|s_1, x|\hat{x}} + \sigma_{\epsilon\tau}^2 + \sigma_{\epsilon_{s_2}}^2}$$

The firms' posterior belief about talent is a weighted average of their beliefs conditional on the first and second signals. The weight is given by β_2 . As in the previous chapter, β_2 will measure the extent to which firms reacts to changes in the talent signal. Since in equilibrium the distribution of the firm's beliefs about investment will be common knowledge, so will be the constant β_2 .

We are now in a position to consider a payment rule for the firm. As in chapter 2, the firm will pay a wage based on their expectations of the talent of an individual. For any lognormally distributed variable z with $E(\ln z) = \mu_z$ and $Var(\ln z) = \sigma_z^2$, the expected value of z is equal to $e^{\mu_z + \frac{\sigma_z^2}{2}}$. It follows that the conditional expectation of talent is,

$$E(\tau|s_1, s_2) = e^{[1-\beta_2][\phi + \alpha_y \mu_{y-1|s_1} + \alpha_x \mu_{x|\hat{x}}] + \beta_2 \ln s_2 + \frac{\beta_2 \sigma_{\epsilon s_2}^2}{2}}$$

And so, if firms pay incomes equal to their conditional expectation of talent, $\ln y$ is

$$\ln y = [1 - \beta_2] [\phi + \alpha_y \mu_{y-1|s_1} + \alpha_x \mu_{x|\hat{x}}] + \beta_2 \ln s_2 + \frac{\beta_2 \sigma_{\epsilon s_2}^2}{2} \quad (3.4)$$

This is the log income which an individual will receive. She will receive a higher income if the firm believes her parent to have been richer, both because of the direct effect of parental income on human capital development and because, given the equilibrium investment rule, richer parents may be believed to invest more. She will also receive a higher income if she looks more talented via the direct signal on talent, s_2 . We will now consider how parents invest and hence what the equilibrium investment rule is.

3.1.2 The investment decisions of parents

Parents choose how much to invest in their child. In order to know how they do that, we must define for a parent a particular value function. This is given by,

$$V_{-1}(y_{-1}) = \max_{c_{-1}} \left\{ \ln c_{-1} + \frac{1}{1+\delta} E[V(y) | y_{-1}, x] \right\} \quad (3.5)$$

subject to

$$y_{-1} \geq c_{-1} + x$$

Parents gain utility from their own consumption and from the expected value which their child will have, and make their consumption-investment choice optimally given their income. δ is the discount rate. Parents know their own income and investment choice and so condition on them when taking expectations of their child's value function. We also include a credit constraint whereby parents are unable to borrow. They can consume themselves or invest in their children, but solely through educational investment. Since they are maximising utility we assume that this constraint is binding.

We proceed with a guess and verify strategy. We guess that the value function is of the following form,

$$V(y) = A \ln y + B$$

Conditional on this proving to be the case, we can find a parent's expectation of their child's value function. Performing the maximisation in equation 3.5 then allows the Euler equation to be found.¹

¹I will presume here the independence of A from $\ln y$ and both A and B from x . The investment choice of parents affects the expected utility of his child only through raising his expected income and

$$\frac{1}{y_{-1} - x} = \frac{E[A]}{1 + \delta} \frac{\partial E[\ln y | y_{-1}, x]}{\partial x}$$

$E[\cdot]$ is the expectations operator. The last term describes the expected return on the investment of the parent. Firms' take as given that parent's invest according to the equilibrium investment rule. There is no way for a parent to signal if they have deviated from it and, by definition, no reason to do so without such a signal (when firms take the rule as given, following the rule solves the parent's maximisation in equation 3.5). Hence the marginal effect of investment in a child is independent of the beliefs that firms have on marginal investment. The only effect which investment has on a child's income is through their talent development which they can signal through s_2 . We can see this by taking expectations of the income process described in equation 3.4.

$$E[\ln y | y_{-1}, x] = [1 - \beta_2] \left[\phi + \frac{\alpha_y}{\sigma_{y_{-1}}^2 + \sigma_{\epsilon s_1}^2} \left[\sigma_{\epsilon s_1}^2 \mu_{y_{-1}} + \sigma_{y_{-1}}^2 \ln y_{-1} \right] + \alpha_x E(\mu_{x|\hat{x}} | y_{-1}) \right] \\ + \beta_2 \left[\phi + \alpha_y \ln y_{-1} + \alpha_x \ln x \right] + \frac{\beta_2 \sigma_{\epsilon s_2}^2}{2}$$

Examining the above expectation of log income, the only place where the chosen investment level plays a role is in the penultimate term. This term is the expected value of $\ln s_2$. A parent expects a marginal increase in investment to raise their child's income only to the extent that they expect it to make them appear more talented. They do not expect it to change the firms' belief about how much they have invested. Firms take the equilibrium investment rule as given and parents have no way of making themselves look richer.

Solving the Euler equation provides the optimal investment choice of a parent. This is the equilibrium investment rule.

not through raising A or B . This will be verified below.

$$\hat{x}(y_{-1}) = \frac{\alpha_x \beta_2 E[A]}{(1 + \delta) + \alpha_x \beta_2 E[A]} y_{-1} \quad (3.6)$$

An equilibrium investment rule is one for which the following statement holds: given that firms expect individuals to invest according to the equilibrium investment rule, the optimal choice of investment is in accordance with the equilibrium investment rule. We have shown so far that given *any* investment rule which firms believe individuals are following, the optimal investment for a parent is a constant fraction of their income. It then follows that investing a constant fraction of income is an equilibrium investment rule: given that firms believe individuals to be investing a fraction λ of their income in their children, the optimal investment for parents to make is a fraction λ of their income.

Furthermore, since investing a fraction λ of their income is the optimal investment strategy irrespective of the rule firms believe they are following, this is a unique investment rule. There is no other rule which firms would ever believe parents to be following, nor is there any other rule which parents would ever choose to follow.

3.1.3 Solving the equilibrium payment and investment rules

We can see that, if the guess about the value function is correct, parents spend a fraction of their income investing in their child. We are calling this fraction λ . Based on this result,

$$\mu_{x|\hat{x}} = \ln \lambda + \mu_{y_{-1}|s_1}$$

and

$$\sigma_{x|\hat{x}}^2 = \sigma_{y_{-1}|s_1, x|\hat{x}} = \sigma_{y_{-1}|s_1}^2$$

The firms belief about how much you are investing in your child is equal to a constant plus their belief about your income, conditional on the signal it receives about your income. By substitution back into equation 3.4 we find that,

$$\ln y = \mu_\tau + \beta_1 (\ln s_1 - \mu_{s_1}) + \beta_2 (\ln s_2 - \mu_{s_2}) + \frac{\beta_2 \sigma_{\epsilon s_2}^2}{2} \quad (3.7)$$

where,

$$\beta_1 = \frac{(\alpha_x + \alpha_y) \sigma_{y-1}^2 \sigma_{\epsilon s_2}^2}{(\alpha_x + \alpha_y)^2 \sigma_{y-1}^2 \sigma_{\epsilon s_1}^2 + (\sigma_{\epsilon \tau}^2 + \sigma_{\epsilon s_2}^2) (\sigma_{y-1}^2 + \sigma_{\epsilon s_1}^2)} \quad (3.8)$$

and,

$$\beta_2 = \frac{(\alpha_x + \alpha_y)^2 \sigma_{y-1}^2 \sigma_{\epsilon s_1}^2 + \sigma_{\epsilon \tau}^2 (\sigma_{y-1}^2 + \sigma_{\epsilon s_1}^2)}{(\alpha_x + \alpha_y)^2 \sigma_{y-1}^2 \sigma_{\epsilon s_1}^2 + (\sigma_{\epsilon \tau}^2 + \sigma_{\epsilon s_2}^2) (\sigma_{y-1}^2 + \sigma_{\epsilon s_1}^2)} \quad (3.9)$$

This defines the optimal payment rule for firms. The mean of $\ln y$ is given by the mean (log) talent, adjusted by the variance in the conditional distribution of τ . Workers will be believed to be worth average (log) income unless they have above or below average signals. An above or below average signal will cause the firm to adjust their beliefs to the extent that the signal is informative. This is captured by the precision of the signals through β_1 and β_2 .

For example, as $\sigma_{\epsilon s_2}^2$ tends to infinity, the second signal is of no use to the firms. They disregard it. Alternatively, if $\sigma_{\epsilon s_2}^2$ were to be zero, the second signal would be perfect and the firms would completely ignore the first signal. In an equivalent fashion to chapter 2, β_1 and β_2 are the coefficients from a multivariate regression of $\ln \tau$ against $\ln s_1$ and $\ln s_2$, the only difference being the use of the log terms. Substituting for $\ln s_1$ and $\ln s_2$ we find the income process.

$$\ln y = \left\{ \mu_\tau + \frac{\beta_2 \sigma_{\epsilon_{s_2}}^2}{2} \right\} + [\beta_1 + (\alpha_x + \alpha_y) \beta_2] (\ln y_{-1} - \mu_{y_{-1}}) + \beta_1 \epsilon_{s_1} + \beta_2 (\epsilon_{s_2} + \epsilon_\tau) \quad (3.10)$$

This is the income that an individual will earn, given his parents income, and the realisation of the shocks to his talent and signals. Inherent in this is that parents invest according to the equilibrium investment rule. The last thing that we need to do then is check that our value function guess was correct and find the equilibrium investment rule.

From equation 3.5 we can see that our value function, given the investment rule, will be of the form,

$$V_{-1}(y_{-1}) = \ln(1 - \lambda) + \ln y_{-1} + \frac{E[A] [\mu_y + [\beta_1 + (\alpha_x + \alpha_y) \beta_2] (\ln y_{-1} - \mu_{y_{-1}})] + E[B]}{1 + \delta}$$

This is of the desired form, $V_{-1}(y_{-1}) = A_{-1} \ln y_{-1} + B_{-1}$, and so verifies that our guess was correct where,

$$A_{-1} = 1 + \frac{1}{1 + \delta} E[A] [\beta_1 + (\alpha_x + \alpha_y) \beta_2]$$

and,

$$B_{-1} = \ln(1 - \lambda) + \frac{E[A]}{1 + \delta} \left\{ \phi + \alpha_x \ln \lambda + \frac{\beta_2 \sigma_{\epsilon_{s_2}}^2}{2} - [\beta_1 - (1 - \beta_2) (\alpha_y + \alpha_x)] \mu_{y_{-1}} \right\} + \frac{1}{1 + \delta} E[B]$$

$E[A]$ and $E[B]$ are the parent's expectation of the parameters of the child's value function. Since the parent cares about the child's utility, and makes an investment choice which can potentially raise or lower the child's income, they must form an expectation of how the child's income impacts on his utility. For example, in the equation for A , the value of an increase in income to the parent is greater to the extent that he can pass it on to his child ($\beta_1 + (\alpha_x + \alpha_y)\beta_2$) and to the extent that he expects doing so to raise his child's utility ($E[A]$). Appendix A shows that $E[A]$ and $E[B]$ are finite valued and how they depend on future discrimination. Since β_1 and β_2 follow deterministic paths, the appendix shows that A and B are actually known with certainty and what those values are equal to.

The equilibrium investment rule is given by equation 3.6. By substituting for A from equation A.1 in Appendix A, we find the investment rule as a function of the parameters of the model and future levels of discrimination.

$$\hat{x}(y_{-1}) = \frac{\alpha_x \beta_2 \left[1 + \sum_{i=1}^{\infty} \frac{\prod_{j=1}^i [\beta_{1,+j} + (\alpha_x + \alpha_y)\beta_{2,+j}]}{(1+\delta)^i} \right]}{(1+\delta) + \alpha_x \beta_2 \left[1 + \sum_{i=1}^{\infty} \frac{\prod_{j=1}^i [\beta_{1,+j} + (\alpha_x + \alpha_y)\beta_{2,+j}]}{(1+\delta)^i} \right]} y_{-1} \quad (3.11)$$

The investment rule depends on β 's in two ways. The term $\alpha_x \beta_2$ governs the channel through which you can influence the income of your child by raising their talent. This is the only way that you can influence the income of your child. In addition, the β terms for every generation from your grandchildren onwards are included, appropriately discounted. They matter because investment in your child increases the opportunity for them to provide advantages to future generations, but only if future generations of firms discriminate based in those advantages. Note that advantages created for generations from your grandchildren onwards can come through raising their level of talent, or through raising the income of their parents (something you cannot do for your own children).

By substitution from the steady state investment rule given in equation A.2 we get

the steady state investment rule.

$$\hat{x}^*(y_{-1}) = \frac{\alpha_x \beta_2^*}{1 + \delta - (\beta_1^* + \alpha_y \beta_2^*)} y_{-1} \quad (3.12)$$

In order for the model to be a reasonable simplification of how individuals act, we must constrain $[\beta_1^* + (\alpha_x + \alpha_y) \beta_2^*]$ to be less than $(1 + \delta)$. This ensures that: an individual's income has a positive, finite effect on his value function in steady state; and there is an interior solution to the consumption-investment problem whereby the fraction of income invested (or consumed) is between 0 and 1. We can see, to some extent, what effect this constraint will have. δ measures the extent to which parents care for their children's expected wellbeing versus their own, α_y and α_x measure the returns to parental income and investment in terms of developing talent, and β_1^* and β_2^* the (endogenous) returns to the parental income and talent signals in terms of income. In order to ensure an interior solution to the parent's consumption-investment problem we constrain their willingness to invest in their child or the advantages of coming from a rich background to be below some upper bound.²

Since we have not established how β_1 and β_2 change with the various parameters of the model, we will investigate how changes in the parameters feed through to the investment rule in the discussion section of this chapter. The parameter δ plays no role in β_1^* and β_2^* so we can say at this stage that increases in the discount rate lead to a smaller share of parental income being invested in steady state. This is very intuitive.

3.2 The investment rule

The equilibrium investment rule is one of the novel features of this chapter. In chapter 2 we considered the effects of marginal decreases in $\sigma_{\epsilon s_1}^2$ and $\sigma_{\epsilon s_2}^2$, exogenously increasing

²The largest value that $\beta_1 + (\alpha_x + \alpha_y) \beta_2$ can take is $\alpha_x + \alpha_y$. This occurs as σ_y^2 grows towards infinity. $\alpha_x + \alpha_y < 1 + \delta$ is a sufficient condition to ensure that $A > 0$ and $0 < x < 1$.

the precision of signals 1 and 2 respectively. The result which we found was that, although β_1 and β_2 did not necessarily shift in the same direction for a given change, the overall effect of an increase in the precision of either signal was an increase in inequality and a fall in mobility. We are now going to consider the effects of these same changes on the payment rule of firms in more detail.

Firms' response to changes in the precision of our signals can be considered in much the same way as consumers respond to a price fall in standard consumer choice theory. According to consumer choice theory with two normal goods, a fall in the price of one good leads to both a substitution effect and an income effect. The good which experiences the price fall unequivocally experiences increased consumption, but there are two offsetting effects on the consumption of the other good: it is now relatively more expensive but the consumer effectively has higher income.

The actions of firms in our model are analogous to that of a consumer in consumer choice theory. The direct effect of an increase in the precision of one signal is a shift toward using that signal at the expense of the other one. This can be seen from equations 3.8 and 3.9 where $\sigma_{y_{-1}}^2$ is fixed. This is our "substitution" effect. We know from chapter 2, however, that the long run effects of such a change are that the economy converges to a higher level of steady state income variance (or in this chapter, variance in log income), endogenously increasing the precision of both signals, and increasing the firms' use of both. This is our "income" effect. The net result is that the signal for which precision exogenously increased should unambiguously be used more by the firm in their payment rule, but the other signal may experience increased or decreased use.

β_1 and β_2 both play important roles in the investment decision of parents. First consider equation 3.6. This illustrates the "direct" role of β_2 . If the firm uses the talent signal more, there is an improved channel through which to influence the income of your child. A parent takes advantage of that channel through investment. There is a further effect of future values of β_1 and β_2 through A , the marginal value of log income. A is increasing in the use of the signals. An additional pound of income is more valuable, if it cannot only be consumed but invested effectively in future generations to raise their

utility. The more effective is this second channel, through greater discrimination by firms in the future, allowing them to identify the effects of increased investment, the more that channel will be used.

To see how this plays out in the model let us take an example. Suppose that there is an exogenous improvement in the precision of signal one. This must increase the use of the parental income signal in the firms' payment rule but, at least at the time of the change, the talent signal will be used less. A parent may change their investment choice as a result of a number of forces which come into play. Most directly, as the talent signal is used less by firms, a parent's ability to influence their child's income is reduced, reducing their incentive to invest. This also reduces the opportunity to improve the prospects of their grandchildren through increasing the funds available for investment of *their* parents. On the other hand, since firms are using the parental income signal more, there is an incentive to attempt to raise the income of your child since the advantages of a privileged background for your grandchildren, great-grandchildren and so on, are improved. This in turn could alter the degree and nature of future discrimination. Any change in the investment choice thus weighs up the reduced effectiveness of investment in generating talent and income in subsequent generations with the increased ability to pass on advantages through birth if such advantages can be established. It is not clear which of these effects will dominate.

3.3 Steady states of the economy

As in chapter 2, I will be primarily interested in how the economy is influenced by exogenous changes in meritocracy and advantage. I will examine this issue again within the new framework of chapter 3. In order to do so, I will deal with three separate cases of the economy. In the first, there will be "weak" inheritance, by which I mean that $\alpha_x + \alpha_y < 1$. The technology of talent development is such that parent's income and investments play a relatively limited role and so who your parents are should be relatively less important. As in chapter 2, there will be no multiplicity of steady states

in this instance. In the second case, inheritance takes on larger values ($\alpha_x + \alpha_y \geq 1$). This may lead to a multiplicity of steady states, and we will consider first the steady state where the firms give relatively little weight to the two signals. Our third case will examine when firm's give relatively more weight to the two signals. As in chapter 2, when inheritance is sufficiently large this will be the only steady state.

3.3.1 Case 1: The steady state of a weak inheritance economy

As mentioned above, a weak inheritance economy will be defined as one for which $\alpha_x + \alpha_y < 1$. We will first focus on two things: the steady state mean log income level; and the steady state value of the variance in log income. These in turn will allow us to calculate the steady state mean income, variance in income, correlation between parents' and children's income, and expected value or utility.

We will begin with the variance in the log of income. In order to find the steady state value, we begin with the income process described in equation 3.10. The variance of this equation provides a relationship between the variance in log income of adjacent generations, and the condition $\sigma_y^2 = \sigma_{y-1}^2$ allows this to be solved for the steady state variance of log income. I will use a * to denote steady state values.

$$\sigma_y^{2*} = \frac{\beta_1^{*2} \sigma_{\epsilon s_1}^2 + \beta_2^{*2} (\sigma_{\epsilon s_2}^2 + \sigma_{\epsilon \tau}^2)}{1 - (\beta_1^* + (\alpha_x + \alpha_y) \beta_2^*)^2} \quad (3.13)$$

β_1^* and β_2^* are the steady state solutions to equations 3.8 and 3.9. It should be noted that finding a steady state of this model requires a solution to the same three equations in three unknowns as finding a steady state to the model of chapter 2, but with the parameter α replaced with $\alpha_x + \alpha_y$.³ The results of chapter 2 hold for the variance of the log of income and the correlation of the log of income across generations

³Were $\beta_1^* + (\alpha_x + \alpha_y) \beta_2^*$ to be greater than or equal to one $\beta_1^* + (\alpha_x + \alpha_y) \beta_2^*$ would equal $\alpha_x + \alpha_y$. A necessary condition for this is $\alpha_x + \alpha_y \geq 1$ and so we can be sure that our weak inheritance economy converges to a finite value steady state variance in log income with $\beta_1^* + (\alpha_x + \alpha_y) \beta_2^* < 1$.

(intergenerational income elasticity).

The steady state value of the mean of log income is found by taking expectations of the log income equation (3.10) and then finding the value μ_y^* for which $\mu_y = \mu_{y-1}$.

$$\mu_y^* = \frac{1}{1 - (\alpha_x + \alpha_y)} \left\{ \phi + \alpha_x \ln \frac{\alpha_x \beta_2^*}{1 + \delta - (\beta_1^* + \alpha_y \beta_2^*)} + \frac{\beta_2^* \sigma_{\epsilon s_2}^2}{2} \right\} \quad (3.14)$$

Equations 3.13 and 3.14 give us the steady state variance and mean of the log of income. We are not primarily interested in these values. They do however allow us to calculate the mean and variance in *income*. From the definition of the mean of a log-normally distributed variable, we can find the steady state mean level of income.

$$E[y]^* = e^{\mu_y^* + \frac{\sigma_y^{2*}}{2}} \quad (3.15)$$

The mean of income is increasing in both the mean and variance of log income. For $\alpha_x + \alpha_y$ less than one, both of these have a constant finite steady state value and so steady state mean income has a constant finite value. Steady state mean investment is given by $\lambda^* E[y]^*$ which is the expected value of equation 3.12.

We can use the mean and variance of log income in a similar way to find the variance of income. Again, appealing to the fact that income is a log-normally distributed variable we find,

$$Var[y]^* = \left(e^{\sigma_y^{2*}} - 1 \right) e^{2\mu_y^* + \sigma_y^{2*}} \quad (3.16)$$

It is also increasing in both the mean and variance of log income and, for $\alpha_x + \alpha_y$

less than one, will have a constant finite value.

In order to find the steady state value of the correlation of income across generations we need to know not just the mean and variance of log income, but the covariance of the log of a parent's and child's income. We can see from equation 3.10 that this is given by $(\beta_1^* + (\alpha_x + \alpha_y)\beta_2^*)\sigma_y^{2*}$.

$$\rho_{y,y-1}^* = \frac{e^{(\beta_1^* + (\alpha_x + \alpha_y)\beta_2^*)\sigma_y^{2*}} - 1}{e^{\sigma_y^{2*}} - 1} \quad (3.17)$$

Finally, we are interested in the mean level of utility or value in the economy in steady state. This is the utility that an individual could expect to receive from behind the veil of ignorance.

$$E[V(y)]^* = A^*\mu_y^* + B^* \quad (3.18)$$

Prior to knowing who their parents will be, an individual's expected utility will be increasing in mean log income, A^* and B^* . Both A^* and B^* are endogenous and given by equations A.2 and A.4 in Appendix A.

Equations 3.15 through 3.18 define the summary statistics of our steady state economy when $\alpha_x + \alpha_y < 1$. Section 3.2 provides some intuition for how the steady state values of β_1 and β_2 will change for an exogenous change in advantage or meritocracy, but we will again leave it to a numerical exercise to investigate these issues further. We will be particularly interested to see if there are cases when mean income increases but expected utility decreases. This could be possible if an exogenous shock caused income variance to increase and hence the risk which a risk averse agent would face from behind the veil of ignorance would be greater. It is possible that given the choice of being born into a country with relatively high or low mean income, an individual may favour the

poorer country because income variance, and hence risk, would be reduced.

3.3.2 Case 2: The low steady state in a strong inheritance economy

If $\alpha_x + \alpha_y \geq 1$ there are two possible steady states which could emerge. In the “low” steady state, $\beta_1 + (\alpha_x + \alpha_y)\beta_2 < 1$. In this case the mean of log income will grow over time while the variance of log income will tend towards a constant finite value.

Let us begin with the variance. From equation 3.10 we can see that the variance in log income is given by,

$$\sigma_y^2 = (\beta_1 + (\alpha_x + \alpha_y)\beta_2)^2 \sigma_{y-1}^2 + \beta_1^2 \sigma_{\epsilon_{s1}}^2 + \beta_2^2 (\sigma_{\epsilon_{s2}}^2 + \sigma_{\epsilon_{\tau}}^2) \quad (3.19)$$

Since $\beta_1 + (\alpha_x + \alpha_y)\beta_2 < 1$ this will tend towards the steady state value given by equation 3.13. In steady state, it will have a constant finite value. We will see, however, that this is not the case for the mean of log income.

The mean of log talent can be found by taking expectations of equation 3.1. Mean log income is then found by taking expectations of the income process in equation 3.10. Doing so, it is immediately apparent that $\alpha_x + \alpha_y$ plays a key role.

$$\mu_y = \phi + \alpha_x \ln \frac{\alpha_x \beta_2}{1 + \delta - (\beta_1 + \alpha_y \beta_2)} + (\alpha_x + \alpha_y) \mu_{y-1} + \frac{\beta_2 \sigma_{\epsilon_{s2}}^2}{2} \quad (3.20)$$

With $\alpha_x + \alpha_y > 1$, the value of μ_y will grow forever. There is steady state growth which will tend towards a constant rate of $\alpha_x + \alpha_y - 1$. Although the variance in log income does reach a constant finite value, and so log income still has a non-degenerate distribution, we can see from equations 3.15 and 3.16 how an ever increasing value of μ_y will affect the mean and variance of income in steady state. As both are increasing

functions of mean log income, they too will increase over time.

In order to examine the effects of an exogenous change in meritocracy or advantage in an environment where the income distribution is changing we will follow a slightly different method to that of case 1. Instead of examining the effects of a shock on our steady state, we will examine the effects of a shock given the current state of our economy. The state variables for case 2 will be μ_{y-1} and σ_{y-1}^2 and these will become additional parameters of our numerical exercise.

One interesting aspect of this change in method is that, since σ_{y-1}^2 is fixed, we will be looking purely at the “substitution effect” of a change in meritocracy or advantage. As one signal is used more, the other will be used less. The effects of an exogenous shock to meritocracy and advantage will be examined for the generation in which the shock occurs. Our economy will be summarised by the same statistics as in case 1: mean income; income variance; the intergenerational correlation of incomes; and expected utility. Other than the fact that we are no longer examining steady state values, the equations for these will be the same as in section 3.3.1 for all but the intergenerational correlation. It will be given by,

$$\rho_{y,y-1} = \frac{e^{(\beta_1 + (\alpha_x + \alpha_y)\beta_2)\sigma_{y-1}^2} - 1}{\sqrt{[e^{\sigma_y^2} - 1][e^{\sigma_{y-1}^2} - 1]}}$$

where σ_y^2 is given in equation 3.19. Mean investment will be given by $\lambda E[y]$ which is the expected value of equation 3.11.

3.3.3 Case 3: The high steady state in a strong inheritance economy

If both $\alpha_x + \alpha_y \geq 1$ and $\beta_1 + (\alpha_x + \alpha_y)\beta_2 \geq 1$, it is clear from equations 3.19 and 3.20 that both the variance and mean of log income will be increasing over time. It is equally evident from equations 3.15 and 3.16 that this will mean that the steady state

Parameter	Values the parameter may take
α_y	0.05, 0.15, ..., 0.75
α_x	0.05, 0.15, ..., 0.75
$\sigma_{\epsilon s_1}^2$	2, 4, ..., 18
$\sigma_{\epsilon s_2}^2$	2, 4, ..., 18
$\sigma_{\epsilon \tau}^2$	1, 2, ..., 7
ϕ	5
δ	0.5

Table 3.1: Parameter values for the numerical exercise in the weak inheritance economy

mean and variance of income will also be increasing. This being the case, the same approach will be followed as in section 3.3.2. For different values of the state variables, a numerical exercise will examine the effects of an exogenous shock to advantage or meritocracy on our economy at the time of the change.

3.4 The weak inheritance economy

As noted in the previous section, when $\alpha_x + \alpha_y < 1$ the mean and variance of the log of income tend towards constant finite values in the long run. This in turn implies constant finite steady state values of all the key statistics which summarise the state of our economy. In what follows, we will examine how shocks to meritocracy and advantage translate into changes in these steady state values. Meritocracy is defined, as in chapter 2, as the precision in the talent signal, while advantage is the precision in the parental income signal. The numerical exercise below considers how marginal increases in both affect the steady state economy given certain values of the parameters. The parameters for this numerical exercise are given in table 3.1. There are 24,381 combinations of the parameters for which $\alpha_x + \alpha_y < 1$.

The parameters ϕ and δ have been fixed since their role in the determination of investment and mean income is known. They do not affect β_1 and β_2 and so an increase in δ , the discount rate, can be seen from equation 3.11 to decrease investment and from equations 3.14 and 3.15 to lead to a decrease in mean income. When parents are investing less in developing talent, the mean income of the economy falls. In addition,

Statistic	Number of cases in which it increased	Percentage
$E[y]^*$	14,515	59.5
$Var[y]^*$	24,381	100.0
$\rho_{y,y-1}^*$	13,058	53.6
$E[V(y)]^*$	8,209	33.7
$E[x]^*$	21,199	86.6
λ^*	24,381	100.0

Table 3.2: The number and proportion of cases where the steady state values increased with an exogenous increase in meritocracy in the weak inheritance economy. Total number of cases = 24,381

since steady state A,B and mean income are all decreasing in δ , it must decrease expected value. ϕ has no effect on investment but, since it is the constant term in the talent process, increases mean income.

3.4.1 The effects of an improvement in meritocracy in a weak inheritance economy

For this numerical exercise, the steady state was found for each of our 24,381 parameter combinations, then the value of $\sigma_{\epsilon s_2}^2$ was lowered by 0.001, simulating an exogenous marginal improvement in the precision of the talent signal. The proportion of cases for which our key statistics increased is given in table 3.2.

There are some interesting findings in the table. Improvements in the information on talent do not universally lead to increases in mean income, and only lead to expected utility improvements in around one-third of cases. In the majority of cases, prior to discovering who their parents are, people would choose to be born in the economy with less precise information. Improvements in meritocracy also do not necessarily lead to increased investment. The rest of this section will discuss some of these findings.

The fall in mean income

The fall in mean income may seem surprising because we know that an improvement in meritocracy will cause the firm to place greater weight on the talent signal. This

tends to increase the share of income that parents invest, which has a positive effect on average income and average talent. What is more, in over half the cases where mean income is falling, firms' use of the parental income signal is increased in steady state, providing further incentive to invest since improvements in children's income will more easily be passed on to subsequent generations.

The circumstances under which these do not lead to an increase in mean income can be seen in equation 3.14. When α_x is small, the effect of investment on mean log income is weakened. When investment does not translate into increased talent, there is no way for parents to increase the income or expected utility of their children and so they do not invest, even if β_2 is high. The driving force of falls in mean income tends to be the last term in equation 3.14. This comes from the variance in the distribution of log talent, conditional on the signals. Additional information provided by the talent signal causes this to fall, and all the more so for large values of β_2 .

An example of an economy likely to exhibit falling steady state mean income would be one with a high value of $\sigma_{\epsilon\tau}^2$ but a low value of α_x . The high value of $\sigma_{\epsilon\tau}^2$ leads to a large amount of variance in talent, and a high value of β_2 . However, the low value of α_x means that parents do not invest to any great extent, despite the high value of β_2 , and do not respond to increases in β_2 with increased investment. This is an economy characterised by low investment, talent distributed to a large extent by luck, and firms using the talent signal a great deal.⁴

It is also the case that mean income may be falling despite increases in mean investment. This happens in close to 6,700 cases. It must be the case in these situations that the share of income being invested in children is increasing but it is not translating to any great extent into improvements in income. Since parents know that this will happen, it is likely that the increases in investment are small. The conditions for which this occurs are the same as for a fall in income in general. A small value of α_x prevents the increased investment share from impacting on income while a large β_2 , induced by

⁴To provide some evidence of this, there are nearly 3,000 cases of falling mean income when $\sigma_{\epsilon\tau}^2$ is 7 compared to only 19 when it is 1, and over 3,000 cases when α_x is 0.05 compared to 217 when it is 0.75.

a large exogenous variance in talent, causes the falling $\sigma_{\epsilon s_2}^2$ to pull down mean income.

A fall in mean investment itself may seem odd when faced with an environment where talent is more easily displayed. This fall in investment is always associated with a fall in mean income, and so it is the fact that mean income can be lower in more meritocratic societies which may push down mean investment. In fact it could be a concern with the model that the constant coefficient of relative risk aversion of one is driving the slight ambiguity over the direction of change in mean investment. It is well known that in a standard intertemporal consumption and savings problem the income and substitution effects cancel out for the log utility function. Agents save neither more nor less in response to a change in the interest rate. This is not what is happening here in response to the change in the level of discrimination. In every case λ^* increases. Parents unambiguously invest a larger fraction of their income in steady state in more meritocratic societies. Falls in steady state mean investment are always driven by falls in mean income and not by agents investing a smaller fraction of their income.

The fall in expected value

The fact that individuals may choose to live in a country where their talent is more difficult to display may also seem somewhat surprising. This is especially so since in over 6,000 cases where expected value is falling, mean income is increasing. From behind the veil of ignorance, people would be happier choosing to be born prior to the improvement in meritocracy into the poorer economy. Why this might occur has to do with the increasing variance in income. The value function displays decreasing marginal utility from income and hence risk or inequality aversion. The increase in mean income has to offset the increased variance in income for an individual to be better off.

The reasons why this might not occur again boil down to the parameter values which tend to accompany it. Once more, the inheritance parameters are relatively low and the exogenous variance in talent high. If these conditions do not cause mean income to fall, they are likely to cause mean income gains to be small. That, coupled with

the difficulty in passing those gains onto children and grandchildren, will make them relatively less valuable and more likely to be overshadowed by the increased variance.

The fall in the intergenerational correlation of income

The fall in the correlation on incomes is a new feature to this chapter. The correlation of the log of income and the log of parental income is increasing, as in chapter 2, but this does not necessarily translate into an increase in the correlation of income and parental income. It is given by equation 3.17. Since a fall in $\sigma_{\epsilon s_2}^2$ will inevitably lead to the firm using the talent signal more, one simple way to nullify the effect of this on the correlation is to have little or no inheritance (a small value of $\alpha_x + \alpha_y$).

There is a common theme that has emerged over the course of this section: occasions where mean income falls while information is more readily available or investment increasing, expected value falls in the face of increasing income, or mobility increases despite increased use of the talent signal, all feature low inheritance (particularly α_x) and high exogenous variance in talent. In some respects these could be thought of as healthy economies. The lack of inheritance will tend to push up intergenerational mobility, and the negative effect which it has on investment will be offset by the high variance in talent propping up β_2 . In other ways they could be thought off as unhealthy. The low value of α_x implies the technology for turning investment into talent (the private education system) is inefficient. In any case, and for various reasons outlined above, these economies display some interesting features.

An interesting case is illustrated in figure 3.1. Here α_x and α_y are 0.35 and $\sigma_{\epsilon\tau}^2$ is 2.5. Higher value of the exogenous variance in talent, and lower values of the inheritance parameters, will tend to cause mean income to be falling across the whole range of meritocracy values. We see here that it is only falling at high values of meritocracy, which coincide with higher values of β_2 . This illustrates the role of β_2 in pulling down mean income as meritocracy increases. The implications of the previous discussion were that, if $\sigma_{\epsilon\tau}^2$ was not large enough, nor inheritance small enough, to cause mean income

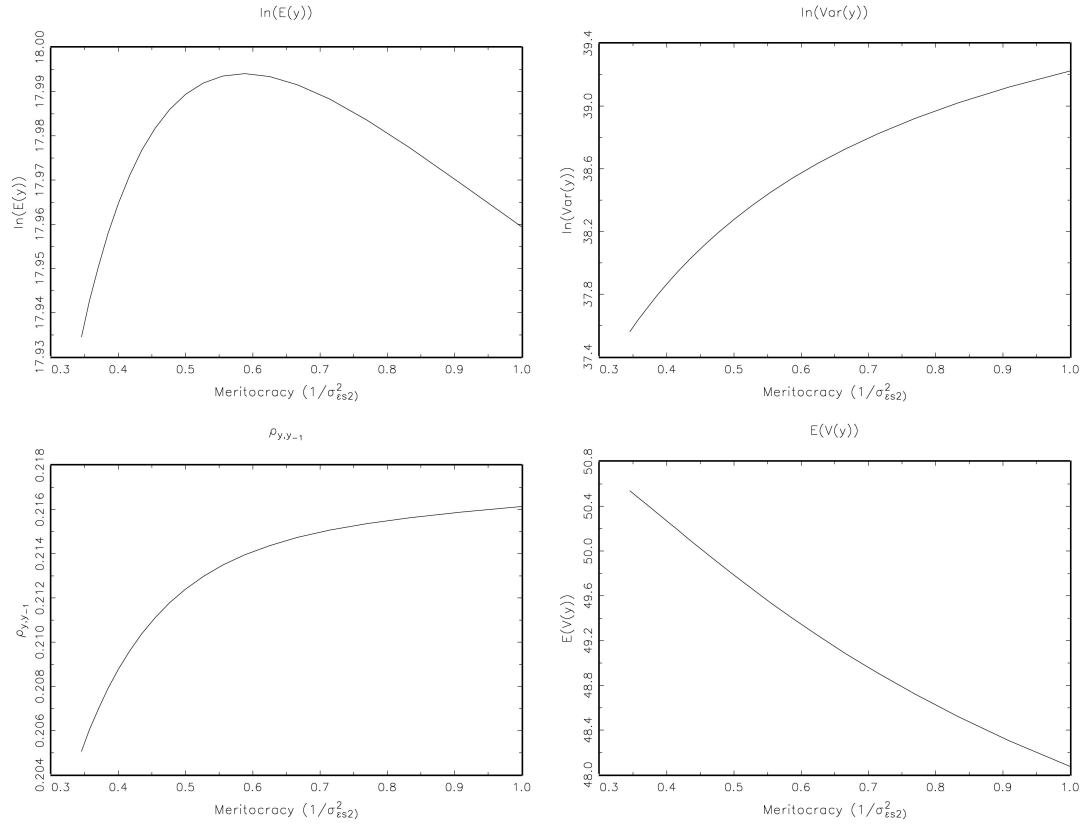


Figure 3.1: The effects of increasing meritocracy on mean income, income variance, the intergenerational correlation of incomes and expected value: $\alpha_x = 0.35$; $\alpha_y = 0.35$; $\sigma_{\epsilon s_1}^2 = 20$; and $\sigma_{\epsilon \tau}^2 = 2.5$

to be falling as meritocracy increases, they could still pull down expected value. This is exactly what happens when meritocracy is increasing from a low value.

3.4.2 The effects of an increase in inherited advantages in a weak inheritance economy

The results of the numerical example investigating the effects of a fall in $\sigma_{\epsilon s_1}^2$ of 0.001 are shown in table 3.3. In all cases firms use the parental income signal more, but the talent signal less. Intuitively the share of income invested by parents could go up or down. However, the intergenerational correlation of incomes always increases. We saw in the previous section that increased precision in the talent signal, which tends to be correlated with parental income, was not always enough to induce falls in mobility.

Statistic	Number of cases in which it increased	Percentage
$E [y]^*$	16,338	67.0
$Var [y]^*$	24,076	98.7
$\rho_{y,y-1}^*$	24,381	100.0
$E [V (y)]^*$	8,890	36.5
$E [x]^*$	20,616	84.6
λ^*	23,701	97.2

Table 3.3: The number and proportion of cases where the steady state values increased with an exogenous increase in inherited advantages in the weak inheritance economy. Total number of cases = 24,381

Improved information about background is, on every occasion, for this parameter set.

The fall in mean income and variance

We know from equations 3.15 and 3.16 that the steady state mean and variance of income are both determined by the steady state mean and variance of the log of income. From chapter 2 we know that the variance in the log of income will be increasing. A fall in the mean or variance in income then requires a fall in the mean of log income.

If the share of income being invested is increasing, a small value of α_x will nullify that, while a large fall in β_2 would push down the last term in equation 3.14. At least for the parameter values we are investigating, a large fall in β_2 is induced by a large value of β_2 . There are at least two reasons why firms may respond a great deal to the talent signal. A large amount of exogenous variance in talent would create an incentive for firms to use it. Also, a high precision would make it more informative, although as a counter to this a small value of $\sigma_{\epsilon_{s2}}^2$ will tend to nullify the effects of the fall in β_2 . A small value of $\sigma_{\epsilon_{s2}}^2$ does lead to falls in the mean and variance of income, although a large amount of exogenous variance in talent may have an even stronger effect.⁵

As in the previous section, a small value of α_x and large amount of exogenous variance in talent are the conditions under which a fall in mean income can occur in

⁵The mean and variance of income fall in 18 and zero cases respectively when $\sigma_{\epsilon_{\tau}}^2$ is 1 but in 2,410 and 94 cases when it is 7. The mean and variance fall in 1,594 and 259 cases respectively when $\sigma_{\epsilon_{s2}}^2$ is 2, but only 419 and zero when it is 18 (in fact for any value of $\sigma_{\epsilon_{s2}}^2$ greater than 6 we find that variance is always increasing).

the face of increasing mean investment. The effect of an increased share of income being invested in children is reduced by an inability to turn that into talent. The lack of a strong investment effect will make it more likely that the fall in β_2 which accompanies an increase in advantage will pull down mean income.

Also as in the previous section, we can consider whether the fact that mean investment can rise or fall is driven by perfectly offsetting effects on the investment choice due to the coefficient of relative risk aversion being equal to one. In over ninety-seven per cent of cases λ increases, so again there does not seem to be a lot of ambiguity in the parent's investment choice: they nearly always invest a larger fraction of their income when information is more readily available to firms and the advantages which investment creates can be passed on to children and grandchildren to a greater extent. In all of the cases where the fraction invested falls, β_2 falls, as it must since β_1 is increasing. A relatively large fall in β_2 will provide an incentive to invest less in children which, as mentioned above, seems to occur at large values of β_2 . We would therefore expect a shock to advantage to produce a large negative effect on the incentive to invest when $\sigma_{\epsilon\tau}^2$ is large and that this would be the circumstances under which λ^* may fall. This is what I find. This story seems to be consistent with agents responding to changes in the incentive to invest rather than the incentive to invest experiencing perfectly offsetting positive and negative effects.

The fall in expected value

There are just under 7,500 cases where expected value falls while mean income is increasing. The conditions for this are the same as in the case of an increase in meritocracy. An example where α_x and α_y are 0.05, and $\sigma_{\epsilon\tau}^2$ 7, is illustrated in figure 3.2. Although the additional information may raise your expected income, it will also raise the uncertainty you face over that income (if mean income is increasing, the variance in income is also increasing). The increase in mean income has to make you sufficiently happier to overcome the increased risk of low income.

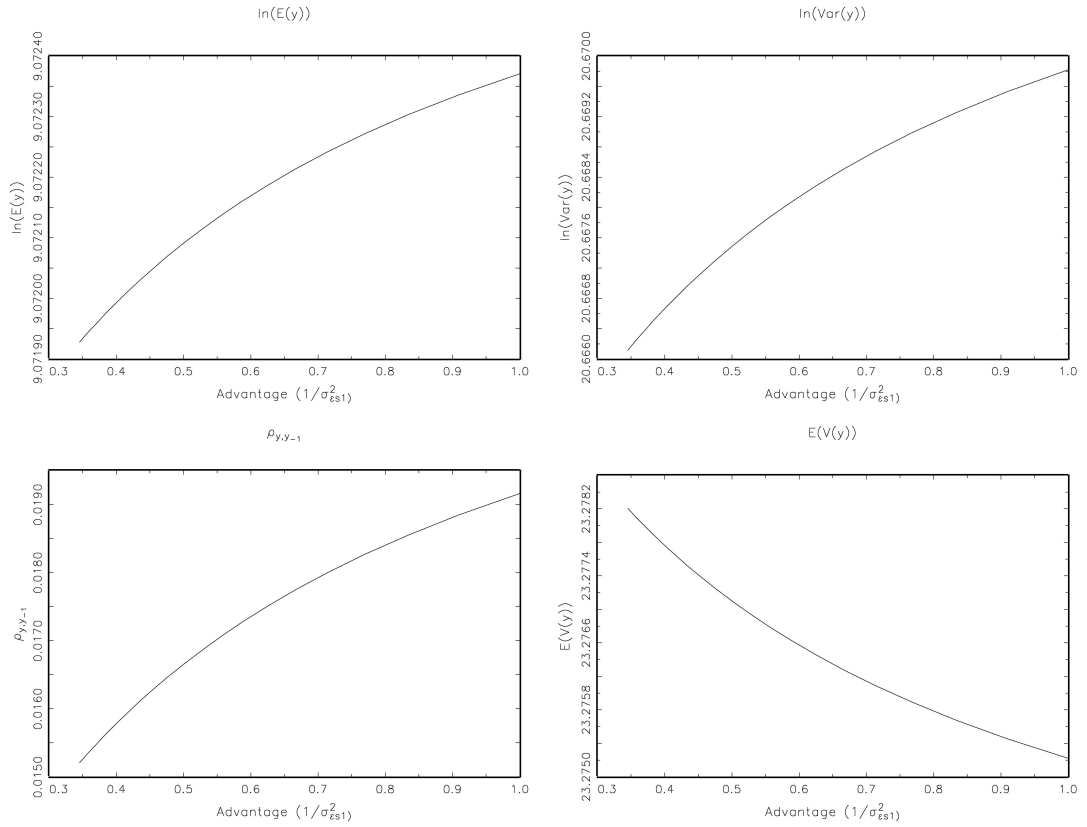


Figure 3.2: The effects of increasing advantage on mean income, income variance, the intergenerational correlation of incomes and expected value: $\alpha_x = 0.05$; $\alpha_y = 0.05$; $\sigma_{\epsilon s_1}^2 = 12$; and $\sigma_{\epsilon \tau}^2 = 7$

3.4.3 A discussion of the weak inheritance economy

Some of the more interesting results derived in this section have had at their heart a great deal of similarity in the parameters which bring them about. One of the key things is a lack of inheritability, particularly a low value of α_x . This implies that private investments made in children do not pay off. The private education system is inefficient. Even when mean investment increases it may not lead to increases in expected income since that investment is ineffective and, even when it does raise mean income, it may not improve the utility which an individual expects to receive since those income gains cannot easily be passed on to the next generation and the improvements in the information available to firms create greater uncertainty. Large values of $\sigma_{\epsilon \tau}^2$ also feed into this. Therefore, in an economy where people display a wide range of talents, but it is difficult to influence the talent of your child and hence pass on advantages,

Parameter	Values the parameter may take
α_y	0.3, 0.4, ..., 0.7
α_x	0.3, 0.4, ..., 0.7
$\sigma_{\epsilon s_1}^2$	2, 6, ..., 18
$\sigma_{\epsilon s_2}^2$	2, 6, ..., 18
$\sigma_{\epsilon \tau}^2$	1, 3, ..., 7
$\mu_{y_{-1}}$	2, 4, ..., 10
$\sigma_{y_{-1}}^2$	2, 4, ..., 10
ϕ	5
δ	0.5

Table 3.4: Parameter values for the numerical exercise in the strong inheritance economy

you may prefer a lower mean income if increases in that mean income reflect increases in meritocracy and advantage that bring about greater income variance.

3.5 The strong inheritance economy

The numerical exercise which I will undertake for the strong inheritance economy ($\alpha_x + \alpha_y > 1$) is slightly different to what we examined in the weak economy case. This is because, although it may converge to a constant finite value of the variance in log income, the economy will not converge to a steady state finite value of the mean and variance in income. Both will experience growth. What we will do instead of examining the steady states is to look at the shift in the level of our key variables when there is an exogenous shock to meritocracy and advantage, *given the current state of the economy*. It is the shift at the time of the change for that generation. For this we need two more parameters which are our state variables: $\mu_{y_{-1}}$, the mean of the log of parental income; and $\sigma_{y_{-1}}^2$ the variance of the log of parental income. The parameters used are given in table 3.4. There are 37,500 combinations of these parameters for which $\alpha_x + \alpha_y$ is greater than one. In this case λ is omitted from the numerical analysis since, as $E[x] = \lambda E[y_{-1}]$ and $E[y_{-1}]$ is given, λ must move in the same direction as $E[x]$.

Statistic	Number of cases in which it increased	Percentage
$E [y]$	7,500	20.0
$Var [y]$	7,500	20.0
$\rho_{y,y-1}$	19,995	53.3
$E [V (y)]$	26,985	72.0
$E [x]$	7,500	20.0

Table 3.5: The number and proportion of cases where the values increased with an exogenous increase in meritocracy in the strong inheritance economy. Total number of cases = 37,500

3.5.1 The effects of an improvement in meritocracy in a strong inheritance economy

The results of an exogenous increase in meritocracy are shown in table 3.5. In all cases β_2 must increase and β_1 must fall.

The fall in mean investment

There are only 7,500 cases where, faced with a higher value of β_2 and a lower value of β_1 , parents invest more in their children. When this occurs, it also leads to increases in mean income and income variance. It only occurs when $\sigma_{\epsilon s_2}^2$ is equal to 2, its minimum value, at which point β_2 will be large and β_1 small. That is to say, only if the primary channel through which parents provide advantages to their children is investment, does investment increase with an improvement in meritocracy.

Looking more closely at λ we can see that it is increasing in β_2 and A . We know that β_2 is increasing, so for mean investment to decrease, A must fall. Therefore A falls in at least 30,000 cases. This says that improvements in meritocracy tend to reduce the marginal value of log income. Improvements in meritocracy push the economy towards providing advantages through investment (at the expense of parental consumption) and away from providing natural advantages through environmental influences without such a sacrifice. What this tells us is that an additional unit of income will be more valuable in a society where you can provide advantages to your children without the sacrifice of your own consumption than in one which requires investment.

The fall in the intergenerational correlation of income

The fall in the intergenerational correlation of incomes was always more likely to occur in this exercise than that performed on the steady states in the previous section. This is because, in the steady state case, the substitution effect which causes a shift towards the talent signal was accompanied by a feedback effect whereby increased variance of the log of income caused both signals to be used more. This feedback effect was likened to the income effect from a fall in an equilibrium price. As it pushes up the use of both signals by the firm, it decreases mobility. In the current case, where we examine the effect in one generation, it is not present. There is therefore less reason to believe mobility should necessarily fall.

The parameter which most strongly affects whether mobility increases or decreases is $\sigma_{\epsilon\tau}^2$. When $\sigma_{\epsilon\tau}^2$ is large it means that there is a lot of exogenous variance in talent which can only be identified by use of the talent signal. As a result, β_2 will tend to be large and β_1 small. Advantages are primarily provided to children through the talent signal, and hence by investment in their development. An increase in β_2 under these circumstances further enhances this opportunity for rich parents to provide advantages to their children and therefore reduces mobility. Another way to think about this is from the firm's point of view: if people are exogenously very different, firms place a high value on information about those differences and being better able to identify them increases this discrimination and reduces mobility.

The fall in expected value

We have already established that in 30,000 cases $E[y]$ is reduced and that in all those cases A is reduced. However in close to 27,000 of these cases $E[V(y)]$ is increasing. It must be that in these cases B has increased. So although the marginal value of log income is reduced in more meritocratic societies, in the majority of cases people expect to have higher utility nonetheless. They expect to derive less utility from their own income but more value generally from living in a more meritocratic society. This

Statistic	Number of cases in which it increased	Percentage
$E [y]$	25,935	69.2
$Var [y]$	37,500	100.0
$\rho_{y,y-1}$	37,500	100.0
$E [V (y)]$	0	0.0
$E [x]$	16,175	43.1

Table 3.6: The number and proportion of cases where the values increased with an exogenous increase in advantages in the strong inheritance economy. Total number of cases = 37,500

is likely to be related to the fall in income variance which the shocks to meritocracy are producing. In the vast majority of cases where income variance is increasing, expected value is decreasing, and in the vast majority of cases where income variance is decreasing, expected value is increasing. This gives additional weight to the argument that changes in uncertainty over income are the driving force behind changes in expected value.

3.5.2 The effects of an increase in inherited advantages in a strong inheritance economy

The result of a fall in inherited advantages is shown in table 3.6.

The increase in variance and fall in mobility

One of the most striking results from this exercise was that variance increased and mobility fell in all cases. Despite the fall in β_2 which inevitably follows an increase in inherited advantages, the increase in discrimination by background was always sufficiently large to have these effects. In order to investigate this further, another numerical analysis was carried out where $\sigma_{\epsilon s_1}^2$ and $\sigma_{\epsilon s_2}^2$ could take values $\{16, 18, \dots, 24\}$ and $\sigma_{\epsilon \tau}^2$ could take values $\{1,2,3,4\}$. Again, in all cases variance increased and mobility fell with an exogenous increase in the extent of inherited advantages.

The fall in expected value

By running the numerical exercise with the alternative parameter sets for the error variances, I was able to investigate whether expected value always fell. It did not, but fell in over 23,000 of the 37,500 cases. From this experiment I was able to reach two further conclusions. Firstly, in all cases where expected value was falling despite expected income increasing, variance was increasing. This adds further weight to my argument that uncertainty over income drives falls in expected utility.

Secondly, there were 4,320 cases where people had higher expected utility after the shock to advantages, yet expected to have lower incomes and experience more uncertainty over their income. This relates to the fact that inherited advantages can be passed on to your descendants without any reduction in your own consumption, making them particularly valuable. In some cases this can lead people to prefer a world in which their income will be lower on average, and more uncertain, but create greater natural advantages for their children.

The increase in mean investment

This also relates to investment decisions. A fall in $\sigma_{cs_1}^2$ inevitably leads to a fall in β_2 and so in the more than 16,000 cases where investment is increasing it must be that A has increased. The marginal value of log income may be larger when that income provides natural advantages to your children to a greater extent, rather than advantages through investment. Furthermore, although the advantages which you provide to your children are limited by your own income, you can create advantages for your grandchildren, great grandchildren, and so on, by investing in the talent development of their parents. This is what is occurring in the more than 16,000 cases where mean investment has increased in spite of a reduction in the extent to which firms discriminate on talent.

3.6 Conclusions

There were two main parts to this chapter. The first was to provide a forward-looking component to the model of chapter 2, and consider the equilibrium investment rule of parents. This brings the model much more closely in line with the existing statistical discrimination literature, although the multiplicity of equilibria present in those models does not feature here. In this model there is a unique equilibrium investment rule, but the multiplicity of steady states found in chapter 2 remains. The results found in that chapter translate directly into increases in the variance of log income, and correlation of the logs on income and parental income, as meritocracy and advantage increase.

We are not primarily interested in the mean, variance or correlations of the log of income. Our main concern is with the distribution of income itself. The second part of this chapter looked at how the mean and variance of income changed with exogenous shocks to meritocracy and advantage within this new framework. It also considered mobility and expected utility.

In many cases, though not all, we found that variance and mobility still moved in opposite directions. This was particularly true for exogenous increases in inherited advantages. There were also a number of interesting results relating to a particular subset of the parameters. These were when inheritance was weak (particularly α_x , such as in the case where the private education system is inefficient) and there was a large amount of exogenous variance in talent. In this case, mean investment might fall despite talent being more easily recognised, mean income may fall despite mean investment increasing, and expected value might decrease despite expected income going up. It is the last of these that is probably the most interesting. Meritocracy and inherited advantages both increase variance and hence, prior to birth, create greater uncertainty over the income you will receive. A risk averse person might prefer to be born into an economy where they expect to be poorer but avoid this increased uncertainty, and so despite raising incomes, meritocracy and inherited advantages may make agents, on average, more unhappy.

Appendix A

The parameters of the value function

In section 3.1.3 I calculated the parameters of the value function of the parent. These were functions of the parent's expected parameters of the value function of their child.

In this appendix I will calculate explicitly the values of A_{-1} and B_{-1} .

Let us begin with A_{-1} .

$$A_{-1} = 1 + \frac{1}{1 + \delta} E_{-1}[A] [\beta_1 + (\alpha_x + \alpha_y) \beta_2]$$

This is the same as in the body of the text but with a subscript placed on the expectations operator to make clear that it is the expectation of the parent.

Moving forward one period,

$$A = 1 + \frac{1}{1 + \delta} E[A_{+1}] [\beta_{1,+1} + (\alpha_x + \alpha_y) \beta_{2,+1}]$$

where $\beta_{1,+1}$ and $\beta_{2,+1}$ are the weights assigned to the signals by firms in the grandchild's generation. Since the β terms follow a deterministic path, everything on the right-hand side is known to the parent. Therefore, $E_{-1}[A] = A$ and so,

$$A_{-1} = 1 + \frac{\beta_1 + (\alpha_x + \alpha_y) \beta_2}{1 + \delta} + E[A_{+1}] \frac{\beta_1 + (\alpha_x + \alpha_y) \beta_2}{1 + \delta} \frac{\beta_{1,+1} + (\alpha_x + \alpha_y) \beta_{2,+1}}{1 + \delta}$$

Since each parent has one child in each generation, families are infinitely lived. As in the main body of the chapter, we will assume that $\beta_1^* + (\alpha_x + \alpha_y) \beta_2^* < 1 + \delta$ where the *'s indicate steady state values. The implications of this constraint are explained briefly in section 3.1.3. It ensures that the sequence below converges to a positive finite value for A_{-1} . If we continue this process through the generations of a family we can see that,

$$A_{-1} = 1 + \frac{\beta_1 + (\alpha_x + \alpha_y) \beta_2}{1 + \delta} + \frac{\beta_1 + (\alpha_x + \alpha_y) \beta_2}{1 + \delta} \frac{\beta_{1,+1} + (\alpha_x + \alpha_y) \beta_{2,+1}}{1 + \delta} + \frac{\beta_1 + (\alpha_x + \alpha_y) \beta_2}{1 + \delta} \frac{\beta_{1,+1} + (\alpha_x + \alpha_y) \beta_{2,+1}}{1 + \delta} \frac{\beta_{1,+2} + (\alpha_x + \alpha_y) \beta_{2,+2}}{1 + \delta} + \dots$$

Each additional term will be smaller than the previous, at least once a steady state value of β_1 and β_2 are reached, which ensures convergence.

By the same logic,

$$A = E_{-1}[A] = 1 + \frac{\beta_{1,+1} + (\alpha_x + \alpha_y) \beta_{2,+1}}{1 + \delta} + \frac{\beta_{1,+1} + (\alpha_x + \alpha_y) \beta_{2,+1}}{1 + \delta} \frac{\beta_{1,+2} + (\alpha_x + \alpha_y) \beta_{2,+2}}{1 + \delta} + \frac{\beta_{1,+1} + (\alpha_x + \alpha_y) \beta_{2,+1}}{1 + \delta} \frac{\beta_{1,+2} + (\alpha_x + \alpha_y) \beta_{2,+2}}{1 + \delta} \frac{\beta_{1,+3} + (\alpha_x + \alpha_y) \beta_{2,+3}}{1 + \delta} + \dots$$

which can be written,

$$A = 1 + \sum_{i=1}^{\infty} \frac{\prod_{j=1}^i [\beta_{1,+j} + (\alpha_x + \alpha_y) \beta_{2,+j}]}{(1 + \delta)^i} \quad (\text{A.1})$$

This is the value of A which is used in the parent's investment function in equation 3.11. The additional utility which a child will receive from an increase in his income depends on the extent to which he can use that income, through investment, to influence the incomes and utilities of future generations. This in turn depends on the level of discrimination which those future generations will face.¹

It is straightforward to show that when $\beta_1 = \beta_{1,+1} = \dots = \beta_1^*$ and $\beta_2 = \beta_{2,+1} = \dots = \beta_2^*$,

$$A^* = \frac{1}{1 - \frac{\beta_1^* + (\alpha_x + \alpha_y) \beta_2^*}{1 + \delta}} = \frac{1 + \delta}{1 + \delta - [\beta_1^* + (\alpha_x + \alpha_y) \beta_2^*]} \quad (\text{A.2})$$

which is the steady state value of A used in section 3.3.1. The assumption $\beta_1^* + (\alpha_x + \alpha_y) \beta_2^* < 1 + \delta$ ensures that in steady state an individual's income has a positive effect on his income.

Moving on to consider B_{-1} from section 3.1.3, we know that B_{-1} is given by,

¹A general version of this equation can be written to provide the value of A for any generation in the future. This is given by,

$$A_{+i} = 1 + \sum_{t=1}^{\infty} \frac{\prod_{j=1}^t [\beta_{1,+i+j} + (\alpha_x + \alpha_y) \beta_{2,+i+j}]}{(1 + \delta)^t}$$

$$\begin{aligned}
B_{-1} &= \ln \left[\frac{1 + \delta}{(1 + \delta) + \alpha_x \beta_2 E_{-1}[A]} \right] + \left[\frac{\alpha_x E_{-1}[A]}{1 + \delta} \right] \ln \left[\frac{\alpha_x \beta_2 E_{-1}[A]}{(1 + \delta) + \alpha_x \beta_2 E_{-1}[A]} \right] \\
&+ \frac{E_{-1}[A]}{1 + \delta} \left\{ \phi + \frac{\beta_2 \sigma_{\epsilon s_2}^2}{2} - [\beta_1 - (1 - \beta_2)(\alpha_y + \alpha_x)] \mu_{y-1} \right\} + \frac{1}{1 + \delta} E_{-1}[B] \\
&= \chi(A, \beta_1, \beta_2) + \frac{E_{-1}[B]}{1 + \delta}
\end{aligned}$$

This is the same equation as was given in section 3.1.3 but having substituted for λ .
By the same logic as for A, we can show that B is given by,

$$B = \sum_{i=1}^{\infty} \frac{\chi(A_{+i}, \beta_{1,+i}, \beta_{2,+i}, \mu_{y,+i-1})}{(1 + \delta)^{i-1}} \quad (\text{A.3})$$

where²

$$\begin{aligned}
\chi(A_{+i}, \beta_{1,+i}, \beta_{2,+i}, \mu_{y,+i-1}) &= \ln \left[\frac{1 + \delta}{(1 + \delta) + \alpha_x \beta_{2,+i} A_{+i}} \right] + \left[\frac{\alpha_x A_{+i}}{1 + \delta} \right] \ln \left[\frac{\alpha_x \beta_{2,+i} A_{+i}}{(1 + \delta) + \alpha_x \beta_{2,+i} A_{+i}} \right] \\
&+ \frac{A_{+i}}{1 + \delta} \left\{ \phi + \frac{\beta_{2,+i} \sigma_{\epsilon s_2}^2}{2} - [\beta_{1,+i} - (1 - \beta_{2,+1})(\alpha_y + \alpha_x)] \mu_{y,+i-1} \right\}
\end{aligned}$$

It follows that in steady state,

$$B^* = \frac{\chi(A^*, \beta_1^*, \beta_2^*, \mu_y^*) (1 + \delta)}{\delta} \quad (\text{A.4})$$

²For small values of λ_{+i} , the proportion of income invested by parents in i generations time, this gives,

$$\begin{aligned}
&\chi(A_{+i}, \beta_{1,+i}, \beta_{2,+i}, \mu_{y,+i-1}) \approx \\
&- \frac{\alpha_x A_{+i} [1 + \beta_{2,+i}]}{(1 + \delta) + \alpha_x \beta_{2,+i} A_{+i}} + \frac{A_{+i}}{1 + \delta} \left\{ \phi + \frac{\beta_{2,+i} \sigma_{\epsilon s_2}^2}{2} - [\beta_{1,+i} - (1 - \beta_{2,+1})(\alpha_y + \alpha_x)] \mu_{y,+i-1} \right\}
\end{aligned}$$

Summary

This thesis has considered how information about an individual's talent or background is used by firms, and how that process is affected by the endogenous determination of the quality of that information. Each chapter attempted to build on the previous one, to look in more detail at some of the features of the economy which emerges as a result of this dynamic process. In the first chapter, talent was a binary variable. The firm had no information about whether a potential employee was talented, but had perfect information about their parental income which defined an individual's "class". This seemed a natural set-up from which to investigate class and the allocation of workers to tasks, but less-so for an investigation of the inequality and mobility which emerge in response to the extent of discrimination which firms carry out. For one thing, there was no mobility in chapter one. Therefore, in chapter two the analysis switched to a model where talent and income were both continuous and jointly determined. In general, neither was perfectly known by firms though they received a noisy signal on each. Discrimination fed inequality, while inequality fed discrimination. Additional realism was then added in chapter three: the income process was log-normal, which more closely matches the observed income distribution than the normal one used in chapter two; risk-averse optimising agents chose how much to invest in their children; and there was potential growth in mean income. The inclusion of optimising agents also allowed some welfare analysis to be brought in. This final section of the thesis will summarise and discuss some of the main conclusions which each of these set-ups allowed me to reach.

The first chapter of this thesis looks at the class divisions which emerge in society.

In the absence of classes, firms lack the information they require to determine the talent of an individual. They have little to base their pay decisions on and do not know which workers to assign to a relatively complex task. The class signal provides them, perhaps imperfectly, with that information. The cost is that the talented group of people created is smaller than the proportion of workers they would like to assign to the complex task. Within the model, where the informational gains exist, they dominate the allocation costs. There are class divisions of society which can be sustained in steady state for which both classes of society are better off than if no such class division existed, and in which output is higher. Instead of everyone having the same probability of being talented, society specialises – the upper class specialise in the complex task and achieve a high marginal productivity. The spillover to the lower classes also raises their marginal productivity.

The ability of firms to offer wages based on class causes those classes of workers to sort appropriately into jobs and raises the economy's output. However, there are many steady state divisions of society which go beyond the point at which these gains are exhausted. The upper class can be squeezed in order to raise their level of talent, and society's productivity, only to the point where they are all talented. Equally, these potential gains should not mask the discrimination which is going on. Even in the absence of any legal impediments to them taking the complex job, the informational asymmetry between them and the firm causes a talented lower class worker to be de facto denied access to the high paying jobs offered to the upper classes, despite the fact that he is able to perform the task while some members of the upper class may not be able to. Class is still very much an imperfect sorting mechanism for firms.

The second and third chapters can be viewed as separate within the same general framework. In the second chapter, means are normalised to be zero and I focus on the variance in income and its correlation across generations. The early part of that chapter is concerned with the feedback effect whereby discrimination is both caused by, and a cause of, greater inequality. The latter part of the chapter considers some perverse effects from improvements in meritocracy: the talented workers identified in meritocratic societies tend to be from rich backgrounds; and since meritocratic institutions

facilitate discrimination amongst workers, they may lead firms to use information of a worker's background more. Both of these feed into lower intergenerational mobility, and this provides a new explanation of the "Great Gatsby Curve" based on incomplete information.

There was one policy discussed in this chapter. It was noted that a policy which breaks the close link between parental income and human capital development would both lower inequality and raise mobility. One way to break this link would seem to be early childhood interventions. There have been arguments put forward previously for how early childhood interventions would effect the economy. Heckman and various coauthors have argued, based on work in neuroscience and childhood development, that early childhood interventions would reduce inequality because of the complementarity between skills produced at different stages of childhood (see for example Cunha & Heckman (2007) and Heckman & Masterov (2007)). This chapter puts forward another explanation for why this might occur based on information extraction by firms and the role of early childhood interventions in limiting the usefulness of information on family background.

The final chapter of the thesis considers in more detail the investment decision underlying the relationship between parental income and a child's human capital development. In the model of that chapter there is a unique equilibrium investment rule whereby parents invest a constant fraction of their income in their child. That fraction is increasing in inheritance (exogenously governed by the elasticities of child's talent with respect to parental income and investment) and the endogenous extent to which firms discriminate. There are two spin-offs of this process which are important innovations in this chapter. The first is that mean income is generally not equal to zero and there may be steady state growth in mean income. The second is that, because we have optimising agents, welfare can be considered and agents are risk-averse in income.

The main result to emphasise from the second part of this chapter was that expected utility might be higher in less meritocratic societies. The increased ability of firms to tell workers apart leads to greater labour market discrimination and inequality. Given

agents are risk averse, they will be relatively unhappy with this income inequality and the uncertainty which meritocracy encourages. In fact, even when meritocracy raises mean income it might lower expected utility because of the inequality which it creates.

There has been one theme which has been carried throughout this thesis – there can be informational gains from discrimination based on background. In chapter 1 this raised productivity through improved sorting and spillover effects, while in chapter 3 it tended to do the same through encouraging investment. However, in chapters 2 and 3 it was the basis for the feedback effect which exaggerated increases in income inequality and the intergenerational correlation of incomes. It is not obvious therefore that these informational gains are beneficial to welfare. The existence of rich and poor classes of workers creates advantages and disadvantages which go beyond their talents. We may need to be careful about how we consider meritocracy, at least as defined by the ability of workers to display their talent and be paid accordingly. Meritocracy leads to discrimination and, with it, inequality and a lack of mobility. If it leads firms to identify and promote the children of the rich, the persistent inequality which this generates may not be viewed in a positive light by a society of risk averse agents. The value to a worker of a privileged background may be great while at the same time the value to society of a privileged class of workers is limited at best. Classes and background allow us to discriminate, they allow us to identify the talented in society, they may even encourage investment in talent and raise mean income, but, when it comes to society's welfare, the uncertainty over income which they create may be sufficient to lower expected utility and cause people to prefer poorer countries in which the income they receive is more certain and the discrimination they face less severe.

Bibliography

- Altonji, J., & Pierret, C. (2001). Employer learning and statistical discrimination. *Quarterly Journal of Economics*, *116*, 313–350.
- Arcidiacono, P., Bayer, P., & Hizmo, A. (2008). Beyond signaling and human capital. *National Bureau of Economic Research, NBER Working Paper 13951*.
- Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter, & A. Rees (Eds.) *Discrimination in Labor Markets*, chap. 1. Princeton University Press.
- Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press, 1st ed.
- Coate, S., & Loury, G. (1993). Will affirmative action policies eliminate negative stereotypes? *American Economic Review*, *83*(5), 1220–1240.
- Corak, M. (2012). Inequality from generation to generation: The united states in comparison. In R. Rycroft (Ed.) *The Economics of Inequality, Poverty, and Discrimination in the 21st Century*. ABC-CLIO forthcoming.
- Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.
- Farber, H. S., & Gibbons, R. (1996). Learning and wage dynamics. *Quarterly Journal of Economics*, *111*, 1007–1047.
- Hassler, J., Rodríguez Mora, J. V., & Zeira, J. (2007). Inequality and mobility. *Journal of Economic Growth*, *12*, 235–259.

- Heckman, J., & Masterov, D. (2007). The productivity argument for investing in young children. *National Bureau of Economic Research, NBER Working Paper 13016*.
- Krueger, A. B. (2012). The rise and consequences of inequality in the united states. Speech to the Center for American Progress.
- Lange, F. (2007). The speed of employer learning. *The Journal of Labour Economics*, 25(1), 1–35.
- Lazear, E. P. (1995). Hiring risky workers. *National Bureau of Economic Research, NBER Working Paper 5334*.
- Lundberg, S. J., & Startz, R. (1983). Private discrimination and social intervention in competitive labor markets. *American Economic Review*, 73(3), 340–347.
- Moro, A. (2003). The effect of statistical discrimination on black-white wage inequality: Estimating a model of multiple equilibria. *International Economic Review*, 44(2), 467–500.
- Moro, A., & Norman, P. (2004). A general equilibrium model of statistical discrimination. *Journal of Economic Theory*, 114, 1–30.
- Norman, P. (2003). Statistical discrimination and efficiency. *Review of Economic Studies*, 70, 615–627.
- Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, 62(4), 659–661.
- Piketty, T. (1995). Social mobility and redistributive politics. *Quarterly Journal of Economics*, 110(3), 551–584.
- Piketty, T. (1998). Self-fulfilling beliefs about social status. *Journal of Public Economics*, 70(1), 115–132.
- Rodgers, W. M. (2008). Understanding the black-white earnings gap: Why do african americans continue to earn less despite dramatic gains in education? *The American Prospect*.

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.