

2015

# 2015-3 Correlation, Consumption, Confusion, or Constraints: Why Do Poor Children Perform So Poorly?

Elizabeth M. Caucutt

Lance J. Lochner

Youngmin Park

Follow this and additional works at: <https://ir.lib.uwo.ca/economicscibc>

 Part of the [Economics Commons](#)

---

## Citation of this paper:

Caucutt, Elizabeth M., Lance J. Lochner, Youngmin Park. "Correlation, Consumption, Confusion, or Constraints: Why Do Poor Children Perform So Poorly?." CIBC Centre for Human Capital and Productivity. CIBC Working Papers, 2015-3. London, ON: Department of Economics, University of Western Ontario (2015).

**Correlation, Consumption, Confusion, or  
Constraints: Why Do Poor Children  
Perform So Poorly?**

by

**Elizabeth M. Caucutt, Lance Lochner and Youngmin Park**

**Working Paper # 2015-3**

**March 2015**



***CIBC Working Paper Series***

Department of Economics  
Social Science Centre  
Western University  
London, Ontario, N6A 5C2  
Canada

This working paper is available as a downloadable pdf file on our website  
<http://economics.uwo.ca/centres/cibc/>

# Correlation, Consumption, Confusion, or Constraints: Why do Poor Children Perform so Poorly?\*

Elizabeth M. Caucutt  
University of Western Ontario,

Lance Lochner  
University of Western Ontario  
NBER

Youngmin Park  
University of Western Ontario

March 5, 2015

## Abstract

The economic and social mobility of a generation may be largely determined by the time it enters school given early developing and persistent gaps in child achievement by family income and the importance of adolescent skill levels for educational attainment and lifetime earnings. After providing new evidence of important differences in early child investments by family income, we study four leading mechanisms thought to explain these gaps: an inter-generational correlation in ability, a consumption value of investment, information frictions, and credit constraints. In order to better determine which of these mechanisms influence family investments in children, we evaluate the extent to which these mechanisms also explain other important stylized facts related to the marginal returns on investments and the effects of parental income on child investments and skills.

## 1 Introduction

Adolescent skill and achievement gaps by parental income explain a substantial share of subsequent differences in educational attainment and lifetime earnings (Cameron and Heckman, 1998; Keane and Wolpin, 1997; Carneiro and Heckman, 2002), suggesting that the economic and social mobility of a generation is largely determined by adolescence. Perhaps more troubling, sizeable differences in achievement by parental income are already evident by very young ages, persisting throughout childhood (Carneiro and Heckman, 2002; Cunha et al., 2006; Cunha and Heckman, 2007; Cunha, 2013). This raises the possibility that a generation's fate may be sealed by the time it enters school.<sup>1</sup> Altogether, this evidence suggests that a complete understanding of intergenerational mobility and its implications for economic and social policy requires convincing answers to the vexing question: Why do poor children perform so poorly?

---

\*For helpful comments, we thank participants at the 2014 HCEO Conference on Social Mobility at the University of Chicago. We also thank Eda Bozkurt and Qian Liu for excellent research assistance. Caucutt and Lochner gratefully acknowledge support from CIGI-INET Research Grants.

<sup>1</sup>Recent studies show that these early achievement and educational attainment gaps have been growing in the United States for decades (Belley and Lochner, 2007; Reardon, 2011).

Given the importance of family investments for early child development (Todd and Wolpin, 2007; Cunha and Heckman, 2008; Cunha et al., 2010; Del Boca et al., 2014; Pavan, 2014), we concentrate on understanding why low-income families invest so much less in their young children compared to higher income families (Guryan et al., 2008; Kaushal et al., 2011). We consider four broad mechanisms often thought to explain early investment and achievement gaps by family income. First, the natural ability of children and parents may be correlated (Becker and Tomes, 1979, 1986). If child achievement is an increasing function of own ability, then a positive inter-generational ability correlation can generate the income – achievement gradients documented in the literature. Second, parents may enjoy making investments in their children. If investments provide a direct benefit to parents above and beyond the future labor market returns to children, parents will choose to invest more as their income rises like they would purchase more of any other normal good. It is also possible that low- and high-income families place different intrinsic value on investments (or human capital more generally).<sup>2</sup> Third, low-income parents may be poorly informed about the productivity of investments in their children (Cunha et al., 2013; Cunha, 2014; Dizon-Ross, 2014). For example, poor parents may incorrectly believe that investments in their young children are unproductive, or they may simply face greater uncertainty in the productivity of or returns to investments. Alternatively, poor parents may recognize the importance of investing in their children, but they may not know which types of investment activities/goods are most productive. Fourth, poor families may be unable to invest efficiently in their children due to limits on their capacity to borrow against their own future income or against the potentially high returns on investments in their children (Becker and Tomes, 1979, 1986; Caucutt and Lochner, 2006, 2012; Cunha et al., 2006; Cunha and Heckman, 2007; Cunha, 2013; Lee and Seshadri, 2014).

We use a simple framework of dynamic human capital investment to formally examine whether these mechanisms are also able to account for other important stylized facts in the literature on child development: (i) the high marginal returns to early investments in economically disadvantaged children, (ii) lower returns on marginal investments in higher income children, (iii) exogenous increases in family income lead to greater investments in children and improved childhood outcomes, and (iv) the impacts of income on child investments, achievement and educational attainment are greater if the income is earned (or received) when children are young. While our analysis does not necessarily reveal which mechanism is most important for explaining income-based achievement gaps, it is useful for helping understand which mechanisms are needed to provide a complete picture of the child development process and the role of family income.<sup>3</sup> This

---

<sup>2</sup>See Lazear (1977) for an early treatment of schooling choices that incorporates a ‘consumption’ value. More recently, Keane and Wolpin (2001), Cunha et al. (2005), and Carneiro et al. (2011) emphasize and estimate the role of heterogeneity in the ‘consumption’ value of schooling for explaining schooling behavior.

<sup>3</sup>See Cunha (2014) for a novel effort to empirically decompose the relative importance of a similar set of

is important, because the different mechanisms can have very different policy implications. For example, if investment and achievement gaps are driven by intergenerational ability correlations or a ‘consumption’ value of investment, then investments in children are likely to be economically efficient and policies designed to improve equity will be inefficient. By contrast, either information-based or credit market frictions can lead to inefficiently low investments in economically disadvantaged children. In this case, it may be possible to improve both equity and efficiency through well-designed policies.

We organize this paper in the following way. In Section 2, we briefly document differences in child achievement and investment levels by family income using data from the Children of the National Longitudinal Survey of Youth (CNLSY). Then, we summarize evidence on other stylized facts from the literature on child development in Section 3. In Section 4, which is the main contribution of this paper, we develop a unified framework of dynamic skill investment that incorporates all four potential mechanisms commonly thought to drive investment and achievement gaps by family income. We use this model to formally examine whether the explanations are also consistent with the stylized facts in Section 3 as well as other evidence in the child development literature. In Section 5, we conclude with a summary of our main results and their implications for future research.

## 2 Child Achievement and Investment Gaps by Family Income

In this section, we document differences in child achievement and investment behavior by family income using data from the CNLSY. A longitudinal survey that links mothers with their children, the CNLSY contains excellent measures of family background and income (starting in 1979), as well as biennial measures of child math and reading achievement and family investments in children (beginning in 1986).<sup>4</sup>

We use background measures of maternal education, race/ethnicity, and “ability” as measured by the Armed Forces Qualifying Test (AFQT).<sup>5</sup> As a medium-run measure of family income, we average all available reports of earnings by the mother and her spouse (if married) from the child’s birth through ages 6 or 7 (depending on which of these ages achievement and investments in children were measured).<sup>6</sup> Child achievement is measured by the Peabody Individual Achieve-

---

mechanisms using unique data on parental perceptions and stated choices about investments in children under different hypothetical budget sets. While we do not empirically evaluate the relative importance of different mechanisms, our theoretical analysis is based on a more general dynamic human capital investment model. We consider a wide range of information frictions and explicitly model intertemporal borrowing constraints.

<sup>4</sup>Our analysis is based on children born to mothers from the random sample.

<sup>5</sup>(Nearly) all CNLSY mothers, born between 1957 and 1964, took the AFQT in 1980 as part of the survey. The AFQT tests basic math and verbal/reading skills.

<sup>6</sup>Before averaging across time, we discount all income back to the child’s birth year using a 5% discount rate, so

ment Tests (PIAT) in math, reading recognition and reading comprehension; these measures are standardized to have a mean of zero and standard deviation of one at each age. A number of child investment activities/inputs are also reported in the CNLSY as we discuss below.

Figure 1 documents sizeable differences in ages 6-7 math and reading achievement by family income quartile. The light (beige) bars represent raw differences in achievement between the reported parental income quartile and the bottom income quartile, while the dark (red) bars report differences after controlling flexibly for maternal race/ethnicity, AFQT, and educational attainment.<sup>7</sup> Raw gaps by income are sizeable: math and reading scores of children with parents in the highest income quartile are all more than half of a standard deviation higher than those with parents in the lowest income quartile. Controlling for other important maternal characteristics substantially reduces these gaps (by as much as three-quarters), but does not eliminate them – parental income still has economically (and statistically) significant effects on child achievement.

Figures 2 and 3 document a number of early childhood family investment measures by parental income at different ages. For all measures except ‘eat with mom and dad daily’ (ages 0-1, 2-3, 4-5) and ‘family meets friends/relatives two or more times per month’ (ages 6-7), investments are monotonically increasing in parental income. For a number of measures, the differences are substantial. For example, mothers of young children from the highest income quartile are over 50% more likely to read to their child three or more times per week compared to mothers from the lowest income quartile. High income mothers with children ages 0-1 are more than twice as likely to have 10 or more books in the home. Among children ages 6-7, those from high income families are more than twice as likely to be enrolled in special lessons or extracurricular activities.

One interpretation of the investment measures reported in Figures 2 and 3 is that they represent different types of investment inputs that influence child development. An alternative interpretation is that they all represent noisy measures of a single underlying ‘investment’. Under the latter interpretation, factor analysis can be used to uncover a more precise measure of the latent investment (Cunha and Heckman, 2008; Cunha et al., 2010).<sup>8</sup> Based on this insight,

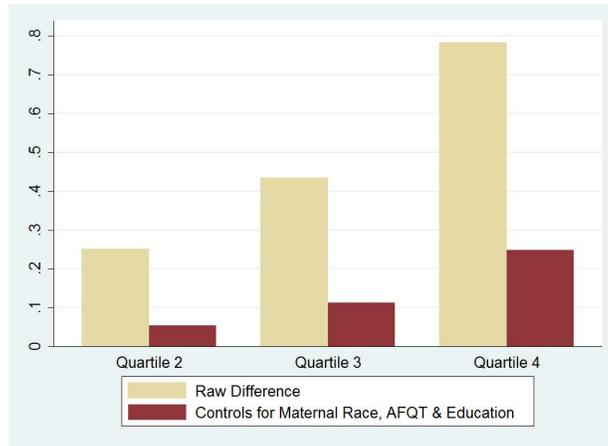
---

our earnings measure reflects average discounted family earnings from the child’s birth through ages 6-7. Individuals are dropped if fewer than 3 income reports are available.

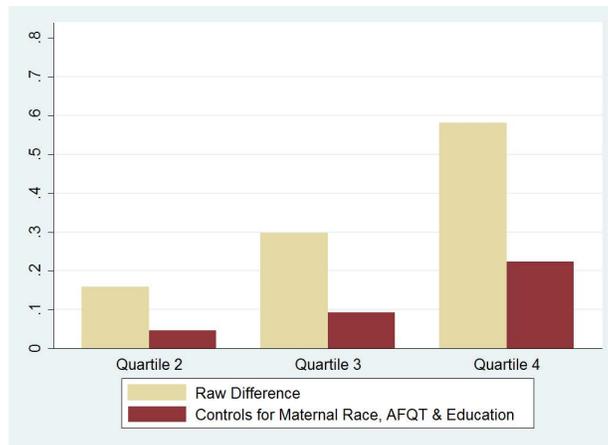
<sup>7</sup>Standard multivariate regressions are used to control for maternal race/ethnicity (white, black, hispanic), AFQT quartiles, and educational attainment (high school dropout, high school graduate, some college, completed college) by including three-way interactions of all three sets of indicators along with indicator variables for parental income quartile. Sample sizes for raw differences by income are 3,449 for math, 3,436 for reading recognition, and 3,267 for reading comprehension. Approximately 80 observations are dropped due to missing covariates when controlling for maternal characteristics.

<sup>8</sup>These interpretations are not necessarily mutually exclusive if families face the same relative prices and productivity of inputs (assuming families maintain correct beliefs about the relative productivity of inputs). In this case, all inputs (in a given period) will be proportional to total child investment expenditure (that period), with all families choosing the same proportional mix. One can then think about the various reported measures as noisy measures of that specific input or of total investment expenditure (multiplied by the factor share for that input). We study the link between multiple early inputs and total early investment expenditures in Section 4.4.3 along

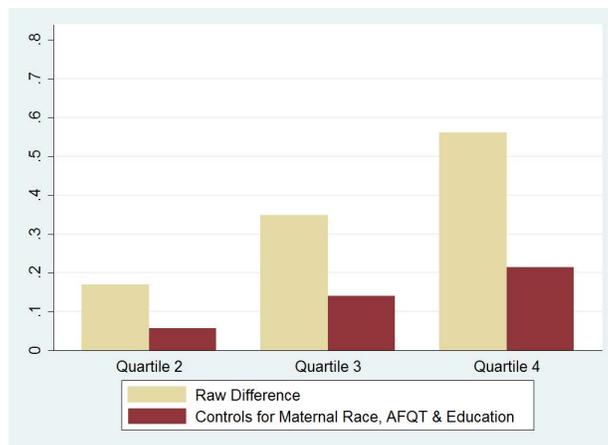
Figure 1: Ages 6-7 Achievement Gaps by Parental Income Quartile (Relative to Quartile 1)



(a) Math Achievement

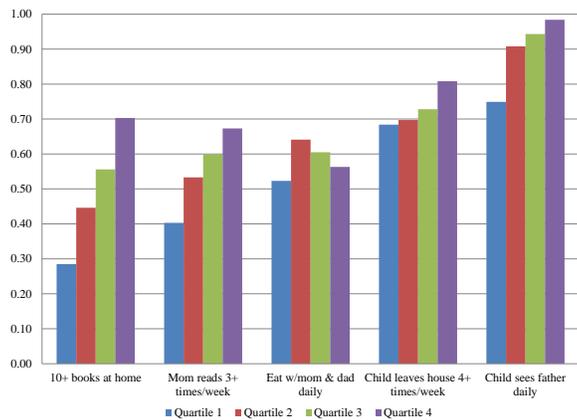


(b) Reading Recognition Achievement

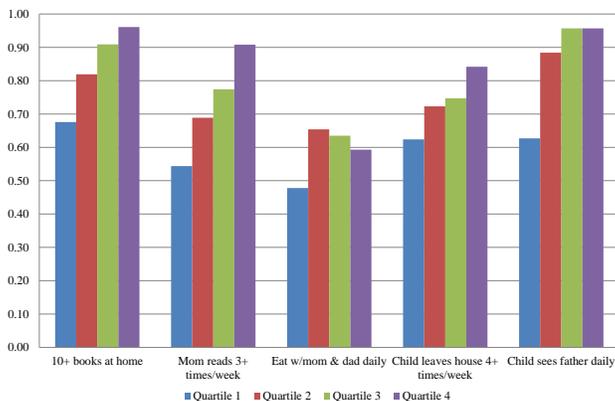


(c) Reading Comprehension Achievement

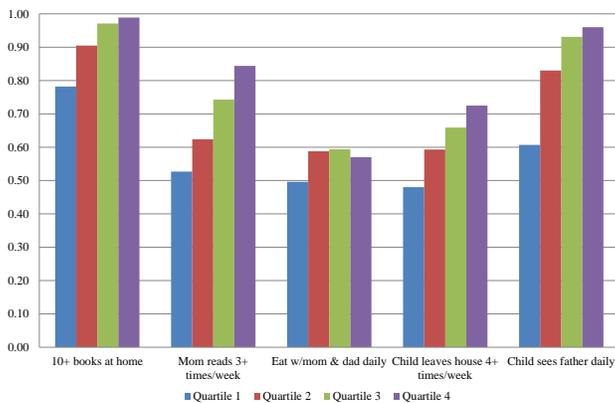
Figure 2: Family Investments in Children Ages 0-1, 2-3, and 4-5 by Parental Income Quartile



(a) Child Ages 0-1

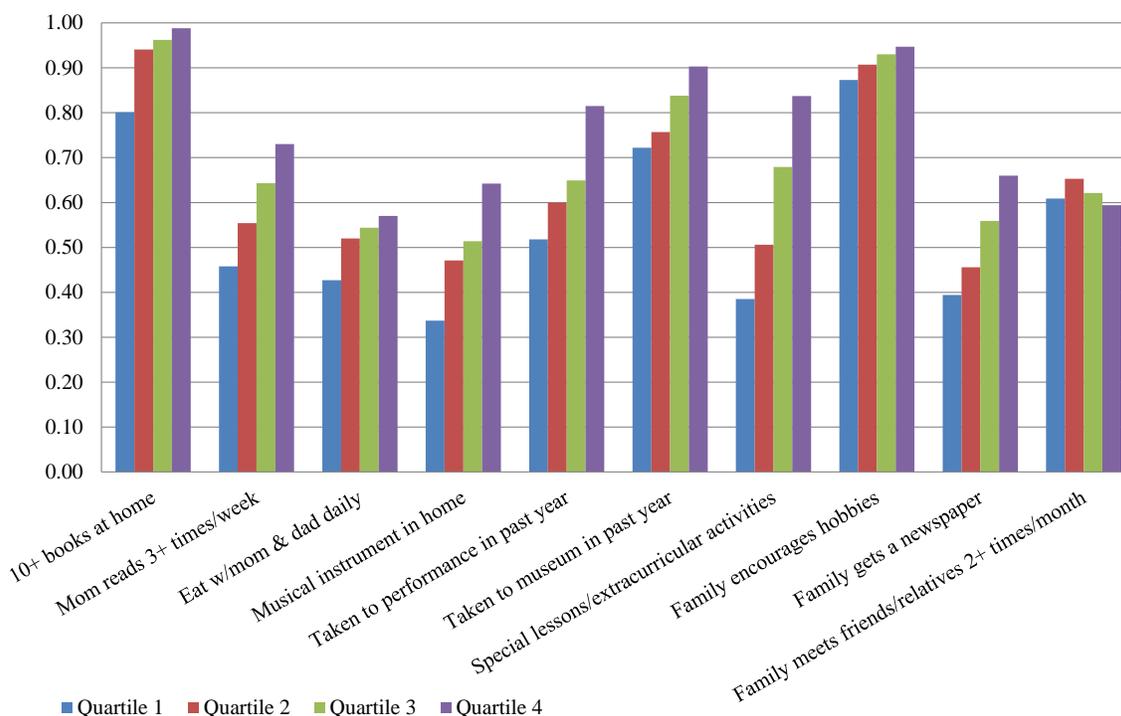


(b) Child Ages 2-3



(c) Child Ages 4-5

Figure 3: Family Investments in Children Ages 6-7 by Parental Income Quartile



we employ principal factor analysis using the measured inputs reported in Figures 2 and 3 to create age-specific predicted investment factor scores for each child. (See Appendix A.) For interpretation purposes, we normalize scores to have a mean of zero and standard deviation of one, plotting average scores by age and parental income quartile in Figure 4. The figure reveals sizeable differences in investment factor scores by parental income that are already evident at very young ages. Investments are roughly a full standard deviation higher among children from high income families relative to low income families. Figure 5 shows that these gaps shrink by as much as 50% but remain sizeable when controlling for maternal race, AFQT, and educational attainment.<sup>9</sup> Comparing Figures 2 and 3 with Figure 5 reveals that maternal characteristics explain a greater share of the income-based gaps in achievement than in investments.

### 3 Additional Stylized Facts on Child Development

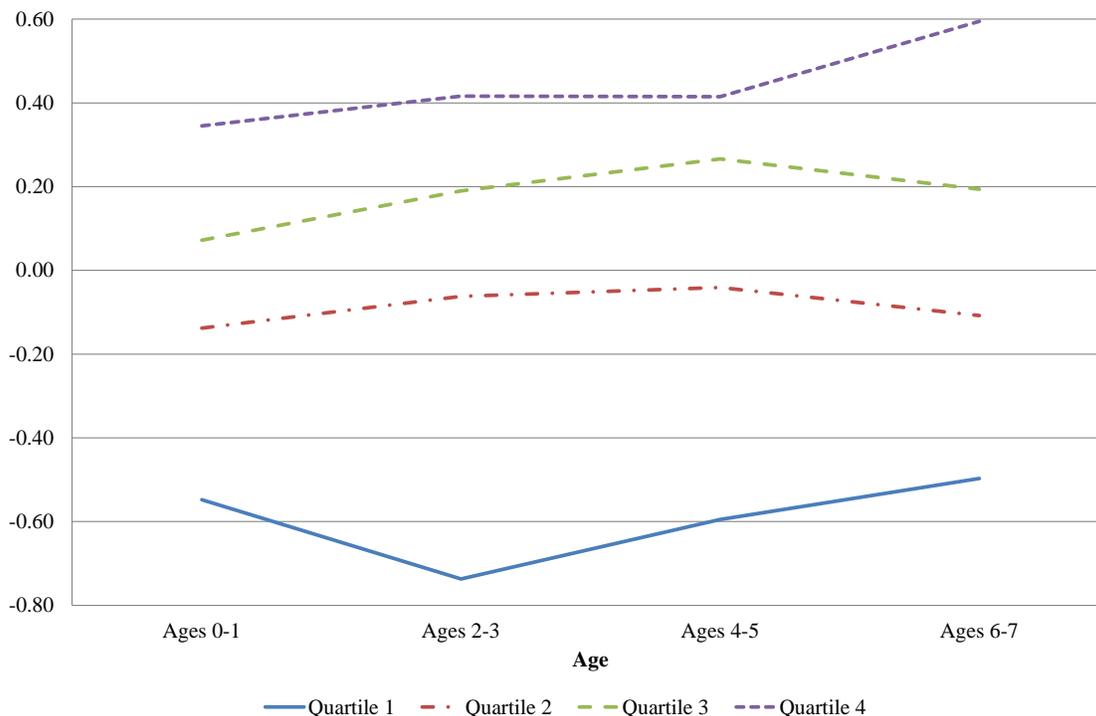
In this section, we discuss four important stylized facts about child development.

Fact 1: *the returns to early marginal investments are higher than the return to savings for*

with the consequences of family mis-perceptions about relative and overall early input productivity levels.

<sup>9</sup>Sample sizes for raw income differences are 2,324 (ages 0-1), 2,749 (ages 2-3), 2,813 (ages 4-5), and 3,493 (ages 6-7). When controlling for maternal characteristics, between 50 and 80 observations are dropped due to missing covariates.

Figure 4: Family Investment Factor Scores by Child Age and Parental Income Quartile



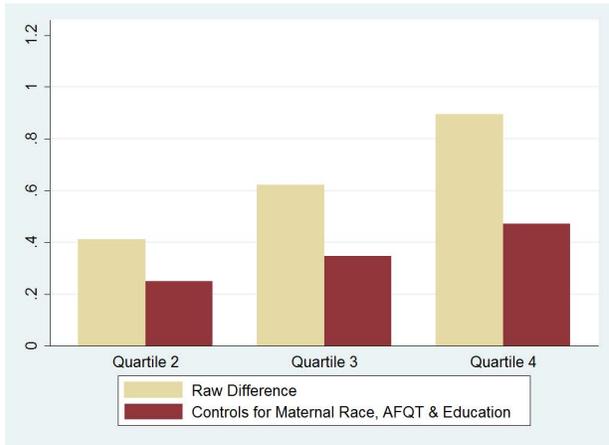
*economically disadvantaged children.* A number of comprehensive surveys (Karoly et al., 1998; Blau and Currie, 2006; Cunha et al., 2006; Almond and Currie, 2011; Heckman and Kautz, 2014; Kautz et al., 2014) document both the short- and long-term impacts of numerous early childhood interventions in the U.S.<sup>10</sup> Most notably, an experimental evaluation of the Perry Preschool program followed participants in the early 1960s through age 40, measuring the program’s impacts on a wide-array of outcomes, including cognitive achievement, educational attainment, earnings, and crime.<sup>11</sup> Based on program costs and impacts measured from ages 15-40, Heckman et al. (2010) estimate a private internal rate of return to Perry Preschool participants of around 8%.<sup>12</sup> The Abecedarian Project offered high quality full-day preschool to a (randomly assigned) sample of mostly African American children born in the mid-1970s who were at risk for delayed intellectual and social development. Follow-up evaluations of Abecedarian through age 21 reveal significant long-term benefits that exceed the program’s costs by a factor of roughly 2.5 (Barnett and

<sup>10</sup>Gertler et al. (2014) analyze the long-run impacts of a Jamaican early childhood intervention that aimed to develop cognitive, language and psychosocial skills. Their estimates suggest that the program significantly increased educational attainment by 0.6 years and earnings by 25% at age 22.

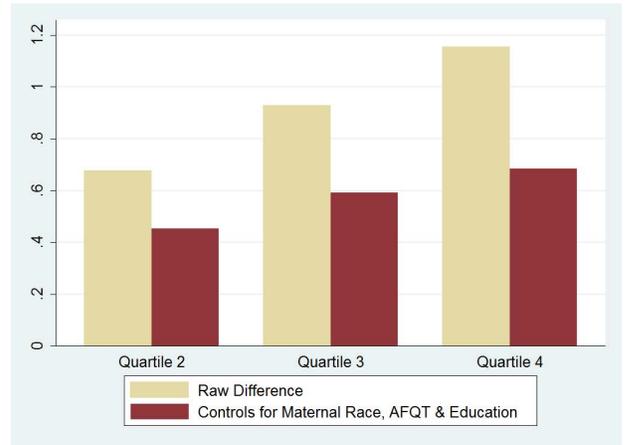
<sup>11</sup>Perry Preschool provided daily high quality preschool (2.5 hours per day) and weekly home visits for two years to children ages 3 and 4. The randomized control trial sample was drawn from low IQ children from families of low socioeconomic status.

<sup>12</sup>Social returns are even higher, largely due to benefits from crime reduction.

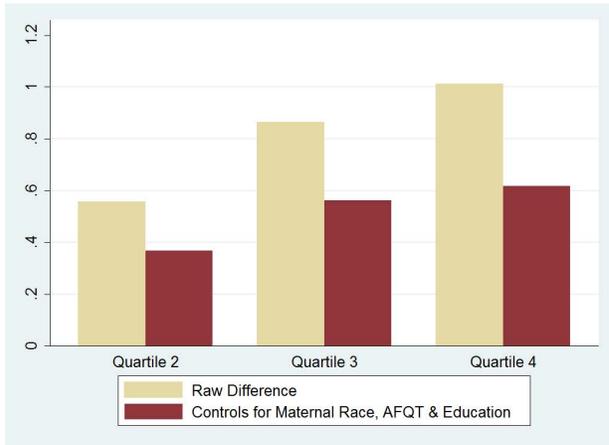
Figure 5: Family Investment Gaps by Parental Income Quartile (Relative to Quartile 1)



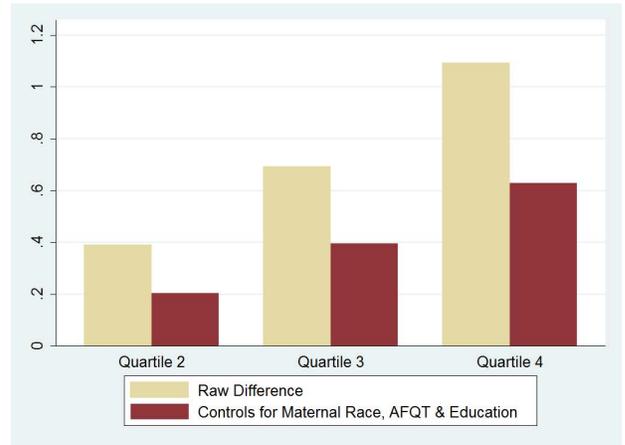
(a) Child Ages 0-1



(b) Child Ages 2-3



(c) Child Ages 4-5



(d) Child Ages 6-7

Masse, 2007).<sup>13</sup> Researchers have also extensively analyzed the long-term impacts of Chicago’s Child-Parent Center (CPC) preschool program, following a sample of low-income, mostly African American participants from the mid-1980s to the present. Rough calculations based on program impacts measured through age 26 suggest an average (private) benefit/cost ratio of 3.6 (Reynolds et al., 2011).<sup>14</sup> Finally, a number of studies document significant long-term impacts of Head Start (Currie and Thomas, 1995, 1999; Garces et al., 2002; Deming, 2009; Ludwig and Miller, 2007; Carneiro and Ginja, 2014); however, its full rate of return has not been systematically estimated.

*Fact 2: the returns to marginal investments are lower for more economically advantaged children.* Because most experimental and government-subsidized early childhood programs serve low-income families, less is known about the lifetime returns to early investments in children from higher income families. However, a number of studies estimate short- and medium-term impacts of early childhood interventions by family income or socioeconomic status (SES). These studies typically report greater benefits for more disadvantaged children. For example, Duncan and Sojourner (2013) estimate that the Infant Health Development Program (IHDP), which provided the Abecedarian preschool curriculum to an economically diverse sample of low-birth weight 1-2 year-olds, yielded significantly greater improvements in age five IQ for the subsample of children from low-income families relative to those from higher income families.<sup>15</sup> A recent analysis of Head Start (Puma et al., 2012) estimates significantly greater impacts on third grade cognitive and learning outcomes for children from ‘high risk’ (i.e. low SES) households relative to lower risk households. Estimated impacts of the Chicago CPC preschool program on educational attainment and earnings at age 28 are also higher for children from ‘high risk’ families (Reynolds et al., 2011). A few studies that estimate the impacts of introducing universal early child care subsidies in Canada and Norway find negligible or even adverse impacts on children from middle- and high-income families, likely due to the substitution of lower quality subsidized/free child care in place of higher quality unsubsidized family or informal care (Baker et al., 2008; Havnes and Mogstad, 2014; Kottelenberg and Lehrer, 2014).<sup>16</sup>

Taking a very different approach, Cunha et al. (2010) apply dynamic factor models using

---

<sup>13</sup>As is common in this literature, these calculations assume an annual real discount rate of 3%. The benefit calculations project lifetime earnings impacts based on average earnings differences by educational attainment and the significant effects of Abecedarian on educational attainment. An age 30 follow-up study (Campbell et al., 2012) estimates a sizeable but statistically insignificant increase in annual earnings for participants (\$33,000 vs. \$21,000). Estimated effects on employment rates and use of public aid at age 30 are statistically significant.

<sup>14</sup>As with the Abecedarian cost-benefit analysis of Campbell et al. (2012), Reynolds et al. (2011) use a 3% discount rate and project lifetime earnings benefits from estimated impacts on age 26 educational attainment; they also incorporate benefits from reductions in child care costs and child abuse/neglect. In a subsequent follow-up, Reynolds et al. (2011) show that the program significantly increased age 28 earnings by 7%.

<sup>15</sup>Brooks-Gunn et al. (1992) estimate greater effects of IHDP on age three IQ for families with lower maternal education.

<sup>16</sup>See (Baker, 2011) for a careful discussion of recent universal early child care initiatives.

multiple noisy measurements of child investments and skill levels to estimate the technology of human capital production from birth through the end of school. Their estimated technology suggests that the most efficient allocation of educational investments would provide more to young disadvantaged children. The fact that actual investments are much lower for disadvantaged children (see Figures 2-5) coupled with diminishing marginal returns, suggests that returns on the margin are higher for the most disadvantaged.

*Fact 3: exogenous increases in parental income lead to greater investments in children and improvements in childhood outcomes.* A number of recent studies attempt to address concerns about endogeneity in estimating the effects of exogenous changes in family income on children. Dahl and Lochner (2012) exploit expansions of the Earned Income Tax Credit (primarily over the mid-1990s) to estimate the effects of additional family income on cognitive achievement. Their instrumental variable estimates suggest that an additional \$1,000 in family income raises combined math and reading scores by 6 percent of a standard deviation. Estimated effects also appear to be larger for children from more disadvantaged families. Milligan and Stabile (2011) estimate that expansions of child tax benefits in Canada led to similar improvements in child cognitive and educational outcomes as well as improvements in child and maternal health. Combining data from ten welfare and anti-poverty experiments, Duncan et al. (2011) attempt to separately identify the effects of changes in family income from employment and other effects induced by different programs. Their analysis reaches similar conclusions regarding the impacts of income on child achievement as Dahl and Lochner (2012) and Milligan and Stabile (2011). Finally, Løken (2010) and Løken et al. (2012) estimate the impact of family income on Norwegian children using regional variation in the economic boom following the discovery of oil as an instrument for income. The latter study estimates that income has sizeable impacts on education and IQ for children from low-income families but much weaker effects for children from higher income families.<sup>17</sup>

Changes in income may affect children in many ways. In this paper, we focus on investment-based theories, so it is important to know whether changes in family income cause families to make different investment choices. A few studies suggest that this is the case. Carneiro and Ginja (2014) estimate models of income dynamics in the U.S., examining the extent to which family investments in children respond to permanent and transitory income shocks. Their results suggest modest positive responses to permanent shocks but negligible responses to transitory shocks. Effects appear to be largest for younger children and those with less-educated parents. Among children whose mothers had not attended college, a 10% increase in permanent income is estimated to

---

<sup>17</sup>Studies on the effects of parental job displacement on children also suggest that family income may have important effects on child schooling and labor market earnings (Oreopoulos et al., 2008; Stevens and Schaller, 2011); however, parental job displacement may also affect child development through other channels (e.g. family dissolution).

increase measures of cognitive stimulation and time investments by about .02 standard deviations. Cunha et al. (2010) and Pavan (2014) estimate both the technology of skill formation for children (from birth through later school ages) and the extent to which family income, as well as maternal and child skills, affect investments in children. Their estimated investment functions suggest that increases in family income lead to significantly higher investments in children.

Fact 4: *the timing of income matters for child development: increases in income at early ages (compared to later ages) lead to larger increases in investments and achievement/educational outcomes.* The estimated child investment functions of Pavan (2014) imply significantly greater effects of family income on investments at very early ages relative to older ages. Other studies estimate the effects of family income received at different child ages on adolescent achievement or educational outcomes. For example, Duncan and Brooks-Gunn (1997), Duncan et al. (1998), and Levy and Duncan (1999) all estimate that income received at earlier ages has a greater impact on adolescent achievement than income received at later ages. Carneiro and Heckman (2002) correctly point out, however, that (undiscounted) early income should have a larger effect than (undiscounted) later income due purely to discounting – something not taken into account in previous analyses.<sup>18</sup> More recent studies address this concern by discounting all income measures back to the year of birth (Carneiro and Heckman, 2002; Caucutt and Lochner, 2006, 2012).<sup>19</sup> Caucutt and Lochner (2006) report results consistent with the earlier literature, finding that income received at young ages has a greater effect than income received at older ages on subsequent child achievement. Caucutt and Lochner (2012) further show that family income earned when children are younger has a significantly greater effect on college attendance than does income earned at later ages; however, Carneiro and Heckman (2002) cannot reject that income has the same effects on college enrolment regardless of the age at which it was received. While both of these studies use data from the CNLSY, the former benefits from a sample size that is roughly twice as large, allowing for greater precision. Furthermore, because Carneiro and Heckman (2002) are more concerned with the importance of borrowing constraints at college-going ages, they control for age 12 math achievement levels, which may absorb much of the effect of early income.

## 4 Understanding Investment and Skill Gaps by Family Income

We now develop a general model of dynamic human capital investment that incorporates four mechanisms thought to generate child investment and skill gaps by family income. Within this framework, we explore the extent to which these mechanisms are also capable of explaining the

---

<sup>18</sup>That is, with perfect credit markets, income received at age 0 should have an effect that is  $(1+r)^a$  times larger than income received at age  $a$ , where  $r$  is the annual interest rate.

<sup>19</sup>These studies all use a 5% rate to discount income back to the year of a child's birth.

additional stylized facts just discussed. The problem is written as a lifecycle problem in which individuals invest in their human capital, while borrowing and saving (potentially subject to borrowing constraints) in an effort to finance investments and smooth consumption over time. However, the problem can also be interpreted as a ‘family’ investment problem, where altruistic parents make investments in their children and family borrowing/saving decisions to smooth family consumption.<sup>20</sup>

## 4.1 A General Model of Child Human Capital Investment

We assume that people live through three stages in their lives. Human capital investment takes place in the first two stages (i.e. ‘childhood’), followed by the final stage, adulthood. Adulthood may last for many periods; however, its length is inconsequential for most of our analysis. We are largely agnostic about the form that investments may take, instead focusing primarily on total investment expenditures at different ages and the dynamic nature of skill production. Conceptually, investments may include various forms of goods inputs like computers and books, parental time in child development activities, formal schooling, and other time inputs by older children.

### 4.1.1 Technology for Human Capital Production

Denote a child’s ability to learn by  $\theta > 0$ . Investment expenditures in periods 1 and 2 are given by  $i_1$  and  $i_2$ , respectively. Early investment produces an interim level of human capital,

$$h_2 = z i_1, \tag{1}$$

where  $z > 0$  is the productivity of early investment. Together, late investment and this interim human capital produce stage 3 (adult) human capital:

$$h_3 = \theta f(h_2, i_2). \tag{2}$$

The human capital production function  $f(\cdot)$  is strictly increasing and strictly concave in both of its arguments. To guarantee appropriate second order conditions hold in the decision problems described below, we assume the following throughout our analysis (without explicit reference):

**Assumption 1.**  $f_{12}^2 < f_{11} f_{22}$  and  $f_{12} > \max \left\{ f_{22} \left( \frac{f_1}{f_2} \right), f_{11} \left( \frac{f_2}{f_1} \right) \right\}$ .

The first condition limits the degree of dynamic complementarity in investments and ensures strict concavity of the production function. The second condition implies that the least costly way to produce additional human capital  $h_3$  is to increase both early and late investments. Most

---

<sup>20</sup>See Caucutt and Lochner (2012) for a direct mapping between this lifecycle problem and a more explicit intergenerational problem.

plausible specifications for human capital production would entail dynamic complementarity (i.e.  $f_{12} \geq 0$ ), satisfying this condition (Cunha et al., 2010; Caucutt and Lochner, 2012); however, the condition holds much more generally.<sup>21</sup> We also assume standard Inada conditions to ensure interior solutions.<sup>22</sup>

At times our analysis will employ a CES human capital production function of the form

$$f(h_2, i_2) = [a^{1-b}h_2^b + (1-a)^{1-b}i_2^b]^{d/b}, \quad (3)$$

where  $a \in (0, 1)$ ,  $b < 1$ , and  $d \in (0, 1)$ ; however, most of our analysis does not rely on any particular functional form. Assumption 1 holds for this production function.

#### 4.1.2 General Decision Problem

We assume a period utility function over consumption,  $u(c)$ , that is strictly increasing, strictly concave and satisfies standard Inada conditions. Tastes for early educational investment,  $i_1$ , are given by  $\nu i_1$ . The time discount rate is  $\beta \in (0, 1)$ , and the gross rate of return on borrowing and saving is  $R > 0$ . Assets saved in period  $j$  are given by  $a_{j+1}$ .

The individual/family receives exogenous income  $y_j$  during childhood periods  $j = 1, 2$ . We will sometimes refer to these as (early and late) parental income; although, it may also include government transfers or earnings while older children are still enrolled in school (in period 2).<sup>23</sup>

Children/families allocate their resources to consumption and skill investment, leaving some assets/debt for when the child grows up:

$$\max_{c_1, c_2, i_1, i_2, a_2, a_3} E [u(c_1) + \nu i_1 + \beta u(c_2) + \beta^2 V(a_3, h_3)] \quad (4)$$

subject to human capital production equations (1) and (2); budget constraints

$$a_{j+1} = Ra_j + y_j - i_j - c_j \quad \text{for } j = 1, 2; \quad (5)$$

initial assets  $a_1$  given; and where  $V(a_3, h_3)$  represents the child's utility in adulthood given  $a_3$  and  $h_3$ .

We now consider the main mechanisms commonly thought to explain income-based gaps in early investment and skill levels. We analyze each mechanism separately, abstracting from the others, in order to highlight the key underlying forces of each mechanism and the extent to which it can explain other stylized facts discussed in Section 3. In the next two subsections, we abstract

<sup>21</sup>For example, the condition holds for homothetic functions (e.g., CES) regardless of the degree of complementarity.

<sup>22</sup>That is,  $\lim_{h_2 \rightarrow 0} f_1(h_2, i_2) = \infty$ ,  $\forall i_2 \geq 0$ , and  $\lim_{i_2 \rightarrow 0} f_2(h_2, i_2) = \infty$ ,  $\forall h_2 \geq 0$ .

<sup>23</sup>Even if one considers the child to be the sole decision maker with  $y_1$  and  $y_2$  reflecting inter vivos transfers from parents, the interpretations in the text regarding parental income carry through as long as transfers are strictly increasing in parental income.

from uncertainty and other information problems, considering the consumption value of investment and the intergenerational correlation of ability. In Section 4.4, we study the implications of uncertainty and mis-information, at which point we describe information sets and variables over which expectations are taken in Equation (4). Finally, in Section 4.5 we introduce restrictions on borrowing of the form  $a_{j+1} \geq -L_j$ , where  $L_j$  is an upper limit on the amount that can be borrowed in period  $j$ . Until then, we assume that borrowing and saving are unrestricted.

## 4.2 Correlated Ability

We begin by studying the implications of a positive intergenerational correlation in ability, which is likely to generate a positive correlation between a child's ability and lifetime parental income, i.e.  $Cov(\theta, Y) > 0$ . To focus on this potential explanation for income-based gaps in investment and achievement, we study the effects of  $\theta$  on investments, marginal returns to investment, and human capital, while abstracting from any consumption value of schooling, uncertainty, and credit constraints. Specifically, we assume  $\nu = 0$  and that families have full and perfect information about the productivity of human capital investments. In the absence of borrowing constraints, the length of adulthood is irrelevant for our analysis, so we simply consider a three-period problem with  $V(a_3, h_3) = u(Ra_3 + h_3)$  and a single lifetime budget constraint. For expositional purposes, we normalize  $z = 1$ .

With these assumptions, the problem can be written as:

$$\max_{c_1, c_2, c_3, i_1, i_2} \{u(c_1) + \beta u(c_2) + \beta^2 u(c_3)\}$$

subject to the lifetime budget constraint:

$$c_1 + R^{-1}c_2 + R^{-2}c_3 = Y - i_1 - R^{-1}i_2 + R^{-2}\theta f(i_1, i_2),$$

where  $Y \equiv Ra_1 + y_1 + R^{-1}y_2$ , which we will often (loosely) refer to as 'lifetime parental income'. Notice,  $Y$  may also include initial family assets and government transfers; however, we can interpret the effects of changes in  $Y$  as changes in parental income holding these constant.

Optimal investments must satisfy the following first order conditions:

$$\theta f_1(i_1, i_2) = R^2, \tag{6}$$

$$\theta f_2(i_1, i_2) = R. \tag{7}$$

When investments are made purely for investment purposes, they are chosen to equate the marginal labor market returns to investment,  $\frac{\partial h_3}{\partial i_j} = \theta f_j(i_1, i_2)$  in both periods  $j = 1, 2$ , with the corresponding return to savings. This is the well-known result of Becker (1975): in the absence of borrowing constraints, uncertainty, and a direct utility value from investment, human

capital investments simply maximize discounted lifetime earnings net of investment expenditures. Importantly, this relationship holds regardless of ability or family income. As such, investments are independent of family income,  $Y$ , given ability. The role of ability is summarized in the following proposition. (All proofs can be found in Appendix B.)

**Proposition 1.** *Optimal investments satisfy the following: (i) the marginal returns to investments are independent of ability; (ii) early and late investments are strictly increasing in ability; (iii) adult human capital  $h_3$  is strictly increasing in ability.*

Not surprisingly, a positive correlation between parental income and child ability would produce a positive correlation between parental income and child investments and skills. Yet, the marginal return on investments should be unrelated to parental income, because investments in all children equate their marginal returns to the interest rate. This is inconsistent with both stylized Facts 1 and 2, which document returns to early investments for poor children that exceed standard interest rates as well as the returns for more economically advantaged children.

Additionally, the model implies no causal relationship between parental income and child investments/skills. Holding the child's ability constant, there should be no correlation between investments and parental income. Furthermore, exogenous changes in parental income should have no effect on investments in children, contradicting stylized Facts 3 and 4.

### 4.3 Consumption Value of Investment

We next explore the implications of a consumption/utility value associated with investment. To focus on this mechanism, we now incorporate tastes for schooling  $\nu \neq 0$  as in Equation (4), while continuing to assume perfect information, no credit constraints with  $V(a_3, h_3) = u(Ra_3 + h_3)$ , and  $z = 1$ .

For simplicity, we assume that  $\beta = R^{-1}$  so that optimal consumption profiles are flat:  $c_t = c = B^{-1}[Y - i_1 - R^{-1}i_2 + R^{-2}\theta f(i_1, i_2)]$  for  $t = 1, 2, 3$ , where  $B = 1 + R^{-1} + R^{-2}$ . In the absence of any consumption value associated with late investment,  $i_2$  is still determined from the first order condition above (Equation (7)); however, optimal early investment must now satisfy:

$$\theta f_1(i_1, i_2) = \left[1 - \frac{\nu}{u'(c)}\right] R^2. \quad (8)$$

When investment provides a direct consumption or utility value to children or their families, this must be taken into account when making investment decisions, driving a wedge between the marginal labor market return to investment and the return to savings. For a positive consumption value ( $\nu > 0$ ), early investment will have a low labor market return on the margin (i.e.  $\theta f_1 < R^2$ ),

because families will want to invest beyond the point where lifetime income is maximized. The opposite is true if families/children dislike investment ( $\nu < 0$ ).

When investment has a non-zero consumption value, the effects of parental income on investments and human capital, as well as the marginal labor market return to early investment, are easily derived from the the first order conditions above and are summarized in the following proposition.

**Proposition 2.** *For  $\nu > 0$  ( $\nu < 0$ ), optimal investments satisfy the following: (i) the marginal return to early investment is strictly less (greater) than the return to savings and is strictly decreasing (increasing) in lifetime parental income  $Y$ ; (ii) early investment is strictly increasing (decreasing) in lifetime parental income  $Y$ ; (iii) later investment is increasing (decreasing) in lifetime parental income  $Y$  if and only if  $f_{12} \geq 0$ ; and (iv) final human capital  $h_3$  is strictly increasing (decreasing) in lifetime parental income  $Y$ .*

These results are intuitive. If families enjoy investing in their children ( $\nu > 0$ ), they will invest beyond their income maximizing amounts and will invest more if their income rises. Thus, the positive relationship between family income and early childhood investment and skills requires a positive consumption value. Proposition 2 shows that if  $\nu > 0$ , the marginal labor market return to early investment should be low (inconsistent with Fact 1) and decreasing in lifetime parental income (consistent with Fact 2). A positive consumption value predicts that early investment should rise with exogenous increases in lifetime parental income (consistent with Fact 3); however, it predicts that the timing of that income is irrelevant (inconsistent with Fact 4).

One might reconcile the high estimated returns for early interventions targeted to economically disadvantaged children (Fact 1) by assuming that low-income parents find investment in their children costly (i.e.  $\nu < 0$  for low-income families). However, this would then imply that investment in young children and their skill levels should decline when poor parents receive additional income, contradicting Fact 3.

#### 4.4 Confusion

Poor families may face greater uncertainty about the returns to investment, or they may simply maintain inaccurate beliefs about the productivity of various investments. We next examine how uncertainty and mis-information influence investment behavior and highlight the importance of accounting for the dynamic nature of skill production. Throughout this subsection, we continue to assume  $\nu = 0$  and  $V(a_3, h_3) = u(Ra_3 + h_3)$  in order to focus on the role of information frictions.

It is useful to define a number of terms as we use them. We use the term *beliefs* to reflect a family's subjective probability distribution for some (productivity) parameter(s). For much of

our analysis, it does not matter whether these beliefs reflect actual variation or simply subjective uncertainty. We use the term *purely objective uncertainty* to refer to the case where beliefs coincide with the actual probability distribution for the parameter(s) of interest. This is also commonly referred to as rational expectations. We use the term *purely subjective uncertainty* to refer to the case where beliefs are non-degenerate even though the actual probability distribution is degenerate (i.e. if the true distribution for parameters were known, there would be no uncertainty). Whether uncertainty is *purely objective* or *purely subjective* is important for the realized marginal labor market returns to investments, because the former implies a distribution of ex post marginal returns for the same investments while the latter does not. We return to this point below.

We begin our analysis by studying uncertainty about  $\theta$ , which can reflect uncertainty about the child's ability to learn, parents' abilities to teach, or even the price of skill in the labor market. Here, we focus on the role of risk aversion and assume  $\theta$  is revealed only after all investments have been made. We next consider uncertainty about  $\theta$ , as well as the marginal productivity of early investments, when that uncertainty is completely resolved after early investments are made but before later investments are made. To focus on the nature of skill production and irreversibility of investments, we abstract from risk aversion here. Finally, we study early investments and their marginal returns when there are many types of early investment activities/inputs, and parents are mis-informed about the relative productivity of those activities/inputs.

#### 4.4.1 Risk Aversion and Uncertain Returns

We begin with a very natural form of *purely objective uncertainty*: both beliefs and the true distribution of  $\theta$  are given by  $\theta \sim \Phi(\theta)$ , with  $\bar{\theta} \equiv E(\theta)$ . For this analysis, we assume that the true value of  $\theta$  is not revealed until after all skill investments have been made. We continue to normalize  $z = 1$ .

If individuals are risk averse, the expected return to risky investments should exceed the return on safe investments if individuals are to hold risky assets at all. With a concave human capital production technology, this means that skill investments will be lower under uncertainty (Levhari and Weiss, 1974).

The first order conditions for investments satisfy

$$f_1(i_1, i_2) = Rf_2(i_1, i_2).$$

Given the separability between  $\theta$  and  $f(\cdot)$ , there is no distortion between early and late investment even if total investment spending is distorted. That is, for a given level of spending  $i_1 + R^{-1}i_2$ , early and late investments are chosen to maximize  $f(i_1, i_2)$ . Assumption 1 ensures that both  $i_1$  and  $i_2$  increase when total investment spending increases.

The level of total investment spending will equate the expected marginal benefit with the marginal cost of investment, so

$$\bar{\theta} f_1(i_1, i_2) + \frac{\text{Cov}(u'(c_3(\theta)), \theta)}{E[u'(c_3(\theta))]} f_1(i_1, i_2) = R^2,$$

where  $c_3(\theta) = Ra_3 + \theta f(i_1, i_2)$  is optimal period 3 consumption in state  $\theta$ . The expected marginal benefit of investment consists of a monetary return (first term) and a utility cost (second term). Because the marginal utility of consumption is low in states with a high return to investment, uncertain returns produce an additional utility cost of investment as reflected in the (negative) covariance term. As such, the expected marginal labor market returns to investment exceed the return to savings:

$$\bar{\theta} f_1(i_1, i_2) = R\bar{\theta} f_2(i_1, i_2) > R^2.$$

Risk averse individuals facing uncertain returns invest less at all ages relative to those who know the return with certainty.

Moreover, if having greater resources makes people less risk averse, then investments are increasing and marginal labor market returns decreasing in lifetime parental income  $Y$ . Because no information about the value of  $\theta$  is revealed until all investments have been made, choices depend only on the discounted present value of income over all investment years and not the timing of that income. The following proposition summarizes these results.<sup>24</sup>

**Proposition 3.** *When there is uncertainty in the final return to investment  $\theta$ , optimal investments satisfy the following: (i) expected marginal returns to investment are strictly greater than the return to savings; (ii) expected marginal returns to investment are strictly decreasing in parental income  $Y$  if  $u(\cdot)$  exhibits decreasing absolute risk aversion; and (iii) early and late investments are strictly increasing in parental income  $Y$  if  $u(\cdot)$  exhibits decreasing absolute risk aversion.*

#### 4.4.2 Uncertain Returns and the Irreversibility of Early Investment

We now consider the case in which families face uncertainty when making early investments in their children; however, that uncertainty is fully resolved before late investments are chosen. While early investments made under uncertainty are irreversible (i.e. families cannot go back in time to modify ex post suboptimal early investment choices), families can base late investment decisions on the realizations of early investments and full knowledge of the human capital

---

<sup>24</sup>These results assume *purely objective uncertainty*; however, they also apply to the case of unbiased *subjective uncertainty*. The only difference in the latter case is that all individuals would experience the same marginal return to investments, given by the expected marginal returns in the case of *purely objective uncertainty*. Thus, the marginal returns to investment exceed the return on savings (and decline with income) regardless of the form of uncertainty.

production process. Thus, families may be able to partially compensate for ex post suboptimal early investment through their choice of late investment. The extent to which this is effective depends crucially on the intertemporal complementarity/substitutability of investments.<sup>25</sup> It also depends on which features of technology are unknown. We consider uncertainty in the productivity of both investments,  $\theta$ , as well as in the productivity of early investments alone,  $z$ , at the time early investments are made.

To focus on the implications of investment irreversibility and the dynamic nature of human capital productivity, we abstract from risk aversion. We continue to focus on *purely objective uncertainty* with the distribution of beliefs over  $(\theta, z)$  reflecting the true variation in these productivity parameters; however, we comment briefly on the implications of *purely subjective uncertainty* at the end of our discussion.

It is useful to begin with the second period investment problem, which conditions on early investment and technology once  $\theta$  and  $z$  are known. Let  $\hat{i}_2(zi_1, \theta)$  denote the optimal second period investment conditional on  $i_1$  and technology state  $(\theta, z)$ :

$$\hat{i}_2(zi_1, \theta) \equiv \underset{i_2}{\operatorname{argmax}} \{ -i_2 + R^{-1}\theta f(zi_1, i_2) \}. \quad (9)$$

Optimal late investment equates the marginal labor market return with the return to savings, i.e.  $\theta f_2(zi_1, i_2) = R$ . From this, it is easy to see that second period investment is increasing in  $\theta$ . However, late investment is increasing in  $z$  if and only if early and late investments are gross complements (i.e.  $f_{12} \geq 0$ ), because  $z$  only affects the marginal return to late investment indirectly through  $h_2 = zi_1$ . Here, we begin to see the distinction between neutral and early-specific productivity as well as the importance of intertemporal complementarity/substitutability of investments. These factors are also important in determining the response of early investment to uncertainty about investment productivity.

Taking the late investment policy  $\hat{i}_2(zi_1, \theta)$  as given, the net realized (or ex post) return to early investment for actual productivity parameters  $(\theta, z)$  is given by:

$$\Pi(i_1, \theta, z) \equiv -i_1 - R^{-1}\hat{i}_2(zi_1, \theta) + R^{-2}\theta f(zi_1, \hat{i}_2(zi_1, \theta)). \quad (10)$$

The following lemma establishes concavity of net realized returns in early investment and is useful for a number of results.

**Lemma 1.** *The net return to early investment  $\Pi(i_1, \theta, z)$  is strictly concave in  $i_1$ .*

---

<sup>25</sup>See Cunha and Heckman (2007) for a discussion of dynamic complementarity and the difficulty of compensating for low early investments by increasing late investments. Cunha et al. (2010) estimate and Caucutt and Lochner (2012) calibrate a strong degree of intertemporal complementarity for investments in children.

Because  $i_1$  must be determined before  $(\theta, z)$  is realized, optimal early investment maximizes the expected net return:

$$\tilde{i}_1 \equiv \operatorname{argmax}_{i_1} E[\Pi(i_1, \theta, z)],$$

where the expectation is taken over the distribution of  $(\theta, z)$ . The first order condition equates the expected marginal labor market return to early investment with the return to savings:

$$E[z\theta f_1(z\tilde{i}_1, \hat{i}_2(z\tilde{i}_1, \theta))] = R^2. \quad (11)$$

With *purely objective uncertainty* and risk neutrality, the expected marginal labor market return to early investment always equals the return to savings regardless of the type (i.e.  $\theta$  or  $z$ ) or extent of uncertainty – the expected marginal return is independent of the  $(\theta, z)$  distribution. In contrast with Facts 1 and 2 of Section 3, the average marginal labor market return should equal the interest rate for children from all backgrounds.

We are also interested in understanding how changes in the distribution of  $(\theta, z)$  affect early investment amounts. We consider two notions of a change in the distribution that are widely used in economics: *first order stochastic dominance* and *mean-preserving spread* (Rothschild and Stiglitz, 1970, 1971). In this framework, how  $\tilde{i}_1$  changes with the distribution of productivity parameters depends on how  $\partial\Pi(i_1, \theta, z)/\partial i_1$  varies with  $(\theta, z)$ . If  $\partial\Pi(i_1, \theta, z)/\partial i_1$  is increasing (decreasing) in  $\theta$ , a first order stochastic dominance shift in  $\theta$  will increase (decrease)  $\tilde{i}_1$ . If  $\partial\Pi(i_1, \theta, z)/\partial i_1$  is concave (convex) in  $\theta$ , a mean-preserving spread in  $\theta$  decreases (increases)  $\tilde{i}_1$ . The same is true for changes in the distribution of  $z$ . We first consider investment when  $\theta$  is unknown, then turn attention to the case with  $z$  unknown. Some of our results assume a CES production function for human capital as defined in Equation (3).

### Neutral Productivity Shock

We now consider uncertainty in the overall ability of a child ( $\theta$ ) that is fully resolved after early investments have been made but before late investments are chosen. We assume  $z$  is known.<sup>26</sup>

**Proposition 4.** (i) *A first order stochastic dominance shift in  $\theta$  increases early investment.*  
(ii) *For the CES production function (3), a mean-preserving spread in  $\theta$  reduces early investment if and only if  $b > d$ .*

It is not surprising that a first order stochastic shift in  $\theta$  unambiguously increases early investment, because  $\theta$  directly raises the marginal return to investment for any given level of early and late investment. The effect of a mean-preserving spread in the distribution of  $\theta$  is more

<sup>26</sup>This case was originally considered by Hartman (1976) in the analysis of firm investment and labor demand under uncertain output prices.

complicated and depends on the degree of complementarity between investments. For a CES human capital production function, an increase in uncertainty about  $\theta$  reduces early investment if and only if early and late investments are gross substitutes (i.e.  $b > d \Leftrightarrow f_{12} < 0$ ). With strong intertemporal substitutability, families facing uncertainty about  $\theta$  will choose to invest little in the first period and wait to learn the productivity of investment. If investment is highly productive, the family can easily compensate for inadequate early investment by investing more in the second period. In the more empirically relevant case where investments are gross complements ( $f_{12} > 0$ ), it is too costly to make up for a lack of early investment by increasing late investment. As a result, early investment increases with the degree of uncertainty.

### Early Investment-Specific Productivity Shock

Families may be more uncertain about the productivity of early investments in their children than they are about later investments like college attendance. To explore this possibility, we now assume  $\theta$  is known from birth and consider the case where  $z$  is initially unknown but revealed after early investments have been made.

A change in  $z$  has two opposing effects on the marginal return to early investment. An increase in  $z$  directly increases the productivity of  $i_1$ , but it also reduces the marginal return because  $h_2 = zi_1$  is subject to diminishing returns in the production of adult human capital  $h_3$ . The latter effect is attenuated by adjustments in late investments when  $f_{12} \neq 0$ . The overall effect of  $z$  depends on the following condition:

**Condition 1.**  $f_1 > \left( \frac{f_{11}f_{22} - f_{12}^2}{-f_{22}} \right) zi_1$ .

If the direct productivity effect (left hand side) is greater than the the diminishing return effect (right hand side), then the marginal return to  $i_1$  is greater for larger  $z$ . For the CES production function given in Equation (3), this condition holds if  $b \geq 0$ .

Condition 1 is appealing, because it is equivalent to requiring that an increase in  $z$  raises the net marginal return to early investment.

**Lemma 2.** *The marginal net return to early investment  $\partial\Pi(i_1, \theta, z)/\partial i_1$  is strictly increasing in  $z$  if and only if Condition 1 holds.*

When Condition 1 is satisfied, a better distribution of  $z$  produces a higher marginal return to  $i_1$ , on average, which makes it profitable to increase early investment. The following proposition formalizes this result and characterizes the effects of a mean-preserving spread in  $z$ .

**Proposition 5.** *For the CES production function (3), (i) a first order stochastic dominance shift in  $z$  increases early investment if  $b \geq 0$ ; (ii) a mean preserving spread in  $z$  reduces early investment if  $b \geq 0$ .*

The effect of a mean-preserving spread in  $z$  is similar to its counterpart for  $\theta$  (Proposition 4), except that uncertainty in  $z$  discourages early investment more than does uncertainty in  $\theta$ . In contrast with an increase in uncertainty about  $\theta$ , an increase in uncertainty about  $z$  can reduce early investment even when early and late investment are gross complements (e.g.  $0 \leq b \leq d$ ). Intuitively, diminishing marginal returns to  $h_2 = zi_1$  in the production of adult human capital lessens the benefits of high  $z$  realizations for the marginal return to early investment. This force moderates the costs of under-investment and discourages early investment when  $z$  is uncertain, but it is absent with uncertainty in  $\theta$ .

### Purely Subjective Uncertainty, Investments, and Marginal Returns

The previous analysis assumes individuals have *purely objective uncertainty* about heterogeneous productivity levels  $(\theta, z)$ . Even if all children have the same productivity levels, they may have different subjective beliefs about the true productivity of investments. For example, the poor may have downward biased beliefs or they may have unbiased beliefs with greater subjective uncertainty. Regardless, Propositions 4 and 5 characterize the effects of changes in beliefs on early investment choices.

More interestingly, *purely subjective uncertainty* has different implications from *purely objective uncertainty* for observed marginal returns in the labor market. In the latter case, families facing the same distribution of productivity levels make the same early investment choices, but they experience different labor market outcomes due to heterogeneous productivity levels. As discussed earlier (see the discussion surrounding Equation (11)), the average realized marginal labor market return under *purely objective uncertainty* always equals the return to savings. The case of *purely subjective uncertainty* is quite different. All families with the same beliefs and ability/productivity will make the same early investment choices and will, therefore, experience the same labor market returns. Strict concavity of  $\Pi(i_1, \theta, z)$  in  $i_1$  (Lemma 1) directly implies that under *purely subjective uncertainty*, the observed marginal labor market return to early investment is strictly decreasing in the level of early investment. This, together with Propositions 4 and 5, directly implies the following corollary.

**Corollary 1.** *Assume the CES production function given in Equation (3). Under purely subjective uncertainty, a mean-preserving spread in the distribution of beliefs about  $\theta$  ( $z$ ) increases the marginal labor market return to early investment if and only if  $b > d$  (if  $b \geq 0$ ).*

Families with greater subjective uncertainty about  $\theta$  will have a higher marginal labor market return to early investment if and only if early and late investments are gross substitutes. Marginal labor market returns will be increasing in the amount of subjective uncertainty about  $z$  under a

modest amount of dynamic complementarity in investments.

### 4.4.3 Biased Beliefs about Human Capital Production

Thus far, we have focused on the extent of uncertainty about the productivity of investments. It is also possible that parents may have little subjective uncertainty about the productivity of investments, but their beliefs may be biased. This possibility seems particularly likely to arise when there are many different potential inputs/activities families may engage in to raise the human capital of their children (e.g. reading to children, taking them to museums, teaching them to play musical instruments, playing with them). Even if parents are correct in gauging the average productivity across different inputs, they might easily misjudge their relative productivity. To explore this issue, we now assume homogeneity in the productivity of different early inputs across families; however, we allow for the possibility that families may hold biased beliefs about the productivity of any or all early child inputs. To simplify the analysis, we abstract away from any form of uncertainty – families are certain but may be wrong. We further assume that families learn the true outcomes of their early investments,  $h_2$ , before they need to make later investment decisions.

Assume early investment consists of  $n$  different ‘activities’  $\mathbf{x} = (x_1, \dots, x_n)$  that produce  $h_2$  according to the following CES production function:

$$h_2 = z \left( \sum_{j=1}^n w_j^{1-\phi} x_j^\phi \right)^{\frac{1}{\phi}}, \quad (12)$$

where  $\phi \in (0, 1)$  and  $\mathbf{w} = (w_1, \dots, w_n) \geq 0$  satisfies  $\sum_{j=1}^n w_j = 1$ .<sup>27</sup> Here,  $z$  reflects the total factor productivity of early investments. Changes in  $z$  have no effect on the relative productivity or optimal composition of different inputs. Productivity weights  $w_j$  determine the relative importance of each input as well as their optimal expenditure shares. It is straightforward to show that demand for input  $x_j$  conditional on total early investment spending  $i_1 = \sum_{j=1}^n x_j$  is given by

$$x_j = w_j i_1.$$

By substituting these conditional demands into the production function, we obtain the indirect production function (as a function of total early expenditure  $i_1$ ) equivalent to that assumed earlier:

$$h_2 = z \left( \sum_{j=1}^n w_j^{1-\phi} (w_j i_1)^\phi \right)^{\frac{1}{\phi}} = z i_1.$$

<sup>27</sup>Vector equality and inequality are defined as follows: (i)  $\tilde{\mathbf{x}} = \mathbf{x}$  if  $\tilde{x}_j = x_j$  for all  $j = 1, \dots, n$ ; (ii)  $\tilde{\mathbf{x}} \neq \mathbf{x}$  if  $\tilde{x}_j \neq x_j$  for some  $j = 1, \dots, n$ ; (iii)  $\tilde{\mathbf{x}} \leq \mathbf{x}$  if  $\tilde{x}_j \leq x_j$  for all  $j = 1, \dots, n$ .

To investigate the implications of incorrect beliefs about early investment productivity, it is useful to distinguish beliefs from actual productivity parameters. Let  $\tilde{z}$  and  $\tilde{\mathbf{w}}$  denote a family's beliefs about  $z$  and  $\mathbf{w}$ , respectively. Without loss of generality, we assume  $\sum_{j=1}^n \tilde{w}_j = 1$ . We say that a belief is *biased* if  $\tilde{z} \neq z$  or  $\tilde{\mathbf{w}} \neq \mathbf{w}$ . When  $\tilde{z} \neq z$ , the bias is *systematic* in the sense that families are, on average, biased about the productivity of early investments. When  $\tilde{z} = z$  but  $\tilde{\mathbf{w}} \neq \mathbf{w}$ , the bias is *non-systematic*, because beliefs are, on average, correct even though they are wrong about the relative productivity of different early inputs.

Let  $(\tilde{\mathbf{x}}, \tilde{i}_1, \tilde{h}_2, \tilde{h}_3)$  be the optimally chosen investments and realized human capital of children with family beliefs  $(\tilde{z}, \tilde{\mathbf{w}})$ . Let  $(\mathbf{x}^*, i_1^*, h_2^*, h_3^*)$  reflect these same variables when beliefs are unbiased. Families first choose investments  $\mathbf{x}$  and  $i_1$  based on their beliefs  $(\tilde{z}, \tilde{\mathbf{w}})$ :

$$\begin{aligned} \tilde{i}_1 &= \operatorname{argmax}_{i_1} \Pi(i_1, \theta, \tilde{z}) \\ \tilde{x}_j &= \tilde{w}_j \tilde{i}_1, \quad \forall j = 1, \dots, n, \end{aligned}$$

where the child's lifetime income net of investments,  $\Pi(\cdot, \cdot, \cdot)$ , is defined by equation (10). Notice that total early investment spending  $\tilde{i}_1$  is only affected by  $\tilde{z}$  and not by  $\tilde{\mathbf{w}}$ , because  $\mathbf{w}$  does not affect total factor productivity.

Next, interim human capital  $\tilde{h}_2$  is realized based on investments choices  $(\tilde{\mathbf{x}}, \tilde{i}_1)$  and the true technology  $(z, \mathbf{w})$ :

$$\tilde{h}_2 = z \left( \sum_{j=1}^n w_j^{1-\phi} \tilde{x}_j^\phi \right)^{\frac{1}{\phi}} = z \tau(\tilde{\mathbf{w}}) \tilde{i}_1,$$

where

$$\tau(\tilde{\mathbf{w}}) \equiv \left( \sum_{j=1}^n w_j^{1-\phi} \tilde{w}_j^\phi \right)^{\frac{1}{\phi}} \leq 1$$

reflects the distortion due to a suboptimal allocation of expenditures across inputs. When  $\tilde{\mathbf{w}} \neq \mathbf{w}$ , early investment spending is less productive than it should be ( $\tau(\tilde{\mathbf{w}}) < 1$ ), so interim human capital is low ( $\tilde{h}_2 < z \tilde{i}_1$ ).

We assume that  $i_2$  is chosen knowing the actual realization for  $\tilde{h}_2$ . That is, families are able to evaluate their child's skill/achievement, effectively learning that their beliefs were mistaken.<sup>28</sup> Given the resulting interim human capital, late investments are determined as in Section 4.4.2 (see Equation (9)). Finally, adult human capital  $\tilde{h}_3$  is produced based on actual  $\tilde{h}_2$  and late investment:

$$\tilde{h}_3 = \theta f(\tilde{h}_2, \hat{i}_2(\tilde{h}_2, \theta)).$$

---

<sup>28</sup>Note that this does not necessarily require that parents learn the true productivity values  $(z, \mathbf{w})$ ; although, it is too late to matter.

where  $\hat{i}_2(\cdot, \cdot)$  is defined in Equation (9).

We first study how *systematic* bias affects early investment and human capital accumulation.

**Proposition 6.** (i)  $\tilde{i}_1$  is strictly increasing in  $\tilde{z}$  if and only if Condition 1 holds for  $(h_2, i_2) = (\tilde{z}\tilde{i}_1, \hat{i}_2(\tilde{z}\tilde{i}_1, \theta))$ . (ii) Suppose that  $\tilde{\mathbf{w}} = \mathbf{w}$ . If and only if  $\tilde{i}_1 \leq i_1^*$ , then:  $\tilde{\mathbf{x}} \leq \mathbf{x}^*$ ,  $\tilde{h}_2 \leq h_2^*$ ,  $\tilde{h}_3 \leq h_3^*$ , and  $z\theta f_1(\tilde{h}_2, \hat{i}_2(\tilde{h}_2, \theta)) \geq R^2$ .

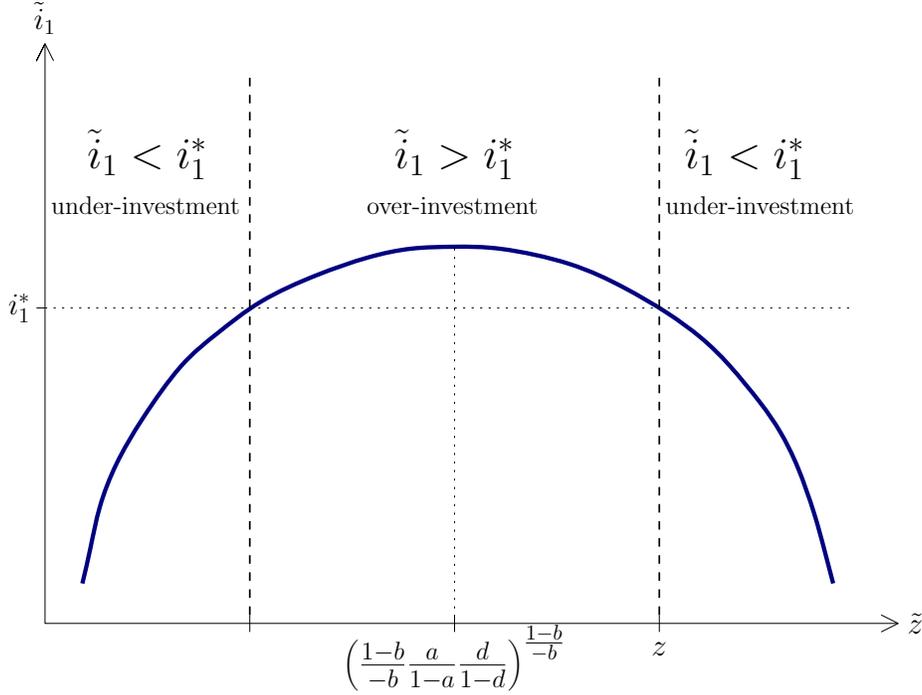
Families with a biased belief about  $z$  behave as if the true productivity of early investment is  $\tilde{z}$  rather than  $z$ . As shown in Lemma 1, higher productivity in early investment does not necessarily lead to more early investment due to the diminishing return effect. However, when this effect is weak so Condition 1 is satisfied (e.g. modest dynamic complementarity or substitutability), individuals with downward biased beliefs under-invest in all early inputs/activities resulting in low levels of human capital. Moreover, under-investment implies a high observed marginal labor market return to early investment. By contrast, when Condition 1 does not hold, families with downward biased beliefs may over-invest (in all inputs) and obtain high levels of human capital. In this case, providing information that shifts beliefs upwards towards the truth would actually *reduce* early investment and human capital. Figure 6 demonstrates this possibility with CES production function (3) and  $b < 0$ . In this example, moving downward biased beliefs  $\tilde{z}$  from the middle of the graph towards the true (higher) value of  $z$  would result in lower (but more efficient) levels of early investment.

Next, consider the effects of *non-systematic* bias.

**Proposition 7.** Suppose that  $\tilde{z} = z$  and  $\tilde{\mathbf{w}} \neq \mathbf{w}$ . Then (i)  $\tilde{x}_j \leq x_j^*$  if and only if  $\tilde{w}_j \leq w_j$ ; (ii)  $\tilde{i}_1 = i_1^*$ ,  $\tilde{h}_2 < h_2^*$ , and  $\tilde{h}_3 < h_3^*$ ; (iii)  $z\tau(\tilde{\mathbf{w}})\theta f_1(\tilde{h}_2, \hat{i}_2(\tilde{h}_2, \theta)) < R^2$  if Condition 1 holds for all  $(h_2, i_2) = (z'i_1^*, \hat{i}_2(z'i_1^*, \theta))$  where  $z' \in [z\tau(\tilde{\mathbf{w}}), z]$ .

Non-systematic bias does not affect total early investment spending  $\tilde{i}_1$ , but it reduces the actual return to early investment due to the misallocation of resources to the wrong inputs. As such, it leads to low levels of human capital. In this case, providing more precise information will not affect total early investment expenditures, but it will lead to more efficient human capital production and, consequently, greater human capital. Because non-systematic bias reduces the productivity of early investment while leaving total investment expenditure unaffected, its effect on the marginal return to early investment depends on Condition 1. When Condition 1 holds, families with non-systematic bias have lower marginal labor market returns to early investment due to misallocation.

Figure 6: Systematic Bias and Early Investment (CES with  $b < 0$ )



#### 4.4.4 Information Problems and the Stylized Facts

The nature of an information problem is important for understanding its effects on human capital investment behavior. While none of the information problems we study are able to explain why the timing of income is important for human capital investment (Fact 4), some are more promising for explaining the other stylized facts in Section 3.

Uncertainty about child ability  $\theta$  (or labor market returns to human capital) coupled with risk aversion causes families to under-invest in their children. With decreasing absolute risk aversion, under-investment is worse among the poor, and an increase in lifetime parental income would be met with an increase in child investments (Fact 3). Expected marginal returns exceed the return to savings (Fact 1) and are especially high for children from low-income families (Fact 2). These results apply whether uncertainty is objective or subjective.

We also explore the implications of uncertainty resolved after early investments have been made but in time for late investment choices to respond. Here, we abstract from risk aversion in order to focus on the role of early investment irreversibility and the technology of skill formation. If early and late investments are mildly complementary, subjective uncertainty about the productivity of early investments,  $z$ , can lead to under-investment in young children and high marginal returns to early investment (relative to interest rates). If poor families face greater subjective uncertainty about  $z$ , then they will invest less in their young children than higher

income families, stopping investment when marginal returns are relatively high.<sup>29</sup> Thus, with modest dynamic complementarity, differences in subjective uncertainty by parental income can help explain Facts 1 and 2. While this form of uncertainty can explain the positive correlation between parental income and child investments, it does not help us understand why changes in parental income lead to contemporaneous changes in child investments and achievement (Facts 3 and 4) unless income brings new information with it.<sup>30</sup> Uncertainty about  $\theta$  (with risk neutrality) resolved after early childhood is inconsistent with a positive parental income – child investment relationship and other stylized facts unless early and late investments are substitutes.<sup>31</sup>

Finally, we consider the possibility that families are simply mistaken about the productivity of early investment activities as documented in Cunha (2014) and Dizon-Ross (2014). Assuming modest dynamic complementarity, we show that poor families that systematically under-estimate the productivity of early investments will under-invest in their young children and have a high marginal return to early investment relative to the return to savings and the marginal return for high income families with accurate beliefs (Facts 1 and 2). By contrast, non-systematic bias (i.e. over-estimation of the productivity of some inputs offset by under-estimation of the productivity of others) has no effect on total early investment expenditures. Instead, it results in a misallocation across early inputs, which tends to reduce the marginal return to early expenditures.<sup>32</sup> Thus, non-systematic bias among poor families cannot explain the basic correlation between family income and child investment/achievement, nor can it explain the high marginal return to investment among the poor (Facts 1 and 2). Neither form of bias helps explain the responsiveness of early investment and achievement to changes in income (Facts 3 and 4).<sup>33</sup>

## 4.5 Borrowing Constraints

Lastly, we explore the possibility that families may be unable to borrow against future earnings to efficiently finance investments in their children. Extending the analysis of Cunha and Heckman (2007), we demonstrate the importance of dynamic complementarity of investments for the impacts of both borrowing constraints and family resources on investment behavior.

We incorporate borrowing constraints by imposing upper limits on the total debts families

---

<sup>29</sup>With purely objective uncertainty and risk neutrality, expected marginal returns to investment always equal the return on savings.

<sup>30</sup> Changes in income may lead to changes in information and, therefore, investment behavior, if information about the productivity value of investment can be purchased by families. We do not explicitly model this possibility.

<sup>31</sup>Available evidence suggests dynamic complementarity for investments (Cunha et al., 2010; Caucutt and Lochner, 2012).

<sup>32</sup>The prediction of over-investment in some inputs and under-investment in others conflicts with Figures 2 and 3, which shows that the poor generally invest less than the rich in nearly all inputs.

<sup>33</sup>Systematic bias in  $z$  may be consistent with Fact 3 if  $b > 0$  and beliefs about  $z$  respond to changes in family income. See footnote 30.

can accumulate in any period. Specifically, we restrict assets carried into any period  $j + 1$  to satisfy the constraint  $a_{j+1} \geq -L_j$ . To focus on the role of borrowing constraints, we consider the problem of Section 4.1.2 (with borrowing constraints) assuming  $z = 1$ ,  $\nu = 0$ , and perfect information.

Because we are not only concerned with borrowing constraints during the investment period, but also later in life, we interpret the continuation utility  $V(a_3, \theta f(h_2, i_2))$  as the solution to the asset allocation problem for individuals entering adulthood, allowing for the possibility of binding future constraints. We assume that individuals live to age  $T$  and that adult earnings depend on human capital acquired through childhood investments  $h_3$ , growing exogenously thereafter with

$$h_j = \Gamma_j h_3, \quad j \in \{4, \dots, T\}. \quad (13)$$

Individuals entering adulthood with human capital  $h_3$  and assets  $a_3$  allocate consumption across their remaining life in the following way:

$$V(a_3, h_3) = \max_{c_3, \dots, c_T} \sum_{j=3}^T \beta^{j-3} u(c_j), \quad (14)$$

subject to budget constraints  $a_{j+1} = Ra_j + h_j - c_j$  for  $j \in \{3, \dots, T\}$ , borrowing constraints  $a_{j+1} \geq -L_j$  for  $j \in \{3, \dots, T-1\}$ , and  $a_{T+1} = 0$ .<sup>34</sup>

Given the value function defined in (14), families solve the maximization problem (4) subject to budget constraints (5), initial assets  $a_1$ , and borrowing constraints  $a_2 \geq -L_1$  and  $a_3 \geq -L_2$ .

#### 4.5.1 Investment Behavior

Consumption allocations satisfy  $u'(c_j) \geq \beta R u'(c_{j+1})$ ,  $\forall j = 1, \dots, T-1$ , where the inequality is strict if and only if the borrowing constraint for period  $j$  ( $a_{j+1} \geq -L_j$ ) binds. First order conditions for investment are given by

$$u'(c_1) = \beta^2 \theta \frac{\partial V(a_3, h_3)}{\partial h_3} f_1(i_1, i_2), \quad (15)$$

$$u'(c_2) = \beta \theta \frac{\partial V(a_3, h_3)}{\partial h_3} f_2(i_1, i_2), \quad (16)$$

where  $\frac{\partial V(a_3, h_3)}{\partial h_3} = \sum_{j=3}^T \beta^{j-3} \Gamma_j u'(c_j) > 0$  with  $\Gamma_3 = 1$ . Taking the ratio of these equations reveals that optimal investment equates the technical rate of substitution in the production of human capital with the marginal rate of substitution for consumption:  $\frac{f_1(i_1, i_2)}{f_2(i_1, i_2)} = \frac{u'(c_1)}{\beta u'(c_2)} \geq R$ .

Unconstrained optimal investments,  $i_1^*$  and  $i_2^*$ , satisfy  $\chi \theta f_1(i_1^*, i_2^*) = R^2$  and  $\chi \theta f_2(i_1^*, i_2^*) = R$ , where  $\chi = \sum_{j=3}^T R^{3-j} \Gamma_j$  reflects the discounted present value of an additional unit of human

<sup>34</sup>Appendix B shows that  $V_2 > 0$ ,  $V_{22} < 0$ , and  $V_{21} < 0$ . These properties are used repeatedly in proving results below.

capital. As in Section 4.2, unconstrained investments maximize the discounted present value of lifetime earnings net of investment costs. They are independent of the marginal utility of consumption and income/transfers, because individuals can optimally smooth consumption across periods. This is not true when borrowing constraints bind as shown in the next proposition.

**Proposition 8.** (i) *If and only if any borrowing constraint binds, then: optimal early investment is strictly less than the unconstrained amount, the marginal return to early investment is strictly greater than the return to savings, and adult human capital is strictly less than the unconstrained level.* (ii) *If any borrowing constraint binds and either (a) the period one constraint does not bind or (b)  $f_{12} > 0$ , then optimal late investment is strictly less than the unconstrained amount.* (iii) *If and only if any borrowing constraint in period two or later binds, then the marginal return to late investment is strictly greater than the return to savings.*

Early investment is always low and its marginal labor market return high (relative to the unconstrained case) when any borrowing constraints bind. Late investment is also low and its marginal return high if constraints at that age or later are binding and either the early constraint does not bind or early and late investments are complementary.

The complementarity of investments across periods plays a central role in determining individual responses to borrowing constraints and changes in parental income. If investments are very substitutable, individuals can shift investment from constrained periods to unconstrained periods with little loss to total acquired human capital. Their ability to do this is diminished as investments become more complementary. In particular, the following dynamic complementarity condition is important for a number of results.

**Condition 2.**  $\frac{f_{12}f}{f_1f_2} > \frac{-V_{22}(-RL_2, h_3)h_3}{V_2(-RL_2, h_3)}$ .

If preferences are given by the constant intertemporal elasticity of substitution (IES) form  $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$  (where  $1/\sigma$  is the IES) and credit constraints are non-binding throughout adulthood, then this condition simplifies to something very intuitive:

$$\underbrace{\frac{f_1 f_2}{f_{12} f}}_{\text{Hicksian elasticity of substitution}} < \underbrace{\frac{1}{\sigma}}_{\text{IES}} \underbrace{\left(1 - \frac{RL_2}{\chi h_3}\right)}_{1 - \frac{\text{maximum debt}}{\text{lifetime income}}}.$$

See Appendix B for details.<sup>35</sup> As the Hicksian elasticity of substitution between early and late investments declines (i.e. investments become more complementary) or the consumption intertemporal elasticity of substitution increases (i.e. individuals become less concerned about maintaining

<sup>35</sup>For the CES production function given in equation (3), the Hicksian elasticity of substitution between early and late investments (the left hand side) is simply  $\frac{d}{d-b}$ . The condition cannot hold for  $d \leq b$ , but this only rules out very strong substitution between early and late investments such that  $f_{12} \leq 0$ .

smooth consumption profiles), this inequality is more likely to hold. More generally, when individual preferences for smooth consumption are strong, Condition 2 requires strong complementarity between early and late investments.

We are now ready to study how family income during early and late childhood affect investment behavior. As noted above, changes in family income have no effect on investments for unconstrained individuals. The following proposition shows how constraints at different stages of child development determine the responsiveness of investment to changes in income at early and late ages. These results highlight how the timing of income/transfers can impact human capital investments and accumulation when individuals are constrained.<sup>36</sup>

**Proposition 9.**

*I. If borrowing constraints bind in late childhood, but not early childhood, then:*

- (i)  $\frac{\partial i_1}{\partial y_1} = R \frac{\partial i_1}{\partial y_2} = \frac{\partial i_1}{\partial (R^{-1}y_2)} > 0;$
- (ii)  $\frac{\partial i_2}{\partial y_1} = R \frac{\partial i_2}{\partial y_2} = \frac{\partial i_2}{\partial (R^{-1}y_2)} > 0;$
- (iii)  $\frac{\partial h_3}{\partial y_1} = R \frac{\partial h_3}{\partial y_2} = \frac{\partial h_3}{\partial (R^{-1}y_2)} > 0.$

*II. If borrowing constraints only bind in early childhood, then:*

- (i)  $\frac{\partial i_1}{\partial y_1} > 0;$  and  $\frac{\partial i_1}{\partial y_2} < 0;$
- (ii)  $\frac{\partial i_2}{\partial y_1} > 0 \iff f_{12} > 0;$  and  $\frac{\partial i_2}{\partial y_2} < 0 \iff f_{12} > 0;$
- (iii)  $\frac{\partial h_3}{\partial y_1} > 0;$  and  $\frac{\partial h_3}{\partial y_2} < 0.$

*III. If borrowing constraints bind during both early and late childhood, then:*

- (i)  $\frac{\partial i_1}{\partial y_1} > 0;$  and  $\frac{\partial i_1}{\partial y_2} > 0 \iff$  Condition 2 holds;
- (ii)  $\frac{\partial i_2}{\partial y_1} > 0 \iff$  Condition 2 holds; and  $\frac{\partial i_2}{\partial y_2} > 0;$
- (iii)  $\frac{\partial h_3}{\partial y_1} > 0$  and  $\frac{\partial h_3}{\partial y_2} > 0.$

There are two key implications of this proposition. First, if the late constraint binds, while the early borrowing constraint is non-binding, investments depend only on the discounted present value of family income  $y_1 + R^{-1}y_2$ , not the timing of income (conditional on discounting  $y_2$ ). Second, when the early constraint binds, both the timing of income and the extent of dynamic complementarity are important factors determining the response of investments to changes in income. For example, early investment is always increasing in  $y_1$ ; however, it is decreasing in  $y_2$  if only the early constraint binds because a late increase in income exacerbates this constraint. When only the early constraint binds, the impacts of income on late investment depend entirely on its effect on early investment and whether early investment raises or lowers the marginal return to late investment. When constraints are binding throughout childhood, increases in

---

<sup>36</sup>See Cunha and Heckman (2007) for a related analysis of the impacts of early vs. late income on the early-to-late investment ratio  $i_1/i_2$  when the early borrowing constraint binds.

income in any period increase investment in both periods if and only if there is sufficient dynamic complementarity.

To better understand the implications of policies aimed at expanding credit for educational investments, we consider the impacts of raising borrowing limits for families at different stages of child development, beginning with limits faced by families with older children (e.g. expanding student loan programs for higher education).

**Proposition 10.** *Assume that the borrowing constraint binds during late childhood (i.e.  $a_3 = -L_2$ ).*

(i) *If the early borrowing constraint does not bind (i.e.  $a_2 > -L_1$ ), then:  $\frac{\partial i_1}{\partial L_2} > 0$ ,  $\frac{\partial i_2}{\partial L_2} > 0$ , and  $\frac{\partial h_3}{\partial L_2} > 0$ .*

(ii) *If the early borrowing constraint also binds (i.e.  $a_2 = -L_1$ ), then:  $\frac{\partial i_1}{\partial L_2} > 0$  if Condition 2 holds;  $\frac{\partial i_2}{\partial L_2} > 0$ ; and  $\frac{\partial h_3}{\partial L_2} > 0$ .*

Relaxing the borrowing constraint during late childhood unambiguously increases late investment. If the early constraint is non-binding or if early and late investments are sufficiently complementary, then any increase in late investment encourages additional early investment as well. Even in the case of strong intertemporal substitutability when early investment may decline, individuals acquire more adult human capital when the late constraint is relaxed.

Next, consider expanded borrowing opportunities for families with young children (e.g. loans for preschool).

**Proposition 11.** *Assume that the borrowing constraint binds during early childhood (i.e.  $a_2 = -L_1$ ).*

(i) *If no other borrowing constraint binds, then:  $\frac{\partial i_1}{\partial L_1} \in (0, 1)$ ;  $\frac{\partial i_2}{\partial L_1} > 0 \iff f_{12} > 0$ ; and  $\frac{\partial h_3}{\partial L_1} > 0$ .*

(ii) *If the late borrowing constraint also binds (i.e.  $a_3 = -L_2$ ) and Condition 2 does not hold, then:  $\frac{\partial i_1}{\partial L_1} > 0$  and  $\frac{\partial i_2}{\partial L_1} < 0$ .*

When individuals are only constrained during early childhood, relaxing that constraint leads to an increase in early investment, which encourages late investment as long as the marginal productivity of  $i_2$  is increasing in  $i_1$ .

When children are constrained in both periods, relaxing the early constraint effectively shifts resources from late to early childhood. If early and late investments are very complementary, they will both tend to move in the same direction. In most cases, investments will increase; however, it is possible that investments could actually *decrease* in both periods. Intuitively, if late investment is very productive, then relaxing the early borrowing constraint can ‘starve’ that

investment. By contrast, if investments are sufficiently substitutable over time, shifting resources from late to early childhood by relaxing the early constraint causes investment to shift from the late to the early period as well.

The stylized facts of Section 3 and other evidence in Caucutt and Lochner (2012) are most consistent with binding early and late constraints and sufficient dynamic complementarity (case III of Proposition 9). In this case, investments increase with additional family income (Fact 3) and the timing of income matters (Fact 4). Because poor children are more likely to be borrowing constrained, Proposition 8 implies that they are likely to have marginal labor market returns that exceed the return to savings as well as the marginal returns for unconstrained children from higher income families (Facts 1 and 2). Finally, Propositions 10 and 11 suggest that policies designed to expand borrowing opportunities (at either stage of child development) can raise the investment and skill levels of children from constrained (i.e. low-income) families, improving both efficiency and equity.

## 5 Conclusions

It is well-known that poor children perform much worse academically and on achievement tests than their more economically advantaged counterparts. The most immediate explanation for these differences is that poor parents invest less in their young children. As we document, poor parents have fewer books in the home, read less to their young children, engage in fewer lessons and extracurricular activities, etc. Important differences in investment activities and achievement by family income remain even after controlling for maternal characteristics like education, achievement, and race. In this paper, we ask the next logical question: why do poor parents invest so much less in their children?

While there are many competing theories for these investment and skill gaps, few studies attempt to sort amongst them.<sup>37</sup> We systematically study four leading investment-based theories/mechanisms thought to drive income-based skill gaps: an intergenerational correlation in ability, a consumption value of investment, information frictions, and credit constraints. In order to help understand which mechanisms drive family investments in children, we consider the extent to which they also explain other important stylized facts related to the marginal returns to investment and the effects of parental income on child investment and skills.

The main lessons from our theoretical analysis are summarized in Table 1, which shows considerable differences in the extent to which each mechanism explains important stylized facts about child development. While a positive intergenerational correlation in ability may be partially

---

<sup>37</sup>Cunha (2014) is an important recent exception.

Table 1: Summary of Results

	Intergen. Ability Correlation	Consum. Value ( $\nu > 0$ )	Uncertainty w/Risk Aversion	Poor more subjective uncertainty, risk neutrality, $i_1$ irreversible		Poor have biased beliefs		
				Uncertainty in $\theta$	Uncertainty in $z$	Systematic downward bias	Non-Systematic bias	Credit Constraints
High MR to $i_1$ for Poor	No	No	Yes	No <sup>†</sup>	Yes <sup>†‡</sup>	Yes <sup>‡</sup>	No <sup>‡</sup>	Yes
Lower MR to $i_1$ for Rich	No	Yes	Yes	No <sup>†</sup>	Yes <sup>†‡</sup>	Yes <sup>‡</sup>	No <sup>‡</sup>	Yes
Increase in Income causes an increase in $i_1$	No	Yes	Yes	No	Only if info. changes with income <sup>‡</sup>	Only if info. changes with income <sup>‡</sup>	No	Yes
Timing of Income	No	No	No	No	No	No	No	Yes

Notes: Where relevant, results in the table assume gross complementarity, i.e.  $f_{12} \geq 0$ . <sup>†</sup> Under purely objective uncertainty, expected marginal returns equal return to savings for everyone. <sup>‡</sup> Assumes that Condition 1 holds or that  $b \geq 0$  in the case of CES production function.

responsible for the relationship between family income and child investment and achievement, it is not helpful for understanding any of the other important stylized facts. A theory based only on a positive consumption value of investment can explain the positive causal effects of income on investment as well as decreasing marginal labor market returns in family income; however, it predicts over-investment in skills such that the labor market returns to investment should be less than the return to savings. This mechanism offers no explanation for the importance of early income relative to late income.

Uncertainty coupled with risk aversion leads to under-investment in human capital and high marginal returns to additional investment. With decreasing absolute risk aversion, investment disincentives are greater for the poor. Thus, this mechanism can explain the qualitative patterns for marginal returns documented in the literature as well as the evidence on causal effects of income on early child investments and achievement. Neither this mechanism nor any other information-based explanation we explore can explain why the timing of income is important. Even in the absence of risk aversion, subjective uncertainty in the productivity of early investments can lead to under-investment and high marginal returns due to the irreversibility of investments. If poor families face greater subjective uncertainty than rich families, then predicted patterns for marginal returns are consistent with empirical evidence. This is also true if poor families simply under-estimate the productivity of early investments compared to higher income families. Unless changes in income directly improve the information of poor families, these misinformation problems only generate a correlation between family income and investment; they cannot explain why changes in income produce changes in investment or achievement.

The inability of poor families to borrow against future income can lead to under-investment in their children, which can further explain high marginal returns to investment among the poor. For children in constrained families, improvements in income lead to increases in investment and higher skill levels. If constraints are binding for families with young children, the timing of income will be important. Thus, binding credit constraints are consistent with all of the stylized facts we consider.

We caution that our comparison of model predictions with the evidence should not be taken as a score sheet, evaluating the importance of each mechanism by the number of facts it explains. For example, a positive intergenerational correlation in ability is almost certainly important given the extent to which maternal characteristics help explain income-based differences in investment and achievement (see Figures 1 and 5); yet, it offers no explanation for any of the other stylized facts. Instead, our results help identify the limits of different mechanisms. For example, only borrowing constraints are capable of explaining why the timing of income may be important. Information or credit market frictions are needed to explain the high marginal returns to early investment among

the poor, suggesting that appropriately designed policies may be able to reduce inequality while improving economic efficiency. More generally, understanding the limits of different mechanisms should be helpful in refining current theories and developing new ones to provide a more complete understanding of the child development process.

## References

- Almond, D. and J. Currie (2011). Human capital development before age five. In *Handbook of Labor Economics*, Volume 4B, Chapter 15, pp. 1315–1486. Amsterdam: Elsevier.
- Baker, M. (2011). Innis lecture: universal early childhood interventions: What is the evidence base? *Canadian Journal of Economics* 44(4), 1069–1105.
- Baker, M., J. Gruber, and K. Milligan (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy* 116(4), 709–745.
- Barnett, W. S. and L. Masse (2007). Comparative benefit-cost analysis of the Abecedarian program and its policy implications. *Economics of Education Review* 26(1), 113–125.
- Becker, G. S. (1975). *Human Capital*, 2nd ed. New York, NY: Columbia University Press.
- Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87(6), 1153–1189.
- Becker, G. S. and N. Tomes (1986). Human capital and the rise and fall of families. *Journal of Labor Economics* 4(3), S1–S39.
- Belley, P. and L. Lochner (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human Capital* 1(1), 37–89.
- Blau, D. and J. Currie (2006). Pre-school, day care, and after-school care: Who’s minding the kids? In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 2, Chapter 20, pp. 1163–1278. Amsterdam: Elsevier.
- Brooks-Gunn, J., R. T. Gross, H. C. Kraemer, D. Spiker, and S. Shapiro (1992). Enhancing the cognitive outcomes of low birth weight, premature infants: For whom is the intervention most effective? *Pediatrics* 89(6), 1209–1215.
- Cameron, S. and J. J. Heckman (1998, April). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy* 106(2), 262–333.

- Campbell, F. A., E. P. Pungello, M. Burchinal, K. Kainz, Y. Pan, B. H. Wasik, O. A. Barbarin, J. J. Sparling, and C. T. Ramey (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian project follow-up. *Developmental Psychology* 48(4), 1033.
- Carneiro, P. and R. Ginja (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy* 6(4), 135–173.
- Carneiro, P., J. Heckman, and E. Vytlačil (2011). Estimating marginal returns to education. *American Economic Review* 101(6), 2754–81.
- Carneiro, P. and J. J. Heckman (2002). The evidence on credit constraints in post-secondary schooling. *Economic Journal* 112(482), 705–734.
- Caucutt, E. and L. Lochner (2012). Early and late human capital investments, borrowing constraints, and the family. NBER Working Paper No. 18493.
- Caucutt, E. and L. J. Lochner (2006). Borrowing constraints on families with young children. In *Innovation in Education*, pp. 39–48. Cleveland: Federal Reserve Bank of Cleveland.
- Cunha, F. (2013). Investments in children when markets are incomplete. Working Paper, University of Pennsylvania.
- Cunha, F. (2014). Gaps in early investments in children. Working Paper, University of Pennsylvania.
- Cunha, F., I. Elo, and J. Culhane (2013). Eliciting maternal expectations about the technology of cognitive skill formation. NBER Working Paper No. 19144.
- Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review* 97(2), 31–47.
- Cunha, F., J. Heckman, and S. Navarro (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers* 57(2), 191–261.
- 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.
- Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, Chapter 12, pp. 697–812. Amsterdam: Elsevier.

- Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78(3), 883–931.
- Currie, J. and D. Thomas (1995). Does Head Start make a difference? *American Economic Review* 85(3), 341–364.
- Currie, J. and D. Thomas (1999). Does Head Start help hispanic children? *Journal of Public Economics* 74(2), 235–262.
- 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), 1927–1956.
- Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies* 81(1), 137–185.
- Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics* 1(3), 111–34.
- Dizon-Ross, R. (2014). Parents’ perceptions and children’s education: Experimental evidence from Malawi. Working Paper.
- Duncan, G. and J. Brooks-Gunn (1997). *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.
- Duncan, G., P. Morris, and C. Rodrigues (2011). Does money really matter? Estimating impacts of family income on young children’s achievement with data from random-assignment experiments. *Developmental Psychology* 47(5), 1263–1279.
- Duncan, G., W.-J. Yeung, J. Brooks-Gunn, and J. Smith (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review* 63(3), 406–423.
- Duncan, G. J. and A. J. Sojourner (2013). Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps? *Journal of Human Resources* 48(4), 945–968.
- Garces, E., J. Currie, and D. Thomas (2002). Longer-term effects of Head Start. *American Economic Review* 92(4), 999–1012.
- Gertler, P., J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* 344(6187), 998–1001.

- Guryan, J., E. Hurst, and M. Kearney (2008). Parental education and parental time with children. *Journal of Economic Perspectives* 22(3), 23–46.
- Hartman, R. (1976). Factor demand with output price uncertainty. *American Economic Review* 66(4), 675–681.
- Havnes, T. and M. Mogstad (2014). Is universal child care leveling the playing field? *Journal of Public Economics*.
- Heckman, J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The rate of return to the high/scope perry preschool program. *Journal of Public Economics* 94(1-2), 114–128.
- Heckman, J. J. and T. Kautz (2014). Fostering and measuring skills interventions that improve character and cognition. In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), *The Myth of Achievement Tests: the GED and the Role of Character in American Life*, Chapter 9, pp. 341–430. Chicago: University of Chicago Press.
- Karoly, L. et al. (1998). *Investing In Our Children: What We Know And Don't Know About The Costs And Benefits Of Early Childhood Interventions*. Santa Monica: The RAND Corporation.
- Kaushal, N., K. Magnuson, and J. Waldfogel (2011). How is family income related to investments in children's learning? In G. Duncan and R. Murnane (Eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, Chapter 9, pp. 187–206. New York: Russell Sage Foundation.
- Kautz, T., J. J. Heckman, R. Diris, B. ter Weel, and L. Borghans (2014). *Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success*. OECD.
- Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy* 105(3), 473–522.
- Keane, M. P. and K. I. Wolpin (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review* 42(4), 1051–1103.
- Kottelenberg, M. and S. Lehrer (2014). Reinvestigating who benefits and who loses from universal childcare in Canada. Working Paper, Queens University.
- Lazear, E. (1977). Education: Consumption or production? *Journal of Political Economy* 85(3), 569–598.
- Lee, S. Y. and A. Seshadri (2014). On the intergenerational transmission of economic status. Working Paper.

- Levhari, D. and Y. Weiss (1974). The effect of risk on the investment in human capital. *American Economic Review* 64(6), 950–963.
- Levy, D. and G. Duncan (1999). Using sibling samples to assess the effect of childhood family income on completed schooling. Working Paper.
- Løken, K. (2010). Family income and children’s education: Using the norwegian oil boom as a natural experiment. *Labour Economics* 17(1), 118–129.
- Løken, K., M. Mogstad, and M. Wiswall (2012). What linear estimators miss: The effects of family income on child outcomes. *American Economic Journal: Applied Economics* 4(2), 1–35.
- Ludwig, J. and D. Miller (2007). Does Head Start improve children’s life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics* 122(1), 159–208.
- Milligan, K. and M. Stabile (2011). Do child tax benefits affect the well-being of children? Evidence from canadian child benefit expansions. *American Economic Journal: Economic Policy* 3(3), 175–205.
- Oreopoulos, P., M. E. Page, and ann Huff Stevens (2008). The intergenerational effects of worker displacement. *Journal of Labor Economics* 26(3), 455–83.
- Pavan, R. (2014). On the production of skills and the birth order effect. Working Paper, Royal Holloway, University of London.
- Puma, M., S. Bell, R. Cook, C. Heid, P. Broene, F. Jenkins, A. Mashburn, and J. Downer (2012). Third grade follow-up to the Head Start Impact Study: Final report. Technical report, Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.
- Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. Duncan and R. Murnane (Eds.), *Whither Opportunity? Rising Inequality, Schools, and Children’s Life Chances*, Chapter 5, pp. 91–116. New York: Russell Sage Foundation.
- Reynolds, A. J., J. A. Temple, S.-R. Ou, I. A. Arteaga, and B. A. White (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science* 333(6040), 360–364.

- Reynolds, A. J., J. A. Temple, B. A. White, S.-R. Ou, and D. L. Robertson (2011). Age 26 cost-benefit analysis of the child-parent center early education program. *Child development* 82(1), 379–404.
- Rothschild, M. and J. E. Stiglitz (1970). Increasing risk: I. A definition. *Journal of Economic Theory* 2(3), 225 – 243.
- Rothschild, M. and J. E. Stiglitz (1971). Increasing risk II: Its economic consequences. *Journal of Economic Theory* 3(1), 66 – 84.
- Stevens, A. H. and J. Schaller (2011). Short-run effects of parental job loss on children’s academic achievement. *Economics of Education Review* 20(3), 289–299.
- Thomson, G. H. (1951). *The Factorial Analysis of Human Ability*. London: University of London Press.
- Todd, P. 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.

## Online Appendix

### A Factor Score Weights

We employ principal factor analysis using the measured inputs reported in Figures 2 and 3 (separately by age) to create age-specific predicted investment factor scores for each child using the Thomson (1951) method. Estimated weights for each factor used in constructing the scores are reported in Table A. For interpretation purposes, scores are normalized to have a mean of zero and standard deviation of one.

Table A: Weights used to Construct Factor Scores

Early Investment Measure	Age Group			
	0-1	2-3	4-5	6-7
10+ Books at Home	0.33	0.24	0.21	0.12
Mom Reads 3+ Times/Week	0.32	0.26	0.22	0.09
Child Eats with Mom and Dad Daily	0.08	0.16	0.19	0.03
Child Leaves House 4+ Times/Week	0.18	0.14	0.15	
Child Sees Father Daily	0.09	0.20	0.24	
Musical Instrument in Home				0.10
Child Taken to a Performance in Past Year				0.18
Child Taken to a Museum in Past Year				0.15
Child Takes Special Lessons/Extracurricular Activities				0.14
Family Gets a Daily Newspaper				0.08
Family Encourages Hobbies				0.08
Family Meets Friends/Relatives 2+ Times/Month				0.02

### B Proofs and Technical Results

#### Proof of Proposition 1

Part (i): Follows directly from the first order conditions (6) and (7).

Parts (ii) and (iii): Applying Cramer's rule to first order conditions (6) and (7) yields

$$\frac{di_1}{d\theta} = \frac{f_2(i_1, i_2)f_{12}(i_1, i_2) - f_1(i_1, i_2)f_{22}(i_1, i_2)}{\theta[f_{11}f_{22} - f_{12}^2]} > 0$$

$$\frac{di_2}{d\theta} = \frac{f_1(i_1, i_2)f_{12}(i_1, i_2) - f_2(i_1, i_2)f_{11}(i_1, i_2)}{\theta[f_{11}f_{22} - f_{12}^2]} > 0,$$

where these effects can be signed using Assumption 1. These imply that the final human capital level,  $h_3$ , is increasing in  $\theta$ .

## Proof of Proposition 2

Part (i): The difference between marginal returns to early investment and the return to savings is immediate from Equation (8).

Consider the effects of parental income on the marginal value of investment,  $\frac{\partial h_3}{\partial i_1} = \theta f_1(i_1, i_2)$ . Differentiating  $f_1(i_1, i_2)$  with respect to  $Y$  at an optimum yields:

$$\begin{aligned} \frac{df_1(i_1, i_2)}{dY} &= f_{11} \frac{di_1}{dY} + f_{12} \frac{di_2}{dY} \\ &= \left[ \frac{\nu u''(c)}{[u'(c)]^2 B} \right] [f_{11}(i_1, i_2) f_{22}(i_1, i_2) - f_{12}^2(i_1, i_2)] / \Delta, \end{aligned}$$

where  $\Delta = \left[ \frac{\nu^2 u''(c)}{[u'(c)]^3 B} \right] f_{22} + \beta^2 \theta [f_{11} f_{22} - f_{12}^2] > 0$  by Assumption 1.  $\frac{df_1(i_1, i_2)}{dY}$  has the opposite sign as  $\nu$ , because the second term in brackets is positive by Assumption 1.

Parts (ii) and (iii): We can determine the effects of  $Y$  on child investments from first order conditions (7) and (8) using Cramer's rule:

$$\begin{aligned} \frac{di_1}{dY} &= \left[ \frac{\nu u''(c)}{[u'(c)]^2 B} \right] f_{22}(i_1, i_2) / \Delta \\ \frac{di_2}{dY} &= \left[ \frac{-\nu u''(c)}{[u'(c)]^2 B} \right] f_{12}(i_1, i_2) / \Delta. \end{aligned}$$

If investment has no consumption value (i.e.  $\nu = 0$ ), then investments are independent of parental income  $Y$ . Otherwise, the effects of parental income on investments depend on whether investment has a positive (i.e.  $\nu > 0$ ) or negative (i.e.  $\nu < 0$ ) consumption value and, in the case of second period investment, the complementarity of early and late investments  $f_{12}(i_1, i_2)$ .

Part (iv): The effect of parental income on final human capital levels is given by

$$\begin{aligned} \frac{dh_3}{dY} &= \theta \left[ f_1 \frac{di_1}{dY} + f_2 \frac{di_2}{dY} \right] \\ &= \theta \left[ \frac{\nu u''(c)}{[u'(c)]^2 B} \right] [f_1(i_1, i_2) f_{22}(i_1, i_2) - f_2(i_1, i_2) f_{12}(i_1, i_2)] / \Delta, \end{aligned}$$

which has the same sign as  $\nu$ , because the second term in brackets is negative by Assumption 1.

## Proof of Proposition 3

When the return to investment is realized after all investments have been made, we can combine the period 1 and 2 budget constraints:

$$\underbrace{(c_1 + R^{-1}c_2)}_{\equiv c} + \underbrace{(i_1 + R^{-1}i_2)}_{\equiv e} + R^{-1}a_3 \leq Y.$$

Notice that only discounted lifetime income  $Y$  enters the budget constraint, so the timing of income,  $y_1$  and  $y_2$ , is irrelevant.

The problem can be decomposed into 3 pieces: investment allocation between periods 1 and 2 (choosing  $i_1$  and  $i_2$  given  $e$ ), consumption allocation between periods 1 and 2 (choosing  $c_1$  and  $c_2$  given  $c$ ), and portfolio choice between human capital investment and savings (choosing  $c$ ,  $e$ , and  $a_3$ ). In proving our results, we first show how  $i_1$  and  $i_2$  (and their marginal returns) depend on the level of total investment spending  $e = i_1 + R^{-1}i_2$  (and its marginal return) and then focus on properties of  $e$  and its marginal return.

First, consider the investment problem for a given level of total investment expenditure  $e$ . Let  $g(e)$  be the maximum  $f(i_1, i_2)$  produced by spending  $e$ :

$$g(e) \equiv \max_{i_1, i_2} \left\{ f(i_1, i_2) \mid i_1 + R^{-1}i_2 \leq e \right\}.$$

Note that  $g(\cdot)$  is strictly increasing and strictly concave because  $f(\cdot)$  is strictly increasing, strictly concave, and the constraint set is convex. The first order conditions for  $i_1$  and  $i_2$ , combined with the envelope condition, are

$$g'(e) = f_1(i_1, i_2) = Rf_2(i_1, i_2), \quad i_1 + R^{-1}i_2 = e. \quad (17)$$

Applying the Implicit Function Theorem, we can determine the effects of  $e$  on investments:

$$\frac{di_1}{de} = \frac{f_{12} - \frac{f_1}{f_2}f_{22}}{2f_{12} - \frac{f_2}{f_1}f_{11} - \frac{f_1}{f_2}f_{22}}, \quad \frac{di_2}{de} = \frac{\frac{f_1}{f_2} \left( f_{12} - \frac{f_2}{f_1}f_{11} \right)}{2f_{12} - \frac{f_2}{f_1}f_{11} - \frac{f_1}{f_2}f_{22}},$$

where we use the first order condition  $R = f_1/f_2$ . Both derivatives are strictly positive by Assumption 1, so spending more on overall education means that both early and late investments increase. We state this result as a Lemma for future reference.

**Lemma 3.** *Both early and late investments are strictly increasing in total investment spending if they are optimally chosen to maximize human capital.*

Next, we consider the intertemporal consumption allocation problem. Define the indirect utility of total consumption spending  $c$  on consumption in periods 1 and 2:

$$U(c) \equiv \max_{c_1, c_2} \left\{ u(c_1) + \beta u(c_2) \mid c_1 + R^{-1}c_2 \leq c \right\}.$$

Note that  $U(\cdot)$  is strictly increasing and strictly concave, because  $u(\cdot)$  is strictly increasing, strictly concave, and the constraint set is convex.

Using the indirect functions  $U(\cdot)$  and  $g(\cdot)$ , we can write the saving and human capital investment problem as follows:

$$\max_{e, s} \left\{ U(Y - s) + \beta^2 E \left[ u(R^2(s - e) + \theta g(e)) \right] \right\},$$

where  $s = e + R^{-1}a_3$  is total savings.

The first order conditions are

$$U'(e) = (R\beta)^2 E[u'(c_3(\theta))] = \beta^2 E[u'(c_3(\theta))\theta]g'(e), \quad (18)$$

where  $c_3(\theta) = R^2(s - e) + \theta g(e)$ .

Part (i): Combining the first order conditions in (18), we get

$$R^2 = \frac{E[u'(c_3(\theta))\theta]}{E[u'(c_3(\theta))]}g'(e) = \bar{\theta}g'(e) + \frac{\text{Cov}(u'(c_3(\theta)), \theta)}{E[u'(c_3(\theta))]}g'(e).$$

Because  $c_3(\theta)$  is strictly increasing in  $\theta$  and  $u(\cdot)$  is strictly concave, the covariance term is strictly negative. This, combined with Equation (17), implies that the expected marginal returns to  $i_1$  and  $i_2$  are greater than the returns on savings.

Parts (ii) and (iii): We first study how the portfolio choice between the safe asset  $a_3$  and the risky human capital investment  $e$  is affected by total savings  $s$ . Specifically, we show that optimal total investment  $e$  is increasing in total savings  $s$  under decreasing absolute risk aversion. We then show that total savings  $s$  is increasing in lifetime parental income  $Y$ .

Optimal human capital investment for a given  $s$  solves

$$\hat{e}(s) \equiv \underset{e}{\operatorname{argmax}} E[u(R^2(s - e) + \theta g(e))].$$

Because the objective function is strictly concave in  $e$ , the Implicit Function Theorem reveals that the sign of  $\hat{e}'(s)$  is equal to the sign of

$$\frac{\partial^2}{\partial e \partial s} \left( E[u(R^2(s - e) + \theta g(e))] \right) \Big|_{e=\hat{e}(s)} = E[u''(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))(-R^2 + \theta g'(\hat{e}(s)))R^2].$$

With decreasing absolute risk aversion, this term is strictly positive and  $\hat{e}'(s) > 0$ . To see this, let  $\hat{\theta} = R^2/g'(\hat{e}(s))$ , so

$$\begin{aligned} & E[u''(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))(-R^2 + \theta g'(\hat{e}(s)))R^2] \\ &= R^2 \int \left[ \frac{u''(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))}{u'(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))} \right] u'(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))(-R^2 + \theta g'(\hat{e}(s)))d\Phi(\theta) \\ &> R^2 \left[ \frac{u''(R^2(s - \hat{e}(s)) + \hat{\theta}g(\hat{e}(s)))}{u'(R^2(s - \hat{e}(s)) + \hat{\theta}g(\hat{e}(s)))} \right] \left\{ \int_{\theta \leq \hat{\theta}} u'(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))(-R^2 + \theta g'(\hat{e}(s)))d\Phi(\theta) \right. \\ &\quad \left. + \int_{\theta \geq \hat{\theta}} u'(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))(-R^2 + \theta g'(\hat{e}(s)))d\Phi(\theta) \right\} \\ &= R^2 \left[ \frac{u''(R^2(s - \hat{e}(s)) + \hat{\theta}g(\hat{e}(s)))}{u'(R^2(s - \hat{e}(s)) + \hat{\theta}g(\hat{e}(s)))} \right] E[u'(R^2(s - \hat{e}(s)) + \theta g(\hat{e}(s)))(-R^2 + \theta g'(\hat{e}(s)))] = 0, \end{aligned}$$

where the inequality used the assumption that  $u''(\cdot)/u'(\cdot)$  is increasing.

Next, we show that total savings  $s$  is increasing with  $Y$ . Define the value of total savings  $s$ :

$$\mathcal{V}(s) \equiv E \left[ \max_{e,a} \left\{ u(R^2a + \theta g(e)) \mid e + a \leq s \right\} \right],$$

and let  $\hat{a}(s) \equiv s - \hat{e}(s)$  reflect the optimal investment in the safe asset conditional on  $s$ .  $\mathcal{V}(s)$  is strictly increasing and strictly concave, because  $u(\cdot)$  is strictly increasing and strictly concave in  $(a, e)$  and the constraint set is convex.

Optimal total savings  $\hat{s}(Y)$  for a given income  $Y$  solves

$$\hat{s}(Y) \equiv \operatorname{argmax}_s \left\{ U(Y - s) + \beta^2 \mathcal{V}(s) \right\}.$$

Applying the Implicit Function Theorem to the first order condition for this problem yields

$$\hat{s}'(Y) = \frac{U''(Y - \hat{s}(Y))}{U''(Y - \hat{s}(Y)) + \beta^2 \mathcal{V}''(\hat{s}(Y))} > 0,$$

because both  $U(\cdot)$  and  $\mathcal{V}(\cdot)$  are strictly concave.

With decreasing absolute risk aversion, both  $\hat{s}(Y)$  and  $\hat{e}(s)$  are strictly increasing, so  $e$  is strictly increasing in  $Y$ . From (17), expected marginal returns to investment are equal to the expected marginal return to total investment spending:  $\bar{\theta} f_1(i_1, i_2) = \bar{\theta} R f_2(i_1, i_2) = \bar{\theta} g'(e)$ . This implies that the expected marginal returns to investment are strictly decreasing in  $Y$  (with decreasing absolute risk aversion), because  $g(\cdot)$  is strictly concave.

## Proof of Lemma 1

We first show the properties of  $\hat{i}_2(z i_1, \theta)$ . The first order condition is:

$$\theta f_2(z i_1, \hat{i}_2(z i_1, \theta)) = R.$$

By applying the Implicit Function Theorem, we get the partial derivatives:

$$\frac{\partial \hat{i}_2(z i_1, \theta, z)}{\partial \theta} = \frac{f_2}{-\theta f_{22}} > 0, \quad \frac{\partial \hat{i}_2(z i_1, \theta, z)}{\partial (z i_1)} = \frac{f_{12}}{-f_{22}} \geq 0 \text{ if and only if } f_{12} \geq 0. \quad (19)$$

Then the partial derivatives of  $\Pi(i_1, \theta, z)$  are

$$\begin{aligned} \frac{\partial \Pi(i_1, \theta, z)}{\partial i_1} &= -1 + R^{-2} z \theta f_1, \\ \frac{\partial^2 \Pi(i_1, \theta, z)}{\partial i_1^2} &= R^{-2} z \theta \left( f_{11} z + f_{12} \frac{\partial \hat{i}_2(z i_1, \theta)}{\partial i_1} \right) = -R^{-2} z^2 \theta \left( \frac{f_{11} f_{22} - f_{12}^2}{-f_{22}} \right) < 0. \end{aligned}$$

## Proof of Lemma 2

The cross-partial derivative is

$$\frac{\partial^2 \Pi(i_1, \theta, z)}{\partial i_1 \partial z} = R^{-2} \theta \left( f_1 + z f_{11} i_1 + z f_{12} \frac{\partial \hat{i}_2(z i_1, \theta)}{\partial z} \right) = R^{-2} \theta \left\{ f_1 - \left( \frac{f_{11} f_{22} - f_{12}^2}{-f_{22}} \right) z i_1 \right\},$$

which is positive if and only if Condition 1 holds.

## Proof of Proposition 4

See Hartman (1976) for proof.

## Proof of Proposition 5

Part (i): A first order stochastic dominance shift in  $z$  will increase  $\tilde{i}_1$  if  $\partial \Pi(i_1, \theta, z) / \partial i_1$  is increasing in  $z$ . For the CES production function (3), the sign of  $\partial^2 \Pi(i_1, z) / (\partial i_1 \partial z)$  is determined by the sign of

$$b \left( \frac{1-d}{1-b} \right) (1-a)^{1-b} (z i_1)^{1-b} i_2 + a^{1-b} d (z i_1) i_2^{1-b},$$

which is positive if  $b \geq 0$ .

Part (ii): A mean-preserving spread in  $z$  reduces  $\tilde{i}_1$  if  $\partial \Pi(i_1, \theta, z) / \partial i_1$  is concave in  $z$ . For the CES production function (3), the sign of  $\partial^3 \Pi(i_1, z) / (\partial i_1 \partial^2 z)$  depends on the sign of

$$- \left\{ b \left( \frac{1-d}{1-b} \right)^2 \left( \frac{a}{1-a} \right)^{b-1} \left( \frac{z i_1}{i_2} \right)^{-b} + d \left( \frac{a}{1-a} \right)^{1-b} \left( \frac{z i_1}{i_2} \right)^b + \left( \frac{1-d}{1-b} \right) \left( \frac{1-d}{1-b} - (1-2d) \right) \right\},$$

which is negative if  $b \geq 0$ .

## Proof of Corollary 1

Follows directly from Lemma 1 and Propositions 4 and 5.

## Proof of Proposition 6

Part (i) follows from Lemma 1 and Lemma 2.

Part (ii):  $\tilde{\mathbf{x}} \leq \mathbf{x}$  and  $\tilde{h}_2 \leq h_2^*$  follows immediately from  $\tilde{x}_j = w_j \tilde{i}_1$  and  $\tilde{h}_2 = z \tilde{i}_1$ .  $\tilde{h}_2 \leq h_2^*$  also implies  $\tilde{h}_3 \leq h_3^*$ . To show this, we differentiate  $\tilde{h}_3$  with respect to  $\tilde{h}_2$  to get

$$\frac{d \tilde{h}_3}{d \tilde{h}_2} = \frac{d}{d \tilde{h}_2} \left( \theta f(\tilde{h}_2, \hat{i}_2(\tilde{h}_2, \theta)) \right) = \theta \left( f_1 + f_2 \frac{\partial \hat{i}_2(\tilde{h}_2, \theta)}{\partial \tilde{h}_2} \right) = \theta \left( f_1 - f_2 \frac{f_{12}}{f_{22}} \right) > 0,$$

where we use Equation (19), and the inequality holds by Assumption 1. We state this result for future reference.

**Lemma 4.**  $\theta f(h_2, \hat{i}_2(h_2, \theta))$  is strictly increasing in  $h_2$ .

Next, the marginal return to early investment  $z\theta f_1(z\tilde{i}_1, \hat{i}_2(z\tilde{i}_1, \theta))$  is strictly decreasing in  $\tilde{i}_1$  by Lemma 1. So we have

$$z\theta f_1(z\tilde{i}_1, \hat{i}_2(z\tilde{i}_1, \theta)) \geq z\theta f_1(z\tilde{i}_1^*, \hat{i}_2(z\tilde{i}_1^*, \theta)) = R^2$$

if and only if  $\tilde{i}_1 \leq \tilde{i}_1^*$ .

### Proof of Proposition 7

Because  $\tilde{i}_1$  depends only on  $\tilde{z}$ ,  $\tilde{i}_1 = \tilde{i}_1^*$  holds when  $\tilde{z} = z$ . From the conditional demand function  $\tilde{x}_j = \tilde{w}_j \tilde{i}_1 = \tilde{w}_j \tilde{i}_1^*$ , we can see that  $\tilde{x}_j \leq x_j^*$  if and only if  $\tilde{w}_j \leq w_j$ .  $\tilde{h}_2 < h_2^*$  directly follows from  $\tilde{i}_1 = \tilde{i}_1^*$  and the definition of optimization, but we explicitly prove this by showing that  $\tau(\tilde{\mathbf{w}}) < 1$ .

$$\tau(\tilde{\mathbf{w}})^\phi = \sum_{j=1}^n w_j \left( \frac{\tilde{w}_j}{w_j} \right)^\phi < \left( \sum_{j=1}^n w_j \left( \frac{\tilde{w}_j}{w_j} \right) \right)^\phi = \sum_{j=1}^n \tilde{w}_j = 1,$$

where the inequality holds due to Jensen's inequality and  $\phi \in (0, 1)$ .

As shown in Lemma 4,  $\tilde{h}_2 < h_2^*$  also implies  $\tilde{h}_3 < h_3^*$ .

Finally, the marginal return to  $i_1$  can be written as

$$z\tau(\tilde{\mathbf{w}})\theta f_1(\tilde{h}_2, \hat{i}_2(\tilde{h}_2, \theta)) = z\tau(\tilde{\mathbf{w}})\theta f_1\left(z\tau(\tilde{\mathbf{w}})\tilde{i}_1^*, \hat{i}_2(z\tau(\tilde{\mathbf{w}})\tilde{i}_1^*, \theta)\right).$$

Because  $z\tau(\tilde{\mathbf{w}}) < z$ , this quantity is smaller than  $R^2 = z\theta f_1(z\tilde{i}_1^*, \hat{i}_2(z\tilde{i}_1^*, \theta))$  if Condition 1 holds for all values of productivity  $z'$  between  $z\tau(\tilde{\mathbf{w}})$  and  $z$  (Lemma 2).

### Simplifying Condition 2 with non-binding constraints during adulthood

Condition 2 simplifies nicely if borrowing constraints are non-binding throughout adulthood and the consumption intertemporal elasticity of substitution is constant. Notice that when constraints are non-binding, we can write  $V(a, h) = v(Ra + \chi h)$  where

$$v(z) = \max_{c_3, \dots, c_T} \sum_{j=3}^T \beta^{j-3} u(c_j) \text{ subject to } \sum_{j=3}^T R^{3-j} c_j = z. \quad (20)$$

With  $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ , it is straightforward to show that  $v(z) = \Psi \frac{z^{1-\sigma}}{1-\sigma}$  where  $\Psi$  is a positive constant that depends on  $\beta$ ,  $R$  and  $T$ .

With these assumptions, Condition 2 can be re-written as:

$$\frac{f_{12}f}{f_1 f_2} > -\frac{v''(-RL_2 + \chi h_3)\chi^2 h_3}{v'(-RL_2 + \chi h_3)\chi} = \frac{\sigma \chi h_3}{-RL_2 + \chi h_3}.$$

Taking the inverse of both sides produces the simplified condition in the text.

## Properties of $V(a_3, h_3)$

Individuals entering adulthood with human capital  $h_3$  and assets  $a_3$  allocate consumption and saving across their remaining life in the following way:

$$V(a_3, h_3) = \max_{c_3, \dots, c_T, a_4, \dots, a_T} \sum_{j=3}^T \beta^{j-3} u(c_j),$$

subject to budget constraints  $a_{j+1} = Ra_j + h_3\Gamma_j - c_j$  for  $j \in \{3, \dots, T\}$ , borrowing constraints  $a_{j+1} \geq -L_j$  for  $j \in \{3, \dots, T-1\}$ , and  $a_{T+1} = 0$ . Let  $\{\hat{c}_j(a_3, h_3)\}_{j=3}^T$  and  $\{\hat{a}_j(a_3, h_3)\}_{j=4}^T$  be the solution to the problem. We sometimes omit the arguments of the policy functions in order to simplify notations.

The value function can be rewritten as:

$$V(a_3, h_3) = u(Ra_3 + \Gamma_3 - \hat{a}_4(a_3, h_3)) + \sum_{j=4}^T \beta^{j-3} u(R\hat{a}_j(a_3, h_3) + h_3\Gamma_j - \hat{a}_{j+1}(a_3, h_3)).$$

Differentiate this with respect to initial human capital level,  $h_3$ :

$$V_2(a_3, h_3) = \sum_{j=3}^T \beta^{j-3} u'(\hat{c}_j) \Gamma_j + \sum_{j=3}^{T-1} \beta^{j-3} \frac{\partial \hat{a}_{j+1}}{\partial h_3} [\beta R u'(\hat{c}_{j+1}) - u'(\hat{c}_j)] = \sum_{j=3}^T \beta^{j-3} u'(\hat{c}_j) \Gamma_j > 0,$$

where the second equality uses the fact that

$$\frac{\partial \hat{a}_{j+1}}{\partial h_3} [\beta R u'(\hat{c}_{j+1}) - u'(\hat{c}_j)] = 0.$$

This follows from the FOC: when the borrowing constraint binds, the derivative  $\frac{\partial \hat{a}_{j+1}}{\partial h_3}$  is zero; and when it does not bind, the Euler equation is zero.

Differentiating  $V_2$  with respect to the initial asset level,  $a_3$ , yields:

$$V_{21}(a_3, h_3) = \sum_{j=3}^T \beta^{j-3} \Gamma_j u''(\hat{c}_j) \frac{\partial \hat{c}_j}{\partial a_3}.$$

If optimal consumption in every period increases with an increase in initial assets, then  $\frac{\partial \hat{c}_j}{\partial a_3} \geq 0$  and  $V_{21} \leq 0$  given strict concavity of  $u(\cdot)$ . If  $\frac{\partial \hat{c}_j}{\partial a_3} > 0$  for at least one  $j \in \{3, \dots, T\}$ , then  $V_{21} < 0$ .

We have two cases:

- I. Suppose no borrowing constraint ever binds. In this case, we can rewrite  $V(a_3, h_3)$  as  $v(ra_3 + \chi h_3)$ , where  $v$  is defined above. All of the proofs carry through with  $V_{21} = \chi v''$ ,  $V_{22} = \chi^2 v''$  and  $V_2 = \chi v'$ .
- II. Let  $j_b \in \{3, \dots, T-1\}$  be the first period that any borrowing constraint binds, i.e.  $a_{j_b+1} = -L_{j_b}$ . In this case, as soon as the constraint binds, an increase in initial assets has no

impact on consumption,  $\frac{\partial \hat{c}_j}{\partial a_3} = 0$ ,  $j \in \{j_b + 1, \dots, T\}$ . Prior to the binding constraint, the increase in initial assets is spread across periods, strictly increasing consumption in all,  $\frac{\partial \hat{c}_j}{\partial a_3} > 0$ ,  $j \in \{3, \dots, j_b\}$ . This must be the case in order for the Euler equations to hold in all periods:  $u'(c_j) = \beta R u'(c_{j+1})$ ,  $j \in \{3, \dots, j_b - 1\}$ . If any  $c_j$  increases, they all must. Note that even if the constraint immediately binds, the increase in initial assets will increase consumption in period 3, and therefore (minimally)  $\frac{\partial \hat{c}_j}{\partial a_3} > 0$ , so  $V_{21} < 0$ .

In a similar fashion:

$$V_{22}(a_3, h_3) = \sum_{j=3}^T \beta^{j-3} \Gamma_j u''(\hat{c}_j) \frac{\partial \hat{c}_j}{\partial (h_3)} < 0.$$

Here, if an individual is constrained in a period and the Euler equation does not hold with equality, the fact that income rises due to the increase in  $h_3$  implies that consumption will rise as well. Again, if the individual faces no future constraints we can rewrite  $V(a_3, h_3)$  as  $v(ra_3 + \chi h_3)$ . All of the proofs carry through with the  $V_{21} = \chi v''$ ,  $V_{22} = \chi^2 v''$  and  $V_2 = \chi v'$ .

## Proof of Proposition 8

Combining FOCs for assets we have:

$$u'(c_1) \geq (\beta R)^{j-1} u'(c_j), \quad j \in \{2, 3, \dots, T\}, \quad (21)$$

where inequalities are strict when the relevant borrowing constraint binds. We can write this as:

$$u'(c_j) \leq (\beta R)^{1-j} u'(c_1), \quad j \in \{2, 3, \dots, T\}. \quad (22)$$

Parts (i): Using equations (15) and (22) we have:

$$u'(c_1) = \beta^2 \theta f_1(i_1, i_2) \left[ \sum_{j=3}^T \beta^{j-3} \Gamma_j u'(c_j) \right] \leq \beta^2 \theta f_1(i_1, i_2) \left[ \sum_{j=3}^T \beta^{j-3} \Gamma_j u'(c_1) (\beta R)^{1-j} \right],$$

which implies  $\theta \chi f_1(i_1, i_2) \geq R^2$ , with strict inequality if any borrowing constraint binds, and equality if no borrowing constraint binds.

We next prove  $\tilde{i}_1 \leq i_1^*$ . Towards a contradiction, suppose  $\tilde{i}_1 > i_1^*$  and let  $\tilde{h}_3 = \theta f(\tilde{i}_1, \tilde{i}_2)$  and  $h_3^* = \theta f(i_1^*, i_2^*)$ . Then the followings hold.

Claim 1:  $\tilde{h}_3 \geq h_3^*$ .

$\tilde{h}_3 < h_3^*$ , together with  $\tilde{i}_1 > i_1^*$ , implies  $f_1(\tilde{i}_1, \tilde{i}_2) < R f_2(\tilde{i}_1, \tilde{i}_2)$ , which contradicts the first order condition  $f_1(\tilde{i}_1, \tilde{i}_2) \geq R f_2(\tilde{i}_1, \tilde{i}_2)$ . To show this, let  $(\underline{i}_1(h_3), \underline{i}_2(h_3))$  be the cost-minimizing investment profile to produce  $h_3$  units of human capital. That is,

$$(\underline{i}_1(h_3), \underline{i}_2(h_3)) \equiv \underset{i_1, i_2}{\operatorname{argmin}} \{i_1 + R^{-1} i_2 \mid \theta f(i_1, i_2) \geq h_3\}.$$

By Assumption 1, both  $\underline{i}_1(h_3)$  and  $\underline{i}_2(h_3)$  are strictly increasing in  $h_3$  (dual of Lemma 3). So  $\tilde{h}_3 < h_3^*$  implies  $\underline{i}_1(\tilde{h}_3) < \underline{i}_1(h_3^*) = i_1^*$  and we have  $\underline{i}_1(\tilde{h}_3) < \tilde{i}_1$  by the assumption  $i_1^* < \tilde{i}_1$ . Both investment profiles  $(\underline{i}_1(\tilde{h}_3), \underline{i}_2(\tilde{h}_3))$  and  $(\tilde{i}_1, \tilde{i}_2)$  produce the same amount of human capital  $\tilde{h}_3$ , but the former is efficient (because it achieves the minimum cost by definition) while the latter uses too much  $i_1$  and too little  $i_2$  compared to the former. Because the isoquant is strictly convex (Assumption 1 implies strict quasi-concavity of  $f(\cdot)$ , which in turn implies strict convexity of the isoquant), the movement along the isoquant in the direction of higher  $i_1$  and lower  $i_2$  decreases the marginal rate of technical substitution:

$$R = \frac{f_1(\underline{i}_1(\tilde{h}_3), \underline{i}_2(\tilde{h}_3))}{f_2(\underline{i}_1(\tilde{h}_3), \underline{i}_2(\tilde{h}_3))} > \frac{f_1(\tilde{i}_1, \tilde{i}_2)}{f_2(\tilde{i}_1, \tilde{i}_2)}.$$

The inequality  $f_1(\tilde{i}_1, \tilde{i}_2) < Rf_2(\tilde{i}_1, \tilde{i}_2)$  contradicts the first order condition  $f_1(\tilde{i}_1, \tilde{i}_2) \geq Rf_2(\tilde{i}_1, \tilde{i}_2)$ , so it must be  $\tilde{h}_3 \geq h_3^*$  if  $\tilde{i}_1 > i_1^*$ .

Claim 2:  $\tilde{i}_2 \geq i_2^*$ .

First, by Claim 1,  $\tilde{h}_3 \geq h_3^*$  and this implies  $\underline{i}_2(\tilde{h}_3) \geq \underline{i}_2(h_3^*) = i_2^*$  (Lemma 3). Second, the first order condition  $f_1(\tilde{i}_1, \tilde{i}_2) \geq Rf_2(\tilde{i}_1, \tilde{i}_2)$  implies  $\tilde{i}_1 \leq \underline{i}_1(\tilde{h}_3)$  and  $\tilde{i}_2 \geq \underline{i}_2(\tilde{h}_3)$  due to the strict convexity of the isoquant. From these two inequalities, we get  $\tilde{i}_2 \geq i_2^*$ .

Combining Claim 1 and Claim 2 together, we have  $\tilde{i}_1 > i_1^*$  and  $\tilde{i}_2 \geq i_2^*$ . This means that the income-maximizing investment  $(i_1^*, i_2^*)$  is affordable, but not chosen by the individual. This contradicts the assumption that  $(\tilde{i}_1, \tilde{i}_2)$  solves the individual's problem. For example, reducing investment to  $(i_1^*, i_2^*)$  and saving the rest to increase only the period 3 consumption is a better strategy, because over-investing in human capital makes the return lower than the return to savings. Thus, if  $\tilde{i}_1$  is the optimal early investment, then it must be smaller than  $i_1^*$ .

Next, we show that  $\tilde{i}_1 \leq i_1^*$  implies  $\tilde{h}_3 \leq h_3^*$ . Let  $\hat{h}_3(i_1) \equiv \theta f(i_1, \hat{i}_2(i_1, \theta))$  be the unconstrained optimum human capital conditional on  $i_1$ , where  $\hat{i}_2(\cdot)$  is defined in (9). By Lemma 4,  $\tilde{i}_1 \leq i_1^*$  implies  $\hat{h}_3(\tilde{i}_1) \leq \hat{h}_3(i_1^*) = h_3^*$ . Moreover,  $\tilde{h}_3 \leq \hat{h}_3(\tilde{i}_1)$  holds because, conditional on  $i_1$ , over-investing in  $i_2$  is never optimal ( $\tilde{i}_2 \leq \hat{i}_2(\tilde{i}_1, \theta)$ ). From these two inequalities, we have  $\tilde{h}_3 \leq h_3^*$ .

We have proved that  $\tilde{i}_1 \leq i_1^*$  and  $\tilde{h}_3 \leq h_3^*$  when  $f_1(\tilde{i}_1, \tilde{i}_2) \geq Rf_2(\tilde{i}_1, \tilde{i}_2)$ . We next show that when some constraints are binding, the inequalities are strict, i.e.  $\tilde{i}_1 < i_1^*$  and  $\tilde{h}_3 < h_3^*$ . Suppose that  $\theta\chi f_1(\tilde{i}_1, \tilde{i}_2) > R^2 = \theta\chi f_1(i_1^*, i_2^*)$  but  $\tilde{i}_1 = i_1^*$ . Obviously, the inequality  $f_1(\tilde{i}_1, \tilde{i}_2) > f_1(i_1^*, i_2^*)$  cannot hold when  $f_{12} = 0$ . When  $f_{12} > 0$ , the inequality implies  $\tilde{i}_2 > i_2^*$ , which is a contradiction because this, together with  $\tilde{i}_1 = i_1^*$ , means  $\tilde{h}_3 > h_3^*$ . When  $f_{12} < 0$ , the inequality implies  $\tilde{i}_2 < i_2^*$ , which in turn implies  $f_1(\tilde{i}_1, \tilde{i}_2) < Rf_2(\tilde{i}_1, \tilde{i}_2)$  due to the strict convexity of the isoquant. This contradicts the first order condition  $f_1(\tilde{i}_1, \tilde{i}_2) \geq Rf_2(\tilde{i}_1, \tilde{i}_2)$ . Thus, it must be  $\tilde{i}_1 < i_1^*$  when  $\theta\chi f_1(\tilde{i}_1, \tilde{i}_2) > R^2$ . In this case,  $\tilde{h}_3 < h_3^*$  also holds because  $\hat{h}_3(i_1)$  is strictly increasing in  $i_1$ .

Part (ii): By Part (i), when any borrowing constraint binds,  $\tilde{i}_1 < i_1^*$  and  $\tilde{h}_3 < h_3^*$ . (a) If the period one constraint does not bind,  $f_1(\tilde{i}_1, \tilde{i}_2) = Rf_2(\tilde{i}_1, \tilde{i}_2)$ , so  $\tilde{i}_1 = \dot{i}_1(\tilde{h}_3)$  and  $\tilde{i}_2 = \dot{i}_2(\tilde{h}_3)$  hold. By Lemma 3,  $\tilde{i}_2 = \dot{i}_2(\tilde{h}_3) < \dot{i}_2(h_3^*) = i_2^*$ . (b) If  $f_{12} > 0$ ,  $\tilde{i}_1 < i_1^*$  implies  $\hat{i}_2(\tilde{i}_1, \theta) < \hat{i}_2(i_1^*, \theta) = i_2^*$ . Because  $\tilde{i}_2 \leq \hat{i}_2(\tilde{i}_1, \theta)$  always holds, we get  $\tilde{i}_2 < i_2^*$ .

Part (iii): A similar analysis to Part (i) shows that  $\theta\chi f_2(i_1, i_2) \geq R$ , with strict inequality if any borrowing constraint in period 2 or later binds, and equality otherwise.

## Proof of Proposition 9

In the decision problem described in Section 4.5, substitute in for  $V(a_3, h_3)$  using equation (14) and in for  $c_1$  and  $c_2$  using the budget constraints in equation (5).

(I) Assuming this person is constrained as an old child, let  $a_3 = -L_2$  and re-write the decision problem as:

$$\max_{i_1, i_2, a_2} u(y_1 - i_1 - a_2) + \beta u(Ra_2 + y_2 - i_2 + L_2) + \beta^2 V(-RL_2, \theta f(i_1, i_2)).$$

First order conditions for  $i_1$ ,  $i_2$  and  $a_2$  are:

$$-u'(c_1) + \beta^2 V_2(-RL_2, \theta f(i_1, i_2)) \theta f_1(i_1, i_2) = 0 \quad (23)$$

$$-\beta u'(c_2) + \beta^2 V_2(-RL_2, \theta f(i_1, i_2)) \theta f_2(i_1, i_2) = 0 \quad (24)$$

$$-u'(c_1) + \beta R u'(c_2) = 0. \quad (25)$$

Together, these first order conditions imply  $f_1 = Rf_2$  at an optimum. Using this with Cramer's rule yields (dropping arguments of  $f(\cdot)$  and  $V(-RL_2, \cdot)$  for expositional purposes):

$$\begin{aligned} \frac{\partial i_1}{\partial y_1} &= \frac{R\beta^3 u''(c_1) u''(c_2) \theta V_2(Rf_{22} - f_{12})}{\Delta_2} > 0, \\ \frac{\partial i_2}{\partial y_1} &= \frac{R\beta^3 u''(c_1) u''(c_2) \theta V_2(f_{11} - Rf_{12})}{\Delta_2} > 0, \end{aligned}$$

$\frac{\partial i_1}{\partial y_2} = R^{-1} \frac{\partial i_1}{\partial y_1}$  and  $\frac{\partial i_2}{\partial y_2} = R^{-1} \frac{\partial i_2}{\partial y_1}$ , where

$$\begin{aligned} \Delta_2 &\equiv \beta^4 V_2 \theta^2 [u''(c_1) + \beta R^2 u''(c_2)] [V_2(f_{11} f_{22} - f_{12}^2) + V_{22} \theta (f_1^2 f_{22} + f_2^2 f_{11} - 2f_1 f_2 f_{12})] \\ &\quad + \beta^3 u''(c_1) u''(c_2) V_2 \theta (f_{11} + R^2 f_{22} - 2Rf_{12}) < 0. \end{aligned} \quad (26)$$

All of these expressions are signed using assumptions on  $V$ , Assumption 1 and  $f_1 = Rf_2$ .

Finally,  $\frac{\partial h_3}{\partial y_j} = f_1 \frac{\partial i_1}{\partial y_j} + f_2 \frac{\partial i_2}{\partial y_j} > 0$  for  $j = 1, 2$ , because all terms in this expression are strictly positive;  $\frac{\partial h_3}{\partial y_1} = R \frac{\partial h_3}{\partial y_2}$  follows directly from the fact that  $\frac{\partial i_j}{\partial y_1} = R \frac{\partial i_j}{\partial y_2}$  for  $j = 1, 2$ .

(II) Let  $a_2 = -L_1$ . With no future constraints, we can rewrite the continuation value  $V(a, h) = v(Ra + \chi h)$  as defined in equation (20). Substitute this into the problem:

$$\max_{i_1, i_2, a_3} u(y_1 - i_1 + L_1) + \beta u(-RL_1 + y_2 - i_2 - a_3) + \beta^2 v(Ra_3 + \chi \theta f(i_1, i_2)),$$

where  $\chi > 0$  is defined above. First order conditions for  $i_1$ ,  $i_2$  and  $a_3$  are:

$$\begin{aligned} -u'(c_1) + \beta^2 v'(Ra_3 + \chi\theta f(i_1, i_2))\chi\theta f_1(i_1, i_2) &= 0 \\ -\beta u'(c_2) + \beta^2 v'(Ra_3 + \chi\theta f(i_1, i_2))\chi\theta f_2(i_1, i_2) &= 0 \\ -\beta u'(c_2) + \beta^2 Rv'(Ra_3 + \chi\theta f(i_1, i_2)) &= 0. \end{aligned}$$

Combining first order conditions, we have  $\chi\theta f_2 = R$ . However,  $f_1 > Rf_2 = R^2/(\chi\theta)$  because  $L_1$  binds (see Proposition 8).

Cramer's rule yields (dropping arguments of  $f(\cdot)$  and  $v(\cdot)$ ):

$$\begin{aligned} \frac{\partial i_1}{\partial y_1} &= \frac{\beta^3 u''(c_1) v' \chi \theta f_{22} [u''(c_2) + \beta R^2 v'']}{\Delta_1} > 0 \\ \frac{\partial i_1}{\partial y_2} &= \frac{-\beta^5 R u''(c_2) v' v'' \chi^2 \theta^2 f_1 f_{22}}{\Delta_1} < 0 \\ \frac{\partial i_2}{\partial y_1} &= \frac{-\beta^3 u''(c_1) v' \chi \theta f_{12} [u''(c_2) + \beta R^2 v'']}{\Delta_1} \\ \frac{\partial i_2}{\partial y_2} &= \frac{\beta^5 R u''(c_2) v' v'' \chi^2 \theta^2 f_1 f_{12}}{\Delta_1} \end{aligned}$$

where

$$\begin{aligned} \Delta_1 &\equiv \beta^3 u''(c_1) v' [u''(c_2) + \beta R^2 v''] \chi \theta f_{22} + \beta^5 u''(c_2) v' v'' \chi^3 \theta^3 f_1^2 f_{22} \\ &\quad + \beta^5 [u''(c_2) + \beta R^2 v''] (v')^2 \chi^2 \theta^2 [f_{11} f_{22} - f_{12}^2] < 0 \end{aligned} \quad (27)$$

by Assumption 1. Clearly,  $\frac{\partial i_2}{\partial y_1} > 0 \iff f_{12} > 0$ , and  $\frac{\partial i_2}{\partial y_2} < 0 \iff f_{12} > 0$ . Because  $\frac{\partial h_3}{\partial y_j} = f_1 \frac{\partial i_1}{\partial y_j} + f_2 \frac{\partial i_2}{\partial y_j}$  for  $j = 1, 2$ , Assumption 1 implies that  $\frac{\partial h_3}{\partial y_1} > 0$  and  $\frac{\partial h_3}{\partial y_2} < 0$ .

(III) Because we are assuming this person is constrained in both childhood periods, let  $a_2 = -L_1$  and  $a_3 = -L_2$ . The decision problem can be written as:

$$\max_{i_1, i_2} u(y_1 - i_1 + L_1) + \beta u(-RL_1 + y_2 - i_2 + L_2) + \beta^2 V(-RL_2, \theta f(i_1, i_2)).$$

The first order conditions for  $i_1$  and  $i_2$  are given by equations (23) and (24), where  $c_1 = y_1 - i_1 + L_1$  and  $c_2 = -RL_1 + y_2 - i_2 + L_2$ .

Cramer's rule yields (dropping arguments of  $f(\cdot)$  and  $V(-RL_2, \cdot)$ ):

$$\begin{aligned} \frac{\partial i_1}{\partial y_1} &= \frac{\beta u''(c_1) [u''(c_2) + \beta V_2 \theta f_{22} + \beta V_{22} \theta^2 f_2^2]}{\Delta_{12}} > 0 \\ \frac{\partial i_1}{\partial y_2} &= \frac{-\beta^3 u''(c_2) \theta [V_2 f_{12} + V_{22} \theta f_1 f_2]}{\Delta_{12}} \\ \frac{\partial i_2}{\partial y_1} &= \frac{-\beta^2 u''(c_1) \theta [V_2 f_{12} + V_{22} \theta f_1 f_2]}{\Delta_{12}} \\ \frac{\partial i_2}{\partial y_2} &= \frac{\beta u''(c_2) \theta [u''(c_1) + \beta^2 V_2 \theta f_{11} + \beta^2 V_{22} \theta f_1^2]}{\Delta_{12}} > 0 \end{aligned}$$

where

$$\begin{aligned}\Delta_{12} \equiv & \beta u''(c_1)u''(c_2) + \beta^2 V_2 \theta [u''(c_1)f_{22} + \beta u''(c_2)f_{11}] + \beta^2 V_{22} \theta^2 [u''(c_1)f_2^2 + \beta u''(c_2)f_1^2] \\ & + \beta^4 (V_2)^2 \theta^2 (f_{11}f_{22} - f_{12}^2) + \beta^4 V_2 V_{22} \theta^3 [f_2(f_2 f_{11} - f_1 f_{12}) + f_1(f_1 f_{22} - f_2 f_{12})] > 0. \quad (28)\end{aligned}$$

Assumptions on  $V(\cdot, \cdot)$  and Assumption 1 ensure that  $\frac{\partial i_1}{\partial y_1}$ ,  $\frac{\partial i_2}{\partial y_2}$ , and  $\Delta_{12}$  are strictly positive. Both  $\frac{\partial i_1}{\partial y_2}$  and  $\frac{\partial i_2}{\partial y_1}$  are strictly positive if and only if Condition 2 holds. Using these results for investments, Assumption 1 implies that  $\frac{\partial h_3}{\partial y_j} > 0$  for  $j = 1, 2$ .  $\square$

### Proof of Proposition 10

Part (i): Based on the problem discussed in the proof of Proposition 9 part (I), we can apply Cramer's rule obtaining:

$$\begin{aligned}\frac{\partial i_1}{\partial L_2} &= \frac{\beta^4 R V_2 V_{21} \theta^2 f_2 [u''(c_1) + \beta R^2 u''(c_2)] (f_1 f_{22} - f_{12}) + \beta^3 u''(c_1) u''(c_2) V_2 \theta (R f_{22} - f_{12})}{\Delta_2} > 0 \\ \frac{\partial i_2}{\partial L_2} &= \frac{\beta^4 R V_2 V_{21} \theta^2 f_2 [u''(c_1) + \beta R^2 u''(c_2)] (f_{11} - R f_{12}) + \beta^3 u''(c_1) u''(c_2) V_2 \theta (f_{11} - R f_{12})}{\Delta_2} > 0,\end{aligned}$$

where  $\Delta_2 < 0$  is defined previously by equation (26). All three of these expressions are signed using assumptions on  $V$ , Assumption 1 and  $f_1 = R f_2$ . Finally,  $\frac{\partial h_3}{\partial L_2} = f_1 \frac{\partial i_1}{\partial L_2} + f_2 \frac{\partial i_2}{\partial L_2} > 0$ , because all terms in this expression are positive.

Part (ii): Based on the problem used in the proof of Proposition 9 part (III), Cramer's rule yields:

$$\begin{aligned}\frac{\partial i_1}{\partial L_2} &= \frac{\beta^4 R V_2 V_{21} \theta^2 (f_1 f_{22} - f_2 f_{12}) + \beta^3 R u''(c_2) V_{21} \theta f_1 - \beta^3 u''(c_2) \theta (V_2 f_{12} + \theta V_{22} f_1 f_2)}{\Delta_{12}} \\ \frac{\partial i_2}{\partial L_2} &= \frac{\beta u''(c_2) [u''(c_1) + \beta^2 V_2 \theta f_{11} + \beta^2 V_{22} \theta^2 f_1^2] + \beta^2 R u''(c_1) V_{21} \theta f_2 + \beta^4 R V_2 V_{21} \theta^2 (f_2 f_{11} - f_1 f_{12})}{\Delta_{12}} > 0,\end{aligned}$$

where  $\Delta_{12} > 0$  is defined previously by equation (28). Using assumptions on  $V$ , and Assumption 1, it is clear that  $\frac{\partial i_1}{\partial L_2} > 0$  if Condition 1 holds, and that  $\frac{\partial i_2}{\partial L_2} > 0$ . Assumption 1 further implies that  $\frac{\partial h_3}{\partial L_2} = f_1 \frac{\partial i_1}{\partial L_2} + f_2 \frac{\partial i_2}{\partial L_2} > 0$ .  $\square$

### Proof of Proposition 11

Part (i): Based on the problem discussed in the proof of Proposition 9 part (II), we can apply Cramer's rule obtaining:

$$\begin{aligned}\frac{\partial i_1}{\partial L_1} &= \frac{f_{22} \{ \beta^3 u''(c_1) v' \chi \theta [u''(c_2) + \beta R^2 v''] + \beta^5 R^2 u''(c_2) v' v'' \chi^2 \theta^2 f_1 \}}{\Delta_1} > 0 \\ \frac{\partial i_2}{\partial L_1} &= \frac{-f_{12} \{ \beta^3 u''(c_1) v' \chi \theta [u''(c_2) + \beta R^2 v''] + \beta^5 R^2 u''(c_2) v' v'' \chi^2 \theta^2 f_1 \}}{\Delta_1}\end{aligned}$$

where  $\Delta_1 < 0$  is defined previously by equation (27). Clearly,  $\frac{\partial i_2}{\partial L_1} > 0 \iff f_{12} > 0$ . Finally,  $\frac{\partial h_3}{\partial L_2} = f_1 \frac{\partial i_1}{\partial L_2} + f_2 \frac{\partial i_2}{\partial L_2} > 0$  by Assumption 1.

Part (ii): Based on the problem used in the proof of Proposition 9 part (III), Cramer's rule yields:

$$\begin{aligned}\frac{\partial i_1}{\partial L_1} &= \frac{\beta u''(c_1)[u''(c_2) + \beta V_2 \theta f_{22} + \beta V_{22} \theta^2 f_2^2] + \beta^3 R u''(c_2) \theta [V_2 f_{12} + V_{22} \theta f_1 f_2]}{\Delta_{12}} \\ \frac{\partial i_2}{\partial L_1} &= \frac{-\beta R u''(c_2)[u''(c_1) + \beta^2 V_2 \theta f_{11} + \beta^2 V_{22} \theta^2 f_1^2] - \beta^2 u''(c_1) \theta [V_2 f_{12} + V_{22} \theta f_1 f_2]}{\Delta_{12}}\end{aligned}$$

where  $\Delta_{12} > 0$  is defined previously by equation (28). Assumptions on  $V$ , and Assumption 1 imply that  $\frac{\partial i_1}{\partial L_1} < 1$ . If Condition 2 does not hold, then  $V_2 f_{12} + V_{22} \theta f_1 f_2 < 0$ , which implies that  $\frac{\partial i_1}{\partial L_1} > 0$  and  $\frac{\partial i_2}{\partial L_1} < 0$ .  $\square$