Investments in Children When Markets Are Incomplete *

Flavio Cunha[†]

September 03, 2013

Abstract

This paper proposes a model of investments in the cognitive skills of children that is based on three features: that it takes time to build a child's stock of cognitive skills; that parental resources evolve stochastically over time; and that parents face constraints that limit their ability to transfer resources across states of nature, across time, and across generations. These constraints distort the equilibrium allocation of investments throughout a child's life and, in turn, produce sub-optimal stocks of cognitive skills. In order to verify the quantitative importance of these distortions, I estimate the model's key parameters and I compute its steady-state equilibrium. I show that the empirically- grounded steady-state explains a variety of facts about cognitive skills, education, and child development. For example, it correctly predicts selection into college by quartiles of family income and terciles of skills measured at adolescent years. Moreover, it is consistent with gaps in cognitive skills that are present at early ages. Finally, it reproduces the pattern of selection into college based on cognitive skills. I use the model to evaluate the impact of different remediation policies on the stationary distribution of cognitive skills and welfare. I analyze the effects of a 50% tuition subsidy, a targeted early investment subsidy, and a targeted early and late investment subsidy that is contingent on parental resources. I show that the policy that subsidizes early and late childhood investments dominates the other policies in welfare, since it is the one that generates the highest equivalent variation across all deciles of permanent income. This also generates a stationary distribution of cognitive skills that first-order stochastically dominates the ones generated by the baseline economy and the other remediation policies.

1 Introduction

The increase in the pecuniary returns of a college degree is partly determined by the differential growth rates of the supply of and the demand for skilled workers (Goldin and Katz, 2008). The reason why the growth rate of supply has decelerated in the recent past is attributed to the fact that credit constraints in college-going years have become tighter (e.g., Belley and Lochner, 2007), that tuition rates have become progressively more expensive (Kane, 1995), or that there is a consistent upward trend in the fraction of the young population from disadvantaged households that have low stocks of skills at the time when they make college their enrollment decisions (Reardon, 2011).

†Department of Economics, University of Pennsylvania.

As Carneiro and Heckman (2003) show, the stocks of skills young individuals have played a very important role in understanding college-enrollment gaps across socio-economic groups. In their analysis of the NLSY/79, they divide the seventeen-year-old white males into twelve different categories. Each category corresponds to quartiles of family income and terciles of cognitive skills as proxied by the Armed Forces Qualification Test (AFQT). Family resources and a child's skills are measured during his or her adolescent years. Within each tercile of the AFQT, the likelihood of enrolling into college increases with family resources. This evidence is consistent with credit constraints affecting an adolescent's decision of going to college. However, family income in adolescent years is correlated with many family characteristics that might lead to higher income and higher educational attainment. For example, parents with higher skills could be more efficient both in the labor market (and thus have higher earnings) as well as in the production of skills in their children. When Carneiro and Heckman (2003) control for parental characteristics, the role of family income in determining college enrollment is substantially weakened. In contrast, the role of the AFQT becomes even more important in explaining college enrollment gaps.

The interpretation of the evidence presented by Carneiro and Heckman (2003) requires caution. The evidence neither proves nor disproves the absence of market failures that cause distortions in the accumulation of skills. The AFQT, which explains a lot about the variation in the college enrollment decision, is itself the outcome of a process that began when the child was still in the womb. If parents cannot be qualified for the loans to finance their child's accumulation of skills, poor parents start investing less than rich parents many years before the college decision is made. As a result, the gaps in cognitive skills across socio-economic groups start to open up at very early ages as documented by Carneiro and Heckman (2003). At the time when the college-going decision is made, the return to college education of a child with a history of low investments is also low, and this demographic group should not benefit much from policies that subsidize this type of investment. The evidence by Reardon (2011) that the gaps in skills have increased in the last 60 years suggest that low skills will become an increasingly important reason for why the U.S. economy will fail to produce a skilled labor force.

With these facts in mind, the goal of the paper is to evaluate the impact and welfare implications of different policies that aim to increase the growth rate of the supply of skilled workers. To be able to address these issues, I develop a model of investment in the cognitive skills of children. My model has three key features. The first feature of the model is that it takes many periods to build an adult person's stock of cognitive skills. This is done to capture the notion of dynamic complementarity proposed in Cunha and Heckman (2007): investments made in the past raise the productivity of current and future investments. In the current paper, complementarity has a dual face: it is very difficult to compensate early neglect with late investments. At the same time, later investments are needed to make early investments pay off. Lifetime credit constraints that affect early investment can explain lower levels of skills among disadvantaged children, which in turn explains much of the gap in college-going across different socioeconomic groups.

The second feature of the model is that parental resources follow a stochastic process that is determined by the parent's own stocks of cognitive skills and exogenous labor productivity shocks. The formulation I adopt in my model is thus consistent with the literature that investigates the properties of labor income processes (e.g., MacCurdy, 1982; Meghir and Pistaferri, 2004).

The third feature of the model is that parents face constraints that limit their ability to transfer resources across states of nature, across time, and across generations. I build on the work of Laitner

(1992), who proposes an economy of overlapping generations that has two distinct market failures: parents cannot buy insurance against their own earnings risk nor can they borrow against the future earnings of their children. At any point in time, parental resources are given by the stochastic labor income and the stocks of a risk-free asset. Unlike the Laitner (1992) model, my model deems labor income a function of the stock of the parent's cognitive skills, which is determined endogenously. Parents can transfer resources to their children by investing in skills or by purchasing risk-free bonds. Again, the types of market failures described in the model are consistent with the findings from the literature that investigates the amount of insurance available to households (e.g., Blundel, Pistaferri, and Preston, 2008; Attanasio and Weber, 2010; Heathcote, Storesletten, and Violante, 2010).

To make the model empirically operational, I estimate the technology for the formation of cognitive skill by following the analysis of Cunha, Heckman, and Schennach (2010). In order to do so, I use the rich measurements of cognitive skills and parental inputs available in the CNLSY/79. After calibrating the model's other key parameters, I compute the model's steady-state equilibrium. I show that it qualitatively reproduces many features documented in the literature on cognitive skill formation and education attainment. For example, the model implies that both ability and family income measured during the adolescent years of the child predict college enrollment. This explains the gaps in cognitive ability that open up at early ages. Agents who enroll in college tend to have higher skills and higher returns to college. Consistent with this evidence, the model predicts that the agents at the top of the high-school distribution of abilities measured at age eighteen tend to respond more strongly to tuition reductions. At the same time, the model predicts that it is not the poorest, but rather especially the children of middle-class parents who will benefit the most from such a policy.[1]

I use my model to analyze the long-run impact of other balanced-budget investment policies which subsidize investments in some or all periods of the accumulation of abilities. Specifically, I consider a subsidy to investments at ages three and four targeted at poor parents of low-ability children, and a policy that subsidizes the investments made by poor parents at multiple periods. For each policy that I analyze, I compute the distribution of costs and benefits across the population.

Among the counterfactuals that I simulate, the policy that subsidizes skill formation throughout the childhood period generates the highest benefits. There are two reasons why this is the case. First, given the fact that in the model labor income is stochastic and markets are incomplete, a subsidy to the price of investment during all periods of childhood serves as a form of insurance against bad labor income shocks. Second, under the market structure proposed by Laitner (1992), low-income parents will underinvest in their children's accumulation of cognitive ability during their entire childhood, adolescence, and early adulthood of their offsprings. Any policy designed to alleviate their underinvestment has to take into account the dynamic complementarity produced by the technology of skill formation. If remediation for disadvantage occurs late, after a series of low initial investments, the marginal productivity of the remediation investment is also low. Late remediation policy may not generate enough incentives to make parents reallocate resources away from consumption and towards the accumulation of their children's skills. Similarly, an intervention that subsidizes or supplements parental investments only during the early stages of child development fails to recognize the complementarity of investments, which implies that for high early investments to pay off as much as possible, they must be followed up by later investments.

---

[1] This is a version of Director's Law as defined by Stigler (1970).

The research presented in this paper relates to other work on the estimation of the production function of human capital (Cunha and Heckman, 2008; Cunha, Heckman, and Schennach, 2010; Todd and Wolpin, 2007). Unlike the previously published papers, I model explicitly how investments are chosen by parents. This is necessary for investigating how policies affect investment choices and the welfare of families in an equilibrium framework.

The paper is also related to the recent work of Del Boca, Flinn, and Wiswall (2013), who estimate a model of parental allocation of time. In their model, families may have one or two children, but they do not allow parents to consume or save. In the model that I propose below, I assume that parents have only one child, but parents also choose how much to borrow against their own future income or save out of their current income in the context of an equilibrium overlapping generations economy.

This paper also relates Carneiro and Ginja (2012), who investigate the extent to which parents are insured against stochastic fluctuations in income. The evidence they present shows that the elasticity of parental investments with respect to permanent shock is 13%. They also show that temporary shocks in income do not affect parental investments. This suggests that parents have access to partial insurance against income fluctuations, which is consistent with the model presented in this paper. Dahl and Lochner (2012) show that changes in income associated with changes in the Earned Income Tax Credit schedule leads to larger changes in investments. Again, if the changes in schedule in EITC correspond to changes in permanent income of the parents, then this is consistent with a model in which poor parents are limited in their capacity to borrow against the future income of their children. Policies that affect their permanent income will translate into changes in investments.

This paper is closely related to Caucutt and Lochner (2012). We both model investments in human capital as seen over many periods of childhood and allow for lifetime liquidity constraints. In their model, parents save not only to smooth consumption during adulthood, but also to finance consumption in old age. Although consumption during retirement is not present in my model, I do consider an equilibrium model and the set of policies that I consider are fully financed in equilibrium, neither of which is present in Caucutt and Lochner (2012). This distinction is important since it makes comparing policies possible not only in terms of their benefits, but also in terms of their costs to society, a difference necessary for estimating welfare implications. Another difference is that the current paper estimates the production function from data related to investments and cognitive skills, while Caucutt and Lochner (2012) calibrate the technology parameters based on the different profiles of labor market earnings.

This paper is organized in the following way. Section 2 presents the model. Section 3 estimates the technology of cognitive skill formation. Section 4 computes the steady state of the model and discuss its properties. Section 5 uses the model to evaluate policies for fostering skill formation. The last section is a conclusion.

2   Model

Laitner (1992) develops a model in which altruistic households face two market imperfections. The first is that parents can only transfer non-negative amounts of financial wealth in the form of a risk-free asset to their children. The second is that parents cannot protect themselves against shocks to their own labor productivity. The heterogeneity in productivity featured in his model is exogenously determined. I build on Laitner (1992) by allowing parents to invest in the cognitive skill of their children. Once

these children become adults themselves, their accrued stock of cognitive skills will partly determine their income as adults. In my model, one generation can transfer resources to the next through risk-free assets or investments in cognitive skills.

The analysis of this section proceeds in the following way. First, I briefly describe the generational structure and how agents evolve through their lifetimes. Second, I introduce and discuss the technology of cognitive skill formation. Third, I write the problem the parents face at each year. Fourth, I describe the two types of firms in this economy: some firms produce consumption goods, while others produce goods used by the parents to invest in the children's cognitive skills. I use the term "education goods" to refer to these goods. Finally, I present the aggregate market structure and discuss how prices are determined in equilibrium.

2.1 Generational Structure

The environment is an overlapping generations economy with an infinite number of periods. Each generation consists of a continuum of agents with mass equal to unity. There is no population growth. Each agent lives for $2T$ periods. During the first T years of life, the agent is a child and by assumption makes no economic decisions. Upon reaching age $T + 1$, the agent becomes an adult and gives birth to a child. The agent dies at the end of the calendar year in which she completes $2T$ years of age and is replaced in the beginning of the next calendar year by the generation of her grandchild. Life goes on in the future in similar fashion.

2.2 The Technology of Skill Formation

I now present the technology of cognitive skill that I use in this paper. In this paper, I use θ_t to denote the stock of cognitive skills of a child at age t. The parental investments in cognitive skill when said child is t years old is represented by x_t. The parental stock of cognitive skill is h.[2] I assume that there are N distinct stages of development. At age t, this child is at developmental stage n_t, $n_t = 1, 2, ..., N$. The developmental stage n_t starts at age $T_{n_t}^A$ and finishes at age $T_{n_t}^B$, $T_{n_t}^A \leq T_{n_t}^B$. The technology for cognitive skill formation when the child is at developmental stage n_t, $n_t = 1, 2, ..., N$, and age t, $T_{n_t}^A \leq t \leq T_{n_t}^B$ is:

$$\ln \theta_{t+1} = \ln \delta_{n_t} + \frac{\rho_{n_t}}{\phi_{n_t}} \ln \left\{ \gamma_{1,n_t} e^{\phi_{n_t} \ln \theta_t} + \gamma_{2,n_t} e^{\phi_{n_t} \ln x_t} + \gamma_{3,n_t} e^{\phi_{n_t} \ln h} \right\}, \tag{1}$$

with $0 < \gamma_{1,n_t}, \gamma_{2,n_t}, \gamma_{3,n_t}, \rho_N < 1, \quad \phi_{n_t} \leq 1, \quad \sum_k \gamma_{k,n_t} = 1$. This technology of skill formation is proposed in Cunha, Heckman, Lochner, and Masterov (2006) and Cunha and Heckman (2007).

Note that the parameters of the technology of skill formation depends on the child's current developmental stage n_t. This allows for investments to have different impacts at different developmental stages. The technology allows for the degree of complementarity between investments x_t, current stocks of skill, θ_t, and parental cognitive skill, h, to vary with the developmental stages of the child.

---

[2]In this paper I assume that the level of cognitive skill does not change once a person reaches adulthood.

2.3 The Problem of the Parents

The parent is assumed to be the decision-maker in the household.[3] The problem she solves depends on the age of the child. When the child is between ages 1 and $T - 1$, he only receives investments in cognitive skills and cannot work.

When the child reaches age T, the parent may invest a minimum level or something beyond that minimum. This minimum level of investment is necessary for the child to attend college. If the parent invests the minimum amount, the child does not attend college but becomes a high-school graduate and works full time. If the parent invests any amount beyond the minimum, the child attends school (college) full-time. At the end of the period, he becomes a college graduate.

2.3.1 The Problem When the Child is Between 1 and T-1 years old

Children between 1 and $T - 1$ years old do not work. They only receive investments in their cognitive skill as provided by his parent. The parental labor supply is perfectly inelastic. At each age t of the child, the parent is subject to productivity innovations ε_t. The shocks ε_t are independently and identically distributed across parents. The shocks follow a first-order Markov process:

$$\ln \varepsilon_{t+1} = \rho_\varepsilon \ln \varepsilon_t + \sigma_\eta \eta_t^\varepsilon \tag{2}$$

Parents are assumed to have positive earnings. I restrict productivity innovations such that there exists $\varepsilon_{\min}$ with the property that $\varepsilon_t \geq \varepsilon_{\min} > 0$ for any $t = T + 1, ..., 2T$.

The labor income of the parents is given by $wh\varepsilon_t$, where w is the wage rate of one efficiency unit. The level of cognitive skill of the parent, h, is the product of investment decisions made by the grandparent. In similar fashion, the level of cognitive skill of the child when an adult, h', will also be the consequence of investments made by the parent, and satisfies $h' = \theta_{T+1}$. The individual state variables for the parents of children who are between "1" and "$T - 1$" years old are $(h, \theta_t, s_t, \varepsilon_t, n_t)$, where n_t is the developmental stage of the child at age t.

Given the state variables, the parent chooses household consumption c_t, savings s_{t+1}, and investments x_t in the cognitive skill of the child. The savings of the parents are in a risk-free asset which pays a rate of interest r. I use p to denote the price of the investment goods in cognitive skill. Following Laitner (1992), the parents cannot leave debts to their children and have negative net worth, so savings are subject to the lower bound equal to $\frac{-wh\varepsilon_{\min}}{(1+r)}$. Let $V(t, h, \theta_t, s_t, \varepsilon_t, n_t)$ denote the value function of the parent of a child at age t, $1 \leq t \leq T - 1$. The problem of the parent is:

$$V(t, h, \theta_t, s_t, \varepsilon_t, n_t) = \max_{c_t, x_t, s_{t+1}} \{u(c_t) + \beta E[V(t+1, h, \theta_{t+1}, s_{t+1}, \varepsilon_{t+1}, n_{t+1}) | \varepsilon_t]\}$$

subject to:

$$c_t + px_t + s_{t+1} = wh\varepsilon_t + (1+r)s_t \tag{3}$$

$$s \geq -(wh\varepsilon_{\min}), \ x_t, c_t \geq 0 \tag{4}$$

and the technology for cognitive skill formation (1).

---

[3]To avoid confusion and repetition, I refer to the parent as she and to the child as he.

2.3.2 The Problem When the Child Reaches T years old

When the child reaches age "T", the parent decides to invest the minimum amount, $\underline{x}$, or something beyond that amount. The parent uses the relevant information to make that decision, which is contained in the vector of state variables $(h, \theta_t, s_t, \varepsilon_t, n_t)$. Her problem can be stated as:

$$V(T, h, \theta_T, s_T, \varepsilon_T, n_T) = \max_{c_T, x_T, s_1'} \left\{ u(c_T) + \beta E \left[V\left(1, h', \theta_1', s_1', \varepsilon_1', n_1'\right)\right]\right\}$$

subject to:

$$c_T + s_1' + p\underline{x} = wh\varepsilon_T + w\theta_T + (1+r)s_T \text{ if } x_T = \underline{x} \tag{5}$$

$$c_T + s_1' + (px_T + \varphi) = wh\varepsilon_T + (1+r)s_T \text{ if } x_T > \underline{x} \tag{6}$$

$$s_T \geq 0 \tag{7}$$

and the technology for the production of skills (1).

The budget constraint (5) states that a child who receives the minimum amount of investments $\underline{x}$ works full time. I will refer to this child as a high-school graduate. Note that the high-school graduate child's earnings are pooled with the rest of the parental resources. For the sake of simplicity, I abstract from productivity shocks for the child before he reaches adulthood.

If the parent decides to invest any amount above the minimum, so that $x_T > \underline{x}$, then the parent must pay the variable cost of the investment, which is p by unit, plus a fixed cost, φ, which I call college tuition. A child who receives more than the minimum amount of investment does not work. This is described by the budget constraint (6).

Note that equation (7) embodies the notion that the parent faces lifetime liquidity constraints. The parent dies and cannot leave debts to the child.

2.4 The Firms

I model two distinct sectors. The first sector uses physical capital and labor, measured in efficiency units, to produce the consumption good. The second sector uses only labor, also measured in efficiency units, to produce the investment good for cognitive skills.

2.4.1 The Consumption Good Sector

I assume that the production function in the consumption good sector exhibits constant returns to scale. This assumption is for convenience. Under constant returns to scale I can work with one firm without loss of generality. In what follows, I use capital letters to denote aggregate levels. Because the focus of the paper is on stationary equilibrium, I do not use time subscripts. As anticipated above, there are two inputs in the production function of goods: physical capital and labor, which is 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:

$$\mathbf{Y} = F(\mathbf{K}, \mathbf{L})$$

The production function of aggregate output presents constant returns to scale, is strictly increasing in all its arguments, satisfies the Inada Conditions, and is twice-continuously differentiable.

The problem of the firm in the good production sector is:

$$\pi_Y = \max \left\{ F\left(\mathbf{K}, \mathbf{L}\right) - w\mathbf{L} - \left(r + \delta\right)\mathbf{K} \right\}$$

The first-order conditions are:

$$w = \frac{\partial F\left(\mathbf{K}, \mathbf{L}\right)}{\partial \mathbf{L}}$$

$$\left(r + \delta\right) = \frac{\partial F\left(\mathbf{K}, \mathbf{L}\right)}{\partial \mathbf{K}}$$

2.4.2 The Education Good Sector

There is an educational sector that produces goods for investment in cognitive skill. Let $\mathbf{E}$ denote the total supply of educational goods. This sector does not use physical capital as input, only labor $\mathbf{U}$. The production technology is

$$\mathbf{E} = \mathbf{U}.$$

The problem of the firm is

$$\pi_E = \max \left\{ p\mathbf{E} - w\mathbf{U} \right\}.$$

As is well known, this problem has a solution with limited, positive production if, and only if

$$p = w.$$

2.5 Market-Clearing Conditions

Let $\zeta_t = \left(h, \theta_t, s_t, n_t, \varepsilon_t\right)$ denote the vector of state variables facing the parents. Let $\zeta = \left(\zeta_1, ..., \zeta_T\right)$. Let $g\left(\zeta\right)$ denote the joint probability density function of the state variables. Let $c_t\left(\zeta_t\right), s_t\left(\zeta_t\right)$ denote the consumption and savings functions when the child is t years old. Let $\mathbf{C}_t$, $\mathbf{S}_t$ denote the aggregate consumption and savings of the household that has a child who is t years old[4], where $t = 1, 2, ..., T$. By definition,

$$\mathbf{C}_t = \int c_t\left(\zeta_t\right) g\left(\zeta\right) d\zeta$$

$$\mathbf{S}_t = \int s_t\left(\zeta_t\right) g\left(\zeta\right) d\zeta.$$

Let $\mathbf{I}$ denote the economy-wide investment in physical capital (conducted by the firm in the consumption good sector). The market clearing in the consumption good sector is given by the condition

$$\sum_{t=1}^{T} \mathbf{C}_t + \mathbf{I} = \mathbf{Y}.$$

---

[4]At this point, I should clarify that the index t does not indicate calendar year, but instead the age of the child. At any calendar year τ there is a mass one of children who are t years-old. These are the children who were born in calendar year $\tau - t$. Because the focus of the paper is on steady states, for the sake of simplicity, I drop the calendar year index τ.

Analogously, equilibrium in the physical capital sector is given by

$$\sum_{t=1}^{T} \mathbf{S}_t = \mathbf{K}.$$

Let $x_t(\zeta_t)$ denote the investments in cognitive skill when the child is t years old. I use $\mathbf{X}_t$ to denote the aggregate investment by households with a t years old child, $t = 1, 2, ..., T$. When the child is t years old, $t = 1, 2, ..., T-1$, this is

$$\mathbf{X}_t = \int x_t(\zeta_t) g(\zeta) d\zeta.$$

When the child is T years old, we have to keep track of the fact that some children receive investments beyond the minimum amount while others do not. The share of the children who receive investments is the share of children who become college graduates. Consequently, aggregate investment by households with a T years old child is given by

$$\mathbf{X}_T = \int_{\{\zeta_T / x_T(\zeta_T)=\underline{x}\}} \underline{x} g(\zeta | x_T(\zeta_T) = \underline{x}) d\zeta +$$
$$+ \int_{\{\zeta_T / x_T(\zeta_T)>\underline{x}\}} x_T(\zeta_T) g(\zeta | x_T(\zeta_T) > \underline{x}) d\zeta.$$

The market clearing condition for this sector is

$$\sum_{t=1}^{T} \mathbf{X}_t = \mathbf{E}.$$

To compute the aggregate stock of efficiency units, let $g_h(h)$ denote the probability density function of adult efficiency units. In households where children are t years old, $t = 1, 2, ..., T-2$, they supply an amount of efficiency units that is given by

$$\mathbf{H}_t = \int h g_h(h) dh.$$

In households where children are T years old, we may have two different types of people supplying efficiency units: the parent and the child who is only receiving the minimum amount of investments, $\underline{x}$. Let $g_\theta(\theta_T)$ denote the probability density function of efficiency unit (determined by cognitive skill) for the children who are T years old.

$$\mathbf{H}_T = \int h g_h(h) dh + \int_{\{\zeta_T / x_T(\zeta_T)=\underline{x}\}} \theta_T g_\theta(\theta_T | x_T(\zeta_T) = \underline{x}) d\theta_T.$$

The total supply of efficiency units in every calendar year in this economy is given by $\mathbf{H}$ which is defined as

$$\mathbf{H} = \sum_{t=1}^{T} \mathbf{H}_t.$$

Let $\mathbf{L}, \mathbf{U}$ denote the aggregate amount of efficiency units allocated to the consumption and education

good sector, respectively. Feasibility of the efficiency units allocation implies

$$\mathbf{L} + \mathbf{U} = \mathbf{H}.$$

2.6 Definition of Stationary Equilibrium

Definition 1 *A Stationary Recursive Competitive Equilibrium is a set of functions* $V(\zeta_t)$, $c_t(\zeta_t)$, $x_t(\zeta_t)$, $s_t(\zeta_t)$, *for* $t = 1, .., T$, *aggregate quantities* $\mathbf{K}$, $\mathbf{L}$, $\mathbf{Y}$, $\mathbf{U}$, *wage rate* w, *interest rate* r, *prices of investment goods* p, *distributions of parents across states,* $g(\zeta)$ *such that:*

(1) *Given prices* w *and* r, *the functions* $\{V(\zeta_t)\}_{t=1}^{T}$, $\{c_t(\zeta_t), x_t(\zeta_t), s_t(\zeta_t)\}_{t=1}^{T}$ *solve the parent's maximization problem.*

(2) *Given prices* w *and* r, $\mathbf{K}$, *and* $\mathbf{L}$ *maximize the consumption-good firm's profits and* $\mathbf{U}$ *maximizes the education-good sector firm's profit.*

(3) *Markets for consumption, investments in education, physical capital and efficiency units clear.*

(4) *The distributions of households across states* $\{\mu_t(\theta_t, h, s_t, \varepsilon_t)\}_{t=1}^{T}$ *are calendar-year invariant and are determined as a fixed point of an operator that maps current-calendar-year distributions into next-calendar-year distributions taking into account the parent's optimal decisions and the evolution of exogenous states.*

3 Estimation and Calibration

As noted by Todd and Wolpin (2007), the identification and estimation of cognitive production functions is a challenging task. One major problem is that both the inputs and outputs can only be proxied, producing the problem of measurement error.

Another problem lies in the fact that one may want to allow for unobservables known to parents and children but not to the observing economist, denoted respectively by π_p and π_c, which can be correlated with parental investments, parental skills, and current skills of the children in a general fashion. I extend the production function (1) by including π_c and π_p, as well as unobservable independent inputs η_{t+1}^{θ}:

$$\ln\theta_{t+1} = \ln\delta_{n_t} + \frac{\rho_{n_t}}{\phi_{n_t}}\ln\left\{\gamma_{1,n_t}e^{\phi_{n_t}\ln\theta_t} + \gamma_{2,n_t}e^{\phi_{n_t}\ln x_t} + \gamma_{3,n_t}e^{\phi_{n_t}\ln h}\right\} + \pi_c + \pi_p + \eta_{t+1}^{\theta} \qquad (8)$$

An important (and challenging) feature of the data I use is the richness of measurements associated with skills of children and parents as well as information on parental investments in child's skills. To reduce the dimensionality of the problem and at the same time use all of the information in the data set without being subject to so many arbitrary choices is to consider state-space models, such as the one used by Geweke (1977), Sargent and Sims (1977), Shumway and Stoffer (1982) and Watson and Engle (1983). This is the approach used by Cunha, Heckman, and Schennach (2010) and the one I use in this paper.[5] In what follows, I briefly describe the estimating equations and the data used in the implementation of their technique.

In this paper, the unobservable latent variables are the skills of the child, θ_t, the skills of the parent, h, the parental investments in child's skills, x_t, the heterogeneity across children, π_c, and across parents, π_p. Each one of these latent variables has an observable counterpart, which in the literature is referred to

---

[5] A more complete treatment of the identification and estimation of the technology of skill formation that I implement in this paper can be found in Cunha, Heckman, and Schennach (2010).

as a "measurement". To fix notation, let $M^{\theta}_{l,t}$ denote the lth measurement associated with the cognitive skill of the child at age t, θ_t, for $l = 1, 2, \ldots, L_{\theta}$ and $t = 1, ..., T$. [6] I model the measurements as separable additive functions of $\ln \theta_t$,

$$M^{\theta}_{l,t} = \mu^{\theta}_{l,t} + \alpha^{\theta}_{l,t} \ln \theta_t + \omega^{\theta}_{l,t} \text{ for } l = 1, \ldots, L_{\theta}, \; t = 1, \ldots, T. \tag{9}$$

The coefficients $\alpha^{\theta}_{l,t}$ are called factor loadings. The error terms $\omega^{\theta}_{l,t}$ are called uniquenesses. The terms $\mu^{\theta}_{l,t}$ are mean functions that may depend on observable regressors, but for ease of notation I keep them implicit.[7] I impose $E\left(\omega^{\theta}_{l,t}\right) = 0$ for $t = 1, \ldots, T$, $l = 1, ..., L_{\theta}$. The separability assumption is not crucial for identification, but I will focus on it because it is the one I use in the estimation reported below.[8] The scale and location of the cognitive skill factor is determined below.

I treat parental investments similarly. I assume that x_t is not directly observed either. Consistent with the notation defined above, I will use the terms $\mu^{x}_{l,t}$, $\alpha^{x}_{l,t}$, and $\omega^{x}_{l,t}$ to denote the mean, factor loading, and uniqueness in the equation of the lth measurement on investment at age t of the child, $M^{x}_{l,t}$. The relationship between the measurement $M^{x}_{l,t}$ and the investment factor x_t is modelled as:

$$M^{x}_{l,t} = \mu^{x}_{l,t} + \alpha^{x}_{l,t} \ln x_t + \omega^{x}_{l,t} \text{ for } l = 1, \ldots, L_x, \; t = 1, \ldots, T. \tag{10}$$

I set $\alpha^{x}_{1,t} = 1$ and $\mu^{x}_{1,t} = 0$ for $t = 1, \ldots, T$. These normalizations set the scale and the location of x_t. In particular, note that the mean of the investment factor, x_t, will not be zero. It will be determined by the mean of measurement $M^{x}_{1,t}$ for all t. I also impose $E\left(\omega^{x}_{l,t}\right) = 0$ for $t = 1, \ldots, T$, $l = 1, ..., L_x$.

Let h denote the mother's latent cognitive skill which is also assumed to be unobserved by the econometrician. Again, I use the terms μ^{h}_{l}, α^{h}_{l}, and ω^{h}_{l} to represent the mean term, the factor loading, and the uniqueness in the equation that relates the measurement M^{h}_{l} to the parental skill factor h. I assume that the relationship is linear and separable:

$$M^{h}_{l} = \mu^{h}_{l} + \alpha^{h}_{l} \ln h + \omega^{h}_{l}. \tag{11}$$

I normalize $\alpha^{h}_{1} = 1$ and $E\left(\ln h\right) = 0$ as well as $E\left(\omega^{h}_{l}\right) = 0$ for $l = 1, ..., L_h$.

I assume that we have access to a set of proxies $M^{\pi_c}_{l}, M^{\pi_p}_{l'}$, $l = 1, ..., L_{\pi_c}$, $l' = 1, ..., L_{\pi_p}$ that are affected by the unobserved heterogeneity across children, π_c, and parents, π_p, respectively. The $M^{\pi_c}_{l}$ proxies are choices or outcomes once the child reaches adulthood: educational attainment, income, employment status, and occupational choice. All of these are potentially affected by unobserved heterogeneity beyond cognitive skill. Consistent with the notation developed above, let $\mu^{\pi_c}_{l}$ denote the mean term. Let $\kappa^{\pi_c}_{l}$ denote the loading on the heterogeneity factor π_c. Let $\omega^{\pi_c}_{l}$ denote the uniqueness. Because I want to capture heterogeneity beyond that in cognitive skills, I will also include the cognitive skill factor of the child at the last age the agent is a child, θ_T, and I denote by $\alpha^{\pi_c}_{l}$ the factor loading associated with it. The equation that relates the outcomes and choices to heterogeneity across children is modelled as:

$$M^{\pi_c}_{l} = \mu^{\pi_c}_{l} + \alpha^{\pi_c}_{l} \ln \theta_T + \kappa^{\pi_c}_{l} \ln \pi_c + \omega^{\pi_c}_{l} \tag{12}$$

---

[6] The number of measurements on cognitive skills, L_{θ}, can vary with the age of the child. To save on notation, I assume it does not although in the empirical implementation below it does so.

[7] In the empirical implementation, the regressors are an intercept and the age of the mother.

[8] For a discussion of nonlinear, nonseparable measurement equations see Cunha, Heckman, and Schennach (2010).

for $l = 1, ..., L_{\pi_c}$. Again, it is necessary to make the location and scale normalization, so I set $\kappa_1^{\pi_c} = 1$, $E(\ln \pi_c) = 0$ and $E(\omega_l^{\pi_c}) = 0$ for $l = 1, ..., L_{\pi_c}$. It is important to point out that π_c need not be orthogonal to θ_T. A similar model is assumed for the parents:

$$M_l^{\pi_p} = \mu_l^{\pi_p} + \alpha_l^{\pi_p} \ln h + \kappa_l^{\pi_p} \ln \pi_p + \omega_l^{\pi_p} \tag{13}$$

I set $\kappa_1^{\pi_p} = 1$, $E(\ln \pi_p) = 0$ and $E(\omega_l^{\pi_p}) = 0$ for $l = 1, ..., L_{\pi_p}$.

Another set of equations in a dynamic factor model describes the evolution of the latent dynamic factors. In this paper, the main transition equation is the technology of cognitive skills (8). It is this equation that links unobserved skills tomorrow, θ_{t+1}, with unobserved current skills, θ_t, unobserved current investments, x_t, unobserved parental skills, h, and unobserved child and parental heterogeneity, π_c and π_p, respectively.

Because test scores for the parents are observed only once, I model the parental skill factor as a static factor. The same will be true for the heterogeneity factors. This implies a strong restriction on their evolution. For example, the parental skill factor h evolves according to the degenerate transition equation:

$$\ln h = m_h (\ln \theta_t, \ln x_t, \ln h, \ln \pi_c, \ln \pi_p) = \ln h$$

and the same holds for the other two static factors, $\ln \pi_c, \ln \pi_p$.

3.1 Anchoring Skills in a Meaningful Metric

As remarked by Cunha, Heckman, and Schennach (2010), an important problem in the estimation of production functions is that test scores do not have a natural metric. To set the scale and location of the factors in a meaningful way, I use an outcome that (a) does have a natural metric and (b) is correlated with cognitive skills. Consider the logarithm of yearly labor income when the person is 25 years old, $M_1^{\pi_c}$. It satisfies condition (a) because it is measured in dollars (or some other currency that can be exchanged for real goods, such as apples). According to the evidence in Heckman, Stixrud, and Urzua (2006), it also satisfies (b) as higher cognitive skills cause $M_1^{\pi_c}$ to increase. I can model the relationship between the natural logarithm of wages at age 25 and cognitive skills at age 14 as:

$$M_1^{\pi_c} = \mu_1^{\pi_c} + \alpha_1^{\pi_c} \ln \theta_T + \ln \pi_c + \omega_1^{\pi_c} \tag{14}$$

Cunha, Heckman, and Schennach (2010) call equations such as (14) anchoring equations and the function:

$$A_\theta (\ln \theta_T) = \mu_1^{\pi_c} + \alpha_1^{\pi_c} \ln \theta_T \tag{15}$$

anchoring functions. These functions can be used to transform information from scores into dollar figures. The anchor functions allow us to estimate the technology function (8) by measuring skills according to a dollar metric, which is meaningful.

We treat parental cognitive skill the same way. Assume that the first measurement on the factor that captures heterogeneity across parents, $M_1^{\pi_p}$, is the natural logarithm of labor income when the parents are 25 years old. I define the anchoring equation and the anchoring function to be, respectively:

$$M_1^{\pi_p} = \mu_1^{\pi_p} + \alpha_1^{\pi_p} \ln h + \ln \pi_p + \omega_1^{\pi_p}$$

$$A_h \left(\ln h \right) = \mu_1^{\pi_p} + \alpha_1^{\pi_p} \ln h$$

I redefine the technology of skill formation in terms of the anchored factor:

$$A_\theta \left(\ln \theta_{t+1} \right) = \ln \delta_{n_t} + \frac{\rho_{n_t}}{\phi_{n_t}} \ln \left\{ \gamma_{1,n_t} e^{\phi_{n_t} A_\theta (\ln \theta_t)} + \gamma_{2,n_t} e^{\phi_{n_t} \ln x_t} + \gamma_{3,n_t} e^{\phi_{n_t} A_h (\ln h)} \right\} + \tag{16}$$

$$+ \gamma_{4,n_t} \ln \pi_c + \gamma_{5,n_t} \ln \pi_p + \eta_t^k,$$

where $\sum_{k=1}^5 \gamma_{k,n_t} = 1$, $0 \le \rho_{n_t} < 1$, $\phi_{n_t} \le 1$ for $n_t = 1, 2, 3$.

3.2   The Estimation Algorithm

Although the discussion about identification does not rely on parametric assumptions, I use parametric maximum likelihood to estimate the model. This decision is mainly driven by the computational costs of solving a high-dimensional dynamic factor model. I now develop the likelihood function. Let I denote the number of children in the sample. To keep the notation simple I define:

$$\ln \lambda_t = \left(\ln \lambda_{1,t}, \ln \lambda_{2,t}, \ln \lambda_3, \ln \lambda_4, \ln \lambda_5 \right)' = \left(\ln \theta_t, \ln x_t, \ln h, \ln \pi_c, \ln \pi_p \right)'.$$

Let $M_{t,l,i}^k$ denote the measurement l associated with the factor k for person i in period t. Let $d_{t,l,i}^k$ take the value one if the measurement $M_{t,l,i}^k$ is observed and zero if it is missing. Let $\omega_{t,l,i}^k$ denote the measurement error associated with the measurement $M_{t,l,i}^k$. Let $p_{\omega_{l,t}^k}$ denote the density function of $\omega_{t,l,i}^k$. Let M contain all the observed measurements, across all agents i, periods t, factors k. Assuming independence across observations, the likelihood $p \left(M \right)$ is

$$p \left(M \right) = \prod_{i=1}^I \int \dots \int p_{\ln \lambda} \left(\ln \lambda \right) \left[\prod_{t=1}^T \prod_{k=1}^5 \prod_{l=1}^{L_k} p_{\omega_{lt}^k} \left(M_{l,t,i}^k - \mu_{l,t}^k - \ln \lambda_{t,i} \alpha_t \right)^{d_{t,l,i}^k} \right] d \ln \lambda. \tag{17}$$

This is maximized subject to a parametric technology constraint and the normalizations on the measurements discussed above. I assume that the measurement error ω_{lt}^k is classical, so that $p_{\omega_{lt}^k | \ln \lambda_t^k} \left(\omega_{lt}^k | \ln \lambda_t^k \right) = p_{\omega_{lt}^k} \left(\omega_{lt}^k \right)$. This considerably reduces the number of terms needed to form the likelihood.

In principle, one can estimate the parameters in $\left\{ \mu_{l,t}^k, \alpha_{l,t}^k \right\}$, the parameters of the technology, and the $p_{\ln \lambda} \left(\ln \lambda \right)$ by maximizing (17) directly. In order to do that, one can approximate $p \left(M \right)$ by computing the integrals numerically in a deterministic fashion. However, if the number of periods and the dimension of $\ln \lambda$ is large, the number of points required to evaluate the integrals is also very large.

A different approach to computation is to consider nonlinear filtering techniques. For this paper, I use the Unscented Kalman Filter (UKF), an extension of the standard Kalman Filter for nonlinear models, developed by Julier and Ulhman (1997) and applied in Cunha, Heckman, and Schennach (2010).[9]

3.3   The Data

I use a sample of 2233 white males taken from the Children of NLSY/79 (CNLSY/79) data set. The children may be as young as 0 years old and as old as 14 years old. The children of the NLSY/79 female

---

[9]The computational costs are nontrivial even after imposing parametric restrictions and assuming that measurement error is classical. The maximization of (17) requires more than ten days in a computer with 32 CPUs.

respondents have been assessed every two years. The assessments I use in this paper are intended to measure their cognitive skills as well as parental inputs for cognitive development measured through the home environment. A convenient feature of the CNLSY/79 data set is that I can also observe cognitive test scores for the parents.

3.3.1 Instruments Designed to Measure Child Skills

I estimate the production function for skills from birth to adolescence. The difficulty in the task is how to measure skills when the children are very young. In what follows, I list the instruments available in the CNLSY/1979 that are used to evaluate the skills children have at different stages of the developmental process. A more detailed description of the measurements can be found in Cunha, Heckman, and Schennach (2010).

At each age, the CNLSY/79 participants' cognitive skills are measured in multiple ways. At early ages, one can combine the Motor and Social Development Scale (MSD), which has items derived from standard measures of child development, such as the Bayley Scales of Infant Development. Another measure of skills for children in their early years are the Parts of the Body assessment and Memory for Locations assessment. The first attempts to measure a one- or two-year-old child's receptive vocabulary knowledge. The second measures a child's short-term memory.

From age three on, one can explore the information from the Peabody Picture Vocabulary Test - Revised (PPVT-R) which measures an individual's receptive (hearing) vocabulary for Standard American English and provides, at the same time, a quick estimate of verbal ability or scholastic aptitude.

Starting at age five, the main measurements of cognitive skills that I use are two components of the Peabody Individual Achievement Test (PIAT): PIAT Mathematics and PIAT Reading Recognition.[10] The PIAT Mathematics measures a child's attainment in mathematics as taught in mainstream education. It consists of eighty-four multiple-choice items of increasing difficulty. The PIAT Reading Recognition subtest measures word recognition and pronunciation ability. Children read a word silently, then say it aloud. The test contains eighty-four items, each with four options, which increase in difficulty from preschool to high school levels.

An important issue is the identification of the initial distribution of the skill factor. I use data on weight at birth for each child. There is an extensive literature that studies the effects of low birth weight on health and economic outcomes (see, e.g., Behrman and Rosenzweig, 2004; Currie and Hyson, 1999; Black, Devereux, and Salvanes, 2007; Case, Fertig, and Paxson, 2005). Another indicator of initial conditions I use is premature birth, which is the leading cause of infant mortality in industrialized societies (see Kramer *et al.*, 2001).

3.3.2 The Measures of Parental Inputs

The measures of quality of a child's home environment that are included in the CNLSY/79 survey are the components of the Home Observation Measurement of the Environment - Short Form (HOME-SF). These are a subset of the measures used to construct the HOME scale designed by Bradley and Caldwell (1980, 1984) to assess the cognitive stimulation these children receive through their home environment, planned

---

[10]I do not use the PIAT Reading Comprehension battery since it is not administered to the children who score low on the PIAT Reading Recognition.

events and family surroundings. These measurements have been used extensively as inputs to explain child characteristics and behaviors (e.g., Todd and Wolpin, 2007). As discussed in Linver, Brooks-Gunn, and Cabrera (2006), some of these items are not useful because they do not vary much among families (i.e., more than 90% to 95% of all families make the same response). For this reason, I limit the analysis to sixteen components of the HOME-SF score. Not all components are surveyed at every age of the child.

One of the problems of the HOME-SF data is that none of the components have a clear scale based on dollars. I obtain external data to scale both the mean and the variance of the investment factor in dollar terms. This is the measurement $M_{1,t}^{x}$ for $t = 1, ..., T$. Since 1960, the U.S. Department of Agriculture (USDA) has published its Annual Report on "Expenditures on Children by Families", which provides estimates of expenditures on children from birth through age 17. The estimates computed by the USDA are used in setting state child support guidelines and foster care payments (see, e.g., Lino, 2000).

The data used in the expenditure estimate on children are from the 1990-92 Consumer Expenditure Survey - Interview portion (CE). The sample consists of 12,850 husband-wife households and 3,395 single-parent households and was weighted to reflect the U.S. population of interest. The figures were estimated according to groups of income level, family size, and age of the youngest child. Estimated family expenditures on child care and education were allocated equally among children.

3.3.3 Instruments Designed to Measure Parental Skills

During the summer and fall of 1980, NLSY79 respondents participated in an effort of the U.S. Departments of Defense and Military Services to update the norms of the Armed Services Vocational Aptitude Battery (ASVAB). Even though the ASVAB consists of a battery of ten tests, I only use the following six measures: (1) arithmetic reasoning; (2) word knowledge; (3) paragraph comprehension; (4) numerical operations; (5) coding speed and (6) mathematics knowledge. The scores were normalized to be mean zero and standard error one in the entire NLSY/79 sample. Note that the typical mother of the white males taken from the CNLSY/79 has higher mean score than the rest of the NLSY/79 sample.

3.3.4 Measurements of Heterogeneity Across Children

To measure heterogeneity across children, I use data on economic outcomes and choices. For children who are at least 21 years old, I use information on labor income and education attainment. For children 19 years old and older, I use information on fertility, birth control practices, crime participation and drinking habits. One potential problem is that we only observe these outcomes for children who are born from very young mothers.

3.3.5 Measurements of Heterogeneity Across Parents

The heterogeneity across families is measured by information on log total family income when the mother is 30 years old, log family income of the mother's spouse, when the mother is also 30 years old, the number of children the mother has had by age 30, and the number of abortions the mother has had by age 30. I abstain from using the panel aspect of log family income since the heterogeneity factor may be strongly correlated with permanent income, which could suffer from colinearity with parental investments.

3.4 Results

I start by determining three different development stages. The first developmental stage starts at age 0 and finishes at age 4. The second development stage starts at age 5 and finishes at age 9. The third developmental stage starts at age 10 and finishes at age 14, which is the last age for which test scores and home inputs are measured in the CNLSY/79. For each developmental stage, I specify the *CES* technology 16.

The measurement equations contain two regressors (age of the mother during each period and a constant). The initial distribution of the factors allows them to be freely correlated. Table 1 presents the result of estimating an anchored technology with the heterogeneity factors π_c and π_p. Self-productivity becomes more important as the children age: the coefficient on past stocks of cognitive skills grow from roughly 0.73 in the first developmental stage, to 0.80 in the second developmental stage, to 0.87 in the third developmental stage. The impact of investments on child cognitive skills is slightly higher at 0.12 early on and around 0.08 later on. When I formally test the difference between the coefficients in the first and second stage, or first and third stage, I find that the p-value is 0.068. This is consistent with evidence that cognitive skills sensitive period is the early developmental stage.

The complementarity parameter indicates a CES specification with a little more curvature than Cobb-Douglas, which cannot be rejected in the estimation. For the purpose of this paper, it is important to remark that the investments at different ages of childhood are not perfect substitutes. This is evidence against the literature that collapses childhood in one period.

3.5 Calibration of Remaining Parameters

The first parameter to be defined is the number of periods, which I set as $T = 19$. That is, each person lives for 38 years, the first 19 of which as a child receiving investments in skills. The remaining 19 years as the parents in a household that has a child.

To estimate the parameter values that describe the law of motion of productivity shocks (2) I use the method described in Cunha, Heckman, and Navarro (2005). This method allows the analyst to separate heterogeneity from unforecastable components in lifecycle earnings (risk). Once I obtain the process that describes the risk agents face, I can compute the implied serial covariance matrix that is only due to risk. I then choose the AR(1) process that provides the best fit to the serial covariance matrix. This results in point estimates $\hat{\rho}_\varepsilon = 0.7910$ and $\hat{\sigma}_\eta^2 = 0.1406$ with standard errors given by $s_{\hat{\rho}_\varepsilon} = 0.0721$ and $s_{\hat{\sigma}_\eta^2} = 0.0235$, respectively.

The AR(1) specification and parameters that I find are different from what is usually estimated in the literature of consumption and income inequality. For example, Blundell, Pistaferri, and Preston model residual income as the summation of a persistent process and a transitory shock that follows an MA(1) process. It is unclear how these different specifications affect the welfare calculations that I perform below. Clearly, the more persistent the shocks, the more important the lack of insurance to shocks. On the other hand, the variance of the shocks that I estimate are three to seven times larger than what is found in the literature. The larger the variance, the larger the welfare costs of the lack of insurance.

The minimum investment x̲ for the children who stop in high-school is 0.2, which is equivalent to U$600.[11] This value is determined endogenously so that the proportion of people who enroll in college in

---

[11] All dollar figures in this paper are deflated to year 2000.

the steady state described below matches the fraction in the NLSY/79 data.

The yearly college tuition φ is set at U\$4,164. This amount is the national average amount paid in 2000 for state residents for a four-year public college.[12]

The utility function is assumed to be described by a Constant Relative Risk Aversion (CRRA) function:

$$u\left(c_t\right) = \frac{c_t^{1-\lambda} - 1}{1 - \lambda}$$

Following Browning, Hansen, and Heckman (1999), I set $\lambda = 2$. This parameter plays an important role in determining the profile of consumption and investment over time. If I used the value $\lambda = 0$, then parents would choose a very smooth profile for investments and a very volatile profile for consumption. If, on the other hand, I used $\lambda \rightarrow \infty$, then parents would choose a consumption profile that is very smooth, and investments would react more strongly to shocks in income.

The aggregate production function $F\left(K, L\right)$ is assumed to be represented by a Cobb-Douglas:

$$F\left(K, L\right) = AK^\alpha L^{1-\alpha}$$

The parameter A is set to 1.81. This is the value so that the parental labor income matches the average in the NLSY/79 data for males who are 30 years old. The share parameter is set to $\alpha = 0.36$ as in Heckman, Lochner, and Taber (1998). Finally, the discount factor is set to $\beta = 0.95$ and the physical capital depreciation parameter $\delta = 4\%$.

4   Computation of the Steady State

In this section, I compute the equilibrium of the economy described in Section 2. I show that the model's steady state equilibrium is consistent with many features documented in the literature that studies cognitive skill accumulation as well as educational decisions.

4.1   Solution Algorithm

The model is solved numerically by iterating the value function. Although the state and control variables are continuous, I choose to discretize the state space. At each age t of the child, I choose a grid for skills θ_t and parental investment x_t. I denote the grid for child cognitive skills by $\vec{\theta}_t = \left\{\bar{\theta}_{t,1}..., \bar{\theta}_{t,K_\theta}\right\}$. The grid for parental investment is $\vec{x}_t = \left\{\bar{x}_1, ..., \bar{x}_{K_x}\right\}$. The remaining state variables have grids that do not depend on the age of the child. Thus, the grid for parental savings is $\vec{s} = \left\{\bar{s}_1, ..., \bar{s}_{K_s}\right\}$, while the grid for parental skill h is $\vec{h} = \left\{h_1, ..., h_{K_h}\right\}$. The grid for productivity shocks is $\vec{\varepsilon} = \left\{\varepsilon_1, ..., \varepsilon_{K_\varepsilon}\right\}$. In this paper, the size of the grid is the following: $K_\theta = 25$, $K_x = 21$, $K_h = 25$, $K_s = 35$, and $K_\varepsilon = 7$.

The algorithm to calculate the steady state can be summarized in the following few steps.

Step 1: Guess equilibrium interest rate and wages r_j, w_j.

Step 2: Given current prices r_j and w_j, given $V_{j,k}\left(t, h, \theta_t, s_t, \varepsilon_t, n_t\right)$, calculate $V_{j,k+1}\left(t, h, \theta_t, s_t, \varepsilon_t, n_t\right)$ by discretizing the state-space and proceeding with value function iteration.

---

[12]This information was obtained in Gillen, Robe, and Garret (2011).

Step 3: If $max\,|V_{j,k}\left(t,h,\theta_t,s_t,\varepsilon_t,n_t\right) - V_{j,k+1}\left(t,h,\theta_t,s_t,\varepsilon_t,n_t\right)| < \epsilon_V$ convergence of value functions has been achieved. If not, go back to Step 2.

Step 4: Draw θ_1 and sequence of shocks $\{\varepsilon_t\}$ from stationary distributions 210,000 times. Discard the first 200,000 to allow for series to converge to stationary distributions of $h, \theta_t, s_t, x_t, \varepsilon_t$ for all t. Given the simulated series, compute aggregate savings, aggregate supply of labor in the goods sector, aggregate labor supply in the education sector, aggregate demand for investments and compute r^{j+1}, w^{j+1}. If $max\,\{|r^j - r^{j+1}|, |w^j - w^{j+1}|\} < \epsilon_{r,w}$ convergence of prices is achieved. If not, go back to Step 1.

4.2   Features of the Steady State Equilibrium

The model predicts many features of the lifecycle skill accumulation process and education decisions. In Table 2, I consider the fraction of children who are at least 19 years old and enroll in any type of college. In the CNLSY/79 data, 51.06% of the children stop studying once they complete high school. The remaining 48.94% go on to college. The model's state steady state equilibrium predicts that 51.43% of the children would stop after completing the high school and 48.57% going on to college. This is the information I use to pin down the minimum value of investment x̲.

Figure 1 is the model's counterpart to the empirical evidence presented by Carneiro and Heckman (2003) discussed in the introduction of this paper, but my model tends to overpredict college enrollment, especially for the top tercile of cognitive skills at age 14. Let θ_{14} denote skill when the child is fourteen years old. Let y_{18} denote the parent's income when the child is eighteen years old. Let T_k denote the terciles of θ_{14}, $k = 1, 2, 3$. Let Q_j denote the quartiles of parent's income, $j = 1, 2, 3, 4$. Let $S = 1$ if at age $t = 19$ the child enrolls in college. Let $S = 0$ otherwise. I use the expression $1\,(x)$ to denote the function that takes the value one if x and zero otherwise. Let $P_{k,j}$ denote the fraction of children in skill tercile k and parent's income quartile j who enroll in college at age 19:

$$P_{k,j} = \sum \frac{1(S = 1, \theta_{14} \in T_k, y_{18} \in Q_j)}{1(\theta_{14} \in T_k, y_{18} \in Q_j)}$$

The steady state equilibrium I compute generates the qualitatively similar gaps in skills across quartiles of permanent family income as are documented by Carneiro and Heckman (2003). The difference is that the model underpredicts the differences in skills measured across family income groups. Figure 2 shows the model's simulations. I construct this figure as follows. At every age t, let θ_t denote skills when the child is t years old. Let μ_t, σ_t denote the mean and standard deviation of θ_t. Define $z_t = \frac{\theta_t - \mu_t}{\sigma_t}$. Let r denote the steady-state equilibrium interest rate, $(r = 4.48\%)$. Let y_t denote the parent's income when the child is t years old. Define the permanent income of the parent, Y, as:

$$Y = \sum_{t=0}^{T} \frac{y_t}{\left(1 + r\right)^t}.$$

I then define the quartiles of permanent family income, $Q_j\left(Y\right)$ for $j = 1, ..., 4$. Let N_j denote the number of children in quartile $Q_j\left(Y\right)$. Then, for each simulated child i, I compute the mean z_t if the child is in

18

quartile $Q_j(Y)$ of family income:

$$\mu(z_t, Q_j) = \frac{1}{N_j} \sum_i z_{i,t} 1 (Q_i = Q_j) \text{ for } j = 1, 2, 3, 4.$$

Figure 3 shows that the model's steady-state equilibrium also predicts college attendance based on cognitive skills measured at adolescent years. At the time of college decisions, children who go on to college tend to have more cognitive skills then the children who decide to stop studying upon completing high school. I check if measures of skills when the children are very young can predict college enrollment or not. It does seem so, according to Figure 4, where I plot the density of skills at age 3 conditional on educational choices at age 18. Abilities at age 3 can already predict educational decisions 15 years later.

In Figure 5, I show that in steady state, the decision to enroll is affected by returns to college. The children who go on to college tend to have higher returns than the children who stop after graduating from high school. This is consistent with the evidence presented by Cunha, Heckman, and Navarro (2006).

As defined above, x_t denotes parental investments in skills when the child is t years old, $t = 1, ..., 19$. The price of each investment good is w. Let X denote the present value of investments (measured in dollars) from birth to age 18, discounted according to the steady-state equilibrium interest rate r:

$$X = \sum_{t=1}^{18} \frac{wx_t}{(1+r)^{t-1}}.$$

Figure 6 compares the density of present value of investments of the children who stop at high school with the ones that go on to college. This figure does not include resources spent on college education, since it only includes investments made before age 19. The solid curve shows the density of X conditional on stopping at high-school, $f(X|S=0)$. The dashed curve shows the density of X conditional on the sample of individuals who enroll in college, $f(X|S=1)$. Clearly, the children who go on to college tend to receive higher lifetime investments than the ones that stop at high school.

4.3 Simulating Different Remediation Policies

The economy whose steady state I just characterized has market failures which distort the parental investment in the cognitive skills development of their children. I use the model to evaluate policies that can compensate for such distortions. In what follows, I present different alternative policies with respect to the point in time of the lifecycle of the child that the intervention occurs. The policies are financed by a flat income tax rate. The tax rate is chosen such that the government budget is balanced at every period. For each economy I present, I will calculate the steady state. I report the implied taxation on income, and I measure the welfare gains in dollars.

4.3.1 A 50% Tuition Subsidy

I consider an economy where the government taxes income and uses the revenue to finance a 50% flat tuition subsidy. This implies changing the budget equation (6) to:

$$c_T + s_1' + (px_T + 0.5\varphi) = (1 - \tau) wh\varepsilon_T + (1 + r) s_T \text{ if } x_T > \underline{x} \tag{18}$$

where, as before, φ denotes the college tuition. I use τ to denote the flat income tax rate. This rate is set in equilibrium so that the government runs a balanced budget.

Let $g_h(.)$ and $g_\varepsilon(.)$ denote the density functions of h and ε, respectively. It is easy to see that the government revenues are given by:

$$RG = (1 - \tau) w \sum_{t=1}^{T} \int \int h\varepsilon_t g_h(h) g_\varepsilon(\varepsilon_t) \, dh d\varepsilon_t. \tag{19}$$

The only government expenditure is the subsidy to college tuition, ST. Let ζ_T denote the state variables when the child is T years old and denote by g_ζ their joint distribution. The total expenditure by the government on the subsidy is:

$$ST = 0.5\varphi \int_{\{\zeta_T / x_T(\zeta_T) > \underline{x}\}} g(\zeta \mid x_T(\zeta_T) > \underline{x}) \, d\zeta.$$

In the definition of equilibrium, I add the condition that the policy has a balanced budget:

$$RG = ST. \tag{20}$$

To accomplish this task, I find the income tax rate that guarantees the equality in (20). As shown in Table 2, the tax rate is roughly 1.2%.

Under the subsidy, there is a large increase in college enrollment. The fraction of children who go on to college rises from 48.57% to 61.26%. In Figure 7, I compare the density of cognitive skills at age 18 conditional on choosing to stop at high school in the baseline against the economy with the tuition subsidy. Note that the latter is mildly dislocated to the left of the former. This indicates that the children who moved from high school to college from the steady state of the baseline economy against that of the tuition subsidy economy are mainly the children who are at the top percentiles in the baseline economy high-school distribution of cognitive skills. However, note that in Figure 8, the density of the college distribution of skills in the tuition economy is to the left of that in the baseline economy. This indicates that the agents who move from high school to college tend to have lower skills than the typical college person in the baseline economy.

Before I analyze the cost and benefit of the tuition subsidy, I describe the other remediation policies that may be implemented by a policymaker interested in increasing the production of cognitive skills in this economy.

4.3.2 Subsidizing Parental Investment in Early Years

Next, I analyze an economy in which the government executes a targeted early intervention. The intervention is targeted at children who are in the bottom quartile of the distribution of cognitive skills at age 3, which I denote $\hat{\theta}$. The parents must also satisfy a maximum income condition: their total resources at age three of the child, R_3, given by the sum of labor income and stock of financial assets, must not exceed U\$40,000. The children who satisfy both conditions are henceforth denoted "disadvantaged children". The government subsidizes 90% of the cost of investment made by the parents of disadvantaged children. It finances this policy by charging an income tax rate.

The budget constraint of each parent in this economy is affected by this policy. At any period t, the resources are now the sum of financial wealth and after-tax labor income: $(1 - \tau) \, wh\varepsilon_t + (1 + r) \, s_t$. For the parents of disadvantaged children, the budget constraint when the children are ages 3 and 4 is now:

$$c_t + 0.1px_t + s_{t+1} = (1 - \tau) \, wh\varepsilon_t + (1 + r) \, s_t.$$

The government revenue is as defined in (19). The government expenditures are given by EY, where:

$$EY = \sum_{t=3}^{4} 0.9p \int_{\left\{\zeta_t / \theta_3 \leq \hat{\theta}, R_3 \leq 40,000\right\}} g_{\zeta_t}\left(\zeta_t | \theta_3 \leq \hat{\theta}, R_3 \leq 40,000\right) d\zeta_t.$$

In steady state equilibrium, the tax rate is such that the government budget is balanced:

$$RG = EY.$$

4.3.3 Subsidizing Parental Investments in Many Periods

I turn now to the description of the last remediation policy to be implemented in this paper. The government subsidizes investments for the parents from age 3 until the period before the parents decide to send the child to college or not. Unlike the early intervention above, this does not condition the subsidy on the level of cognitive skills of the child. It only considers the resources the parents have at each age of the child. Let $\underline{R}$ denote the maximum amount of resources that qualify the parents to receive the subsidy. Let χ denote the subsidy rate. Let R_t denote parental resources when the child is t years old. I denote by $\chi(R_t)$ the subsidy schedule proposed by the government:

$$\chi(R_t) = \chi \max\left\{0, \frac{\underline{R} - R_t}{\underline{R}}\right\}.$$

In this economy, I set $\chi = 0.4$ and $\underline{R} = $ U\$40,000. This means that parents who have net resources that are zero, get a 40% subsidy. Parents that have resources above U\$40,000 receive no subsidy. Again, the government collects an income tax, according to a flat tax rate. The budget constraint for the parent becomes:

$$c_t + (1 - \chi(R_t)) \, px_t + s_{t+1} = (1 - \tau) \, wh\varepsilon_t + (1 + r) \, s_t$$

The government revenue is as described by equation (19). The government expenditure is MP, where:

$$MP = \sum_{t=1}^{T-1} \int \chi(R_t) \, px_t g_{\zeta_t}\left(\zeta_t |\right) d\zeta_t.$$

Again, the tax rate is such that in steady state the government budget is balanced:

$$RG = MP.$$

4.3.4 Discussion

In this final section, I calculate the welfare under each different policy. Let $c_t^B, c_t^T, c_t^E, c_t^M$ denote the consumption of a household when a child is t years old in the baseline economy, the economy with the 50% tuition subsidy, the early subsidy economy and the economy with the multiple period subsidy, respectively. Define the ex-post lifetime utility of a household in the baseline economy as V^B :

$$V^B = \sum_{t=1}^{T} \beta^{t-1} u\left(c_t^B\right) + \beta^T E\left[V^{B\prime}\right]$$

where $V^{B\prime}$ is the expected lifetime utility of the descendent. For any other alternative economy k, $k = T, E, M$; define the ex-post lifetime utility of the household in the alternative economy k as:

$$V^k = \sum_{t=1}^{T} \beta^{t-1} u\left(c_t^k\right) + \beta^T E\left[V^{k\prime}\right] \text{ for } k = T, E, M.$$

The equivalent variation is the constant dollar amount λ^k that the household must receive in order to be just as well in the baseline economy as in alternative economy k :

$$\sum_{t=1}^{T} \beta^{t-1} u\left(c_t^B + \lambda\right) + \beta^T E\left[V^{B\prime}\right] = \sum_{t=1}^{T} \beta^{t-1} u\left(c_t^k\right) + \beta^T E\left[V^{k\prime}\right]$$

Let r^B denote the steady-state equilibrium interest rate in the baseline economy. The present value of the equivalent variation Λ is defined as:

$$\Lambda = \sum_{t=1}^{T} \frac{\lambda}{\left(1 + r^B\right)^{t-1}}.$$

In Figure 9, I show the average present value of the equivalent variation by deciles of the present value of lifetime earnings. For example, agents in the first decile of present value of lifetime earnings must get roughly U\$1,250.00 in the baseline economy to be indifferent to an economy with a 50% tuition subsidy. Under the tuition subsidy, the agents in deciles 6 and 7 are the ones that tend to gain the most. Under the early subsidy regime, the winners are mainly the parents at the lowest income deciles. Under the multiple period subsidy policy, the parents at deciles 4 and 5 tend to gain the most.

These findings highlight the importance of dynamic complementarity in the production of skills when markets are incomplete. Given the large variance of uninsurable shocks in income, bad realizations of shocks could be permanently costly in terms of cognitive skill formation. This fact amplifies the incentives for parents to engage in precautionary savings which they normally have in models such as in Laitner (1992). The policies that subsidize investments throughout the entire childhood period thus serve as a form of insurance against bad realizations in income and also reduce the need for parental precautionary savings.

The policy that subsidizes early investments is expensive and targets poor parents with low-skill children. The size of the subsidy is non-trivial, and they shift parental incentives from consumption towards investments in children and higher savings. Because of the dynamic complementarity in investments, the higher early investment raises the investments in later periods, and part of the increase in parental savings

is due to self-insurance motives to finance this higher investment.

The policy that subsizes tuition, on the other hand, has high benefits for middle-income parents who are likely to benefit from it since the majority of these households have children who are at the margin of whether or not to attend college. Given the large uncertainty that parents face in this economy, poor parents tend to discount the benefit of the policy far more than middle- and higher- income parents. For this reason, they do not benefit very much from a policy that intervenes so late and before so much uncertainty has been revealed.

In Figure 10, I plot the stationary densities of adult skills for the baseline, early subsidy, and early and late subsidy economy. I find that the economy in which investments of poor parents are subsidized during many periods generates a distribution of skills that first-order stochastically dominates the remaining.[13] Consistent with these welfare effects of the policies, the policies differ in their effectiveness in promoting the college attendance of disadvantaged children (Table 3). The policy that subsidizes investments across all periods of the child's lifetime tend to disproportionately increase the college enrollment of poor children.

5   Conclusion

This paper formulates, identifies, estimates, and interprets a multistage model of the production of skills in childhood. Investments made at different ages of the child are not perfect substitutes as assumed in the previous literature. In the model and in the estimation provided, I allow the technology to vary according to childhood developmental stages. The effects of parental investments are allowed to be different at different stages so that I can capture the notion of critical and sensitive periods which are much discussed in the literature on child development.

Parents are subject to lifetime constraints. They cannot leave debts to their children. They also face uninsurable productivity shocks in labor income. These market failures distort the equilibrium allocation of investments in the cognitive development of their children. I estimate the parameters describing the evolution of productivity shocks.

The model's empirically grounded steady-state equilibrium explains a variety of facts about cognitive skills, education, and child development. It predicts the observed selection into college by quartiles of family income and terciles of skills measured at adolescent years. Moreover, it is consistent with gaps in cognitive skills that are present at very early ages. Finally, it reproduces the pattern of selection into college based on cognitive skills.

I evaluate the impact of different remediation policies on the stationary distribution of cognitive skills and welfare. I consider a 50% tuition subsidy, a targeted early investment subsidy, and a multiperiod subsidy investment that is contingent on parental resources. I show that the policy that starts early and follows up later dominates the other policies in welfare, since it is the one that generates the highest equivalent variation across all deciles of permanent income. Indeed, it also generates a stationary distribution of cognitive skills that first-order stochastically dominates the ones generated by the baseline economy and other remediation policies.

---

[13]The density associated with the tuition economy falls between that taken from the baseline and that taken from the early subsidy economy.
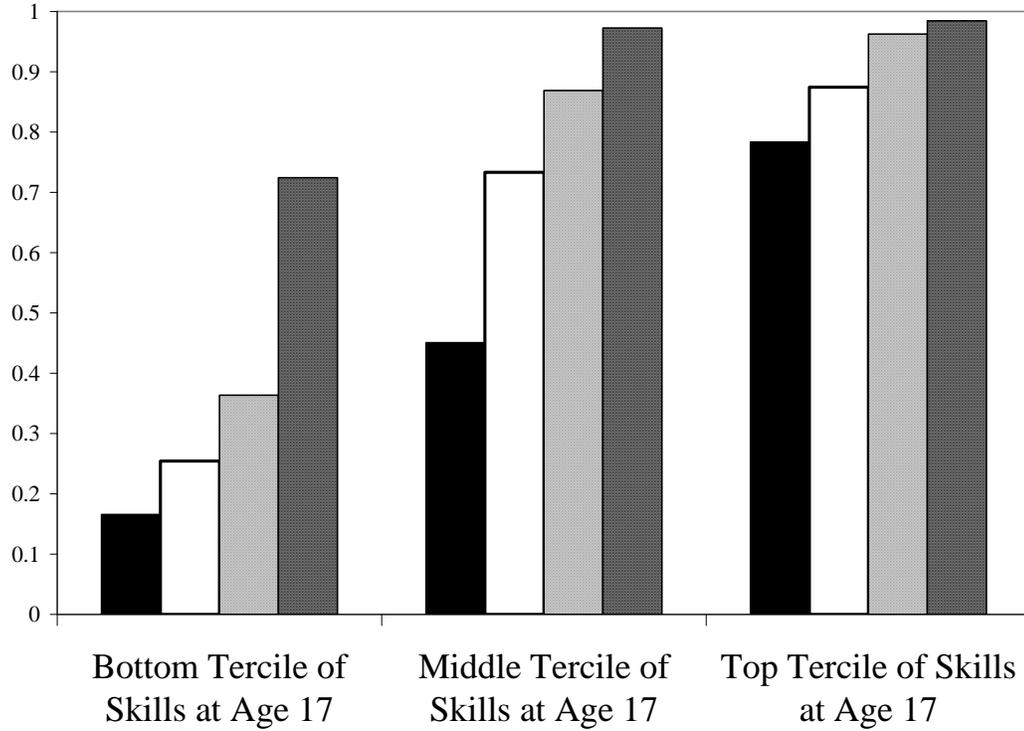
Figure 1: Percentage Enrolled in College, Model Prediction.

This figure is the output of the simulation of the model's steady state. Cognitive skills, θ, is measured by the skills at age 14. Family income, y, is the parental income when the child is 17 years old. Let T_k denote the terciles of θ, $k = 1, 2, 3$. Let Q_j denote the quartiles of parent's income, y, $j = 1, 2, 3, 4$. Let $S = 1$ if the NLSY/79 respondent enrolls in college and $S = 0$ otherwise. In this figure, Carneiro and Heckman (2003) plot the fraction of people in skill tercile k and family income quartile j who enroll in college, $P_{k,j}$, $P_{k,j} = \sum \frac{1(S=1, \theta \in T_k, y \in Q_j)}{1(\theta \in T_k, y \in Q_j)}$.
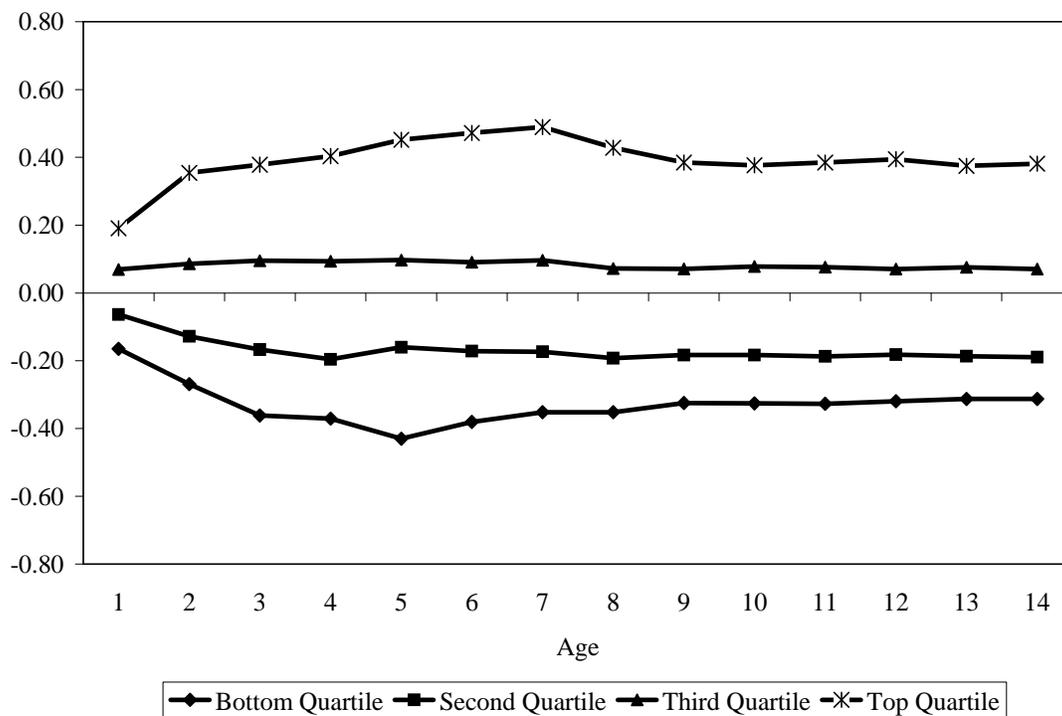
Figure 2: Mean Standardized Skills at Each Age, by Quartile of Permanent Family Income.

At every age t let θ_t denote skills when the child is t years old. Let μ_t, σ_t denote the mean and standard deviation of θ_t. Define $z_t = \frac{\theta_t - \mu_t}{\sigma_t}$.

Let r denote the steady-state equilibrium interest rate, $(r = 4.48\%)$. Let y_t denote the parent's income when the child is t years old. Let Y denote permanent income. It is defined as $Y = \sum_{t=0}^{T} \frac{y_t}{(1+r)^t}$. In this figure, I plot the mean z_t by quartiles of Y.
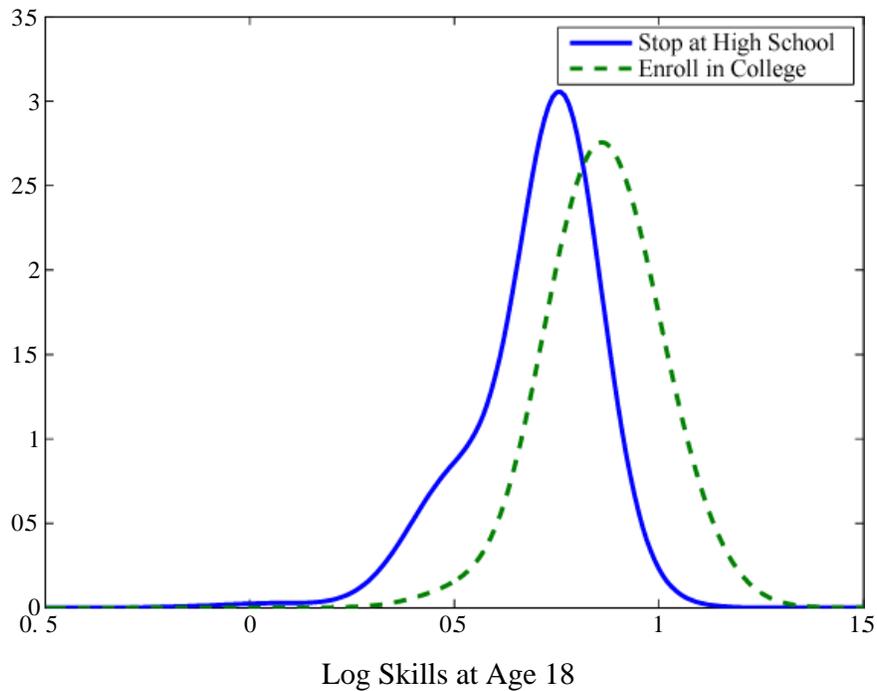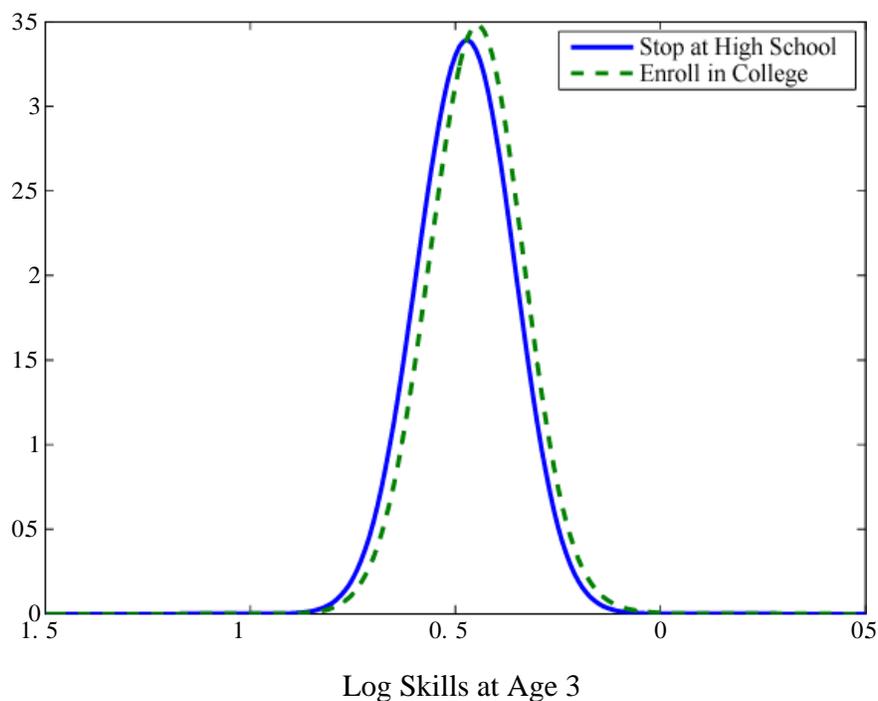
Figure 3: Densities of Log Skills at Age 18 by School Choice.

During the last period of childhood, parents decide whether or not to enroll their child in college. Let S take the value zero if the parents decide to stop the education of the child at the end of high school. Let $S = 1$ if the child is enrolled in college. Let θ_{18} denote the skills at age 18 of the child. In this figure, the solid curve shows the density of θ_{18}, conditional on the sample of those who stop their education at the high school level, $f(\theta_{18}|S=0)$. The dashed curve shows the density of θ_{18}, conditional on the sample who enrolls in college, $f(\theta_{18}|S=1)$.

Figure 4: Densities of Log Skills at Age 3 by School Choice.

During the last period of childhood, parents decide whether or not to enroll their child in college. Let S take the value zero if the parent decides to stop the child's education at the end of high school. Let $S = 1$ if the child enrolls in college. Let θ_3 denote the skills at age 3 of the child. In this figure, the solid curve shows the density of θ_3 conditional on the sample of those who stop at high school, $f(\theta_3 | S = 0)$. The dashed curve shows the density of θ_3 conditional on the sample of those who enroll in college, $f(\theta_3 | S = 1)$.
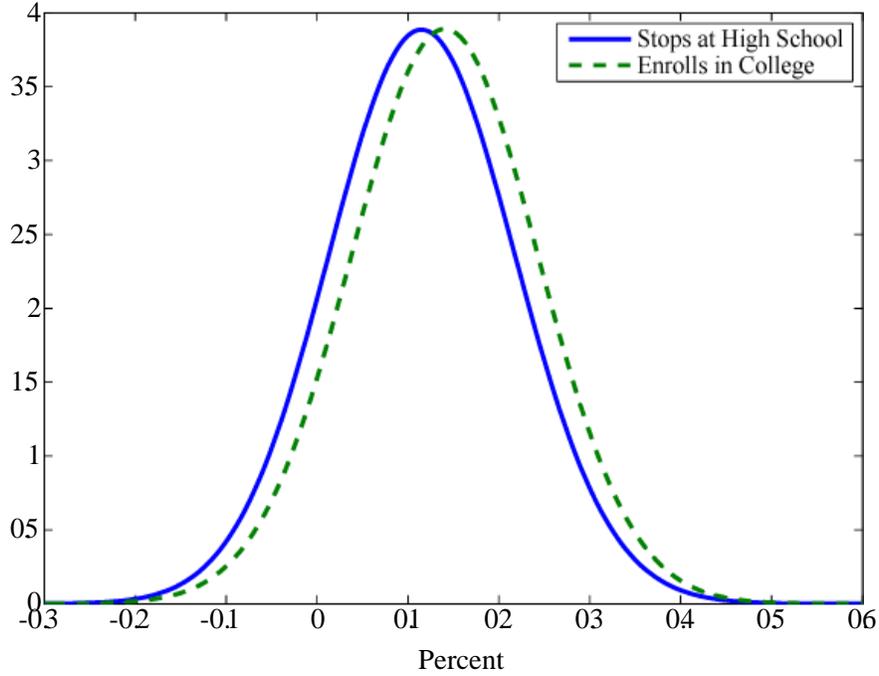
Figure 5: Densities of Returns for Each Year of College, by School Choice.

During the last period of childhood, parents decide whether or not to enroll their child in college. Let S take the value zero if the parents decide to stop the education of the child at the end of high school. Let $S = 1$ if the child is enrolled in college. Let $y_{s,t}$ denote the income of an adult at schooling level s. Define $Y_s = \sum_{t=T+1}^{2T} \frac{y_{s,t}}{(1+r)^{t-(T+1)}}$. The returns to one year of college, R, are defined as $R = \frac{Y_1 - Y_0}{Y_0}$. In this Figure, I plot the distribution of R for the subsamples $S = 0$ (the solid curve) and $S = 1$ (the dashed curve).
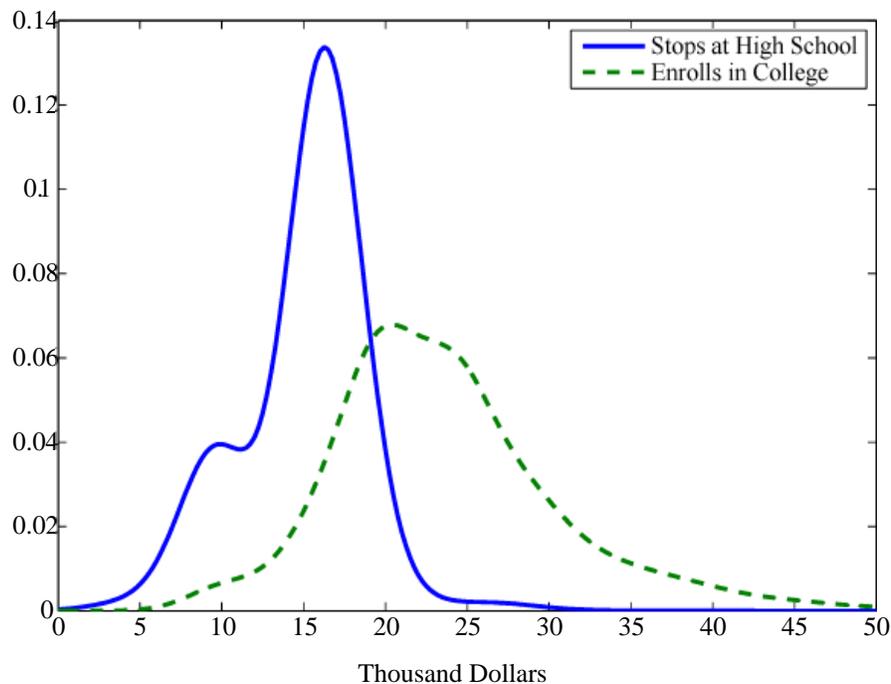
Figure 6: Densities of Present Value of Investments, by School Choice.

Let x_t denote the parental investments in skills when the child is t years old, $t = 1, ..., 18$. The price of each investment good is w. Let X denote the present value of total investments (measured in dollars), discounted according to the steady-state equilibrium interest rate r, $X = \sum_{t=1}^{18} \frac{wx_t}{(1+r)^{t-1}}$. Let S take the value zero if the child stops at high school and $S = 1$ if the child enrolls in college. The solid curve shows the density of X for high school, $f(X|S=0)$. The dashed curve shows the density of X for college, $f(X|S=1)$.
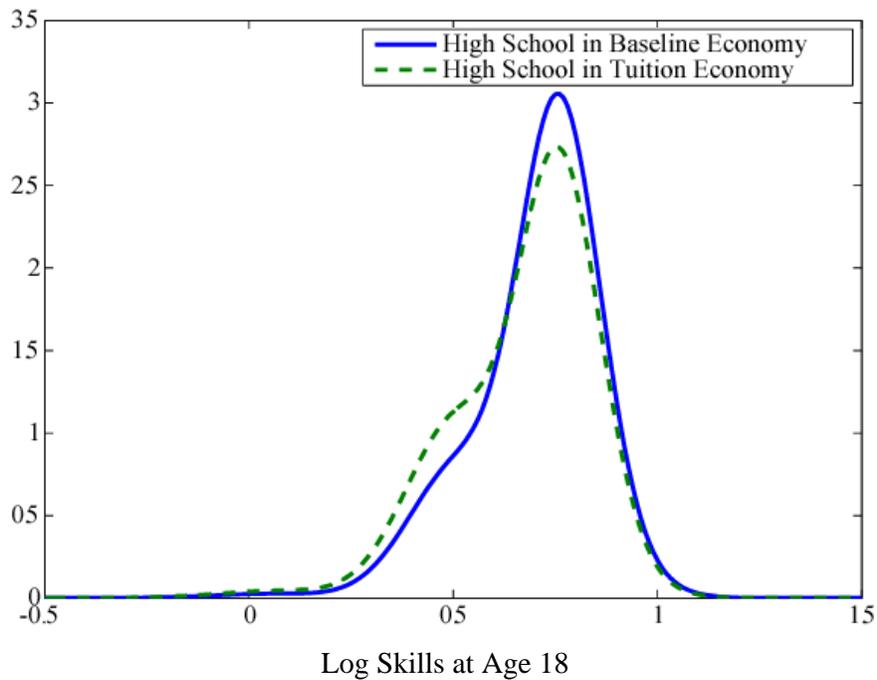
Figure 7: Densities of High School Log Skills at Age 18: Baseline versus Tuition Economy.

Let θ denote skills when the child is 18 years old. Let S take the value zero if the parent stops the child's education right after high school and $S = 1$ if the child goes on to college. I compare selection into the high-school sector in the baseline against an economy with a 50% tuition subsidy (roughly U\$2,000). In this figure, the solid curve shows the density of θ for high school in the baseline economy, $f\left(\theta\middle|\, S = 0, \text{baseline}\right)$. The dashed curve shows the density of θ for high school in the 50% tuition subsidy economy, $f\left(\theta\middle|\, S = 0, \text{tuition}\right)$.
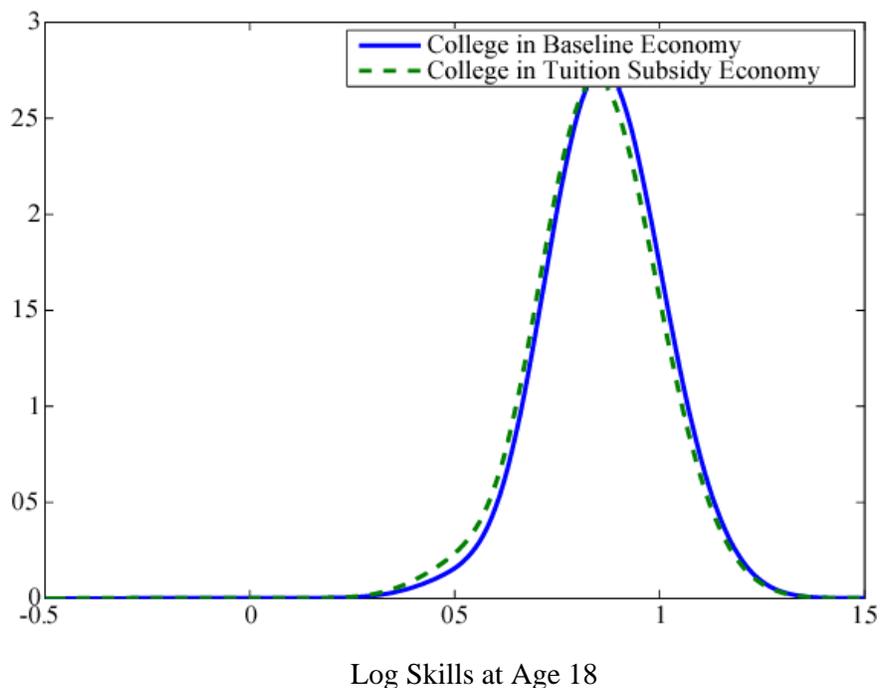
Figure 8: Densities of College Log Skills at Age 18: Baseline versus Tuition Economy.

Let θ denote skills when the child is 18 years old. Let S take the value zero if the parent stops the child's education in high school and $S = 1$ in enrolled in college. I compare selection into the enrollment in college in the baseline against an economy with a 50% tuition subsidy (roughly U\$2,000). In this figure, the solid curve shows the density of θ for college in the baseline economy, $f(\theta|S = 1, \text{baseline})$. The dashed curve shows the density of θ for college with a 50% subsidy on tuition, $f(\theta_{18}|S = 1, \text{tuition})$.
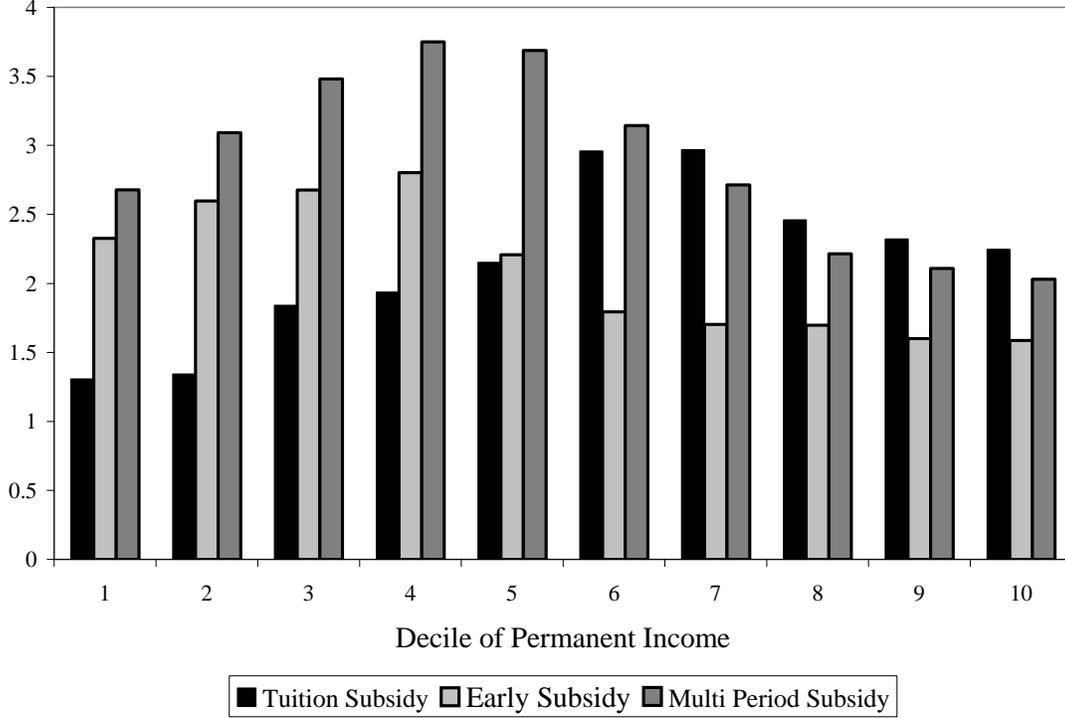
Figure 9: Mean Present Value of the Equivalent Variation.

Let c_t^B and c_t^k denote the consumption when a child is t years old in the baseline economy and an alternative economy. Define the ex-post lifetime utility of a household in the baseline economy as V^B where $V^B = \sum_{t=1}^{T} \beta^{t-1} u\left(c_t^B\right) + \beta^T E\left[V^{B\prime}\right]$ where $V^{B\prime}$ is the lifetime utility of the descentent. For any other economy k let V^k denote the ex-post lifetime utility of the household. The equivalent variation is the constant dollar amount λ^k the amount that the household must receive in order to be just as well in the baseline economy as in alternative economy k: $V^B = V^k$. Let r denote the steady-state equilibrium interest rate in the baseline economy. Let $\Lambda = \sum_{t=1}^{T} \frac{\lambda}{(1+r)^{t-1}}$. The figure above shows mean Λ by deciles of lifetime earnings.
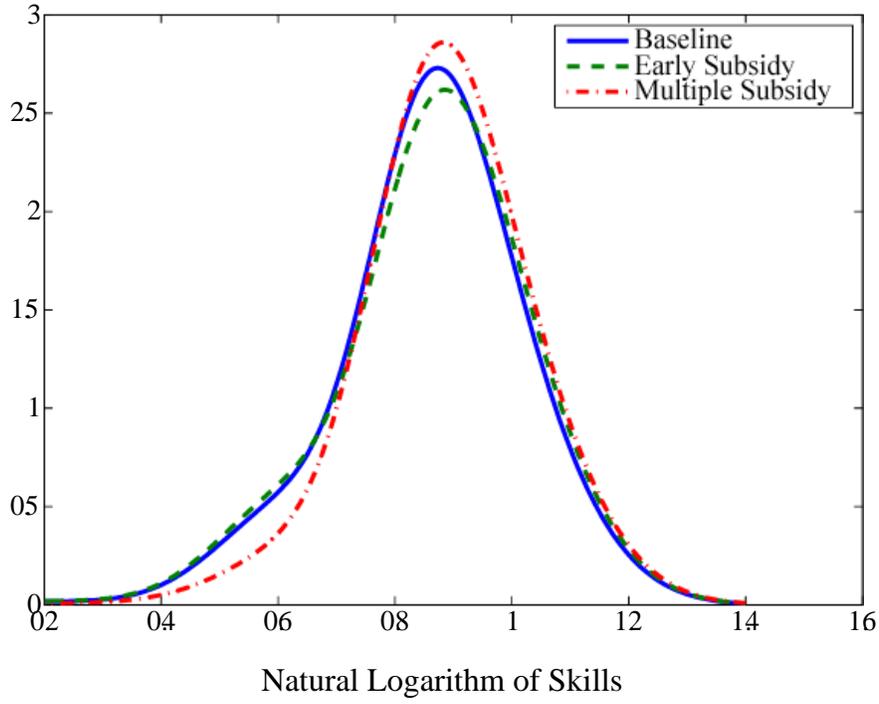
Figure 10: Stationary Densities of Adult Log Skills: Comparison Across Different Economies.

Let h^B, h^E, h^M denote the parental cognitive skills level in the baseline economy, the economy with an early subsidy, and the economy with the multiple period subsidy, respectively. Let g_k denote the stationary distribution of the natural logarithm of parental cognitive skills, $\ln h^k$, for $k = B, E, M$. In this Figure, I plot the densities $g_B \left(\ln h^B \right)$ against $g_E \left(\ln h^E \right)$ and $g_M \left(\ln h^M \right)$.

Parameter	Parameter Description	Stage 1	Stage 2	Stage 3
γ_{1,n_t}	Current Skill, θ_t	0.7353**	0.7989**	0.8723**
γ_{2,n_t}	Investment, x_t	0.1237*	0.0773*	0.0803*
γ_{3,n_t}	Mother's Skills, h	0.1410**	0.1237*	0.0473*
γ_{4,n_t}	Child Heterogeneity, π_c	0.1423*	0.1476*	0.1497*
γ_{5,n_t}	Family Heterogeneity, π_p	0.1642	0.1632	0.1671
ρ_{n_t}	Scale Parameter	0.6448**	0.8857**	0.9201**
ϕ_{n_t}	Complementarity Parameter	−0.2087	−0.1090	−0.1283
$\sigma^2_{n_t}$	Variance of η^θ_t	0.3456*	0.2637*	0.1020*

*Statistically signitifant at 5%, **Statistically signitifant at 1%

Table 1: CES technology, anchored cognitive skill factors, and unobserved heterogeneity.

In this table I show the estimated coefficients on the CES specification for three developmental stages n_t:

$$
\ln\theta_{t+1} = \frac{1}{\alpha_1^{\pi_c}}\frac{\rho_{n_t}}{\phi_{n_t}}\ln\left[\gamma_{1,n_t}e^{\phi_{n_t}\alpha_1^{\pi_c}\ln\theta_t} + \gamma_{2,n_t}e^{\phi_{n_t}\ln x_t} + \gamma_{3,n_t}e^{\phi_{n_t}\alpha_1^{\pi_p}\ln h}\right] +
$$
$$
+ \frac{1}{\alpha_1^{\pi_c}}\gamma_{4,n_t}\pi_c + \frac{1}{\alpha_1^{\pi_c}}\gamma_{5,n_t}\pi_p + \eta^\theta_t
$$

The first developmental stage starts at age zero and finishes at age four. The second developmental stage starts at age five and finishes at age nine. The third developmental stage starts at age ten and finishes at age fourteen.

The anchor for θ_t is the natural logarithm of labor income. I use labor income for the children who are at least 22 years old with nonzero labor income. The anchor for h is the natural logarithm of labor income when the parents are 30 years old.

	Stop at High School	Enroll in College	Income Tax Rate	Equilibrium Interest Rate
CNLSY/1979	51.06	48.94		
Baseline Economy	51.43	48.57		4.48%
50% Tuition Subsidy	38.74	61.26	1.189%	4.66%
Early Subsidy Program	37.36	62.64	0.831%	4.68%
Multiperiod Subsidy Program	38.88	61.12	0.517%	4.99%

Table 2: Equilibrium Schooling Choices, Interest Rate, and Tax Rate.

In this table, I show the percentage of individuals who stop studying at the high school level or continue to college in the white male sample from the CNLSY/79 (first row) and compare it to the model's prediction in the baseline economy (second row), an economy with a 50% tuition subsidy (third row), an economy with a subsidy to parental investments for disadvantaged children when they are three and four years old (fourth row), and a multiple period subsidy (last row). For each economy, I show the equilibrium interest rate (third column) and the equilibrium income tax rate (fourth column).

	Baseline	Tuition	Early Subsidy Economy	Multiple Period Subsidy
Stop at High School	79.41	71.61	68.19	64.72
Enroll in College	20.59	28.39	31.81	35.28

Table 3: Schooling Choices of Disadvantaged Children.

In this table, I show the percentage of individuals who stop their studies at the high school level or continue to college for the children who are part of the disadvantaged children sample. A disadvantaged child is one who is in the bottom quartile of the distribution of cognitive skills at age three and the parent's total resources at age three of the child does not exceed U\$40,000.

References

[1] Attanasio, O. and G. Weber (2010). Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy. *Journal of Economic Literature*, 48(4): 693–751.

[2] Berhman, J. R. and M. R. Rosenzweig (2004). Returns to Birthweight. *Review of Economics and Statistics* 86(2), 586-601.

[3] Belley, Phillipe and L. L. Lochner (2007). The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital*, 1(1), 37-89.

[4] Black, S. E., P. J. Devereux, and K. Salvanes (2007). From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. *Quarterly Journal of Economics* 122(1), 409-439.

[5] Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption Inequality and Partial Insurance. *American Economic Review*, 98(5), 1887–1921.

[6] Bradley, R. H. and B. M. Caldwell (1980). The relation of Home Environment, Cognitive Competence, and IQ among Males and Females. *Child Development* 51(4), 1140-1148.

[7] Bradley, R. H. and B. M. Caldwell (1984). The relation of Infant's Home Environments to Achievement Test Perfomance in First Grade: A Follow-up Study. *Child Development* 55(3), 803-809.

[8] Browning, M., L. P. Hansen, and J. J. Heckman (1999). Micro Data and General Equilibrium Models. In J. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1A, Amsterdam: Elsevier, 543-663.

[9] Carneiro, P. and R. Ginja (2012). Partial Insurance and Investments in Children. University of Uppsala Working Paper.

[10] Carneiro, P., K. Hansen, and J. J. Heckman (2003). Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economic Review* 44(2), 361-422.

[11] Carneiro, P. and J. J. Heckman (2003). Human Capital Policy. In J. J. Heckman, A. B. Krueger and B. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies?* Cambridge, MA: MIT Press.

[12] Case, A., A. Fertig, and C. Paxson (2005). The Lasting Impact of Childhood Health and Circumstance. *Journal of Health Economics* 24(2), 365-389.

[13] Caucutt, E. M. and L. J. Lochner (2012). Early and Late Human Capital Investments, Borrowing Constraints, and the Family, University of Western Ontario Working Paper.

[14] Cunha, F. and J. J. Heckman (2007). The Technology of Skill Formation. *American Economic Review,* 97(2), 18-35.

[15] Cunha, F. and J. J. Heckman (2008). Formulating, Identifying, and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43(4), 738-782.

[16] Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov (2006). Interpreting the Evidence on Life Cycle Skill Formation. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 12, 697-812. Amsterdam: North-Holland.

[17] Cunha, F., J. J. Heckman, S. Navarro (2005). Separating Heterogeneity from Uncertainty in Life Cycle Earnings. *Oxford Economic Papers* 57(2), 191-261.

[18] Cunha, F., J. J. Heckman, S. Navarro (2006). Counterfactual Analysis of Inequality and Social Mobility. In S. L. Morgan, D. B. Grusky, and G. S. Fields (Eds.), *Mobility and Inequality: Frontiers of Research in Sociology and Economics,* Chapter 4, 290-348. Stanford, CA: Stanford University Press.

[19] Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, 78(3), 883-931.

[20] Currie, J. and R. Hyson (1999). Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight. *American Economic Review* 89(2), 245-250.

[21] Dahl, G. and L. Lochner (2012). The Impact of Family Income On Child Achievement: Evidence from the Earned Income Tax Credit, *American Economic Review* 102(5), 1927-56.

[22] Del Boca, D., C. Flinn, and M. Wiswall (2013). Household Choices and Child Development. *Review of Economic Studies*, forthcoming.

[23] Gillen, Andrew, Jonathan Robe, and Daniel Garrett (2011). Net Tuition and Net Price Changes, Center for College Affordability and Productivity Policy Paper, available at http://centerforcollegeaffordability.org/uploads/Net_Tuition_Trends_2011.pdf

[24] Geweke, J. (1977). The Dynamic Factor Analysis of Economic Time Series Models. In D. J. Aigner and A. S. Goldberger (Eds.), *Latent Variables in Socio-Economic Models*, Chapter 5, 365-383. Amsterdam: North-Holland.

[25] Goldin, C. and Katz, L. F. (2008). *The Race between Education and Technology*, Belknap Press of Harvard University Press.

[26] Heathcote, J., K. Storesletten, and G. Violante (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States, *Journal of Political Economy* 118(4), 681-722.

[27] Heckman, J. J., A. B. Krueger, and B. M. Friedman (2003). *Inequality in America: What Role for Human Capital Policies?* Cambridge, MA: MIT Press.

[28] Heckman, J. J., L. J. Lochner, and C. Taber (1998). Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents. *Review of Economic Dynamics* 1(1), 1-58.

[29] Heckman, J. J., J. Stixrud, and S. Urzua (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics* 24(3), 411-482.

[30] Julier, S. J. and J. K. Uhlmann (1997). A New Extension of the Kalman Filter to Nonlinear Systems. *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defence Sensing, Simulation and Controls.*

[31] Kane, Thomas (1995). Rising Public College Tuition and College Entry: How Well Do Public Subsidies Promote Access to College? NBER Working Paper 5164.

[32] Kramer, M. S., L. Goulet, J. Lydon, L. Séguin, H. McNamara, C. Dassa, R. W. Platt, M. F. Chen, H. Gauthier, J. Genest, S. Kahn, M. Libman, R. Rozen, A. Masse, L. Miner, G. Asselin, A. Benjamin, J. Klein, and G. Koren (2001). Socio-economic Disparities in Preterm Birth: Causal Pathways and Mechanisms. *Paediatric and Perinatal Epidemiology* 15 (Supplement 2), 104-123.

[33] Laitner, J. (1992). Random Earnings Differences, Lifetime Liquidity Constraints, and Altruistic Intergenerational Transfers. *Journal of Economic Theory* 58(2), 135-170.

[34] Lino, M. (2000). Expenditures on Children by Families, 2000 Annual Report. Miscellaneous Publication 1528-2000, U. S. Department of Agriculture, Center for Nutrition Policy and Promotion.

[35] Linver, M. R., J. Brooks-Gunn, and N. Cabrera (2004). The Home Observation for Measurement of the Environment (HOME) Inventory: The Derivation of Conceptually Designed Subscales. *Parenting: Science & Practice* 4(2/3), 99-114.

[36] MaCurdy, T. E. (1982). The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis. *Journal of Econometrics*, 18(1): 83–114.

[37] Meghir, C. and L. Pistaferri (2004). Income Variance Dynamics and Heterogeneity. *Econometrica*, 72(1), 1–32.

[38] Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. Sean F. Reardon. In R. Murnane and G. Duncan (Eds.), *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*, New York: Russell Sage Foundation Press.

[39] Sargent, T. J. and C. A. Sims (1977). Business Cycle Modeling Without Much a Priori Economic Theory. In *New Methods in Business Cycle Research: Proceedings from a Conference.* Minneapolis, MN: Federal Reserve Bank of Minneapolis.

[40] Schennach, S. M. (2004). Estimation of Nonlinear Models with Measurement Error. *Econometrica* 72(1), 33-75.

[41] Shumway, R. H. and D. S. Stoffer (1982). An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm. *Journal of Time Series Analysis* 3(3), 253-264.

[42] Stigler, G. J. (1970). Director's Law of Public Income Redistribution. *Journal of Law and Economics* 13(1), 1-10.

[43] Todd, P. E. and K. I. Wolpin (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, 1(1), 91-136.

[44] Watson, M. W. and R. F. Engle (1983). Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC, and Varying Coefficient Regression Models, *Journal of Econometrics* 23(3), 385-400.