Marital Matching and the Distribution of Innate Ability*

Chris Bidner[†] and John Knowles[‡]

May 7, 2019

Abstract

We analyse how inequality and social mobility are influenced by incomplete information about the inherited ability of potential spouses. When motivated by dynastic altruism, singles rank potential spouses according to beliefs about their expected ability, *independently* of their earnings potential. Due to the heritability of innate ability, such beliefs are informed not only by noisy signals based on the observed performance of individuals, but also by family 'status', a scalar that summarizes the information revealed by each agent's ancestors. In the equilibrium matching, the strength of sorting on actual ability depends on the quality of information. This makes social mobility sensitive to parameters, such as the return to ability, that alter the noisiness of the signals. Redistributional policies, because they are common knowledge, have little effect on inequality of ability, because they do not alter the ranking of potential spouses. Matching on beliefs also implies a new 'status motive' for parental investment; the illusion of higher ability of the offspring generates genuinely higher ability among the grandchildren.

Keywords: Family Economics, Inequality, Household Formation, Marriage.
JEL Classification: D10, D19, D31, D80, D83, H31, H52, I24, J11, J12, J18

*We are grateful for constructive comments from presentations at the University of Calgary, Mannheim, Sciences Po, the 2018 NBER Summer Institute and SFU. This research was supported by the Social Sciences and Humanities Research Council of Canada.

†Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, B.C. Canada V5A 1S6. cbidner@sfu.ca

‡Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, B.C. Canada V5A 1S6. jaknowle@sfu.ca

1 Introduction

Even in the most meritocratic of societies, individuals' economic prospects are circumscribed by their family background. Parents actively shape the economic prospects of offspring in various ways, for instance by investing in human capital. However, parents also *passively* transmit a range of productive capabilities–"innate ability" for short–to their offspring.[1]

There are instrumental reasons for concern about inequality of innate ability, because it has direct implications for income inequality and social mobility.[2] But the inequality of innate ability is also intrinsically important because heritability makes it robust to standard policy tools–innate ability cannot readily be redistributed, subsidized or publicly provided.[3] This resistance to policy intervention only exacerbates fears surrounding the formation of a 'cognitive elite' and the associated detrimental social consequences.[4] For instance, Arrow et al. [2000] describe settings where "position is secured by the possession of advanced mental skills and passed on from generation to generation through the genetic inheritance of these skills" as among the "specters of intractable inequality [that] now haunt the public imagination". Beyond economic inequality, such processes are also feared to generate problematic social cleavages.[5]

Despite this passive nature of ability transmission and its robustness to policy, economic choices are still relevant, because the degree of parent-child transmission of ability depends on who marries whom. When the ability of both parents contribute to offspring ability, the intergenerational persistence of ability hinges on the strength of marital sorting on the ability dimension. Even when everyone prefers a spouse

[1]The term "innate ability" is chosen because of the emphasis it places on the passive nature of transmission. That is, it refers to the aggregate of those productive capabilities which resist the active interventions of parents (or anyone else). We are primarily motivated by the genetic component of cognitive ability, but we are by no means restricted to this interpretation: the term is completely analogous to the "endowments" of Becker and Tomes [1979], which include cognitive abilities but also "family reputation and "connections", and knowledge, skills, and goals provided by their family environment" (p.1153). The term is also completely analogous to the "social competence" of Clark [2014].

[2]For instance, it is central in understanding the 'unusually' strong persistence of economic status documented over multiple generations (Clark [2014]).

[3]Even beyond standard policy tools, evidence from Clark [2014] suggests that this passive transmission process is unaffected by wildly different institutional settings.

[4]See Murray [2012], Arrow et al. [2000], Herrnstein and Murray [1994], and Young [1958].

[5]For instance, David Brooks worries that the "members of the educated class use their intellectual, financial and social advantages to pass down privilege to their children, creating a hereditary elite that is ever more insulated from the rest of society" (From "The Strange Failure of the Educated Elite", by David Brooks in the New York Times May 28 2018). Such insulation is problematic because the "fragile web of civility, mutual regard, and mutual obligations at the heart of any happy society begins to tear" (Herrnstein and Murray [1994]).

with higher ability, the outcome is not obvious, because of an important information friction–innate ability is not observed. Sorting of singles into couples must rely on beliefs about an individual's ability, which may be formed, not only from observation of their outcomes prior to matching, but also, given the hereditary nature of ability, from their family history.

Our goal in this paper is to model how such beliefs are formed, how they influence marital sorting, and how the resulting forces cause inequality and social mobility to depend on the policies and technology of a given society.[6] The simple model presented in section 2 incorporates these features in a tractable manner. The key assumptions are that singles are altruistic towards (potential) descendants and that ability is both heritable and unobservable . Each individual is born with an ability that depends on their parents' average ability plus a random component. This ability gets combined with a random 'luck' component to form the individual's human capital, which in turn influences beliefs about their ability, and hence their attractiveness in the marriage market, independently of the effect of earnings. The heritability of ability therefore requires that we track beliefs across generations in light of these observed human capital realizations. We solve analytically for the husband-wife ability correlations in the steady-state equilibrium, and derive the key statistics describing inequality and intergenerational mobility of both ability and income: the parent-child ability correlations and the population variances.

In order to focus on the central issue of marital sorting on beliefs about innate ability, we abstract in the main model from all other frictions in the marriage market.[7] In the equilibrium of our model, marriages are perfectly sorted on observables, and we show that this leads to perfect sorting on *expected* ability. The strength of sorting on *actual* ability depends on the precision of beliefs–the quality of information. We show how the quality of information evolves over time, and globally converges to a unique steady-state value. The quality of information will, via its effect on sorting, determine the intergenerational persistence of ability and the dispersion of ability in society.

When the marital sorting is exogenously fixed, intergenerational transmission of innate ability is independent of technology, social policy and other features of the eco-

[6] As will become clear, we are abstracting from evolutionary forces stemming from differential fitness (we treat fertility as independent of ability). Instead, our focus is on endogenous marital sorting on the ability dimension. This sorting will have implications for the variability of ability in the population, but *not* for average ability.

[7] We discuss frictional marriage markets and provide a version with imperfect sorting on human capital in the extensions section of the model. We show that such an extension modifies our key results in a transparent manner yet does not undermine the key mechanisms we emphasize.

nomic environment.[8] However, in our model, marital sorting responds to elements of the economic environment that affect the quality of information, such as the role of ability versus luck, in the formation of human capital, and so these elements will have an impact on the transmission of ability in our setting. We also examine the conditions under which policy variables–such as those governing the equality of outcomes (redistribution of income) and opportunity (redistribution of parental inputs)–will have an impact on the distribution of innate ability.

Our analysis introduces the role of the family as a repository of information. The experiences of a generation are incorporated into their family history, which then in turns shapes the experiences of future generations. We show how the relevant family history can be summarized by a scalar, which we call *family status*, and derive the simple law of motion that governs it. The model conveniently delivers an endogenous measure of the importance of family background; the weight that agents optimally place on an individual's family status relative to their human capital during the process of belief updating.

The full model presented in section 2 generalizes the analysis in various ways. We endogenize parental investment in human capital, and parameterize the return to ability and parental investment. We consider several departures from strict 'meritocracy' in the model, such as the possibility of nepotism, where wealthy parents can divert the income of others to their own offspring, as well as the role of progressive redistribution of income to poorer households. We also we allow for a policy that makes society more meritocratic, by redistribution of education investment.

The equilibrium has the property that perceptions of ability in one generation influence actual ability in suceeeding generations. This raises two more questions: the impact of luck in prior generations, and the motivation for parental investments. We show that the economic luck experienced by ancestors percolates down through the generations via a novel channel–the effect on expected innate ability of descendants. Luck raises the ancestors' perceived ability, which allows them to marry spouses of higher expected ability. The higher-ability of spouses in turn contributes to the actual ability (on average) of descendants. This impact is persistent, since the original luck becomes embedded into family status, which has an independent effect on the marital prospects of all descendants. Similarly, matching on status distorts the investment incentives facing parents. If investments are not publicly observed, then parents will be motivated

[8] This is a direct consequence of passive transmission. The economic environment could shape the distribution of ability (but not its transmission conditional on sorting) via standard evolutionary dynamics if there were differential fertility. In any case, we abstract from differential fertility to focus in on the sorting issue.

to invest, in part, by an attempt to raise the marriage market's assessment of their offspring's ability. This investment will of course not affect the ability of offspring (as it is determined by parental abilities, not investment), but it will affect the expected ability of further generations. Again, it allows offspring to attract a spouse of higher ability on average, and this will raise the expected ability of grandchildren. Furthermore, the higher assessment of offspring ability will be preserved in higher family status which, will help all future generations attract higher-ability spouses.

Our work is clearly related to a large literature on inherited inequality. The bulk of this literature ignores marital matching so that households can be treated like individuals (e.g. Becker and Tomes [1979, 1986], Lee and Seshadri [2019], Solon [2004], Clark [2014], Benabou [2002]. Sorting is considered, but fixed exogenously, in Kremer [1998]). Some work in this literature does incorporate marital sorting, but imposes perfect information about the relevant characteristics (e.g. Aiyagari et al. [2000], Fernández et al. [2005], Cole et al. [1992], Anderson [2015], Zak and Park [2002]).

Our steady-state analysis of the informational role of family status relate to a smaller literature on matching on unobserved characteristics. Previous work in this area abstracted from generational structure, so the analysis was either static or held constant the unobserved type over time (e.g. Hoppe et al. [2009], Cole et al. [1995], Hopkins [2012], Bidner et al. [2016], Bidner [2010], Rege [2008], Anderson and Smith [2010], Bergstrom and Bagnoli [1993], Chade [2006]).[9]

We also contribute to research concerned with the role of parents in shaping the economic opportunities of their children. In addition to analyzing the role of human-capital investments, this body of work emphasizes the *active* role of parents in making decisions about where to live (Benabou [1996], Durlauf [1996], Fernandez and Rogerson [1996]), which cultural traits to inculcate (Bisin and Verdier [2001], Francois and Zabojnik [2005], Tabellini [2008]), and which parenting style to adopt (Doepke and Zilibotti [2017]).

1.1 Discussion of 'Innate Ability'

As noted, our notion of 'innate ability' permits a variety of interpretations; transmission need not be genetic in nature. Indeed, our model accommodates cultural transmission, whereby ability is encoded in cultural traits such as values and beliefs (e.g. Cavalli-Sforza and Feldman [1981], Boyd and Richerson [2005], Bisin and Verdier [2001], and Dawkins [2006]).

[9] Unobserved ability is also considered in Comerford et al. [2017] but they are concerned with labour-market discrimination rather than marital matching.

We have three reasons for describing our model in terms of the genetic channel. First, our premises are consistent with mainstream models in quantitative genetics. These premises are that transmission (i) is passive, and (ii) necessarily involves the contribution of both parents.[10]

Second, the existence of a significant genetic channel is supported by a rapidly increasing volume of available evidence. The evidence from behavioural genetics, typically grounded in a comparison of identical and fraternal twins or of biological and adopted siblings, suggests important genetic components for a variety of relevant traits from IQ to financial literacy (see Polderman et al. [2015] and Sacerdote [2011] for a review of this vast literature). These findings have been further supported in recent years by evidence from molecular genetics, whereby an individual's relevant traits are associated directly with their genome (e.g. see Okbay et al. [2016]). In addition, recent evidence on inter-generational social mobility over multiple generations suggests that the transmission process is remarkably stable across wildly different economic environments, and is also consistent with a variety of other signatures of genetic transmission (see Clark [2014], especially Chapter 7).

Third, if sorting is to be of quantitative relevance, then the transmission of ability between parents and offspring must be of sufficiently high fidelity (Kremer [1998]). The genetic channel offers a far higher parent-child fidelity because, unlike genes, the transmission of culture is not restricted to parent-child pairs and can operate with equal force within groups of peers, colleagues and extended family. Of course, none of this is to suggest that cultural transmission is unimportant (only that it will have little quantitative impact on our objects of study, such as parent-child correlations, unless it operates in a manner that mimics genetic transmission), or that human capital is unresponsive to environmental influences (we explicitly include environmental contributions in the model). In short, our intention is to exploit a deliberately simple model that delivers insights which transparently extend beyond specific interpretations.

[10]The additive genetic model (Fisher , 1919), which is widely used in the field of polygenic inheritance, also supports our more restrictive assumptions: the effects of genes are additive, and hence the effects on the offspring of the genes of one parent are independent of the effects of the genes of the other. We do not however impose the controversial assumptions that the genes are uncorrelated with environmental effects or that the latter are uncorrelated across agents.

2 Model

2.1 Fundamentals

2.1.1 Population

We consider an infinite-horizon discrete-time economy with time indexed by $t = 0, 1, 2, ...$. Each date t starts with a continuum of households, indexed by i, each consisting of one male and one female parent. Household (i, t) has two offspring, one male and one female. The offspring from each household are then matched with an individual of the opposite sex,forming the households of the following period. The parents of each household die at the end of the period, and the household index is retained by the male offspring.[11] We denote the household index of his wife by i'. That is, household $(i, t+1)$ is formed by the marriage of the male offspring from household (i, t) with his spouse, the female from household (i', t). By construction, the measure of males, females and households is the same in each period, which we normalize to unity.

2.1.2 Parental Characteristics

The parents in family (i, t) each have a predetermined *ability*, denoted by the scalars $(\theta_{i,t-1}, \theta_{i',t-1})$. Ability is unobserved, but there are commonly-held beliefs about the ability of each parent, described by the probability-density functions $(\psi_{i,t-1}, \psi_{i',t-1})$. Parents also have predetermined human-capital levels, denoted in logarithmic terms by the scalars $(x_{i,t-1}, x_{i',t-1})$.

2.1.3 Household Income

The human capital of parents in household (i, t) is used to produce household output. Since we are interested in the role of economic policy, we allow for redistribution of household output in the determination of household income. Redistribution arises from two sources. First, we allow for a 'plutocratic' effect: the redistribution of output towards households formed by the offspring of high-income parents. Second, we allow for progressive taxation : the redistribution of income towards low-productivity households. The structural equations underlying these policies (described in the Appendix), imply a reduced-form in which the income of household (i, t), denoted y_{it}, is a log-linear function

[11] Since there is no asymmetry by sex, the female's index would work just as well. The emphasis on the male lineage is motivated by the empirical literature which focuses on traits that are easier to measure for the male lineage, such as last names and earnings.

of the human capital and the parental income of the spouses:

$$y_{i,t} = \beta_0 + \beta_1 \cdot [x_{i,t-1} + x_{i',t-1}]/2 + \beta_2 \cdot [y_{i,t-1} + y_{i',t-1}]/2. \tag{1}$$

The extent of meritocracy is thus determined by the parameters β_1 and β_2. For instance, equality of outcomes is achieved when $\beta_1 = \beta_2 = 0$. The constant β_0 is determined by the resource constraint (total output equals total income).

2.1.4 Offspring's Ability

Offspring are born after household income is generated. The offspring in household (i, t) are endowed with an unobserved *ability*, denoted θ_{it}.[12] This ability is partly inherited from the parents, according to:

$$\theta_{it} = b \cdot [\theta_{i,t-1} + \theta_{i',t-1}]/2 + \upsilon_{it} \tag{2}$$

where $b \in (0, 1)$ and $\upsilon_{it} \sim N(0, \sigma_\upsilon^2)$ is an idiosyncratic component.

2.1.5 Offspring's Human Capital

The *human capital* of offspring depends on ability, but also on parental investment and luck. Parental investment must be financed from family income. While other agents may correctly infer the equilibrium investment investment from parental income, deviations from this value are not publicly observed. Again, since we are interested in the role of economic policy, we allow for the effective redistribution of parental investments. The structural model, described in the Appendix, implies that human capital of offspring in household (i, t), denoted x_{it}, is log-linear in parental investment h_{it} and in the idiosyncratic 'luck' component $\varepsilon_{i,t}$, where $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$. This reduced-form equation is given by:

$$x_{i,t} = \alpha_0' + \alpha_1 \cdot \theta_{i,t} + \alpha_2 \cdot h_{i,t} + \varepsilon_{i,t}. \tag{3}$$

The value of σ_ε^2 parameterizes the extent of luck. The parameter α_1 is the return to ability–i.e. it measures the extent to which ability matters for human-capital production. The parameter α_2 measures the extent to which parental investment matters for human-capital production, and is thus declining in the extent of parental investment redistribution. For instance, we can think of 'equality of opportunity' arising when

[12] Our assumption that siblings have the same ability is for simplicity–it is not essential but allows us to abstract from the possibility that sibling outcomes are informative about one's ability.

$\alpha_2 = \beta_2 = 0$. The constant, α_0', ensures that redistribution of investment leaves total expenditure on human-capital investment unchanged.

2.1.6 Preferences

Parents have preferences over their own consumption and the infinite sequence of utilities of their descendants, discounted geometrically by discount factor $\delta \in (0, 1)$:

$$U_{i,t} = u(C_{i,t}) + \mathbb{E}_t \left[\sum_{\tau=1}^{\infty} \delta^{\tau} \cdot u(C_{i,t+\tau}) \right] \tag{4}$$

where $C_{i,t}$ is the consumption of household (i, t), and $u(C) = \ln C$. Since spouses each have the same utility function, the same consumption and the same descendants, utility in our model is non-transferable. [13]

2.1.7 Beliefs and Status

Before the offspring's human capital is observed, there are commonly-held prior beliefs about the ability of the offspring of household (i, t), denoted by the probability density function $\bar{\psi}_{it}$. The expected ability of offspring under these prior beliefs defines *family status*, denoted $\bar{\phi}_{it}$:

$$\bar{\phi}_{it} \equiv \mathbb{E}[\theta_{it} | \bar{\psi}_{it}] = \int \theta \bar{\psi}_{it}(\theta) d\theta. \tag{5}$$

Beliefs about the ability of offspring are updated on the basis of the offspring's realized human capital. Given these commonly-held posterior beliefs, denoted by the probability-density function ψ_{it}, we can define the *individual status* of the offspring from household (i, t), denoted ϕ_{it}, as their expected ability:

$$\phi_{it} \equiv \mathbb{E}[\theta_{it} | \psi_{it}] = \int \theta \psi_{it}(\theta) d\theta. \tag{6}$$

We can now define the *quality of information* associated with offspring of household (i, t) as γ_{it}^{-1} , where

$$\gamma_{it} \equiv \mathbb{E}[(\theta_{it} - \mathbb{E}[\theta_{it} | \psi_{it}])^2 | \psi_{it}] = \int (\theta - \phi_{it})^2 \psi_{it}(\theta) d\theta \tag{7}$$

[13] Strictly speaking, this says that parents care about the consumption of descendants down the male line. Given symmetry across the sexes, this is equivalent to caring about an average across all descendants within each subsequent generation. We adopt this specification purely for notational convenience.

is the variance of ability, conditional on beliefs.

2.1.8 Marital Matching

At each date t, the offspring from household $(i, t-1)$ enter the marriage market, characterized by their human capital, the income of their parents, and by publicly-held posterior beliefs about their ability: $\omega_{it} \equiv \{x_{it}, y_{it}, \psi_{it}\}$. We refer to ω_{it} as the agent's *marriage type*. The marriage market assigns males and females to pairs, forming the households of the next period. The assignment is subject to the commonly-held posterior beliefs. Since the information about agent ability is imperfect and symmetric, the abilities of paired agents will be conditionally uncorrelated; being matched provides no information about the spouse's abilities $\{\theta, \theta'\}$ beyond that described by beliefs:

$$\mathbb{E}[(\theta_{it} - \phi_{it})(\theta_{i't} - \phi_{i't})|\psi_{it}, \psi_{i't}] \equiv \int \int (\theta - \phi_{it})(\theta' - \phi_{i't})\psi_{it}(\theta)\psi_{i't}(\theta')d\theta d\theta' = 0. \quad (8)$$

This implies that any correlation of the ability of husband and wife is due solely to sorting on the observable marriage types ω_{it}.

2.1.9 Initial Conditions

At $t = 0$ each household i is endowed with income $Y_{i0} > 0$ for all i. Nature endows the offspring in household i with ability $\theta_{i,0} \sim N(\bar{\phi}_{i,0}, \bar{\gamma}_{i,0})$, where $\bar{\phi}_{i,0} \in \mathbb{R}$ and $\bar{\gamma}_{i0} > 0$ for all i. Prior beliefs about the ability of offspring in family i at $t = 0$ therefore coincide with this normal distribution.

2.2 Equilibrium and Steady State

2.2.1 Equilibrium Conditions

An equilibrium of our model is a description of how beliefs are formed, of how much households invest in human capital, and of how spouses are assigned in the marriage market. We require equilibrium beliefs to be rational, given the assignment and the investment policy. Given match formation and beliefs, we also require the equilibrium investment policy to be optimal and the assignment to be both feasible and stable. These conditions are described below; a more formal definition of the Bayesian equilibrium can be found in the appendix.

To be rational, given prior beliefs, human capital and parental income, posterior beliefs must satisfy Bayes' rule:

$$\psi_{it}(\theta) = \frac{\bar{\psi}_{it}(\theta) \cdot f(x_{it} \mid \theta, y_{it})}{\int \bar{\psi}_{it}(\tilde{\theta}) \cdot f(x_{it} \mid \tilde{\theta}, y_{it}) d\tilde{\theta}}, \tag{9}$$

where $\bar{\psi}_{it}(\theta)$ are prior beliefs and f is the conditional distribution of log human capital[14] that is implied by (3).[15]

Prior beliefs must satisfy rationality given (posterior) beliefs about parental ability, $(\psi_{i,t-1}, \psi_{i',t-1})$, the conditional independence of parental ability (8), and the ability transmission process (2). Rationality of prior beliefs thus requires that, for $t > 0$, we have:

$$\bar{\psi}_{it}(\theta) = \int \int \tilde{\psi}_{i,t-1}(r_1) \cdot \tilde{\psi}_{i',t-1}(r_2 - r_1) \cdot f^{\upsilon}(\theta - r_2) dr_1 dr_2, \tag{10}$$

where $\tilde{\psi}_{i,t-1}(r) \equiv (2/b) \cdot \psi_{i,t-1}((2/b) \cdot r)$ is the density of $(b/2) \cdot \theta_{i,t-1}$ given parental beliefs, and f^{υ} is the density associated with the distribution of the random component of ability, υ_{it}.

Optimality of investments, given match formation and beliefs, requires that h_{it} be chosen to maximize (4), taking as given the investment strategies of others (including future generations).

The assignment is one-to-one, meaning that each agent is assigned either one partner or none. To be feasible, the assignment must ensure that the number of assigned women of each type be no greater than the number of female offspring of that type, and that the converse holds for the men. Stability further requires that no two agents would strictly prefer to marry each other than remain in their assigned match. Partner preferences are evaluated according to (4) , taking belief formation and the investment policy as given.

2.2.2 Steady State

Equilibrium behaviour implies a distribution of ability (and income) across individuals within a generation and across individuals within lineages. With respect to the

[14] We could allow agents to use other variables beyond household income to form this conditional distribution (e.g. beliefs about parental abilities), but doing so would be redundant in the cases we consider. This is because optimal parental investment is only a function of household income (plus, in extensions, an idiosyncratic component).

[15] For instance, if optimal investment implies $h_{it} = h^*(y_{it}) + \varepsilon^h_{it}$, where ε^h_{it} is a mean-zero belief error, then $f(x \mid \theta, y_{it}) = f^{\varepsilon^x}(x - \alpha'_0 - \alpha_1 \cdot \theta - \alpha_2 \cdot h^*(y_{it}))$, where f^{ε^x} is the density of the compound stochastic component, $\varepsilon^x_{it} \equiv \alpha_2 \cdot \varepsilon^h_{it} + \varepsilon_{it}$.

distribution of abilities, our primary objects of interest are (i) the variance of ability

$$\sigma^2_{\theta,t} \equiv \mathrm{Var}[\theta_{i,t}],$$

(ii) the husband-wife ability correlation

$$\rho^{HW}_{\theta,t} \equiv \mathrm{Cor}[\theta_{i,t}, \theta_{i',t}],$$

and (iii) the parent-child ability correlation

$$\rho^{PC}_{\theta,t} \equiv \mathrm{Cor}[\theta_{i,t}, \theta_{i,t-1}].$$

In a steady state (stationary equilibrium), these moments are time-invariant, and denoted σ^2_θ, ρ^{HW}_θ and ρ^{PC}_θ.[16]

The 'income' counterparts of these moments are also time-invariant in the steady state. These consist of: (i) inequality, as measured by the variance of log income

$$\sigma^2_{y,t} \equiv \mathrm{Var}[y_{i,t}],$$

and (ii) social mobility, as measured (inversely) by the parent-child income correlation

$$\rho^{PC}_{y,t} \equiv \mathrm{Cor}[y_{i,t}, y_{i,t-1}].$$

Since income is a joint property of the household, we ignore the income counterpart of the husband-wife ability correlation.[17]

3 Analysis and Results

3.1 Segregation and Proportional Investment

The model's linear structure and Gaussian distribution generates equilibria in which the matching outcomes and optimal investments are straightforward to characterize. This allows us to focus squarely on our primary object of interest–how beliefs determine the strength of marital sorting on unobserved ability.

[16]The analysis also allows us to analyse a variety of other correlations of interest, including the ability correlation of individuals $k > 1$ generations apart, as well as the correlation of offspring ability with parental income. These correlations are derived and discussed in the Appendix.

[17]In the appendix we also derive and analyze other correlations of interest. For instance, see proposition 11 for the correlation between offspring ability and parental income, proposition 9 for the k-generation ability correlation, and proposition 10 for the k-generation income correlation.

The equilibrium sorting takes an intuitive form: *segregation*. This means that each individual has the same marriage type as their partner (i.e. $\omega_i = \omega_{i'}$); in other words, the assignment features positively-assorted matching along each observable dimension.[18] This is particularly useful because it allows us to analyze equilibrium ability while holding fixed the sorting patterns along other dimensions (such as human capital and parental income). In this equilibrium, it is optimal for all households invest the same proportion z of their income in children's human capital. Our strategy in analysing the model is to first take these two features as given and then to verify them once we have derived their implications for the belief-updating process.[19]

One implication of segregation, arising from $\psi_{it} = \psi_{i't}$, is that the husband and wife will be of equal status (expected ability):

$$\phi_{it} = \phi_{i't}. \tag{11}$$

Another implication, arising from $x_{it} = x_{i't}$ and $y_{it} = y_{i't}$, is that we can simplify the income equation (1) to:

$$y_{i,t} = \beta_0 + \beta_1 \cdot x_{i,t-1} + \beta_2 \cdot y_{i,t-1}. \tag{12}$$

Due to the proportionality of the investment policy, the human capital equation (3) simplifies to:

$$x_{i,t} = \alpha_0 + \alpha_1 \cdot \theta_{i,t} + \alpha_2 \cdot y_{i,t} + \varepsilon_{i,t}, \tag{13}$$

where $\alpha_0 = \alpha_0' + \alpha_2 \cdot \ln z$. Using (13) in (12) yields a reduced-form relationship between the offspring's income and their ability, parents' income, and luck:

$$y_{i,t+1} = \pi_0 + \pi_1 \cdot \theta_{i,t} + \pi_2 \cdot y_{i,t} + \varepsilon^y_{i,t}$$

where $\pi_0 \equiv \beta_0 + \beta_1\alpha_0$, $\pi_1 \equiv \beta_1\alpha_1$, $\pi_2 \equiv \beta_1\alpha_2 + \beta_2$, and $\varepsilon^y_{it} \equiv \beta_1 \cdot \varepsilon_{it}$. Thus, the reduced-form luck component, $\varepsilon^y_{i,t}$, is normally distributed with mean zero and variance $\sigma^2_{\varepsilon^y} \equiv \beta_1^2 \cdot \sigma^2_\varepsilon$.

Another implication of proportional investment, arising from $u(C_{it}) = \ln C_{it} = \ln(1-z) + y_{it}$, is that utility over future generations can be expressed in terms of the

[18] This should be interpreted as a large-population idealization of marrying someone (from a different family) with 'similar' characteristics.

[19] We verify that segregation is indeed feasible and stable in section D.1.1 of the appendix, and that households optimally invest the same proportion of income in section 5.

discounted sum of log incomes:

$$U_{i,t} = \mathbb{E}_t \left[\sum_{\tau=0}^{\infty} \delta^{\tau} \cdot y_{i,t+\tau} \right] + U_0(z), \tag{14}$$

where $U_0(z) \equiv [1/(1-\delta)] \cdot \ln(1-z)$ is a constant.

Notice too that proportional investment means that the equilibrium value of the investment rule z only appears in the constant terms, and will therefore not be relevant for deriving the correlations of interest.

3.2 The Quality of Information

3.2.1 Rational Posterior Beliefs

Posterior beliefs about offspring ability are formed by updating prior beliefs on the basis of the offspring's human capital, x_{it}, and parental income, y_{it}. While this sort of Bayesian updating can quickly become complicated, our model reduces this to a very tractable form. By re-arranging the human-capital equation (13), we derive $s_{i,t}$, a noisy signal of ability:

$$s_{i,t} \equiv \frac{x_{it} - \alpha_0 - \alpha_2 \cdot y_{i,t}}{\alpha_1} = \theta_{it} + \xi_{i,t}, \tag{15}$$

where $\xi_{it} \equiv \varepsilon_{it}/\alpha_1$. This in turn implies that the signal is normally distributed:

$$s_{it} \sim N(\theta_{it}, \sigma_{\xi}^2), \tag{16}$$

where $\sigma_{\xi}^2 = \sigma_{\varepsilon}^2/\alpha_1^2$ denotes the variance of the signal noise.

If the prior belief about the ability of offspring is normal (as is the case at $t = 0$), then since the signal is also normally distributed, standard results imply that the posterior will also be normal. In particular, suppose that prior beliefs are normal with a variance of $\bar{\gamma}_{it}$ and a mean of $\bar{\phi}_{it}$. Then, posterior beliefs will also be normal, with variance γ_{it} given by:

$$\gamma_{it} \equiv \frac{\sigma_{\xi}^2 \cdot \bar{\gamma}_{it}}{\sigma_{\xi}^2 + \bar{\gamma}_{it}}. \tag{17}$$

The mean of posterior beliefs depends on a weighted average of the mean of prior beliefs and the signal:

$$\phi_{it} \equiv \lambda_{it} \cdot \bar{\phi}_{it} + (1 - \lambda_{it}) \cdot s_{it}. \tag{18}$$

As one might expect,the relative weight λ_{it} placed on the prior is decreasing in the variance of the prior $\bar{\gamma}_{it}$:

$$\lambda_{it} \equiv \frac{\sigma_\xi^2}{\sigma_\xi^2 + \bar{\gamma}_{it}}. \tag{19}$$

The value of λ_{it} provides a convenient measure of the equilibrium importance of family background; note also that λ_{it} evolves over time, as a function of the variance of beliefs.

To summarize, rationality of the posterior beliefs implies that ability will be normally distributed around the posterior mean:

$$\theta_{it} \sim N(\bar{\phi}_{it}, \bar{\gamma}_{it}) \;\Rightarrow\; \theta_{it} \mid s_{it} \sim N(\phi_{it}, \gamma_{it}). \tag{20}$$

3.2.2 Rational Prior Beliefs

Prior beliefs about offspring ability are formed by combining the beliefs about the parent's ability with the ability-transmission equation, (2). Given segregation, it is straightforward to show that the distribution of prior beliefs has a mean of

$$\bar{\phi}_{it} = b \cdot \phi_{i,t-1}, \tag{21}$$

and a variance of

$$\bar{\gamma}_{it} = \frac{b^2}{2} \cdot \gamma_{i,t-1} + \sigma_v^2. \tag{22}$$

Furthermore, if posterior beliefs about parental types are normally distributed, then it follows from (2) that prior beliefs will also be normal. That is:

$$\theta_{i,t-1} | s_{i,t-1} \sim N(\phi_{i,t-1}, \gamma_{i,t-1}) \;\Rightarrow\; \theta_{it} \sim N(\bar{\phi}_{it}, \bar{\gamma}_{it}). \tag{23}$$

3.2.3 Dynamics of Beliefs

We now consider the implications of the above analysis for the dynamics of beliefs. The key property, arising from linearity of the reduced form, and the assumption that noise is Gaussian, is that prior and posterior beliefs are normally distributed for all (i, t). This follows from (20), (23) and the initial condition that priors are normal at $t = 0$.

That is, we can iteratively apply the principle that normality of the priors at t implies normality of the posteriors at t, which in turn implies normality of the priors at $t+1$.

The way in which beliefs evolve over time can be inferred by using (15), (18), and (21). Together, these equations imply that perceptions of expected ability–i.e. individual status–evolve in response to both actual ability and signal noise:

$$\phi_{it} = \lambda_{it} \cdot b \cdot \phi_{i,t-1} + (1-\lambda_{it}) \cdot \theta_{i,t} + (1-\lambda_{it}) \cdot \xi_{i,t}. \tag{24}$$

The dynamics of family status can be similarly derived.[20] Thus, status is in part inherited and in part sensitive to individual outcomes. A larger λ_{it} implies that status is more persistent–i.e. less subject to individual outcomes. The implication that signal noise is transmitted to descendants will become especially relevant when it comes to understanding incentives to invest in offspring and the nuanced ways in which economic fortunes persist.

The variance of beliefs also evolves over time; using (22) and (17), we can show that the variance of posterior beliefs for family i evolves according to the following difference equation:

$$\gamma_{it} = \frac{\sigma_\xi^2 \cdot [(b^2/2) \cdot \gamma_{i,t-1} + \sigma_v^2]}{\sigma_\xi^2 + (b^2/2) \cdot \gamma_{i,t-1} + \sigma_v^2}. \tag{25}$$

It is straightforward to see that this implies global convergence to the value of γ which is the unique positive solution to

$$\gamma = \frac{\sigma_\xi^2 \cdot [(b^2/2) \cdot \gamma + \sigma_v^2]}{\sigma_\xi^2 + (b^2/2) \cdot \gamma + \sigma_v^2}. \tag{26}$$

The fact that this limit is independent of i further implies that, in the long run, the quality of information is the same for all individuals.

Proposition 1 *In the long run, the quality of information is the same for all individuals:* $\lim_{t\to\infty} \gamma_{it}^{-1} = \gamma^{-1}$ *where* γ *satisfies (26). This long-run quality of information,* γ^{-1}*, is (i) decreasing in signal noise* $\sigma_\xi^2 \equiv \sigma_\varepsilon^2/\alpha_1^2$*, (ii) decreasing in the heritability parameters,* b *and* σ_v^2*, and is (iii) independent of the policy parameters* α_2*,* β_1*, and* β_2*.*

The proof follows from the above discussion with the comparative statics being easily derived from (26).[21] Intuitively, the steady-state quality of information is reduced as

[20] In particular, $\bar{\phi}_{i,t+1} = b \cdot [\lambda_{it} \cdot \bar{\phi}_{i,t} + (1-\lambda_{it}) \cdot \theta_{i,t} + (1-\lambda) \cdot \xi_{i,t}]$.

[21] The right side of (26) is increasing in σ_ξ^2 (which is $\sigma_\varepsilon^2/\alpha_1^2$), implying that so too is γ. The right side of (26) is increasing in σ_v^2 and in b, implying that so too is γ.

the signal from human capital becomes noisier (i.e. as σ_ξ^2 increases). Notice in particular that this implies that the quality of information is increasing in the return to ability, α_1. Similarly, the quality of information decreases as the noise embedded in prior beliefs increases (i.e. as b and σ_v^2 increase). Finally, the policy parameters do not matter for the quality of information since agents 'filter out' the impact of policy when updating beliefs.

Since the long-run relative importance of family background, λ_{it} is determined by γ^{-1} , we can now use equations (19) and (22) to derive the analogous comparative statics for λ_{it}.

Proposition 2 *In the long run, the relative importance of family background is the same for all individuals:* $\lim_{t\to\infty}\lambda_{it} = \lambda \equiv \frac{\sigma_\xi^2}{\sigma_\xi^2+(b^2/2)\cdot\gamma+\sigma_v^2}$ *where* γ *satisfies (26). This long-run relative importance of family background,* λ*, is (i) increasing in signal noise* $\sigma_\xi^2 \equiv \sigma_\varepsilon^2/\alpha_1^2$*, (ii) decreasing in the heritability parameters* b *and* σ_v^2*, and (iii) independent of* α_2*,* β_1*, and* β_2*.*

The proof follows from the above discussion, with the comparative statics being easily derived using (26).

Intuitively, agents rely more on prior beliefs as the signal becomes noisier. This implies that family background becomes less important as the return to ability increases. Similarly, agents rely less on prior beliefs as these beliefs become noisier (i.e. as the heritability parameters increase). To illustrate the above results, Figure 1 plots the effect of the return to ability on the steady-state quality of information and relative importance of family background for a given set of parameters, listed in the caption.

To summarize, a noisier signal reduces the long-run quality of information and raises the relevance of family background. A noisier prior also reduces the long-run quality of information, but reduces the importance of family background. Perhaps counter-intuitively, this implies that a higher return to ability lowers the relevance of family background (despite the fact that the only systematic component of ability is heritable), as does a stronger relationship between ability and parental abilities (i.e. a higher b).

3.3 Steady-State Distributions

Deriving the quality of information in the steady-state is valuable because it allows us to derive the steady-state distributions of ability and income as follows.

The ability-transmission equation (2) alone implies a steady-state relationship between the variance of ability and the parent-child ability correlation as functions of the

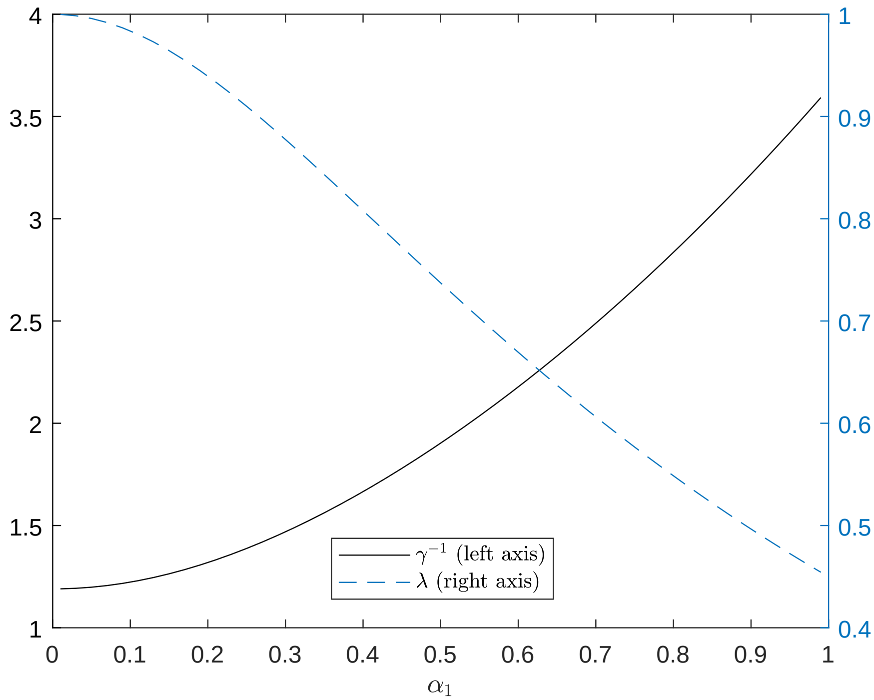

Figure 1: Steady-State Quality of Information and Importance of Family Background

Note: The solid line is the steady-state quality of information, γ^{-1}, and is measured on the left axis. The dashed line is the relative importance of family background, λ, and is measured on the right axis. Parameter values are $b = 0.9, \sigma_\upsilon^2 = 0.5, \sigma_\varepsilon^2 = 0.5$.

husband-wife ability correlation:[22]

$$\sigma_\theta^2 = \frac{\sigma_\upsilon^2}{1 - \frac{b^2}{2} \cdot (1 + \rho_\theta^{HW})} \tag{27}$$

$$\rho_\theta^{PC} = \frac{b}{2} \cdot (1 + \rho_\theta^{HW}) \tag{28}$$

Intuitively, a stronger husband-wife correlation implies less 'mixing' of parental ability. This in turn reduces the expected difference in the abilities of a parent and child, increasing the expected difference in ability between two random unrelated individuals.

We now proceed to closing this system by deriving the steady-state husband-wife ability correlation, ρ_θ^{HW}.

[22] These equations are analogous to those in Kremer [1998], where the husband-wife correlation is taken as exogenous. For instance, if we had instead assumed that type were observed then segregation would imply $\rho_\theta^{HW} = 1$ and therefore $\sigma_\theta^2 = \sigma_\upsilon^2/[1 - b^2]$ and $\rho_\theta^{PC} = b$.

Lemma 1 *In the steady state, the husband-wife ability correlation satisfies*

$$\rho_\theta^{HW} = 1 - \frac{\gamma}{\sigma_\theta^2} \tag{29}$$

where γ^{-1} is the steady-state quality of information.

Intuitively, given matching involves segregation on beliefs, a greater quality of information means that actual spousal abilities will tend to better resemble another. Since $\gamma \in [0, \sigma_\theta^2]$,[23] it follows that γ/σ_θ^2 lies in $[0, 1]$ and therefore provides a natural 'standardized' measure of how uninformative beliefs are. The expression on the right side of (29) therefore is a standardized measure of how informative beliefs are.

The steady-state values of $(\rho_\theta^{HW}, \sigma_\theta^2, \rho_\theta^{PC})$ are therefore given by the solution to equations (27), (28) and (29), as we report in Proposition 3.[24]

Proposition 3 *In the steady-state equilibrium, given quality of information, γ^{-1}, the dispersion, persistence and sorting properties of ability are given by:*

$$\sigma_\theta^2 = \frac{\sigma_\upsilon^2 - \gamma \cdot \frac{b^2}{2}}{1 - b^2} \tag{30}$$

$$\rho_\theta^{PC} = b \cdot \frac{\sigma_\upsilon^2 - \frac{\gamma}{2}}{\sigma_\upsilon^2 - b^2 \cdot \frac{\gamma}{2}} \tag{31}$$

$$\rho_\theta^{HW} = \frac{\sigma_\upsilon^2 - \gamma \cdot (1 - \frac{b^2}{2})}{\sigma_\upsilon^2 - b^2 \cdot \frac{\gamma}{2}}. \tag{32}$$

Each of these outcomes is increasing in the quality of information.

The proof is in appendix section C.1. Intuitively, a greater quality of information strengthens husband-wife sorting and this in turn raises the variance and intergenerational persistence of ability as outlined above.

Having analysed the forces that drive the sorting, persistence, and dispersion of ability, we now consider the consequences of these forces for income inequality and social mobility.

[23]Intuitively, in the extreme case where there is no available information, agents' beliefs would coincide with those arising from taking a random draw from the population–and thus $\gamma = \sigma^2$. On the other extreme, information is so good that there is no uncertainty, so that $\gamma = 0$.

[24]The fact that $\gamma \in [0, \sigma_\theta^2]$ ensures that each of the quantities in the proposition falls within the appropriate ranges (i.e. variance is non-negative and correlations are between -1 and 1. Specifically, $\gamma \in [0, \sigma_\theta^2]$ ensures that $\rho_\theta^{HW} \in [0, 1]$, that $\rho_\theta^{PC} \in (b/2, b)$, and that $\sigma_\theta^2 \in [\sigma_\upsilon^2/(1 - b^2/2), \sigma_\upsilon^2/(1 - b^2)]$.

Proposition 4 *Given the steady state quality of information, γ^{-1}, the steady state dispersion and persistence properties of income are given by:*

$$\sigma_y^2 = \frac{\left[\frac{1+b\pi_2}{1-b\pi_2}\right]\pi_1^2 \cdot \sigma_\theta^2(\gamma) + \sigma_{\varepsilon^y}^2}{1-\pi_2^2} \tag{33}$$

$$\rho_y^{PC} = \pi_2 + b \cdot \frac{(1-\pi_2^2)\cdot[\pi_1^2 \cdot \sigma_\theta^2(\gamma)]}{(1+b\pi_2)\cdot[\pi_1^2\cdot\sigma_\theta^2(\gamma)] + (1-b\pi_2)\cdot[\sigma_{\varepsilon^y}^2]}, \tag{34}$$

where $\sigma_\theta^2(\gamma)$ is given by (30). Both of these outcomes are increasing in the quality of information, γ^{-1}.

The proof is contained in the derivations in appendix section C.1. The intuition is straightforward–a higher quality of information raises the inequality and persistence of ability, and this translates directly into greater inequality and persistence of income.

The expression for the persistence of income, ρ_y^{PC}, clarifies the role of heritable ability in social mobility: a positive parent-offspring income correlation would be observed *even if* parental income had no causal impact on income (i.e. even if $\pi_2 = 0$).

4 Discussion of Results

We now turn to our main results: the implications of these propositions for marital sorting on ability and for the persistence and inequality of ability and income.

4.1 The Steady-State Distributions of Ability and Income

4.1.1 Impact of the Economic Environment

By construction, inherited ability is not affected by an agent's environment. Nevertheless, the environment shapes the dispersion and persistence of ability in a society, because the environment determines the availability of information about ability, thereby influencing sorting. In other words, only those aspects of the environment that determine the information available for matching can be expected to have an impact on the dispersion and persistence of ability.

Corollary 1 *An increase in α_1 or a decrease in σ_ε^2 will strengthen sorting, persistence and dispersion of ability.*

The return to ability, α_1, has an impact on how similar an individual's ability is to that of their spouse and their parents and children, because an increase in the return to ability makes an agent's human capital a more reliable signal of ability, which facilitates stronger sorting in the marriage market. These relationships are illustrated in Figure 2.

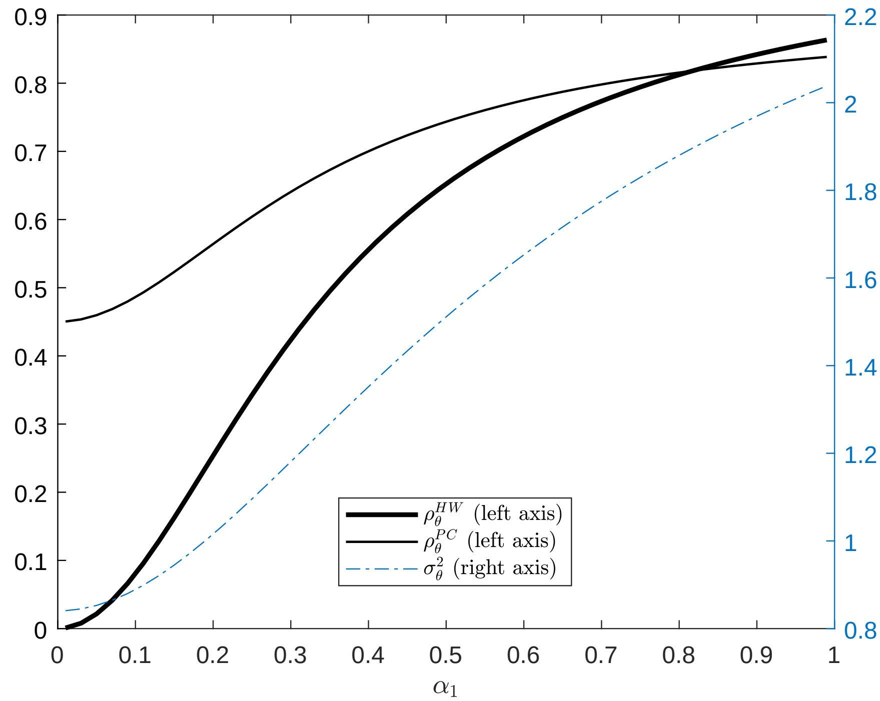

Figure 2: The Return to Ability and Features of the Ability Distribution

Note: The thick solid line is husband-wife ability correlation, the thin solid line is the parent-child ability correlation, and the dashed line is the population variance of ability. The solid lines are measured on the left axis, the dashed line is measured on the right axis. Parameter values are $b = 0.9, \sigma_v^2 = 0.5, \sigma_\varepsilon^2 = 0.5$.

Consider a competing explanation for this effect: in a model with frictional matching, a greater return to ability provides incentives to search more intensely for a high-ability partner. Note that this competing explanation would also predict that sorting systematically varies with the policy environment (e.g. lower redistributive taxation should also provide incentives to search more intensely for a high ability partner). We show in the next section that redistribution does not affect sorting in our model. Thus the evidence from Clark [2014], that inter-generational ability transmission appears to be relatively robust across different social regimes, is supportive of our mechanism against this alternative.

In terms of income, the variables $(\alpha_1, \sigma_\varepsilon^2)$ will have a direct effect on inequality and social mobility (i.e. holding ability sorting fixed) and an indirect sorting effect.

Corollary 2 *The direct effect of α_1 on social mobility and on inequality is exacerbated* *by the sorting effect. The direct effect of σ_ε^2 on social mobility is augmented by the sorting effect, whereas the direct effect of σ_ε^2 on inequality is mitigated by the sorting effect.*

Intuitively, if we hold ability sorting fixed, a larger return to ability lowers social mobility and raises inequality. But it also facilitates stronger marital sorting on ability, and thereby raises the persistence and dispersion of ability, which in turn lowers social mobility and raises inequality further. Figure 3 illustrates this by examining the total effect of α_1 on ρ_y^{PC} (left) and σ_y^2 (right) and comparing these to the values arising when sorting is fixed at the initial value. The lower of the two lines in each panel represent the direct effect of α_1 and the difference between the lines in each panel represents the indirect sorting effect.

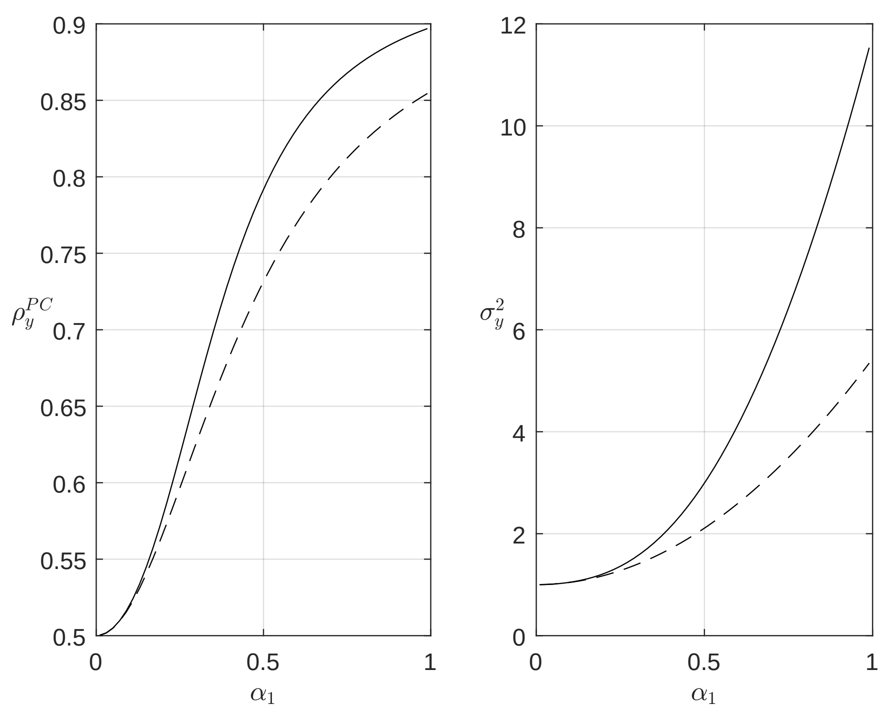

Figure 3: Return to Ability, Income Distribution and the Sorting Effect

Note: The left panel is the parent-child income correlation and the right panel is the variance of log income. The solid lines indicate the equilibrium values, whereas the dashed lines hold fixed the husband-wife sorting on ability (at the initial level). The difference in the height of the lines indicates the size of the sorting effect. Parameter values are $b = 0.9, \sigma_v^2 = 0.5, \sigma_\varepsilon^2 = 0.5, \alpha_2 = 0.5, \beta_1 = 1, \beta_2 = 0$.

Similarly, if we hold ability sorting fixed, a larger luck component raises social mobility and raises inequality. But it also weakens marital sorting on ability, and thereby lowers the persistence and dispersion of ability, which in turn raises social mobility further but also lowers inequality. This leads to the possibility that luck will have a non-monotonic effect on income inequality: luck raises inequality holding ability sorting fixed, but lowers inequality by weakening ability sorting. This is illustrated in Figure 4. Here the upper line represents the direct effect and the lower line incorporates the counteracting indirect sorting effect.

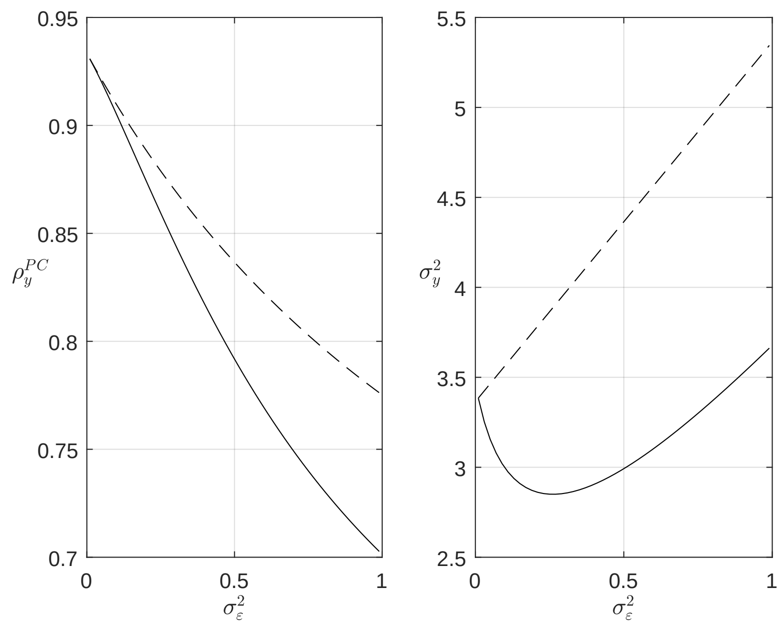

Figure 4: Luck, Income Distribution and the Sorting Effect

Note: The left panel shows the parent-child income correlation and the right panel shows the variance of log income. The solid lines indicate the equilibrium values, whereas the dashed lines hold fixed the husband-wife sorting on ability (at the initial level). The difference in the height of the lines indicates the size of the sorting effect. Parameter values are $b = 0.9, \sigma_v^2 = 0.5, \alpha_1 = 0.5, \alpha_2 = 0.5, \beta_1 = 1, \beta_2 = 0$.

4.1.2 Institutional/Policy Environment

The institutional/policy environment depends on the parameters α_2, β_1, and β_2. Recall that these describe the extent to which parental human-capital inputs are redistributed, as well as the extent to which income is redistributed via taxation and departures from meritocracy. In terms of income, these parameters clearly have an impact on social mobility and inequality (see proposition 4). However, in the base model at least, they have no impact on the extent of *ability* sorting, persistence or dispersion.

Corollary 3 *The institutional/policy environment variables, as measured by the reduced-form parameters α_2, β_1, and β_2, have no effect on the sorting, persistence or dispersion of ability.*

This result highlights the difficulty in inferring (unobservable) changes in the sorting, persistence and dispersion of ability from (observable) changes in the sorting, persistence and dispersion of income. The evidence from Clark [2014], that parent-child correlations of income fail to predict the strength of the correlations across many generations, is supportive of this implication of our model.

4.1.3 Heritability Environment

The 'heritability' variables, b and σ_v^2, will clearly have a direct impact on the dispersion and persistence of the heritable characteristic (see proposition 3). But, less obviously, they will also have an effect on sorting, and thus an indirect sorting effect on the dispersion and persistence of ability, via their effect on the precision of steady-state beliefs.

Corollary 4 *The direct effect of b on the persistence and dispersion of ability is* weakened *by the sorting effect. The direct effect of σ_v^2 on the persistence and dispersion of ability is* strengthened *by the sorting effect.*

4.2 The Role of Luck and Prior Generations

The expected ability of an individual's descendants will clearly depend not only on that individual's ability, but also (and equally) on the ability of that individual's eventual spouse. Since an individual's prospects for attracting a high-ability spouse depend on the individual's *perceived* ability (with no independent effect of the individual's *actual* ability), the individual's *appearance* of high ability helps ensure the *reality* of high ability among the individual's descendants.

What shapes the perception of an individual's ability, holding fixed their actual ability? Our model reveals two related channels. The first is economic luck: good luck raises human capital and thus generates a more positive signal of underlying ability. The second is family status: a higher family status represents a higher prior belief about ability, producing a higher posterior belief following any given signal. Although logically independent, these channels are related insofar as one determinant of family status is the economic luck experienced by ancestors.

To clarify these channels, we first define *dynastic ability*, as the expected discounted sum of descendants' abilities:

$$V_{it}^{\theta} \equiv \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}_t \left[\theta_{i,t+\tau} \right], \tag{35}$$

The present discounted value of dynastic incomes (and thus utility, by (14)) can be expressed as an affine function of dynastic ability and parental income (see Corollary 6 in the appendix). In order to isolate the effect of luck and family status on an individual's dynastic ability, we focus on *conditional dynastic ability,* $\mathbb{E}_t\left[V_{it}^{\theta}|\theta_{it}\right]$.

Proposition 5 *Conditional dynastic ability is increasing in economic luck:*

$$\frac{d\mathbb{E}_t\left[V_{it}^{\theta}|\theta_{it}\right]}{d\varepsilon_{it}} = \frac{b\delta}{1-b\delta}\frac{1-\lambda}{2-b\delta\lambda}\cdot\frac{1}{\alpha_1} > 0. \tag{36}$$

This result may appear counter-intuitive: luck has no direct impact on the ability of agents or their descendants, and no indirect effect through income or investment, and yet luck has an equilibrium effect on the abilities of the descendants.

The magnitude of this effect is decreasing in the equilibrium importance of family background, λ. Intuitively, luck affects the signal and therefore the effect of luck is stronger when agents optimally place less weight on the prior relative to the signal. Since λ is increasing in σ_{ε}^2 (from Proposition 2), we have that the effect of luck on conditional dynastic ability is decreasing in σ_{ε}^2.

There are two competing effects associated with the return to ability, α_1. First, a higher return to ability lowers λ (from Proposition 2) and thus raises the sensitivity of dynastic ability to the signal by the previous argument. However, the final fraction reveals an offsetting direct effect: a higher return to ability makes the signal less sensitive to luck. The effect of α_1 on this marginal effect may be non-monotonic; given the parameters used in the previous figures, for instance, the luck sensitivity of the signal exhibits an inverse U-shape.

Proposition 6 *Conditional dynastic ability is increasing in family status. In particular,*

$$\frac{d\mathbb{E}_t\left[V_{it}^{\theta}|\theta_{it}\right]}{d\bar{\phi}_{it}} = \frac{b\delta}{1-b\delta}\frac{\lambda}{2-b\delta\lambda} > 0. \tag{37}$$

This result highlights an under-appreciated role of family background in shaping future fortunes. A higher family status $\bar{\phi}_{it}$ raises the economic prospects of descendants, in this case, purely by raising their expected ability (again, a characteristic that is not directly affected by family status and for which transmission is entirely passive). Thus, an individual's family status is a relevant determinant of future economic success, even if the individual's ability is held fixed, and even if parental income has no direct effect on offspring income (i.e. $\pi_2 = 0$).

The magnitude of this effect is increasing in the equilibrium importance of family background, λ. Intuitively, family status affects the prior, and will therefore have a stronger effect when agents optimally place more weight on the prior relative to the signal.

These results suggest a novel role for parents and grandparents. When singles experience economic luck, the expected ability of their future children is increased (Proposition 5). For instance, the expected ability of a child is increasing in shocks to their parent's income, even if we conditioned on father's (or mother's) ability. But shocks to the income of grandparents will also have an independent effect. This is because when singles experience luck, the family status enjoyed by children rises. Regardless of the actual ability realized by the individual's children, their greater family status allows them to attract a higher-ability spouse, and therefore to produce grandchildren with a higher expected ability (Proposition 6). This implies that the expected ability of a child will be increasing in the paternal grandparent's income, even after conditioning on the parents' income and the father's ability.

These sorts of predictions cannot arise in models where marital sorting on the ability dimension is is ignored or treated as exogenous, as is the case for previous models of inter-generational transmission. In such models, conditional on the father's ability, the mother's ability would not be correlated with the income of parents or grandparents, and so these income variables would not be conditionally correlated with the child's ability.

These results contribute to existing theory by proposing new incentives for parental investment. If the offspring's ability is affected by parental luck, then grandparents have an incentive to 'manufacture' economic luck by augmenting their human-capital investment in their children (i.e. the future parents). Although such investment cannot influence the ability of their offspring, it will influence the ability of their grandchildren (since, by the above logic, the investment will help their offspring attract higher-ability spouses). These insights will be formalized in the next section.

5 Optimal Parental Investment

We now consider the optimal investment problem facing each household, and verify that our proportional investment rule is indeed optimal. Recall that parental investment is not directly observed. As a result, parents have two motivations for investing: a standard one of raising the income-generating capacity of offspring (Becker and Tomes [1986, 1979]), and a novel one of manipulating the market's assessment of their offspring's ability. By raising the market's assessment, offspring are able to secure partners with higher expected ability. Since matching is assortative on human capital and parental income, this has no impact on offspring income. However, it *will* have an impact on the income of grandchildren (and subsequent generations) because it raises their expected ability.

The cost of investing is that consumption is reduced. If household (i, t) makes an expenditure on human-capital investment equal to proportion z_{it} of their income, then their consumption, in logs, is given by:

$$c_{it} = \ln(1 - z_{it}) + y_{it}. \tag{38}$$

The most direct benefit of investing is that investment raises the expected income of offspring. Recalling (3), the offspring will have a human capital equal to (in logs):

$$x_{i,t} = \alpha_0' + \alpha_1 \cdot \theta_{i,t} + \alpha_2 \cdot [\ln z_{it} + y_{it}] + \varepsilon_{i,t},$$

so that the expected income of offspring is therefore:

$$\mathbb{E}_t[y_{i,t+1}] = \pi_0' + \pi_1 \cdot \mathbb{E}_t[\theta_{i,t}] + \pi_2 \cdot y_{i,t} + \beta_1 \alpha_2 \cdot \ln z_{it}. \tag{39}$$

The indirect benefit of investing is that it will raise the status of offspring. The public observes human capital and parental income and has rational expectations about the investment that was made by each agents' parents. If the public expects an investment share of z_{it}^*, and a parent deviates from expectations by choosing a different investment share z_{it}, then the relevant signal is distorted:

$$s_{i,t} \equiv \frac{x_{i,t} - \alpha_0' - \alpha_2 \cdot y_{i,t}}{\alpha_1} = \theta_{i,t} + \xi_{i,t} + \frac{\alpha_2}{\alpha_1} \cdot (\ln z_{it} - \ln z_{it}^*).$$

The signal translates into individual status according to (18). That is:

$$\begin{aligned}
\mathbb{E}_t[\phi_{i,t}] &= \lambda \cdot \bar{\phi}_{i,t} + (1 - \lambda) \cdot \mathbb{E}_t[s_{i,t}] \\
&= \lambda \cdot \bar{\phi}_{i,t} + (1 - \lambda) \cdot \mathbb{E}_t[\theta_{i,t}] + (1 - \lambda) \cdot \frac{\alpha_2}{\alpha_1} \cdot (\ln z_{it} - \ln z_{it}^*)
\end{aligned}$$

Thus the signal translates into family status according to:

$$\begin{aligned}
\mathbb{E}_t[\bar{\phi}_{i,t+1}] &= b \cdot \mathbb{E}_t[\phi_{i,t}] \\
&= b\lambda \cdot \bar{\phi}_{i,t} + b(1 - \lambda) \cdot \mathbb{E}_t[\theta_{i,t}] + b(1 - \lambda) \frac{\alpha_2}{\alpha_1} \cdot (\ln z_{it} - \ln z_{it}^*).
\end{aligned} \tag{40}$$

Recall that household (i,t)'s payoffs are given by

$$U_{it} = c_{it} + \mathbb{E}_t \left[\sum_{\tau=1}^{\infty} \delta^\tau c_{i,t+\tau} \right].$$

Under the optimal investment strategy, the expectation term above is linear in $\mathbb{E}_t[y_{i,t+1}]$ and $\mathbb{E}_t[\bar{\phi}_{i,t+1}]$, as we show in the appendix. From (39) and (40) we see that these expectations are linear in $\ln z_{it}$. This, along with (38), tells us that household (i,t)'s investment problem boils down to a simple problem of the form:

$$\max_{z_{it}\in[0,1]} \left\{ \ln(1 - z_{it}) + \zeta_1 \cdot \ln z_{it} + \zeta_2 \cdot \ln z_{it} \right\}.$$

This expression allows a clear view of the relevant forces at play. The first term represents the cost of investment, whereas the second and third terms represent two distinct benefits: ζ_1 reflects the standard motivation to invest, based on raising the earning capacity of offspring, whereas ζ_2 reflects the new, information-driven motivation to invest to raise the offspring's status.

Proposition 7 *All families optimally invest the same fraction of their income:*

$$z_{it}^* = z^* = \frac{\zeta_1 + \zeta_2}{1 + \zeta_1 + \zeta_2},$$

where

$$\zeta_1 \equiv \frac{\delta \beta_1 \alpha_2}{1 - \delta[\beta_1 \alpha_2 + \beta_2]}$$
$$\zeta_2 \equiv \zeta_1 \cdot \frac{b\delta}{1 - b\delta} \cdot \frac{1 - \lambda}{2 - b\delta\lambda}.$$

The term ζ_1 represents the sort of incentives analyzed in standard models such as Becker and Tomes [1986, 1979]. The new force that we identify here is the ζ_2 term, and in particular the final component, $(1 - \lambda)/(2 - b\delta\lambda)$, which measures the 'status-based' incentive to invest. Straightforward calculation reveals that this term is decreasing in the relevance of family background, λ. This overall effect is composed of two opposing effects: a higher λ lowers investment incentives since status becomes less sensitive to investment efforts, however this is partially offset by the fact a higher status will persist for longer. We note that the 'status' motive for investment can be quantitatively important relative to the standard motive. That is, $\zeta_2 > \zeta_1$ for $b\delta$ large enough and λ small enough.

Since higher parental investment is associated theoretically and empirically with economic development, the analysis suggests a new mechanism through which economic development is hindered in societies where family background plays a central concern in the marriage market. Intuitively, in such cases it is difficult to shift the market's beliefs about offspring ability when the market places little weight on offspring performance relative to the prior.

Corollary 5 *Parental investment, z^*, is increasing in the return to ability, α_1, and is decreasing in the importance of luck, σ_ε^2.*

This follows from proposition 2 and the fact that ζ_2 is decreasing in λ. Note that the effect of these variables operates entirely via the new status channel that we identify. Indeed, a higher return to ability has no *direct* effect on investment incentives (investment does not raise ability directly, nor is complementary to it)–rather, a higher return to ability makes the market's beliefs about ability more sensitive to human capital (and thus parental investment efforts) in steady state. Similarly, a greater role for luck has no *direct* effect on investment incentives (such uncertainty is additive).

Finally, we note a role for policy. The status motive for investment is socially wasteful; thus there is excessive investment in human capital. The policy parameters $(\alpha_2, \beta_1, \beta_2)$ continue to have no impact on the 'status-based' incentive to invest (since they do not affect the quality of information) although they will of course have an effect on the standard 'income-based' incentive to invest. For instance, greater redistribution of income will lower the standard 'income-based' incentive to invest, and thus will act to raise welfare (whilst simultaneously raising equity).

6 Extensions

6.1 Imperfect Marital Sorting

In order to maintain our focus on equilibrium marital sorting on the ability dimension, it has been convenient that marital sorting on other dimensions (human capital, status and parental income) is characterized by marital segregation: perfect assortment of observable variable. In this section we consider a version in which frictions induce a departure from perfect segregation in the marriage market. Doing so affords us some empirical realism, allowing us to make predictions about other ways in which societies differ when they differ in the degree of marital sorting on human capital. The generalization also demonstrates that nothing in the main analysis hinges on perfect segregation and that our main results generalize naturally to more complex environments where the equilibrium is characterized by imperfect sorting.

Suppose that prior to marriage formation, agents observe a noisy measure of human capital, generated according to:

$$\hat{x}_{it} \equiv x_{it} + \nu_{it} \tag{41}$$

where $\nu_{it} \sim N(0, \sigma_\nu^2)$. This noisy view of human capital is then used to update prior beliefs about ability, $\bar{\psi}_{it}$, to form *interim* beliefs, $\hat{\psi}_{it}$. An individual's status at the time of matching is derived from the interim beliefs; since these beliefs incorporate the noise of the signal, perfect segregation on the observables generates imperfect segregation on actual human capital.[25] After the formation of marriages, human capital x_{it} is publicly observed and posterior beliefs, ψ_{it} are formed. These posterior beliefs then form the basis of the prior beliefs inherited by the next generation, as in the main model.

Essentially the same updating procedure as in the main model applies. Indeed, since human capital x_{it} is observed after marriages are formed, the market updates beliefs on the basis of x_{it} since the signal $\hat{x}_{it}$ does not provide any additional information about ability over-and-above that provided by x_{it}. As such, belief updating occurs as in the main model, and the quality of information converges to γ^{-1} in the steady state. However, it is the quality of information associated with *interim* beliefs that matters for sorting.

The quality of information associated with *interim* beliefs can be derived using a simple generalization of the above analysis. Since the noisy observation of human capital is given by

$$\hat{x}_{it} = \alpha_0 + \alpha_1\theta_{it} + \alpha_2 \cdot y_{i,t} + \varepsilon_{it} + \nu_{it}, \tag{42}$$

it follows that the relevant signal now includes an additional noise component:

$$\hat{s}_{it} \equiv \frac{\hat{x}_{it} - \alpha_0 - \alpha_2 \cdot y_{i,t}}{\alpha_1} = \theta_{it} + \frac{\varepsilon_{it} + \nu_{it}}{\alpha_1}. \tag{43}$$

[25] This approach (also used in Bidner [2010] for a similar purpose) is not the only way in which to generate imperfect sorting. For instance, we could instead have assumed (along the lines of Fernández et al. [2005]) that a proportion of families at each date are randomly matched (for exogenous reasons), and the remaining proportion are segregated as in the main model. Our approach seems more natural in our context given our central focus on the role of information.

The error component of this signal is

$$\hat{\xi}_{it} \equiv \frac{\varepsilon_{it} + \nu_{it}}{\alpha_1} \tag{44}$$

which has a variance of

$$\sigma^2_{\hat{\xi}} = \frac{\sigma^2_{\varepsilon} + \sigma^2_{\nu}}{\alpha^2_1}. \tag{45}$$

Similar arguments to those in the main model apply here, so that the steady-state quality of information, $\hat{\gamma}^{-1}$, is derived from:

$$\hat{\gamma} \equiv \frac{\sigma^2_{\hat{\xi}}[\frac{b^2}{2} \cdot \gamma + \sigma^2_{\upsilon}]}{\sigma^2_{\hat{\xi}} + \frac{b^2}{2} \cdot \gamma + \sigma^2_{\upsilon}}. \tag{46}$$

Since $\sigma^2_{\hat{\xi}} > \sigma^2_{\xi}$ whenever there is noise ($\sigma^2_{\nu} > 0$), it naturally follows that the quality of information deteriorates (i.e. $\hat{\gamma}^{-1} < \gamma^{-1}$) in such cases. The expressions for husband-wife ability correlation, the variance of ability and the parent-child ability correlation are the same as those derived before, with γ replaced with $\hat{\gamma}$. In short, the new element introduced by imperfectly observed human capital is reflected in $\sigma^2_{\hat{\xi}}$.

Since this departure from perfect segregation has no qualitative impact on our model, the main results derived earlier also apply to a more realistic world with imperfect sorting on observables. For instance, the institutional/policy parameters still exert no influence on the strength of ability sorting or persistence. Quantitatively, the noise on human capital reduces the quality of information, and this in turn weakens the ability correlation within husband-wife and parent-child pairs.

The new empirical prediction arising from this generalization is that societies with a stronger husband-wife human capital correlation will have a stronger husband-wife ability correlation and thus a stronger parent-child ability correlation. In terms of income, this implies that societies with strong husband-wife human capital sorting will also be societies with low income mobility (high parent-child income correlations). This latter prediction is shared by models in which there is no unobserved ability (e.g. Fernández et al. [2005], Kremer [1998]). Unlike these models, the prediction would remain in our setting even in the absence of parental human-capital investments.

6.2 When Do Policy Variables Matter?

Policy variables will not affect the distribution of ability unless they affect the steady state quality of information. One way to generalize the model to achieve this is to suppose that parental contributions are a stochastic function of parental investment. That is, allocating a proportion z of income to human capital investment translates into an effective (log) contribution of

$$h_{it} = \ln z + y_{it} + \varepsilon_{it}^h \tag{47}$$

where $\varepsilon_{it}^h \sim N(0, \sigma_{\varepsilon^h}^2)$. The human capital of offspring is therefore:

$$x_{it} = \alpha_0 + \alpha_1 \cdot \theta_{it} + \alpha_2 \cdot y_{i,t} + \alpha_2 \cdot \varepsilon_{it}^h + \varepsilon_{it}, \tag{48}$$

where, as before, $\alpha_0 \equiv \alpha_0' + \alpha_2 \cdot z$. By defining $\varepsilon_{it}^x \equiv \alpha_2 \cdot \varepsilon_{it}^h + \varepsilon_{it}$ as 'aggregate' luck, we can express human capital in a generalized form:

$$x_{it} = \alpha_0 + \alpha_1 \cdot \theta_{it} + \alpha_2 \cdot y_{it} + \varepsilon_{it}^x. \tag{49}$$

This generalization is useful because our results need only be adjusted by replacing ε_{it} with ε_{it}^x. In particular, when considering $\sigma_{\varepsilon^h}^2 > 0$, the variance of human capital luck generalizes beyond σ_ε^2 to $\sigma_{\varepsilon^x}^2 \equiv \sigma_\varepsilon^2 + \alpha_2^2 \cdot \sigma_{\varepsilon^h}^2$. In particular, we have the following.

Proposition 8 *If parental investment has a stochastic component (i.e. if $\sigma_{\varepsilon^h}^2 > 0$), a greater equality of opportunity (i.e. a lower α_2) will raise the steady state variation in ability (σ_θ^2), the husband-wife ability correlation (ρ_θ^{HW}), and the parent-child ability correlation (ρ_θ^{PC}).*

The intuition is that stochastic parental investment adds a layer of noise to observed human capital. This added noise acts to weaken sorting on ability and thus reduce the variability and persistence of ability. However, as human capital is made less sensitive to parental investment, this additional noise layer is diminished and sorting is strengthened.

This result is counter-intuitive: by making offspring human capital less sensitive to parental investment, offspring ability becomes *more* closely related to parental ability. The main implication of this is that (observed) changes in the intergenerational persistence of income will not reliably track (unobserved) changes in the intergenerational persistence of ability. For instance a reduction in α_2 will raise ability persistence but can lower income persistence. This is because of the direct effect of human capital being less sensitive to parental investment. On the other hand, the two persistence outcomes will shift in the same direction in response to changes in α_1 for example.

7 Conclusions

This paper contributes to the theory of social mobility by proposing a model of the transmission of innate ability under incomplete information. The marriage market plays a central role because ability depends equally on both parents. A satisfactory model of such a marriage market must contend with the unobserved nature of ability and its mutability across generations, and with the way in which beliefs are shaped by observed characteristics of participants as well as their family histories. In its full generality, this problem would require a model of multi-dimensional matching with incomplete information; our model circumvents the daunting complexity of the general approach by generating an equilibrium with perfect sorting on observables.

Our key theoretical result is that a simple scalar variable, family status, turns out to be a sufficient statistic for the information revealed by countless generations of family history, encompassing ancestors from countless lineages. We show that this variable follows over time a simple linear law of motion. We then show precisely how inequality and mobility of both income and ability are determined by the marital sorting on ability, which depends in turn on the precision of the information encoded by family status and by the noisy signals derived from current outcomes. We also show, by extending the model to allow for noisy observations, that these results do not rely on perfect sorting.

Our model reveals a new indirect channel through which various variables influence inequality and social mobility. Any parameter shift that raises the equilibrium quality of information will strengthen husband-wife sorting on ability and therefore also strengthen the parent-child ability correlation and raise the dispersion of ability in the population. Conversely, redistribution policies which did not alter information had no effect on the distribution or transmission of ability in our model, a result that echoes the 'policy invariance' of mobility as reported by Clark [2014].

The role of the family as a repository of information implies that idiosyncratic shocks to current status will affect the income of descendants; the actual ability of descendants will be sensitive to the *perceived* ability of the ancestor, independently of the ancestor's *actual* ability. This logic suggests a new link between the fortunes of different generations, and also introduces a status-based motive for parental investment in offspring.

The model rests on various simplifying assumptions in order to gain tractability;

relaxing some of these offers scope for additional insight in future work. For instance, allowing for asymmetries by gender, embedding fertility decisions, and adding an explicit political economy layer (which influences and is influenced by the distribution of ability), and incorporating other matching frictions all seem promising avenues.

We are hopeful that the model will prove useful for more direct empirical work in the near future. The model makes clear predictions about the distribution of ability that are testable in principal. The key barrier is obtaining reliable, large-scale data on the innate ability of individuals, their spouse and children. To the extent that ability is interpreted as being genetic in nature, the increasing availability of genetic markers in large datasets will, we hope, soon fulfill this requirement.

References

S. Rao Aiyagari, Jeremy Greenwood, and Nezih Guner. On the state of the union. *Journal of Political Economy*, 108(2):213–244, 2000.

Axel Anderson. A dynamic generalization of becker's assortative matching result. *Journal of Economic Theory*, 159(PA):290–310, 2015. URL https://EconPapers.repec.org/RePEc:eee:jetheo:v:159:y:2015:i:pa:p:290-310.

Axel Anderson and Lones Smith. Dynamic matching and evolving reputations. *The Review of Economic Studies*, 77(1):3–29, 2010. doi: 10.1111/j.1467-937X.2009.00567.x. URL +http://dx.doi.org/10.1111/j.1467-937X.2009.00567.x.

Kenneth Arrow, Samuel Bowles, and Steven Durlauf. *Introduction*, chapter 0, pages ix–xv. Princeton University Press, 2000. ISBN 0-691-00467-6.

Gary Becker and Nigel Tomes. Human capital and rise and fall of families. *Journal of Labor Economics*, 4:S1–S39, 1986.

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

Roland Benabou. Heterogeneity, stratification, and growth: Macroeconomic implications of community structure and school finance. *American Economic Review*, 86(3): 584–609, June 1996.

Roland Benabou. Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica*, 70(2):481–

517, 2002. ISSN 1468-0262. doi: 10.1111/1468-0262.00293. URL http://dx.doi.org/10.1111/1468-0262.00293.

Ted Bergstrom and Mark Bagnoli. Courtship as a waiting game. *Journal of Political Economy*, 101(1):185–202, 1993. URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:101:y:1993:i:1:p:185-202.

Chris Bidner. Pre-match investment with frictions. *Games and Economic Behavior*, 68(1):23 – 34, 2010. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2009.06.004. URL http://www.sciencedirect.com/science/article/pii/S0899825609001286.

Chris Bidner, Guillaume Roger, and Jessica Moses. Investing in skill and searching for coworkers: Endogenous participation in a matching market. *American Economic Journal: Microeconomics*, 8(1):166–202, 2016. doi: 10.1257/mic.20140110. URL http://www.aeaweb.org/articles.php?doi=10.1257/mic.20140110.

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2):298–319, April 2001. URL http://ideas.repec.org/a/eee/jetheo/v97y2001i2p298-319.html.

R. Boyd and P.J. Richerson. *The Origin and Evolution of Cultures*. Evolution and Cognition. Oxford University Press, 2005. ISBN 9780199883127. URL https://books.google.ca/books?id=LoZMCAAAQBAJ.

L.L. Cavalli-Sforza and M.W. Feldman. *Cultural Transmission and Evolution: A Quantitative Approach*. Monographs in population biology. Princeton University Press, 1981. ISBN 9780691082837. URL https://books.google.ca/books?id=pBGvyNXkWYcC.

Hector Chade. Matching with noise and the acceptance curse. *Journal of Economic Theory*, 127(1):81–113, July 2006.

Gregory Clark. *The Son Also Rises: Surnames and the History of Social Mobility*. Princeton University Press, Princeton,N.J., 2014.

Harold L. Cole, George J. Mailath, and Andrew Postlewaite. Social norms, savings behavior, and growth. *Journal of Political Economy*, 100:1092–1125, 1992.

Harold L. Cole, George J. Mailath, and Andrew Postlewaite. Incorporating concern for relative wealth into economic models. *Quarterly Review*, 1932(Sum):12–21, 1995.

David Comerford, Jose V Rodriguez Mora, and Michael J Watts. The rise of meritocracy and the inheritance of advantage. Working Papers 1716, University of Strathclyde Business School, Department of Economics, July 2017. URL https://ideas.repec.org/p/str/wpaper/1716.html.

Richard Dawkins. *The Selfish Gene*. Oxford University Press, 30th anniversary edition, 2006. ISBN 6610870179.

Matthias Doepke and Fabrizio Zilibotti. Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5):1331–1371, 2017. doi: 10.3982/ECTA14634. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA14634.

Steven N Durlauf. A theory of persistent income inequality. *Journal of Economic Growth*, 1(1):75–93, March 1996.

Raquel Fernandez and Richard Rogerson. Income distribution, communities, and the quality of public education. *The Quarterly Journal of Economics*, 111(1):135–64, February 1996.

Raquel Fernández, Nezih Guner, and John Knowles. Love and money: A theoretical and empirical analysis of household sorting and inequality. *The Quarterly Journal of Economics*, 120(1):273–344, 2005.

Patrick Francois and Jan Zabojnik. Trust, social capital, and economic development. *Journal of the European Economic Association*, 3(1):51–94, 03 2005. URL http://ideas.repec.org/a/tpr/jeurec/v3y2005i1p51-94.html.

Richard J. Herrnstein and Charles Murray. *The Bell Curve*. The Free Press, New York, 1994.

Ed Hopkins. Job market signaling of relative position, or becker married to spence. *Journal of the European Economic Association*, 10(2):290–322, 2012. ISSN 1542-4774. doi: 10.1111/j.1542-4774.2010.01047.x. URL http://dx.doi.org/10.1111/j.1542-4774.2010.01047.x.

Heidrun C. Hoppe, Benny Moldovanu, and Aner Sela. The theory of assortative matching based on costly signals. *The Review of Economic Studies*, 76(1):253–281, 2009.

Michael Kremer. How much does sorting increase inequality? *Quarterly Journal of Economics*, 113(4):1137–1167, 1998.

Sang Yoon (Tim) Lee and Ananth Seshadri. On the intergenerational transmission of economic status. *Journal of Political Economy*, 2019. doi: 10.1086/700765. URL https://doi.org/10.1086/700765. forthcoming.

Charles Murray. *Coming Apart: The State of White America, 1960-2010*. Crown Publishing Group, 2012. ISBN 9780307453440. URL https://books.google.ca/books?id=o6fUPokvyR0C.

Aysu Okbay, Jonathan P. Beauchamp, and Mark Alan Fontana. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature Letters*, 533(7604):539–542, 2016.

Tinca J C Polderman, Beben Benyamin, and Christiaan A de Leeuw. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47:702–709, 2015.

Mari Rege. Why do people care about social status? *Journal of Economic Behavior & Organization*, 66(2):233–242, May 2008.

Bruce Sacerdote. Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees? In Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, editors, *Handbook of Social Economics*, volume 1, pages 1 – 30. North-Holland, 2011. doi: https://doi.org/10.1016/B978-0-444-53187-2.00001-2. URL http://www.sciencedirect.com/science/article/pii/B9780444531872000012.

Gary Solon. *A Model of Intergenerational Mobility Variation over Time and Place*, pages 38–47. Cambridge University Press, Cambridge, 01 2004.

Guido Tabellini. The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, 123(3):905–950, August 2008. URL http://ideas.repec.org/a/tpr/qjecon/v123y2008i3p905-950.html.

Michael Young. *The Rise of the Meritocracy 1870-2033: an Essay on Education and Society*. Thames and Hudson, London, 1958.

Paul J. Zak and Kwang Woo Park. Population genetics and economic growth. *Journal of Bioeconomics*, 4(1):1–38, Jan 2002. ISSN 1573-6989. doi: 10.1023/A:1020604724888. URL https://doi.org/10.1023/A:1020604724888.

A Structural Model

In this section we develop the structural model that underlies the reduced-form equations describing human capital and income (i.e. equations (1), (3), and (13)).

A.1 Human Capital

Let X_{it} denote the human capital of a single agent of family i of generation t. We assume that X_{it} is determined by the contribution of ability, $G_{it} \equiv \exp(\theta_{it})$, effective parental inputs, $P_{i,t-1}$, and luck, $\varepsilon_{it} \sim N\left(0, \sigma_{\varepsilon}^2\right)$:

$$X_{it} = A_1 \cdot G_{it}^{\alpha_1} \cdot P_{i,t-1}^{\chi} \cdot \exp(\varepsilon_{it}),$$

where A_1, α_1 and χ are parameters with positive values. Effective parental inputs depend on parental investment, $H_{i,t-1}$, and a public input, $\hat{P}_{i,t-1}$:

$$P_{i,t-1} \equiv H_{i,t-1}^{\chi_1} \hat{P}_{i,t-1}^{\chi_2}.$$

Parental investment depends on household income, $Y_{i,t-1}$, and an idiosyncratic component, $\varepsilon_{it}^h \sim N(0, \sigma_{\varepsilon^h}^2)$:

$$H_{it} = z \cdot Y_{i,t-1} \cdot \exp(\varepsilon_{it}^h),$$

where $z \in [0, 1]$. In the main model, z is exogenous and $\sigma_{\varepsilon^h}^2 = 0$ for all (i, t). We endogenize z and allow $\sigma_{\varepsilon^h}^2 > 0$ in extensions. The public input available to i depends on their parental investment according to a policy parameter, σ:

$$\hat{P}_{i,t-1} = \frac{H_{i,t-1}^{1-\sigma}}{\int H_{j,t-1}^{1-\sigma} dj} \cdot B_t,$$

where B_t is the total available public input (the value of which does not matter since it will be absorbed into the constant). Higher values of σ correspond to more progressive distributions of the public input. For instance $\sigma = 1$ corresponds to equal access, $\sigma < 1$ implies a regressive system whereby access to public inputs is biased toward those making greater parental investment, and $\sigma > 1$ implies a progressive system whereby the bias is toward those making lower parental investments. Together then we have:

$$X_{it} = A_1 \cdot G_{it}^{\alpha_1} \cdot \tilde{B}_{t-1}^{\chi} \cdot [H_{i,t-1}^{\chi_1 + \chi_2 \cdot (1-\sigma)}]^{\chi} \cdot \exp(\varepsilon_{it}),$$

where $\tilde{B}_{t-1} \equiv B_t \cdot [\int H_{j,t-1}^{1-\sigma} dj]$.

By taking logs we get the following reduced form relationship:

$$x_{it} = \alpha_{0t} + \alpha_1 \cdot \theta_{it} + \alpha_2 \cdot y_{i,t-1} + \varepsilon^x_{it}, \tag{50}$$

where

$$\alpha_2 \equiv \chi \cdot [\chi_1 + \chi_2 \cdot (1 - \sigma)], \tag{51}$$

$\varepsilon^x_{it} \equiv \varepsilon_{it} + \alpha_2 \cdot \varepsilon^h_{it}$ and $\alpha_{0t} \equiv \ln A_1 + \chi\chi_2 \cdot \ln \tilde{B}_{t-1} + \alpha_2 \ln z$. The value of α_{0t} depends only on the distribution of income, and thus will be a constant, α_0, in the steady state. If access to public inputs is sufficiently progressive, then greater parental investment will actually reduce offspring human capital. To avoid such cases we restrict attention to policies for which $\sigma \leq \bar{\sigma} \equiv 1 + (\chi_1/\chi_2)$. Here we see that more progressive access to public resources (higher σ) implies a reduction in α_2. This in turn lowers the variance of ε^x_{it}.

A.2 Income

Household income, Y_{it}, results from the redistribution of pre-tax income, Y^{pre}_{it}:

$$Y_{it} = [Y^{\text{pre}}_{it}]^{1-\tau}\, \hat{Y}_t$$

where $\tau \in [0, 1]$ parameterizes *income redistribution* (Benabou (2002)), and $\hat{Y}_t$ ensures that the resource constraint holds:

$$\int Y_{it} di = \int Y^{\text{pre}}_{it} di.$$

That is,

$$\hat{Y}_t = \frac{\int Y^{\text{pre}}_{it} di}{\int [Y^{\text{pre}}_{it}]^{1-\tau}\, di}.$$

Pre-tax income, Y^{pre}_{it}, depends on household output, Q_{it}, and parental income, $\bar{Y}_{i,t-1}$:

$$Y^{\text{pre}}_{it} = Q^{\mu}_{it} \bar{Y}^{1-\mu}_{i,t-1} \cdot \hat{Y}^{\text{pre}}_t$$

where μ parameterizes *meritocracy* and $\hat{Y}_t^{\text{pre}}$ ensures that the resource constraint holds:

$$\int Y_{it}^{\text{pre}} di = \int Q_{it} di.$$

That is,

$$\hat{Y}_t^{\text{pre}} = \frac{\int Q_{it} di}{\int Q_{it}^{1-\mu} di}.$$

Parental income, $\bar{Y}_{i,t-1}$, depends on the parental income of both household members:

$$\bar{Y}_{i,t-1} = Y_{i,t-1}^{1/2} Y_{i',t-1}^{1/2}.$$

Finally, household output, Q_{it}, is produced by the human capital of both household members:

$$Q_{it} = X_{it}^{1/2} X_{i't}^{1/2}$$

Taking logs gives us the following relationship:

$$y_{it} = \beta_{0t} + \beta_1 \cdot [x_{i,t} + x_{i',t}]/2 + \beta_2 \cdot [y_{i,t-1} + y_{i',t-1}]/2,$$

where $\beta_{0t} \equiv \ln \hat{Y}_t + (1-\tau) \cdot \ln \hat{Y}_t^{\text{pre}}$, $\beta_1 \equiv (1-\tau) \cdot \mu$, and $\beta_2 \equiv (1-\tau) \cdot (1-\mu)$. The value of β_{0t} depends only on the distribution of income and output, and thus will be a constant, β_0, in the steady state.

B Marriage Market Details

The offspring from households at date t are characterized by their *marriage type* $\omega_{it} \equiv \{x_{it}, y_{it}, \psi_{it}\}$. The marriage market is described by a matching set, which we require to be feasible and stable. Formally, let Ω denote the set of possible marriage types. Define a *matching set* M as a subset of Ω^2 such that $M_m(A) \equiv \{\omega_f | \omega_m \in A \text{ and } (\omega_m, \omega_f) \in M\}$ and $M_f(A) \equiv \{\omega_m | \omega_f \in A \text{ and } (\omega_m, \omega_f) \in M\}$ are non-empty for each $A \subseteq \Omega$. A matching set describes how males and females are to be paired on the basis of their marriage types. Males with marriage types in A are to be paired with females who have marriage types in $M_m(A)$, and similarly females with marriage types in A are to be paired with males who have marriage types in $M_f(A)$. A matching set is *feasible* if the measure of males with marriage types in A equals the measure of females with marriage types in $M_m(A)$ and the measure of females with marriage types in A equals the measure of males with marriage types in $M_f(A)$ for each measurable $A \subseteq \Omega$. A matching set is *stable* if no unmatched pair prefer to marry than remain with their assigned partner.

Segregation requires that agents marry someone with identical characteristics: $M_m(A) = A$ and $M_f(A) = A$ for all $A \subseteq \Omega$. Segregation is trivially feasible, and we prove below (section D.1.1) that it is stable.

C Additional Results

C.1 Deriving Correlations

Proof of Lemma 1. Decompose ability as follows:

$$\theta_{it} = \mathbb{E}[\theta_{it}|\psi_{it}] + (\theta_{it} - \mathbb{E}[\theta_{it}|\psi_{it}]) \tag{52}$$

so that

$$\theta_{it}^2 = \mathbb{E}[\theta_{it}|\psi_{it}]^2 + (\theta_{it} - \mathbb{E}[\theta_{it}|\psi_{it}])^2 + 2\mathbb{E}[\theta_{it}|\psi_{it}](\theta_{it} - \mathbb{E}[\theta_{it}|\psi_{it}]) \tag{53}$$

and

$$\begin{aligned} \theta_{it}\theta_{i't} =& \mathbb{E}[\theta_{it}|\psi_{it}]\mathbb{E}[\theta_{i't}|\psi_{i't}] + (\theta_{it} - \mathbb{E}[\theta_{it}|\psi_{it}])(\theta_{i't} - \mathbb{E}[\theta_{i't}|\psi_{i't}]) & (54)\\ &+ \mathbb{E}[\theta_{it}|\psi_{it}](\theta_{i't} - \mathbb{E}[\theta_{i't}|\psi_{i't}]) + \mathbb{E}[\theta_{i't}|\psi_{i't}](\theta_{it} - \mathbb{E}[\theta_{it}|\psi_{it}]) & (55) \end{aligned}$$

Taking expectations of (53), applying the law of iterated expectations,[26] and applying the definition of γ_{it} gives

$$\mathbb{E}[\theta_{it}^2] = \mathbb{E}[\mathbb{E}[\theta_{it}|\psi_{it}]^2 + \gamma_{it}]. \tag{56}$$

Taking expectations of (54), applying the law of iterated expectations and the conditional independence of spouse abilities to get

$$\begin{aligned} \mathbb{E}[\theta_{it}\theta_{i't}] &= \mathbb{E}[\mathbb{E}[\theta_{it}|\psi_{it}]\mathbb{E}[\theta_{i't}|\psi_{i't}]] & (57)\\ &= \mathbb{E}[\mathbb{E}[\theta_{it}|\psi_{it}]^2] & (58) \end{aligned}$$

[26]That is, use the fact that $\mathbb{E}[\mathbb{E}[\theta_{it}|\psi_{it}](\theta_{it} - \mathbb{E}[\theta_{it}|\psi_{it}])]$ equals $\mathbb{E}[\mathbb{E}[\mathbb{E}[\theta_{it}|\psi_{it}](\theta_{it} - \mathbb{E}[\theta_{it}|\psi_{it}])|\psi_{it}]]$ which is zero.

where the final equality is an implication of segregation. It then follows that

$$\mathbb{E}[\theta_{it}^2] = \mathbb{E}[\theta_{it}\theta_{i't}] + \mathbb{E}[\gamma_{it}]. \tag{59}$$

Since $\mathbb{E}[\theta_{it}] = \mathbb{E}[\theta_{i't}] = 0$, it then follows that in the steady state:

$$\begin{aligned} \sigma_\theta^2 &= \mathrm{Cov}[\theta_{it}\theta_{i't}] + \mathbb{E}[\gamma_{it}] && (60) \\ &= \rho_\theta^{HW} \cdot \sigma_\theta^2 + \gamma, && (61) \end{aligned}$$

where the final equality uses the fact that in the steady state we have $\gamma_{it} = \gamma$ and $\mathrm{Var}[\theta_{it}] = \mathrm{Var}[\theta_{i't}] = \sigma_\theta^2$. Simple manipulation yields the expression given in Lemma 1
□

C.2 Ability

Proof of Proposition 3. The result follows once we establish (27) and (28) in the text. To this end, notice that in the steady state $\mathbb{E}[\theta_{it}] = 0$. Recalling that if r_1 and r_2 are mean-zero random variable then $\mathrm{Cov}(r_1, r_2) = \mathbb{E}[r_1 r_2]$, we can use (2) to get:

$$\mathrm{Cov}(\theta_{it}, r_{it}) = \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i,t-1}, r_{it}) + \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i',t-1}, r_{it}) + \mathrm{Cov}(\upsilon_{i,t}, r_{it}) \tag{62}$$

where r_{it} is any mean-zero random variable. Thus, we have the following system:

$$\begin{aligned} \mathrm{Cov}(\theta_{it}, \theta_{it}) &= \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i,t-1}, \theta_{it}) + \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i',t-1}, \theta_{it}) + \sigma_\upsilon^2 && (63) \\ \mathrm{Cov}(\theta_{it}, \theta_{i,t-1}) &= \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i,t-1}, \theta_{i,t-1}) + \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i',t-1}, \theta_{i,t-1}) && (64) \\ \mathrm{Cov}(\theta_{it}, \theta_{i',t-1}) &= \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i,t-1}, \theta_{i',t-1}) + \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i',t-1}, \theta_{i',t-1}) && (65) \end{aligned}$$

Using the steady state conditions $\mathrm{Cov}(\theta_{it}, \theta_{it}) = \mathrm{Cov}(\theta_{i,t-1}, \theta_{i,t-1}) = \mathrm{Cov}(\theta_{i',t-1}, \theta_{i',t-1}) = \sigma_\theta^2$ and $\mathrm{Cov}(\theta_{i,t-1}, \theta_{i',t-1}) = \mathrm{Cov}(\theta_{i,t}, \theta_{i',t})$, and the symmetry property $\mathrm{Cov}(\theta_{it}, \theta_{i',t-1}) = \mathrm{Cov}(\theta_{it}, \theta_{i,t-1})$, simplifies this to:

$$\begin{aligned} \sigma_\theta^2 &= b \cdot \mathrm{Cov}(\theta_{i,t-1}, \theta_{it}) + \sigma_\upsilon^2 && (66) \\ \mathrm{Cov}(\theta_{it}, \theta_{i,t-1}) &= \frac{b}{2} \cdot \sigma_\theta^2 + \frac{b}{2} \cdot \mathrm{Cov}(\theta_{i',t}, \theta_{i,t}) && (67) \end{aligned}$$

These tell us the ability variance and the intergenerational ability covariance as a function of the spousal covariance. Dividing both by σ_θ^2 and solving yields (27) and (28) in the text. □

C.2.1 Multi-Generational Ability Correlations

Using the ability transmission equation and symmetry gives, for $k = 2, 3, ...$:

$$\text{Cov}(\theta_{it}, \theta_{i,t-k}) = b \cdot \text{Cov}(\theta_{it}, \theta_{i,t-(k-1)}). \tag{68}$$

Using the steady state conditions, the ability correlation between family members $k \in \{2, 3, ..\}$ generations apart is thus given by:

$$\rho_{\theta,k}^{PC} \equiv \frac{\text{Cov}(\theta_{it}, \theta_{i,t-k})}{\sigma_\theta^2} = b^{k-1} \cdot \rho_\theta^{PC} = b^k \cdot \left[\frac{\sigma_\upsilon^2 - \frac{\gamma}{2}}{\sigma_\upsilon^2 - b^2 \cdot \frac{\gamma}{2}} \right]. \tag{69}$$

Proposition 9 *The multi-generation ability correlation implied by extrapolating the parent-child ability correlation always* understates *the true multi-generation ability correlation:* $(\rho_\theta^{PC})^k < \rho_{\theta,k}^{PC}$ *for all* $k = 2, 3, ...$

Proof. This is a straightforward consequence of:

$$\frac{(\rho_\theta^{PC})^k}{\rho_{\theta,k}^{PC}} = \left[\frac{\sigma_\upsilon^2 - \frac{\gamma}{2}}{\sigma_\upsilon^2 - b^2 \cdot \frac{\gamma}{2}} \right]^{k-1} \in (0, 1). \tag{70}$$

□

The extent of the bias increases in k as this ratio goes to zero as k increases. Furthermore this bias is endogenous in our setting, as the ratio is decreasing in γ.

The divergence between extrapolated and actual long correlations is in this case due entirely to the omission of the other parent's ability. Intuitively, the one-generation correlation does not capture the fact that both the parent and offspring abilities are positively correlated with the other parent's ability.

C.3 Income

Proof of Proposition 4. Notice that in the steady state $\mathbb{E}[\theta_{it}] = 0$ and $\mathbb{E}[y_{it}] = \frac{\pi_0}{1-\beta_2}$. For what follows, take y to be the de-meaned counterpart (to save on notation). In de-meaned terms, we have

$$y_{i,t+1} = \pi_1 \cdot \theta_{i,t} + \pi_2 \cdot y_{i,t} + \varepsilon_{it}^y. \tag{71}$$

Using a method identical to that for ability, we get the following system:

$$\text{Cov}(y_{i,t+1}, y_{i,t+1}) = \pi_1 \cdot \text{Cov}(\theta_{it}, y_{i,t+1}) + \pi_2 \cdot \text{Cov}(y_{i,t}, y_{i,t+1}) + \sigma^2_{\varepsilon^y} \tag{72}$$

$$\text{Cov}(y_{i,t+1}, \theta_{it}) = \pi_1 \cdot \text{Cov}(\theta_{it}, \theta_{it}) + \pi_2 \cdot \text{Cov}(y_{i,t}, \theta_{it}) \tag{73}$$

$$\text{Cov}(y_{i,t+1}, y_{i,t}) = \pi_1 \cdot \text{Cov}(\theta_{it}, y_{i,t}) + \pi_2 \cdot \text{Cov}(y_{i,t}, y_{i,t}). \tag{74}$$

From (2) and $\text{Cov}(\theta_{i',t-1}, y_{i',t}) = \text{Cov}(\theta_{i,t-1}, y_{i,t})$ we have:

$$\text{Cov}(\theta_{it}, y_{i,t}) = \frac{b}{2} \cdot \text{Cov}(\theta_{i,t-1}, y_{i,t}) + \frac{b}{2} \cdot \text{Cov}(\theta_{i',t-1}, y_{i,t}) \tag{75}$$

$$= b \cdot \text{Cov}(\theta_{i,t-1}, y_{i,t}). \tag{76}$$

Applying the steady state conditions gives the following system:

$$\text{Cov}(y_{i,t}, y_{i,t}) = \pi_1 \cdot \text{Cov}(\theta_{it}, y_{i,t+1}) + \pi_2 \cdot \text{Cov}(y_{i,t}, y_{i,t-1}) + \sigma^2_{\varepsilon^y} \tag{77}$$

$$\text{Cov}(y_{i,t+1}, \theta_{it}) = \pi_1 \cdot \text{Cov}(\theta_{it}, \theta_{it}) + \pi_2 \cdot \text{Cov}(y_{i,t}, \theta_{it}) \tag{78}$$

$$\text{Cov}(y_{i,t}, y_{i,t-1}) = \pi_1 \cdot \text{Cov}(\theta_{it}, y_{i,t}) + \pi_2 \cdot \text{Cov}(y_{i,t}, y_{i,t}) \tag{79}$$

$$\text{Cov}(\theta_{it}, y_{i,t}) = b \cdot \text{Cov}(\theta_{i,t}, y_{i,t+1}). \tag{80}$$

Solving gives us the two covariances of interest:

$$\text{Cov}(y_{it}, y_{it}) \equiv \sigma^2_y = \frac{\left(\frac{1+b\pi_2}{1-b\pi_2}\right)\pi_1^2 \cdot \sigma^2_\theta + \sigma^2_{\varepsilon^y}}{1-\pi_2^2} \tag{81}$$

$$\text{Cov}(y_{it}, y_{i,t-1}) = \frac{\left(\frac{b+\pi_2}{1-b\pi_2}\right)\pi_1^2 \cdot \sigma^2_\theta + \pi_2 \cdot \sigma^2_{\varepsilon^y}}{1-\pi_2^2}, \tag{82}$$

where σ^2_θ was derived above in the ability section. The correlation of interest is:

$$\rho_y^{PC} \equiv \frac{\text{Cov}(y_{it}, y_{i,t-1})}{\text{Cov}(y_{it}, y_{it})} = \frac{\left(\frac{b+\pi_2}{1-b\pi_2}\right)\pi_1^2 \cdot \sigma^2_\theta + \pi_2 \cdot \sigma^2_{\varepsilon^y}}{\left(\frac{1+b\pi_2}{1-b\pi_2}\right)\pi_1^2 \cdot \sigma^2_\theta + \sigma^2_{\varepsilon^y}} \tag{83}$$

$$= \pi_2 + b \cdot \Phi(\gamma), \tag{84}$$

where $\Phi(\gamma)$ is defined as

$$\Phi(\gamma) \equiv \frac{(1-\pi_2^2) \cdot [\pi_1^2 \cdot \sigma^2_\theta(\gamma)]}{(1+b\pi_2) \cdot [\pi_1^2 \cdot \sigma^2_\theta(\gamma)] + (1-b\pi_2) \cdot [\sigma^2_{\varepsilon^y}]}. \tag{85}$$

The result follows from (81) and (83). $\square$

Here we see that income would be persistent *even if income did not depend on parental income* (i.e. if $\pi_2 = 0$).

Since Φ is increasing in γ^{-1}, we have that α_1 raises the persistence of income both directly (via π_1) and indirectly (via sorting, γ). Similarly, σ_ε^2 lowers the persistence of income both directly (via $(\sigma_{\varepsilon^y}^2)$ and indirectly (via sorting, γ).

We also see that ρ_y^{PC} depends on π_2, whereas ρ_θ^{PC} was independent of π_2. Thus, policy that affects the sensitivity of human capital to parental inputs or meritocracy in general, will have an effect on income mobility but will have no impact on ability mobility. As such, (i) changes in persistence of observed characteristics need not be informative about changes in the persistence of unobserved characteristics, and (ii) the effect of such policy will be limited by the fact that income will persist even if parental income has no direct effect on income.

C.3.1 Multi-Generational Income Correlations

To work out $\text{Cov}(y_{it}, y_{i,t-k})$ for $k = 2, 3, ...$ note:

$$\begin{aligned}\text{Cov}(y_{it}, y_{i,t-k}) &= \pi_1 \cdot \text{Cov}(\theta_{it}, y_{i,t-k}) + \pi_2 \cdot \text{Cov}(y_{i,t-1}, y_{i,t-k}) &(86)\\ &= \pi_1 \cdot \text{Cov}(\theta_{it}, y_{i,t-k}) + \pi_2 \cdot \text{Cov}(y_{i,t}, y_{i,t-(k-1)}) &(87)\end{aligned}$$

and

$$\begin{aligned}\text{Cov}(\theta_{it}, y_{i,t-k}) &= b \cdot \text{Cov}(\theta_{i,t-1}, y_{i,t-k}) &(88)\\ &= b \cdot \text{Cov}(\theta_{i,t}, y_{i,t-(k-1)}) &(89)\\ &= b^k \cdot \text{Cov}(\theta_{i,t}, y_{i,t}) &(90)\\ &= b^k \cdot \frac{\pi_1}{1 - b\pi_2} \cdot \sigma_\theta^2. &(91)\end{aligned}$$

Thus, letting $\rho_{y,k}^{PC} \equiv \text{Cov}(y_{it}, y_{i,t-k})/\sigma_y^2$ be the k-generation income correlation, we have

$$\begin{aligned}\rho_{y,k}^{PC} &= b^k \cdot \left[\frac{\pi_1^2}{1 - b\pi_2} \cdot \frac{\sigma_\theta^2}{\sigma_y^2}\right] + \pi_2 \cdot \rho_{y,k-1}^{PC} &(92)\\ &= b^k \cdot \Phi(\gamma) + \pi_2 \cdot \rho_{y,k-1}^{PC}, &(93)\end{aligned}$$

where $\Phi(\gamma)$ is defined in (85).

This expression can be used to compare the k-generation correlation with that implied by the geometric extrapolation of the 1-generation correlation. In particular, the multi-generation correlation can be over- or under-stated by the geometric extrapolation of the 1-generation correlation depending on the relative size of ρ_1 and b.

Proposition 10 *Extrapolating the single-generation income correlation* understates *the true multi-generation income correlation (i.e. $\rho_1^k < \rho_k$) if*

$$\pi_2 < b \cdot (1 - \Phi(\gamma)). \tag{94}$$

The extrapolation overstates *the true correlation (i.e. $\rho_1^k > \rho_k$) if the inequality is reversed. The extrapolation* equals *the true correlation (i.e. $\rho_1^k = \rho_k$) if the inequality is replaced with an equality.*

Proof. We first establish that $\rho_1 < b$ implies $\rho_1^k < \rho_k$ for all $k \in \{2, 3, ...\}$. Since $\rho_1 = b\Phi(\gamma) + \pi_2$, the condition $\rho_1 < b$ is the same as the one stated in the proposition. To this end, we first we show that if (i) $\rho_1 < b$ and (ii) $\rho_1^k < \rho_k$ for some $k \in \{2, 3, ...\}$, then $\rho_1^{k+1} < \rho_{k+1}$. To see this note

$$\begin{aligned} \rho_1^{k+1} &= \rho_1^k \cdot [b\Phi + \pi_2] && (95) \\ &< b^{k+1}\Phi + \rho_1^k \cdot \pi_2 && (96) \\ &< b^{k+1}\Phi + \rho_k \cdot \pi_2 = \rho_{k+1}, && (97) \end{aligned}$$

where the first inequality comes from (i) and the second from (ii).

Second we show that (i) implies that (ii) holds for $k = 2$. This follows since

$$\begin{aligned} \rho_1^2 &= \rho_1 \cdot [b\Phi + \pi_2] && (98) \\ &< b^2\Phi + \rho_1 \cdot \pi_2 = \rho_2, && (99) \end{aligned}$$

where the inequality comes from (i).

By induction we therefore have that $\rho_1 < b$ implies $\rho_1^k < \rho_k$ for all $k \in \{2, 3, ...\}$. Thus long run persistence is larger than that implied by the short run persistence if $\rho_1 < b$ (equivalently, $\pi_2 < b \cdot (1 - \Phi(\gamma))$. The same logic applies when "<" in (i) is replaced with "=" or ">". □

In other words, the extrapolation can either overstate or understate the true persistence of income across multiple generations. The condition on parameters that determines which case arises relates to the relative strength of parental transmission of income and ability. For instance, the extrapolation *overstates* the true persistence in Becker and Tomes [1979] but *understates* it in Clark [2014]. In the former, there is no luck component, so that $\Phi = 1$ and the condition in the proposition can never hold. In the latter, parental income does not matter ($\pi_2 = 0$), so that the condition is always satisfied (as long as there is a luck component).[27]

C.3.2 Ability and Parental Income

How are offspring ability and parental income related in steady state? That is, to what extent to richer parents tend to have higher ability offspring? One way to get at this is to consider the steady state linear relationship between ability and parental income:

$$\theta_{it} = \varrho_0 + \varrho_1 \cdot y_{it} + e_{it}, \tag{100}$$

where $\mathbb{E}[e_{it} y_{it}] = 0$.

Proposition 11 *In the steady state, estimating a regression of ability on parental income (100) will yield a coefficient ϱ_1 equal to $\frac{b}{\pi_1} \cdot \Phi(\gamma)$ and an R-squared of $\frac{b^2}{1-b\pi_2} \cdot \Phi(\gamma)$, where Φ is defined in (85). Both of these quantities are increasing in the steady state quality of information, γ^{-1}.*

Proof. To get at this, first use the same notation as above to express the variance of income as follows:

$$\sigma_y^2 = \frac{1}{1 - b\pi_2} \frac{\pi_1^2 \sigma_\theta^2}{\Phi(\gamma)}. \tag{101}$$

Using (78) and (80) we have

$$\mathrm{Cov}(\theta_{it}, y_{i,t}) = b \cdot \frac{\pi_1}{1 - b\pi_2} \cdot \sigma_\theta^2. \tag{102}$$

As such, the correlation is

$$\mathrm{Cor}(\theta_{it}, y_{i,t}) = b \cdot \frac{1}{1 - b\pi_2} \cdot \frac{\pi_1 \sigma_\theta}{\sigma_y} = b \cdot \sqrt{\frac{\Phi(\gamma)}{1 - b\pi_2}}. \tag{103}$$

The value of ϱ_1 is the OLS coefficient on parental income and thus is

$$\varrho_1 = \frac{\mathrm{Cov}(\theta_{it}, y_{i,t})}{\sigma_y^2} = b \cdot \frac{\pi_1}{1 - b\pi_2} \cdot \frac{\sigma_\theta^2}{\sigma_y^2} = \frac{b}{\pi_1} \cdot \Phi(\gamma). \tag{104}$$

[27] Solon (2014) points out that the long run persistence of measured income can be understated by the short run persistence if there is measurement error. In contrast, the presence of luck does not require any measurement error. The two are not equivalent since the measurement error is not transmitted to offspring whereas luck is.

Note too that the R-squared from such a regression–i.e. the proportion of the variance of ability explained by parental income–is given by the squared correlation coefficient:

$$[\text{Cor}(\theta_{it}, y_{i,t})]^2 = \frac{b^2}{1 - b\pi_2} \cdot \Phi(\gamma). \tag{105}$$

Both of these measures of association are increasing in the quality of information, γ^{-1}, because Φ is decreasing in γ. $\square$

That is, better information in the steady state will raise the tendency for those with above average incomes to have children of above average ability. Again, the indirect sorting effect of α_1 and σ_ε^2 will exacerbate the direct effects. For instance, an increase in α_1 will mean that income is more sensitive to ability and therefore we would expect a high parental income to be more strongly associated with a high parental ability and thus with a high child ability. In addition, however, a higher α_1 raises the strength of sorting, implying a stronger association between parental ability and child ability.

D Values

The purpose of this section is to lay the groundwork that will ultimately allow us to derive expressions for the present value of dynastic income and dynastic ability. In particular these quantities will be expressed in terms of variables under which household (i, t) has control: the expected income and family status of the offspring in household (i, t). This will be useful for describing preferences in the marriage market, as potential partners will affect both of these expectations. It will also be useful for analysing optimal investment, as parental investment will also influence both of these expectations.

Consider the following system of expectations:

$$\mathbb{E}_t\left[y_{i,t+\tau+1}\right] = \pi_0 + \pi_1 \cdot \mathbb{E}_t\left[\theta_{i,t+\tau}\right] + \pi_2 \cdot \mathbb{E}_t\left[y_{i,t+\tau}\right] \tag{106}$$

$$\mathbb{E}_t\left[\theta_{i,t+\tau}\right] = (b/2) \cdot \mathbb{E}_t\left[\theta_{i,t+(\tau-1)}\right] + (1/2) \cdot \mathbb{E}_t\left[\bar{\phi}_{i,t+\tau}\right] \tag{107}$$

$$\mathbb{E}_t\left[\bar{\phi}_{i,t+\tau+1}\right] = b \cdot \left[\lambda \cdot \mathbb{E}_t\left[\bar{\phi}_{i,t+\tau}\right] + (1-\lambda) \cdot \mathbb{E}_t\left[\theta_{i,t+\tau}\right]\right] \tag{108}$$

The first of these is from the reduced-form income equation. The second is from the ability transmission equation, noting that $\mathbb{E}_t\left[\theta_{i',t+(\tau-1)}\right] = \mathbb{E}_t\left[\phi_{i,t+(\tau-1)}\right]$ and that $\mathbb{E}_t\left[\bar{\phi}_{i,t+\tau}\right] = b \cdot \mathbb{E}_t\left[\phi_{i,t+(\tau-1)}\right]$. The third comes from the relationship between family status and individual status, and the expression for individual status. For $\tau = 1, 2, ...,$

define

$$\mathbf{w}_\tau \equiv \begin{bmatrix} \mathbb{E}_t\left[y_{i,t+\tau}\right] - \frac{\pi_0}{1-\pi_2} \\ \mathbb{E}_t\left[\theta_{i,t+\tau-1}\right] \\ \mathbb{E}_t\left[\bar{\phi}_{i,t+\tau}\right] \end{bmatrix} \tag{109}$$

so that the system can be written

$$\mathbf{A}_1 \mathbf{w}_{\tau+1} = \mathbf{A}_2 \mathbf{w}_\tau \tag{110}$$

where

$$\mathbf{A}_1 \equiv \begin{bmatrix} 1 & -\pi_1 & 0 \\ 0 & 1 & 0 \\ 0 & -(1-\lambda)\cdot b & 1 \end{bmatrix}, \ \mathbf{A}_2 \equiv \begin{bmatrix} \pi_2 & 0 & 0 \\ 0 & b/2 & 1/2 \\ 0 & 0 & \lambda \cdot b \end{bmatrix}. \tag{111}$$

This gives

$$\mathbf{w}_{\tau+1} = \mathbf{A}\mathbf{w}_\tau \tag{112}$$

where $\mathbf{A} \equiv \mathbf{A}_1^{-1}\mathbf{A}_2$. Therefore we have

$$\mathbf{w}_\tau = \mathbf{A}^{\tau-1}\mathbf{w}_1 \tag{113}$$

so that

$$\mathbf{V} \equiv \sum_{\tau=1}^{\infty} \delta^{\tau-1}\mathbf{w}_\tau \tag{114}$$

$$= \left[\sum_{\tau=1}^{\infty} \delta^{\tau-1}\mathbf{A}^{\tau-1}\right]\mathbf{w}_1 = \left[\sum_{\tau=0}^{\infty} \delta^{\tau}\mathbf{A}^{\tau}\right]\mathbf{w}_1 = \varphi\mathbf{w}_1, \tag{115}$$

where $\varphi \equiv [\mathbf{I_3} - \delta\mathbf{A}]^{-1}$ (where $\mathbf{I_3}$ is the 3×3 identity matrix).

D.1 Dynastic Income

Define

$$V_{it}^y \equiv \mathbb{E}_t\left[\sum_{\tau=1}^{\infty} \delta^{\tau-1} y_{i,t+\tau}\right] \tag{116}$$

The main result here is the following.

Lemma 2 *We have*

$$V_{it}^{y} = \varphi_0 + \varphi_{11} \cdot \mathbb{E}_t \left[y_{i,t+1} \right] + \varphi_{12} \cdot \mathbb{E}_t \left[\theta_{i,t} \right] + \varphi_{13} \cdot \mathbb{E}_t \left[\bar{\phi}_{i,t+1} \right] \tag{117}$$

where $\{\varphi_0, \varphi_{11}, \varphi_{12}, \varphi_{13}\}$ *are positive constants given by*

$$\varphi_0 = \frac{\delta \pi_0}{1-\delta} \cdot \frac{1}{1-\delta\pi_2} \tag{118}$$

$$\varphi_{11} = \frac{1}{1-\delta\pi_2} \tag{119}$$

$$\varphi_{12} = \frac{1}{1-\delta\pi_2} \cdot \frac{\delta\pi_1}{(1-b\delta)(2-b\delta\lambda)} \cdot b(1-b\delta\lambda) \tag{120}$$

$$\varphi_{13} = \frac{1}{1-\delta\pi_2} \cdot \frac{\delta\pi_1}{(1-b\delta)(2-b\delta\lambda)}. \tag{121}$$

To see this note that, by definition, the first element of $\mathbf{V}$ (defined in (114)) is

$$\mathbf{V}_{11} = \sum_{\tau=1}^{\infty} \delta^{\tau-1} \mathbb{E}_t \left[y_{i,t+\tau} \right] - \frac{\pi_0}{1-\pi_2} \sum_{\tau=1}^{\infty} \delta^{\tau-1} \tag{122}$$

$$= V_{it}^{y} - \frac{\pi_0}{1-\pi_2} \frac{1}{1-\delta}. \tag{123}$$

From (115) we have

$$\mathbf{V}_{11} = \varphi_{11} \cdot (\mathbb{E}_t \left[y_{i,t+1} \right] - \frac{\pi_0}{1-\pi_2}) + \varphi_{12} \cdot \mathbb{E}_t \left[\theta_{i,t} \right] + \varphi_{13} \cdot \mathbb{E}_t \left[\bar{\phi}_{i,t+1} \right] \tag{124}$$

where φ_{rc} is the row r column c element of φ. Therefore

$$V_{it}^{y} = \varphi_0 + \varphi_{11} \cdot \mathbb{E}_t \left[y_{i,t+1} \right] + \varphi_{12} \cdot \mathbb{E}_t \left[\theta_{i,t} \right] + \varphi_{13} \cdot \mathbb{E}_t \left[\bar{\phi}_{i,t+1} \right] \tag{125}$$

where $\varphi_0 \equiv \frac{\pi_0}{1-\pi_2} \left[\frac{1}{1-\delta} - \varphi_{11} \right]$. The values given in the lemma are revealed by direct calculation of φ.

D.1.1 Segregation is Stable and Feasible

Segregation is trivially feasible. For stability we note that the attractiveness of individuals in the marriage market is summarized by an index of their observed characteristics. In particular, potential partners are evaluated according to $\mathbb{E}_t U_{i,t+1}$, which is precisely

V_{it}^{y} from the previous section. Partner characteristics matter because

$$\mathbb{E}_t\left[y_{i,t+1}\right] = \beta_0 + \beta_1 \cdot [x_{i,t} + x_{i',t}] + \beta_2 \cdot [y_{it} + y_{i't}]/2 \tag{126}$$

$$\mathbb{E}_t\left[\bar{\phi}_{i,t+1}\right] = b \cdot [\phi_{i,t} + \phi_{i',t}]. \tag{127}$$

It therefore follows that the attractiveness of the spouse from household (j, t) is given by:

$$a_{j,t} \equiv (\varphi_{11}\beta_1/2) \cdot x_{j,t} + (\varphi_{11}\beta_2/2) \cdot y_{j,t} + (\varphi_{13}b) \cdot \phi_{j,t}. \tag{128}$$

Stability in the period t marriage market requires segregation on a_{it}, which is indeed achieved by segregation. That is, if an agent from household (i, t) were to strictly prefer to marry an agent from household (j, t) to their assigned partner under segregation, (i.e. an agent from household (i', t), where $a_{i',t} = a_{it}$), then it must be that $a_{i',t} < a_{j,t}$. But then this implies $a_{i,t} < a_{j,t} = a_{j',t}$, so that j would strictly prefer to *not* match with i over their assigned partner under segregation j'. Thus, segregation is indeed stable and feasible.

D.1.2 Optimal Parental Investment

Proof of Proposition 7.

Assuming all other households invest a fraction z^* of their income, the preferences of household (i, t) are given by

$$\ln(1 - z_{it}) + y_{it} + \delta \cdot V_{it}^{y}, \tag{129}$$

where V_{it}^{y} is defined above. Investment affects V_{it}^{y} in two ways, since:

$$\mathbb{E}_t\left[y_{i,t+1}\right] = \text{constants} + (\beta_1\alpha_2) \cdot \ln z_{it} \tag{130}$$

$$\mathbb{E}_t\left[\bar{\phi}_{i,t+1}\right] = b \cdot \mathbb{E}_t[\phi_{i,t}] = \text{constants} + b(1 - \lambda)\frac{\alpha_2}{\alpha_1} \cdot \ln z_{it}. \tag{131}$$

Therefore, ignoring constants, preferences over z_{it} are given by

$$\ln(1 - z_{it}) + \zeta_1 \cdot \ln z_{it} + \zeta_2 \cdot \ln z_{it} \tag{132}$$

where

$$\zeta_1 \equiv \delta \cdot \varphi_{11} \cdot (\beta_1 \alpha_2) = \frac{\delta \beta_1 \alpha_2}{1 - \delta[\beta_1 \alpha_2 + \beta_2]} \tag{133}$$

$$\zeta_2 \equiv \delta \cdot \varphi_{13} \cdot b(1-\lambda) \frac{\alpha_2}{\alpha_1} = \zeta_1 \cdot \frac{b\delta}{1 - b\delta} \cdot \frac{1-\lambda}{(2 - b\delta\lambda)} \tag{134}$$

Maximizing with respect to z_{it} delivers the expression claimed in the proposition.□

D.2 Dynastic Ability

Lemma 3 *Dynastic ability, $V_{it}^{\theta} \equiv \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}_t \left[\theta_{i,t+\tau}\right]$, can be written*

$$V_{it}^{\theta} = \varphi_{22} \cdot \mathbb{E}_t \left[\theta_{it}\right] + \varphi_{23} \cdot b \cdot \phi_{it}, \tag{135}$$

where $\varphi_{22} = \frac{2 - b\delta(1+\lambda)}{(1-b\delta)(2-b\delta\lambda)}$ and $\varphi_{23} = \frac{\delta}{(1-b\delta)(2-b\delta\lambda)}$.

Proof. To show this, use the vector $\mathbf{V}$ from (114). By definition, the second element is:

$$\mathbf{V}_{21} = \sum_{\tau=1}^{\infty} \delta^{\tau-1} \mathbb{E}_t \left[\theta_{i,t+\tau-1}\right] = V_{it}^{\theta}. \tag{136}$$

As such, $V_{it}^{\theta} = \varphi_{21} \cdot \mathbb{E}_t \left[y_{i,t+1}\right] + \varphi_{22} \cdot \mathbb{E}_t \left[\theta_{it}\right] + \varphi_{23} \cdot \mathbb{E}_t \left[\bar{\phi}_{i,t+1}\right]$, where direct computation of φ yields $\varphi_{21} = 0$ and the values of φ_{22} and φ_{23} are those given in the lemma. The result follows from $\bar{\phi}_{i,t+1} = b \cdot \phi_{it}$ □
This characterization, via the following corollary, demonstrates that dynastic ability is a meaningful measure because of its close relationship to the present discounted value of family incomes.

Corollary 6 *Dynastic income can be expressed as an affine function of Dynastic ability and parental income:*

$$V_{it}^{y} = \left[\frac{\pi_0}{1 - \delta\pi_2} \frac{1}{1-\delta}\right] + \left[\frac{\pi_2}{1 - \delta\pi_2}\right] \cdot y_{it} + \left[\frac{\pi_1}{1 - \delta\pi_2}\right] \cdot V_{it}^{\theta}. \tag{137}$$

Proof. Use the fact that $\mathbb{E}_t \left[y_{i,t+1}\right] = \pi_0 + \pi_1 \cdot \mathbb{E}_t \left[\theta_{i,t}\right] + \pi_2 \cdot y_{it}$ in Lemma 2, collect terms and simplify. □
A direct consequence of Lemma 3 is that

$$\mathbb{E}_t[V_{it}^{\theta} | \theta_{it}] = \varphi_{22} \cdot \theta_{it} + \varphi_{23} \cdot b \cdot \phi_{it}. \tag{138}$$

This tells us that an individual's dynastic ability depends on their *actual* and *expected* ability. From the definitions of φ_{22} and φ_{23} we can write

$$\mathbb{E}_t[V_{it}^{\theta}|\theta_{it}] = \frac{1}{1-b\delta} \cdot \left[\left(1 - \frac{b\delta}{2-b\delta\lambda}\right) \cdot \theta_{it} + \frac{b\delta}{2-b\delta\lambda} \cdot \phi_{it}\right]. \tag{139}$$

This tells us that $b\delta/[2-b\delta\lambda]$ acts as the weight placed on 'appearances' relative to 'reality' when it comes to shaping dynastic ability. This weight is increasing in λ, so that a greater steady state emphasis on family background implies that dynastic ability is more sensitive to expected ability relative to actual ability (i.e. appearances are more important).

In any case, from belief updating we have

$$\phi_{it} = \lambda \cdot \bar{\phi}_{it} + (1-\lambda) \cdot s_{it} = \lambda \cdot \bar{\phi}_{it} + (1-\lambda) \cdot (\theta_{it} + \xi_{it}). \tag{140}$$

Putting this together yields:

$$\mathbb{E}_t[V_{it}^{\theta}|\theta_{it}] = [\varphi_{22} + \varphi_{23} \cdot b \cdot (1-\lambda)] \cdot \theta_{it} + [\varphi_{23} \cdot b \cdot \lambda] \cdot \bar{\phi}_{it} + [\varphi_{23} \cdot b \cdot (1-\lambda)] \cdot \xi_{it}. \tag{141}$$

Proof of Proposition 5. From (141) we have

$$\frac{d}{d\varepsilon}\mathbb{E}_t[V_{it}^{\theta}|\theta_{it}] = [\varphi_{23} \cdot b \cdot (1-\lambda)] \cdot \frac{d}{d\varepsilon}\xi_{it}. \tag{142}$$

The result follows from the computed value of φ_{23} and the fact that $\frac{d}{d\varepsilon}\xi_{it} = 1/\alpha_1$. □

Proof of Proposition 6. From (141) we have

$$\frac{d}{d\bar{\phi}_{it}}\mathbb{E}_t[V_{it}^{\theta}|\theta_{it}] = \varphi_{23} \cdot b \cdot \lambda \tag{143}$$

The result follows from the computed value of φ_{23}. □