

The Technology of Skill Formation

Pedro Carneiro

University College London and

Institute for Fiscal Studies

Flavio Cunha

University of Chicago

James J. Heckman*

University of Chicago and

The American Bar Foundation

December 29, 2003

*This research was supported by NICHD R01-34598-03.

Abstract

This paper presents formal models of child development that capture the essence of recent findings from the empirical literature on child development. The goal is to provide theoretical frameworks for interpreting the evidence from a vast empirical literature, for guiding the next generation of empirical studies and for formulating policy. We start from the premise that skill formation is a life-cycle process. It starts in the womb and goes on throughout most of the adult life. Families and firms have a role in this process that is at least as important as the role of schools. There are multiple skills and multiple abilities that are important for adult success. Abilities are both inherited and created, and the traditional debate of nature versus nurture is outdated and scientifically obsolete. The technology of skill formation has two additional important characteristics. The first one is that IQ and behavior are more plastic at early ages than at later ages. Furthermore, behavior is much more malleable than IQ as individuals age. The second is that human capital investments are complementary over time. Early investments increase the productivity of later investments. Early investments are not productive if they are not followed up by later investments. The returns to investing early in the life cycle are high. Remediation of inadequate early investments is difficult and very costly.

JEL Code: J31

Pedro Carneiro	Flavio Cunha	James J. Heckman
Department of Economics	Department of Economics	Department of Economics
University College London	University of Chicago	University of Chicago
Gower Street	1126 E. 59th Street	1126 E. 59th Street
London WC1E 6BT	Chicago IL 60637	Chicago IL 60637
UK	USA	USA
Phone: (44-20) 76795824	Phone: (773) 256-6367	Phone: (773) 702-0634
Fax: (44-20) 79162775	Fax: (773) 256-6313	Fax: (773) 702-8490
Email: p.carneiro@ucl.ac.uk	Email: flavio@uchicago.edu	Email: jjh@uchicago.edu

1 Introduction

This paper presents formal models of child development that capture the essence of recent findings from the empirical literature on child development. The goal is to provide theoretical frameworks for interpreting the evidence from a vast empirical literature, for guiding the next generation of empirical studies and for formulating policy.

Recent empirical research in a variety of fields has substantially improved our understanding of how skills and abilities are formed over the life cycle. The early human capital literature (Becker, 1964, Mincer, 1974) viewed human capital as a rival explanation for human ability and emphasized that acquired human capital could explain many features of earnings distributions and earnings dynamics that models of innate cognitive ability could not. This point of view underlies many recent economic models of family influence (*e.g.* Aiyagari, Greenwood, Seshadri, 2002; Laitner, 1992, 1997). Later work (Ben-Porath, 1967 and Griliches, 1977) emphasized that innate ability was an input into the production of human capital, although it was ambiguous about its effect on human capital accumulation. More innate ability could lead to less schooling if all schooling did was teach one what an able person could learn without formal schooling. On the other hand, more innate ability might make learning easier and promote schooling. The signalling literature made the latter interpretation in developing models of schooling that emphasized that higher levels of schooling signalled higher innate ability. In one extreme form, this literature suggested that there was no learning content in schooling.

The entire literature assumed that ability is an innate, scalar, invariant measure of cognitive skill. Except for work by Marxist economists (see, *e.g.* Bowles and Gintis, 1976 and Edwards, 1976), noncognitive traits like motivation, persistence, time preference and self control were neglected and treated as peripheral to the skill formation and earnings determination process.

The literature in economics focuses on liquidity constraints and heritability as the principal sources

of parental influence on child development. Becker and Tomes (1979, 1986) initiated a large literature that emphasizes the importance of credit constraints, family income and inherited ability on the schooling and earnings of children. Important developments of this work by Laitner (1992, 1997), Benabou (2000, 2002), Aiyagari, Greenwood, Seshadri (2002) and Seshadri and Yuki (2003), emphasize the role of credit constraints and altruism in forming the skills of children. Ability is treated as exogenously determined and the lifecycle of the child at home is collapsed into a single period so that there is no distinction between early and late investments in children. Becker and Tomes (1986) suggest that there may be no trade-off between equity and efficiency in government transfer policy because the return to human capital investment is high due to the presence of credit constraints.

Recent research, summarized in Carneiro and Heckman (2003), presents a much richer picture of schooling, life cycle skill formation and earnings determination. It recognizes the importance of both cognitive and noncognitive abilities in explaining schooling and socioeconomic success. These abilities are themselves produced by family and personal actions. Both genes and environments produce these abilities and environments affect genetic transmission mechanisms (See Turkheimer et al., 2003). This interaction has important theoretical and empirical implications.

The following conclusions emerge from the recent empirical literature on child development. Cognitive ability is affected by environmental influences (including in utero experiences) and is formed relatively early (by age 8 or so). It is hard to change IQ after this age. Noncognitive skills (motivation, self-discipline, time preference) associated with development of the child's prefrontal cortex can also be affected by environmental interventions. These skills remain more malleable at later ages than cognitive skills. Noncognitive skills are valued in the market place and also affect academic and social achievement.

Complementarity of investments and self productivity, two distinct ideas folded into one in our previous analyses, are essential features of the skill and ability formation process.¹ Skill begets skill; ability begets

¹Heckman, Lochner and Taber (1998) develop a model in which ability determines schooling and both ability and schooling

ability. Strong complementarity leads to a trade-off between efficiency and equity in considering investments in human capital. Diminishing returns would argue in favor of equalization of investment across persons. Complementarity and self productivity are forces toward specialization of investments made after the early years to certain groups. Disadvantaged young adults with low levels of cognitive and noncognitive skills have lower rates of return to schooling and job training than more advantaged young adults. Due to complementarity, remediation for neglected investment is costly, and may be prohibitively so for the most disadvantaged.

One contribution of our analysis is to place the child development process in a multiperiod context, disaggregating the one period of family influence assumed in a variety of current models into multiple periods. Complementarity and self-productivity of human capital imply an equity-efficiency trade-off for late child investments but no equity-efficiency trade-off for early investments. This has important consequences for the design and evaluation of public policies toward families. In particular, the returns to late childhood investment and remediation for persons from disadvantaged backgrounds is low.

A second contribution of our analysis is to emphasize the secondary importance of credit constraints in the college going years, as traditionally conceived in applied economics in explaining child schooling attainment. Permanent income plays an important role, not income in adolescent years. Carneiro and Heckman (2002, 2003) present evidence for American society that only a small fraction (at most 8%) of American children are credit constrained in making college decisions. The important constraints facing children are ones on their early environment-parental background and motivation and the like. The important market failure is the inability of children to buy their parents and not the inability of families to secure loans for a child's education. This has major implications for the way family policy should be designed, and how to remedy deficits in low income and disadvantaged populations.

determine post school investment. While Ben-Porath (1967) emphasized the self-productivity of human capital, he assumed human capital was homogeneous and did not develop models of heterogeneous skills and abilities.

Controlling for cognitive ability, in American society with current meritocratic policies in place, family income plays only a minor role in determining college enrollment decisions although much public policy is predicated on the opposite point of view. Yet ability itself seems to be determined by early family environments. Permanent income matters in determining schooling and ability, but “cash in advance” credit constraints facing parents in the child’s teenage years do not. Ability has both environmental and genetic components, and environments affect the expression of the genes. Evidence from interventions on disadvantaged populations demonstrate that interventions can raise measured ability but their major impact is on noncognitive abilities. These features are missing from the current literature in economics on child development and our aim is to redress these gaps. They are also ignored in current empirical studies of family and genetic influence. Measured ability is determined in part by environmental factors.

The paper proceeds in the following way. In the next section we present the main characteristics of the technology of skill formation. Then we construct a simple model illustrating the use of this technology. The last section concludes.

This draft is incomplete. In this version, we present the evidence that motivates our work and some very simple models that capture some of the key features of the empirical literature. It is more of a progress report than a finished paper. It sets forth our agenda and the main contours of our work. Much remains to be done and we welcome suggestions. Nonetheless, any finished model that is faithful to the evidence summarized in this paper will stress that (a) parental endowments are key constraints governing family influence in American society; (b) early child investments must be distinguished from late child investments and that an equity-efficiency trade-off exists for late investments but not for early investments and (c) abilities are created, not solely inherited, and are multiple in variety. These insights change the way we interpret the evidence and design public policy. Point (a) is emphasized in many papers. Point (b) is ignored by models that consider only one period of childhood investment. Point (c) has received scant attention in the formal literature on childhood investment.

2 The Technology of Skill Formation

In this section we emphasize some features of the human capital accumulation technology that are important. Some of them have not yet been fully incorporated in economic models. We provide some empirical examples that illustrate the empirical importance of these features. A more complete review of this evidence is provided by Carneiro and Heckman (2003).

Human capital accumulation and skill formation are dynamic processes. The skills acquired in one stage of the life cycle affect both the initial conditions and the technology of learning at the next stage. Human capital is produced over the life cycle by families, schools, and firms, although most discussions of skill formation focus on schools as the major producer of abilities and skills, despite a substantial body of evidence that families and firms are also major producers of abilities and skills. Skill formation starts in the womb and takes place throughout the whole life of the individual. Over one half of lifetime human capital is acquired through post-school investments (Heckman, Lochner and Taber, 1998).

A major determinant of successful schools is successful families. Schools work with what parents bring them. They operate more effectively if parents reinforce them by encouraging and motivating children. Job training programs, whether public or private, work with what families and schools supply them and cannot remedy twenty years of neglect. Children from disadvantaged families may suffer from a lack of resources invested in them, or they may have parents that lack the information necessary to make adequate investments in their children, even if resources are made available (for example, through state programs), because of poor education or the like. It is easier to compensate for low current funds (if parents borrow against future consumption to finance current investments in their children) than against low parental human capital.

Abilities are both inherited and created. As summarized in Shonkoff and Phillips (2000), the “long standing debate about the importance of nature versus nurture, considered as independent influences,

is overly simplistic and scientifically obsolete”. They write: “Scientists have shifted their focus to take account of the fact that genetic and environmental influences work together in dynamic ways over the course of development. At any time, both are sources of human potential and growth as well as risk and dysfunction. Both genetically determined characteristics and those that are highly affected by experience are open to intervention. The most important questions now concern how environments influence the expression of genes and how genetic make-up, combined with children’s previous experiences, affects their ongoing interactions with their environments during the early years and beyond.” Hansen, Heckman and Mullen (2003) show that schooling affects cognitive ability. Becker and Mulligan (1997) argue that parents can invest in and manipulate their children’s discount rate, which can be broadly interpreted as another type of ability.

A study of human capital policy grounded in economic and scientific fundamentals improves on a purely empirical approach to policy evaluation that relies on evaluations of the programs and policies in place or previously experienced. Although economic policy analysis should be grounded in data, it is important to recognize that the policies that can be evaluated empirically are only a small subset of the policies that might be tried. If we base speculation about economic policies on economic fundamentals, rather than solely on estimated “treatment effects” that are only weakly related to economic fundamentals, we are in a better position to think beyond what has been tried to propose more innovative solutions to human capital problems.

Carneiro and Heckman (2003) investigate the study of human capital policy by placing it in the context of economic models of life cycle learning and skill accumulation rather than focusing exclusively on which policies have “worked” in the past. This paper extends their analysis by presenting formal models of the investment process.

Figure 1 summarizes the major finding of Carneiro and Heckman and the motivation for this paper. It plots the rate of return to human capital at different stages of the life cycle for a person of given

abilities. The horizontal axis represents age, which is a surrogate for the agent's position in the life cycle. The vertical axis represents the rate of return to investment assuming the same amount of investment is made at each age. *Ceteris paribus* the rate of return to a dollar of investment made while a person is young is higher than the rate of return to the same dollar made at a later age. Early investments are harvested over a longer horizon than those made later in the life cycle (Becker, 1964). In addition, because early investments raise the productivity (lower the costs) of later investments, human capital is synergistic. Learning begets learning; skills (both cognitive and noncognitive) acquired early on facilitate later learning. Early deficits make later remediation difficult. Finally, young children's cognition and behavior are more easily malleable than cognition and behavior in adults: even in the absence of dynamic complementarity, early investments are more productive than late investments. For an externally specified opportunity cost of funds r (represented by the horizontal line with intercept r in figure 1), an optimal investment strategy is to invest less in the old and more in the young. At any age, investment is more profitable for persons with higher innate ability. Figure 2 presents the optimal investment quantity counterpart of figure 1.

Carneiro and Heckman (2003) develop an alternative interpretation of figure 1 as an empirical description of the economic returns to investment at current levels of spending in the American economy. The return to investment in the young is high; the return to investments in the old and less able is quite low. A socially optimal investment strategy would equate returns across all investment levels. A central empirical conclusion of their analysis is that at current investment levels, efficiency in public spending would be enhanced if human capital investment were directed more toward the young and away from older, less-skilled, and illiterate persons for whom human capital is a poor investment.

2.1 Multiple Skills, Plasticity, Self-Productivity and Dynamic Complementarity

In the rest of this section we examine in more detail three important features of the technology of skill formation: 1) multiple skills; 2) plasticity; 3) self-productivity and dynamic complementarity. By multiple skills we mean that there exists a multiplicity of skills which are important for an individual's success in life. By plasticity, we mean that the malleability an individual's IQ and behavior traits changes (decreases) as people age. By self-productivity and dynamic complementarity we mean that skill begets skill. Late investments are complements with early investments in the production of human capital. Without early investments late investments are unproductive. Conversely, complementarity also implies that early investments that are not followed up by later investments may not be productive either.

The analysis in Carneiro and Heckman (2003) and in this paper challenges the conventional point of view that equates skill with intelligence. It draws on a body of research that demonstrates the importance of both cognitive and noncognitive skills in determining socioeconomic success. Heckman, Hsee and Rubinstein (2001) and Heckman and Rubinstein (2002) provide evidence of the importance of noncognitive skills from an analysis of the GED program. GED recipients are high school dropouts who get a high school certification through the GED. In terms of cognitive ability, they are as smart as regular high school graduates. This is shown in figure 3, that plots AFQT distributions for high school graduates and GED recipients for different demographic groups in the NLSY. However, table 1 presents the coefficients of a log wage regression on GED recipiency and high school graduation and shows that GED recipients have much lower wages than high school graduates. Furthermore, they have lower wages than regular high school dropouts with the same level of cognitive ability. This means they lack some other skill, which we interpret as a non-cognitive skill. Table 2 shows that GED recipients are also more likely to exhibit disruptive behavior in school and work, and higher turnover rates on the job, than either high school

graduates or high school dropouts. They lack skills such as motivation and discipline. These skills are important in the labor market. Gaps in non-cognitive measures (such as anti-social behavior) by family income appear very early in the life-cycle, as documented in figure 4 and in the work of Carneiro and Heckman (2003).

Current educational policy and economic analysis focuses on tested academic achievement as the major output of schools. Proposed systems for evaluating school performance are often premised on this idea. Economic models of signaling and screening assume that predetermined cognitive ability is an important determinant, if not the most important determinant, of academic and economic success. Recent evidence challenges this view. No doubt, cognitive ability is an important factor in schooling and labor market outcomes. At the same time, noncognitive abilities, although harder to measure, also play an important role.

Recent studies in child development (e.g. Shonkoff and Phillips 2000) emphasize that different stages of the life cycle are critical to the formation of different types of abilities. When the opportunities for formation of these abilities are missed, remediation is costly, and full remediation is often prohibitively costly. These findings highlight the need to take a comprehensive view of skill formation over the life cycle that is grounded in the best science and economics so that effective policies for increasing the low level of skills in the workforce can be devised.

Both cognitive and noncognitive skills are affected by families and schools, but they differ in their malleability over the life cycle, with noncognitive skills being more malleable than cognitive skills at later ages. This finding is supported by studies of early childhood interventions that primarily improve noncognitive skills, with substantial effects on schooling and labor market outcomes, but only weakly affect cognitive ability. Table 3 shows that the well known early childhood programs have short lasting effects on IQ but long lasting effects on achievement and behavioral outcomes of disadvantaged children. Mentoring programs in the early teenage years can also affect these skills (see Carneiro and Heckman, 2003). Current

analyses of skill formation focus too much on cognitive ability and too little on noncognitive ability in evaluating human capital interventions, and in formalizing the skill formation process.

Differences in levels of cognitive and noncognitive skills by family income and family background emerge early and persist. If anything, schooling widens these early differences. The work of Carneiro, Heckman and Masterov (2003) on sources of racial skill differential is illustrative of this claim. As shown in figure 5, test score gaps across race groups emerge very early (the graph displays the density of math scores at age 5 for white males in different race groups, using the Children of NLSY). Figure 6 plots the effect of schooling on test scores for different demographic groups in the NLSY. Test scores grow at a much slower rate for blacks than for whites as children from both race groups progress through school.

The idea of self-productivity of human capital investments is rather old in economics and is developed in the work of Ben-Porath (1967) who specifies a production function where the stock human capital increases the productivity of additional investments in human capital: human capital is a crucial input in the production of more human capital. Becker and Tomes (1986) specify a production function where innate ability increases the productivity of parental investments in the child's human capital. The stock of ability and human capital, and further investments in human capital are complementary inputs in the production of skill. Complementarity also means that the costs of remediating the neglect of early investments in human capital can be very high, if remediation investments have no solid (human capital) base to build on. It also means that if early investments are not followed up by later investments then their effect on the amount of skill accumulated by early adulthood may be small.

Carneiro and Heckman (2003) summarize a body of evidence that suggests that complementarity is empirically important. Table 4, from their paper, shows that white males in the High School and Beyond with higher levels of cognitive ability have higher returns to college than individuals with lower levels of cognitive ability.² Those who know more to start with benefit more from the college experience. This

²These estimates correct for the endogeneity of schooling and account for heterogeneity in the returns to schooling, both

finding is replicated in other datasets. Table 5 shows that individuals with higher ability and education are more likely to participate in company training than those with lower ability and education levels. Individuals with higher levels of human capital receive higher investments through company training than those with low levels of human capital. This is a common finding in the job training literature. A final example comes from the work of Currie and Thomas (1995) who study the Head-Start program and conclude that the overall effects of this program on test scores are lower for black than for white children, as seen in table 6 (the relevant parameter is the coefficient on Head Start participation from the set of columns that include mother fixed effects, in panel A). In fact, the panel B of table 7 shows that the effect of the program on test scores at the age the program ends is about the same for blacks and whites (the direct effect of Head Start). There is no difference on the effect of Head Start participation on PPVT scores between blacks and whites at the age they leave the program (see the third column of the first line of panel B of this table). However the fade out effects after exit from Head Start are much larger for black children. That is why a few years after these children have left the program we still see some impact on test scores for whites but no impact of Head Start on test scores for blacks (as shown in table 6). These fade out effects are estimated from the interaction of Head Start participation with age, and presented in the second line of panel B of this table. In another paper, Currie and Thomas (2000) suggest that these differential fade out effects may be due to the fact that black Head Start children go on to attend much lower quality schools than white Head Start children. Head Start investments are followed up by very poor schooling for black children and therefore it is not surprising that the final effect of Head Start on test scores of blacks is small.³ The productivity of early investments that are not followed up by later investments can be very small. There is another aspect to complementarity that should be emphasized:

in terms of observable and unobservable variables.

³In other analysis of the Head Start data, Currie, Garces and Thomas (2003) show that Head Start has important effects on high school graduation, wages and criminal behavior of adults. The effect on criminal behavior is very strong for blacks. Although the program had a small effect on black test scores it had a large effect on black adult outcomes through its effect on behavioral skills.

early deficits are hard to remediate with later investments, and the cost of remediation can be prohibitively high because the productivity of late investments is very small in the absence of early investments. The whole literature on public job training shows that it is hard to remediate the neglect of skill investment in childhood and adolescence (see *e.g.* Lalonde, 1995, Heckman, Lalonde and Smith, 1999, and Carneiro and Heckman, 2003).

The ideas put forth so far can be formalized in a simple two period CES production function (easily generalizable to multiple periods):

$$H = A \left[\gamma_0 (K_0)^\phi + \gamma_1 (K_1)^\phi \right]^{\frac{\rho}{\phi}} h^\alpha$$

where H is the final human capital of the child, A is ability, h is the human capital of parents, K_0 and K_1 are early and late investments. Later we can allow H , K_0 and K_1 to be vectors of skills and vectors of investments and therefore have multiple skills. The Ben-Porath (1967) technology is a special case of the one we have here. γ_0 and γ_1 are the plasticity parameters. If γ_1 is smaller than γ_0 then plasticity is smaller at later ages than at early ages. The term $\frac{1}{1-\phi}$ is the elasticity of substitution. When ϕ is zero we have a Cobb-Douglas technology. As ϕ approaches $-\infty$ the technology gets closer and closer to the Leontieff function. In the appendix we embed this technology in a dynastic model of human capital investment and simulate the model. Figure 7 (which comes from the simulation of this model) illustrates how the costs of late remediation of poor early investments change when the elasticity of substitution changes. When complementarity increases, remediation costs increase as well.

We use this technology in a parental investment model where we allow investment to take place in multiple periods. This is a simple but important extension of the traditional model of Becker and Tomes (1979, 1986). In the appendix we present an overlapping generations model with altruism, with human

capital investment, uncertainty and credit constraints. In this model parents are altruistic and can invest in children over two (or more) periods: early childhood and adolescence. Parental human capital (and family and neighborhood environments) is an input into the production of the child's human capital, as are the child's innate ability and the resources invested in the child in both periods. Early investments may be limited by several reasons, such as low parental human capital, or low availability of funds for early investments. Scarcity of funds at early ages can be compensated if parents face rising income and can postpone their consumption until the end of the early childhood of their child (substitute present and future consumption).⁴ Low parental human capital cannot be easily substituted at early ages. A family can be credit constrained in both investment periods, or in only one of them. This model operationalizes the idea of short run and long run credit constraints of Cameron and Heckman (1998, 2001) and Carneiro and Heckman (2002). The government can intervene to remedy poor investments and poor environments in disadvantaged families. Interventions can come in early childhood and in late adolescence. Remediation of poor early investments in these families is very costly but may be granted on the grounds of equity. Interventions in early childhood may be both efficient and equitable. The model presented in the appendix is very incomplete but is illustrative of our current work (Carneiro, Cunha and Heckman, 2003).⁵

⁴In this model each parent only has one child, although this assumption can be relaxed.

⁵Consider the case in which parents can insure perfectly against idiosyncratic innovations in income, but cannot buy financial claims contingent on realizations of the ability shock, which follows a first-order Markov process. Given the child's ability a , the parental human capital h and the value of financial claims the parent receives as bequest b , the parent decides how much early-investment x to perform on the child, how much to consume and how much to buy in securities $s(\varepsilon)$. Notice that $s(\varepsilon)$ pays one unit of consumption good if the innovation in income is ε and zero otherwise. Then, given the realization of the innovation in income ε , the financial claims $s(\varepsilon)$, the parental human capital h , the early investment x , and the child's ability a , the parent decides how much late-investment to perform on the child. In this simple version, we assume that the parent can either send the child to college or not. Let $z = 1$ if the child is sent to college, and $z = 0$ otherwise. This simple model replicates closely some empirical regularities consistent with the literature summarized above. First, the lifetime returns to late-investment (i.e. $z = 1$) are about 22% for those who actually are invested on in the late stage. The figure for those who do not get invested on is only around 10% as shown in table 1 in the appendix. Table 2 in the appendix breaks up the above calculation per ability group. We denote group 1 the lowest ability group, and group 15 the highest ability group. Notice that the returns to investment in the late stage are roughly increasing in ability and around the range of 19%-27%, while those who do not get any investment face returns in the range of 5%-14%. Table 3 in the appendix shows the intergenerational mobility table. The element a_{ij} of this table reports the probability that a child is in the j -th decile of the present value of earnings distribution given that the parent is in the i -th decile. The table shows very little persistence, which is a result of the complete set of insurance contracts against the innovations in income.

3 Conclusion

This paper presents formal models of child development that capture the essence of recent findings from the empirical literature on child development. The goal is to provide theoretical frameworks for interpreting the evidence from a vast empirical literature, for guiding the next generation of empirical studies and for formulating policy. We start from the premise that skill formation is a life-cycle process. It starts in the womb and goes on throughout most of the adult life. Families and firms have a role in this process that is at least as important as the role of schools. There are multiple skills and multiple abilities that are important for adult success. Abilities are both inherited and created, and the traditional debate of nature versus nurture is outdated and scientifically obsolete. The technology of skill formation has two additional important characteristics. The first one is that IQ and behavior are more plastic at early ages than at later ages. Furthermore, behavior is much more malleable than IQ as individuals age. The second is that human capital investments are complementary over time. Early investments increase the productivity of later investments. Early investments are not productive if they are not followed up by later investments. The returns to investing early in the life cycle are high. Remediation of inadequate early investments is difficult and very costly.

References

- [1] Aiyagari, S.R, J.Greenwood and A.Seshadri, "Efficient Investment in Children," *Journal of Economic Theory*, 2002, Vol.102, pp.290-321.
- [2] Becker, Gary. 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. New York: Columbia University Press.
- [3] Becker, Gary and Casey Mulligan, 1997, "On the Endogenous Determination of Time Preference", *The Quarterly Journal of Economics*, Vol. 112, No 3, pp 729-758.
- [4] Becker, G. and N.Tomes, 1979. "An Equilibrium Theory of The Distribution of Income and Intergenerational Mobility," *Journal of political Economy*, Vol 87, pp.1153-1189.
- [5] Becker, G. and N.Tomes, 1986. "Human Capital and The Rise and Fall of Families," *Journal of Labor Economics* Vol.4(3): S1-S39.
- [6] Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review* 90, March 2000, 96-129.
- [7] Benabou, Roland. 2002. "Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica* 70(2), March 2002, 481-517.
- [8] Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle Earnings," *Journal of Political Economy*, 75, no.4 (part 1): 352-365.
- [9] Bowles, Samuel and Howard Gintis. 1976. *Schooling in Capitalist America*. New York: Basic.
- [10] Cameron, S., and J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males" *Journal of Political Economy*, 106, no.2: 262-333.

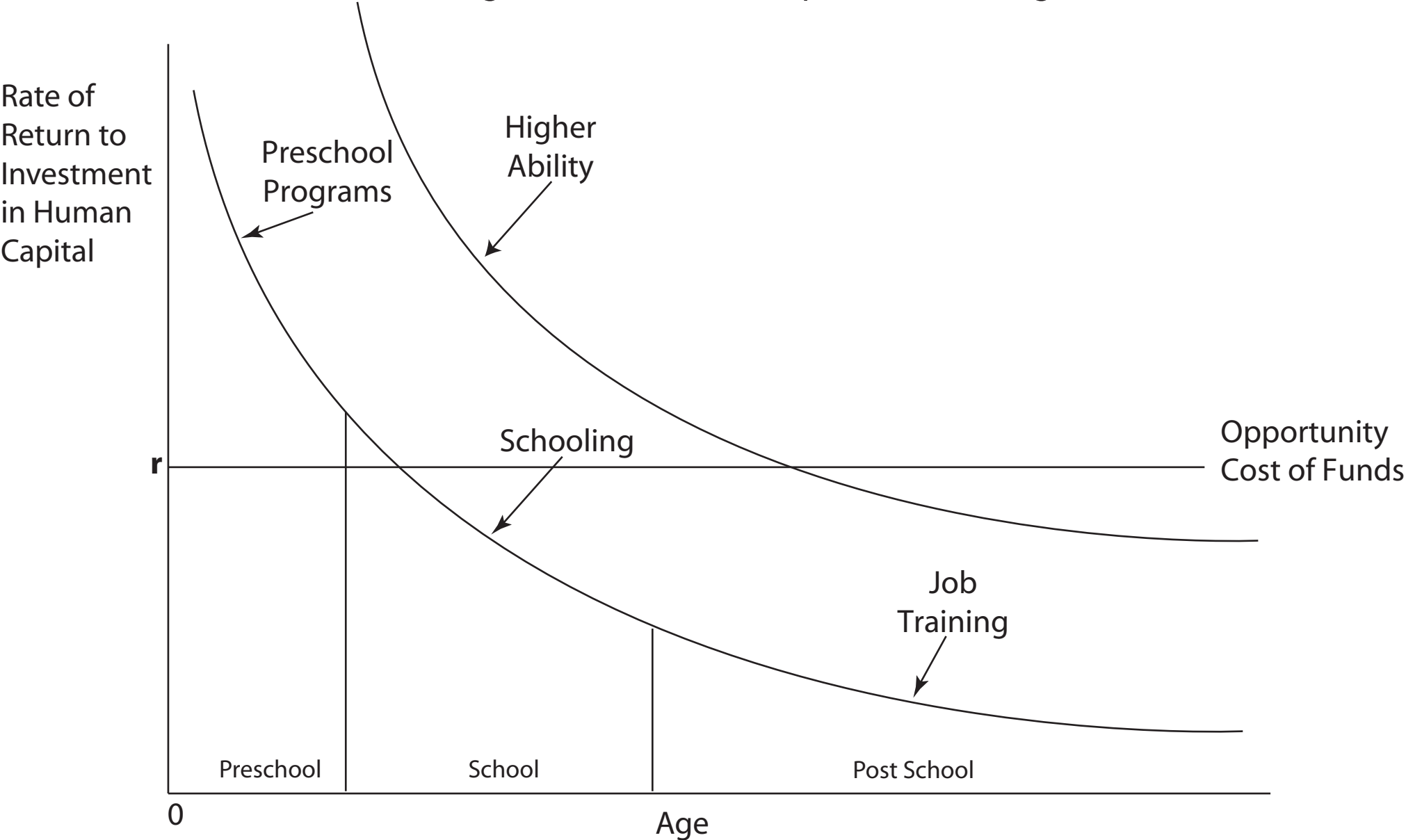
- [11] Cameron, S., and J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males" *Journal of Political Economy* 109, no.3: 455-499.
- [12] Carneiro, Pedro, Flavio Cunha and James J. Heckman, 2003, "Interpreting the Evidence of Family Influence on Child Development", paper presented at the Economics of Early Childhood Conference, Minneapolis Fed, October 2003.
- [13] Carneiro, Pedro, and James J. Heckman. 2002. "The Evidence on Credit Constraints in Post-secondary Schooling." *Economic Journal* 112, no.482: 705-734.
- [14] Carneiro, Pedro and James J. Heckman, (2003). "Human Capital Policy," Forthcoming in J. Heckman and A. Krueger (eds.), *Inequality in America: What Role for Human Capital Policy?* (Cambridge, MA: MIT Press).
- [15] Carneiro, Pedro, James J. Heckman and Dimitriy Masterov (2003) "Racial Gaps in Pre-Market Factors", NBER working paper no 10068.
- [16] Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review* 85, no.3: 341-364.
- [17] Currie, Janet and Duncan Thomas, 2000, "School Quality and the Long Term Effects of Head Start", *Journal of Human Resources*, Vol 35, No 4, pp 755-774.
- [18] Edwards, Richard. 1976. "Individual Traits and Organizational Incentives: What Makes a Good Worker?" *Journal of Human Resources* 11, no.1: 51-68.
- [19] Garces, Eliana, Duncan Thomas and Janet Currie. 2002. "Longer-Term Effects of Head Start." *American Economic Review* 92, no.4: 999-1012.

- [20] Griliches, Zvi. 1977. "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica* 45, no.1: 1-22.
- [21] Hansen, Karsten, James J. Heckman, and Kathleen Mullen. 2003. "The Effect of Schooling and Ability on Achievement Test Scores," Forthcoming, *Journal of Econometrics*.
- [22] Heckman, James J., Jingjing Hsee, and Yona Rubinstein. 2002. "The GED is a 'Mixed Signal': The Effect of Cognitive and Noncognitive Skills on Human Capital and Labor Market Outcomes." University of Chicago working paper.
- [23] Heckman, James J., Robert LaLonde, and Jeffrey Smith. 1999. "The Economics and Econometrics of Active Labor Market Programs." In *Handbook of Labor Economics*, vol. 3A, Orley Ashenfelter and David Card, eds. Amsterdam: Elsevier.
- [24] Heckman, James J., Lance Lochner and Christopher Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1, no. 1: 1-58.
- [25] Heckman, James J., and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91, no.2: 145-149.
- [26] Laitner, J. 1992, "Random Earnings Differences, Lifetime Liquidity Constraints and Altruistic Intergenerational Transfers," *Journal of Economic Theory*, Vol.58, pp.135-170.
- [27] Laitner, J. 1997, "Intergenerational and Interhousehold Economic Links," in M.Rosenzweig and Oded Stark, eds, *Handbook of Population and Family Economics*, Amsterdam: North Holland.
- [28] LaLonde, Robert, 1995, "The Promise of Public Sector-Sponsored Training Programs," *Journal of Economic Perspectives*, 9(2): 149-68.

- [29] Mincer, Jacob, 1974, *Schooling, Experience and Earnings*, New York: Columbia University Press.
- [30] Seshadri, A. and K.Yuki, "Equity and Efficiency Effects of Redistributive Policies," University of Wisconsin, Madison, June, 2003.
- [31] Shonkoff, Jack, and Deborah Phillips, eds. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, D.C.: National Academy Press.
- [32] Turkheimer, E., A.Haley, M.Waldron, B.D'Onofrio and I. Gottesman 2003, "Socioeconomic Status Modifies heritability of IQ in Young Children," *Psychological Science* Vol.14, #6 pp.623-628.

Figure 1

Rates of Return to Human Capital Investment Initially
Setting Investment to be Equal Across all Ages



Rates of Return to Human Capital Investment Initially Setting Investment to be Equal Across all Ages

Figure 2

Optimal Investment Levels

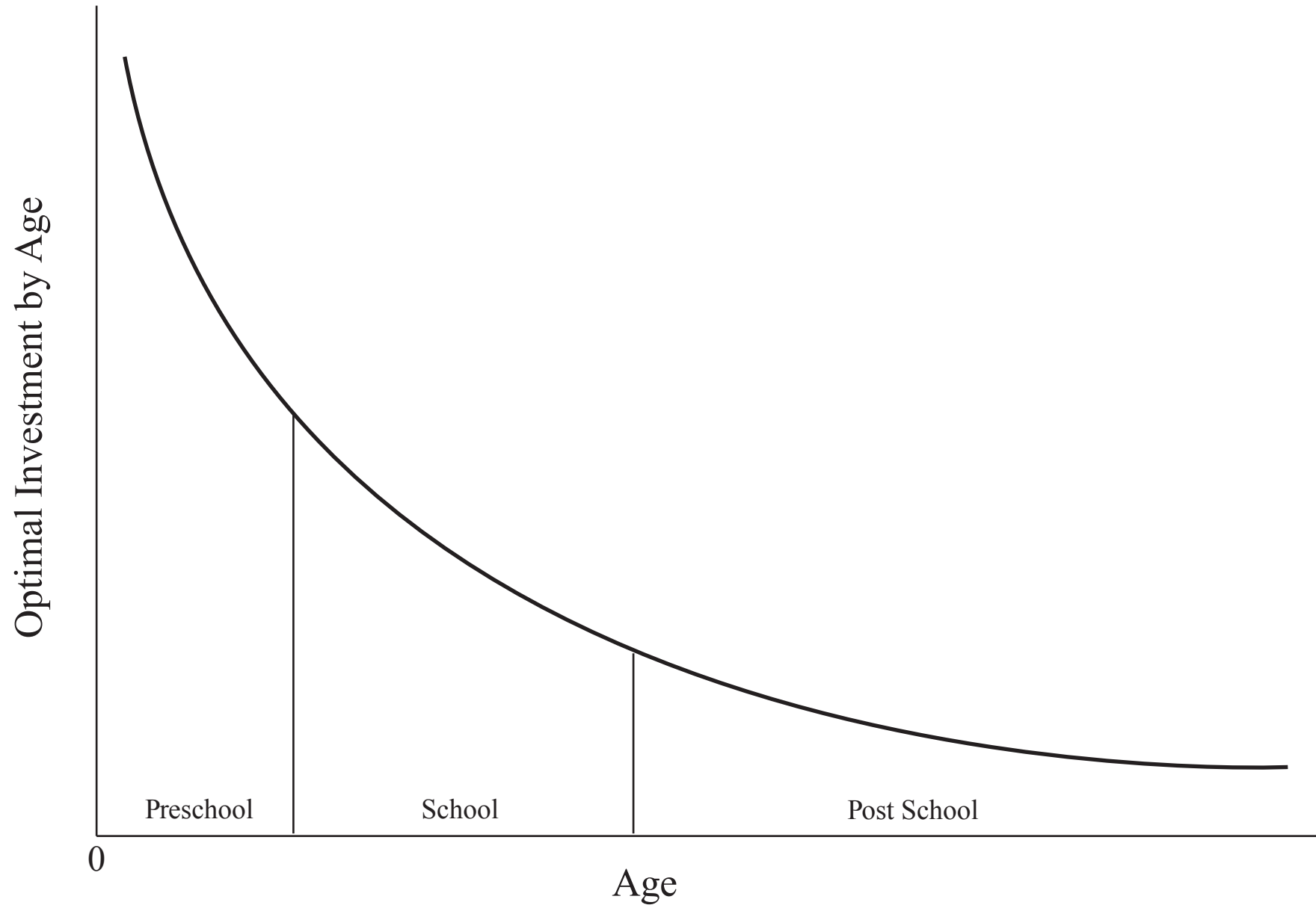
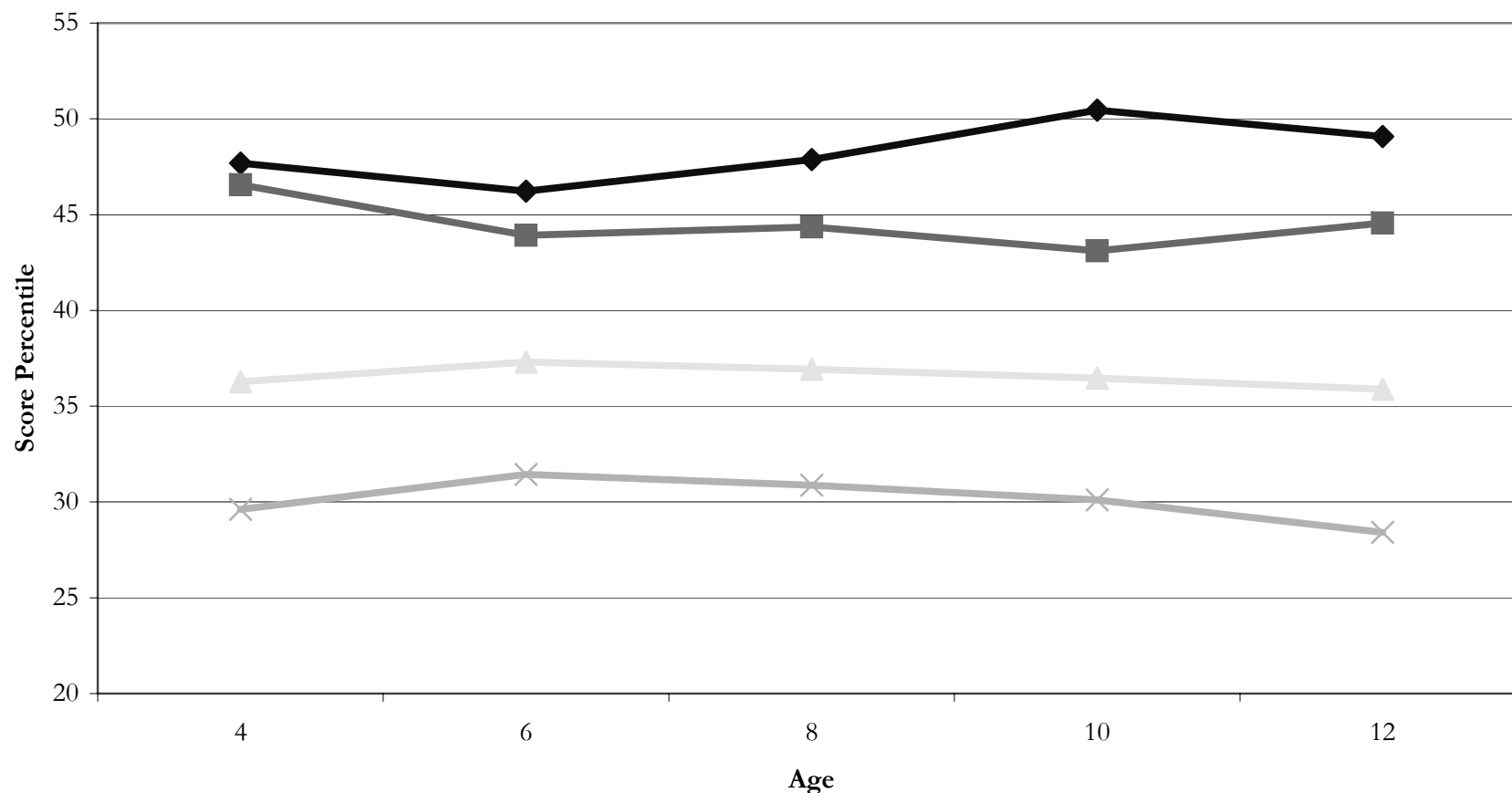


Figure 3

Children of NLSY

Average Percentile Rank on Anti-Social Score, by Income Quartile*

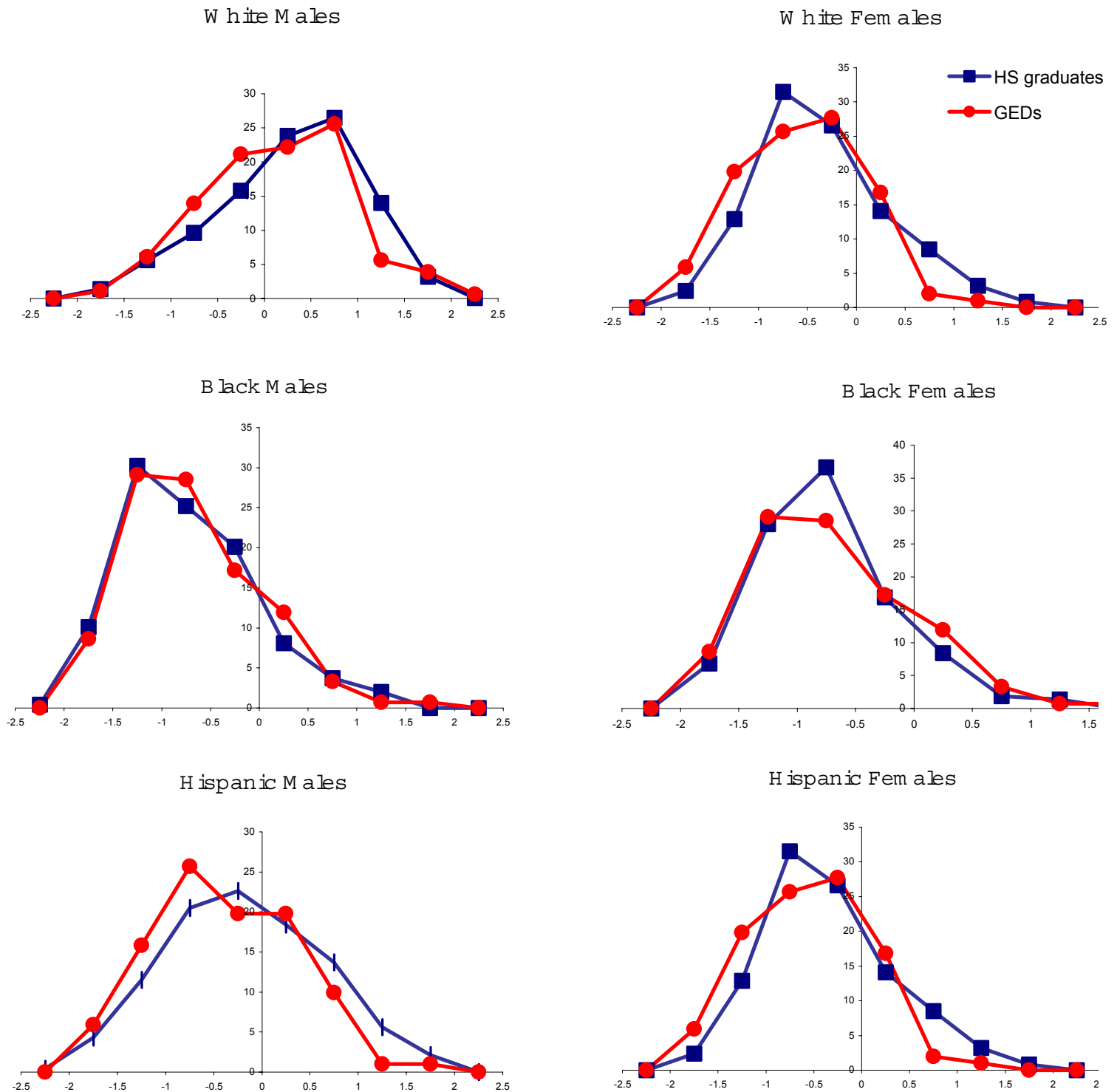


*The income measure we use is average family income between the ages of 6 and 10. Income quartiles are then computed from this measure of income

The higher the anti-social score the worse is the behavior of the child.

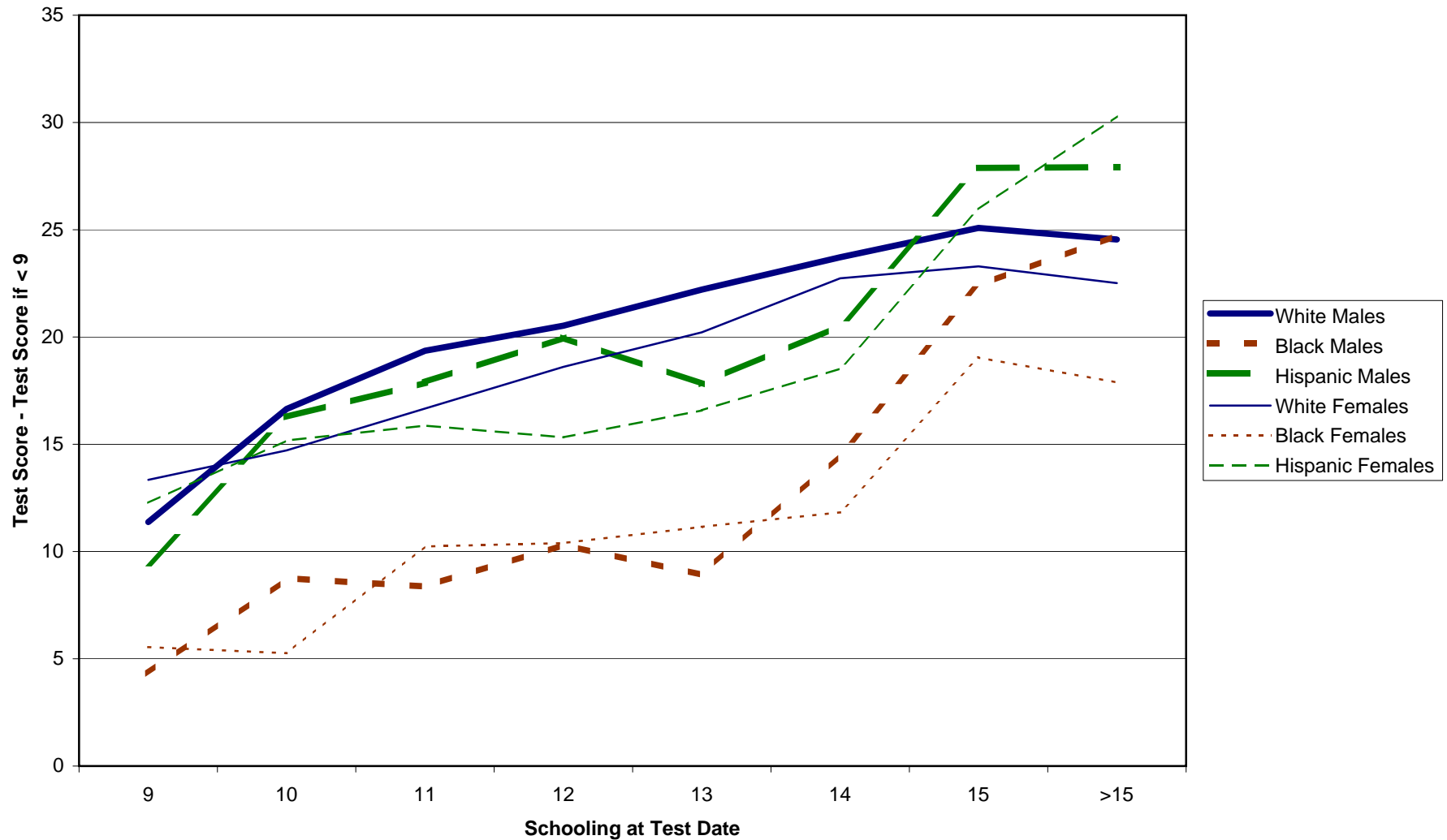
—◆— Lowest Income Quartile —■— Second Income Quartile —▲— Third Income Quartile —x— Highest Income Quartile

Figure 4
Density of Age Adjusted AFQT Scores,
GED Recipients and High School Graduates with Twelve Years of Schooling



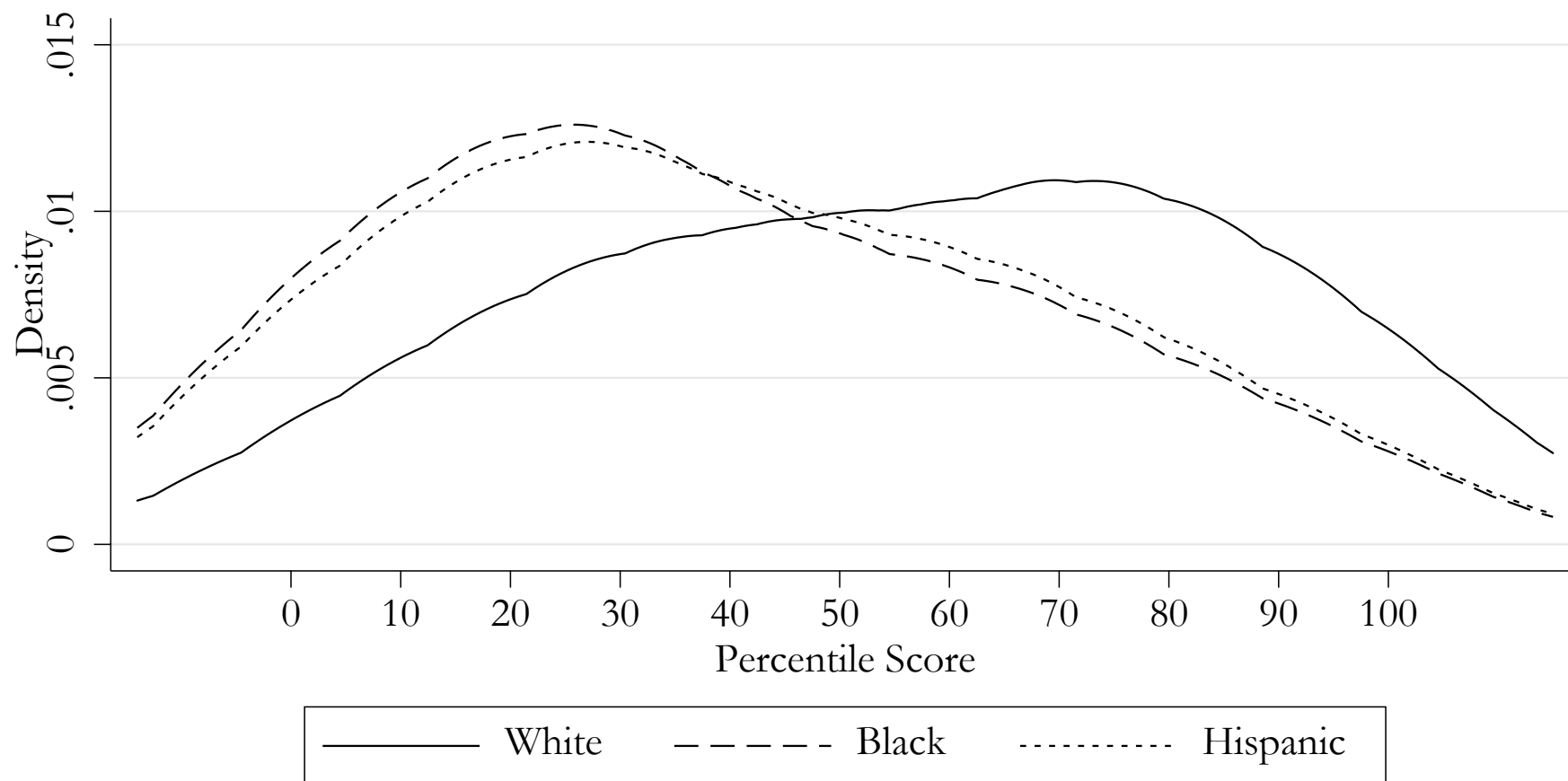
Source: Heckman, Hsee and Rubinstein (2001)

Figure 5 - Effect of Schooling on AFQT for Different Demographic Groups, NLSY79
Coefficients from a Regression of AFQT on Schooling at Test Date and Completed Schooling



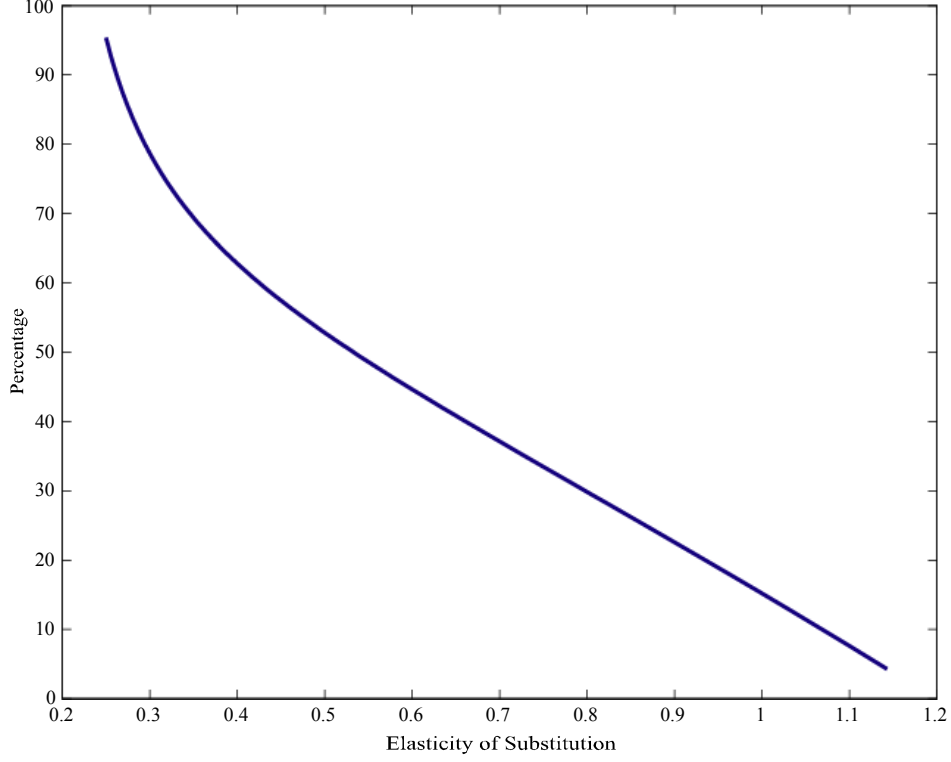
This graph shows the effect of schooling at test date on AFQT scores for different demographic groups in the NLSY. It plots the coefficients on schooling at test date of a regression of AFQT scores on schooling at test date and complete schooling (see Hansen, Heckman and Mullen). The baseline category is 8 years of schooling. For example, white males with 9 years of schooling at test date score 12 points higher on the AFQT than white males with 8 years of schooling. White males with 15 years of schooling score 25 points higher on the AFQT than white males with 8 years of schooling.

Figure 6
Density of Percentile PIAT Math Scores at Ages 5–6
CNLSY 79 Males



This test measures the child's attainment in mathematics as taught in mainstream education. It consists of 84 multiple-choice questions of increasing difficulty, beginning with recognizing numerals and progressing to geometry and trigonometry. The percentile score was calculated separately for each sex at each age.

Figure 7: Percentage Increase in Investment in Period 1 Relative to Unconstrained Amount of Investment in Period 1 Needed to Remedy Low Investment at Period 0



Let $K_1(\xi)$ and $K_1^*(\xi)$ denote the optimal and remediation investments in period 1. In this figure we plot $\frac{K_1^*(\xi) - K_1(\xi)}{K_1(\xi)}$. For each value of the elasticity of substitution ξ , we compute the steady state stock of human capital $H(\xi)$. We take this as the target. We then set the parental human capital $H_p(\xi)$ 2.5% below $H(\xi)$. We then compute K_0 , the investment in period 0, by approximating the policy function $g(H_p(\xi))$ linearly around the steady state. We then use the production function to determine the remediation investment in period 1 that is needed to obtain $H(\xi)$ given initial conditions $H_p(\xi)$ and $K_0 = g(H_p(\xi))$.

Table 1:
How Do Labor Markets Treat the GED Recipients?
A First Glance at the Data
High School Dropouts, GED Recipients and High School Graduates

Variable	OLS		
	(i)	(ii)	(iii)
High school dropout	-0.273 (0.024)	-0.193 (0.026)	-0.022 (0.033)
GED degree	-0.181 (0.039)	-0.187 (0.038)	-0.107 (0.038)
Armed Forces Qualifying Test*		0.106 (0.013)	0.074 (0.014)
Years of schooling			0.070 (0.011)
Training			0.029 (0.005)
GED-HSD <i>F</i> -test: probability>F: GED=HSD	0.092 {0.029}	0.006 {0.876}	-0.085 {0.039}
Observations	12824	12824	12824
Individuals	1288	1288	1288
R-square	0.140	0.161	0.183

Notes:

The table reports results for a sub-sample of white males aged 20-36 from the NLSY

The sub-sample excludes GED recipients who got their degree at age 16 or 17.

All specifications include control for: (1) experience, (2) county level unemployment rate, (3) region of residence, (4) and cohort of birth.

* Age-adjusted to 0 mean in the population sample

High school dropouts are those who dropped out of school and did not get a GED diploma

GED recipients are those who dropped out of school and get a GED diploma.

High school graduates who graduated high school and did not take further schooling.

() Standard errors in parenthesis.

Table 2:
Illicit and Delinquent Activity by Whites,
Shown Separately for High School Dropouts, GED Recipients, and High School Graduates.

	Males [^]			Females ^{^^}		
	HS Dropouts	GED Recipients	HS Graduates	HS Dropouts	GED Recipients	HS Graduates
Index of illicit activity (ILA) ~	0.11 (0.012)	0.18* (0.017)	0.05 (0.006)	-0.01 (0.013)	0.05* (0.015)	-0.04 (0.004)
<u>Particular questions:</u>						
Skipped school in last year	0.13 (0.023)	0.10 (0.030)	0.00 (0.011)	0.00 (0.030)	0.13* (0.035)	0.00 (0.011)
Shoplifted last year	0.05 (0.027)	0.15* (0.039)	0.01 (0.014)	0.00 (0.038)	0.17* (0.045)	-0.03 (0.014)
Smoked pot last year	0.14 (0.029)	0.26* (0.037)	0.03 (0.016)	0.05 (0.044)	0.27* (0.043)	0.03 (0.017)
Used drugs last year	0.10 (0.026)	0.26* (0.039)	0.03 (0.013)	0.09 (0.038)	0.24* (0.045)	0.03 (0.013)
Ever stopped by police	0.16 (0.028)	0.25* (0.039)	0.09 (0.014)	-0.03 (0.030)	0.00 (0.035)	-0.09 (0.009)

Notes: The table shows means (with standard errors in parenthesis) from the NLSY
~ ILA is the average score on the 22 yes/no questions regarding illicit and delinquent behavior.

Responses are age-adjusted and standardized to 0 mean in the population sample

[^] The male sample excludes males reporting being in prison, for any period of time, in the years 1979-1994

^{^^} The female sample excludes teenage mothers

HSD = high school dropouts who do not get a GED degree

GED = GED recipients

HSG = high school graduates who do not take further schooling (12 years of schooling)

* Significantly different from HSD figures at the 5 percent level

Table 3
Outcomes of Early Intervention Programs

Program (Years of Operation)		Outcome	Followed Up to Age	Age of Treatment Effect*	Control Group	Change in Treated Group
Cognitive Measures	Early Training Project (1962 - 1965)	IQ	16 - 20	6	82.8	+12.2
	Perry Preschool Project (1962 - 1967)	IQ	27	7	87.1	+4.0
	Houston PCDC (1970 - 1980)	IQ	8 - 11	2	90.8	+8.0
	Syracuse FDRP (1969 - 1970)	IQ	15	3	90.6	+19.7
	Carolina Abecedarian (1972 - 1985)	IQ	21	12	88.4	+5.3
	Project CARE (1978 - 1984)	IQ	4.5	3	92.6	+11.6
	IHDP (1985 - 1988)	IQ (HLBW sample)	8	8	92.1	+4.4
Educational Outcomes	Early Training Project	Special Education	16 - 20	18	29%	-26%
	Perry Preschool Project	Special Education	27	19	28%	-12%
		High School Graduation		27	45%	+21%
	Chicago CPC (1967 - present)	Special Education	20	18	25%	-10%
		Grade Retention		15	38%	-15%
		High School Graduation		20	39%	+11%
	Carolina Abecedarian	College Enrollment	21	21	14%	+22%
Economic Outcomes	Perry Preschool Project	Arrest Rate	27	27	69%	-12%
		Employment Rate		27	32%	+18%
		Monthly Earnings		27	\$766	+ \$453
		Welfare Use		27	32%	-17%
	Chicago CPC (preschool vs. no preschool)	Juvenile Arrests	20	18	25%	-8%
	Syracuse FDRP	Probation Referral	15	15	22%	-16%
	Elmira PEIP (1978 - 1982)	Arrests (HR sample)	15	15	0.53	-.029

Notes: HLBW = heavier, low birth weight sample; HR = high risk. *Age when treatment effect was last statistically significant.

Cognitive measures include Stanford-Binet and Weshler Intelligence Scales, California Achievement Tests, and other IQ and achievement tests measuring cognitive ability. All results significant at .05 level or higher.

Source: Karoly (2001) For a discussion of the specific treatments offered under each program, see Heckman (2000) and Karoly (2001).

Table 4
Return to one year of college for individuals
at different percentiles of the math test score distribution
White males from High School and Beyond

	5%	25%	50%	75%	95%
Average return in the population	0.1121 (0.0400)	0.1374 (0.0328)	0.1606 (0.0357)	0.1831 (0.0458)	0.2101 (0.0622)
Return for those who attend college	0.1640 (0.0503)	0.1893 (0.0582)	0.2125 (0.0676)	0.2350 (0.0801)	0.2621 (0.0962)
Return for those who do not attend college	0.0702 (0.0536)	0.0954 (0.0385)	0.1187 (0.0298)	0.1411 (0.0305)	0.1682 (0.0425)
Return for those at the margin	0.1203 (0.0364)	0.1456 (0.0300)	0.1689 (0.0345)	0.1913 (0.0453)	0.2184 (0.0631)

Wages are measured in 1991 by dividing annual earnings by hours worked per week multiplied by 52. The math test score is an average of two 10th grade math test scores. There are no dropouts in the sample and the schooling variable is binary (high school - college). The gross returns to college are divided by 3.5 (average difference in years of schooling between high school graduates that go to college and high school graduates that do not in a sample of white males in the NLSY). To construct the numbers in the table we proceed in two steps. First we compute the marginal treatment effect using the method of local instrumental variables as in Carneiro, Heckman and Vytlačil (2001). The parameters in the table are different weighted averages of the marginal treatment effect. Therefore, in the second step we compute the appropriate weight for each parameter and use it to construct a weighted average of the marginal treatment effect (see also Carneiro, 2002). Individuals at the margin are indifferent between attending college or not.

Table 5
Average marginal effect of AFQT, family income, grade completed and father's education
on participation in company training

Variables	Average marginal effect					
	White males		Black males		Hispanic males	
	(1)	(2)	(1)	(2)	(1)	(2)
Age-adjusted AFQT	0.0149 (0.0024)	- -	0.0182 (0.0033)	- -	0.0066 (0.0037)	- -
Family income in 1979 (in \$10,000)	-0.0021 (0.0012)	-0.0005 (0.0011)	-0.0047 (0.0024)	-0.0019 (0.0023)	0.0011 (0.0024)	0.0015 (0.0023)
Grade completed	0.0382 (0.001)	- -	0.0060 (0.0014)	- -	0.0036 (0.0014)	- -
Father's education	-0.0014 (0.0006)	0.0007 (0.0005)	0.0003 (0.0008)	0.0010 (0.0008)	0.0002 (0.0007)	0.0008 (0.0007)
	White females		Black females		Hispanic females	
	(1)	(2)	(1)	(2)	(1)	(2)
	(1)	(2)	(1)	(2)	(1)	(2)
Age-adjusted AFQT	0.0076 (0.0025)	- -	0.0169 (0.0038)	- -	0.0159 (0.0045)	- -
Family income in 1979 (in \$10,000)	-0.0007 (0.0011)	0.0001 (0.0011)	-0.0006 (0.0024)	0.0014 (0.0023)	-0.0065 (0.0031)	-0.0043 (0.0029)
Grade completed	0.0027 (0.0010)	- -	0.0014 (0.0016)	- -	0.0013 (0.0016)	- -
Father's education	0.0001 (0.0006)	0.0009 (0.0006)	0.0015 (0.0008)	0.0021 (0.0008)	-0.00001 (0.0009)	0.0007 (0.0008)

Note: The panel data set was constructed using NLSY79 data from 1979-1994. Data on training in 1987 is combined with 1988 in the original data set. Company training consists of formal training conducted by employer, and military training excluding basic training.

Specification (1) includes a constant, age, father's education, mother's education, number of siblings, southern residence at age 14 dummy, urban residence at age 14 dummy, and year dummies.

Specification (2) drops age-adjusted AFQT and grade completed. Average marginal effect is estimated using average derivatives from a probit regression. Standard errors are reported in parentheses.

Table 6

EFFECT OF PARTICIPATION IN HEAD START AND PRESCHOOL ON PPVT SCORE
AND ABSENCE OF GRADE REPETITION

Variable	OLS-unadjusted			OLS-adjusted			Mother fixed effects		
	White (i)	African- American (ii)	Difference (iii)	White (iv)	African- American (v)	Difference (vi)	White (vii)	African- American (viii)	Difference (ix)
<i>A. Dependent Variable: PPVT Score</i>									
Head Start ^a	-5.621 (1.570)	1.037 (1.223)	-6.658 (1.990)	-0.383 (1.453)	0.739 (1.135)	-1.122 (1.844)	5.875 (1.520)	0.247 (1.358)	5.628 (2.038)
Other preschool ^b	9.077 (1.275)	2.007 (1.481)	7.070 (1.955)	1.679 (1.171)	-0.790 (1.311)	2.469 (1.759)	1.173 (1.296)	0.615 (1.296)	0.557 (1.833)
Constant	31.512 (0.783)	13.762 (0.823)	17.749 (1.136)	-106.706 (16.306)	-49.201 (15.846)	-57.505 (22.737)			
<i>F</i> (Head Start = preschool)	75.38 [0.00]	0.40 [0.53]	36.22 [0.00]	1.56 [0.21]	1.21 [0.27]	2.77 [0.10]	7.45 [0.01]	0.06 [0.81]	4.81 [0.03]
<i>F</i> (all covariates)	43.62 [0.00]	0.99 [0.37]	133.49 [0.00]	71.51 [0.00]	15.70 [0.00]	79.78 [0.00]	3.75 [0.00]	3.13 [0.00]	4.31 [0.00]
<i>R</i> ²	0.03	0.01	0.14	0.27	0.19	0.34	0.73	0.68	0.75
Sample size	2,319	1,158	3,477	2,319	1,158	3,477	2,319	1,158	3,477
<i>B. Dependent Variable: Probability Never Repeated Grade</i>									
Head Start ^a	-0.035 (0.058)	-0.010 (0.061)	-0.025 (0.084)	0.004 (0.061)	0.000 (0.064)	-0.004 (0.088)	0.473 (0.122)	0.008 (0.098)	0.465 (0.158)
Other preschool ^b	0.029 (0.062)	-0.069 (0.085)	0.098 (0.104)	-0.005 (0.063)	0.100 (0.088)	0.095 (0.106)	0.061 (0.099)	0.163 (0.125)	-0.102 (0.158)
Constant	0.654 (0.031)	0.537 (0.043)	0.118 (0.052)	0.487 (0.810)	0.049 (0.882)	0.572 (1.191)			
<i>F</i> (Head Start = preschool)	0.76 [0.38]	0.47 [0.49]	1.20 [0.27]	0.02 [0.90]	1.30 [0.26]	0.61 [0.44]	8.40 [0.01]	1.22 [0.27]	8.05 [0.01]
<i>F</i> (all covariates)	0.39 [0.68]	0.34 [0.72]	2.82 [0.02]	2.50 [0.00]	1.15 [0.32]	2.21 [0.00]	3.57 [0.00]	1.26 [0.28]	2.35 [0.01]
<i>R</i> ²	0.01	0.01	0.01	0.08	0.05	0.08	0.62	0.59	0.61
Sample size	414	314	728	414	314	728	414	314	728

Notes: Standard errors are reported in parentheses below the coefficients; *p* values are given in brackets below the *F* statistics. Variance-covariance matrices were estimated by the method of infinitesimal jackknife for PPVT scores. OLS-adjusted regressions include controls for child age, gender, and whether first born, (log) household permanent income, mother's education, mother's AFQT score, mother's height, number of siblings when the mother was age 14, and grandmother's education. Fixed-effect models include controls for child age, gender, whether first born, and household income at age 3.

^aDummy variable = 1 if participated in Head Start.

^bDummy variable = 1 if participated in other preschool.

Table 7 — FIXED-EFFECTS ESTIMATES OF IMPACT OF HEAD START AND PRESCHOOL ON CHILD WELL-BEING, INCLUDING INTERACTIONS WITH MATERNAL HUMAN CAPITAL AND CHILD AGE

Variable	Dependent variable: PPVT score			Dependent variable: probability never repeated grade			Dependent variable: probability of measles immunization		
	White (i)	African- American (ii)	Difference (iii)	White (iv)	African- American (v)	Difference (vi)	White (vii)	African- American (viii)	Difference (ix)
<i>A. Include interactions with AFQT of mother:</i>									
Head Start ^a	4.826 (2.136)	-0.462 (1.821)	5.288 (2.807)	0.123 (0.186)	-0.006 (0.146)	0.130 (0.239)	0.046 (0.047)	0.083 (0.050)	-0.036 (0.069)
Head Start × AFQT of mother	2.032 (3.352)	2.103 (4.810)	-0.072 (5.863)	0.831 (0.323)	0.040 (0.316)	0.791 (0.452)	0.060 (0.062)	0.030 (0.099)	0.029 (0.119)
Other preschool ^b	2.278 (2.170)	-1.300 (1.483)	3.578 (2.628)	0.217 (0.204)	0.210 (0.192)	0.007 (0.281)	0.086 (0.044)	0.048 (0.049)	0.038 (0.067)
Other preschool × AFQT of mother	-1.396 (2.724)	4.545 (3.764)	-5.941 (4.647)	-0.203 (0.246)	-0.135 (0.419)	-0.068 (0.473)	0.045 (0.044)	0.007 (0.082)	0.038 (0.095)
<i>F</i> (Head Start and interaction)	7.72 [0.00]	0.10 [0.91]	3.39 [0.03]	11.48 [0.00]	0.01 [0.99]	5.39 [0.01]	4.04 [0.02]	4.00 [0.02]	0.16 [0.85]
<i>F</i> (Preschool and interaction)	0.74 [0.48]	0.74 [0.48]	1.04 [0.35]	0.59 [0.56]	0.89 [0.41]	0.02 [0.98]	14.14 [0.00]	1.12 [0.33]	0.87 [0.42]
<i>F</i> (all covariates)	3.74 [0.00]	3.12 [0.00]	4.29 [0.00]	3.79 [0.00]	0.95 [0.48]	2.26 [0.00]	154.10 [0.00]	80.26 [0.00]	117.00 [0.00]
<i>R</i> ²	0.73	0.68	0.75	0.63	0.59	0.62	0.69	0.68	0.69
<i>B. Include Interactions with Age of Child:</i>									
Head Start ^a	6.878 (2.397)	6.845 (1.933)	0.033 (3.080)	0.266 (0.311)	0.218 (0.295)	0.048 (0.429)	0.266 (0.045)	0.258 (0.048)	0.008 (0.067)
Head Start × age of child ^c	-0.192 (0.410)	-1.278 (0.309)	1.086 (0.513)	0.025 (0.036)	-0.025 (0.033)	0.050 (0.049)	-0.043 (0.008)	-0.035 (0.007)	-0.008 (0.011)
Other preschool ^b	0.165 (1.832)	2.970 (1.863)	-2.805 (2.613)	0.173 (0.350)	0.726 (0.461)	-0.553 (0.572)	0.128 (0.031)	0.045 (0.046)	0.083 (0.057)
Other preschool × age of child ^c	0.264 (0.362)	-0.467 (0.386)	0.731 (0.529)	-0.014 (0.041)	-0.074 (0.059)	0.061 (0.071)	-0.002 (0.006)	0.002 (0.009)	-0.004 (0.011)
<i>F</i> (Head Start and interaction)	7.89 [0.00]	8.86 [0.00]	5.26 [0.01]	7.68 [0.00]	0.29 [0.75]	4.78 [0.01]	18.53 [0.00]	15.00 [0.00]	0.48 [0.617]
<i>F</i> (Preschool and interaction)	0.64 [0.53]	1.27 [0.28]	0.96 [0.38]	0.25 [0.78]	1.69 [0.19]	0.50 [0.61]	13.73 [0.00]	1.21 [0.30]	1.46 [0.23]
<i>F</i> (all covariates)	3.74 [0.00]	3.19 [0.00]	4.31 [0.00]	2.76 [0.01]	1.17 [0.32]	1.92 [0.02]	160.23 [0.00]	85.57 [0.00]	122.61 [0.00]
<i>R</i> ²	0.73	0.68	0.75	0.62	0.59	0.61	0.69	0.69	0.69

Notes: Standard errors are reported in parentheses below the coefficients; *p* values are given in brackets below the *F* statistics. The variance-covariance matrix for PPVT models was calculated by the method of infinitesimal jackknife. All models include controls for child age, gender, whether first born, and household income at age 3.

^a Dummy variable = 1 if participated in Head Start.

^b Dummy variable = 1 if participated in other preschool.

^c Age of child is expressed as years since age 5.

1 Appendix - Embedding the Skill Technology in a Dynastic Model with Human Capital Investments with Uncertainty and Credit Constraints

1.1 Generational Structure

The environment is an economy with an infinite number of periods, each one denoted $t \in \{0, 1, 2, \dots\}$. In each period there are generations that overlap. Each generation consists of a continuum of agents that live for four periods. We denote these periods by child, adolescent, young, old. At the end of each period t the old adults die and they are replaced by the children spawned by the young adults. These children will then become adolescents at period $t+1$, young adults at period $t+2$, and old adults at period $t+3$. Life goes on in the future in similar fashion.

1.2 Ability and Human Capital

Each agent is born with innate ability a . We assume that child's ability a follows a first-order Markov process of the form:

$$a = \mu a_{-1} + \nu$$

with $\mu \in (0, 1)$ and $\nu \sim N(0, \sigma_\nu^2)$. We denote by A the invariant distribution of ability and by \mathcal{A} its support. The distribution function A is a primitive of the model. The invariance assumption generates the implication that the cross-sectional distribution of ability will be given by $A_t = A$ for all t . The ability of a child born in period t is perfectly known.

Adults differ in terms of their human capital h . Parents can influence the productivity of their offspring by investing in the education of their kid during childhood and adolescent periods. Consider a period- t parent. That means that the parent was born at period $t-2$ and his kid was born in the current period t . Let h denote the human capital of the parent. Assume that the parent invests resources x, z during the childhood and adolescent years of his offspring. The kid will grow up and become an adult person with human capital h' as described by:

$$h' = H(a, h, x, z) \tag{1}$$

The function H is taken to be a primitive of the model.

Assumption A1 The production function of human capital has the following properties:

H is strictly increasing in all its arguments.

H is a strictly concave function in h .

One example that satisfies assumption A1 is the Constant-Elasticity-of-Substitution production function:

$$h' = \left[\gamma_a a^\phi + \gamma_h h^\phi + \gamma_x x^\phi + \gamma_y z^\phi \right]^{\frac{\rho}{\phi}}, \quad \phi \leq 1, 0 < \rho < 1$$

Another example is:

$$h' = \begin{cases} e^a h^\lambda [\gamma_x x]^\rho, & \text{if } z = 0 \\ e^a h^\lambda [\gamma_x x^\phi + \gamma_y z^\phi]^{\frac{\rho}{\phi}}, & \text{if } z > 0 \end{cases}$$

for $0 < \lambda < 1, \phi \leq 1, 0 < \rho < 1$.

1.3 The Problem of the Agent

Adults are subject to income innovations when young and old. Let ε, η denote the shocks in income of the young and the old. We assume that the shocks are independently and identically distributed. We denote by G^ε, G^η the distributions of ε, η , respectively.

The Budget Constraint We start by assuming that the agents can buy full insurance against the idiosyncratic shocks in income. However, grandparents cannot buy insurance against shocks in the ability of the grandchild. Let $s(\eta)$ denote the amount of securities bought by the young parent that pays one unit of consumption good if the idiosyncratic shock in the income of the agent when old parent is η and zero otherwise. Let $q_s(\eta)$ denote the price of such security. Accordingly, let $b(\varepsilon)$ denote the amount of securities that the old parent leaves as bequest to his kid and that pays one unit of consumption good if the child's income idiosyncratic shock when young parent is ε and zero otherwise. We denote by $q_b(\varepsilon)$ the price of such claim. Let c_t^{t-2}, h_t^{t-2} denote the consumption and human capital stock of the period- t young parent. Let x_t^t denote the investment in the human capital of period- t child made by his parent. We assume that x_t^t can take on a discrete number of values that are in the set $X = \{x_1, x_2, \dots, x_n\}$ with $x_1 < x_2 < \dots < x_n$. Let w denote the wage rate. Since we focus on steady states, we do not index w according to time. The budget constraints the agent faces when young is:

$$c_t^{t-2} + w x_t^t + \int s(\eta) q_s(\eta) d\eta = w h_t^{t-2} + \varepsilon_t + b_t(\varepsilon_t)$$

To describe the old parent's budget constraint, we note that there are fixed costs of investing in the human capital of the kid when he is adolescent. We assume that this fixed cost is constant over time. Let c_{t+1}^{t-2} be the consumption of the period- $(t+1)$ old parent (who is the period- t young parent). Let p denote the fixed cost of investment in the human capital of the period- $(t+1)$ adolescent (again, the period- t child). Let z_{t+1}^t denote the amount invested in the period- $(t+1)$ adolescent. We assume that z_{t+1}^t can take on values in the set $Z = \{0, z_1, z_2, \dots, z_m\}$ with $z_1 < z_2 < \dots < z_m$. The budget constraint of the old parent is:

$$c_{t+1}^{t-2} + (w z_{t+1}^t + p) + \int b(\varepsilon) q_b(\varepsilon) d\varepsilon = w h_t^{t-2} + \eta_t + s_{t+1}(\eta_t), \text{ if } z_{t+1}^t > 0$$

$$c_{t+1}^{t-2} + \int b(\varepsilon) q_b(\varepsilon) d\varepsilon = w h_t^{t-2} + \eta_t + s_{t+1}(\eta_t), \text{ if } z_{t+1}^t = 0$$

Preferences We assume that preferences are represented by the CRRA utility function:

$$u(c) = \frac{c^{1-\sigma} - 1}{1-\sigma}$$

The Problem of the Parent The problem of the period- t young parent is:

$$V(\varepsilon_t, a_t, b_t, h^{t-2}) = \max \left\{ u(c_t^{t-2}) + \beta \int W(\eta_{t+1}, a_t, s_{t+1}, h^{t-2}, x_t^t) dG^\eta(\eta_{t+1}) \right\}$$

subject to:

$$c_t^{t-2} + wx_t^t + \int s(\eta_{t+1}) q_s(\eta_{t+1}) d\eta_{t+1} = wh_t^{t-2} + \varepsilon_t + b_t(a_t, \varepsilon_t)$$

$$x_t^t \in X$$

$$\eta_{t+1} \sim iid G^\eta$$

and the problem of the period- $t+1$ old parent is

$$W(\eta_{t+1}, a_t, s_{t+1}, h_t^{t-2}, x_t^t) = \max \left\{ u(c_t^{t-2}) + \beta \theta \int \int V(\varepsilon_{t+2}, a_{t+2}, b_{t+2}, h^t) \right\}$$

subject to:

$$c_{t+1}^{t-2} + (wz_{t+1}^t + p) + \int b(\varepsilon) q_b(\varepsilon) d\varepsilon = wh^{t-2} + \eta_t + s_{t+1}(\eta_t), \text{ if } z_{t+1}^t > 0$$

$$c_{t+1}^{t-2} + \int b(\varepsilon) q_b(\varepsilon) d\varepsilon = wh^{t-2} + \eta_t + s_{t+1}(\eta_t), \text{ if } z_{t+1}^t = 0$$

$$z_{t+1}^t \in Z$$

$$h^t = H(a_t, x_t^t, z_{t+1}^t, h^{t-2})$$

$$a_{t+2} \sim iid A$$

$$\varepsilon_{t+2} \sim iid G^\varepsilon$$

1.4 Goods Production

There are two inputs in the production function of goods: physical capital and labor, which measured in efficiency units. Let \mathbf{k}, \mathbf{l} denote the aggregate quantities of physical capital and labor, respectively. Let \mathbf{y} denote the aggregate output. The production technology is represented by the production function f :

$$\mathbf{y} = F(\mathbf{k}, \mathbf{l})$$

Assumption A2 The production function of aggregate output has the following properties:
 F is twice-continuously differentiable.
 F is strictly increasing in all its arguments.
 F satisfies the Inada Conditions.
 F presents constant returns to scale.
 F is a strictly concave function.

1.5 Educational Sector

There is an educational sector that produces goods for investment in human capital. This sector does not use physical capital as input, only labor. The production technology is represented by the production function J :

$$\mathbf{e} = J(\mathbf{u})$$

Assumption A3 The production function of education goods has the following properties:
 J is linear.

1.6 Notation

In what follows, let $a \in \mathcal{A}$ denote the ability of a child born in period $t - 1$ from a parental with human capital $h \in \mathcal{H}$. Let $x_{t-1}^{t-1}(a, h), z_t^{t-1}(a, h)$ denote the early and late investments received by such a child born in period $t - 1$. The distribution of efficiency units h' of period $t + 1$ young parents is given by:

$$G_{t+1}^h(h') = \Pr \{(a, h) \in \mathcal{A} \times \mathcal{H} / H[a, h, x_{t-1}(a, h), z_t(a, h)] \leq h'\}$$

The set of period $t + 1$ young adults become skilled is given by:

$$\mathcal{S}_{t+1} = \{(a, h) \in \mathcal{A} \times \mathcal{H} / z_t(a, h) > 0\}$$

Let \mathbf{x}_t denote the period- t aggregate level of early investment in human capital of the children born in period t . Accordingly, G_t^h is the distribution of parental human capital h in period t . Then:

$$\mathbf{x}_t = \int_{\mathcal{A}} \int_{\mathcal{H}} x_t^t(a, h) dA(a) dG_t^h(h)$$

Let \mathbf{z}_t denote the period- t aggregate level of early investment in human capital of the children born in period $t - 1$. Accordingly, G_{t-1}^h is the distribution of parental human capital h in period t . Then:

$$\mathbf{z}_t = \int_{\mathcal{A} \cap \mathcal{S}_t} \int_{\mathcal{H} \cap \mathcal{S}_t} z_t^{t-1}(a, h) dA(a) dG_{t-1}^h(h)$$

1.7 Feasibility

The aggregate output of the goods sector in period t , \mathbf{y}_t , may be used as aggregate consumption of the young parents, \mathbf{c}_t^{t-2} , aggregate consumption of the old parents, \mathbf{c}_t^{t-3} , or aggregate investment \mathbf{i}_t :

$$\mathbf{c}_t^{t-2} + \mathbf{c}_t^{t-3} + \mathbf{i}_t = \mathbf{y}_t$$

Let $\mathbf{b}_{t+1}^{t-3}, \mathbf{s}_{t+1}^{t-2}$ denote the period- t aggregate old parents' aggregate bequest and young parents' savings, in period $t + 1$ dollars, respectively. The equilibrium condition of the physical capital market implies:

$$\frac{\mathbf{b}_{t+1}^{t-3} + \mathbf{s}_{t+1}^{t-2}}{1 + r_{t+1}} = \mathbf{k}_{t+1}$$

Let $\mathbf{h}_t^{t-3}, \mathbf{h}_t^{t-2}$ denote the period- t aggregate stock of efficiency units of old and young parents, respectively. We know that:

$$\mathbf{h}_t^{t-3} = \int_{\mathcal{H}} h dG_{t-1}^h(h)$$

$$\mathbf{h}_t^{t-2} = \int_{\mathcal{H}} h dG_t^h(h)$$

Normalizing the size of each cohort to 1 it follows that the period- t aggregate stock of human capital is:

$$\mathbf{h}_t = \mathbf{h}_t^{t-3} + \mathbf{h}_t^{t-2}$$

The aggregate stock of physical capital may be allocated to goods or educational sectors:

$$\mathbf{l}_t + \mathbf{u}_t = \mathbf{h}_t$$

The equilibrium in the educational sector requires that:

$$\mathbf{x}_t + \mathbf{z}_t = \mathbf{e}_t$$

1.8 Definition of Stationary General Equilibrium with Insurable Income Shocks

The individual state variable for the young parents are the income shock ε_t , the realization of the ability shock of the child a_t , the bequest they inherit from their parents conditioned on the realization of income shock ε_t , b_{ε}^t , the stock of human capital they have h^t . The control variables are the consumption when young c_t^{t-2} , the amount of Arrow-Debreu securities that they buy s_{η}^t , and the amount of early investment they do on the human capital of the kid x_t^t . When an old parent (at period $t + 1$), the individual state variables are the income shock when η_{t+1} , the amount of financial claims they possess given the realization of η_{t+1} , s_{η}^t , their stock of human

capital h^t , the ability of the child a_t , the amount of early investment x_t^t . The control variables are consumption when old c_{t+1}^{t-2} , amount of Arrow-Debreus securities they leave as bequest b_ε^{t+1} and the late investment z_{t+1}^t . For both young and old parentes, the aggregate state variables are the distribution of human capital, bequest and savings G_t^h, G_t^b , and G_t^s , respectively. In what follows, we denote by $\zeta_t^y = (\varepsilon_t, a_t, b_\varepsilon^t, h^t, G_t^h, G_t^b, G_t^s)$ and $\zeta_{t+1}^o = (\eta_{t+1}, a_t, s_\eta^t, h^t, x_t^t, G_{t+1}^h, G_{t+1}^k, G_{t+1}^s)$ the set of state variables for the young and old parents, respectively.

We are ready to define the concept of equilibrium.

Definition 1 *We say that the decision rules $c^y(\zeta^y), s_\eta(\zeta^y), x(\zeta^y), c^o(\zeta^o), b_\varepsilon(\zeta^o), z(\zeta^o)$, corresponding value functions V and W , functions for aggregate factors of production $(\mathbf{k}(G^k, G^h), \mathbf{l}(G^k, G^h), \mathbf{u})$ and prices $(w(G^k, G^h), r(G^k, G^h), \{q_b(\varepsilon)\}, \{q_s(\eta)\})$ constitute a stationary general equilibrium with insurable income shocks if*

(1) *Given prices, transition rules for stocks of physical and human capital, transition rules for income shocks ε and η , then the decision rules solve the maximization problem of the agent.*

(2) *Given prices the allocation $(\mathbf{k}(G^k, G^h), \mathbf{l}(G^k, G^h), \mathbf{u}(G^k, G^h))$ solves:*

$$\max \{F(\mathbf{k}, \mathbf{l}) - (r + \delta) \mathbf{k} - w \mathbf{l}\}$$

$$\max \{J(\mathbf{u}) - w \mathbf{u}\}$$

(3) *The aggregate factors of production are generated by the aggregation of the decision rules of the individuals.*

(4) *The allocation is feasible*

(5) *There are no arbitrage opportunities:*

$$q_s(\eta) = \frac{g^\eta(\eta)}{(1+r)}$$

$$q_b(\varepsilon) = \frac{g^\varepsilon(\varepsilon)}{(1+r)}$$

(6) *The distribution of asset holdings G^k, G^s and human capital G^h are generated by the decision rules of the agents and are stationary (i.e. $G_{t+1}^b = G_t^b, G_{t+1}^s = G_t^s, G_{t+1}^h = G_t^h, \forall t$)*

1.9 Characterization of the Stationary General Equilibrium with Insurable Income Shocks

1.9.1 The First-Order Conditions

The first-order conditions for contingent financial claims $s_t(\eta_{t+1})$ of the young parent in period t are given by:

$$\frac{\partial u}{\partial c^y} q_s(\eta) = \beta \frac{\partial W(\eta, a, s, h, x)}{\partial s} g^\eta(\eta) d\eta \quad (2)$$

The first-order conditions for contingent bequests $b_{t+2}(\varepsilon_{t+2})$ of the old parent in period $t+1$ are given by:

$$\frac{\partial u}{\partial c^o} q_b(\varepsilon') = \beta \theta g^\varepsilon(\varepsilon') \int \frac{\partial V(\varepsilon', a', h', b')}{\partial b'} g^a(a') da' \quad (3)$$

We start by deriving the steady-state Euler equations. To carry out this task, we use the Envelope conditions on V and W . Note that:

$$\frac{\partial W(\eta, a, s, h, x)}{\partial s} = \frac{\partial u}{\partial c^o} \quad (4)$$

Now, plugging (4) into (2) we get that:

$$\frac{\partial u}{\partial c^y} q_s(\eta) = \beta \frac{\partial u}{\partial c^o} g^\eta(\eta)$$

Using the fact that the insurance market is competitive (i.e., $q_s(\eta) = \frac{g^\eta(\eta)}{1+r}$) it follows that:

$$\frac{\partial u}{\partial c^y} = \beta (1+r) \frac{\partial u}{\partial c^o} \quad (5)$$

Consider now the Envelope Condition for bequests (in financial claims contingent on shock ε'):

$$\frac{\partial V(\varepsilon, a, h, b)}{\partial b} = \frac{\partial u}{\partial c^y} \quad (6)$$

Plugging (6) into (3) it follows that:

$$\frac{\partial u}{\partial c^o} q_b(\varepsilon') = \beta \theta g^\varepsilon(\varepsilon') \int \frac{\partial u}{\partial c^y} g^a(a') da'$$

Again, competition in insurance market implies $q_b(\varepsilon') = \frac{g^\varepsilon(\varepsilon')}{1+r}$ and thus:

$$\frac{\partial u}{\partial c^o} = \beta \theta (1+r) \int \frac{\partial u}{\partial c^y} g^a(a') da' \quad (7)$$

1.9.2 The Deterministic Steady State:

Consider the version without uncertainty. In this case, equation (7) becomes:

$$\frac{\partial u}{\partial c^o} = \beta \theta (1+r) \frac{\partial u}{\partial c^y}$$

but using (5) it follows that the steady state interest rate is determined by:

$$\beta^2 \theta (1+r)^2 = 1$$

Assuming that the steady-state interest rate is positive and $\theta = 1$:

$$1+r = \frac{1}{\beta}$$

In this case, the steady state interest rate is the same as in the steady state of the neoclassical growth model.

Computation Given the steady state interest rate r , the capital-labor ratio is determined, because the production function satisfies constant-returns to scale. Given the capital-labor ratio, we can determine the steady state wage rate w . Given prices, we can then solve the problem of the agents by discretizing the state space and obtain the policy functions using the method of Bellman Iteration. We then simulate the series by sampling shocks on ability and income and applying the policy functions. We discard the first 100000 samples to allow the Markov Chain to converge. We use Monte-Carlo integration to compute the expectations that are required to check the market clearing conditions.

We approximate the Markov process for ability according to the procedure developed in Tauchen(1986). We use a 15-point grid to approximate the process:

$$a_{t+1} = 0.5a_t + \nu_{t+1}$$

$$\nu_{t+1} \sim N(0, 0.5)$$

we calculate the invariant distribution to sample agents for the simulated series. Figure 1 shows the histogram associated with the invariant distribution of ability.

The production function for human capital is:

$$h' = \begin{cases} \exp\{\gamma_a a\} h^{\gamma_h} x^\rho, & \text{if } z = 0 \\ \exp\{\gamma_a a\} h^{\gamma_h} [\gamma_x x^\phi + \gamma_z z^\phi]^{\frac{\rho}{\phi}}, & \text{if } z > 0 \end{cases}$$

and we use $\gamma_a = 0.5$, $\gamma_h = 0.1$, $\rho = 0.6$, $\gamma_x = 0.6$, $\gamma_z = 0.4$, $\phi = -2$.

We take $X = \{0.1, 0.2, 0.3\}$ and $Z = \{0, 1\}$. Again, the child is labeled skilled if $z = 1$, and unskilled otherwise. Figure 2 shows the histogram associated with the steady state distribution of human capital. We note that the distribution of human capital is skewed even though the distribution of ability is symmetric around zero.

The aggregate production function is:

$$f(k, l) = k^\alpha l^{1-\alpha}$$

with $\alpha = 0.36$. In figure 3 we plot the histogram associated with the stationary distribution of bequests. We note that many parents leave negative bequests to their children. These children tend to receive a lot of investment in human capital, and the parents borrow against the income of the children to finance such investments. It is also interesting to note that the distribution is also skewed and not centered around zero. A few parents leave relatively large bequests to their children in equilibrium. In figure 4, we plot the histogram of savings. Saving decisions are made by the young parents. First, the distribution of savings is lightly skewed, with positive, close to zero, mean.

Table 1 shows the intergenerational mobility matrix. The element m_{ij} of this matrix can be read as the probability that a child will be in the j -th decile of the present value of earnings distribution given that his parent is in the i -th decile. This table shows that there is some persistence that is partly caused by genetics of ability and partly caused by the effect of parental human capital in the production function of skills.

Table 2 shows the average treatment effect, the treatment on the treated and the treatment on the untreated estimators of the returns to skills. First, note from column 2 from table 2 that the treatment on the treated is positive and about 24% over the lifetime, while the treatment on the untreated is negative and also around -24%. The average treatment effect is the weighted mean of these two estimators. The average treatment effect is negative because the majority of the kids become unskilled (about 58%).

Table 3 shows the treatment on the treated and the treatment on the untreated estimators by ability group. The TT estimator is fairly increasing with ability, and always above 19%. The TU estimator is also increasing with ability, but it is always negative. The variance within each ability group is caused by the variation in parental human capital.

1.10 Uninsurable Income Shocks and Liquidity Constraints

We now change the budget constraint of the agent in the following manners. First, the agents cannot buy insurance against the idiosyncratic shock in income that they face. Second, the agents face liquidity constraints: they cannot leave negative bequests to the kids, and cannot carry on negative assets from young adulthood to old age. The problem of the parent becomes:

$$V(\varepsilon, a, b, h) = \max \left\{ u(c^y) + \beta \int W(\eta, a, s, h, x) dG^\eta(\eta) \right\}$$

subject to:

$$c^y + wx + \frac{s}{1+r} = wh + \varepsilon + b$$

$$s \geq 0$$

$$x_t^t \in X$$

$$\eta \sim iid G^\eta$$

and the problem of the period- $t + 1$ old parent is

$$W(\eta, a, s, h, x) = \max \left\{ u(c^o) + \beta \theta \int \int V(\varepsilon', a', b', h') dG^a(a') dG^\varepsilon(\varepsilon') \right\}$$

subject to:

$$c^o + wz + p + \frac{b'}{1+r} = wh + \eta + s, \text{ if } z > 0$$

$$c + \frac{b'}{1+r} = wh + \eta + s, \text{ if } z = 0$$

$$b' \geq 0$$

$$z_{t+1}^t \in Z$$

$$h' = H(a, x, z, h)$$

$$a \sim iid A$$

$$\varepsilon \sim iid G^\varepsilon$$

The first-order conditions for savings s are:

$$\frac{\partial u}{\partial c^y} \left(\frac{1}{1+r} \right) = \beta \int \frac{\partial W(\eta, a, s, h, x)}{\partial s} dG^\eta(\eta), \text{ if } s > 0$$

The first-order conditions for bequests b are:

$$\frac{\partial u}{\partial c^o} \left(\frac{1}{1+r} \right) = \beta \theta \int \int \frac{\partial V(\varepsilon', a', b', h')}{\partial b'} dG^a(a') dG^\varepsilon(\varepsilon'), \text{ if } b' > 0$$

The envelope condition for s and b are:

$$\frac{\partial W(\eta, a, s, h, x)}{\partial s} = \frac{\partial u}{\partial c^o}$$

$$\frac{\partial V(\varepsilon', a', b', h')}{\partial b'} = \frac{\partial u}{\partial c^y}$$

Using these facts, we conclude that:

$$\frac{\partial u}{\partial c^y} \left(\frac{1}{1+r} \right) = \beta \int \frac{\partial u}{\partial c^o} dG^\eta(\eta), \text{ if } s > 0$$

$$\frac{\partial u}{\partial c^o} \left(\frac{1}{1+r} \right) = \beta \theta \int \int \frac{\partial u}{\partial c^y} dG^a(a') dG^\varepsilon(\varepsilon'), \text{ if } b' > 0$$

Figure 1: The Histogram of Ability

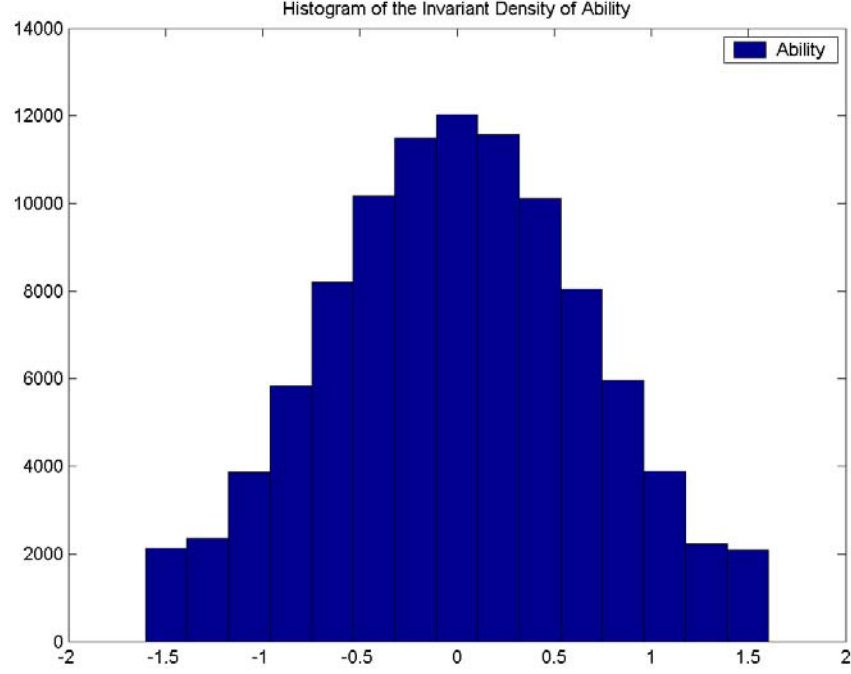


Figure 1 displays the histogram associated with the Invariant Distribution of Ability. We use Tauchen (1986) procedure to approximate such distribution by a 15-point discrete Markov Process. We assume that ability follows the process:

$$a_{t+1} = \mu a_t + \nu_{t+1}$$

with $\mu \in (0, 1)$ and $\nu_{t+1} \sim N(0, \sigma_\nu^2)$. Based on these facts, the invariant distribution of ability a_{t+1} is $N\left(0, \frac{\sigma_\nu^2}{1-\mu^2}\right)$.

Figure 2: The Histogram of Human Capital

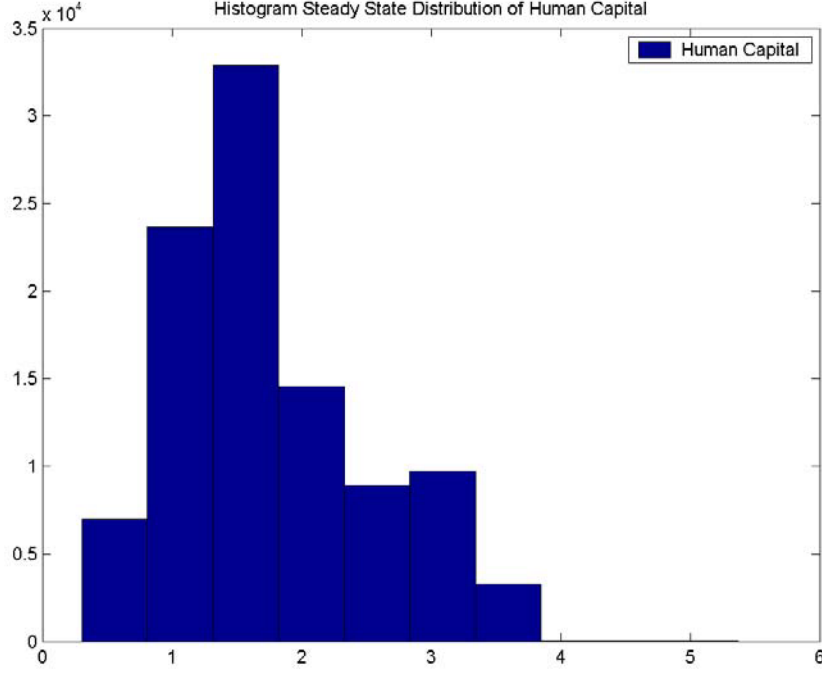


Figure 2 is the histogram of the Steady State Distribution of Human Capital in the model with full insurance against idiosyncratic innovations in income that are uncorrelated with human capital. We remind the reader that the law of motion of human capital is given by the production function

$$h' = H(a, h, x, z)$$

where h' denotes the human capital of the child when an adult, a denotes the child's innate ability, h is the parental human capital, x is the early investment in human capital and z is the investment in human capital that takes place in adolescent years. To obtain this histogram, we proceed in the following manner. First, we solve the model as described in the text by discretizing the state space and applying the Bellman Iteration principle. This process allows us to obtain the decision rules of savings, bequest and investment in human capital as functions of the child's ability, parental human capital and parental wealth. We then simulate a series of data by drawing ability from its stationary distribution. We discard the first 100,000 realizations of such series to allow the Markov chain to converge. Then, we save the next 100,000 realizations from which we compute the histogram above.

Figure 3: The Histogram of Present Value of Earnings

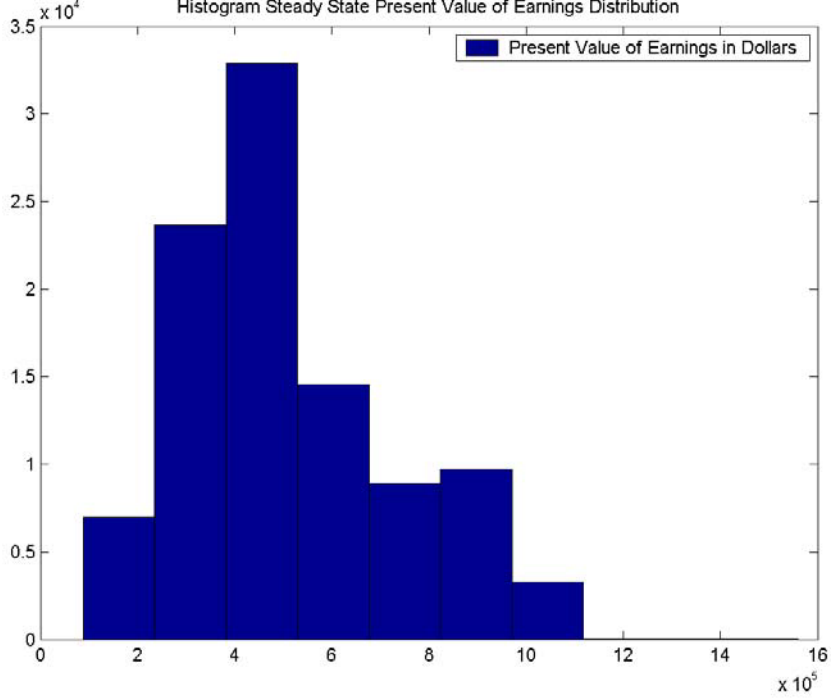


Figure 3 plots the the histogram of the Steady State Distribution of Present Value of Earnings (in hundred thousand dollars) in the model with full insurance against idiosyncratic innovations in income that are uncorrelated with human capital. We remind the reader of two points: first, the agent supplies labor inelastically during young and old adulthood (children never work), and second that there is no depreciation of human capital h over the lifecycle of the agent. Thus, the present value of earnings y is given by:

$$y = wh + \frac{w}{1+r}h$$

where w, r are the steady-state wage and interest rates. To obtain this histogram, we proceed in the following manner. First, we solve the model as described in the text by discretizing the state space and applying the Bellman Iteration principle. This process allows us to obtain the decision rules of savings, bequest and investment in human capital as functions of the child's ability, parental human capital and parental wealth. We then simulate a series of data by drawing ability from its stationary distribution. We discard

the first 100,000 realizations of such series to allow the Markov chain to converge. Then, we save the next 100,000 realizations from which we compute the histogram above.

Figure 4: Table 1 - The Returns to Skill

Table			
Average Returns			
	Overall	Conditioned on Being Skilled	Conditioned on Being Unskilled
Average Returns	0.1714	0.221	0.0972
Standard Deviation	0.0774	0.0579	0.0278

Table 1 presents the returns to becoming skilled as implied by the model with full insurance against idiosyncratic innovations in income that are uncorrelated with human capital. Column 1 presents the average returns for the population, column 2 presents the implied returns for only those who become skilled and the last column shows the returns for those who do not become skilled. Formally, let y_s , y_u denote the factual lifetime present value of earnings if skilled and unskilled, respectively. Let τ denote the fixed cost of becoming skilled, let x denote the early investment in human capital. According to the model if the agent becomes skilled then $z = 1$ while if it is unskilled then $z = 0$. For each person with ability a , parental human capital h , and early investment x we observe whether $z = 1$ or $z = 0$. Let h^s, h^u denote the human capital of the skilled and unskilled, respectively. Let \tilde{h}^u, \tilde{h}^s denote the counterfactual human capital for the skilled as unskilled, and unskilled as skilled, respectively. Then,

$$y_s = wh^s + \frac{w}{1+r}h^s$$

$$y_u = wh^u + \frac{w}{1+r}h^u$$

$$\tilde{y}_u = w\tilde{h}^u + \frac{w}{1+r}\tilde{h}^u$$

$$\tilde{y}_s = w\tilde{h}^s + \frac{w}{1+r}\tilde{h}^s$$

that is \tilde{y}_u, \tilde{y}_s denote the counterfactual lifetime present value of earnings if skilled and unskilled, respectively. The average return to becoming skilled for those who are skilled is given by:

$$E(\text{returns} \mid \text{skilled}) = E\left[\frac{y_s - \tilde{y}_u - \tau}{\tilde{y}_u + \tau} \mid \text{skilled}\right]$$

which is shown in column 2 shows. Column 3 shows:

$$E(\text{returns} \mid \text{unskilled}) = E\left[\frac{\tilde{y}_s - y_u - \tau}{y_u + \tau} \mid \text{unskilled}\right]$$

Let p denote the proportion of agents that become skilled. Column 1 shows:

$$E(\text{returns}) = pE(\text{returns} \mid \text{skilled}) + (1 - p)E(\text{returns} \mid \text{unskilled})$$

Figure 5: Table 2 - The Returns to Skill per Ability Group

Table Average Returns by Ability Group				
Ability Group	Average Returns Conditioned on Being Skilled	Variance of Returns Conditioned on Being Skilled	Average Returns Conditioned on Being Unskilled	Variance of Returns Conditioned on Being Unskilled
Group 1	0.197752	0.001277	0.055550	0.000014
Group 2	0.159974	0.002802	0.065188	0.000222
Group 3	0.218582	0.001319	0.084675	0.000302
Group 4	0.178326	0.003059	0.068183	0.000125
Group 5	0.166669	0.003181	0.087984	0.000479
Group 6	0.248078	0.000088	0.078313	0.000256
Group 7	0.250214	0.000212	0.106608	0.000088
Group 8	0.186507	0.003670	0.107686	0.000306
Group 9	0.228918	0.003176	0.083991	0.000304
Group 10	0.260372	0.000444	0.078347	0.000069
Group 11	0.203515	0.004248	0.116352	0.000667
Group 12	0.274707	0.000096	0.114489	0.000722
Group 13	0.276833	0.000049	0.128190	0.000320
Group 14	0.271600	0.000638	0.137184	0.000067
Group 15	0.277610	0.000019	0.139753	0.000054

Table 2 presents the returns to becoming skilled **per ability group** as implied by the model with full insurance against idiosyncratic innovations in income that are uncorrelated with human capital. Column 1 presents the average returns for the group of agents that are skilled in equilibrium, column 2 presents the variance of such returns. Column 3 presents the implied returns for only those who are unskilled in equilibrium and the last column shows the variance of such returns. Formally, let $y_s(a)$, $y_u(a)$ denote the factual lifetime present value of earnings of a person with ability a if skilled and unskilled, respectively. Let τ denote the fixed cost of becoming skilled, let x denote the early investment in human capital. According to the model if the agent becomes skilled then $z = 1$ while if it is unskilled then $z = 0$. For each person with ability a , parental human capital h , and early investment x we observe whether $z = 1$ or $z = 0$. Let $h^s(a)$, $h^u(a)$ denote the human capital of the skilled and unskilled of a person with ability a , respectively. Let $\tilde{h}^u(a)$, $\tilde{h}^s(a)$ denote the counterfactual human capital for the skilled as unskilled, and unskilled as skilled of a person with ability a , respectively. Then,

$$y_s(a) = wh^s(a) + \frac{w}{1+\tau}h^s(a)$$

$$y_u(a) = wh^u(a) + \frac{w}{1+r}h^u(a)$$

$$\tilde{y}_u(a) = w\tilde{h}^u(a) + \frac{w}{1+r}\tilde{h}^u(a)$$

$$\tilde{y}_s(a) = w\tilde{h}^s(a) + \frac{w}{1+r}\tilde{h}^s(a)$$

that is $\tilde{y}_u(a), \tilde{y}_s(a)$ denote the counterfactual lifetime present value of earnings if skilled and unskilled of a person with ability a , respectively. The average return to becoming skilled for those who are skilled is given by:

$$E(\text{returns} \mid a, \text{skilled}) = E\left[\frac{y_s(a) - \tilde{y}_u(a) - \tau}{\tilde{y}_u(a) + \tau} \mid a, \text{skilled}\right]$$

which is shown in column 1 shows. Column 3 shows:

$$E(\text{returns} \mid a, \text{unskilled}) = E\left[\frac{\tilde{y}_s(a) - y_u(a) - \tau}{y_u(a) + \tau} \mid \text{unskilled}\right]$$

Figure 6 : The Average Returns to Skill Conditioned on Being Skilled

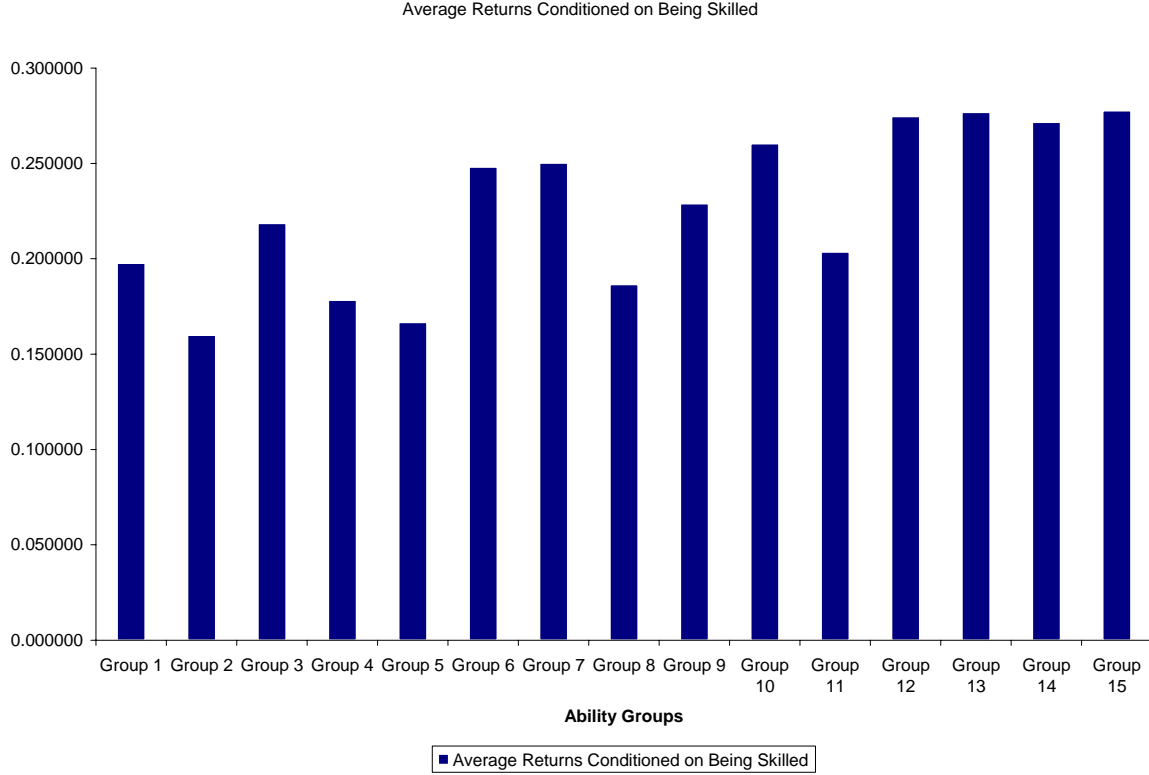


Figure 6 presents the returns to becoming skilled conditioned on being skilled, **per ability group**, as implied by the model with full insurance against idiosyncratic innovations in income that are uncorrelated with human capital. Column 1 presents the average returns for the group of agents that are skilled in equilibrium, column 2 presents the variance of such returns. Column 3 presents the implied returns for only those who are unskilled in equilibrium and the last column shows the variance of such returns. Formally, let $y_s(a)$, $y_u(a)$ denote the factual lifetime present value of earnings of a person with ability a if skilled and unskilled, respectively. Let τ denote the fixed cost of becoming skilled, let x denote the early investment in human capital. According to the model if the agent becomes skilled then $z = 1$ while if it is unskilled then $z = 0$. For each person with ability a , parental human capital h , and early investment x we observe whether $z = 1$ or $z = 0$. Let $h^s(a)$, $h^u(a)$ denote the human capital of the skilled and unskilled of a person with ability a , respectively. Let $\tilde{h}^u(a)$, $\tilde{h}^s(a)$ denote the counterfactual human capital for the skilled as unskilled, and unskilled as skilled of a person with ability a , respectively. Then,

$$y_s(a) = wh^s(a) + \frac{w}{1+r}h^s(a)$$

$$y_u(a) = wh^u(a) + \frac{w}{1+r}h^u(a)$$

$$\tilde{y}_u(a) = w\tilde{h}^u(a) + \frac{w}{1+r}\tilde{h}^u(a)$$

$$\tilde{y}_s(a) = w\tilde{h}^s(a) + \frac{w}{1+r}\tilde{h}^s(a)$$

that is $\tilde{y}_u(a)$, $\tilde{y}_s(a)$ denote the counterfactual lifetime present value of earnings if skilled and unskilled of a person with ability a , respectively. The average return to becoming skilled for those who are skilled is given by:

$$E(\text{returns} \mid a, \text{skilled}) = E\left[\frac{y_s(a) - \tilde{y}_u(a) - \tau}{\tilde{y}_u(a) + \tau} \mid a, \text{skilled}\right]$$

which is shown in column 1 shows. Column 3 shows:

$$E(\text{returns} \mid a, \text{unskilled}) = E\left[\frac{\tilde{y}_s(a) - y_u(a) - \tau}{y_u(a) + \tau} \mid \text{unskilled}\right]$$

Figure 7: The Returns to Skill Conditioned on Being Unskilled

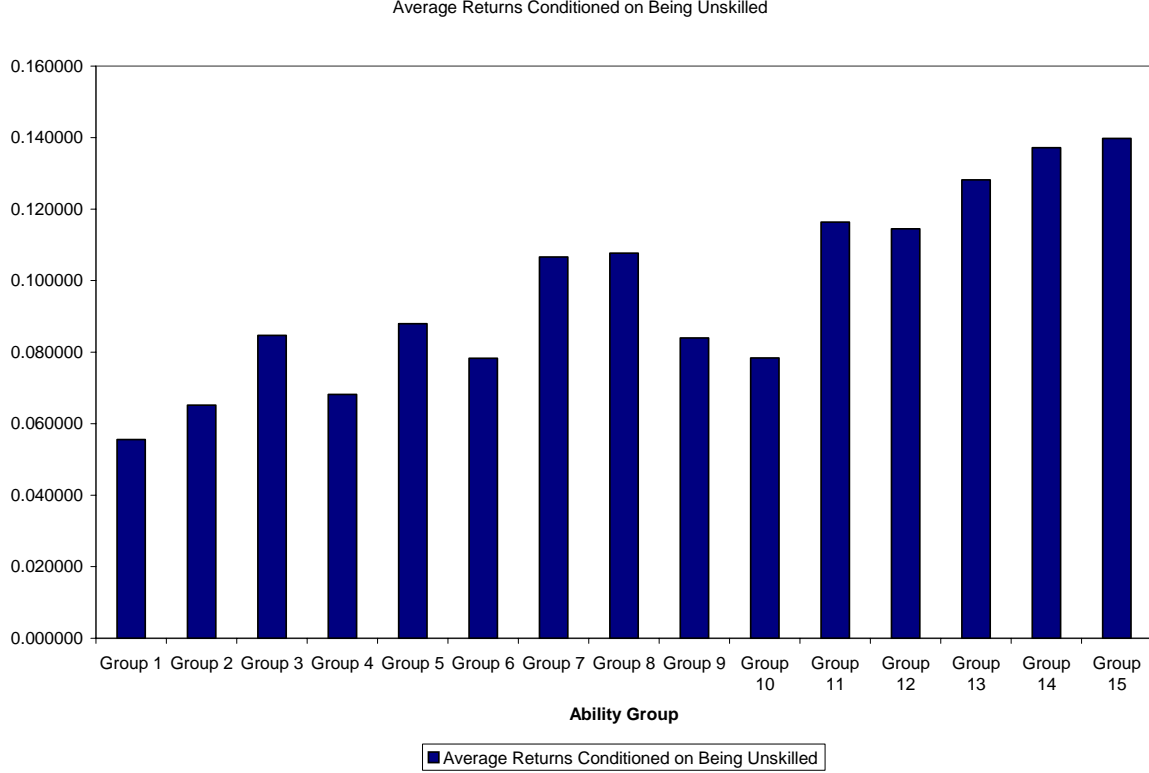


Figure 7 presents the returns to becoming skilled conditioned on being skilled, **per ability group**, as implied by the model with full insurance against idiosyncratic innovations in income that are uncorrelated with human capital. Column 1 presents the average returns for the group of agents that are skilled in equilibrium, column 2 presents the variance of such returns. Column 3 presents the implied returns for only those who are unskilled in equilibrium and the last column shows the variance of such returns. Formally, let $y_s(a)$, $y_u(a)$ denote the factual lifetime present value of earnings of a person with ability a if skilled and unskilled, respectively. Let τ denote the fixed cost of becoming skilled, let x denote the early investment in human capital. According to the model if the agent becomes skilled then $z = 1$ while if it is unskilled then $z = 0$. For each person with ability a , parental human capital h , and early investment x we observe whether $z = 1$ or $z = 0$. Let $h^s(a)$, $h^u(a)$ denote the human capital of the skilled and unskilled of a person with ability a , respectively. Let $\tilde{h}^u(a)$, $\tilde{h}^s(a)$ denote the counterfactual human capital for the skilled as unskilled, and unskilled as skilled of a person with ability a , respectively. Then,

$$y_s(a) = wh^s(a) + \frac{w}{1+r}h^s(a)$$

$$y_u(a) = wh^u(a) + \frac{w}{1+r}h^u(a)$$

$$\tilde{y}_u(a) = w\tilde{h}^u(a) + \frac{w}{1+r}\tilde{h}^u(a)$$

$$\tilde{y}_s(a) = w\tilde{h}^s(a) + \frac{w}{1+r}\tilde{h}^s(a)$$

that is $\tilde{y}_u(a)$, $\tilde{y}_s(a)$ denote the counterfactual lifetime present value of earnings if skilled and unskilled of a person with ability a , respectively. The average return to becoming skilled for those who are skilled is given by:

$$E(\text{returns} \mid a, \text{skilled}) = E \left[\frac{y_s(a) - \tilde{y}_u(a) - \tau}{\tilde{y}_u(a) + \tau} \mid a, \text{skilled} \right]$$

which is shown in column 1 shows. Column 3 shows:

$$E(\text{returns} \mid a, \text{unskilled}) = E \left[\frac{\tilde{y}_s(a) - y_u(a) - \tau}{y_u(a) + \tau} \mid \text{unskilled} \right]$$

Figure 8: Table 3 - The Intergenerational Mobility Table

Table										
Intergenerational Mobility Table in the Model with Full Insurance Against Idiosyncratic Shocks in Income										
Decile	1	2	3	4	5	6	7	8	9	10
1	0.1413	0.1375	0.1270	0.0969	0.1858	0.0893	0.0276	0.0422	0.1194	0.0331
2	0.2649	0.0989	0.0769	0.0084	0.0411	0.1824	0.1405	0.0003	0.1612	0.0254
3	0.2019	0.0841	0.1400	0.1060	0.0048	0.0851	0.2335	0.0103	0.1067	0.0277
4	0.0377	0.1117	0.0810	0.0674	0.1197	0.0278	0.1245	0.1743	0.1720	0.0839
5	0.0162	0.1084	0.1299	0.2124	0.1514	0.0123	0.0299	0.1493	0.1393	0.0509
6	0.0058	0.1437	0.0848	0.1617	0.1427	0.0104	0.0226	0.2034	0.1456	0.0792
7	0.0030	0.1511	0.0858	0.1710	0.1627	0.0078	0.0138	0.2117	0.1363	0.0569
8	0.0527	0.0717	0.1017	0.1344	0.1121	0.1322	0.0312	0.0719	0.1124	0.1797
9	0.2196	0.0080	0.0916	0.0479	0.0878	0.1630	0.1721	0.0426	0.0235	0.1439
10	0.0615	0.0786	0.0687	0.0565	0.0922	0.1669	0.1633	0.1146	0.0448	0.1530

Table 3 displays the steady-state intergenerational mobility table implied by the model with full insurance against idiosyncratic innovations in income that are uncorrelated with human capital. The element a_{ij} of the table is the probability that a child will be in the j -th decile of the lifetime present value of earnings given that his parent is in the i -th decile. For example, $a_{11} = 0.1413$ means that with probability 14.13% a child with a parent in the first decile of the lifetime present value of earnings will also be in the first-decile of lifetime present value of earnings, whereas with probability 3.31% it will be in the highest decile. To obtain this table we proceed as follows. First, we solve the model as described in the text by discretizing the state space and applying the Bellman Iteration principle. This process allows us to obtain the decision rules of savings, bequest and investment in human capital as functions of the child's ability, parental human capital and parental wealth. We then simulate a series of data by drawing ability from its Markov process (and not from the stationary distribution, as parental ability matters to determine child's ability). We discard the first 100,000 realizations of such series to allow the Markov chain to converge. Then, we save the next 100,000 realizations from which we compute the table above.