Web Appendices for "The Technology of Skill Formation"

Flavio Cunha and James Heckman

Article published in *American Economic Review* **97**(2), May 2007

Contents

A Deriving the Technology of Skill Formation from its Primitives

A Model of Skill Formation

In the models presented in this section of the appendix and in CHLM, parents make decisions about their children. We ignore how the parents get to be who they are and the decisions of the children about their own children. Appendix C develops a generationally consistent model.

Suppose that there are two periods in a child's life, "1" and "2", before the child becomes an adult. Adulthood comprises distinct third and fourth periods. The child works for two periods after the two periods of childhood. Models based on the analysis of Becker and Tomes (1986) assume only one period of childhood. We assume that there are two kinds of skills: θ^C and θ^N. For example, θ^C can be thought of as cognitive skill and θ^N as noncognitive skill. Our treatment of ability is in contrast to the view of the traditional literature on human capital formation that views IQ as innate ability. In our analysis, IQ is just another skill. What differentiates IQ from other cognitive and noncognitive skills is that IQ is subject to accumulation during critical periods. That is, parental and social interventions can increase the IQ of the child, but they can do so successfully only for a limited time.

Let I_t^k denote parental investments in child skill k at period t, $k = C, N$ and $t = 1, 2$. Let h' be the level of human capital as the child starts adulthood. It depends on both components of (θ_2^C, θ_2^N). The parents fully control the investment of the child. A richer model incorporates, among other features, investment decisions of the child as influenced by the parent through preference formation processes (see Carneiro, Cunha, and Heckman, 2003, and Cunha and Heckman, 2007a).

We first describe how skills evolve over time. Assume that each agent is born with initial conditions $\theta_1 = \left(\theta_1^C, \theta_1^N\right)$. These can be determined by parental environments. At each stage t let $\theta_t = \left(\theta_t^C, \theta_t^N\right)$ denote the vector of skill or ability stocks. The technology of production of skill k at period t is

$$\theta_{t+1}^k = f_t^k\left(\theta_t, I_t^k\right), \tag{A-1}$$

for $k = C, N$ and $t = 1, 2$. We assume that f_t^k is twice continuously differentiable, increasing and concave. In this model, stocks of both skills and abilities produce next period skills and the productivity of investments. Cognitive skills can promote the formation of noncognitive skills and vice versa.

Let θ_3^C, θ_3^N denote the level of skills when adult. We define adult human capital h' of the child as a combination of different adult skills:

$$h' = g\left(\theta_3^C, \theta_3^N\right). \tag{A-2}$$

The function g is assumed to be continuously differentiable and increasing in $\left(\theta_3^C, \theta_3^N\right)$. This model assumes that there is no comparative advantage in the labor market or in life itself.[1]

To fix ideas, consider the following specialization of our model. Ignore the effect of initial conditions and assume that first period skills are just due to first period investment:

$$\theta_2^C = f_1^C\left(\theta_1, I_1^C\right) = I_1^C$$

and

$$\theta_2^N = f_1^C\left(\theta_1, I_1^C\right) = I_1^N,$$

where I_1^C and I_1^N are scalars. For the second period technologies, assume a CES structure:

$$\begin{aligned} \theta_3^C &= f_2^C\left(\theta_2, I_2^C\right) \\ &= \left\{\gamma_1\left(\theta_2^C\right)^\alpha + \gamma_2\left(\theta_2^N\right)^\alpha + (1-\gamma_1-\gamma_2)\left(I_2^C\right)^\alpha\right\}^{\frac{1}{\alpha}} \quad \text{where} \quad \begin{array}{l} 1 \geq \gamma_1 \geq 0, \\ 1 \geq \gamma_2 \geq 0, \\ 1 \geq 1-\gamma_1-\gamma_2 \geq 0, \end{array} \end{aligned} \tag{A-3}$$

and

$$\begin{aligned} \theta_3^N &= f_2^N\left(\theta_2, I_2^N\right) \\ &= \left\{\eta_1\left(\theta_2^C\right)^\nu + \eta_2\left(\theta_2^N\right)^\nu + (1-\eta_1-\eta_2)\left(I_2^N\right)^\nu\right\}^{\frac{1}{\nu}} \quad \text{where} \quad \begin{array}{l} 1 \geq \eta_1 \geq 0, \\ 1 \geq \eta_2 \geq 0, \\ 1 \geq 1-\eta_1-\eta_2 \geq 0, \end{array} \end{aligned} \tag{A-4}$$

where $\frac{1}{1-\alpha}$ is the elasticity of substitution in the inputs producing θ_3^C and $\frac{1}{1-\nu}$ is the elasticity of substitution of inputs in producing θ_3^N where $\alpha \in (-\infty, 1]$ and $\nu \in (-\infty, 1]$. Notice that I_2^N and I_2^C are direct complements with (θ_2^C, θ_2^N)

[1]Thus we rule out one potentially important avenue of compensation that agents can specialize in tasks that do not require the skills in which they are deficient. In CHLM we briefly consider a more general task function that captures the notion that different tasks require different combinations of skills and abilities.

irrespective of the substitution parameters α and ν, except in limiting cases.

The CES technology is well known and has convenient properties. It imposes direct complementarity even though inputs may be more or less substitutable depending on α or ν. We distinguish between direct complementarity (positive cross partials) and CES-substitution/complementarity. Focusing on the technology for producing θ_3^C, when $\alpha = 1$, the inputs are perfect substitutes in the intuitive use of that term (the elasticity of substitution is infinite). The inputs θ_2^C, θ_2^N and I_2^C can be ordered by their relative productivity in producing θ_3^C. The higher γ_1 and γ_2, the higher the productivity of θ_2^C and θ_2^N respectively. When $\alpha = -\infty$, the elasticity of substitution is zero. All inputs are required in the same proportion to produce a given level of output so there are no possibilities for technical substitution, and

$$\theta_3^C = \min\left\{\theta_2^C, \theta_2^N, I_2^C\right\}.$$

In this technology, early investments are a *bottleneck* for later investment. Compensation for adverse early environments through late investments is impossible.

The evidence from numerous studies reviewed in the text and in CHLM cited shows that IQ is no longer malleable after ages 8-10. Taken at face value, this implies that if θ^C is IQ, for all values of I_2^C, $\theta_3^C = \theta_2^C$. Period 1 is a critical period for IQ but not necessarily for other skills and abilities. More generally, period 1 is a critical period if

$$\frac{\partial \theta_{t+1}^C}{\partial I_t^C} = 0 \text{ for } t > 1.$$

For parameterization (A-3), this is obtained by imposing $\gamma_1 + \gamma_2 = 1$.

The evidence on adolescent interventions surveyed in CHLM shows substantial positive results for such interventions on noncognitive skills (θ_3^N) and at most modest gains for cognitive skills. Technologies (A-3) and (A-4) can rationalize this pattern. Since the populations targeted by adolescent intervention studies tend to come from families with poor backgrounds, we would expect I_1^C and I_1^N to be below average. Thus, θ_2^C and θ_2^N will be below average. Adolescent interventions make I_2^C and I_2^N relatively large for the treatment group in comparison to the control group in the intervention experiments. At stage 2, θ_3^C (cognitive ability) is essentially the same in the control and treatment groups, while θ_3^N (noncognitive ability) is higher for the treated group. Large values of $(\gamma_1 + \gamma_2)$ (associated with a small coefficient on I_2^C) or small values of $(\eta_1 + \eta_2)$ (so the coefficient on I_2^N is large) and high values of α and ν can produce this pattern. Another case that rationalizes the evidence is when $\alpha \to -\infty$ and $\nu = 1$. Under these conditions,

$$\theta_3^C = \min\left\{\theta_2^C, \theta_2^N, I_2^C\right\}, \tag{A-5}$$

while

$$\theta_3^N = \eta_1 \theta_2^C + \eta_2 \theta_2^N + (1 - \eta_1 - \eta_2) I_2^N. \tag{A-6}$$

The attainable period 2 stock of cognitive skill $\left(\theta_3^C\right)$ is limited by the minimum value of $\theta_2^C, \theta_2^N, I_2^C$. In this case, any level of investment in period 2 such that $I_2^C > \min\left\{\theta_2^C, \theta_2^N\right\}$ is ineffective in incrementing the stock of cognitive skills. Period 1 is a bottleneck period. Unless sufficient skill investments are made in θ_C in period 1, it is not possible to raise skill θ_C in period 2. This phenomenon does not appear in the production of the noncognitive skill, provided that $(1 - \eta_1 - \eta_2) > 0$. More generally, the higher ν and the larger $(1 - \eta_1 - \eta_2)$, the more productive is investment I_2^N in producing θ_2^N.

To complete the CES example, assume that adult human capital h' is a CES function of the two skills accumulated at stage two:

$$h' = \left\{\tau\left(\theta_3^C\right)^{\phi} + (1 - \tau)\left(\theta_3^N\right)^{\phi}\right\}^{\frac{\rho}{\phi}}, \tag{A-7}$$

where $\rho \in (0, 1)$, $\tau \in [0, 1]$, and $\phi \in (-\infty, 1]$. In this parameterization, $\frac{1}{1-\phi}$ is the elasticity of substitution across different skills in the production of adult human capital. Equation (A-7) reminds us that the market, or life in general, requires use of multiple skills. Being smart isn't the sole determinant of success. In general, different tasks require both skills in different proportions. One way to remedy early skill deficits is to make compensatory investments. Another way is to motivate people from disadvantaged environments to pursue tasks that do not require the skill that deprived early environments do not produce. A richer theory would account for this choice of tasks and its implications for remediation.[2] For the sake of simplifying our argument, we work with equation (A-7) that captures the notion that skills can trade off against each other in producing effective people. Highly motivated, but not very bright, people may be just as effective as bright but unmotivated people. That is one of the lessons from the GED program. (See Heckman and Rubinstein, 2001, and Heckman, Stixrud, and Urzua, 2006.)

The analysis is simplified by assuming that investments are general in

[2] See the appendix in CHLM.

nature: $I_1^C = I_1^N = I_1$, $I_2^C = I_2^N = I_2$.[3] Cunha and Heckman (2007a,b) develop the more general case of skill-specific investments which requires more notational complexity.

With common investment goods, we can solve out for θ_2^C and θ_2^N in terms of I_1 to simplify (A-3) and (A-4) to reach

$$\theta_2^C = \{(\gamma_1 + \gamma_2)(I_1)^\alpha + (1 - \gamma_1 - \gamma_2)(I_2)^\alpha\}^{\frac{1}{\alpha}} \tag{A-8}$$

and

$$\theta_2^N = \{(\eta_1 + \eta_2)(I_1)^\nu + (1 - \eta_1 - \eta_2)(I_2)^\nu\}^{\frac{1}{\nu}} . \tag{A-9}$$

If we then substitute these expressions into the production function for adult human capital (A-7), we obtain

$$h' = \left\{\tau \left[\tilde{\gamma}(I_1)^\alpha + (1 - \tilde{\gamma})(I_2)^\alpha\right]^{\frac{\phi}{\alpha}} + (1 - \tau)\left[\tilde{\eta}(I_1)^\nu + (1 - \tilde{\eta})(I_2)^\nu\right]^{\frac{\phi}{\nu}}\right\}^{\frac{\rho}{\phi}}, \tag{A-10}$$

where $\tilde{\gamma} = \gamma_1 + \gamma_2$, $\tilde{\eta} = \eta_1 + \eta_2$. Equation (A-10) expresses adult human capital as a function of the entire sequence of childhood investments in human capital. Current investments in human capital are combined with the existing stocks of skills in order to produce the stock of next period skills.

A conveniently simple formulation of the problem arises if we assume that $\alpha = \nu = \phi$ so that CES substitution among inputs in producing outputs and CES substitution among skill in producing human capital are the same. This produces the convenient and familiar-looking CES expression for adult human capital stocks:

$$h' = \left\{\gamma I_1^\phi + (1 - \gamma) I_2^\phi\right\}^{\frac{\rho}{\phi}}, \tag{A-11}$$

where $\gamma = \tau\tilde{\gamma} + (1 - \tau)\tilde{\eta}$ and $\phi = \alpha = \nu$. The parameter γ is *a skill multiplier*. It arises because I_1 affects the accumulation of θ_2^C and θ_2^N. These stocks of skills in turn affect the productivity of I_2 in forming θ_3^C and θ_3^N. Thus γ captures the net effect of I_1 on h' through both self-productivity and direct

[3] Thus when a parent buys a book in the first period of childhood, this book may be an investment in all kinds of skills. It is an investment in cognitive skills, as it helps the child get exposure to language and new words. It can also be an investment in noncognitive skills, if the book may contain a message on the importance of being persistent and patient.

complementarity.[4] $\frac{1}{1-\phi}$ is a measure of how easy it is to substitute between I_1 and I_2 where the substitution arises from both the task performance (human capital) function in equation (A-7) and the technology of skill formation. Within the CES technology, ϕ is a measure of the ease of substitution of inputs. In this analytically convenient case, the parameter ϕ plays a dual role. First, it informs us how easily one can substitute across different skills in order to produce one unit of adult human capital h'. Second, it also represents the degree of complementarity (or substitutability) between early and late investments in producing skills. In this second role, the parameter ϕ dictates how easy it is to compensate for low levels of stage 1 skills in producing late skills.

In principle, compensation can come through two channels: (i) through skill investment or (ii) through choice of market activities, substituting deficits in one skill by the relative abundance in the other through choice of tasks. We do not develop the second channel of compensation in this appendix, deferring it to later work. It is discussed in Carneiro, Cunha, and Heckman (2003).

When ϕ is small, low levels of early investment I_1 are not easily remediated by later investment I_2 in producing human capital. The other face of CES complementarity is that when ϕ is small, high early investments should be followed with high late investments. In the extreme case when $\phi \to -\infty$, (A-11) converges to $h' = (\min\{I_1, I_2\})^\rho$. We analyzed this case in CHLM. The Leontief case contrasts sharply with the case of perfect CES substitutes, which arises when $\phi = 1$: $h' = [\gamma I_1 + (1-\gamma) I_2]^\rho$. When we impose the further restriction that $\gamma = \frac{1}{2}$, we generate the model that is implicitly assumed in the existing literature on human capital investments that collapses childhood into a single period. In this special case, only the total

[4]Direct complementarity between I_1 and I_2 arises if

$$\frac{\partial^2 h}{\partial I_1 \partial I_2} > 0.$$

As long as $\rho > \phi$, I_1 and I_2 are direct complements, because

$$\operatorname{sign}\left(\frac{\partial^2 h}{\partial I_1 \partial I_2}\right) = \operatorname{sign}(\rho - \phi).$$

This definition of complementarity is to be distinguished from the notion based on the elasticity of substitution between I_1 and I_2, which is $\frac{1}{1-\phi}$. When $\phi < 0$, I_1 and I_2 are sometimes called complements. When $\phi > 0$, I_1 and I_2 are sometimes called substitutes. When $\rho = 1$, I_1 and I_2 are always direct complements, but if $1 > \phi > 0$, they are CES substitutes.

amount of human capital investments, regardless of how it is distributed across childhood periods, determines adult human capital. In the case of perfect CES substitutes, it is possible in a physical productivity sense to compensate for early investment deficits by later investments, although it may not be economically efficient to do so.

When $\rho = 1$, we can rewrite (A-11) as

$$h' = I_1 \left\{\gamma + (1-\gamma)\,\omega^{\phi}\right\}^{\frac{1}{\phi}},$$

where $\omega = I_2/I_1$. Fixing I_1 (early investment), an increase in ω is the same as an increase in I_2. The marginal productivity of late investment is

$$\frac{\partial h'}{\partial \omega} = (1-\gamma)\, I_1 \left\{\gamma + (1-\gamma)\,\omega^{\phi}\right\}^{\frac{1-\phi}{\phi}} \omega^{\phi-1}.$$

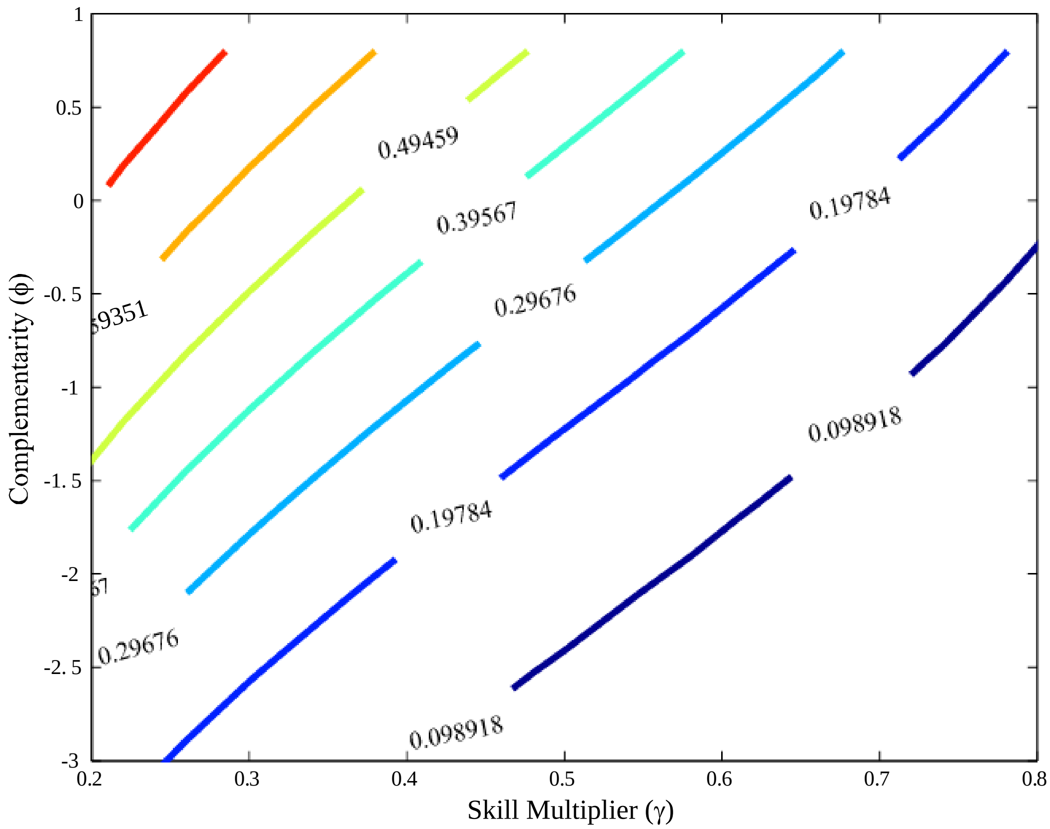
For $\omega > 1$ and $\gamma < 1$, marginal productivity is increasing in ϕ and $(1-\gamma)$. Thus, provided that late investments are greater than earlier investments, the more substitutable I_2 is with I_1 (the higher ϕ) and the lower the skill multiplier γ, the more productive are late investments. Figure A1 graphs the isoquants for $\frac{\partial h'}{\partial \omega}$ when $\omega = 2$. It shows that a high ϕ trades off with a high γ. As $\left(\phi, 1-\gamma\right)$ increases along a ray, $\frac{\partial h'}{\partial \omega}$ increases. For a fixed skill multiplier γ, the higher ϕ, the higher the marginal productivity of second period investment.

If, however, $\omega < 1$ as in Figure A2, then $\frac{\partial h'}{\partial \omega}$ could be decreasing as $\left(\phi, 1-\gamma\right)$ increases along a ray and the trade-off between ϕ and $(1-\gamma)$ along a $\left(\frac{\partial h'}{\partial \omega}, \omega\right)$ isoquant is reversed. If I_1 is large relative to I_2 (i.e., $\omega < 1$), for a fixed γ the marginal product of I_2 is decreasing in ϕ. More CES complementarity implies greater productivity (see Figure A2).[5] The empirically relevant case for the analysis of investment in disadvantaged children is $\omega > 1$ as shown in Figure C3, so greater CES-substitutability and a smaller skill multiplier produce a higher marginal productivity of remedial second period investment.

[5]One can show that at sufficiently low values of ϕ, the marginal productivity is no longer increasing in ϕ.

Figure A1

The indifference curves of the marginal productivity of the ratio of late to early investments as a function of ϕ and γ when $I_2/I_1 = 2$.



Define $\omega = \frac{I_2}{I_1}$, the ratio of late to early investments in human capital. From the homogeneity of degree one we can rewrite the technology as:

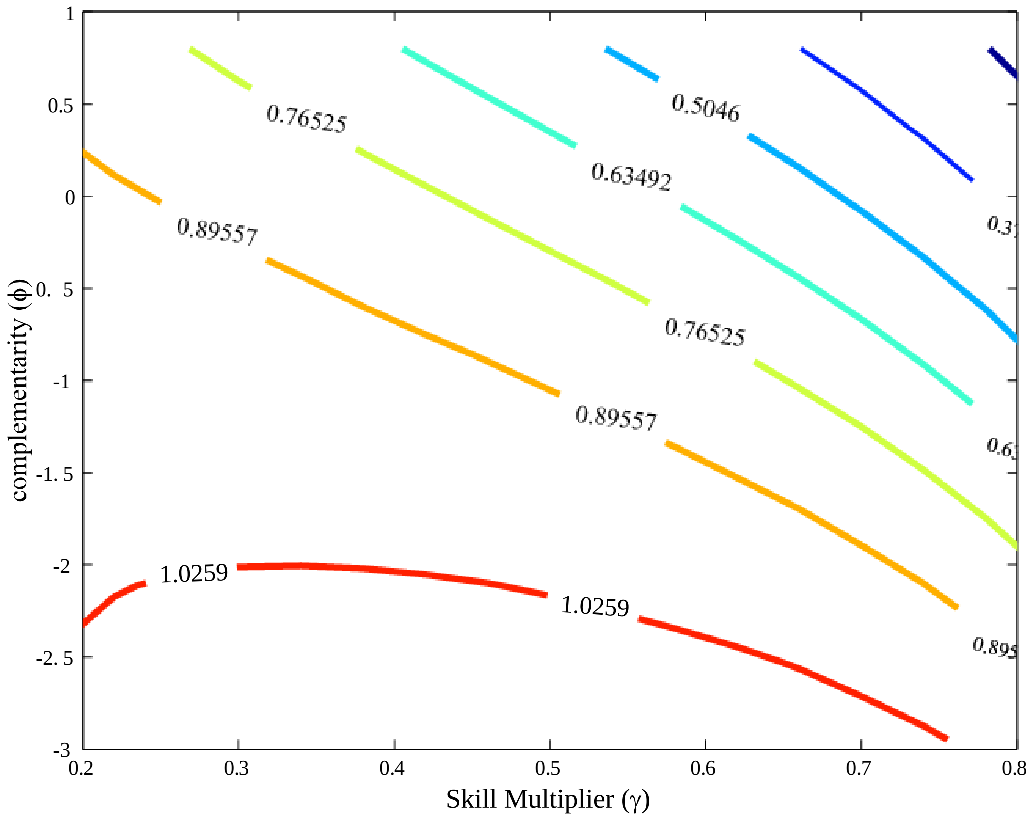
$$h = I_1 \left[\gamma + (1-\gamma)\,\omega^{\phi}\right]^{\frac{1}{\phi}}.$$

The marginal product of the ratio of late to early investment, ω, holding early investment constant, is

$$\frac{\partial h}{\partial \omega} = (1-\gamma)\, I_1 \left[\gamma + (1-\gamma)\,\omega^{\phi}\right]^{\frac{1-\phi}{\phi}} \omega^{\phi-1}$$

This figure displays the indifference curves of $\frac{\partial h}{\partial \omega}$ when $\omega = 0.5$. Each indifference curve shows the corresponding level of $\frac{\partial h}{\partial \omega}$. Note that for a given value of γ the value of the function tends to decrease as we increase ϕ. The function also increases as we decrease γ.

Figure A2
The indifference curves of the marginal productivity of the ratio of late to early investments as a function of ϕ and γ when $I_2/I_1 = 1/2$.



Consider the CES specification for the technology of human capital formation:

$$h = \left[\gamma I_1^{\phi} + (1-\gamma) I_2^{\phi}\right]^{\frac{1}{\phi}}$$

Define $\omega = \frac{I_2}{I_1}$, the ratio of late to early investments in human capital. From the homogeneity of degree one we can rewrite the technology as:

$$h = I_1 \left[\gamma + (1-\gamma)\,\omega^{\phi}\right]^{\frac{1}{\phi}}.$$

The marginal product of the ratio of late to early investment, ω, holding early investment constant, is

$$\frac{\partial h}{\partial \omega} = (1-\gamma)\, I_1 \left[\gamma + (1-\gamma)\,\omega^{\phi}\right]^{\frac{1-\phi}{\phi}} \omega^{\phi-1}$$

This figure displays the indifference curves of $\frac{\partial h}{\partial \omega}$ when $\omega = 0.5$. Each indifference curve shows the corresponding level of $\frac{\partial h}{\partial \omega}$. Note that for a given value of γ the value of the function tends to decrease as we increase ϕ. However, the function may not be monotonic with respect to γ.

Table 7. The Ratio of Optimal Early and Late Investments $\frac{I_1}{I_2}$ Under Different Assumptions About the Skill Formation Technology

	Low Self-Productivity: $\gamma < \frac{(1+r)}{(2+r)}$	High Self-Productivity: $\gamma > \frac{(1+r)}{(2+r)}$
High Degree of Complementarity: $\phi < 0$	$\frac{I_1}{I_2} \to 1$ as $\phi \to -\infty$	$\frac{I_1}{I_2} \to 1$ as $\phi \to -\infty$
Low Degree of Complementarity: $0 \leq \phi \leq 1$	$\frac{I_1}{I_2} \to 0$ as $\phi \to 1$	$\frac{I_1}{I_2} \to \infty$ as $\phi \to 1$

Note: This table summarizes the behavior of the ratio of optimal early to late investments according to four cases: I_1 and I_2 have high complementarity, but self-productivity is low; I_1 and I_2 have both high complementarity and self-productivity; I_1 and I_2 have low complementarity and self-productivity; and I_1 and I_2 have low complementarity, but high self-productivity. When I_1 and I_2 exhibit high complementary, complementarity dominates and is a force towards equal distribution of investments between early and late periods. Consequently, self-productivity plays a limited role in determining the ratio $\frac{I_1}{I_2}$ (row 1). On the other hand, when $I_1$and I_2 exhibit a low degree of complementarity, self-productivity tends to concentrate investments in the *late* period if self-productivity is low, but in the *early* period if it is high (row 2).

B Comparing the Technology of Skill Formation to the Ben-Porath Model

B.1 The General Technology of Skill Formation

Let θ_t be a $L \times 1$ vector of skills or abilities at stage t. Included are pure cognitive abilities (e.g. IQ) as well as noncognitive abilities (time preference, self control, patience, judgment). The notation is sufficiently flexible to include acquired skills like general education or a specific skill. Agents start out life with vector θ_1 of skills (abilities). The θ_1 are produced by genes and *in utero* environments which are known to affect child outcomes (see, e.g., the essays in Keating and Hertzman, 1999).

Let I_t be a $K \times 1$ vector of investments at stage t. These include all inputs invested in the child including parental and social inputs. The technology of skill formation can be written as

$$\theta_{t+1} = f_t\left(\theta_t, I_t\right)$$

where f_t is a stage-t function mapping skill (ability) levels and investment at stage t into skill (ability) levels at the end of the period. For simplicity we assume that f_t is twice continuously differentiable in its arguments. Its domain of definition is the same for all inputs. The inputs may be different at different stages of the life cycle, so the inputs in I_t may be different from the inputs at period τ different from t.

Universal self-productivity at stage t is defined as

$$\frac{\partial \theta_{t+1}}{\partial \theta_t} = \frac{\partial f_t}{\partial \theta_t} > 0.$$

In the general case this is a $L \times L$ matrix. More generally, some components of this matrix may be zero at all stages while other components may always be positive. In principle, some skills could have negative effects in some periods. At some stages, some components may be zero while at other stages they may be positive.

Universal direct complementarity at stage t is defined by the $L \times K$ matrix:

$$\frac{\partial^2 \theta_{t+1}}{\partial \theta_t \partial I_t'} > 0.$$

Higher levels of θ_t raise the productivity of I_t. Alternatively, higher levels of

I_t raise the productivity of θ_t. Again, in the general case, some components at some or all stages may have zero effects, and some may have negative effects. They can switch signs across stages.

This notation is sufficiently general to allow for the possibility that some components of skill are produced only at certain critical periods. Period t is critical for skill (ability) j if

$$\frac{\partial \theta_{t+1,j}}{\partial I_t} \neq 0,$$

for some levels of θ_t, $I_t = i_t$, but

$$\frac{\partial \theta_{t+k+1,j}}{\partial I_{t+k}} = 0, \qquad k > 0,$$

for all levels of θ_t, $I_t = i_t$.

Sensitive periods might be defined as those periods where, at the same level of input θ_t, I_t, the $\frac{\partial \theta_{t+1}}{\partial I_t}$ are high. More formally, letting $\theta_t = \vartheta$, $I_t = i$, t is a sensitive period for skill (or ability) j if

$$\left.\frac{\partial \theta_{t+k+1,j}}{\partial I_{t+k}}\right|_{\theta_{t+k}=\vartheta, I_{t+k}=i} < \left.\frac{\partial \theta_{t+1,j}}{\partial I_t}\right|_{\theta_t=\vartheta, I_t=i}, \qquad \text{for } k \neq 0.$$

Clearly there may be multiple sensitive periods, and there may be sensitivity with respect to one input that is not true of other inputs.

An alternative definition of critical and sensitive periods works with a version of the technology that solves out $\theta_{t+1,j}$ as a function of lagged investments and initial conditions $\theta_1 = \theta_1$:

$$\theta_{t+1,j} = M_{t,j}\left(I_t, I_{t-1}, \ldots, I_1, \theta_1\right), \quad j = 1, \ldots, J.$$

Stage t^* is a *critical period* for $\theta_{t+1,j}$ if investments are productive at t^* but not at any other stage $k \neq t^*$. Formally,

$$\frac{\partial \theta_{t+1,j}}{\partial I_k} = \frac{\partial M_{t,j}\left(I_t, I_{t-1}, \ldots, I_1, \theta_1\right)}{\partial I_k} \equiv 0, \quad k \neq t^*, \quad j = 1, \ldots, J,$$

for all $\theta_1, I_1, \ldots, I_t$, but

$$\frac{\partial \theta_{t+1,j}}{\partial I_{t^*}} = \frac{\partial M_{t,j}\left(I_t, I_{t-1}, \ldots, I_1, \theta_0\right)}{\partial I_{t^*}} > 0, \quad j = 1, \ldots, J,$$

for some $\theta_1, I_1, \ldots, I_t$.

Stage t^* is a *sensitive period* for $\theta_{t+1,j}$ if at the same level of inputs, investment is more productive at stage t^* than at stage t. Formally, t^* is a sensitive period for $\theta_{t+1,j}$ if for $k \neq t^*$,

$$\left.\frac{\partial \theta_{t+1,j}}{\partial I_k}\right|_{\theta_1=\theta_1, I_k=i_k, k=1,\ldots,t, k\neq t^*} \leq \left.\frac{\partial \theta_{t+1,j}}{\partial I_{t^*}}\right|_{\theta_1=\theta_1, I_k=i_k, k=1,\ldots,t}.$$

The inequality is strict for at least one period $k = 1, \ldots, t$, $k \neq t^*$.

This definition of critical periods agrees with the previous one. Our second definition of sensitive periods may not agree with the previous one, which is defined only in terms of the effect of investment on the next period's output. The second definition fixes the period at which output is measured and examines the marginal productivity of inputs in producing the output. It allows for feedback effects of the investment in j on output beyond j through self-productivity in a way that the first definition does not.

At each stage t, agents can perform certain tasks. The level of performance in task l at stage t is $T_{l,t} = T_{l,t}(\theta_t)$. For some tasks, and some stages, components of θ_t may be substitutes or complements. Thus we can, in principle, distinguish complementarity or substitution in skills (abilities) in stage t in task performance from complementarity or substitution in skill production. Agents deficient in some skills may specialize in some tasks. This is an alternative form of remediation compared to remediation through skill investment (see CHLM).

B.2 Relationship with the Ben-Porath (1967) Model

The conventional formulation of the technology of skill formation is due to Ben-Porath, who applied the concept of a production function to the formation of adult skills. Our analysis has many distinctive features which we elaborate after first reviewing his analysis. Let θ_t be scalar human capital. This corresponds to a model with one skill (general human capital). His model postulates that human capital at time $t+1$ depends on human capital at t, invariant ability (denoted κ), and investment at t, I_t. I_t may be a vector. The same type of investments are made at each stage. Skill is measured in the same units over time. His specification of the investment technology is

$$\theta_{t+1} = f(I_t, \theta_t, \kappa)$$

where f is concave in I_t. The technology is specialized further to allow for depreciation of scalar human capital at rate ρ. Thus we obtain

$$\theta_{t+1} = g\left(I_t, \theta_t, \kappa\right) + \left(1 - \rho\right)\theta_t.$$

When $\rho = 0$, there is no depreciation. "θ_t" is carried over (not fully depreciated) as long as $\rho < 1$.

Self-productivity in his model arises when $\frac{\partial \theta_{t+1}}{\partial \theta_t} = \frac{\partial g(\kappa, \theta_t, I_t)}{\partial \theta_t} + (1 - \rho) > 0$. This comes from two sources: a carry over effect, $(1 - \rho) > 0$, arising from the human capital that is not depreciated, and the effect of θ_t on gross investment $\left(\frac{\partial g(\kappa, \theta_t, I_t)}{\partial \theta_t} > 0\right)$. If $g\left(I_t, \theta_t, \kappa\right) = \phi_1\left(\theta_t, \kappa\right) + \phi_2\left(I_t, \kappa\right)$, there is no essential distinction between $\left(1 - \rho\right)\theta_t$ and $g\left(I_t, \theta_t, \kappa\right)$ as sources of self-productivity if we allow ρ to depend on κ $\left(\rho\left(\kappa\right)\right)$.

Complementarity of all inputs is defined as

$$\frac{\partial^2 g\left(I_t, \theta_t, \kappa\right)}{\partial \theta_t \partial I_t'} > 0.$$

In a more general case, some components of this vector may be negative or zero. In the case of universal complementarity, the stock of θ_t raises the marginal productivity of I_t. Direct complementarity and self-productivity, singly and together, show why skill begets skill. Our model generalizes the Ben-Porath model by (a) allowing for different skill formation technologies at different stages to capture the notion of critical and sensitive periods; (b) allowing qualitatively different investments at different stages; (c) allowing for both skill and ability formation and (d) considering the case of vector skills and abilities.

His model features the opportunity cost of time as an essential ingredient. His "neutrality assumption", $\left(\theta_{t+1} = f\left(I_t \theta_t, \kappa\right)\right)$, guarantees that productivity in the market (opportunity costs) increases at the same rate as productivity of human capital in self production. For an analysis of parental investment in young children, child time and its opportunity costs are not relevant. Thus, his neutrality assumption is not relevant. In the original Ben-Porath paper, a Cobb-Douglas technology is used. We allow for more general substitution possibilities among investments.

C An OLG/Complete Markets Model

Consider the following formulation of the problem in the text in a complete markets framework. An individual lives for $2T$ years. The first T years the individual is a child of an adult parent. From age $T+1$ to $2T$ the individual lives as an adult and is the parent of a child. The individual dies at the end of the period in which he is $2T$ years-old, just before his child's child is born. This is an overlapping generations model in which at every calendar year there are an equal and large number of individuals of every age $t \in \{1, 2, ..., 2T\}$.

A household consists of an adult parent, born in generation g, and his child, will be the next generation of this dynasty, generation $g+1$. In what follows, we will use the subscript g to denote the state and control variables of the parent, and $g+1$ to denote those of the child.[6] Furthermore, note that when the child is t years-old, the parent is $T+t$ years-old. It suffices to keep track of the age of the child.

Children are assumed to make no decisions. The parents invest in their children because of altruism. We first consider the case in which all parents have the same preferences and supply labor inelastically. Let $I_{g,t}$ denote parental investments in child skill when the child is t years-old, where $t = 1, 2, \ldots, T$. The output of the investment process is a skill vector.

We now describe how skills evolve over generations. Assume that each agent is born with initial conditions $\theta_{g+1,1}$. Let h_g denote the parental characteristics (e.g., their IQ, education, etc.). Let h_{g+1} denote the level of skills as the child starts adulthood.[7] The technology of skill formation is

$$h_{g+1} = m_2\left(\theta_{g+1,1}, h_g, I_{g,1}, I_{g,2}\right). \tag{C-1}$$

At the beginning of adulthood, the parents of generation g draw two random variables: the initial level of skill of the child, $\theta_{g+1,1}$ and a permanent shock ε_g to their human capital endowment. We use $p\left(\theta_{g+1,1}, \varepsilon_g\right)$ to denote the joint density of these random variables. Let $q\left(\theta_{g+1,1}, \varepsilon_g\right)$ denote the price of a claim that delivers one unit of consumption good if the initial skill level is $\theta_{g+1,1}$ and the permanent shock is ε_g and zero otherwise. Upon reaching adulthood, the parents receive a bequest that contains $b_g\left(\theta_{g+1,1}, \varepsilon_g\right)$

[6] Note that if the parent is born in year y the child will be born in year $y+T$. We define generations in terms not of year of birth but order of birth within a dynasty.

[7] In the text, we use h' to denote this value. Here we use a generationally consistent notation.

of such claims. For every generation g, the state variables for the parent are described by the vector $\psi_g = \left(\varepsilon_g, \theta_{g+1,1}, h_g, b_g\left(\theta_{g+1,1}, \varepsilon_g\right)\right)$.

Let $c_{g,1}$ and $c_{g,2}$ denote the consumption of the household in the first and second period of the lifecycle of the child. The parents decide how to allocate the resources among consumption and investments at different periods as well as bequests in claims $b_{g+1}(\theta, \varepsilon)$ which may be positive or negative. Assuming that human capital (parental and child) is scalar, the budget constraint is

$$c_{g,1}+I_{g,1}+\frac{c_{g,2}+I_{g,2}}{(1+r)}+\frac{\int b_{g+1}(\theta,\varepsilon)\, q(\theta,\varepsilon)\, d\theta d\varepsilon}{(1+r)^2} = wh_g\varepsilon_g+\frac{wh_g\varepsilon_g}{(1+r)}+b_{g+1}\left(\theta_{g+1,1},\varepsilon_g\right). \tag{C-2}$$

Let β denote the utility discount factor and δ denote the parental altruism toward the child. Let $u(.)$ denote the utility function. The recursive formulation of the problem of the parent is

$$V\left(\varepsilon_g, \theta_{g+1,1}, h_g, b_g\left(\theta_{g+1,1}, \varepsilon_g\right)\right) = \max\left\{\begin{array}{c} u\left(c_{g,1}\right)+\beta u\left(c_{g,2}\right)+ \\ +\beta^2\delta E\left[V\left(\varepsilon_{g+1}, \theta_{g+2,1}, h_{g+1}, b_g\left(\theta_{g+2,1}, \varepsilon_{g+1}\right)\right)\right]\end{array}\right\}. \tag{C-3}$$

The problem of the parent is to maximize (C-3) subject to the budget constraint (C-2) and the technology of skill formation (C-1). The first-order conditions for investments are:

$$\delta\beta^2 E\left(\frac{\partial V}{\partial h_{g+1}}\right)\frac{\partial h_{g+1}}{\partial I_{g,1}} = u'\left(c_{g,1}\right) \tag{C-4}$$

$$\delta\beta^2 E\left(\frac{\partial V}{\partial h_{g+1}}\right)\frac{\partial h_{g+1}}{\partial I_{g,2}} = \frac{u'\left(c_{g,1}\right)}{1+r}. \tag{C-5}$$

The levels of investments do not depend on parental resources. Bad draws for the productivity shock ε_g are protected by the financial markets through a larger bequest $b\left(\theta_{g+1,1}, \varepsilon_g\right)$. The early and late investments are determined so that the marginal cost of an extra unit of investment is equal to the marginal expected payoff of that same unit:

$$\delta\beta^2 \sum_{\tau=1}^{\infty} \frac{1}{(1+r)^{2\tau}} E\left[\left(1+\frac{1}{1+r}\right) w\varepsilon_{g+\tau}\frac{\partial h_{g+\tau}}{\partial h_{g+1}}\right]\frac{\partial h_{g+1}}{\partial I_{1,g+1}} = 1 \tag{C-6}$$

$$\delta\beta^2 \sum_{\tau=1}^{\infty} \frac{1}{(1+r)^{2\tau}} E\left[\left(1+\frac{1}{1+r}\right) w\varepsilon_{g+\tau} \frac{\partial h_{g+\tau}}{\partial h_{g+1}}\right] \frac{\partial h_{g+1}}{\partial I_{2,g+1}} = \frac{1}{1+r}. \qquad \text{(C-7)}$$

Derivations for the Dynamic Complete Market Case

To see how to obtain equations (C-6) and (C-7), consider the first-order condition for $b_{g+1}\left(\theta_{g+2,1}, \varepsilon_{g+1}\right)$:

$$q\left(\theta_{g+2,1}, \varepsilon_{g+1}\right) \frac{u'\left(c_{g,1}\right)}{(1+r)^2} = \delta\beta^2 \frac{\partial V}{\partial b_{g+1}\left(\theta_{g+2,1}, \varepsilon_{g+1}\right)} p\left(\theta_{g+2,1}, \varepsilon_{g+1}\right)$$

Suppose that the price of the claim is actuarially fair, so that $q\left(\theta_{g+2,1}, \varepsilon_{g+1}\right) = p\left(\theta_{g+2,1}, \varepsilon_{g+1}\right)$. If we use the Benveniste-Scheinkman Theorem (Benveniste and Scheinkman, 1979) to compute $\frac{\partial V}{\partial b_{g+1}\left(\theta_{g+2,1}, \varepsilon_{g+1}\right)}$ it follows that

$$\frac{u'\left(c_{g,1}\right)}{(1+r)^2} = \delta\beta^2 u'\left(c_{g+1,1}\right). \qquad \text{(C-8)}$$

Let $A_g = w + \frac{w}{(1+r)}$. We apply the Benveniste-Scheinkman Theorem to compute $\frac{\partial V}{\partial h_g}$:

$$\frac{\partial V}{\partial h_g} = u'\left(c_{g,1}\right) A_g \varepsilon_g + \delta\beta^2 E\left[\frac{\partial V}{\partial h_{g+1}} \frac{\partial h_{g+1}}{\partial h_g}\right]. \qquad \text{(C-9)}$$

A recursive application of the Benveniste-Scheinkman Theorem on $\frac{\partial V}{\partial h_{g+1}}$ shows that

$$\frac{\partial V}{\partial h_{g+1}} = u'\left(c_{g+1,1}\right) A_{g+1} \varepsilon_{g+1} + \delta\beta^2 E\left[\frac{\partial V}{\partial h_{g+2}} \frac{\partial h_{g+2}}{\partial h_{g+1}}\right]. \qquad \text{(C-10)}$$

Replacing (C-10) in (C-9) we obtain

$$\frac{\partial V}{\partial h_g} = u'\left(c_{g,1}\right) A_g \varepsilon_g + \delta\beta^2 E\left[u'\left(c_{g+1,1}\right) A_{g+1} \varepsilon_{g+1} \frac{\partial h_{g+1}}{\partial h_g}\right] + \delta^2\beta^4 E\left[\frac{\partial V}{\partial h_{g+2}} \frac{\partial h_{g+2}}{\partial h_g}\right].$$

Continuining with the recursion we conclude that

$$\frac{\partial V}{\partial h_g} = \sum_{\tau=0}^{+\infty} \left(\delta\beta^2\right)^{\tau} E\left[u'\left(c_{g+\tau,1}\right) A_{g+\tau}\varepsilon_{g+\tau}\frac{\partial h_{g+\tau}}{\partial h_g}\right].$$

Now, use (C-8) to obtain

$$\frac{\partial V}{\partial h_g} = u'\left(c_{g,1}\right)\sum_{\tau=0}^{+\infty}\frac{1}{(1+r)^{2\tau}} E\left[A_{g+\tau}\varepsilon_{g+\tau}\frac{\partial h_{g+\tau}}{\partial h_g}\right]. \tag{C-11}$$

By replacing (C-11) into (C-4) and (C-5) we obtain (C-6) and (C-7).

D Program Definitions, Tables and Figures

Descriptions of Intervention Programs Discussed in the Text

Head Start. Head Start is a national program targeted to low-income preschool aged children (ages 3–5) that promotes school readiness by enhancing their social and cognitive development through the provision of educational, health, nutritional, social and other services to enrolled children and families. There is a new program, Early Head Start, that begins at age 1.

Perry Preschool Program. The Perry preschool experiment was an intensive family enhancement preschool program administered to randomly selected disadvantaged black children enrolled in the program over five different waves between 1962 and 1967. Children were enrolled $2\frac{1}{2}$ hours per day, 5 days a week, during the school year and there were weekly $1\frac{1}{2}$-hour home visits. They were treated for 2 years, ages 3 and 4. A control group provides researchers with an appropriate benchmark to evaluate the effects of the preschool program.

The Abecedarian Project. The Abecedarian Project recruited children born between 1972 and 1977 whose families scored high on a "High Risk" index. It enrolls and enriches the family environments of disadvantaged children beginning a few months after birth and continuing until age 5. At age 5—just as they were about to enter kindergarten—all of the children were reassigned to either a school age intervention through age 8 or to a control group. The Abecedarian program was more intensive than the Perry program. Its preschool program was a year-round, full-day intervention.

Chicago Parent-Child Center. The CPC was started in 1967, in selected public schools serving impoverished neighborhoods of Chicago. Using federal funds, the center provided half-day preschool program for disadvantaged 3- and 4-year-olds during the 9 months that they were in school. In 1978, state funding became available, and the program was extended through third grade and included full-day kindergarten.

Tables

Figures

Table D1 Economic Benefits And Costs

	Perry	Chicago CPC
Child Care	986	1,916
Earnings	40,537	32,099
K-12	9,184	5,634
College/Adult	-782	-644
Crime	94,065	15,329
Welfare	355	546
FG Earnings	6,181	4,894
Abuse/Neglect	0	344
Total Benefits	150,525	60,117
Total Costs	16,514	7,738
Net Present Value	134,011	52,380
Benefits-To-Costs Ratio	9.11	7.77

Notes: All values discounted at 3% and are in $2004. Numbers differ slightly from earlier estimates because FG Earnings for Perry and Chicago were estimated using the ratio of FG Earnings Effect to Earnings Effect (about 15%) that was found in Abecedarian

Source: Barnett, 2004.

Table D2

Test Scores and the Timing of Income

White Males, CNLSY/1979

	Ability Factor at ages 12-13[a]		Ability Factor at ages 12-13[a]		Ability Factor at ages 12-13[a]		Ability Factor at ages 12-13[a]	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
Family Permanent Income 0-14[b]	**0.1899**	4.1300	**0.2963**	3.5600	**0.2877**	2.8200	**0.1879**	2.2400
Family Permanent Income 0-4[c]			-0.1224	-1.5800				
Family Permanent Income 5-9[d]					-0.0527	-0.6400		
Family Permanent Income 10-14[e]							0.0346	0.5900
Mother's Ability[f]	**0.2729**	9.5700	**0.2742**	9.4600	**0.2607**	8.9500	**0.2606**	8.8100
Mother's Age at Test Date	0.0070	1.1500	0.0069	1.0700	0.0059	0.9400	0.0053	0.8400
Constant	**-0.9315**	-3.8600	**-0.8847**	-3.5100	**-1.0599**	-4.2400	**-0.9916**	-4.0000
Number of Observations	883		855		871		860	
R^2	0.1748		0.1718		0.1777		0.1739	

[a]The Ability Factor at ages 12-13 is obtained in the following way. First, we standardize the scores in PIAT Math and PIAT Reading Recognition so that at each age each score has mean zero and variance one. Then, we factor analyze the scores at each age and extract one factor. The factor values at age 12 and age 13 are combined to form only one factor. Let f_{12}, f_{13} denote the factors at age 12 and 13, respectively. Let d_{12} take the value one if the factor at age 12 is nonmissing. Let d_{13} take the value one if the factor at age 13 is nonmissing. Then we construct the factor f as $f = d_{12}*f_{12} + (1-d_{12})*d_{13}*f_{13}$.

[b]Family Permanent Income from age 0 to age 14 of the child. It is the average (inflation-adjusted) family income from age 0 to age 14 of the child.

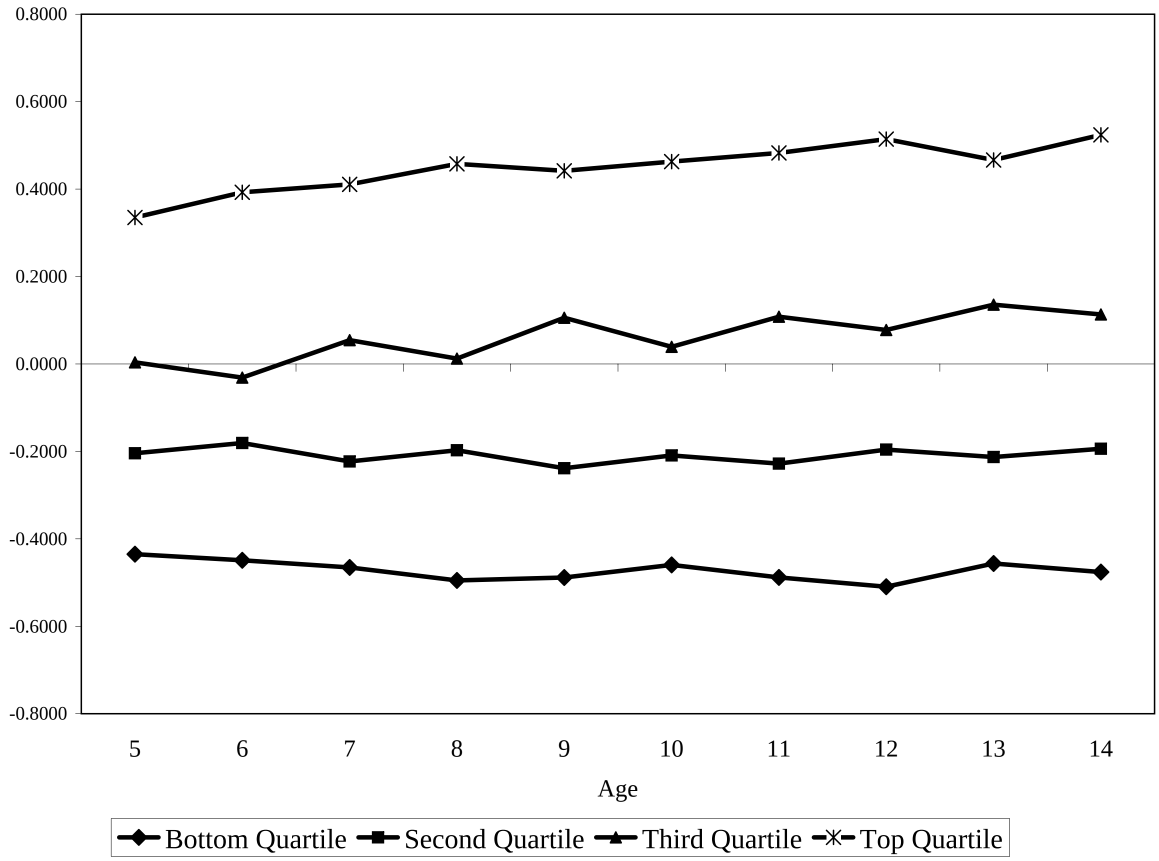
[c]Family Permanent Income from age 0 to age 4 of the child. It is the average (inflation-adjusted) family income from age 0 to age 4 of the child.

[d]Family Permanent Income from age 5 to age 9 of the child. It is the average (inflation-adjusted) family income from age 5 to age 9 of the child.

[e]Family Permanent Income from age 10 to age 14 of the child. It is the average (inflation-adjusted) family income from age 10 to age 14 of the child.

[f]AFQT of the Mother (NLSY/79).

Figure 1
Children of the NLSY
Average Standardized Score
PIAT Math by Permanent Income Quartile



This figure shows the average standardized score in the PIAT Math test from ages 5 to 14 by quartile of family permanent income. The sample consists of all Children of NLSY/79. Family permanent income is the mean family income from age 0 to age 18 of the child. At each age, we standardize the PIAT math score so it has mean zero and variance one. That is, let $m_{i,t}$ denote the score of child i at age t. Let μ_t, σ_t^2 denote the mean and variance of the PIAT-Math score at age t. We construct the variable $z_{i,t}$ as:

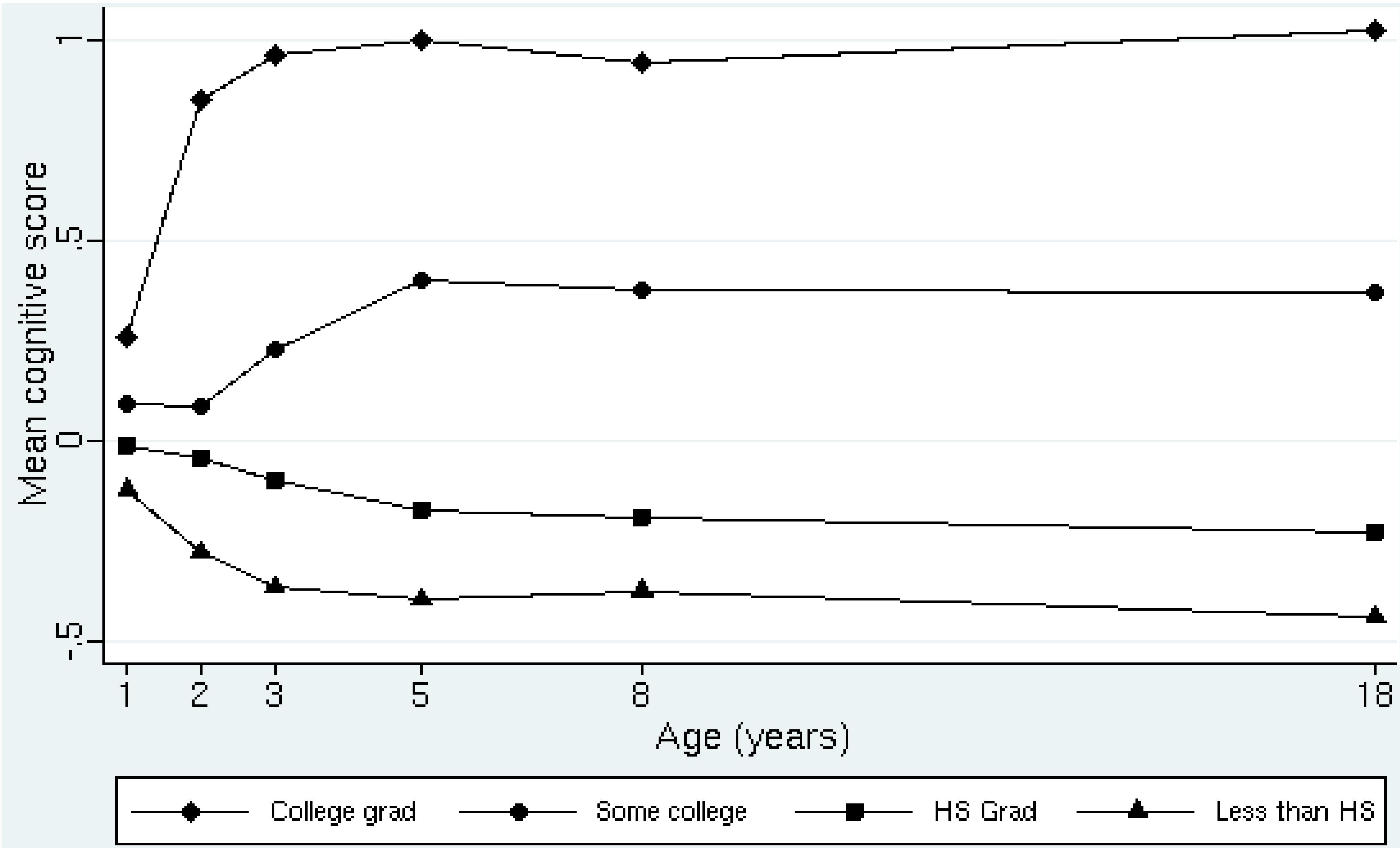
$$z_{i,t} = \frac{m_{i,t} - \mu_t}{\sigma_t}$$

We then proceed by calculating the mean $z_{i,t}$ by quartile of family income. Let $1(q_i = Q_j)$ denote the function that takes the value one if the family permanent income of child i is in quartile Q_j and zero otherwise. Let $\bar{z}_{j,t}$ denote the mean standardized score at age t of the children whose permanent income is in quartile Q_j :

$$\bar{z}_{j,t} = \frac{\sum_i z_{i,t} 1(q_i = Q_j)}{\sum_i 1(q_i = Q_j)}$$

Figure D0

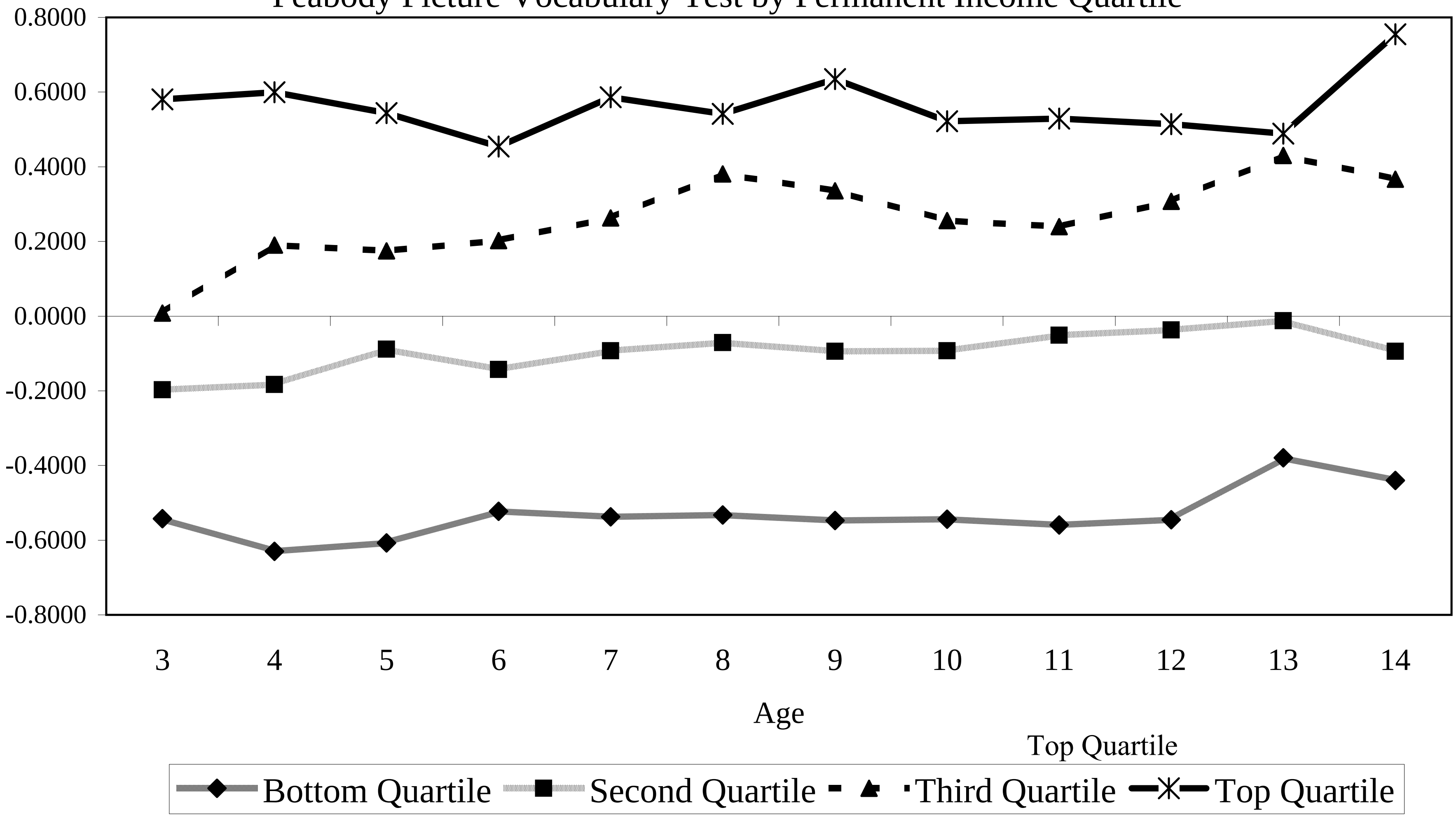
Trend in Mean Cognitive Score by Maternal Education



Source: Brooks-Gunn et al., (2006).

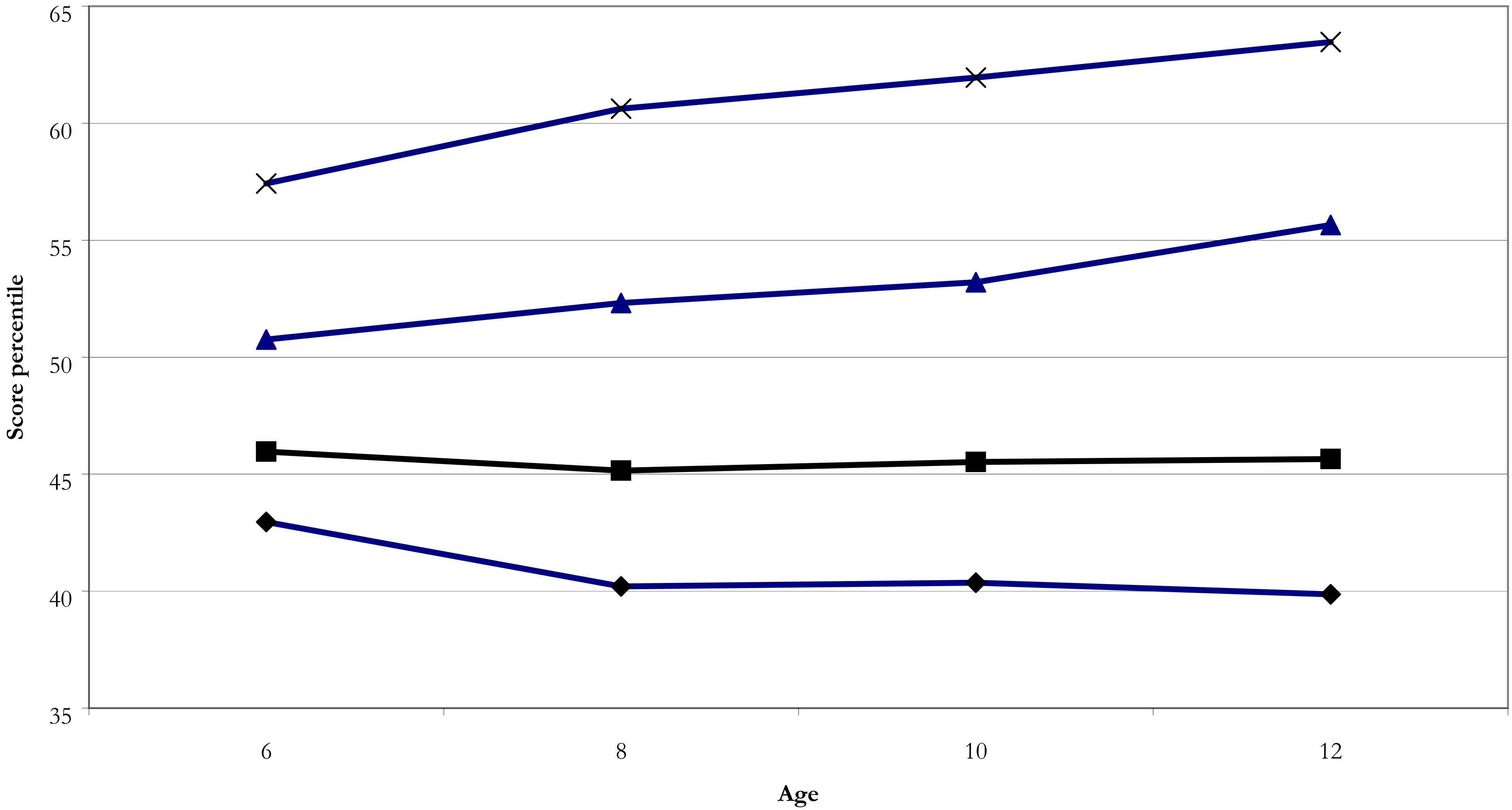
The dramatic results on the importance of the early years in creating differences among children arise if “Bayley scores” are used as a measure of cognition at age 1. As Michael Lewis and Harry McGurk (1972) point out, this is illegitimate since the Bayley score tests other aspects of child development in addition to cognition.

Figure D00
Children of NLSY
Average Standardized Score
Peabody Picture Vocabulary Test by Permanent Income Quartile



Source: Full Sample of Children of the National Longitudinal Survey of Youth.
Please see our website for a full explanation of this figure.

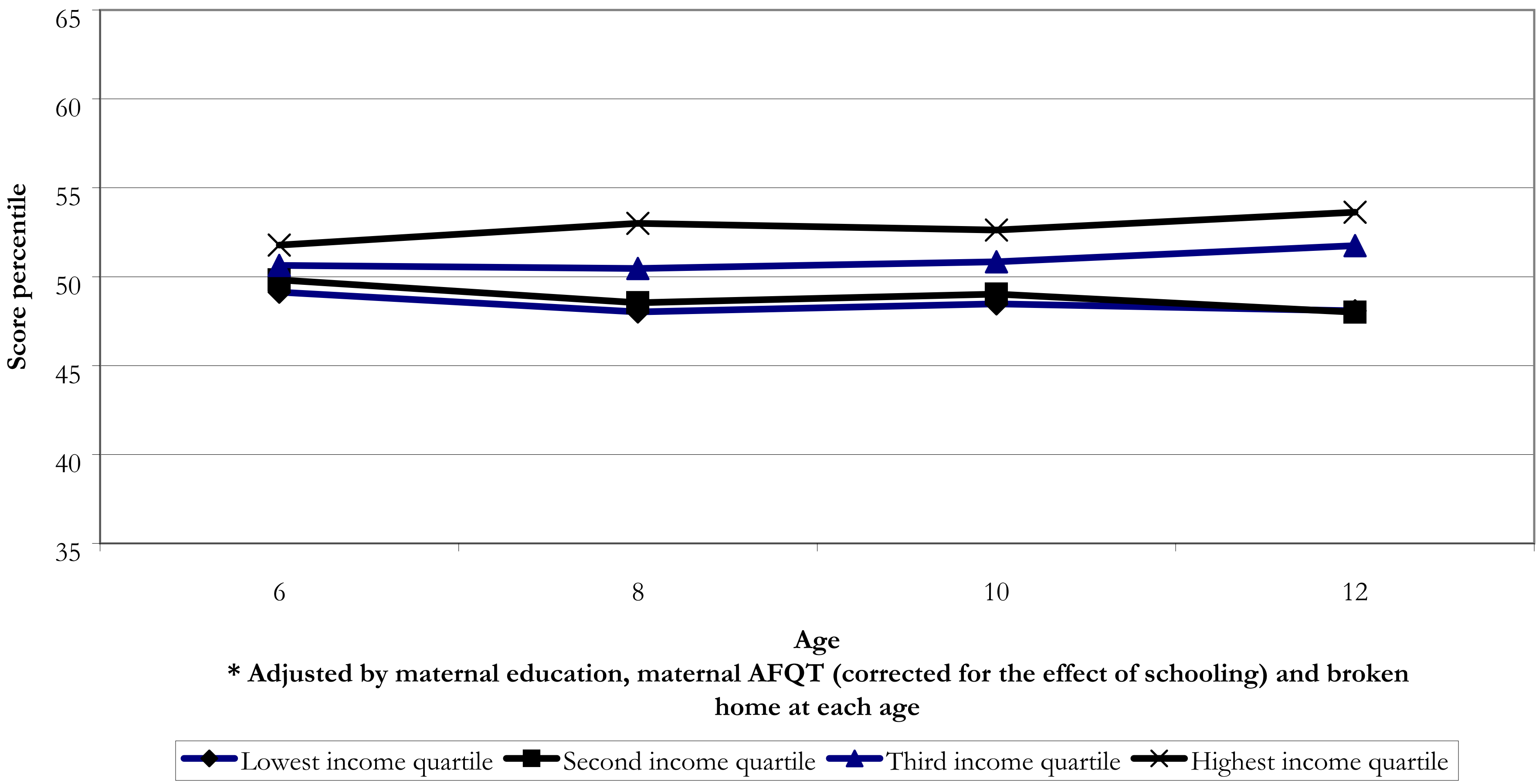
Figure D1a
Children of NLSY
Average percentile rank on PIAT Math score, by income quartile*



***Income quartiles are computed from average family income between the ages of 6 and 10.**

Figure D1b
Children of NLSY

Adjusted average PIAT Math score percentiles by income quartile*



*** Adjusted by maternal education, maternal AFQT (corrected for the effect of schooling) and broken home at each age**

Figure D2a
Children of NLSY

Average percentile rank on anti-social behavior score, by income quartile*

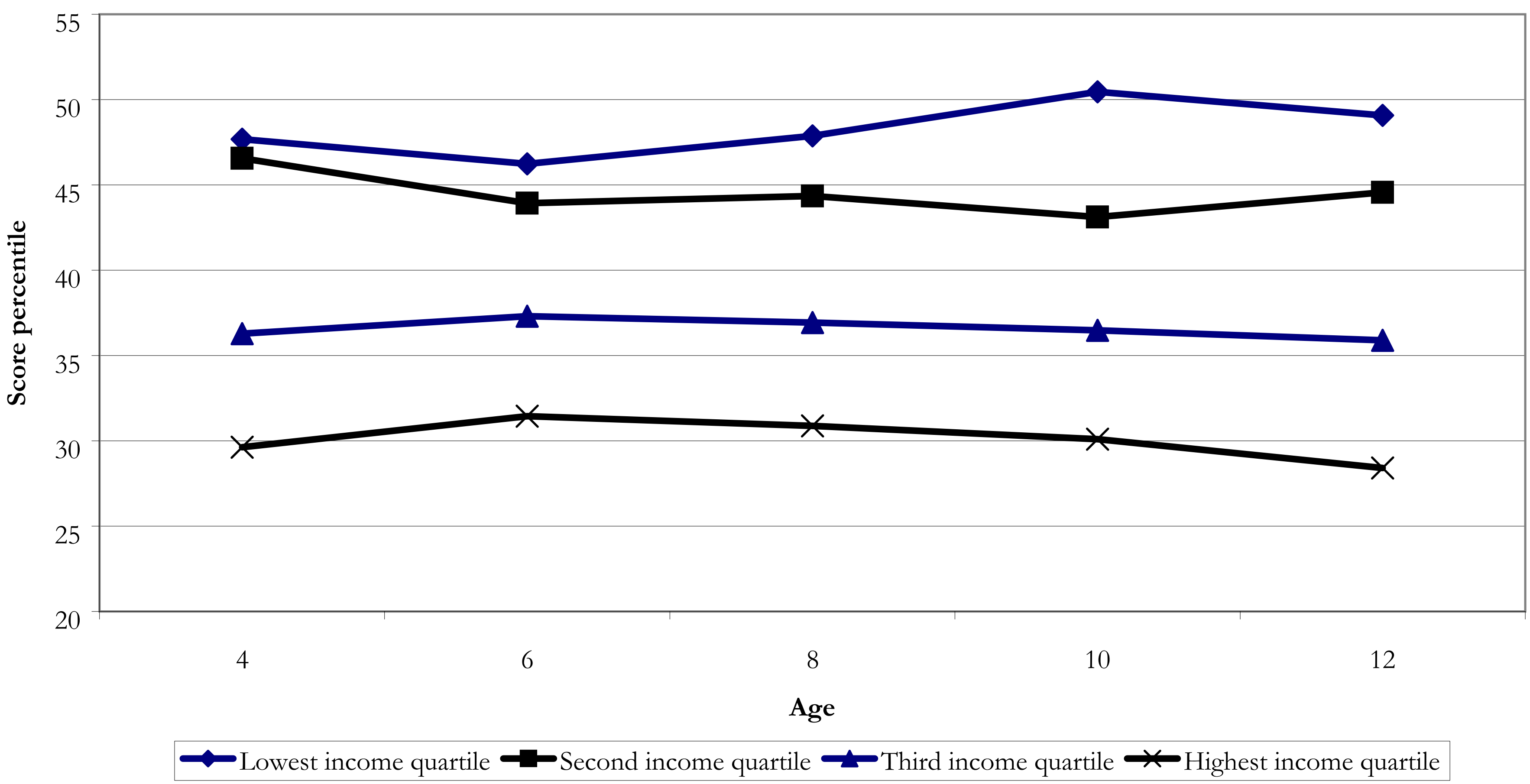
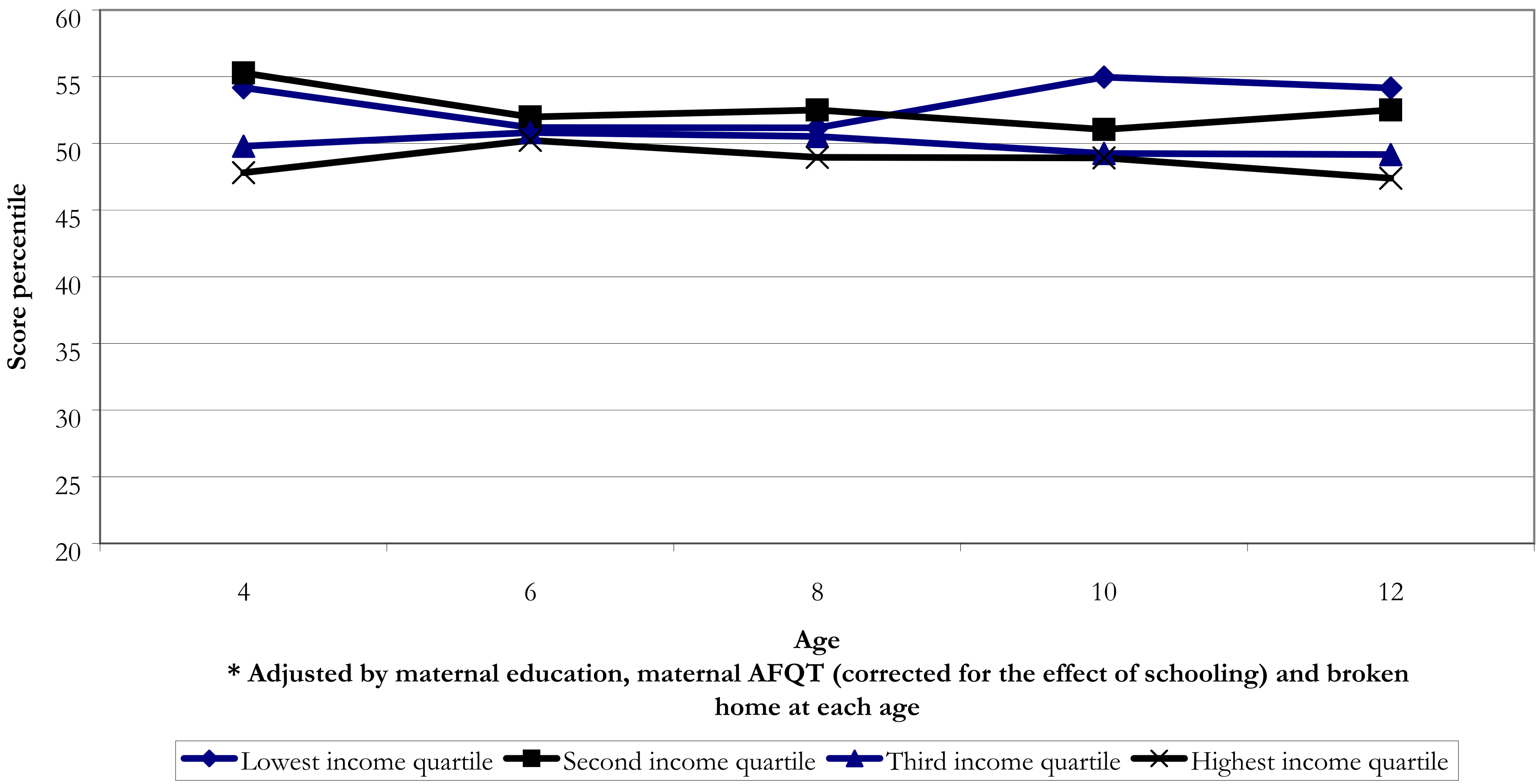


Figure D2b
Children of NLSY

Adjusted average anti-social behavior score percentile by income quartile*



*** Adjusted by maternal education, maternal AFQT (corrected for the effect of schooling) and broken home at each age**

Figure D3a

Average percentile rank on anti-social behavior score, by race

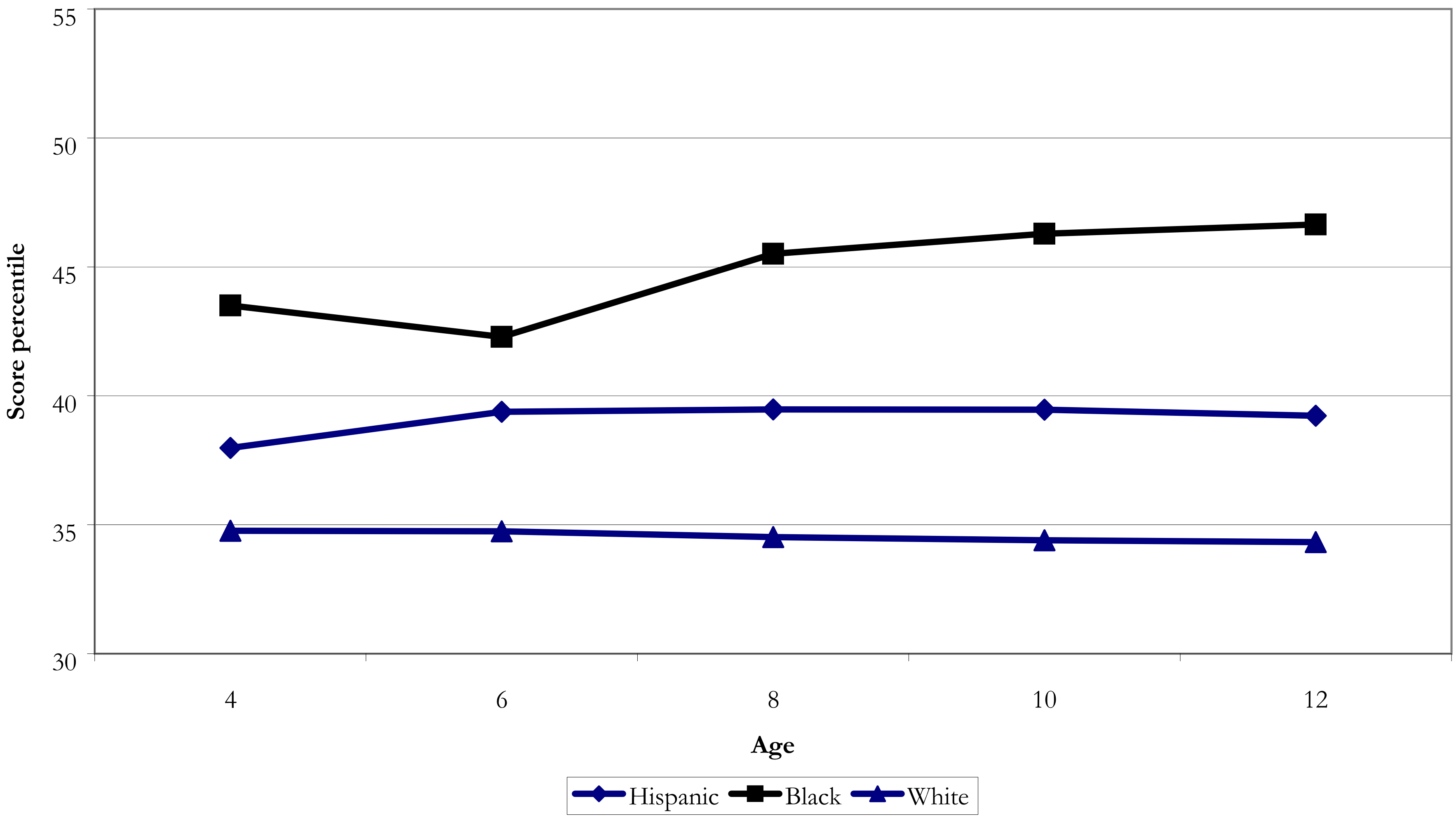
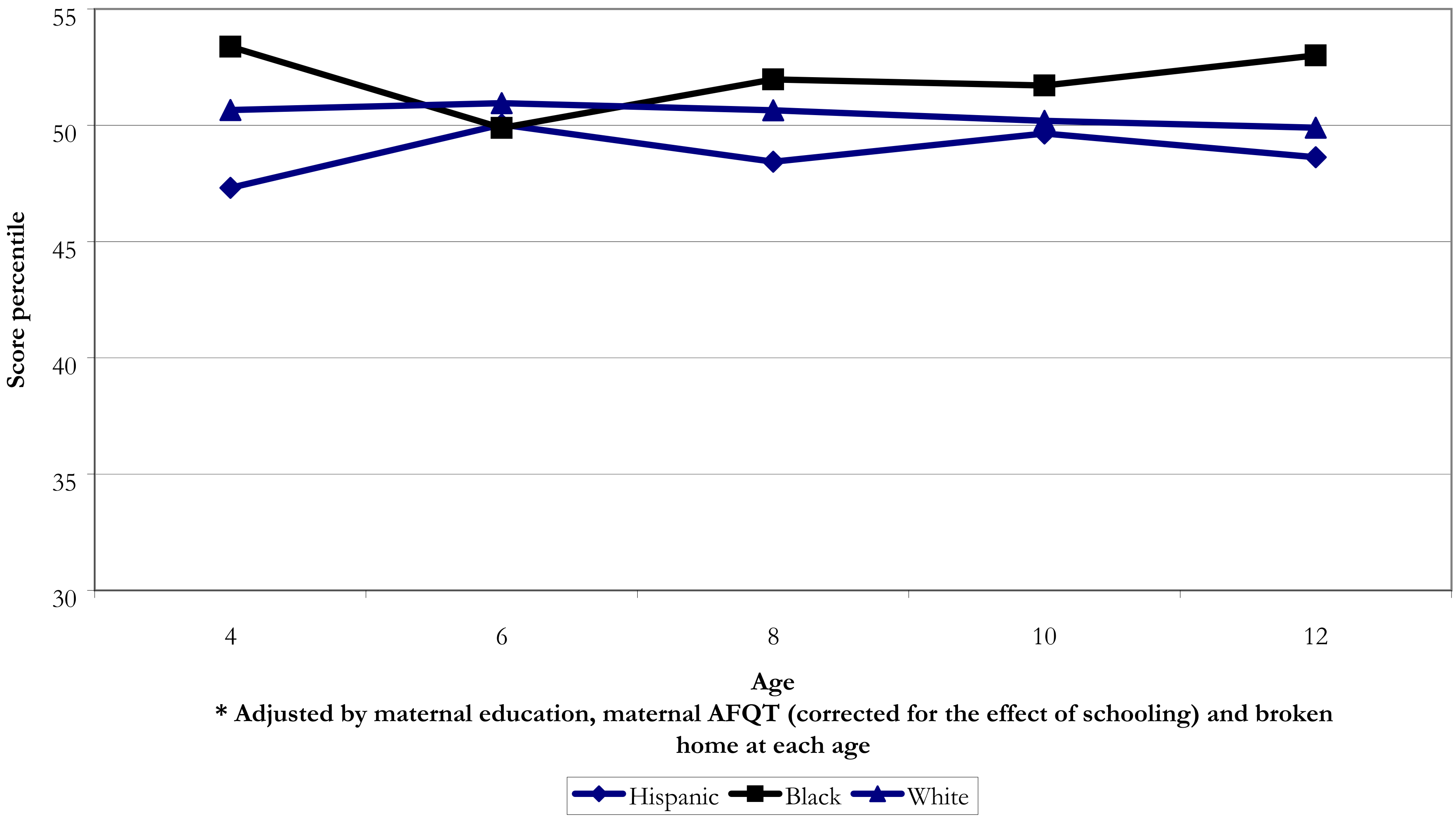


Figure D3b

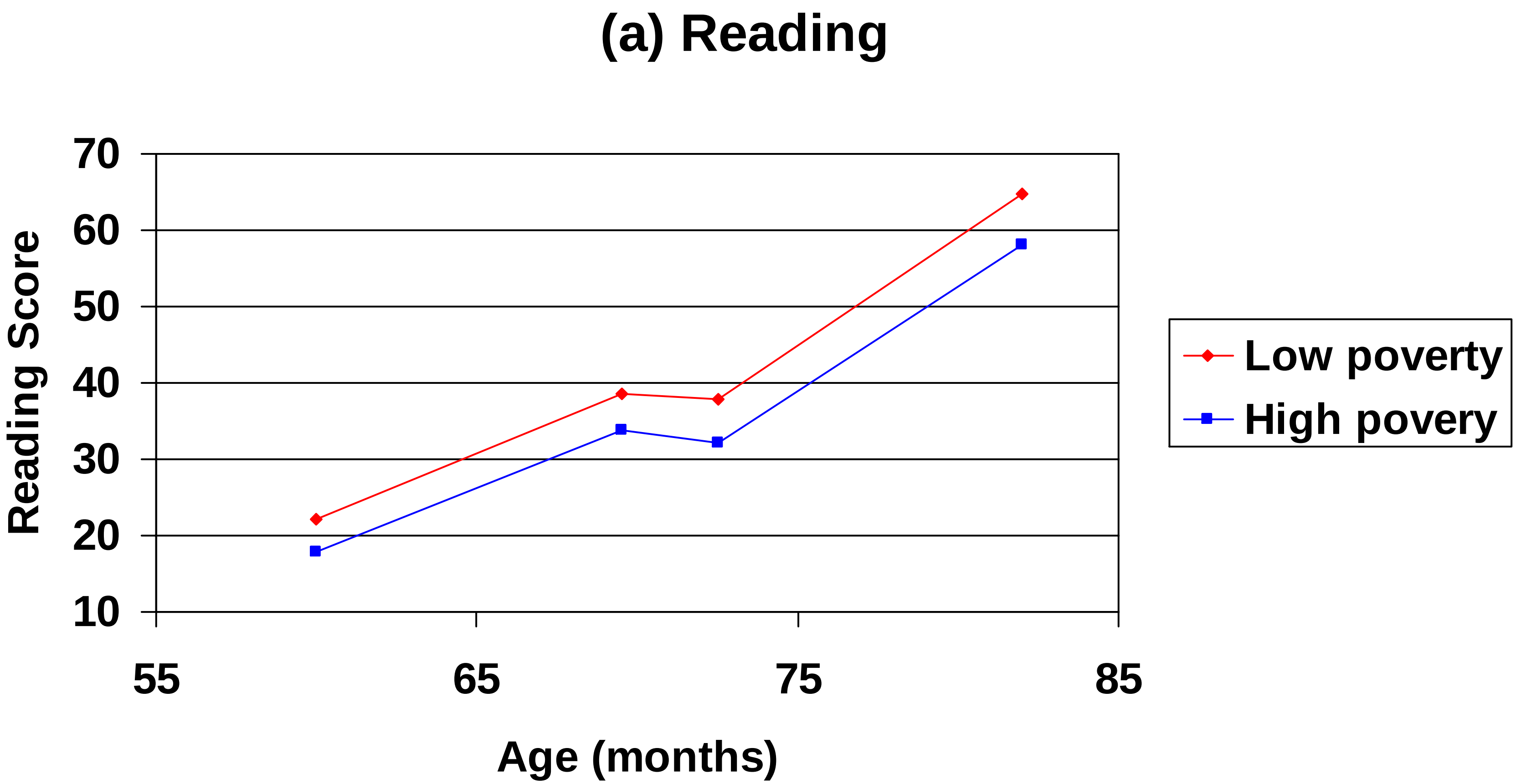
Adjusted average anti-social behavior score percentile by race*



*** Adjusted by maternal education, maternal AFQT (corrected for the effect of schooling) and broken home at each age**

Figure D4a

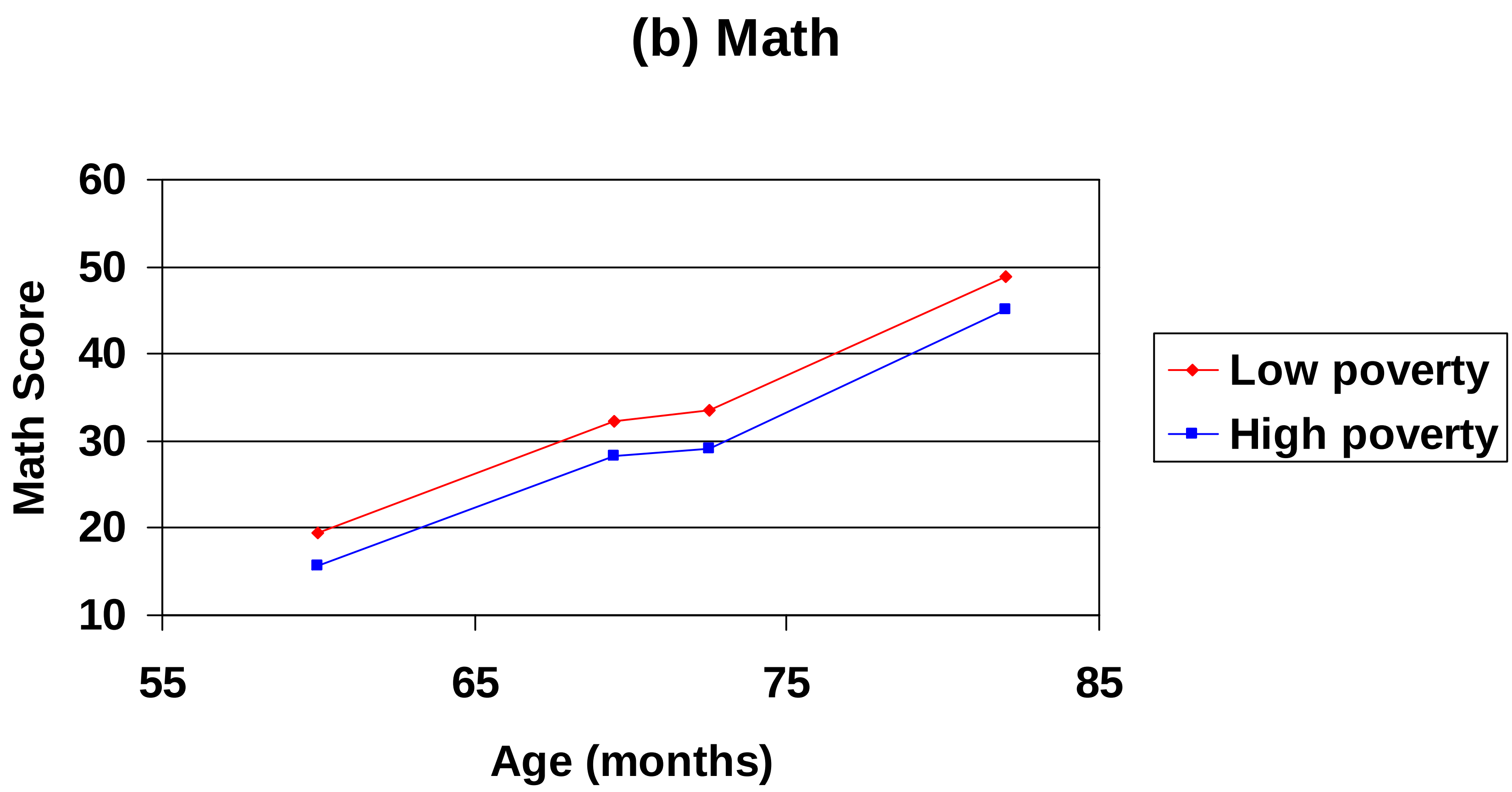
Early Childhood Longitudinal Study (ECLS)



Source: Raudenbush (2006)

Figure D4b

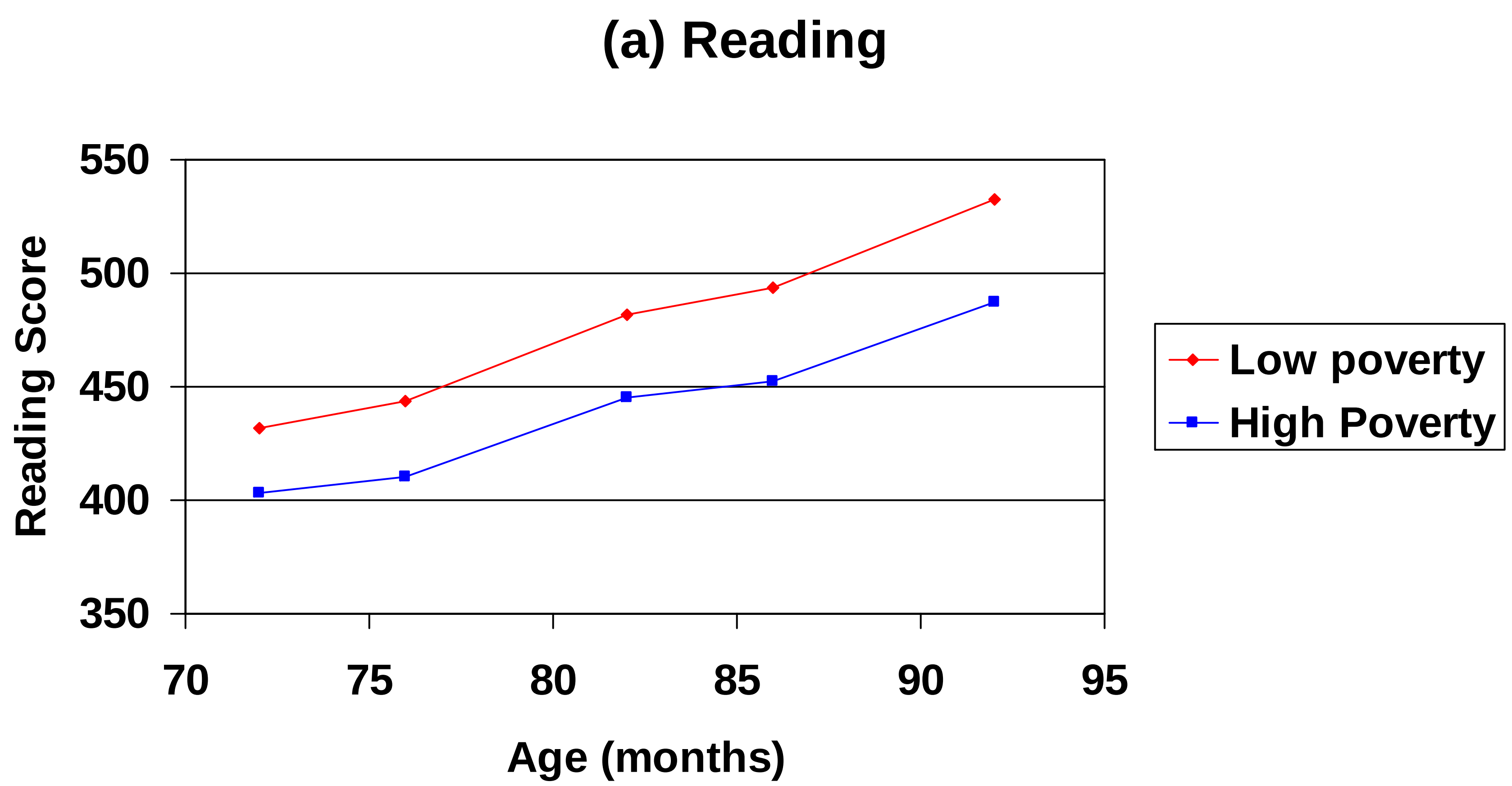
Mean trajectories, high and low poverty schools (ECLS)



Source: Raudenbush (2006)

Figure D5a

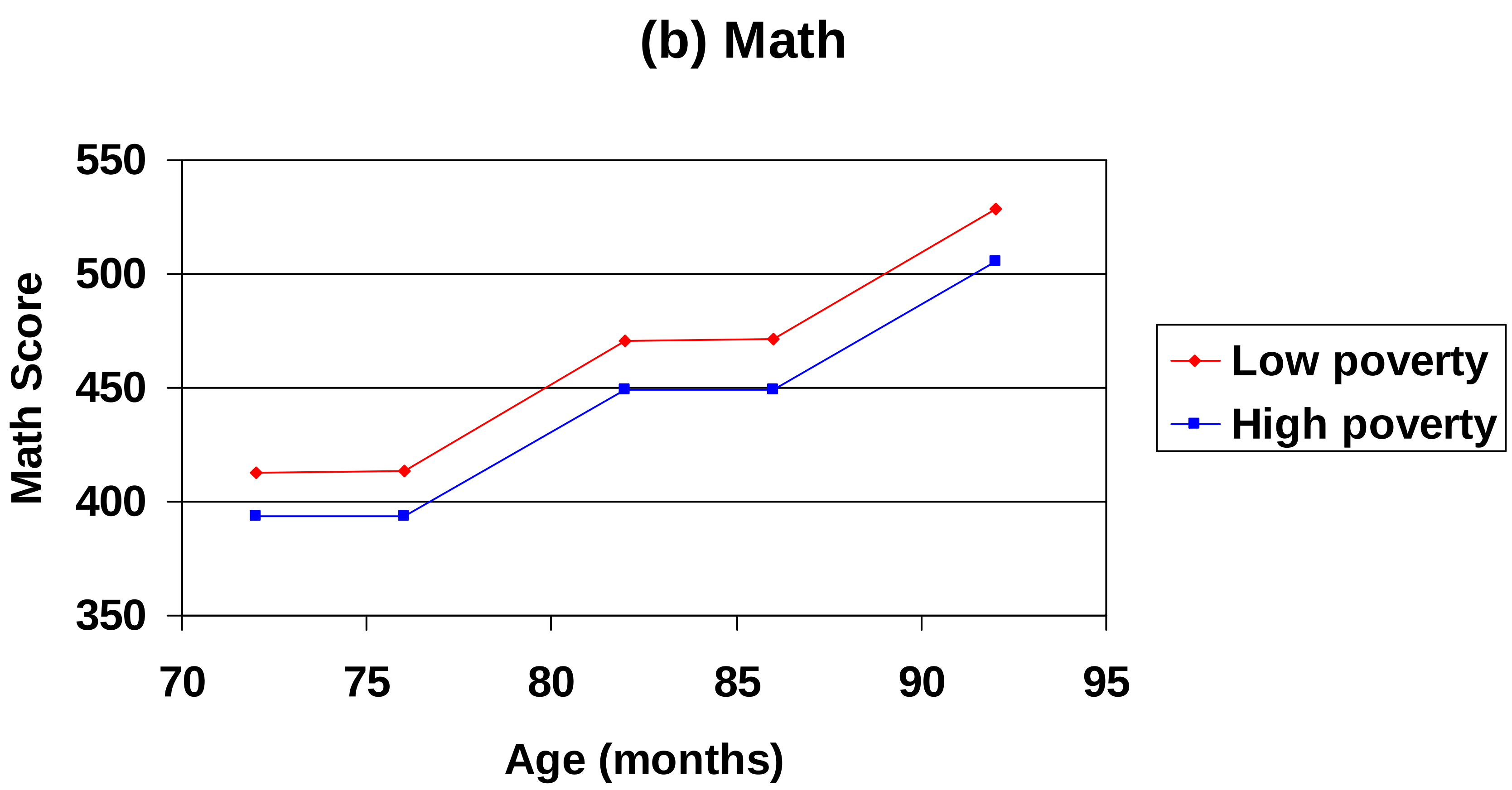
Average trajectories, Grades 1-3, high and low poverty schools (Sustaining Effects Study)



Source: Raudenbush (2006)

Figure D5b

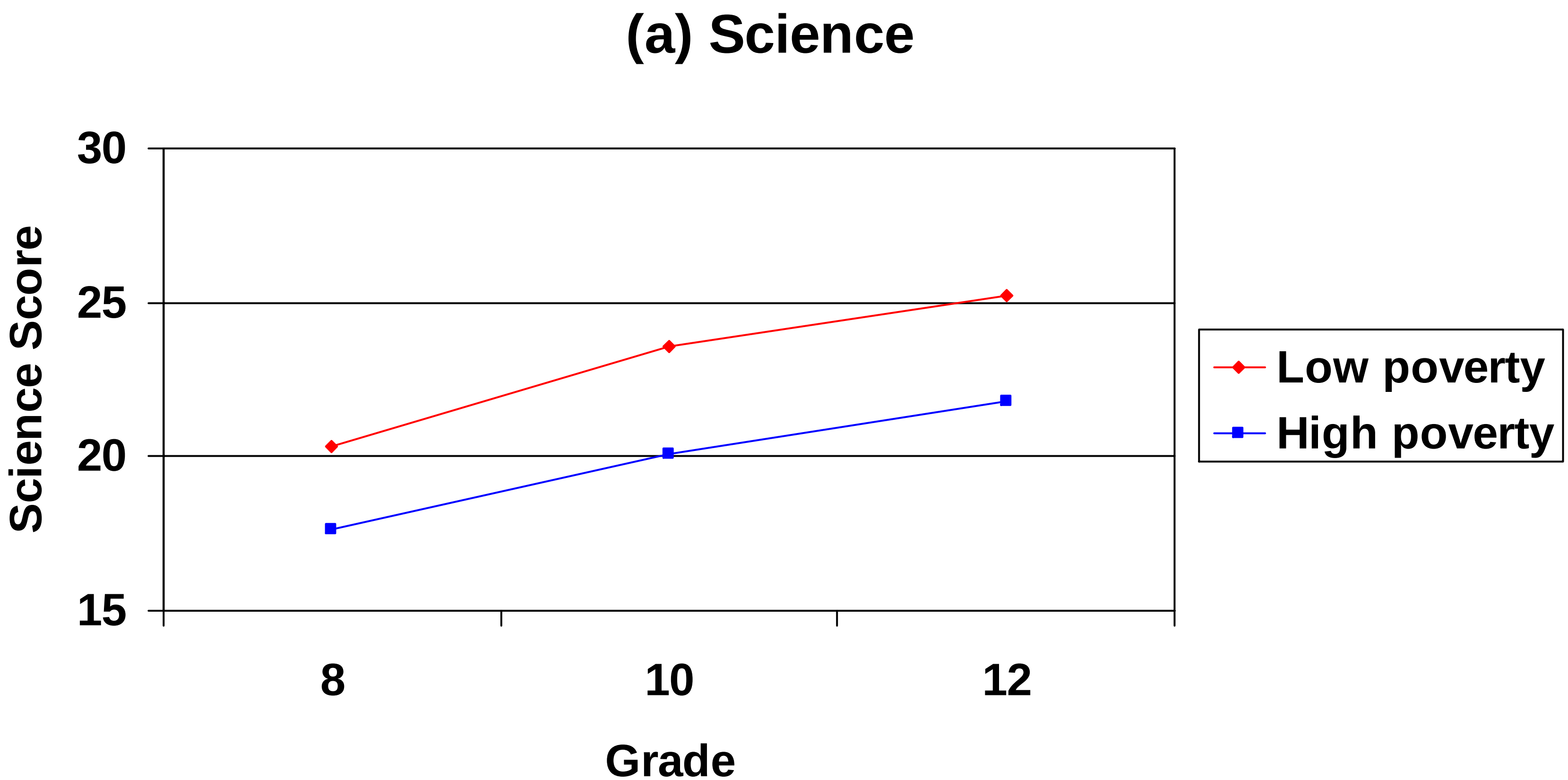
Average trajectories, Grades 1-3, high and low poverty schools (Sustaining Effects Study)



Source: Raudenbush (2006)

Figure D6a

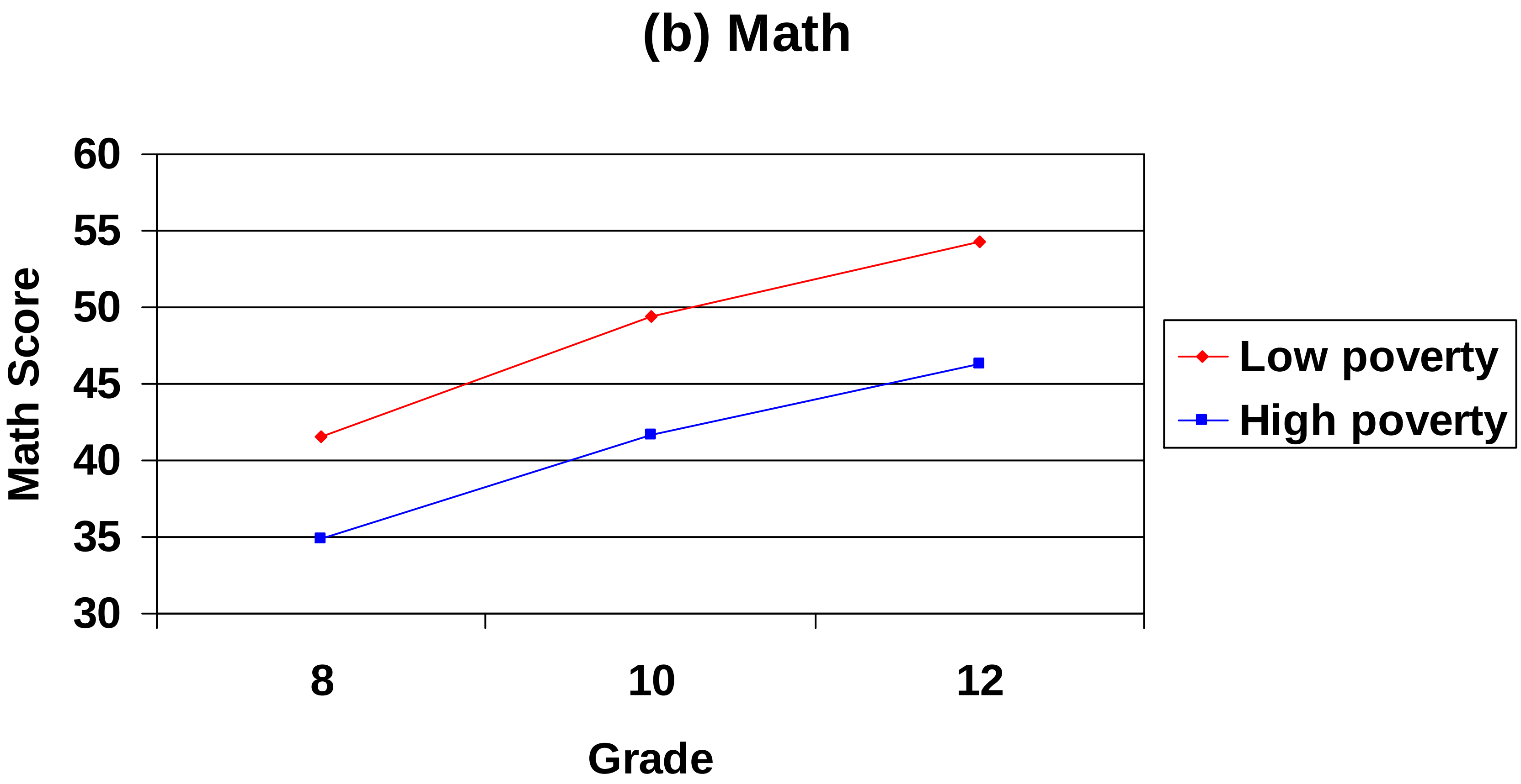
Average achievement trajectories, Grades 8-12 (NELS 88).



Source: Raudenbush (2006)

Figure D6b

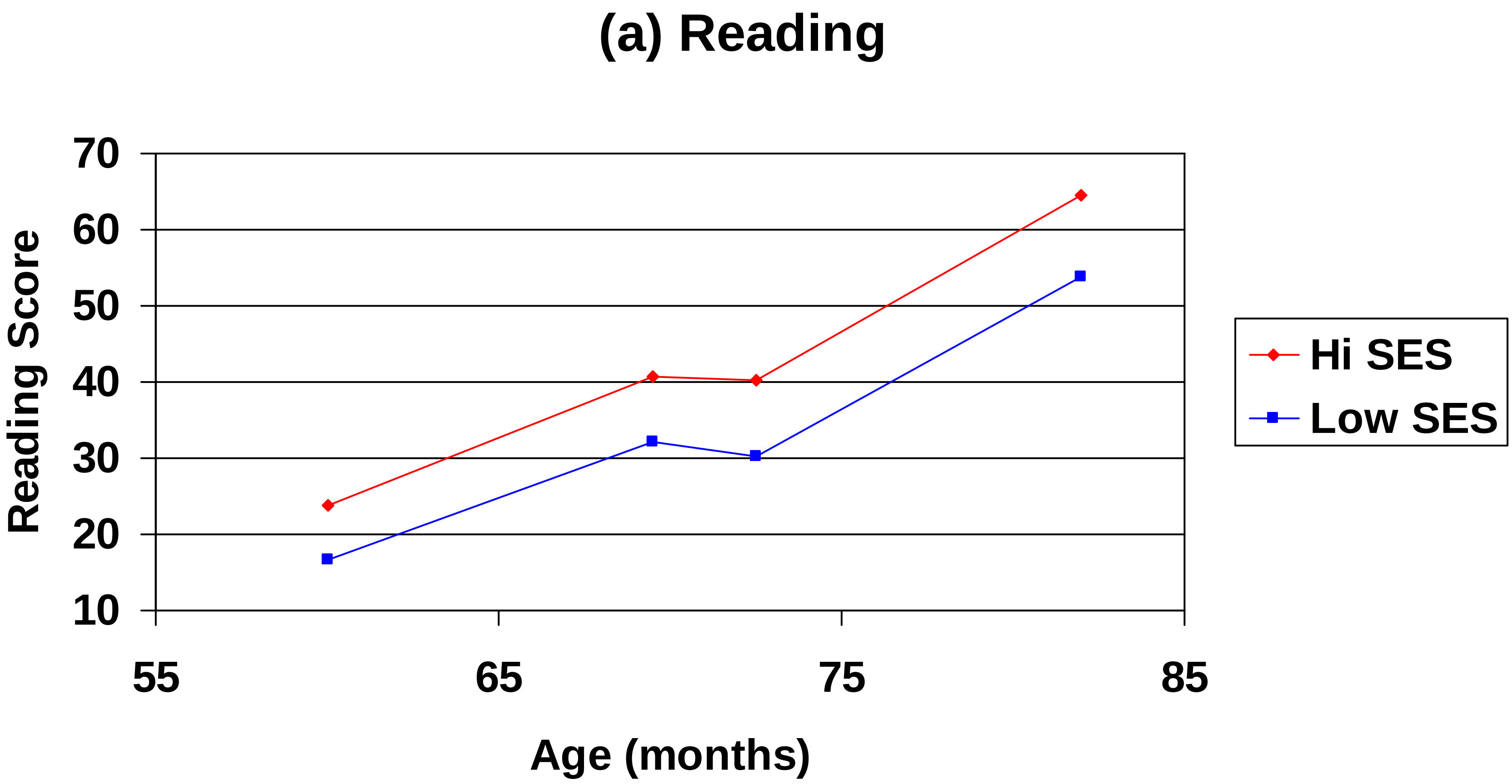
Average achievement trajectories, Grades 8-12 (NELS 88).



Source: Raudenbush (2006)

Figure D7a

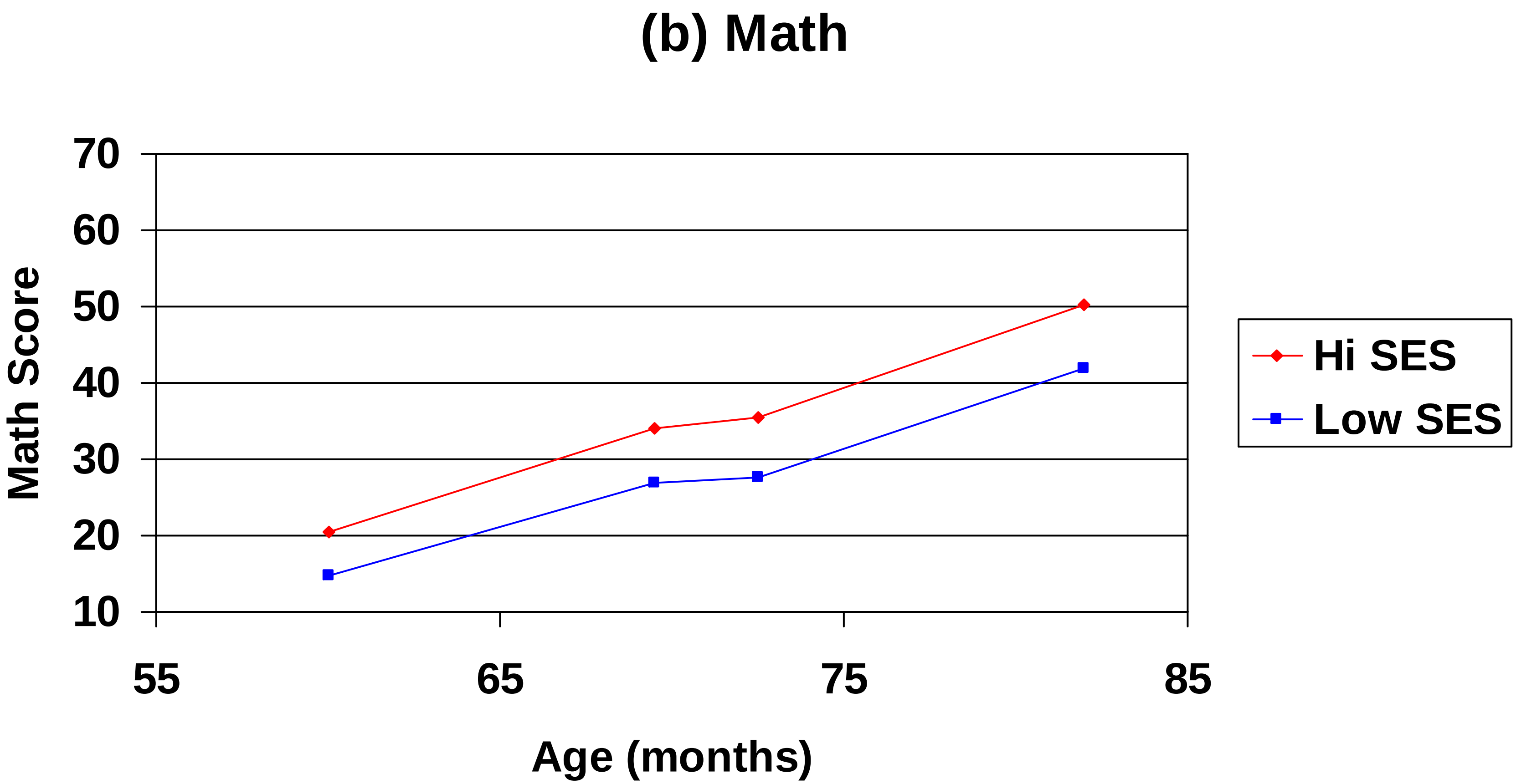
Growth as a function of student social background: ECLS



Source: Raudenbush (2006)

Figure D7b

Growth as a function of student social background: ECLS

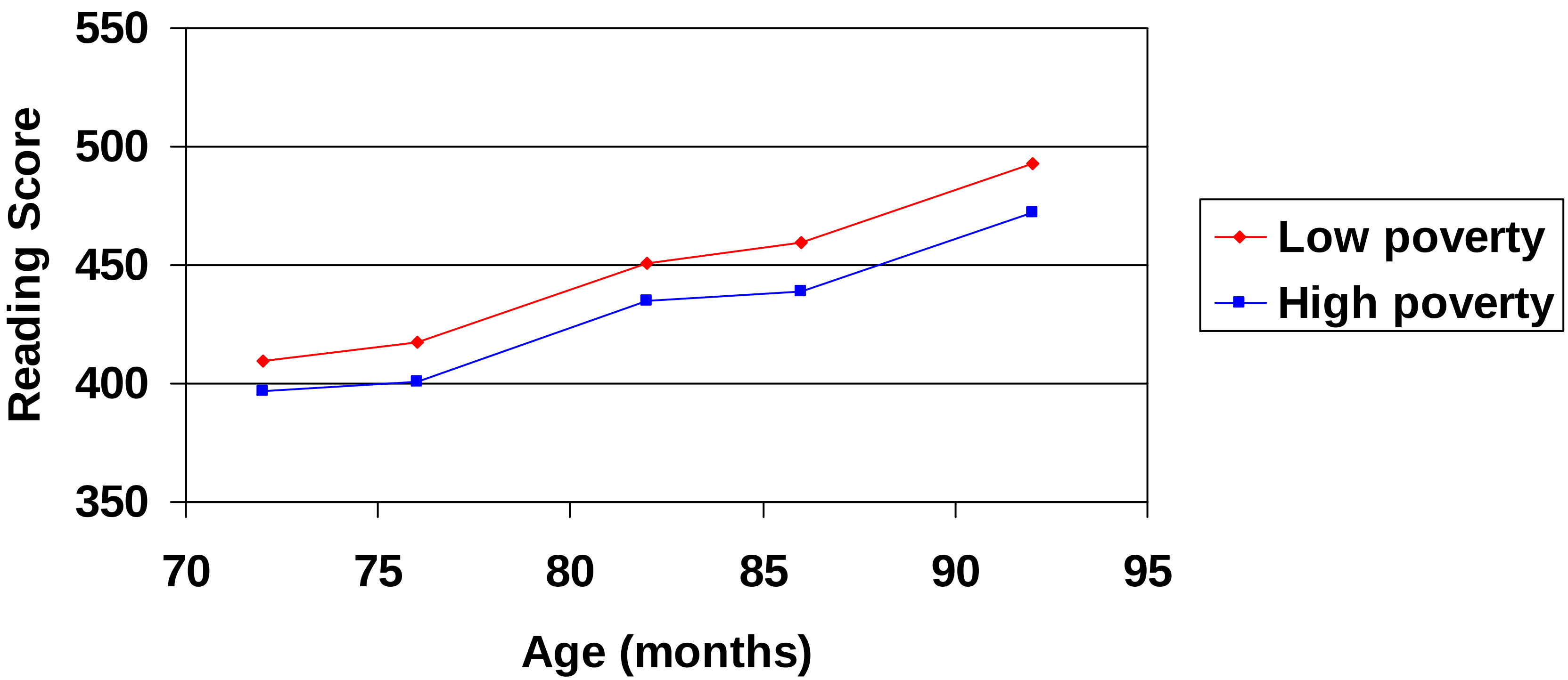


Source: Raudenbush (2006)

Figure D8a

Growth as a Function of School Poverty for Poor Children: Sustaining Effects Data

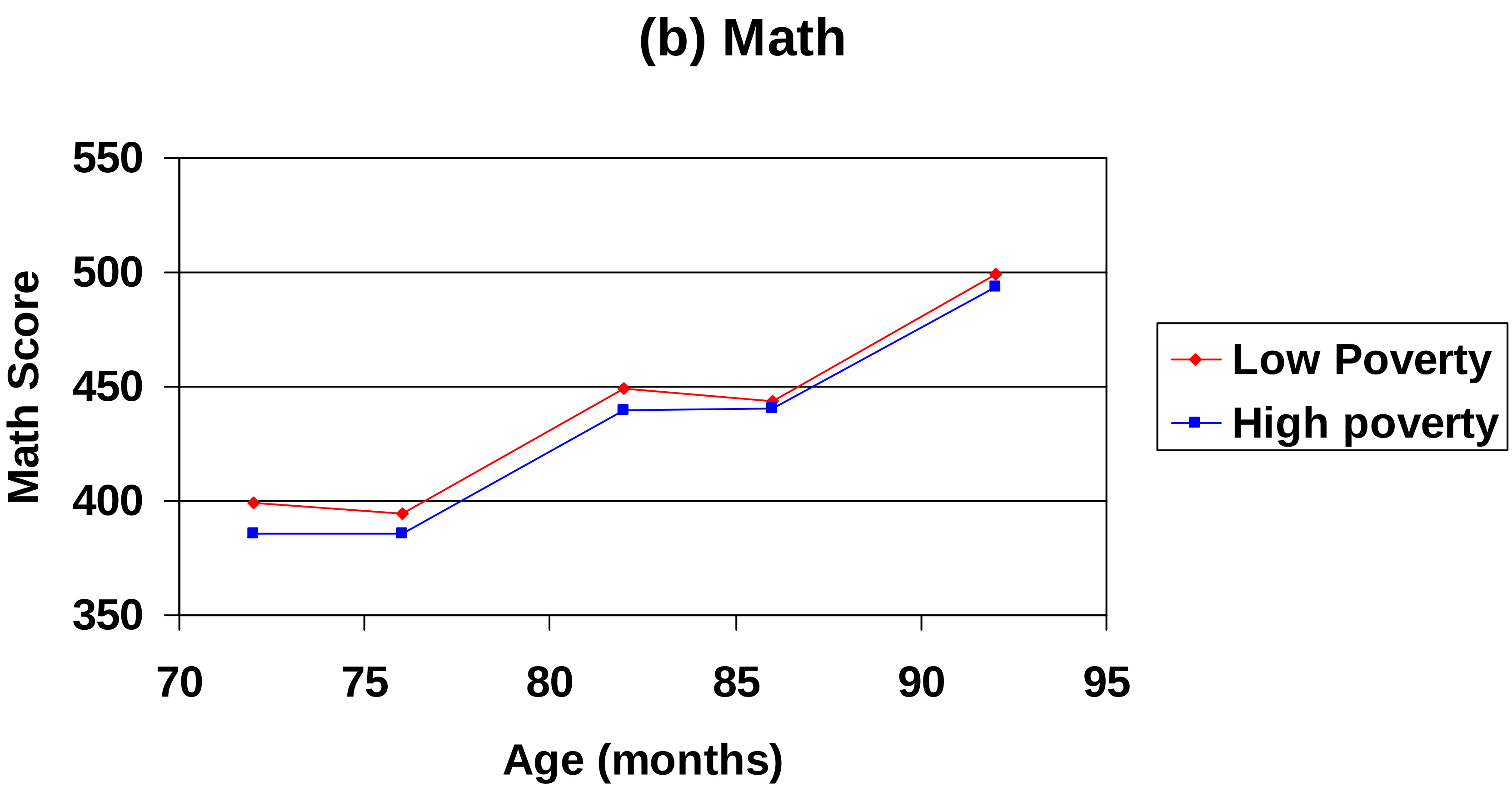
(a) Reading



Source: Raudenbush (2006)

Figure D8b

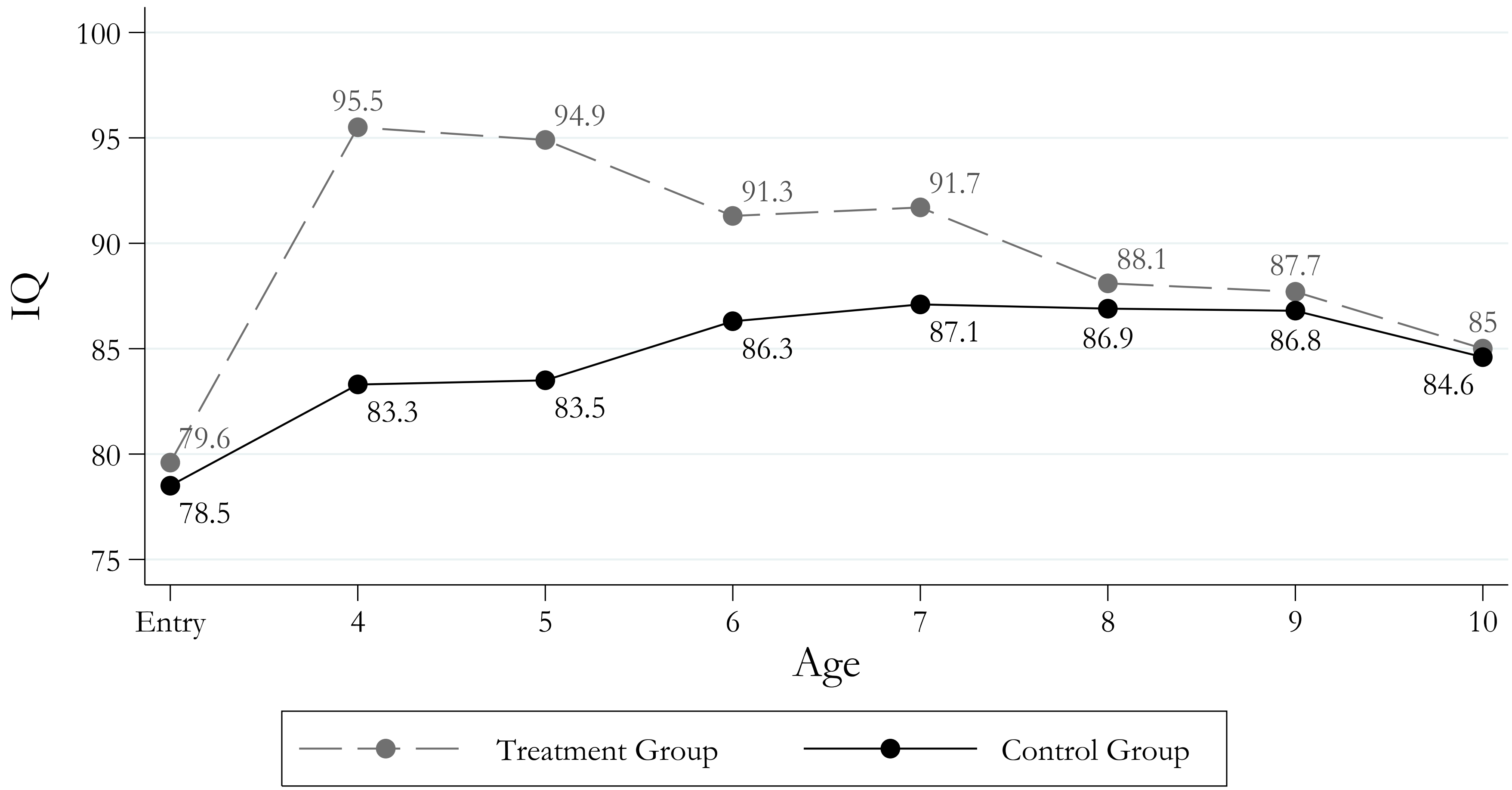
Growth as a Function of School Poverty for Poor Children: Sustaining Effects Data



Source: Raudenbush (2006)

Figure D9a

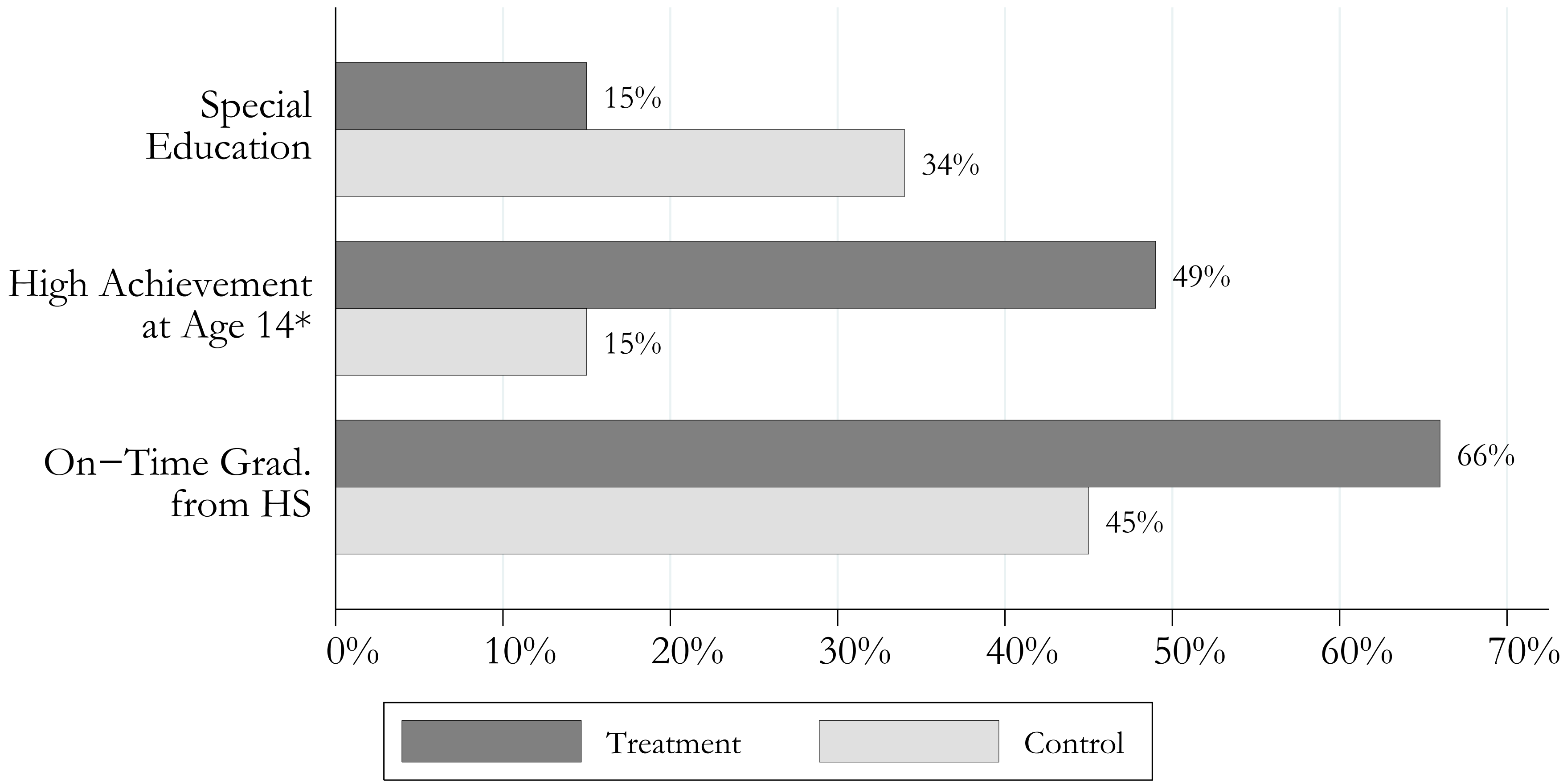
Perry Preschool Program: IQ, by Age and Treatment Group



Source: Perry Preschool Program. IQ measured on the Stanford−Binet Intelligence Scale (Terman & Merrill, 1960). Test was administered at program entry and each of the ages indicated.

Figure D9b

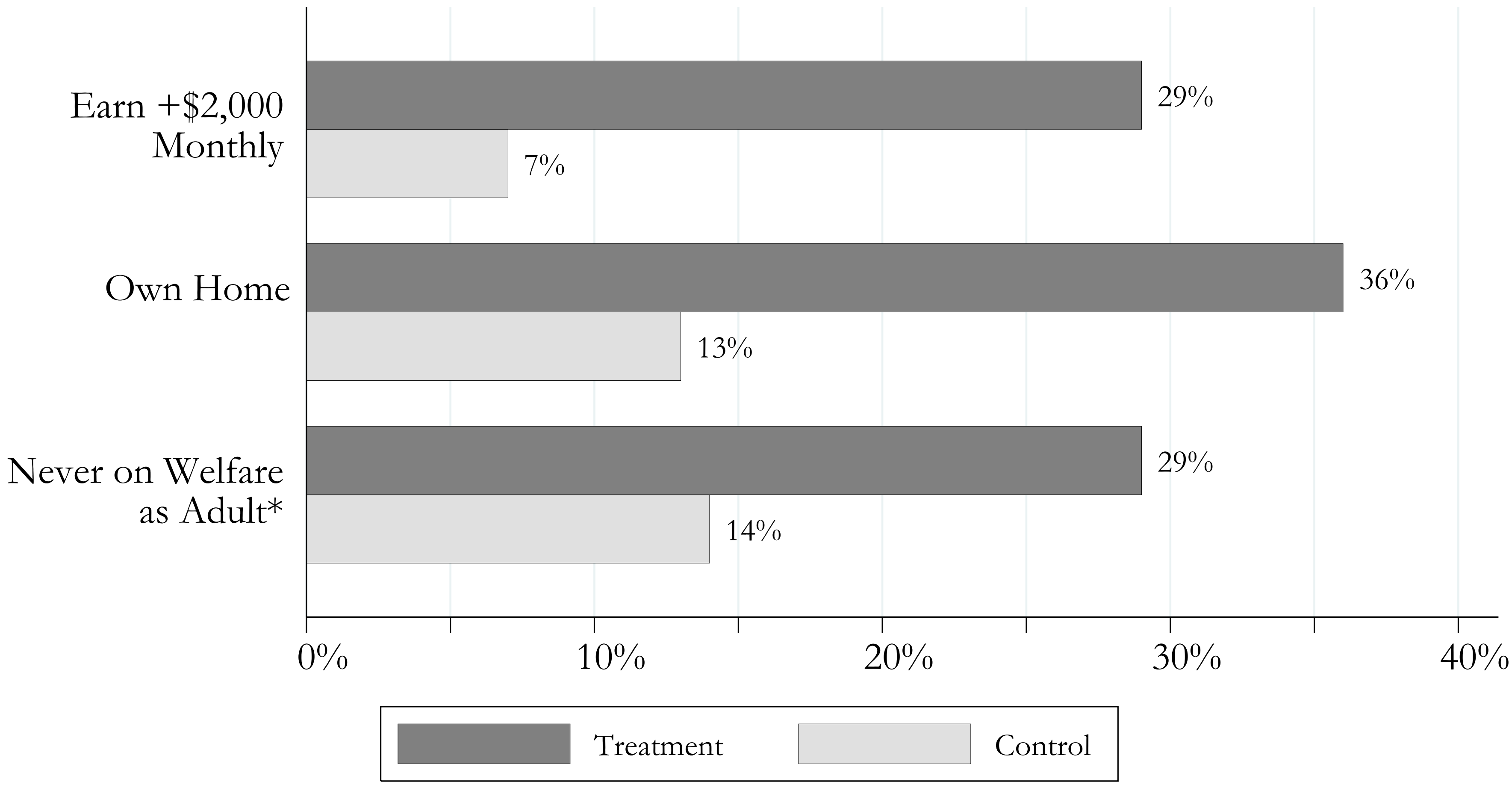
Perry Preschool Program: Educational Effects, by Treatment Group



Source: Barnett (2004).
Notes: *High achievement defined as performance at or above the lowest 10th percentile on the California Achievement Test (1970).

Figure D9c

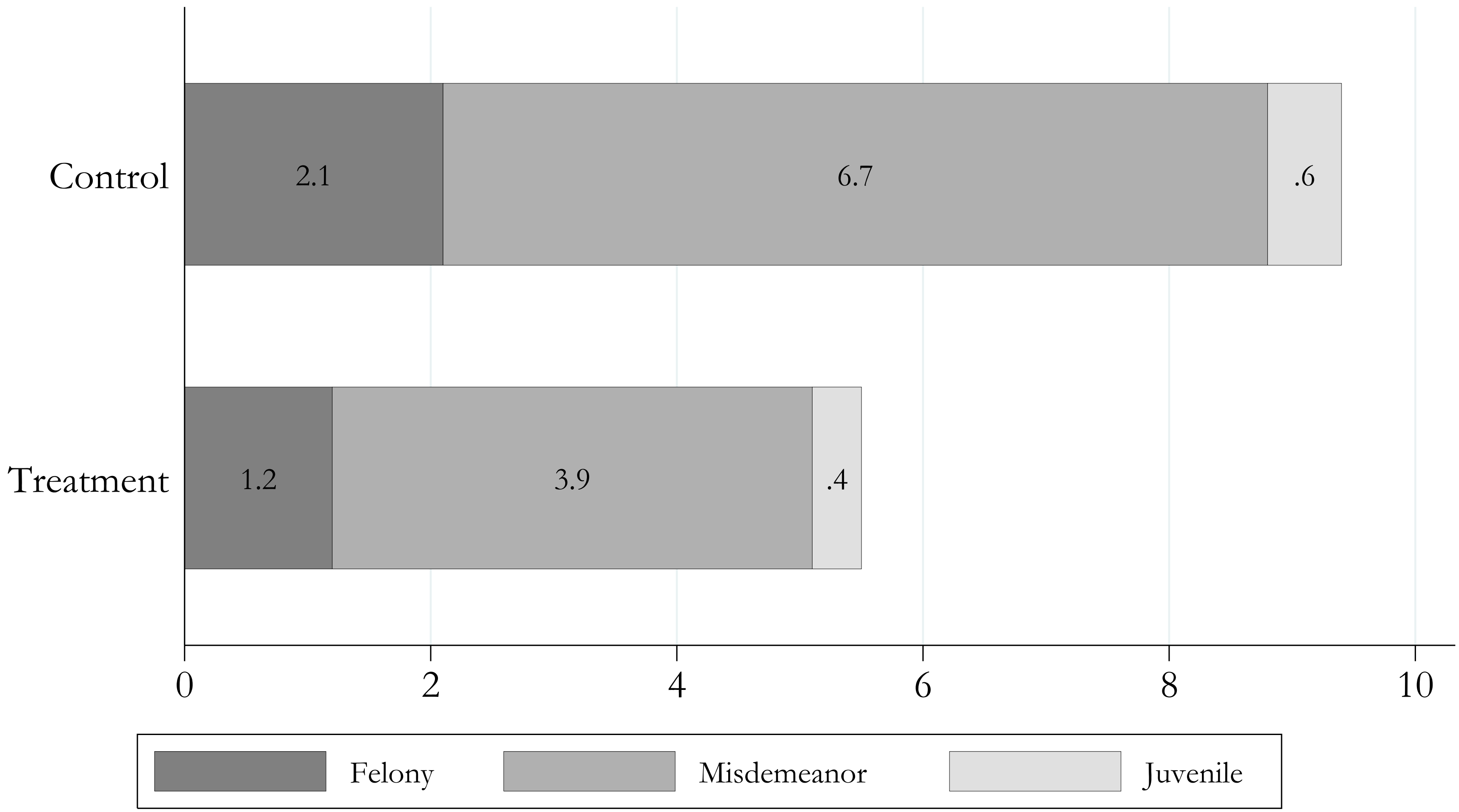
Perry Preschool Program: Economic Effects at Age 27, by Treatment Group



Source: Barnett (2004). *Updated through Age 40 using recent Perry Preschool Program data, derived from self−report and all available state records.

Figure D9d

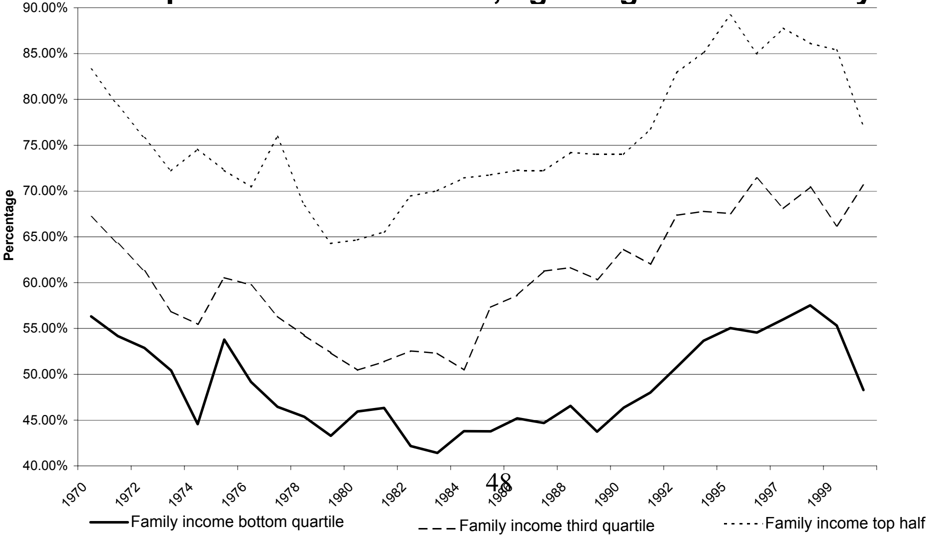
Perry Preschool Program: Arrests per Person before Age 40, by Treatment Group



Source: Perry Preschool Program. Juvenile arrests are defined as arrests prior to age 19.

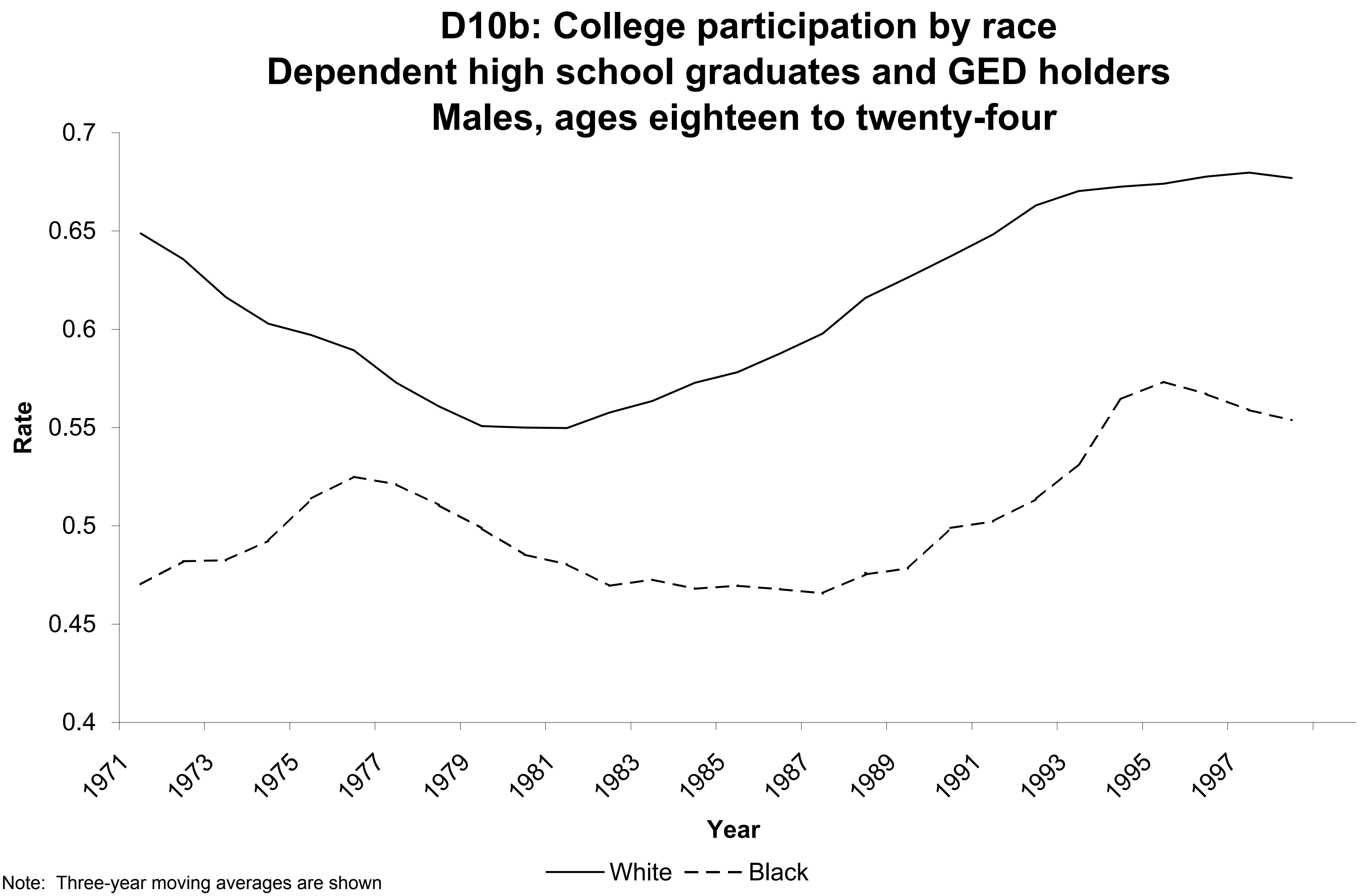
D10a: College participation of HS graduates and GED holders

Dependent* white males, ages eighteen to twenty-four



Source: Computed from the CPS P-20 School Reports and the October CPS.

*Dependent is living at parental home or supported by parental family while at college.



Note: Three-year moving averages are shown

Source: Computed from the CPS P-20 School Reports and the October CPS.

*Dependent is living at parental home or supported by parental family while at college.

References

Becker, G. S. and N. Tomes (1986, July). Human capital and the rise and fall of families. *Journal of Labor Economics 4*(3, Part 2), S1–S39.

Ben-Porath, Y. (1967, August). The production of human capital and the life cycle of earnings. *Journal of Political Economy 75*(4, Part 1), 352–365.

Benveniste, L. M. and J. A. Scheinkman (1979, May). On the differentiability of the value function in dynamic models of economics. *Econometrica 47*(3), 727–732.

Brooks-Gunn, J., F. Cunha, G. Duncan, J. J. Heckman, and A. Sojourner (2006). A reanalysis of the IHDP program. Unpublished manuscript, Infant Health and Development Program, Northwestern University.

Carneiro, P., F. Cunha, and J. J. Heckman (2003, October 17). Interpreting the evidence of family influence on child development. In *The Economics of Early Childhood Development: Lessons for Economic Policy*, Minneapolis, Minnesota. The Federal Reserve Bank. Presented at "The Economics of Early Childhood Development: Lessons for Economic Policy Conference," Minneapolis Federal Reserve Bank, Minneapolis, MN. October 17, 2003.

Cunha, F. and J. J. Heckman (2007a). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. Unpublished manuscript, University of Chicago, Department of Economics. Forthcoming, *Journal of Human Resources*.

Cunha, F. and J. J. Heckman (2007b, May). The technology of skill formation. *American Economic Review 97*(2). Forthcoming.

Heckman, J. J. and Y. Rubinstein (2001, May). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review 91*(2), 145–149.

Heckman, J. J., J. Stixrud, and S. Urzua (2006, July). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Keating, D. P. and C. Hertzman (Eds.) (1999). *Developmental Health and the Wealth of Nations: Social, Biological, and Educational Dynamics*. New York: Guilford Press.