Appendix: Estimating the Technology of Children's Skill Formation

Francesco Agostinelli
W.P. Carey School of Business,
Arizona State University

Matthew Wiswall
W.P. Carey School of Business,
Arizona State University
University of Wisconsin-Madison
NBER

July 14, 2016

A Proofs

A.1 Proof of Lemma 1

Proof. First, we note that with $Z_{t,m}$, $G_t(\theta_t, I_t)$, $\mu_{t,m}$ and $\lambda_{t,m}$ known, we then identify the distribution of the measurement error, given by $F_{\epsilon_{t,m}}(\epsilon) = pr(\epsilon_{t,m} \leq \epsilon)$. Define $\widetilde{Z}_{t,m} = \ln\theta_t + \frac{\epsilon_{t,m}}{\lambda_{t,m}}$ and its characteristic function $\varphi_{\widetilde{Z}_{t,m}}(x) = E\left[e^{ix\left(\frac{Z_{t,m}-\mu_{t,m}}{\lambda_{t,m}}\right)}\right]$. Define $\varphi_{\ln\theta_t}(x) = E\left[e^{ix\ln\theta_t}\right]$ to be the characteristic function of $\ln\theta_t$. Given the independence between $\epsilon_{t,m}$ and $\ln\theta_t$ (Assumption 1), we can rewrite the characteristic function for $\frac{\epsilon_{t,m}}{\lambda_{t,m}}$ to be:

$$\varphi_{\frac{\epsilon_{t,m}}{\lambda_{t,m}}}(x) = \frac{\varphi_{\widetilde{Z}_{t,m}}(x)}{\varphi_{\ln\theta_{t,m}}(x)}$$

Given the one-to-one mapping between characteristic functions and distributions, we identify the marginal density of $\frac{\epsilon_{t,m}}{\lambda_{t,m}}$. Since $\lambda_{t,m}$ is known, we also identify the marginal density of $\epsilon_{t,m}$, $F_{\epsilon_{t,m}}(\epsilon)$.

Next, consider the following conditional expectation:

$$E(Z_{t+1,m}|\ln\theta_t = a, \ln I_t = \ell) = \mu_{t+1} + \lambda_{t,m}E(\ln\theta_{t+1}|\ln\theta_t = a, \ln I_t = \ell)$$
$$+E(\epsilon_{t+1,m}|\ln\theta_t = a, \ln I_t = \ell)$$

where $E(\epsilon_{t+1,m}|\ln\theta_t = a, \ln I_t = \ell) = 0$ given Assumption 1 ($\epsilon_{t+1,m}$ independent of $\ln\theta_t$ and $\ln I_t$).

Iterating expectations and substituting for $\ln\theta_t = \frac{Z_{t,m}-\mu_{t,m}-\epsilon_{t,m}}{\lambda_{t,m}}$, we have the following:

$$E(Z_{t+1,m}|\ln\theta_t = a, \ln I_t = \ell) = \int E(Z_{t+1,m}|\frac{Z_{t,m}-\mu_{t,m}-\epsilon}{\lambda_{t,m}} = a, \ln I_t = \ell, \epsilon)dF_{\epsilon_{t,m}}(\epsilon)$$

Again applying Assumption 1 ($\epsilon_{t,m}$ independent of $Z_{t+1,m}$), we have

$$= \int E(Z_{t+1,m}|\frac{Z_{t,m}-\mu_{t,m}-\epsilon}{\lambda_{t,m}} = a, \ln I_t = \ell)dF_{\epsilon_{t,m}}(\epsilon)$$

Re-writing again, we have

$$= \int E(Z_{t+1,m}|Z_{t,m} = \lambda_{t,m}a + \mu_{t,m} + \epsilon, \ln I_t = \ell)dF_{\epsilon_{t,m}}(\epsilon)$$

$$= \int E(Z_{t+1,m}|Z_{t,m} = b(\epsilon), \ln I_t = \ell) dF_{\epsilon_{t,m}}(\epsilon)$$

Note that for each realization of $\epsilon_{t,m} = \epsilon$, we have $Z_{t,m} = b(\epsilon)$, where $b(\epsilon)$ is known given $\mu_{t,m}, \lambda_{t,m}$, and a are known. We identify the conditional expectation $E(Z_{t+1,m}|Z_{t,m} = b(\epsilon), \ln I_t = \ell)$ from the observed distribution of $Z_{t+1,m}$ and $Z_{t,m}$ measures. Because the distribution of measurement errors $F_{\epsilon_{t,m}}(\epsilon)$ is identified, we identify $E(Z_{t+1,m}|\ln\theta_t = a, \ln I_t = \ell)$.

■

Example 1 *Consider the case where* $\epsilon_{t,m} \sim N(0, \sigma^2_{t,m})\ \forall t$. *We identify* $\sigma^2_{t,m}$ *from* $V(Z_{t,m}) = \lambda^2_{t,m} V(\ln\theta_t) + V(\epsilon_{t,m})$ *since we have already identified* $V(\ln\theta_t)$ *and* $\lambda_{t,m}$. *The idea of the proof of Lemma 1 is that the value of the current latent skills (*$\ln\theta_t = a$*) comes both from observable measure (*$Z_{t,m}$*) and unobservable measurement error (*$\epsilon_{t,m}$*). Since we identify the distribution of the unobservable, we are able to integrate out each possible realization of that unobservable random variable. Indeed, if* ϵ *takes value 0, because we are fixing* $\ln\theta_t$ *to be equal to* a*, this implies that* $Z_{t,m}$ *would equal:*

$$Z_{t,m} = \lambda_{t,m} \cdot a + \mu_{t,m} = b(0)$$

where both $\lambda_{t,m}$ *and* $\mu_{t,m}$ *are known. Hence weight* $E(Z_{t+1,m}|Z_{t,m} = b(0), \ln I_t = \ell)$ *with the likelihood of the event that* ϵ *takes the value of zero. Because* $\epsilon_{t,m} \sim N(0, \sigma^2_{t,m})$*, we have that the marginal density of the measurement error is*

$$f_{\epsilon_{t,m}}(\epsilon) = \frac{1}{\sigma_{t,m}\sqrt{2\pi}} e^{-\frac{\epsilon^2}{2\sigma^2_{t,m}}}$$

and

$$\int E(Z_{t+1,m}|Z_{t,m} = b(\epsilon), \ln I_t = \ell)\, f_{\epsilon_{t,m}}(\epsilon) d\,\epsilon$$

Because $\epsilon_{t,m}$ *is a continuous random variable, we integrate over all the values to find* $E(Z_{t+1,m}|\ln\theta_t = a, \ln I_t = \ell)$. *This approach would be similar in the case where investment is also a latent variable. In this case, we would integrate over the support of the measurement error terms of both variables.*

A.2 Proof of Theorem 1

Proof.

Given $G_t(\theta_t, I_t)$ and the measurement parameters for period t, $\mu_{t,m}$ and $\lambda_{t,m}$, are known, we use Lemma 1 to identify $E(\ln \theta_{t+1} | \ln \theta_t = a, \ln I_t = \ell)$ from $E(Z_{t+1,m} | \ln \theta_t = a, \ln I_t = \ell)$ for any $a \in \mathbb{R}$ and $\ell \in \mathbb{R}$. We then use the following transformation:

$$\frac{E(Z_{t+1,m}| \ln \theta = a_1, \ln I_t = \ell_1) - E(Z_{t+1,m}| \ln \theta = a_2, \ln I_t = \ell_2)}{E(Z_{t+1,m}| \ln \theta = a_3, \ln I_t = \ell_3) - E(Z_{t+1,m}| \ln \theta = a_2, \ln I_t = \ell_2)} = \frac{\ln f_t(e^{a_1}, e^{\ell_1}) - \ln f_t(e^{a_2}, e^{\ell_2})}{\ln f_t(e^{a_3}, e^{\ell_3}) - \ln f_t(e^{a_2}, e^{\ell_2})}$$

Because the function f_t satisfies the known location and scale definition, then for the points (a_2, ℓ_2) and (a_3, ℓ_3) the function evaluated at those points, $f_t(e^{a_2}, e^{\ell_2})$ and $f_t(e^{a_3}, e^{\ell_3})$, where $f_t(e^{a_2}, e^{\ell_2}) \neq f_t(e^{a_3}, e^{\ell_3})$, is known. Call these known points, $f_t(e^{a_2}, e^{\ell_2}) = \alpha_2$ and $f_t(e^{a_3}, e^{\ell_3}) = \alpha_3$.

$$\frac{E(Z_{t+1,m}| \ln \theta_t = a_1, \ln I_t = \ell_1) - E(Z_{t+1,m}| \ln \theta_t = a_2, \ln I_t = \ell_2)}{E(Z_{t+1,m}| \ln \theta_t = a_3, \ln I_t = \ell_3) - E(Z_{t+1,m}| \ln \theta_t = a_2, \ln I_t = \ell_2)} = \frac{\ln f_t(e^{a_1}, e^{\ell_1}) - \alpha_2}{\alpha_3 - \alpha_2}$$

We identify the function $\ln f_t(\theta_t, I_t)$ over its support by varying $a_1 \in \mathbb{R}$ and $\ell_1 \in \mathbb{R}$. We cannot of course use this transformation to identify the function at the point (a_2, ℓ_2), but the function evaluated at this point $f_t(e^{a_2}, e^{\ell_2})$ is already known by Definition 1. ■

A.3 Derivation of Example with CES Technology (Example 2)

$$\Delta_1 = \frac{\ln f_0(a_1, 0) - \ln f_0(1, 1)}{\ln f_0(e^1, e^1) - \ln f_0(1, 1)}$$

$$\Delta_1 = \frac{\ln(\gamma_0 \, a_1) - 0}{\ln(e^1) - 0}$$

$$\Delta_1 = \frac{\ln(\gamma_0 \, a_1)}{1}$$

$$e^{\Delta_1} = \gamma_0 \, a_1$$

$$\gamma_0 = \frac{e^{\Delta_1}}{a_1}$$

Once we have γ_0, we can use the same ratio as before taking $a_1 \neq \{0,1\}$, $a_3 \neq 0$, $\ell_1 = 1$, $a_2 = a_4 = \ell_2 = \ell_4$=1 and taking the limit $\ell_3 \to 0$ we have:

$$\Delta_2 = \frac{\ln f_0(a_1,1) - \ln f_0(1,1)}{\ln f_0(a_3,0) - \ln f_0(1,1)}$$

$$\Delta_2 = \frac{\ln f_0(a_1,1) - 0}{\ln f_0(a_3,0) - 0}$$

$$\Delta_2 = \frac{\ln f_0(a_1,1)}{\ln f_0(a_3,0)}$$

$$\Delta_2 = \frac{\ln(\gamma_0\, a_1^{\phi_0} + 1 - \gamma_0)}{\ln(\gamma_0\, a_3)}$$

$$\ln(\gamma_0\, a_3)\Delta_2 = \ln(\gamma_0\, a_1^{\phi_0} + 1 - \gamma_0)$$

$$(\gamma_0\, a_3)^{\Delta_2} = \gamma_0\, a_1^{\phi_0} + 1 - \gamma_0$$

$$(a_1)^{\phi_0} = \frac{(\gamma_0\, a_3)^{\Delta_2} - 1 + \gamma_0}{\gamma_0}$$

$$\phi_0 \ln(a_1) = \ln\left(\frac{(\gamma_0\, a_3)^{\Delta_2} - 1 + \gamma_0}{\gamma_0}\right)$$

$$\phi_0 = \frac{\ln\left(\frac{(\gamma_0\, a_3)^{\Delta_2} - 1 + \gamma_0}{\gamma_0}\right)}{\ln(a_1)}$$

A.4 Technologies and Output Elasticities

One rationale for the choice of a technology specification with non-constant returns to scale is the flexibility this specification offers with respect to the implied output elasticity. We consider the output elasticity with respect to investment defined as

$$\epsilon_I \equiv \frac{\partial \ln \theta_{t+1}}{\partial \ln I_t}$$

This elasticity is key to quantifying the effects of policy interventions.

In the general CES case, with technology given by

$$\theta_{t+1} = \left[\gamma_t \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}\right]^{\frac{\psi_t}{\phi_t}},$$

the output elasticity is given by

$$\epsilon_I = \frac{\psi_t}{\phi_t}\left[\gamma \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}\right]^{\frac{\psi_t}{\phi_t}-1} \phi(1-\gamma_t) I_t^{\phi_t - 1} \cdot \frac{I_t}{\left[\gamma_t \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}\right]^{\frac{\psi_t}{\phi_t}}}$$

$$= \frac{\psi_t (1-\gamma_t) I_t^{\phi_t}}{\gamma_t \theta_t^{\phi_t} + (1-\gamma_t) I_t^{\phi_t}} \in [0, \infty)$$

In the special case of constant returns to scale (CRS), $\psi_t = 1$, and $\epsilon_I \in (0,1)$. CRS implies this elasticity is bounded from above by 1. The general, non-constant returns to scale, case allows a larger than unit elastic response.

Similarly, the general translog technology,

$$\ln \theta_{t+1} = \alpha_{1t} \ln \theta_t + \alpha_{2t} \ln I_t + \alpha_{3t} \ln \theta_t \ln I_t$$

with elasticity

$$\epsilon_I = \alpha_{1t} + \alpha_{3t} \ln \theta_t$$

also allows general higher than unit elastic elasticities.

The main insight we want to underline is that the CES technology with constant return to scale restricts the output elasticity to be between 0 and 1: a one percent change in investment leads to a less than one percent change in next period skills. This prediction is independent of data, hence it can potentially be very restrictive in the context of child development and skills formation.

B Additional Tables and Figures

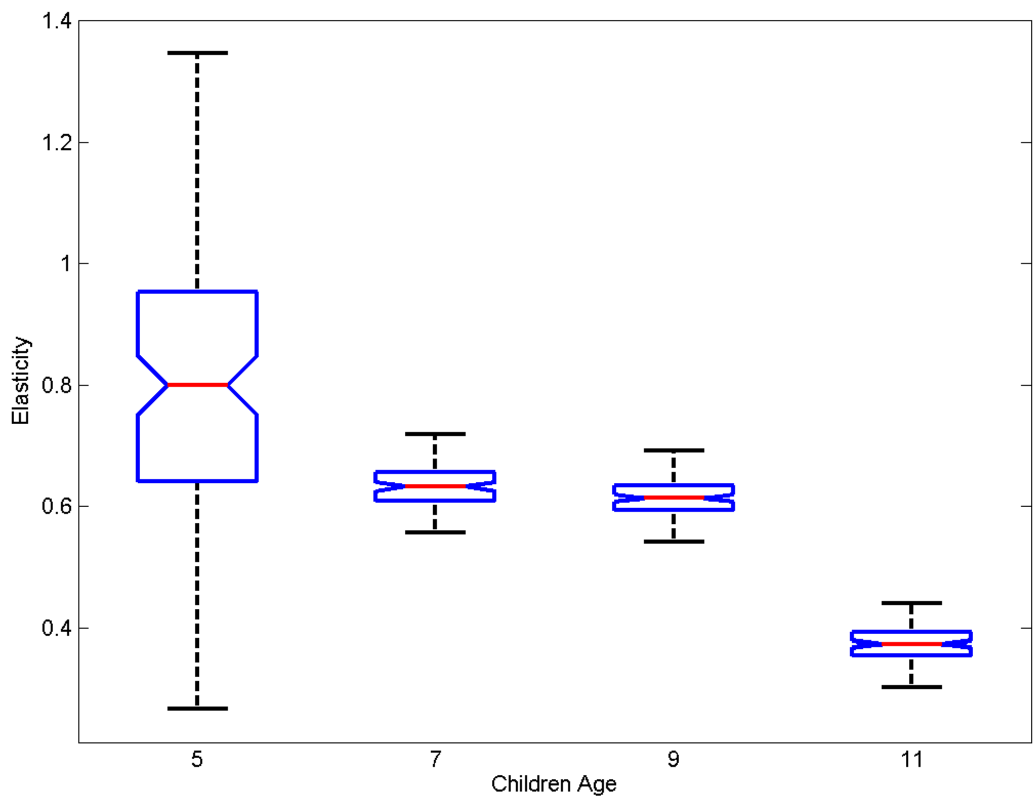
B.1 Additional Tables for Model 1 Corrected for Measurement Error

Table B.1-1: Estimates for Income Process

Constant	0.377
	(0.013)
Log Family Income t-1	0.753
	(0.008)
Variance Innovation	0.579
	(0.008)

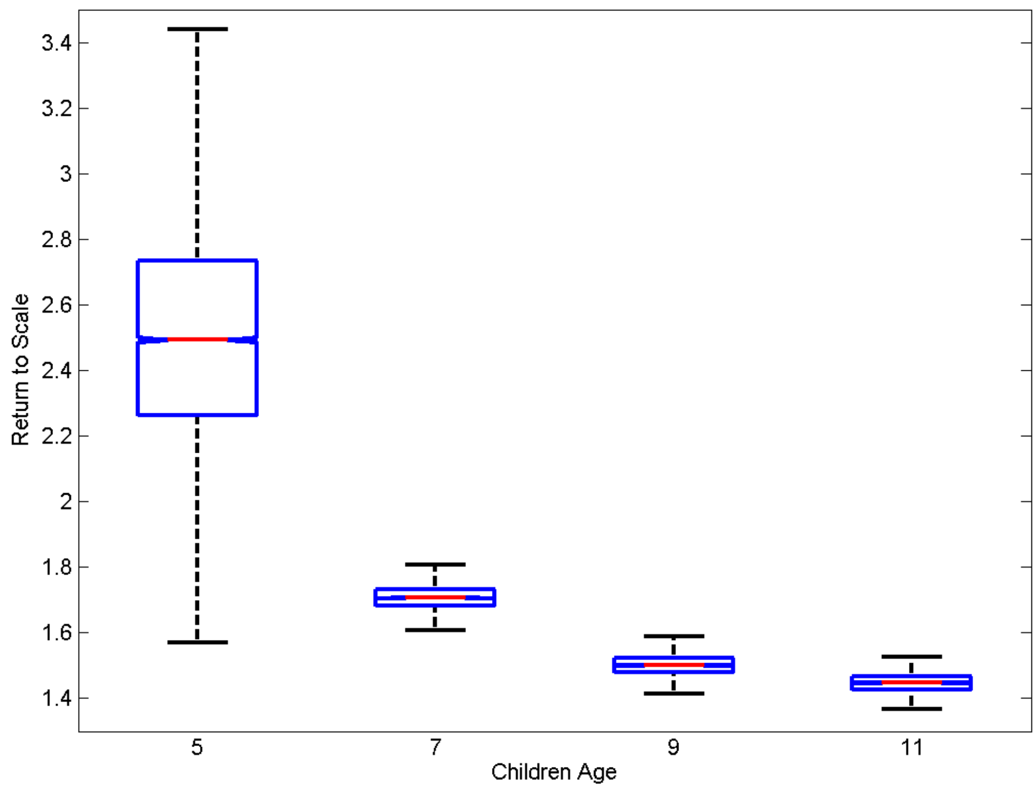
Notes: This table shows the estimates for the income process. The dependent variable is log family income at time t. Log Family Income $t-1$ is log family income two years prior. Standard errors in parenthesis are computed using a cluster bootstrap.

Figure B.1-1: Distribution of Elasticity of Next Period Skills with respect to Investment by Age



Notes: This figure shows the box plot for the elasticity of next period skills with respect to investment by different ages in the estimated Model 1 controlling for measurement error. The box plot is constructed as follow: the "central box" represents the central 50% of the data. Its lower and upper boundary lines are at the 25^{th} and 75^{th} quantile of the data. The central line indicates the median of the data while the two extreme lines (the top and the bottom ones) represents the 5^{th} and 95^{th} percentiles.

Figure B.1-2: Distribution of Technology Return to Scale by Age



Notes: This figure shows the box plot for the technology return to scale by different ages in the estimated Model 1 controlling for measurement error. The box plot is constructed as follow: the "central box" represents the central 50% of the data. Its lower and upper boundary lines are at the 25^{th} and 75^{th} quantile of the data. The central line indicates the median of the data while the two extreme lines (the top and the bottom ones) represents the 5^{th} and 95^{th} percentiles.

B.2 Descriptive Statistics

Table B.2-1: Children's Skills Measures

Measures	Range Values	Age Range	Scoring Order
(The Peabody Individual Achievement Test):			
Math	0-84	5-14	Positive
Recognition	0-84	5-14	Positive
Comprehensive	0-84	5-14	Positive

Notes: This table shows the features of children cognitive measures. The first column indicate each type of children skills measure we use to estimate our model. The second column shows the minimum and maximum value that each measure takes. The third column shows the minimum and maximum children age at which each measure is available. The last column indicates whether the measure is ordered positively (the higher score tend to reveal higher skills) or negatively (the lower score tend to reveal higher skills).

Table B.2-2: Mothers Cognitive Skills Measures

Measures	Range Values	Scoring Order
Arithmetics	0-30	Positive
Word Knowledge	0-35	Positive
Paragraph Composition	0-15	Positive
Numeric Operations	0-50	Positive
Coding Speed	0-84	Positive
Math Knowledge	0-25	Positive

Notes: This table shows the features of mother cognitive measures. The first column indicate each type of mother cognitive skills measure we use to estimate our model. The second column shows the minimum and maximum value that each measure takes. The last column indicates whether the measure is ordered positively (the higher score tend to reveal higher skills) or negatively (the lower score tend to reveal higher skills).

Table B.2-3: Mothers Noncognitive Skills Measures

Type of variables	Range Values	Label	Scoring Order
Mother Noncognitive Measures			
(Rosenberg indexes): I am a person of worth I have a number of good qualities I am able to do things as well as most other people I take a positive attitude toward myself	1-4	1= Strongly agree 2= Agree 3=Disagree 4=Strongly disagree	Negative
I am inclined to feel that I am a failure I felt I do not have much to be proud of I wish I could have more respect for myself I certainly feel useless at times At times I think I am no good at all	1-4	1= Strongly agree 2= Agree 3=Disagree 4=Strongly disagree	Positive
(Rotter Indexes): Rotter 1 (Life is in control or not)	1-4	1= In Control and closer to my opinion 2= In control but slightly closer to my opinion 3= Not in control but slightly closer to my opinion 4= Not in control and closer to my opinion	Negative
Rotter 2 (Plans work vs Matter of Luck)	1-4	1= Plans work and closer to my opinion 2= Plans work but slightly closer to my opinion 3= Matter of Luck but slightly closer to my opinion 4= Matter of Luck and closer to my opinion	Negative
Rotter 3 (Luck not a factor vs Flip a coin)	1-4	1= Luck not a factor and closer to my opinion 2=Luck not a factor but slightly closer to my opinion 3= Flip a coin but slightly closer to my opinion 4= Flip a coin and closer to my opinion	Negative
Rotter 4 (Luck big role vs Luck no role)	1-4	1= Luck big role and closer to my opinion 2=Luck big role but slightly closer to my opinion 3= Luck no role but slightly closer to my opinion 4= Luck no role and closer to my opinion	Positive

Notes: This table shows the features of mother noncognitive measures. The first column indicate each type of mother cognitive skills measure we use to estimate our model. The second column shows the minimum and maximum value that each measure takes. The third column shows the type of answers associated with each measure value. The last column indicates whether the measure is ordered positively (the higher score tend to reveal higher skills) or negatively (the lower score tend to reveal higher skills).

Table B.2-4: Descriptive Statistics about Children's Cognitive Skills Measures

Measures	Mean	Std	Min	Max	Number of Values
Age 5-6					
PIAT Math	11.858	4.278	0.000	37.000	32.000
PIAT Recognition	12.864	5.048	0.000	57.000	35.000
PIAT Comprehensive	12.770	4.930	0.000	49.000	35.000
Age 7-8					
PIAT Math	23.016	8.681	0.000	74.000	58.000
PIAT Recognition	25.748	8.774	0.000	80.000	67.000
PIAT Comprehensive	24.099	8.142	0.000	69.000	60.000
Age 9-10					
PIAT Math	38.720	10.832	0.000	84.000	71.000
PIAT Recognition	40.825	11.487	0.000	84.000	76.000
PIAT Comprehensive	37.540	10.231	0.000	78.000	64.000
Age 11-12					
PIAT Math	48.184	10.543	0.000	84.000	78.000
PIAT Recognition	51.079	13.278	0.000	84.000	74.000
PIAT Comprehensive	45.732	11.272	0.000	84.000	72.000
Age 13-14					
PIAT Math	53.767	11.387	0.000	84.000	78.000
PIAT Recognition	58.670	14.262	0.000	84.000	74.000
PIAT Comprehensive	51.015	12.229	0.000	84.000	74.000

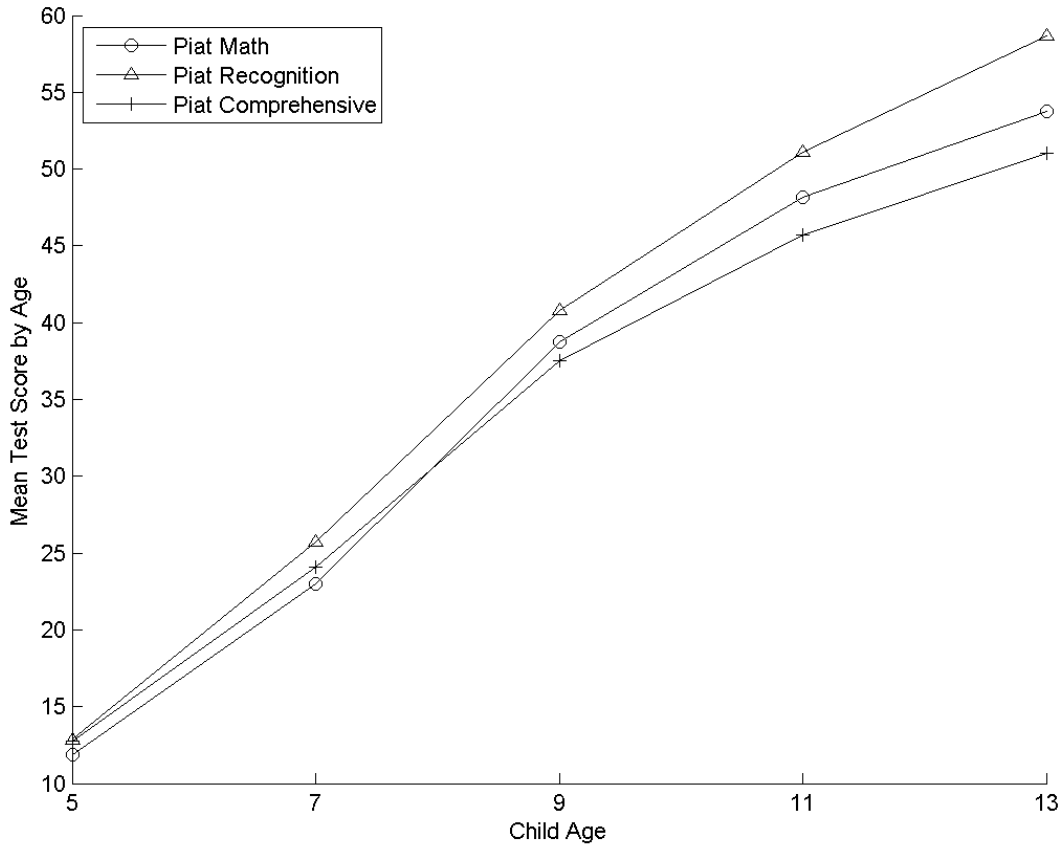
Notes: This table shows main sample statistics of children cognitive skills measures by children age.

Table B.2-5: Descriptive Statistics of Mother Cognitive and Noncognitive Skills Measures

Measures	Mean	Std	Min	Max	Number of Values
Mother Cognitive Skills					
Mom‘s Arithmetic Reasoning Test Score	13.946	6.603	0.000	30.000	31.000
Mom‘s Word Knowledge Test Score	21.773	8.562	0.000	35.000	36.000
Mom‘s Paragraph Composition Test Score	9.620	3.778	0.000	15.000	16.000
Mom‘s Numerical Operations Test Score	31.044	11.831	0.000	50.000	51.000
Mom‘s Coding Speed Test Score	42.953	17.468	0.000	84.000	85.000
Mom‘s Mathematical Knowledge Test Score	10.853	5.867	0.000	25.000	26.000
Mother Non Cognitive Skills					
Mom‘s Self-Esteem: ”I am a person of worth”	2.461	0.549	0.000	3.000	4.000
Mom‘s Self-Esteem: ” I have good qualities”	2.338	0.539	0.000	3.000	4.000
Mom‘s Self-Esteem: ”I am a failure”	3.379	0.618	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I am as capable as others”	2.291	0.567	0.000	3.000	4.000
Mom‘s Self-Esteem: ”I have nothing to be proud of”	3.360	0.669	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I have a positive attitude”	2.183	0.619	0.000	3.000	4.000
Mom‘s Self-Esteem: ”I wish I had more self-respect”	2.796	0.817	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I feel useless at times”	2.650	0.770	1.000	4.000	4.000
Mom‘s Self-Esteem: ”I sometimes think I am no good”	3.039	0.802	1.000	4.000	4.000
Mom‘s Rotter Score:”I have no control”	2.863	1.058	1.000	4.000	4.000
Mom‘s Rotter Score: ”I make no plans for the future”	2.386	1.192	1.000	4.000	4.000
Mom‘s Rotter Score: ”Luck is big factor in life”	3.205	0.856	1.000	4.000	4.000
Mom‘s Rotter Score: ”Luck plays big role in my life”	2.594	1.024	1.000	4.000	4.000

Notes: This table shows main sample statistics of mother cognitive skills measures.

Figure B.2-1: Descriptive Statistics: Mean of PIATs over the Childhood



Notes: This figure shows the mean Piat Math, Recognition and Comprehensive test scores by age. The x-axis shows children age. Child age of 5 is age 5-6, 7 is age 7-8, and so on.

B.3 Measurement Parameter Estimates

Table B.3-1: Measurement Parameter Estimates for Children's Cognitive Measures

Measures	μ	λ	Signal	Noise
Age 5-6				
PIAT Math	11.858	1.000	0.270	0.730
PIAT Recognition	12.864	2.238	0.972	0.028
PIAT Comprehensive	12.770	2.159	0.948	0.052
Age 7-8				
PIAT Math	11.858	1.000	0.757	0.243
PIAT Recognition	15.592	0.906	0.608	0.392
PIAT Comprehensive	15.014	0.802	0.554	0.446
Age 9-10				
PIAT Math	11.858	1.000	0.779	0.221
PIAT Recognition	10.297	1.136	0.894	0.106
PIAT Comprehensive	12.273	0.936	0.765	0.235
Age 11-12				
PIAT Math	11.858	1.000	0.803	0.197
PIAT Recognition	2.107	1.347	0.918	0.082
PIAT Comprehensive	6.129	1.089	0.833	0.167
Age 13-14				
PIAT Math	11.858	1.000	0.927	0.073
PIAT Recognition	8.556	1.195	0.845	0.155
PIAT Comprehensive	9.041	1.002	0.806	0.194

Notes: This table shows the measurement error parameters and associated statistics for children cognitive measures. The first two columns shows the measurement parameters (μ and λ) while the last two columns shows the signal and noise variance decomposition for the children cognitive measures.

Table B.3-2: Measurement Parameter Estimates for Mother Cognitive and Noncognitive Measures

Measures	μ	λ	Signal	Noise
Mother Cognitive Skills				
Mom‘s Arithmetic Reasoning Test Score	13.946	1.000	0.692	0.308
Mom‘s Word Knowledge Test Score	21.773	1.345	0.745	0.255
Mom‘s Paragraph Composition Test Score	9.620	0.584	0.722	0.278
Mom‘s Numerical Operations Test Score	31.044	1.720	0.638	0.362
Mom‘s Coding Speed Test Score	42.953	2.308	0.527	0.473
Mom‘s Mathematical Knowledge Test Score	10.853	0.854	0.639	0.361
Mother Non Cognitive Skills				
Mom‘s Self-Esteem: ”I am a person of worth”	2.461	1.000	0.152	0.848
Mom‘s Self-Esteem: ” I have good qualities”	2.338	1.263	0.252	0.748
Mom‘s Self-Esteem: ”I am a failure”	3.379	1.612	0.311	0.689
Mom‘s Self-Esteem: ”I am as capable as others”	2.291	1.127	0.181	0.819
Mom‘s Self-Esteem: ”I have nothing to be proud of”	3.360	1.746	0.312	0.688
Mom‘s Self-Esteem: ”I have a positive attitude”	2.183	1.474	0.260	0.740
Mom‘s Self-Esteem: ”I wish I had more self-respect”	2.796	2.080	0.297	0.703
Mom‘s Self-Esteem: ”I feel useless at times”	2.650	1.861	0.268	0.732
Mom‘s Self-Esteem: ”I sometimes think I am no good”	3.039	2.096	0.313	0.687
Mom‘s Rotter Score:”I have no control”	2.461	1.000	0.092	0.908
Mom‘s Rotter Score: ”I make no plans for the future”	2.338	1.263	0.140	0.860
Mom‘s Rotter Score: ”Luck is big factor in life”	3.379	1.612	0.118	0.882
Mom‘s Rotter Score: ”Luck plays big role in my life”	2.291	1.127	0.044	0.956

Notes: This table shows the measurement error parameters and associated statistics for mother cognitive and noncognitive measures. The first two columns shows the measurement parameters (μ and λ) while the last two columns shows the signal and noise variance decomposition for the mother measures.

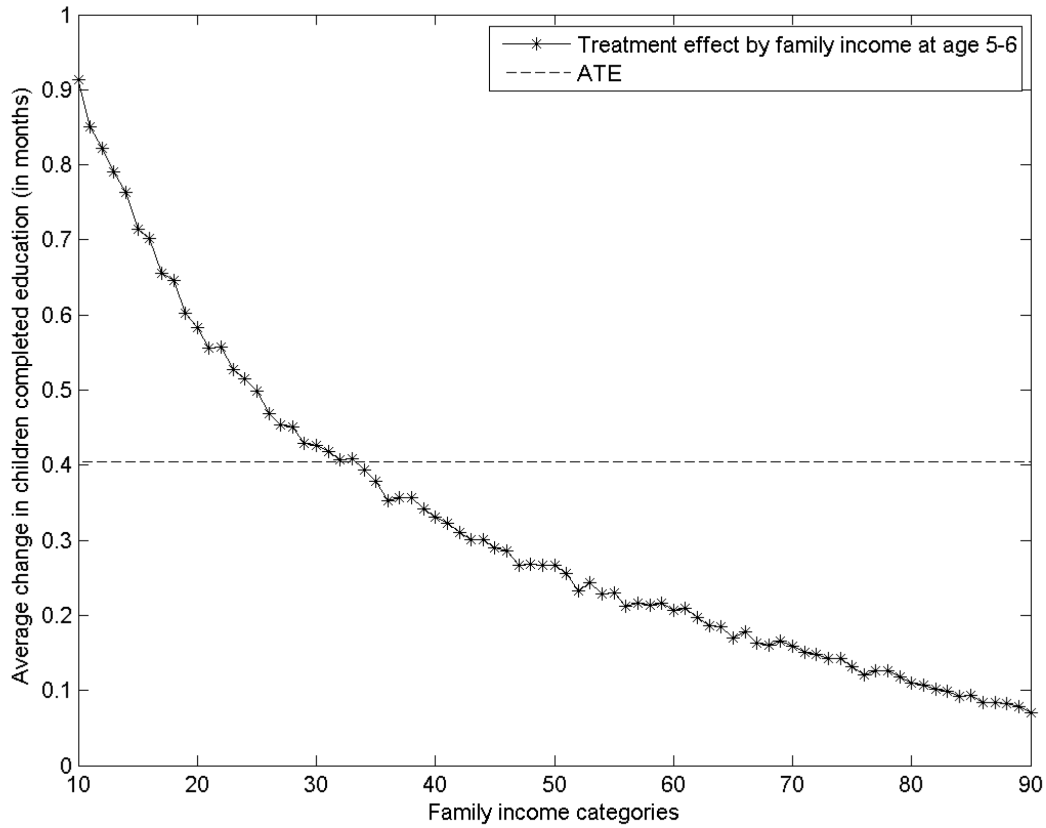
B.4 Estimates and Results for Model 2 with Measurement Error Corrected Estimator

Table B.4-1: Estimates for Investment (Model 2)

Parameter	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	0.230	0.069	0.068	0.065
	(0.059)	(0.021)	(0.029)	(0.030)
Log Mother Cognitive Skills	0.071	0.004	0.011	-0.005
	(0.022)	(0.009)	(0.014)	(0.012)
Log Mother Noncognitive Skills	0.359	0.711	0.660	0.678
	(0.131)	(0.059)	(0.084)	(0.084)
Log Family Income	0.341	0.217	0.261	0.262
	(0.076)	(0.054)	(0.072)	(0.082)
Variance Shocks	1.186	0.969	0.831	1.028
	(0.232)	(0.134)	(0.211)	(0.259)

Notes: This table shows the measurement error corrected estimates for the investment equation for Model 2. Each column shows the coefficients of the investment equation at the given ages. The dependent variable is investment in period t which is determined by the covariates at time t . For example, the first column shows the coefficients at age 5-6 for parental investments and child's skill and family income at age 5-6 as well. Standard errors in parenthesis are computed using a cluster bootstrap.

Figure B.4-1: Heterogeneity in Policy Effects by Age 5 Household Income (Outcome: Schooling at Age 23)



Notes: This figure plots the heterogeneous effect of a \$1,000 income transfer at age 5-6 on completed months of schooling by the percentile of initial (age 5-6) family income for the estimated Model 2, controlling for measurement error. Each income category is defined as the people contained between n^{th} and the $n-1^{th}$ of the percentiles of the income distribution. For example, Income category 10 in the graph means the people who belong between the 9^{th} and 10^{th} percentile of the income distribution. In the estimated income distribution for our sample, income categories 10, 50, and 90 contain families with about \$14,000, \$45,000, and \$145,000 of annual family income. This figure also plots the average effect over the income distribution.

B.5 Estimates and Results without Measurement Error Correction (Model 1 and Model 2)

Table B.5-1: Estimates for Investment (Model 1 and Model 2)

	Model 1 (Free Return to Scale Technology and TFP Dynamics)				Model 2 (Restricted Return to Scale Technology and No TFP Dynamics)			
Parameter	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	0.083	0.032	0.024	0.015	0.083	0.045	0.030	0.014
	(0.023)	(0.009)	(0.009)	(0.007)	(0.023)	(0.012)	(0.011)	(0.007)
Log Mother Cognitive Skills	0.082	0.010	0.010	-0.002	0.082	0.010	0.010	-0.002
	(0.019)	(0.011)	(0.014)	(0.011)	(0.019)	(0.011)	(0.014)	(0.011)
Log Mother Noncognitive Skills	0.248	0.454	0.442	0.553	0.248	0.448	0.440	0.553
	(0.093)	(0.073)	(0.098)	(0.074)	(0.093)	(0.073)	(0.098)	(0.074)
Log Family Income	0.587	0.504	0.524	0.434	0.587	0.498	0.521	0.435
	(0.074)	(0.070)	(0.095)	(0.077)	(0.074)	(0.069)	(0.095)	(0.078)
Variance Shocks	1.635	1.522	1.537	1.535	1.635	1.504	1.529	1.537
	(0.224)	(0.172)	(0.364)	(0.327)	(0.224)	(0.168)	(0.360)	(0.329)

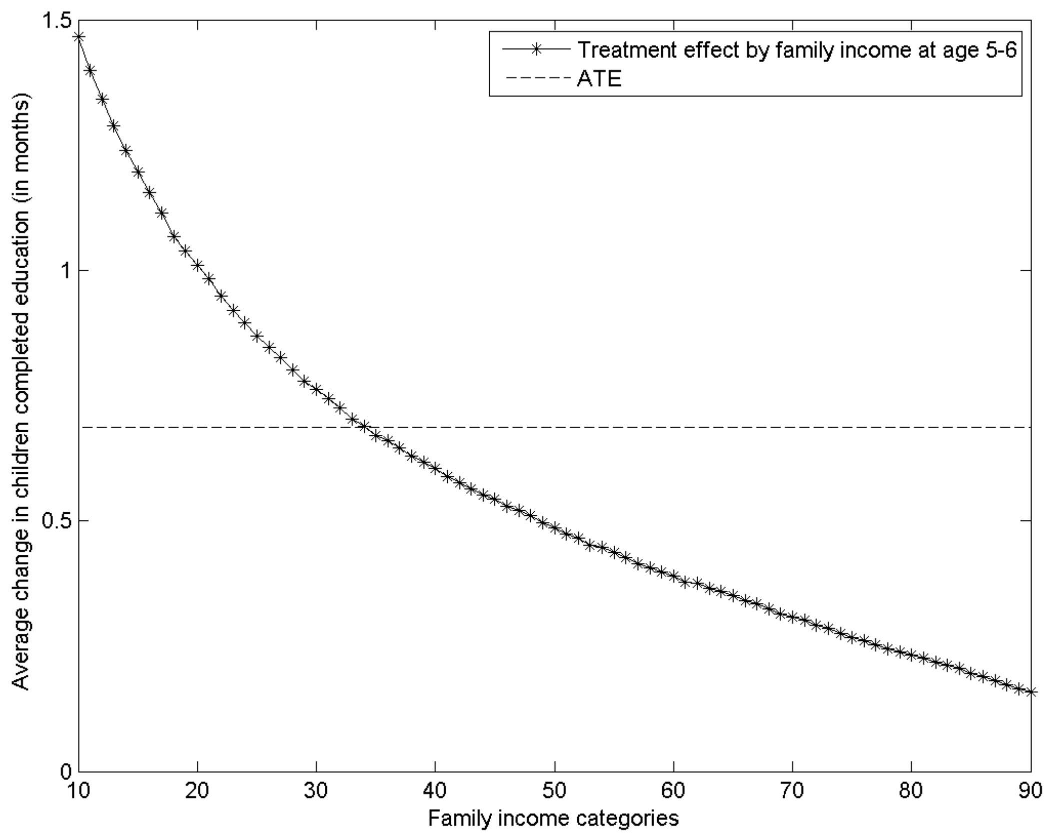
Notes: This table shows the estimates (not corrected for measurement error) for the investment equation for both Model 1 and Model 2. Each column shows the coefficients of the investment equation at the given ages. The dependent variable is investment in period t which is determined by the covariates at time t . For example, the first column shows the coefficients at age 5-6 for parental investments and child's skill and family income at age 5-6 as well. Standard errors in parenthesis are computed using a cluster bootstrap.

Table B.5-2: Estimates for Skill Technology (Model 1 and Model 2)

	Model 1 (Free Return to Scale Technology and TFP Dynamics)				Model 2 (Restricted Return to Scale Technology and No TFP Dynamics)			
Parameter	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	0.875	0.771	0.669	0.770	0.625	0.868	0.897	0.880
	(0.057)	(0.022)	(0.017)	(0.018)	(0.047)	(0.039)	(0.039)	(0.052)
Log Investment	0.518	0.069	0.042	0.325	0.370	0.125	0.101	0.127
	(0.089)	(0.066)	(0.061)	(0.099)	(0.045)	(0.038)	(0.039)	(0.052)
(Log Skills *	0.006	0.007	0.002	-0.006	0.005	0.008	0.002	-0.007
Log Investment)	(0.012)	(0.003)	(0.002)	(0.002)	(0.009)	(0.004)	(0.002)	(0.003)
Return to scale	1.399	0.846	0.713	1.089	1.000	1.000	1.000	1.000
	(0.098)	(0.072)	(0.063)	(0.096)	(-)	(-)	(-)	(-)
Variance shocks	7.490	7.673	6.716	7.382	5.354	6.155	7.211	9.092
	(0.127)	(0.145)	(0.192)	(0.220)	(0.386)	(0.565)	(0.769)	(0.980)
Log TFP	12.789	18.491	18.477	14.011	0.000	0.000	0.000	0.000
	(0.215)	(0.299)	(0.444)	(0.690)	(-)	(-)	(-)	(-)

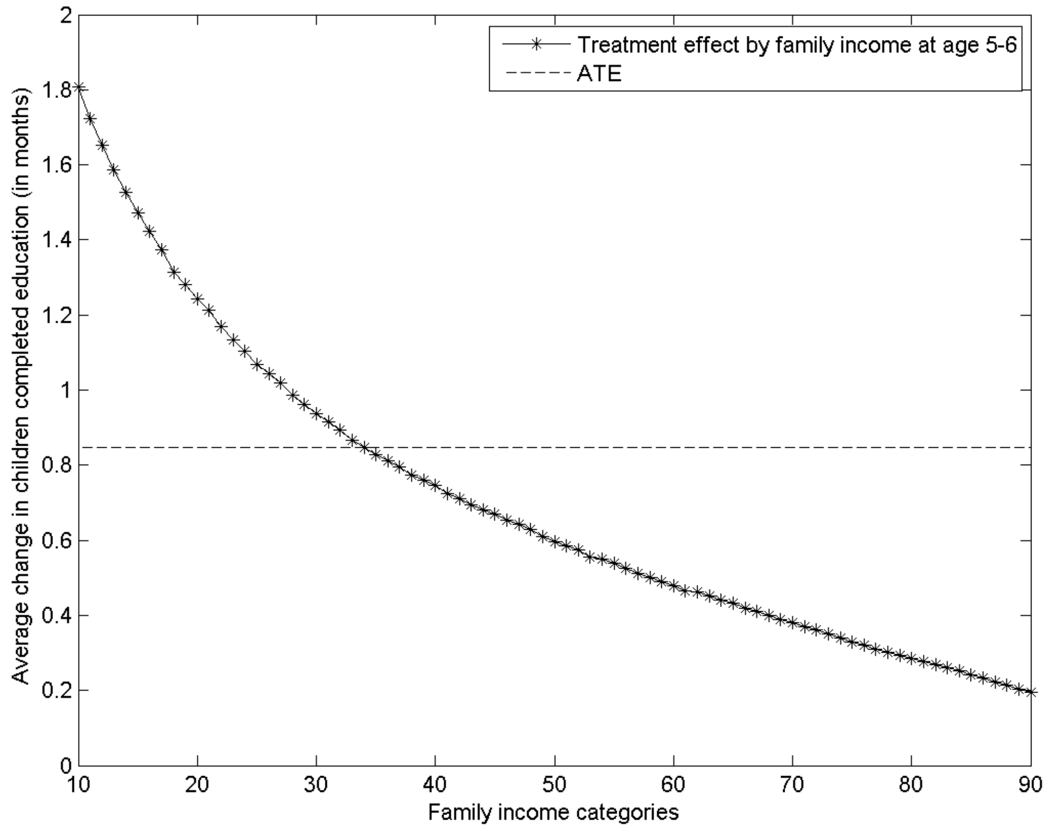
Notes: This table shows the estimates (not corrected for measurement error) for the technology of skills formation and the technology return to scale (i.e. the sum of the share parameters for each input) for not measurement error corrected estimates of both Model 1 and Model 2. Each column shows the coefficients of the technology of skills formations at the given age. The dependent variable is log skills in the next period t+1 while the covariates (inputs) are at time t. For example, the first column shows the coefficients for the skills inputs at age 5-6 which lead to log skills at age 7-8. Standard errors in parenthesis are computed using a cluster bootstrap.

Figure B.5-1: Heterogeneity in Policy Effects by Age 5 Household Income (Outcome: Schooling at Age 23, Model 1)



Notes: This figure plots the heterogeneous effect of a $1,000 income transfer at age 5-6 on completed months of schooling by the percentile of initial (age 5-6) family income for the estimated Model 1, not controlling for measurement error. Each income category is defined as the people contained between n^{th} and the $n-1^{th}$ of the percentiles of the income distribution. For example, Income category 10 in the graph means the people who belong between the 9^{th} and 10^{th} percentile of the income distribution. In the estimated income distribution for our sample, income categories 10, 50, and 90 contain families with about $14,000, $45,000, and $145,000 of annual family income. This figure also plots the average effect over the income distribution.

Figure B.5-2: Heterogeneity in Policy Effects by Age 5 Household Income (Outcome: Schooling at Age 23, Model 2)



Notes: This figure plots the heterogeneous effect of a $1,000 income transfer at age 5-6 on completed months of schooling by the percentile of initial (age 5-6) family income for the estimated Model 2, not controlling for measurement error. Each income category is defined as the people contained between n^{th} and the $n-1^{th}$ of the percentiles of the income distribution. For example, Income category 10 in the graph means the people who belong between the 9^{th} and 10^{th} percentile of the income distribution. In the estimated income distribution for our sample, income categories 10, 50, and 90 contain families with about $14,000, $45,000, and $145,000 of annual family income. This figure also plots the average effect over the income distribution.

B.6 Skills measures in CNLSY79

Measures for Cognitive Skills

- **Peabody Picture Vocabulary Test**

  The Peabody Picture Vocabulary Test, revised edition (PPVT-R) "measures an individual's receptive (hearing) vocabulary for Standard American English and provides, at the same time, a quick estimate of verbal ability or scholastic aptitude" (see Dunn and Dunn, 1981). The PPVT was designed for use with individuals aged 2 to 40 years. The English language version of the PPVT-R consists of 175 vocabulary items of generally increasing difficulty. The child listens to a word uttered by the interviewer and then selects one of four pictures that best describes the word's meaning. The PPVT-R has been administered, with some exceptions, to NLSY79 children between the ages of 3-18 years of age until 1994, when children 15 and older moved into the Young Adult survey. In the current survey round, the PPVT was administered to children aged 4-5 and 10-11 years of age, as well as to some children with no previous valid PPVT score.

  The first item, or starting point, is determined based on the child's PPVT age. Starting at an age-specific level of difficulty is intended to reduce the number of items that are too easy or too difficult, in order to minimize boredom or frustration. The suggested starting points for each age can be found in the PPVT manual (see Dunn and Dunn, 1981).

  Testing begins with the starting point and proceeds forward until the child makes an incorrect response. If the child has made 8 or more correct responses before the first error, a "basal" is established. The basal is defined as the last item in the highest series of 8 consecutive correct answers. Once the basal is established, testing proceeds forwards, until the child makes six errors in eight consecutive items. If, however, the child gives an incorrect response before 8 consecutive correct answers have been made, testing proceeds backwards, beginning at the item just before the starting point, until 8 consecutive correct responses have been made. If a child does not make eight consecutive responses even after administering all of the items, he or she is given a basal of one. If a child has more than one series of 8 consecutive correct answers, the highest basal is used to compute the raw score.

  A "ceiling" is established when a child incorrectly identifies six of eight consecutive items. The ceiling is defined as the last item in the lowest series of eight consecutive items with six incorrect responses. If more than one ceiling is identified, the lowest ceiling is used to compute the raw score. The assessment is complete once both a basal and a ceiling have been established. The ceiling is set to 175 if the child never makes six errors in eight consecutive items.

A child's raw score is the number of correct answers below the ceiling. Note that all answers below the highest basal are counted as correct, even if the child answered some of these items incorrectly. The raw score can be calculated by subtracting the number of errors between the highest basal and lowest ceiling from the item number of the lowest ceiling.

- **The Peabody Individual Achievement Test (PIAT): Math**

  The PIAT Mathematics assessment protocol used in the field is described in the documentation for the Child Supplement (available on the Questionnaires page). This subscale measures a child's attainment in mathematics as taught in mainstream education. It consists of 84 multiple-choice items of increasing difficulty. It begins with such early skills as recognizing numerals and progresses to measuring advanced concepts in geometry and trigonometry. The child looks at each problem on an easel page and then chooses an answer by pointing to or naming one of four answer options.

  Administration of this assessment is relatively straightforward. Children enter the assessment at an age-appropriate item (although this is not essential to the scoring) and establish a "basal" by attaining five consecutive correct responses. If no basal is achieved then a basal of "1" is assigned (see PPVT). A "ceiling" is reached when five of seven items are answered incorrectly. The non-normalized raw score is equivalent to the ceiling item minus the number of incorrect responses between the basal and the ceiling scores.

- **The Peabody Individual Achievement Test (PIAT): Reading Recognition**

  The Peabody Individual Achievement Test (PIAT) Reading Recognition subtest, one of five in the PIAT series, measures word recognition and pronunciation ability, essential components of reading achievement. Children read a word silently, then say it aloud. PIAT Reading Recognition contains 84 items, each with four options, which increase in difficulty from preschool to high school levels. Skills assessed include matching letters, naming names, and reading single words aloud.

  The only difference in the implementation procedures between the PIAT Mathematics and PIAT Reading Recognition assessments is that the entry point into the Reading Recognition assessment is based on the child's score in the Mathematics assessment, although entering at the correct point is not essential to the scoring.

The scoring decisions and procedures are identical to those described for the PIAT Mathematics assessment.

- **The Peabody Individual Achievement Test (PIAT): Reading Comprehension**

  The Peabody Individual Achievement Test (PIAT) Reading Comprehension subtest measures a child's ability to derive meaning from sentences that are read silently. For each of 66 items of increasing difficulty, the child silently reads a sentence once and then selects one of four pictures that best portrays the meaning of the sentence.

  Children who score less than 19 on Reading Recognition are assigned their Reading Recognition score as their Reading Comprehension score. If they score at least 19 on the Reading Recognition assessment, their Reading Recognition score determines the entry point to Reading Comprehension. Entering at the correct location is, however, not essential to the scoring.

  Basals and ceilings on PIAT Reading Comprehension and an overall nonnormed raw score are determined in a manner identical to the other PIAT procedures. The only difference is that children for whom a basal could not be computed (but who otherwise completed the comprehension assessment) are automatically assigned a basal of 19. Administration instructions can be found in the assessment section of the Child Supplement.

C Alternative Measures

One of the characteristics of the data used to study child development is the rich variety skill measures. The previous sections considered identification where the skill measures are in a "raw" form: each measure is a linear function of the latent log skill. This measurement system, while commonly assumed in the prior literature, is in some respects a "best case."

In this section, we consider alternative forms of measures and re-examine whether we can identify the same types of production technologies using these alternative measures. We consider four classes of measures which are frequently encountered empirically: (i) *age-standardized* measures where the raw measures are transformed ex post to have mean 0 and standard deviation 1 for the sample at hand; (ii) *relative* measures where the measures reflect not the level of a child's skill but the child's skill relative to the population mean; (iii) *ordinal* measures which provide a discrete ranking of children's skills; and iv) *censored* measures where the measures are truncated with a "floor" (finite minimum value) and/or a "ceiling" (finite maximum value). For each type of measure, we discuss which of our prior identification results still hold, if any, and what auxiliary assumptions would be sufficient to restore our identification results.

C.1 Age-Standardized Measures

Age-standardized measures are defined as the following transformation of raw measures $Z_{t,m}$:

$$Z^S_{t,m} = \frac{Z_{t,m} - E(Z_{t,m})}{V(Z_{t,m})^{1/2}}. \tag{C-1}$$

By construction, these measures are mean 0 and standard deviation 1 for all child ages.

Our main identification result using standardized measures (Theorem 1) continues to hold if the technology of skill formation has known scale and location functions (KLS, Definition 1). To show this, we can re-write the standardized measures as a linear function of the latent variable:

$$Z^S_{t,m} = \mu^S_{t,m} + \lambda^S_{t,m} \ln \theta_t + \epsilon^S_{t,m}$$

where the measurement parameters and measurement error are

$$\mu^S_{t,m} = -\lambda^S_{t,m}(V(\ln \theta_t)) \cdot E(\ln \theta_t)$$

$$\lambda^S_{t,m} = \frac{\lambda_{t,m}}{V(Z_{t,m})^{1/2}} = \frac{\lambda_{t,m}}{(\lambda^2_{t,m}V(\ln\theta_t) + V(\epsilon_t))^{1/2}}$$

$$\epsilon^S_{t,m} = \frac{\epsilon_{t,m}}{V(Z_{t,m})^{1/2}} = \frac{\epsilon_{t,m}}{(\lambda^2_{t,m}V(\ln\theta_t) + V(\epsilon_t))^{1/2}}$$

These expressions show that the standardized measurement parameters are linear functions of the underlying moments of the latent skill distribution.[1] The reason for the invariance of our identification result to the use of standardized or raw measures is that any measurement parameters are "transformed away" as shown in Lemma 1. More generally, identification of the KLS production technologies is invariant to any increasing linear transformation of the original raw measures, say $Z'_{t,m} = a + bZ_{t,m}$ for $a \in R$ and $b \in R^+$.[2]

However, the use of age-standardized measures may not be cost free in the sense that age-standardized measures, which are constructed to be age-stationary in their first and second moments, contain no information about skill dynamics in these moments. For example, standardizing *age-invariant* measures, as defined in the previous section, so that the mean and variance of these measures is equal at all ages, would essentially "throw away" information regarding the average skill development of children across ages. This loss of information prevents the identification of the broader classes of technology of skills formation discussed above, the unknown Total Factor Productivity (TFP) functions (as in equation 14) or unknown scale functions (as in equation 15).

To see this point, recall that the identification of TFP or scaling parameter are based on additional information of the dynamics of measurement parameters. In the case of raw measures, those parameters are fully free parameters. On the other hand, when we use standardized measures, the new measurement parameters ($\mu^S_{t,m}$ and $\lambda^S_{t,m}$) are no longer free parameters but functions of the moments of the latent distribution. Hence, restricting the dynamics of the measurement parameters in this case (imposing Assumption 2 and Assumption 3) is equivalent to restricting the dynamics of the latent skills, and can restrict the possible classes of technologies. While age-standardizing measures may provide some descriptive value, in the

[1] It is important to recognize that the use of standardized measures does not necessarily imply that any particular restriction on the underlying latent variables such as $E(\ln\theta_t) = 0$ or $V(\ln\theta_t) = 1$. The standardizations are necessarily in terms of the observed measures, not the unobserved latent variables.

[2] One caveat deserves mention. Recall that because the initial conditions are normalized to a particular measure, using standardized rather than raw measures can affect the normalized location and scale of the latent skills, and in general affect the values of the production parameters which are identified up to the normalized initial period measure.

context of identifying dynamic production technologies, there is simply no point to transforming the measures in this way and throwing away potentially important identifying information.

C.2 Relative Measures

Some of the proxies used to measure children outcomes come from surveys where observers (often mothers, fathers, or other caregivers) provide assessments of the child. It can be plausible then that these observers are actually evaluating the child with respect to their perceptions of the average in the population. We call this type of measure a *relative measures*. In this case, these measures can be written as:

$$Z^R_{t,m} = \mu^R_{t,m} + \lambda^R_{t,m}(\ln\theta_t - E(\ln\theta_t)) + \epsilon^R_{t,m}. \tag{C-2}$$

where $(\ln\theta_t - E(\ln\theta_t)$ is the latent variable being measured by $Z^R_{t,m}$, which we model as the deviation of the actual level of the child's skill $\ln\theta_t$ relative to the mean value in the population $E(\ln\theta_t)$. Relative measures are not ordinal ranking measures (which we discuss below) but a continuous measure of skills relative to the population mean. As with the age-standardized measures, the relative measures are an increasing linear function of the underlying latent variable, and therefore the main identification result in Theorem 1 continues to hold as the measurement parameters are "transformed away."

C.3 Ordinal Measures

We define ordinal measures the measures which are based on children rankings: this child has higher skills than another child. Let's assume that we observe in data children's skill rank. Let $Z_t = \{1, 2, \ldots, J\}$ be the child's human capital rank, with 1 highest level, and J lowest level. The observer (or us forming ranks from test scores) forms rank according to this ordinal model:

$$Z^O_{t,m} = \begin{cases} J \text{ if } \lambda^O_{t,m}\ln\theta_t + \epsilon^O_{t,m} < \kappa_{J,t,m} \\ J-1 \text{ if } \kappa_{J,t,m} < \lambda^O_{t,m}\ln\theta_t + \epsilon^O_{t,m} < \kappa_{J-1,t,m} \\ \vdots \\ 2 \text{ if } \kappa_{3,t,m} < \lambda^O_{t,m}\ln\theta_t + \epsilon^O_{t,m} < \kappa_{2,t,m} \\ 1 \text{ if } \lambda^O_{t,m}\ln\theta_t + \epsilon^O_{t,m} > \kappa_{2,t,m} \end{cases} \tag{C-3}$$

where the $\kappa_2, \ldots, \kappa_J$, with $\kappa_2 > \kappa_3, \ldots, \kappa_J$, are measurement parameters which provide the mapping from latent skills $\ln\theta_t$ and measurement error $\epsilon^O_{t,m}$ to the observed ordinal ranking values $Z^O_{t,m}$. The probability a child is ranked first $(j = 1)$ is then

$$pr(Z^O_{t,m} = 1) = pr(\lambda^O_{t,m} \ln \theta_t + \epsilon^O_{t,m} > \kappa_{2,t,m})$$
$$= F_\epsilon(\lambda^O_{t,m} \ln \theta_t - \kappa_{2,t,m})$$

where F_ϵ is the distribution function for the measurement error $\epsilon^O_{t,m}$.

With ordinal ranking measures the non-parametric identification result no longer holds. There is no longer a one-to-one mapping between a child's latent skills θ_t and expected measures, as multiple values of θ_t are consistent with a child having a certain rank. Without additional assumptions beyond Assumption 1 (independence of measures), ordinal measures of skills do not allow non-parametric identification of the continuous skill production function.

If the researcher were to assume a particular known distribution for the measurement errors F_ϵ, then under this assumption for an ordinal measure of $t+1$ skills we would have:

$$F^{-1}_\epsilon(pr(Z^O_{t+1,m} = 1 | \ln \theta_t, \ln I_t)) = \lambda_{t+1,m} f_t(I_t, \theta_t) - \kappa_{2,t+1,m}$$

where $pr(Z^O_{t+1,m} = 1 | \ln \theta_t, \ln I_t)$ is the probability the child receives rank 1 at age $t+1$ given inputs θ_t, I_t at age t. This expression shows that with a known distribution for measurement errors, we can then apply Theorem 1 to identify a KLS technology $f_t(I_t, \theta_t)$ up to this assumed distribution.

C.4 Censored Measures

Censored measures are defined as

$$Z^C_{t,m} = \begin{cases} \overline{Z} \text{ if } Z_{t,m} \geq \overline{Z} \\ Z_{t,m} \text{ if } \underline{Z} < Z_{t,m} < \overline{Z} \\ \underline{Z} \text{ if } Z_{t,m} < \underline{Z} \end{cases} \tag{C-4}$$

where $Z_{t,m} = \mu_{t,m} + \lambda_{t,m} \ln \theta_t + \epsilon_{t,m}$ is the "latent" measure, and $\overline{Z}$ ("ceiling") and $\underline{Z}$ ("floor"), with $\overline{Z} > \underline{Z}$, are the truncation points. Censoring occurs, for example, when a test score used as the measure has a maximum score (answering all questions correctly) and a minimum score (say answering none of the questions correctly). In practice, researchers can ascertain whether censoring is an important issue empirically by investigating what proportion of the sample actually has measured skills at the floor or ceiling points of the measure. Because censored measures do not have full support, the non-parametric identification result of Theorem 1 appears no longer to hold. As with the ordinal measures, auxiliary assumptions could be used to achieve identification up these additional assumptions (for a complete analyze of the problem, see Wang et al. 2009, Koedel and Betts, 2010)

D Monte Carlo Exercise for Model 1 and Measurement Error Correction

We implement a Monte Carlo exercise to examine the properties of our estimator. The true data generating process is assumed to be the estimated (measurement error corrected) Model 1 with some additional parametric assumptions about the measurement error process. In order to simulate the dataset, we use the both the estimated measurement parameters and the joint distribution of children skills and investments. In addition, we assume that all the measurement noises are Normally distributed.[3] We generate a simulated longitudinal dataset of 10,000 children ranging from age 5-6 to age 13-14. In particular, the Monte Carlo analysis is performed estimating the model on 200 simulated data sets. In the following tables we show the mean estimates over the 200 estimates of the coefficients.

We focus only on estimates of skills technology, investment process and children's skills measurement parameters. Tables D-1-D-3 show true and mean estimated parameters. Overall, the estimator is able to recover the true parameters with minimal bias.

[3] We assume that the standard deviation of the error terms for all the skills measures are 0.5 (children and mothers) while we fix to 0.1 the standard deviation of the error terms for all the investment measures.

Table D-1: Monte Carlo Estimates for Investment Process

Parameter	True Parameters Age 5-6	Age 7-8	Age 9-10	Age 11-12	Monte Carlo Estimates Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	0.230	0.027	0.020	0.018	0.249	0.026	0.020	0.018
Log Mother Cognitive Skills	0.071	0.004	0.012	-0.005	0.077	0.002	0.008	-0.011
Log Mother Noncognitive Skills	0.359	0.742	0.694	0.712	0.322	0.748	0.700	0.700
Log Family Income	0.341	0.227	0.274	0.275	0.352	0.224	0.272	0.292
Variance Shocks	1.186	1.019	0.868	1.087	1.263	0.993	0.827	1.103

Notes: This table shows the both the true estimates (reported also in Table 3) and the mean Monte Carlo estimates for the investment equation. Each column shows the coefficients of the investment equation at the given ages. The dependent variable is investment in period t which is determined by the covariates at time t . For example, the first column shows the coefficients at age 5-6 for parental investments and child's skill and family income at age 5-6 as well.

Table D-2: Monte Carlo Estimates for Skill Technology

	True Parameters				Monte Carlo Estimates			
Parameter	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 5-6	Age 7-8	Age 9-10	Age 11-12
Log Skills	1.966	1.086	0.897	1.065	1.955	1.091	0.897	1.071
Log Investment	0.799	0.695	0.713	0.252	0.759	0.700	0.839	0.502
(Log Skills * Log Investment)	-0.105	-0.005	-0.003	0.003	-0.092	-0.005	-0.005	-0.002
Return to scale	2.660	1.776	1.606	1.320	2.623	1.786	1.731	1.571
Variance shocks	5.612	4.519	3.585	4.019	5.613	4.520	3.586	4.018
Log TFP	13.067	14.747	11.881	2.927	13.060	14.689	11.801	2.594

Notes: This table shows the both the true estimates (reported also in Table 4) and the mean Monte Carlo estimates for the technology of skills formation. Each column shows the coefficients of the technology of skills formations at the given age. The dependent variable is log skills in the next period $t+1$ while the covariates (inputs) are at time t. For example, the first column shows the coefficients for the skills inputs at age 5-6 which lead to log skills at age 7-8.

Table D-3: Monte Carlo Estimates for Measurement Parameters

	True Constant (μ)					Monte Carlo Constant (μ) Estimates				
Parameter	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14
PIAT Math	11.858	11.858	11.858	11.858	11.858	11.858	11.858	11.858	11.858	11.858
PIAT Recognition	12.864	15.592	10.297	2.107	8.556	12.864	15.592	10.298	2.110	8.555
PIAT Comprehensive	12.770	15.014	12.273	6.129	9.041	12.770	15.013	12.270	6.132	9.040

	True Factor Loadings (λ)					Monte Carlo Factor Loadings (λ) Estimates				
Parameter	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14	Age 5-6	Age 7-8	Age 9-10	Age 11-12	Age 13-14
PIAT Math	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
PIAT Recognition	2.238	0.906	1.136	1.347	1.195	2.238	0.905	1.136	1.347	1.196
PIAT Comprehensive	2.159	0.802	0.936	1.089	1.002	2.159	0.802	0.936	1.089	1.002

Notes: This table shows the both the true estimates (reported also in Table B.3-1) and the mean Monte Carlo estimates for the measurement parameters of children skills measures equation. Each column shows the parameters at the given ages for each test score.