Models of Intergenerational Mobility

Steven N. Durlauf

Second Annual Summer School on Socioeconomic Inequality

University of Chicago

2013

Overview

I will describe four approaches to modelling intergenerational mobility.

1. Family investment income
2. Skills
3. Neighborhoods
4. Genetics

Theories are mutually compatible.

I will also discuss efforts to parse out sources of intergenerational persistence.

i. Family Income/Investment Models

Becker and Tomes (1993) and Loury (1989) are the sources of what may be regarded as the classical intergenerational mobility models.

These models focus on income as the outcomes whose intergenerational persistence is of interest.

Substantively, the model treats the transmission process as intrafamily. Parents invest in the human capital of their offspring.

Demography

I distinct family dynasties, denoted by i.

Individuals are assumed to live 2 periods. The first period is childhood, denoted as c and the second period is adulthood, denoted as a. t denotes time of birth.

Individuals reproduce asexually; this allows us to ignore issues of intermarriage across dynasties and is inessential for the purposes of this discussion.

Education/Income Relationship

In childhood, an individual in dynasty i born at time t makes no choices but receives a human capital investment $I_{i,c,t}$ from her parent.

This investment produces a level of education attainment for the child $E_{i,c,t}$.

Realized education is determined by the process

$$E_{i,c,t} = e\left(I_{i,c,t}, \zeta_{i,c,t}\right)$$

$\zeta_{i,c,t}$ denotes unobserved heterogeneity across children; can think of as latent ability. Treat as observable to parent. It is natural to allow it to be correlated within families for both genetic and environmental reasons.

Assume that is $\zeta_{i,c,t}$ identically distributed across all agents with probability measure $\mu_{\zeta}\left(\cdot\right)$. Assume that parents observe $\zeta_{i,c,t}$ when they make their investment decisions; this assumption is nontrivial as it allows parents to adjust behavior for child "quality".

As an adult, this same individual i born at t works and receives income $Y_{i,a,t+1}$; note that the time index has changed as we have moved from childhood to adulthood. Income is determined by the process

$$\log Y_{i,a,t+1} = f\left(E_{i,c,t}, \varepsilon_{i,a,t+1}\right)$$

In this equation, $\varepsilon_{i,a,t+1}$ represents unobserved heterogeneity associated with adults; it is assumed to be identically distributed across agents with probability measure $\mu_{\varepsilon}(\cdot)$. Given separate modelling of unobserved childhood heterogeneity, perhaps $\varepsilon_{i,a,t+1}$ is most naturally interpreted as labor market luck.

Substantive assumption is made that parents have no information on $\varepsilon_{i,a,t+1}$ when investment decisions in children are made at time t.

Formally, if $F_{i,a,t}$ is information set of parent i at time t,

$$\mu\left(\varepsilon_{i,a,t+1} \middle| F_{i,a,t}\right) = \mu\left(\varepsilon_{i,a,t+1}\right)$$

Model is in fact simple enough that $F_{i,a,t}$ is generated by $Y_{i,a,t}$ and $\zeta_{i,a,t}$

Parental Investment

Where does choice appear in this model? Choice appears via parental investment decisions. Each adult splits her income between the consumption, $C_{i,a,t+1}$ and the human capital investment in the child $I_{i,c,t+1}$, i.e.

$$Y_{i,a,t+1} = C_{i,a,t+1} + I_{i,c,t+1}$$

This is more than an identity as it means that a parent cannot borrow to raise the human capital investment of her child. To do so would require that a parent can create a legal obligation for a child to pay her debts.

(Obviously, the parent will not be around to repay any loan!) Hence all investment in a child must come from the parent's income.

This structure is sufficient to provide a description of intergenerational income transmission. Since each adult agent is solving an identical decision problem, if a solution exists for the optimal human capital investment for each value of $Y_{i,a,t+1}$ and $\zeta_{i,y,t+1}$ it must be the case that it can be expressed as

$$I_{i,c,t+1} = g\left(\log Y_{i,a,t+1}, \zeta_{i,c,t+1}\right)$$

Parental Preferences

Nothing has been assumed (so far) about the nature of parental preferences.

Loury (1981) considers an environment in which

$$V_{i,a,t} = u\left(c_{i,a,t}\right) + \gamma V_{i,a,t+1} = \sum_{j=0}^{\infty} \gamma^{j} u\left(c_{i,a,t+j}\right)$$

Other approaches include

$$V_{i,a,t} = u\left(c_{i,a,t}\right) + \gamma Y_{i,a,t+1}$$

or

$$V_{i,a,t} = u\left(c_{i,a,t}\right) + \gamma E_{i,a,t+1}$$

These forms simplify the analysis as parents do not solve infinite horizon problems.

Dynamics

The law of motion for family income generated by this model is

$$\log Y_{i,a,t+1} = f\left(e\left(g\left(\log Y_{i,a,t}, \zeta_{i,c,t}\right), \zeta_{i,c,t}\right), \varepsilon_{i,a,t+1}\right)$$

This equation provides a description of the evolution of income across generations.

The nature of the evolution will depend on the three functions $e(\cdot,\cdot)$, $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$.

The first function $e(\cdot,\cdot)$ characterizes the relationship between human capital investment and education.

The second function, $f(\cdot,\cdot)$ characterizes how education is converted into income.

These two functions represent the technology of the economy.

The third function, $g(\cdot,\cdot)$, will depend on the technology of the economy as well as the preferences of adults.

Notice that this general structure does *not* produce the regression that is the basis of the IGE literature.

To do produce the linear IGE model one needs additional assumptions. I now provide functional form assumptions on primitives that generate the standard IGE equation.

Becker and Tomes (1979) via Solon (2013)

First, assume that human capital is additive in investment and ability.

$$E_{i,c,t} = \theta \log I_{i,c,t} + \zeta_{i,c,t}$$

As before $\zeta_{i,c,t}$ denotes unobserved ability heterogeneity. Recall that parents are assumed to know the value of $\zeta_{i,c,t}$ when human capital investment decisions are made.

Second, income is determined by

$$Y_{i,a,t+1} = \mu + w \log E_{i,c,t} + \varepsilon_{i,a,t+1}$$

so that

$$Y_{i,a,t+1} = \mu + w\theta \log E_{i,c,t} + w\zeta_{i,c,t} + \varepsilon_{i,a,t+1}$$

Third, the utility of the adult i born at time $t-1$ is

$$U_{i,a,t} = (1-\pi)\log C_{i,a,t} + \pi \log Y_{i,a,t+1}$$

Equilibrium law of motion

$$\log Y_{i,a,t+1} = \mu + \gamma \log\left(\frac{\pi\theta w}{1-\pi\left(1-\theta w\right)}\right) + \pi \log Y_{i,a,t-1} + w\varsigma_{i,c,t+1} + \varepsilon_{i,a,t+1}$$

Note that it is not natural to assume that $\zeta_{i,c,t+1}$ is independent across time if there is a genetic component to ability, for example. Suppose that

$$\zeta_{i,c,t+1} = \delta + \lambda \zeta_{i,c,t} + v_{i,a,t+1}$$

where $v_{i,a,t+1}$ is uncorrelated, then for the regression

$$\log Y_{i,a,t+1} = \alpha + \beta \log Y_{i,a,t-1} + \xi_{i,a,t+1}$$

One can show that

$$\beta = \frac{w\theta + \lambda}{1 + \lambda w \theta}$$

Credit Constraints

This model is often said to embody credit constraints. This claim is based on the fact that parental investment is determined by the parental budget constraint.

In other words, parents cannot borrow against children's future income.

Loury (1981) first emphasized this restriction. That said, it is not a capital market imperfection as conventionally understood. It derive from the idea that a parent cannot make children liable for parents' debts.

Imperfection is variation of classical OLG.

ii. Skills/Lifecourse Dynamics Approach to Intergenerational Mobility

The frontier in understanding family and other influences on individuals moves away from the focus on income towards cognitive skills and personality traits which together represent a broad conception of skills.

This work has been pioneered by James Heckman; the surveys Borghans, Duckworth, Heckman, and B. ter Weel (2008) and Almlund, Duckworth, Heckman, and Kautz (2011) are comprehensive surveys.

Focus is to a large extent on family, but other factors are naturally included.

Early childhood investment is one obvious example.

Note that the objects of interest in intergenerational mobility are now changed.

The new economics of skills has two critical features.

First, it employs a broad definition of skills. In particular, it differentiates between cognitive and noncognitive skills. In this respect, the economics of skills has followed the psychology literature, in which intelligence and personality studied as distinct aspects of the mind.

Many psychologists dislike the term “noncognitive” skills since these skills are also part of the mind, and so in their view are cognitive. Nevertheless, I will follow the language of economists.

Second, the literature focuses on development across the childhood and adolescence.

It therefore directly challenges the 2-period overlapping generations paradigm.

Cognitive skills are those that one associates with intelligence and human capital. Both of these components, in turn, may be understood as possessing subcomponents.

Crystal intelligence refers to the ability of an individual to draw on his knowledge and experience to solve recurring problems.

Fluid intelligence refers to the ability to draw on knowledge and experience to solve novel problems.

Noncognitive skills may be thought of as personality traits. In the psychology literature, a standard way of conceptualizing personality is the so-called 5-factor model:

Openness (which captures curiosity and receptivity to new situations)

Conscientiousness (which includes whether one is well organized and/or efficient)

Extraversion (which includes friendliness and whether one is high energy)

Agreeableness which includes friendliness and compassion), and

Neuroticism (which includes self-confidence and sensitivity to stress).

The acronym for these is OCEAN.

From the perspective of understanding intergenerational mobility, the importance of this broad conception of skills is that an individual is associated with a vector of cognitive skills, θ_i^C and a vector of noncognitive skills θ_i^{NC} which evolve over childhood and adolescence.

While they do not stop at age 18, the family's role can be understood as influencing the evolution during this time.

Heckman has developed formal models of skill evolution in a large number of papers; in terms of notation I essentially follow Cunha and Heckman (2007) which is an especially accessible treatment; the main difference is that I index investments so that an investment at t affects skills at t.

The formal model of skill evolution, in abstract terms, is simply a difference equation that describes the values of $\theta_{i,t}^{C}$ and a vector of noncognitive skills $\theta_{i,t}^{NC}$ as the outcomes of dynamic processes; for $t > 0$

$$\theta_{i,t}^{C} = f_t^{C}\left(\theta_{i,t-1}^{C}, \theta_{i,t-1}^{NC}, I_{i,t}, h_i\right)$$

and

$$\theta_{i,t}^{NC} = f_t^{NC}\left(\theta_{i,t-1}^{C}, \theta_{i,t-1}^{NC}, I_{i,t}, h_i\right)$$

In these equations, $I_{i,t}$ denotes a vector that captures everything that affects individual i at time t. Cunha and Heckman refer to these effects as investments, which corresponds to the idea that $I_{i,t}$ creates a change in the stock of skills.

Under this very broad definition, $I_{i,t}$ would therefore include everything from parental inputs to schooling to luck (e.g. whether one has been healthy during the time period). The term h_i denotes a vector of initial conditions for i. Cunha and Heckman mention parental characteristics such as IQ and education as components of h_i.

There are two things to observe about the dynamic processes.

First, notice that $\theta_{i,t}^{C}$ is a function of $\theta_{i,t-1}^{NC}$ and $\theta_{i,t}^{NC}$ is a function of $\theta_{i,t-1}^{C}$. This captures the idea that one type of skills facilitates the acquisition of the other type. It is easy to think of intuitive examples. A smart child who is not conscientious may not acquire much knowledge.

Second, the functions f_t^{C} and f_t^{NC} have time indices. This captures the idea that mapping from the arguments of the functions to skill levels can depend on age. This is one way, at least in terms of notation, to allow for the possibility that the plasticity of skills changes over childhood and adolescence. We have talked about the standard example of differential plasticity: language acquisition ability.

Cunha and Heckman show that if one engages in recursive substitution, to eliminate the dependence of $\theta_{i,t}^{C}$ and $\theta_{i,t}^{NC}$ on $\theta_{i,t-1}^{C}$ and $\theta_{i,t-1}^{NC}$, one can reformulate the dynamic skills models as

$$\theta_{i,t}^{C} = m_{t}^{C}\left(I_{i,t},..,I_{i,1},h_{i}\right)$$

and

$$\theta_{i,t}^{NC} = m_{t}^{NC}\left(I_{i,t},...,I_{i,1},h_{i}\right)$$

This formulation emphasizes the idea that the stocks of skills at t are determined by initial conditions and the series of investments up to that time.

The properties of the investment are important in thinking about policy.

One such property is dynamic complementarity. Suppose that $I_{i,j,t}$ is an investment that the government can augment. If

$$\frac{\partial^2 f_t^C\left(\theta_{i,t-1}^C, \theta_{i,t-1}^{NC}, I_{i,t}, h_i\right)}{\partial \theta_{i,t-1}^C \partial I_{i,j,t}} > 0$$

$$\frac{\partial^2 f_t^C\left(\theta_{i,t-1}^C, \theta_{i,t-1}^{NC}, I_{i,t}, h_i\right)}{\partial \theta_{i,t-1}^{NC} \partial I_{i,j,t}} > 0$$

then skills exhibit dynamic complementarities in the sense that the marginal product of increasing $I_{i,j,t}$ is increasing in the stocks of skills.

It may be that the efficacy of later investments, e.g. high school, depends on skills acquired earlier in life. A related notion of dynamic complementarity is

$$\frac{\partial^2 m_t^C\left(I_{i,t},..,I_{i,1},h_i\right)}{\partial I_{i,j,t}\partial I_{i,t-k}} > 0,\ 0 < k \leq t-1$$

which involves complementarities across investments. This captures the idea that schooling investments at different ages are interrelated.

Cunha and Heckman also define the notion of critical periods of development.

Let $I_{i,-j,t}$ denote the vector of investments in i at t other than investment type j.

A critical development period with respect to investment type j is a time T such that

$$\frac{\partial m_t^C\left(I_{i,j,t},\overline{I}_{i,-j,t},\overline{I}_{i,t-1},..,\overline{I}_{i,1},\overline{h}_i\right)}{\partial I_{i,j,t}} \gg 0, \text{ if } t = T$$

and

$$\frac{\partial m_t^C\left(I_{i,j,t},\overline{I}_{i,-j,t},\overline{I}_{i,t-1},..,\overline{I}_{i,1},\overline{h}_i\right)}{\partial I_{i,j,t}} \approx 0, \text{ if } t \neq T .$$

Notice that this is a distinct idea from that of dynamic complementarity.

The critical period idea says that there are particular times when investments are efficacious.

Complementarity suggests that investments are interconnected.

Skills and Intergenerational Mobility

The skills approach is qualitatively different from the family income investment approach.

1. Different objects of interest and richer set of mechanisms.
2. Life cycle is explicit.
3. Parent and offspring skill lifecycles interact in ways not available to 2 generation OG. Example (due to Heckman ES Pres. Lecture). Credit constraints of parents in early childhood are key.

iii. Neighborhood Models

A different approach to understanding intergenerational mobility focuses on social influences on children and adults.

An obvious example is education. For most individuals, education is publically provided. Further, school quality is determined by peer quality as well as the level of local influence on educational expenditure. Similarly, local communities are sources of information and role models.

From this vantage point, what is important is that parents affect memberships in neighborhoods.

Durlauf (1996)

This model studies an environment in which neighborhoods are the sole mechanism for transmitting economic status across generations.

Continuum of agents in population: will rule out isolated families.

$N_{n,t}$ = collection of families in neighborhood n in period t

$\mu\left(N_{n,t}\right)$ = population size of neighborhood n in period t

$\hat{F}_{Y,n,t}$ = empirical income distribution of neighborhood n in period t

Preferences

$$EU_{i,a,t} = (1-\alpha)\log C_{i,a,t} + \alpha E\left(\log Y_{i,a,t+1} \middle| F_t\right)$$

Only difference from Solon model is that parents cannot forecast future income of offspring.

Income/Education relationship

$$\log Y_{i,a,t} = \kappa + \log H_{n,t-1} + \xi_{i,a,t}$$

$\log H_{n,t-1}$ denotes human capital generated in neighborhood in which a child grows up.

$$\xi_{i,a,t} = \nu_{n,a,t} + \gamma_{i,a,t}$$

This allows for neighborhood- and individual-specific unobservables.

Budget Constraint

$$Y_{i,t} = C_{i,t} + T_{i,t}$$

$T_{i,t}$ denotes taxes. No private human capital investment. Taxes are linear in income

$$T_{i,t} = \tau_{n,t} Y_{i,t}$$

Human Capital and Education

$ED_{n,t}$ =educational quality level

$TE_{n,t}$ = total expenditure need for quality level

$$TE_{n,t} = \lambda_1 ED_{n,t} + \lambda_2 \mu\left(N_{n,t}\right) ED_{n,t}$$

Incentive for heterogeneous neighborhoods because of fixed costs.

Human Capital Formation Process

$$H_{n,t} = \Theta\left(\hat{F}_{Y,n,t}\right) ED_{n,t}$$

$\Theta(\cdot)$ is increasing in distribution (first order dominance sense)

Neighborhood Formation Rule

Core/Club: families can form neighborhoods and exclude others.

This can be relaxed.

Neighborhood Equilibrium

1. Neighborhoods will stratify by income

2. Constant tax rate in each neighborhood equal to α

First follows from tax revenue and social interaction effects.

Second follows from Cobb-Douglas preferences.

Dynamics of Inequality

Proposition

Let $g_{n,t}$ denote expected income growth among offspring in neighborhood n at t

If n is a higher income neighborhood than n' has at least as large a population than and $\mu\left(N_{n,t}\right) \geq \mu\left(N_{n',t}\right)$, then

$$g_{n,t} > g_{n',t}$$

Intuition

Educational quality and social interactions create the disparity.

Population requirement avoids returns to scale offsetting these benefits to higher income.

Permanent Inequality

Denote the highest and lowest income adults at time t as i and i'. Call this pair-specific ratio ρ_{i,i',t_0}. For $t > t_0$ elapses, $\rho_{i,i',t}$ follows the descendants of these adults.

Proposition

With positive probability, $\rho_{i,i',t}$ will never decrease.

Intuition

Income growth rates of the families differ. Log income difference behaves as a random walk with drift away from an absorbing barrier, in this case $\rho_{i,i',t}$.

Technology assumption is critical.

This is a case in which standard convergence measure would fail in cross country growth literature.

Role of Assumptions

Cobb-Douglas is critical. It means that there is no heterogeneity in preferred tax rates between more and less affluent.

Core equilibrium concept is not essential, if one keeps Cobb-Douglas. One can show willingness to pay differentials will support stratification.

Neighborhoods and Assortative Matching

The neighborhood model hints at a general idea: matching may contribute to inequality and intergenerational persistence of economic status.

Assortative mating is a standard theory of intergenerational mobility. Note that this may refer to parental income, or the educational status of individuals, etc.

One can move beyond marriage. If parents, thought of as workers are matched to firms, the matching rule will determine wages, which, following Becker and Tomes reasoning, will matter for offspring investment.

Separately, matching can imply a social determinant to outcomes in the sense that the skills transmitted from parents to offspring will have rewards determined by the matching of the adult offspring to others in the economy.

Link to Intergenerational Mobility

Approach matters for intergenerational mobility as it elucidates mechanisms.

Parental disadvantage may place child in a social environment in which “traps” exist.

If matching traps parents, then Becker/Heckman mechanisms come into play

Complementarity

Complementarity is a fundamental concept in the social interactions literature because it characterizes how individuals affect one another and presence or absence of complementarities is critical in understanding how social influences affect individuals and, in consequence, inequality in its various dimensions.

Consider a group of I individuals. Each individual is associated with a vector of attributes x_i. At this stage, the attributes may be choices, characteristics, or some combination of each; x_{ia} denotes an element of x_i. Each vector of attributes produces a payoff

$$\Phi_i\left(x_i, x_{-i}\right)$$

where x_{-i} denotes the vector of attributes of members of the group other than i.

What features of this interdependence of attributes in agent i's payoff function of interest? Two features warrant explicit definition. The payoff function is said to exhibit positive spillovers with respect to a particular attribute a if

$$\frac{\partial \Phi_i\left(x_i, x_{-i}\right)}{\partial x_{ja}} \geq 0,\ j \neq i$$

The payoff function is said to exhibit complementarities with respect to attribute a if

$$\frac{\partial^2 \Phi_i\left(x_i, x_{-i}\right)}{\partial x_{ib} \partial x_{ja}} \geq 0,\ j \neq i$$

Positive spillovers tell us about the marginal payoff consequences for *i* associated with changes in the attributes of others. These are important in calculating welfare consequences of payoff interdependence for an individual given his group.

Complementarities matter for understanding how the marginal payoff changes in *i*'s attributes are affected by changes in the attributes of others. Intuitively, since choices are determined by first order conditions, one expects complementarities to matter in determining equilibrium choices as well as the implications of alternative configurations of a population into groups.

Increasing Differences and Supermodularity

In the modern literature, the notion of complementarity has been generalized to account for richer state spaces and payoff functions. The function $f(x,y)$, which maps some set S (the joint support of x and y) to R, is said to exhibit to exhibit increasing differences with respect to x and y if, for any $\overline{x} \geq \underline{x}$ and $\overline{y} \geq \underline{y}$,

$$f(\overline{x},\overline{y}) - f(\underline{x},\overline{y}) \geq f(\overline{x},\underline{y}) - f(\underline{x},\underline{y})$$

A concept that is closely related to increasing differences is supermodularity. Let $z=(x,y)$, i.e. combine the elements of x and y into a vector z. Define the meet of z and z' as

$$(z \wedge z')_a = \min\{z_a, z'_a\}$$

The meet operator thus forms a new vector based on pairwise minima across the original vectors. Similarly, one can define the join of z and z' as

$$(z \vee z')_a = \max\{z_a, z'_a\}$$

The function $f(z)$ is supermodular if

$$f(z \vee z') + f(z \wedge z') \geq f(z) + f(z')$$

Supermodularity is a stronger property than increasing differences in the sense that supermodularity always implies increasing differences, while the converse does not hold. However, if the support of z can be written as the Cartesian product of completely ordered sets, then one can show that supermodularity of a function is equivalent to increasing differences with respect to all of its arguments, i.e. $f(z)$ exhibits increasing differences for any partition of z into two vectors.

Note that these supermodularity and increasing differences do not require continuity, let alone differentiability, of $f(z)$.

If second order differentiability does hold, increasing differences in all arguments of $f(z)$ is equivalent to the condition that $\frac{\partial^2 f(z)}{\partial z_a \partial z_b} \geq 0$, a $\neq b$, and so corresponds to classical definition of complementarity.

Complementarities in Characteristics and Assortative Matching

I now focus on complementarities in characteristics. Here the economic ideas involve the way that groups form.

Once groups form, there are no subsequent choices.

Obviously, a complete theory of the social determinants of inequality needs to address both how groups (or richer social structures form) as well as the choices that occur after group formation occurs.

Becker Marriage Model

To see how complementarities in characteristics affect group formation, I start with a classic analysis due to Becker (1973) that relates the efficiency of assortative matching to complementarity. Consider a population of *I* men and *I* women. Suppose that the product of a marriage depends on scalar characteristics m_i and w_j of the men and women respectively. Suppose that the product of a given match is $\Phi(m,w)$. Becker (1973) establishes the following.

Proposition. Optimality of assortative matching.

If $\dfrac{\partial^2 \Phi(m,w)}{\partial m \partial w} \geq 0$ then assortative matching maximizes the sum of the products of marriages.

For assortative matching to be *in*efficient, there must exist, pairs $\left(\underline{m},\bar{w}\right)$ and $\left(\bar{m},\underline{w}\right)$, $\bar{m} > \underline{m}$ and $\bar{w} > \underline{w}$ such that

$$\Phi\left(\underline{m},\bar{w}\right)+\Phi\left(\bar{m},\underline{w}\right) > \Phi\left(\bar{m},\bar{w}\right)+\Phi\left(\underline{m},\underline{w}\right)$$

or

$$\Phi\left(\underline{m},\bar{w}\right)-\Phi\left(\underline{m},\underline{w}\right)-\Phi\left(\bar{m},\bar{w}\right)+\Phi\left(\bar{m},\underline{w}\right) > 0$$

$$\int_{\underline{w}}^{\bar{w}}\frac{\partial\Phi\left(\underline{m},w\right)}{\partial w}-\int_{\underline{w}}^{\bar{w}}\frac{\partial\Phi\left(\bar{m},w\right)}{\partial w} > 0$$

$$-\int_{\underline{m}}^{\bar{m}}\frac{\partial^2\Phi\left(m,w\right)}{\partial m\partial w} > 0$$

The last inequality is inconsistent with the complementarity assumption, which proves the proposition.

If one is working in an environment in which increasing differences is equivalent to supermodularity, then Becker’s result follows virtually by definition.

Hence, Becker’s finding holds for “marriages” of any size; it also does not depend on second-order differentiability.

Becker’s result represents a true “equity-efficiency payoff”. It is evident that the most efficient payoff (in the sense of maximizing total payoffs) also maximizes the differences between the highest- and lowest-payoff marriages.

A separate question is whether assortative matching occurs in equilibrium even when the male and female attributes are substitutes and not complements. The answer depends on the rules by which a husband and wife split the output of the marriage. If agents are paid their marginal products (so to speak) the equilibrium matching will correspond to the social planner's solution.

On the other hand, if the marital output is equally split between partners, then the only equilibrium that can occur is one with assortative matching, even if it is not efficient.

This example illustrates how a lack of markets can increase inequality.

Notice that the Becker result and the definition of supermodularity take the location of agents in the output function seriously; in other words the first argument of the function refers to men and the second argument refers to women.

Other optimal matching problems may not have this feature. Here is one example.

Suppose that a social planner has NK agents with scalar characteristics a_i who must be organized into K-tuples, each of which produces some payoff. In this case, one cannot immediate equate supermodularity with the efficiency of assortative matching, since supermodularity takes the order of the agents as given. In order to preserve the equivalence, it is necessary to add an assumption Durlauf and Seshadri (2003) call permutation invariance. Permutation invariance means that if a is a K-tuple of characteristics and a' is a permutation of a, then

$$\Phi(a) = \Phi(a')$$

In this case, one can show that assortative matching is also efficient. To see why, consider any configuration. Reorder the vectors so that the elements in each run from largest to smallest. If the vectors do not exhibit assortative matching, replace them with their join and meet. These must produce at least as much as the original vectors. Rank order the join and meet and repeat. Eventually, you will produce assortatively matched sets of agents. See Durlauf and Seshadri (2003) for the formal argument.

Permutation invariance may make sense in some contexts. If a firm is assigned K workers, the firm's manager will assign the workers to tasks in order to maximize total output. The order in which the workers characteristics are reported does not matter to the manager. When one considers contexts with permutation invariance, assortative matching is equivalent to stratification of agents across groups with respect to the characteristic a.

By stratification, I mean that the supports of the characteristics can be completely ordered using weak inequalities.

The efficiency of assortative matching for this context does depend on the assumption that all groups are of equal size. In other words, the comparisons of the configurations of alternative group compositions in which supermodularity implies the efficiency of assortative matching presupposes that the arguments of the payoff functions have the same dimension. Durlauf and Seshadri (2003) gives an example in which assortative matching, breaks down when group sizes differ.

Multiple Equilibria and Complementarities

I next describe a simple complementarity game. The exposition is based Cooper and John (1988). The Cooper-John model is useful in illustrating some of the implications of complementarity in a simple and intuitive way. In their model, individuals make choices ω_i in order to maximize

$$\Phi\left(\omega_i, \bar{\omega}_{-i}\right)$$

where $\bar{\omega}_{-i}$ denotes the average choice of agents other than i.

This formulation limits the form of complementarities in two ways that have economic content.

First, it means that any permutation of the distribution of choices by others leaves the payoff of the agent unchanged. This means that the identities of the agents are irrelevant.

Second, the only moment of the distribution of others' choices that matters to i is the average of the choices. It is easy to think of cases where other moments matter. These restrictions are useful as they allow one to focus on symmetric equilibria, i.e. equilibria in which all agents choose the same level of ω.

In a noncooperative (specifically, Nash) game, it is easy to see that each agent choice solves

$$\frac{\partial \Phi\left(\omega_i, \overline{\omega}_{-i}\right)}{\partial \omega_i} = 0$$

(Here and elsewhere I assume that all first order conditions are met with equality.) A symmetric equilibrium is defined by a common choice level ω^{NC} such that

$$\frac{\partial \Phi\left(\omega^{NC}, \omega^{NC}\right)}{\partial \omega_i} = 0$$

In contrast, a cooperative solution is a common choice level is defined by

$$\frac{\partial \Phi\left(\omega^{c}, \omega^{c}\right)}{\partial \omega_{i}}+\frac{\partial \Phi\left(\omega^{c}, \omega^{c}\right)}{\partial \overline{\omega}_{-i}}=0$$

Why? This is the first order condition that the population will solve if individuals coordinate their decisions with one another.

I now impose two additional assumptions. First,

$$\frac{\partial^2 \Phi\left(\omega_i, \overline{\omega}_{-i}\right)}{\partial \omega_i^2} < 0$$

This looks like a decreasing returns type assumption and is imposed to help ensure a bounded solution for equilibrium choices. Second,

$$\frac{\partial \Phi\left(\omega_i, \overline{\omega}_{-i}\right)}{\partial \overline{\omega}_{-i}} > 0$$

i.e. positive spillovers.

Under these assumptions, it must be the case that

$$\omega^{NC} < \omega^{C}$$

Idea: complementarities in choices may lead to lower choice levels than optimal. School effort?

Multiple Equilibria

From the noncooperative first order condition one can construct the reaction function $\omega_i = \phi\left(\overline{\omega}_{-i}\right)$.

Noncooperative equilibria are fixed points of this map.

In order for there to be multiple equilibria, it is necessary that $\frac{d\phi\left(\overline{\omega}_{-i}\right)}{d\overline{\omega}_{-i}} > 0$ over some part of the support of $\overline{\omega}_{-i}$ (which is of course the same as the support of ω_i).

To calculate this slope, substitute the reaction function into the first order condition

$$\frac{\partial \Phi\left(\phi\left(\overline{\omega}_{-i}\right), \overline{\omega}_{-i}\right)}{\partial \omega_i} = 0$$

and totally differentiate with respect to $\overline{\omega}_{-i}$.

$$\frac{\partial^2 \Phi\left(\phi\left(\overline{\omega}_{-i}\right),\overline{\omega}_{-i}\right)}{\partial \omega_i^2}\frac{d\phi\left(\overline{\omega}_{-i}\right)}{d\overline{\omega}_{-i}}+\frac{\partial^2 \Phi\left(\phi\left(\overline{\omega}_{-i}\right),\overline{\omega}_{-i}\right)}{\partial \omega_i \, \partial \overline{\omega}_{-i}}=0 \Rightarrow$$

$$\frac{d\phi\left(\overline{\omega}_{-i}\right)}{d\overline{\omega}_{-i}}=-\frac{\dfrac{\partial^2 \Phi\left(\phi\left(\overline{\omega}_{-i}\right),\overline{\omega}_{-i}\right)}{\partial \omega_i \, \partial \overline{\omega}_{-i}}}{\dfrac{\partial^2 \Phi\left(\phi\left(\overline{\omega}_{-i}\right),\overline{\omega}_{-i}\right)}{\partial \omega_i^2}}.$$

Hence, complementarity over part of the support of the choice variable is necessary for multiple equilibria to be possible in this model.

Why is the multiple equilibrium property of such interest? The reason is simple. Multiplicity of equilibria means that two groups of individuals with identical preferences and constraints can have very different aggregate outcome.

Hence multiple equilibria is one way to understand group level inequality.

And to the extent that history determines equilibrium selection, we have a basis for understanding how group level inequality is reinforcing.

A Multinomial Logit Approach to Social Interactions

(Brock and Durlauf (2006))

Assumptions

1. Each agent faces a common choice set with L discrete possibilities, i.e. $\Omega_i = \{0,1,\ldots,L-1\}$.

2. Each choice l produces a payoff for i according to:

$$V_{i,l} = h_{i,l} + Jp_{i,l}^{e} + \varepsilon_{i,l}$$

Shorthand for Bayes-Nash. Inessential in this context.

3. Random utility terms $\varepsilon_{i,l}$ are independent across *i* and *l* and are doubly exponentially distributed with index parameter β,

$$\mu\left(\varepsilon_{i,l} \leq \varsigma\right) = \exp\left(-\exp\left(-\beta\varsigma + \gamma\right)\right)$$

where γ is Euler's constant.

Characterizing Choices

These assumptions may be combined to produce a full description of the choice probabilities for each individual.

$$\mu\left(\omega_i = l \middle| h_{i,j}, p_{i,j}^e \ \forall j\right) =$$
$$\mu\left(\operatorname{argmax}_{j\in\{0\ldots L-1\}} h_{i,j} + Jp_{i,j}^e + \varepsilon_{i,j} = l \middle| h_{i,j}, p_{i,j}^e \ \forall j\right)$$

The double exponential assumption for the random payoff terms leads to the canonical multinomial logit probability structure

$$\mu\left(\omega_i = l \middle| h_{i,j}, p^e_{i,j}\ \forall j\right) = \frac{\exp\left(\beta h_{i,l} + \beta J p^e_{i,l}\right)}{\sum_{j=0}^{L-1} \exp\left(\beta h_{i,j} + \beta J p^e_{i,j}\right)}$$

so the joint probabilities for all choices may be written as

$$\mu\left(\omega_1 = l_1, \ldots, \omega_I = l_I \middle| h_{i,j}, p^e_{i,j}\ \forall i, j\right) =$$
$$\prod_i \frac{\exp\left(\beta h_{i,l_i} + \beta J p^e_{i,l_i}\right)}{\sum_{j=0}^{L-1} \exp\left(\beta h_{i,j} + \beta J p^e_{i,j}\right)}$$

Self-Consistency of Beliefs

Self-consistent beliefs imply that the subjective choice probabilities p_l^e equal the objective expected values of the percentage of agents in the group who choose l, p_l, the structure of the model implies that

$$p_{i,l}^e = p_l = \int \frac{\exp\left(\beta h_{i,l} + \beta J p_l\right)}{\sum_{j=0}^{L-1} \exp\left(\beta h_{i,j} + \beta J p_j\right)} dF_h$$

where F_h is the empirical probability distribution for the vector of deterministic terms $h_{i,l}$.

It is straightforward to verify that under the Brouwer fixed point theorem, at least one such fixed point exists, so this model always has at least one equilibrium set of self-consistent aggregate choice probabilities.

Equilibrium reflects interplay of intensity of social interactions effect J, dispersion of unobserved heterogeneity, β, and distribution of private incentives, dF_h

To say more, I examine a special case.

Characterizing Equilibria

To understand the properties of this model, it is useful to focus on the special case where $h_{i,l} = 0 \; \forall i, l$. For this special case, the choice probabilities (and hence the expected distribution of choices within a group) are completely determined by the compound parameter βJ.

An important question is whether and how the presence of interdependencies produces multiple equilibria for the choice probabilities in a population.

In order to develop some intuition as to why the number of equilibria is connected to the magnitude of βJ, it is helpful to consider two extreme cases for the compound parameter, namely $\beta J = 0$ and $\beta J = \infty$.

For the case $\beta J = 0$, one can immediately verify that there exists a unique equilibrium for the aggregate choice probabilities such that $p_l = \frac{1}{L}$ $\forall l$. This follows from the fact that under the assumption that all individual heterogeneity in choices come from the realizations of $\varepsilon_{i,l}$, a process whose elements are independent and identically distributed across choices and individuals. Since all agents are ex ante identical, the aggregate choice probabilities must be equal.

The case $\beta J = \infty$ is more complicated. The set of aggregate choice probabilities $p_l = \frac{1}{L}$ is also an equilibrium if $\beta J = \infty$ since conditional on these probabilities, the symmetries in payoffs associated with each choice that led to this equilibrium when $\beta J = 0$ are preserved as there is no difference in the social component of payoffs across choices.

However, this is not the only equilibrium. To see why this is so, observe that for any pair of choices l and l' for which the aggregate choice probabilities are nonzero, it must be the case that

$$\frac{p_l}{p_{l'}} = \frac{\exp\left(\beta J p_l\right)}{\exp\left(\beta J p_{l'}\right)}$$

for any βJ. This follows from the fact that each agent is ex ante identical. Thus, it is immediate that any set of equilibrium probabilities that are bounded away from 0 will become equal as $\beta J \Rightarrow \infty$. This condition is necessary as well as sufficient, so any configuration such that $p_l = \frac{1}{b}$ for some subset of b choices and $p_l = 0$ for the other $L - b$ choices is an equilibrium. Hence, if $J = \infty$, there exist

$$\sum_{b=1}^{L} \binom{L}{b} = 2^L - 1$$

different equilibrium probability configurations.

Recalling that β indexes the density of random utility and J measures the strength of interdependence between decisions, this case, when contrasted with $\beta J = 0$ illustrates why the strength of these interdependences and the degree of heterogeneity in random utility interact to determine the number of equilibria.

These extreme cases may be refined to produce a more precise characterization of the relationship between the number of equilibria and the value of βJ.

Proposition. Multiple equilibria in the multinomial logit model with social interactions

Assume that $h_{i,l} = k \ \forall i,l$. Then there will exist at least three self-consistent choice probabilities if $\frac{\beta J}{L} > 1$.

Comments

1. There is an interplay of the degree of unobserved heterogeneity and the strength of social interactions that determines the number of equilibria.

2. This is an example of a phase transition.

3. The threshold for multiplicity depends on the number of choices. Perhaps suggests red bus/blue bus problem.

Multinomial Choice Under Alternative Error Assumptions

The basic logic of the multinomial model is straightforward to generalize. This can be seen if one considers the preference structure

$$V_{i,l} = h_{i,l} + Jp^{e}_{i,l} + \beta^{-1}\varepsilon_{i,l}$$

This is the same preference structure we worked with earlier, except that β is handled differently. We assume that these unobserved utility terms are independent and identically distributed with a common distribution function $F_{\varepsilon}(\cdot)$.

For this model, the probability that agent i makes choice l is

$$\mu\left(\begin{array}{l}\varepsilon_{i,0}-\varepsilon_{i,l}\leq\beta\left(h_{i,l}-h_{i,0}\right)+\beta J\left(p_{i,l}^{e}-p_{i,0}^{e}\right),...,\\ \varepsilon_{i,L-1}-\varepsilon_{i,l}\leq\beta\left(h_{i,l}-h_{i,L-1}\right)+\beta J\left(p_{i,l}^{e}-p_{i,L-1}^{e}\right)\end{array}\right)$$

As is standard, conditional on a realization of $\varepsilon_{i,l}$, the probability that l is chosen is

$$\prod_{j\neq i}F_{\varepsilon}\left(\beta h_{i,l}-\beta h_{i,j}+\beta Jp_{i,l}^{e}-\beta Jp_{i,j}^{e}+\varepsilon_{i,l}\right)$$

The probability of the choice l without conditioning on the realization of $\varepsilon_{i,l}$ is

$$p_{i,l} = \int \prod_{j \neq l} F_{\varepsilon}\left(\beta h_{i,l} - \beta h_{i,j} + \beta J p^{e}_{i,l} - \beta J p^{e}_{i,j} + \varepsilon\right) dF_{\varepsilon}$$

This structure of this multinomial choice model whose structure is fully analogous to the multinomial logit structure developed under parametric assumptions. Under self-consistency, the aggregate choice probabilities of this general multinomial choice model are the solutions to

$$p_l = \int\int \prod_{j \neq l} F_{\varepsilon}\left(\beta h_l - \beta h_j + \beta J p_l - \beta J p_j + \varepsilon\right) dF_{\varepsilon} dF_h$$

As in the multinomial logit case, the compound parameter βJ plays a critical role in determining the number of self-consistent equilibrium choice probabilities p_l.

Proposition. Uniqueness versus multiplicity of self-consistent equilibria in multinomial choice models with social interactions

Assume that $h_{i,l} = 0 \ \forall i, l$ and $\varepsilon_{i,l}$ are independent across i and l. There exists a threshold T such that if $\beta J < T$, then there is a unique self-consistent equilibrium, whereas if $\beta J > T$ there exist at least three self-consistent equilibria.

The relationship between βJ and the number of equilibria is less precise than was found for the multinomial logit case, as this proposition does not specify anything about the way in which L, the number of available choices, affects the number of equilibria.

This lack of precision is to be expected since we did not specify the distribution of the errors.

How Should One Think About Poverty Traps?

A low (average) outcome equilibrium occurs when private incentives are weak and social incentives are strong.

Hence private and social theories of poverty traps are complements, not substitutes.

iv. Behavioral Genetics

One possible source of intergenerational persistence of socioeconomic status is via genes. There is, unsurprisingly, little question that an individual’s cognitive skills play an important role in determining socioeconomic outcomes.

That said, there is great controversy (and in my judgment no clear evidence) on the role of genes in determining cognitive and noncognitive skills.

This is not to say that there is no role for genes in transmitting socioeconomic status from parents to children, but rather that empirical social science, despite many strong claims, has failed to establish the empirical salience of genes in explaining intergenerational mobility and cross-sectional inequality.

Further, it is important to keep in mind that even if genes play a first order role, this has no bearing on whether government policies can affect inequality. To use a famous example due to Arthur Goldberger (1979), eyesight may be purely determined by genes, but this does not affect the efficacy of wearing glasses.

The classical model used to measure the respective roles of nature, nurture, and luck can be described as follows. Note that the calculations are expressed in terms of variances instead of covariances; the latter is more standard in the literature, but there is no substantive difference between the approaches. The use of variances makes explicit how contrasts between different types of individuals are the basis of measuring how genes matter.

Define the following variables

$\omega_i =$ outcome of interest

$A_i =$ genetic component

$C_i =$ family or shared environment component, which means that $C_i = C_j$ if i and j are members of the same family

$E_i =$ idiosyncratic component, often called non-shared environment component, in contrast to C_i

All variables are measured as deviations from some mean, and so have expected value 0. Further, since the three components are all latent, without loss of generality, one can assume that

$$\operatorname{var}(A_i) = \operatorname{var}(C_i) = \operatorname{var}(E_i) = 1$$

The classical ACE models of genes and environment is based on the linear relationship

$$\omega_i = aA_i + cC_i + eE_i$$

The object of the literature is to identify the contributions of the three factors to overall variance, in particular the genetic coefficient, a, since this coefficient is the basis for measuring the role of nature versus nurture.

To understand how the literature measures the role of genes, it is first assumed that the different determinants of ω_i are uncorrelated with one another, i.e.

$$\operatorname{cov}\left(A_i, C_i\right) = \operatorname{cov}\left(A_i, E_i\right) = \operatorname{cov}\left(C_i, E_i\right) = 0.$$

so that

$$\operatorname{var}\left(\omega_i\right) = a^2 + c^2 + e^2$$

This makes a variance decomposition meaningful.

One also assumes

$$\operatorname{cov}\left(A_i, C_{i'}\right) = \operatorname{cov}\left(A_i, E_{i'}\right) = \operatorname{cov}\left(C_i, E_{i'}\right) = 0 \text{ if } i \neq i'$$

These no correlation assumptions are important.

For a random sample of individuals it is evident that the variance contributions of the three factors are not identified, since the only observable object is $\mathrm{var}(\omega_i)$. So how does the literature proceed?

The basic idea is to employ twins data and to distinguish between monozygotic (identical) and dizygotic (fraternal) twins.

For notational simplicity, m and d will be used to denote twin pair types and t denotes twins raised together and s denotes twins raised separately.

To see how data on twins can allow for identification, consider data taken from identical twins i and i' raised in different families. Under the orthogonality assumptions, it must be the case that for all pairs (i, i') of separated twins that

$$\mathrm{var}\left(\omega_i - \omega_{i'} \middle| m, s\right) = 2c^2 + 2e^2$$

The observables $\text{var}(\omega_i)$ and $\text{var}(\omega_i - \omega_{i'} | m, s)$ may be employed to calculate the variance contribution of genes to overall socioeconomic outcomes.

$$\frac{a^2}{\text{var}(\omega_i)} = \frac{\text{var}(\omega_i) - .5\,\text{var}(\omega_i - \omega_{i'} | m, s)}{\text{var}(\omega_i)} =$$
$$1 - \frac{\text{var}(\omega_i - \omega_{i'} | m, s)}{2\,\text{var}(\omega_i)}$$

In other words, the genetic contribution to outcome variance is identified.

When twins are not separated, then one needs to use the differences between monozygotic and dizygotic twins in order to make the analogous calculation. For twins raised together, the monozygotic and dizygotic differences have variances

$$\operatorname{var}\left(\omega_i - \omega_{i'} \middle| m,t\right) = 2e^2$$

and

$$\operatorname{var}\left(\omega_i - \omega_{i'} \middle| d,t\right) = a^2 + 2e^2$$

respectively.

These formula imply that

$$a^2 = \mathrm{var}\left(\omega_i - \omega_{i'} \middle| d,t\right) - \mathrm{var}\left(\omega_i - \omega_{i'} \middle| m,t\right)$$

Note that c^2 and e^2 are identified as well.

Where does the empirical literature stand? These types of calculations have been summarized (simplistically) as suggesting

$$\frac{a^2}{\operatorname{var}(\omega_i)} \approx .5$$

$$\frac{c^2}{\operatorname{var}(\omega_i)} \approx 0$$

$$\frac{e^2}{\operatorname{var}(\omega_i)} \approx .5$$

Claims with respect to IQ typically involve higher heritability estimates.

Is this approach to evaluating the role of genes in socioeconomic outcomes credible?

Econometricians who have evaluated the use of twins data to uncover the role of genes, most notably Arthur Goldberger, have concluded that the answer is no. The reason for this rejection of the twins data methodology, which nevertheless continues to be employed in both economics and other social sciences, involves the assumptions that are made to allow for the variance decomposition that lies at the heart of the exercise.

In my judgment, the assumption $\operatorname{cov}\left(A_i, C_i\right) = 0$ is especially problematic

This assumption says that a child's genetic inheritance and parenting experience are uncorrelated. Since the genetic inheritance is of course correlated with the parents' genes, $\operatorname{cov}\left(A_i, C_i\right) = 0$ requires the parent's genes to be uncorrelated with the family environment they create.

This is obviously untenable since the point of the exercise is to understand how genes affect behavior.

There are also reasons to question the assumption that separated twins do not share a common environment. Careful examination of separated twins data has revealed that separation often involves being raised by neighbors or relatives. In other words, it is not the case that separated twins necessarily have uncorrelated family environments.

Periodically, claims have emerged in the social science literature that IQ differences between blacks and whites are determined, to a large extent, by genetic differences.

Such studies are all subject to the critiques that Goldberger and others have raised.

Further, the analyses that make such claims are self-contradictory.

If an analysis assumes that all individuals have the same unconditional mean, which is necessary to justify the removal of the mean from the data, then the analysis has assumed there are no racial differences in IQ.

If race-specific IQ means have been removed, then the analysis has assumed the conclusion.

The problem with the various genetic claims about racial differences is that they confuse first and second moments. The racial difference assertion is about the first moments of the data; the variance decompositions that have been described all involve second moments.

Beyond ACE

It is possible to relax some assumptions in the classical ACE studies by combining the two calculations made above.

Consider the system

$$\operatorname{var}(\omega_i) = a^2 + c^2 + e^2$$

$$\operatorname{var}(\omega_i - \omega_{i'} \mid m,t) = 2e^2$$

$$\operatorname{var}(\omega_i - \omega_{i'} \mid m,s) = 2c^2 + 2e^2$$

$$\operatorname{var}(\omega_i - \omega_{i'} \mid d,t) = a^2 + 2e^2$$

$$\operatorname{var}(\omega_i - \omega_{i'} \mid d,s) = a^2 + 2c^2 + 2e^2$$

This is a system of 5 equations in 3 unknowns. ACE parameters are overidentified. One can relax covariance assumptions. (Ao and Durlauf, in progress). Conti and Heckman earlier (2010) proposed the same idea for measurement systems.

Nature/Nurture Interactions

Modern research into the role of genetic influences has moved beyond the ACE model to consider gene-environment interactions; these are often called $G \times E$ models. Purcell (2002) is a useful introduction to how researchers are proceeding, but unfortunately the new work, in my judgment, introduces interactions in a mechanical way. For example, Purcell (2002) assumes that gene-environment interactions occur via some observable variable M_i and that the form of this interaction leads to

$$\omega_i = \left(a + \beta_A M_i\right) A_i + \left(c + \beta_C M_i\right) C_i + \left(e + \beta_E M_i\right) E_i$$

This functional form is ad hoc. Further, the approach assumes

$$\operatorname{cov}\left(A_i, C_i \middle| M_i\right) = \operatorname{cov}\left(A_i, E_i \middle| M_i\right) = \operatorname{cov}\left(C_i, E_i \middle| M_i\right) = 0.$$

which is simply the questionable orthogonality assumption written in terms of conditional covariances. Again, in my judgment, it is no more plausible because of the conditioning.

Epigenetics

One of the most exciting developments in genetic research involves what is called epigenetics.

Epigenetics refers to the way that the environment influences the expression of genes.

Particular attention has focused on how the environment experienced by a mother affects gene expression during fetal development.

A famous example of this concerns the effects of a famine in the Netherlands during 1945 on fetal development. The important finding is that as adults, the offspring of mothers whose pregnancies overlapped with the famine exhibit greater obesity than others.

The explanation of this finding is that the expression of genes in the developing fetus was influenced by the fact that the mother was experiencing a calorie-deprived environment. This experience caused genes to be expressed that led offspring to crave calories. The triggering mechanism can be explained by evolutionary arguments if our distant ancestors evolved in an environment in which famines, for example, were multigenerational, which seems plausible.

Remarkably, there is even evidence that epigenetic effects can be transgenerational in some species, although nothing is known at this point with respect to humans; see Youngson and Whitelaw (2008) for a survey. But the evidence suggests the possibility that environmental effects on a mother can have persistent consequences across generations.

Genome-Wide Association Studies

My discussion so far has treated A as unobservable. With the emergence of genomic data, social scientists are beginning to examine whether such data are predictive of socioeconomic outcomes. The use of genetic data in this way is called a genome-wide association study (GWAS).

Unfortunately, the GWAS methodology cannot identify gene complexes, which one would expect are the source of emergent properties such as intelligence. The problem is that the DNA sequence is so complicated, that the identification of gene-gene interactions represents the frontier of the literature.

This barrier exists in understanding the genetics of complex diseases such as tuberculosis. In fact, for a number of diseases, geneticists have referred to the lack of GWAS evidence as the missing heritability problem.

A second problem is that attributes such as intelligence are, in my view, likely to be emergent properties from overlapping gene complexes, so studies relating a single gene to politics, etc. are not interpretable.

Bottom Line

There are good reasons to believe that genes matter, yet social scientists have yet to develop persuasive ways to measure the influence.

The development of more credible ways to conceptualize and measure the role of genes is, in my view, very important!

v. Empirical Work on Evaluating Roles of these Mechanisms

While there is an enormous amount of research measuring intergenerational mobility, there is relatively little that attempts to evaluate the relative empirical salience of different mechanisms.

However, a few analyses attempt to address this.

Bowles and Gintis (2002)

Bowles and Gintis interpret

$$\log Y_{i,o} = \alpha + \beta \log Y_{i,p} + \varepsilon_{i,o}$$

via following system in which IQ denotes intelligence, ED denotes education, NC denotes personality (fatalism), W denotes wealth, and ETH denotes race.

$$\log Y_{i,o} = \kappa + \gamma_{IQ} IQ_{i,o} + \gamma_{ED} ED_{i,o} + \gamma_{NC} NC_{i,o} + \gamma_{W} W_{i,o} + \gamma_{ETH} ETH_{i,o} + \xi_{i,o}$$

$$IQ_{i,o} = \alpha_{IQ} + \lambda_{IQ} Y_{i,p} + \varepsilon_{IQ,i,o}$$

$$ED_{i,o} = \alpha_{ED} + \lambda_{ED} Y_{i,p} + \varepsilon_{ED,i,o}$$

$$W_{i,o} = \alpha_{ED} + \lambda_{W} Y_{i,p} + \varepsilon_{W,i,o}$$

$$ETH_{i,o} = \alpha_{ED} + \lambda_{ETH} Y_{i,p} + \varepsilon_{ETH,i,o}$$

Table 3

The Main Causal Channels of Intergenerational Status Transmission in the U.S.

Channel	*Earnings*	*Income*
IQ, conditioned on schooling	0.05	0.04
Schooling, conditioned on IQ	0.10	0.07
Wealth		0.12
Personality (fatalism)	0.03	0.02
Race	0.07	0.07
Total Intergenerational Correlation Accounted For	0.25	0.32

Notes: For each channel, the entry is the correlation of parent income with the indicated predictor of offspring income, multiplied by its normalized regression coefficient in an earnings or income equation. The total is the intergenerational correlation resulting from these channels, in the absence of a direct effect of parents' status on offspring status.
Source: Calculations described in text and Bowles and Gintis (2001).

Blanden, Gregg, and MacMillan (2007)

BGM interpret β

$$\log Y_{i,o} = \alpha + \beta \log Y_{i,p} + \varepsilon_{i,o}$$

the following system, in which cognitive skills are denoted by C, noncognitive skills denoted by NC and and education denoted by ED

$$\log Y_{i,o} = \kappa + \gamma_C C_{i.o} + \gamma_{NC} NC_{i,o} + \gamma_{ED} ED_{i,o} + \xi_{i,o}$$
$$C_{i,o} = \alpha_C + \lambda_C Y_{i,p} + \varepsilon_{NC,i,o}$$
$$NC_{i,o} = \alpha_{NC} + \lambda_{NC} Y_{i,p} + \varepsilon_{NC,i,o}$$
$$ED_{i,o} = \alpha_{ED} + \pi_C C_{i,o} + \gamma_{NC} NC_{i,o} + \lambda_{ED} Y_{i,pi,o} + \varepsilon_{ED,i,o}$$

Comparison of 1958 and 1970 UK cohorts find IGE increased from .205 to .291; model can explain 80% of change; “large part” due to changes in coefficients in equations for C, NC, ED.

Noncog relation to parental income is found to be increasing, but operates via education.