The Option Value of Educational Choices
And the Rate of Return to Educational Choices

James Heckman
University of Chicago

Sergio Urzúa
Northwestern University

University College Dublin

Institute for Computational Economics Lecture
University of Chicago
August 1, 2008
Introduction

- Conventional models of rates of return to schooling follow Becker (1964).
- Assume perfect certainty.
- Compare two earnings streams associated with schooling $s$ and $s'$, $s < s'$:

  \[
  Y(s, t) = \text{earnings at schooling } s \text{ at age } t \\
  Y(s', t) = \text{earnings at schooling } s' \text{ at age } t
  \]

- Usually years of schooling are assumed to be ordered.
Introduction

- Rate of return is computed from pairwise comparisons of earnings streams.
- Define costs of going from \( s \rightarrow s' \) at time \( t \) as \( C(s', s, t) \),

\[
0 = \sum_{t=0}^{T} \frac{Y(s', t) - Y(s, t) - C(s', s, t)}{(1 + \rho(s', s))^{t}}
\]
- Assumes pairwise earnings profiles cross, but only once.
- \( \rho(s', s) \) is the pairwise rate of return.
The “Mincer” model of earnings approximates $\rho(s, s')$ under special conditions which are tested and rejected in U.S. data (Heckman, Lochner and Todd, 2006, 2008).

Main reasons for rejection:

(i) additive separability in $\log Y(s, t)$ between work experience and schooling is violated,

(ii) the costs of schooling are more than earnings foregone and earnings in school do not cover tuition costs,

(iii) Cunha, Heckman and Navarro (2005) and Cunha and Heckman (2007,2008) document that huge “psychic costs” are required to rationalize schooling choices in an expected income maximizing model.
As noted by Weisbrod (1962), pairwise comparisons between earnings streams associated with \( s \) and \( s' \) miss an important component of the return to the transition \( s \) to \( s' \).

Getting to \( s' \) means you have the option to go on to \( s'' > s' \).

The **option value** as defined by Weisbrod is the return that arises from not having to stop at \( s' \) (to go onto higher levels of education).

There is some confusion because Weisbrod’s “option value” is not the value of an (American option).

Option values as defined by Weisbrod can arise even in an environment of prefect certainty.

True rates of return are underestimated by internal rates of return.
Heckman, Lochner and Todd (2006) show that the internal rate of return, *ex ante* is not in general the proper rate of return criterion in a multiperiod ($T \geq 3$) model with uncertainty and more than two schooling choices even if earnings profiles cross only once in terms of age.

Discounted alternative earnings streams associated with value function branches can cross multiply even if, for pairwise schooling levels, they cross only once.
Need a more general rate of return concept based on value functions.

The rate of return cannot be defined independently of the interest rate as in the simple Becker model.

$\text{IRR} \geq r$ does not answer the question whether there is under-investment or over-investment in schooling.
Summary of Current State of the Empirical Literature on Returns to Education

- Preoccupation of most of the empirical literature with internal rates of return — IRR — in an environment of perfect certainty.

- IRR *a la* Becker misses learning and option values arising from learning and nonlinearity in payoffs of education in years of schooling.

- Mincer model seeks to approximate IRR and in general fails to identify even IRR.

- Need a model to capture these features plus psychic costs of schooling.
Need a model that recognizes that many educational choices are not simply summarized by “years of schooling” as in the Mincer model.

Thus \{s\} is not necessarily ordered: job training, etc.

In addition, people can drop in and drop out of school.

There are multiple decisions:

(a) **Whether** to move to a feasible schooling state
(b) **When** to move

Both create options and we can define option values for each.
Theoretical Contributions of this Paper

- A dynamic sequential model of educational choices among discrete states with option values arising from learning and nonlinearity of reward functions at different stages of the life cycle.
- We build a model of schooling connecting high school dropping out, GED attainment, delay, college choices and returns.
- Define the correct concept of the rate of return to schooling in a dynamic model with uncertainty, nonlinearity and delay.
- Like Arcidiacono (2004), we model learning about persistent shocks (see also Miller, 1984; Pakes, 1986; and others).
Our Model:

- Agents are risk neutral.
- Model is identified semiparametrically:
  
  (i) non-parametric identification of distributions of unobservables that are *serially persistent*;
  
  (ii) earnings equations parametric (but flexible functional forms).
Empirical Contributions of This Paper

- Estimate true rates of return and compare with IRR.
- Decompose option values by stages (educational choices and times choices are made; account for delay).
- Estimate at each stage the respective contributions of non-linearity and learning to option values and rates of return.
- Estimate contributions of both cognitive and noncognitive skills to returns and costs.
- We analyze jointly high school dropout and GED returns, as well as returns to two year and four year colleges (Eckstein-Wolpin, 1999).
Relationship to Previous Work

- Like Weisbrod (1962) and Altonji (1993), we recognize the option value that comes from educational choices.
- Like Levhari and Weiss (1974) and Keane and Wolpin (1997), we recognize uncertainty in post-educational earnings.
- Like Altonji (1993), Arcidiacono (2004) and Santos (2008), we recognize the learning value of schooling.
- Unlike Keane and Wolpin (1997, 2001), we consider serially persistent shocks which agents learn about (as in Miller 1984 and Pakes 1986)
- This produces much greater estimated option values than an independent shock model.
- We define and estimate the appropriate rate of return for a dynamic model with serially persistent shocks, nonlinearity and learning.
Consider a simple economic model as prologue.

The model we estimate is much richer. This simple model motivates our analysis.

Periods and schooling levels are assumed to be the same.
Each schooling level $s$ characterized by a shock, $\epsilon_s$.

More precisely, suppose that there is uncertainty about net earnings conditional on $s$, so that actual discounted lifetime earnings for someone with $s$ years of school are

$$Y_s = \left[ \sum_{x=0}^{T} (1 + r)^{-x} Y(s, x) \right] \epsilon_s.$$
s is ordered; $s' > s$ means more schooling in $s$. 
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- $s$ is ordered; $s' > s$ means more schooling in $s$.
- A one time, schooling (state) specific shock.
s is ordered; \( s' > s \) means more schooling in \( s \).

A one time, schooling (state) specific shock.

Assume that \( E(\epsilon_s \mid \mathcal{I}_{s-1}) = 1 \) and define expected earnings associated with schooling \( s \) conditional on current schooling \( s - 1 \),

\[
\bar{Y}_s = E(Y_s \mid \mathcal{I}_{s-1}).
\]
The decision problem for a person with $s$ years of schooling given the sequential revelation of information is to complete another year of schooling if

$$Y_s \leq \frac{E(V_{s+1} | I_s)}{1 + r},$$

so the value of schooling level $s$, $V_s$, is

$$V_s = \max \left\{ Y_s, \frac{E(V_{s+1} | I_s)}{1 + r} \right\}$$

for $s < \bar{S}$, the maximum number of years of schooling.
The decision problem for a person with $s$ years of schooling given the sequential revelation of information is to complete another year of schooling if

$$Y_s \leq \frac{E(V_{s+1} | I_s)}{1 + r},$$

so the value of schooling level $s$, $V_s$, is

$$V_s = \max \left\{ Y_s, \frac{E(V_{s+1} | I_s)}{1 + r} \right\}$$

for $s < \bar{S}$, the maximum number of years of schooling.

At $\bar{S}$, after all information is revealed, $V_{\bar{S}} = Y_{\bar{S}} = \bar{Y}_{\bar{S}} \in \bar{S}$. 
Endogenously determined probability of going on from school level $s$ to $s + 1$:

$$p_{s+1,s} = Pr \left( \epsilon_s \leq \frac{E(V_{s+1} | \mathcal{I}_s)}{(1 + r) \bar{Y}_s} \right),$$

where $E(V_{s+1} | \mathcal{I}_s)$ may depend on $\epsilon_s$ because it enters the agent’s information set.
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- Endogenously determined probability of going on from school level $s$ to $s + 1$:

$$ p_{s+1,s} = Pr \left( \epsilon_s \leq \frac{E(V_{s+1} | \mathcal{I}_s)}{(1 + r) \bar{Y}_s} \right), $$

where $E(V_{s+1} | \mathcal{I}_s)$ may depend on $\epsilon_s$ because it enters the agent's information set.

- The average earnings of a person who stops at schooling level $s$ are

$$ \bar{Y}_s E \left[ \left( \epsilon_s \mid \epsilon_s > \frac{E(V_{s+1} | \mathcal{I}_s)}{(1 + r) \bar{Y}_s} \right) \right]. \quad (1) $$
The expected value of schooling level $s$ as perceived at current schooling $s-1$ is:

$$E(V_s \mid I_{s-1})$$

$$= (1 - E(p_{s+1,s} \mid I_{s-1})) \bar{Y}_s E \left[ \left( \epsilon_s \mid \epsilon_s > \frac{E_s(V_{s+1} \mid I_s)}{(1 + r)\bar{Y}_s} \right) \mid I_{s-1} \right]$$

$$+ E(p_{s+1,s} \mid I_{s-1}) \left( \frac{E(V_{s+1} \mid I_{s-1})}{1 + r} \right).$$
The expected value of schooling level $s$ as perceived at current schooling $s - 1$ is:

$$E(V_s \mid I_{s-1}) = (1 - E(p_{s+1,s} \mid I_{s-1})) \bar{Y}_s E \left[ \left( \epsilon_s \mid \epsilon_s > \frac{E_s(V_{s+1} \mid I_s)}{(1 + r)\bar{Y}_s} \right) \right| I_{s-1} \right]$$

$$+ E(p_{s+1,s} \mid I_{s-1}) \left( \frac{E(V_{s+1} \mid I_{s-1})}{1 + r} \right).$$

The first component is the direct return. The second component arises from the option to go on to higher levels of schooling.
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

If schooling choices are irreversible, the \textbf{option value} of schooling $s$, as perceived after completing $s - 1$ levels of schooling is

$$O_{s,s-1} = E \left( [V_s - Y_s] \mid I_{s-1} \right).$$
If schooling choices are irreversible, the option value of schooling \( s \), as perceived after completing \( s - 1 \) levels of schooling is

\[
O_{s,s-1} = E ([V_s - Y_s] \mid I_{s-1}) .
\]

Value of schooling is \( E(V_s \mid I_{s-1}) \).
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- The *ex ante* rate of return to schooling \( s \) as perceived at the end of stage \( s - 1 \), before the information is revealed, is

\[
R_{s,s-1} = \frac{E(V_s | I_{s-1}) - Y_{s-1}}{Y_{s-1}}.
\]  
(2)
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- The \textit{ex ante} rate of return to schooling $s$ as perceived at the end of stage $s-1$, before the information is revealed, is

$$R_{s,s-1} = \frac{E(V_s | I_{s-1}) - Y_{s-1}}{Y_{s-1}}. \quad (2)$$

- This expression assumes no direct costs of schooling.
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- If there are up-front direct costs of schooling, $C_{s-1}$, to advance beyond level $s - 1$, the *ex ante* return is

$$\tilde{R}_{s,s-1} = \frac{E(V_s | I_{s-1}) - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}}.$$ 

- This expression assumes that tuition or direct costs are incurred up front and that returns are revealed one period later.
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

\( \tilde{R}_{s,s-1} \) is an appropriate *ex ante* rate of return concept because if

\[
Y_{s-1} + C_{s-1} \leq \frac{E(V_s \mid I_{s-1})}{1 + r},
\]

i.e.,

\[
r \leq \frac{E(V_s \mid I_{s-1}) - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}} = \tilde{R}_{s,s-1},
\]

then it would be optimal to advance one more year of schooling (from \( s-1 \) to \( s \)) given the assumed certain return on physical capital \( r \).
Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- \( \tilde{R}_{s,s-1} \) is an appropriate *ex ante* rate of return concept because if

\[
Y_{s-1} + C_{s-1} \leq \frac{E(V_s | I_{s-1})}{1 + r},
\]

i.e.,

\[
r \leq \frac{E(V_s | I_{s-1}) - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}} = \tilde{R}_{s,s-1},
\]

then it would be optimal to advance one more year of schooling (from \( s-1 \) to \( s \)) given the assumed certain return on physical capital \( r \).

- The *ex post* return as of period \( s \) is

\[
\frac{V_s - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}}.
\]
The distinction between \textit{ex ante} and \textit{ex post} returns to schooling is an important one that is not made in the conventional literature on “returns to schooling” surveyed in Willis (1986) or Katz and Autor (1999).
The distinction between *ex ante* and *ex post* returns to schooling is an important one that is not made in the conventional literature on “returns to schooling” surveyed in Willis (1986) or Katz and Autor (1999).

Levhari and Weiss (1974) and Altonji (1993) make this distinction.
Illustrate the role of uncertainty and non-linearity of log earnings in terms of schooling.
Example

- Illustrate the role of uncertainty and non-linearity of log earnings in terms of schooling.
- Simulate a five schooling-level version of the model with uncertainty.
Example

- Illustrate the role of uncertainty and non-linearity of log earnings in terms of schooling.
- Simulate a five schooling-level version of the model with uncertainty.
- Assume an interest rate of $r = 0.1$. 
Example

- Illustrate the role of uncertainty and non-linearity of log earnings in terms of schooling.
- Simulate a five schooling-level version of the model with uncertainty.
- Assume an interest rate of $r = 0.1$.
- Further assume that $\epsilon_s$ is independent and identically distributed log-normal: $\log(\epsilon_s) \sim N(0, \sigma)$ for all $s$. 
Example

- Illustrate the role of uncertainty and non-linearity of log earnings in terms of schooling.
- Simulate a five schooling-level version of the model with uncertainty.
- Assume an interest rate of $r = 0.1$.
- Further assume that $\epsilon_s$ is independent and identically distributed log-normal: $\log(\epsilon_s) \sim \mathcal{N}(0, \sigma)$ for all $s$.
- Assume that $\sigma = 0.1$ in the results presented in the tables.
Simulated Returns under Uncertainty with Option Values

(Log Wages Linear in Schooling: $\bar{Y}_{s+1} = (1 + r)\bar{Y}_s$)

| Educ. Level | Transition Probability in $\bar{Y}$ | Proportional Increase in Observed Earnings | Option/Total Value $E(R_{s,s-1} | I_s)$ | Avg. Return on Treatment | Treatment on Treated | Treatment on Untreated |
|-------------|-------------------------------------|-------------------------------------------|-------------------------------------|---------------------------|----------------------|------------------------|
| 2           | 0.796                               | 0.100                                      | 0.086                               | 0.075                     | 0.201                | 0.242                  | 0.041                  |
| 3           | 0.746                               | 0.100                                      | 0.082                               | 0.060                     | 0.182                | 0.231                  | 0.037                  |
| 4           | 0.669                               | 0.100                                      | 0.072                               | 0.038                     | 0.155                | 0.216                  | 0.032                  |
| 5           | 0.520                               | 0.100                                      | 0.016                               | 0.000                     | 0.111                | 0.196                  | 0.019                  |

OLS (Mincer) estimate of the rate of return is 0.063.

(Sheepskin Effects: $Y_{s+1} = (1 + \rho_{s+1})Y_s$ with $\rho_2 = 0.1$, $\rho_3 = 0.3$, $\rho_4 = 0.1$, $\rho_5 = 0.2$)

| Educ. Level | Transition Probability in $\bar{Y}$ | Proportional Increase in Observed Earnings | Option/Total Value $E(V_{s} | I_{s-1})$ | Avg. Return on Treatment | Treatment on Treated | Treatment on Untreated |
|-------------|-------------------------------------|-------------------------------------------|-------------------------------------|---------------------------|----------------------|------------------------|
| 2           | 0.997                               | 0.100                                      | 0.101                               | 0.239                     | 0.459                | 0.460                  | 0.068                  |
| 3           | 0.997                               | 0.300                                      | 0.116                               | 0.100                     | 0.459                | 0.460                  | 0.068                  |
| 4           | 0.846                               | 0.100                                      | 0.092                               | 0.093                     | 0.224                | 0.257                  | 0.045                  |
| 5           | 0.822                               | 0.200                                      | 0.041                               | 0.000                     | 0.212                | 0.249                  | 0.043                  |

OLS (Mincer) estimate of the rate of return is 0.060.
The simulated model assumes lifetime earnings for someone with $s$ years of school equal $\bar{Y}_s \epsilon_s$ where $\epsilon_s$ are independent and identically distributed $\log(\epsilon_s) \sim N(0, 0.1)$. An interest rate of $r = 0.10$ is assumed. The transition probability from $s - 1$ to $s$ is given by

$$p_{s,s-1} = \Pr_{s-1} \left( \epsilon_{s-1} \leq \frac{E(V_s | I_s)}{(1 + r) \bar{Y}_{s-1}} \right),$$

where the subscript means that the agent conditions his/her information on that available at $s - 1$. Observed earnings for someone with $s$ years of school are

$$\bar{Y}_s E \left[ \left( \epsilon_s \mid \epsilon_s > \frac{E_s(V_{s+1} | I_s)}{(1 + r) \bar{Y}_s} \right) \left| I_{s-1} \right. \right],$$

and option values are $E_{s-1}(V_s - Y_s)$. The return to school year $s$ for someone with earnings $Y_{s-1}$ is $R_{s,s-1} = \frac{E(V_s | I_{s-1}) - Y_{s-1}}{Y_{s-1}}$. 
Average returns reflect the expected return over the full distribution of $Y_{s-1}$, or $E_{s-1}[R_{s,s-1}]$. “Treatment on Treated” reflects returns for those who continue to grade $s$, or

$$E \left[ R_{s,s-1} \mid \epsilon_{s-1} \leq E_{s-1}(V_s \mid I_s) \frac{Y_{s-1}}{(1 + r)} \right] I_{s-1}.$$ 

“Treatment on Untreated” reflects returns for those who do not continue to grade $s$, or

$$E \left[ R_{s,s-1} \mid \epsilon_{s-1} > E_{s-1}(V_s \mid I_s) \frac{Y_{s-1}}{(1 + r)} \right] I_{s-1}.$$ 

The marginal treatment effect equals $r = 0.10$. OLS (Mincer) estimate is the coefficient on schooling in a log earnings regression (the Mincer return).
A More General Model with Delay, Dropout and Return

Our Data: The Decision Tree Agents Face

- Educational choice $s$ is made at time $t$, $t \in \{1, \ldots, T\}$.
- The set $\{s(1), \ldots, s(T)\}$ is not necessarily ordered.
Given High School Enrollment at Age 18 — Top of Tree

Figure 1. Evolution of Education Achieved
NLSY79 - White Males

EHS = Enrolled in High School
HSG = High School Graduate
HSD = High School Dropout
ECO = Enrolled in College
SCO = Some College Completed
GED = GED Completed

Note: All figures are in 2000 US Dollars.
Figure 1. Evolution of Education Achieved
NLSY79 - White Males

14s1
17s1
18s2
19s1 –
20s2
22s2
24s2
26s2

EHS (1954)

HSD (210)
GED (22)

HSD (188)
GED (16)

EHS (1744)
HSD (159)
GED (48)
ECO (111)
HSG (274)
ECO (142)
HSG (1311)

Note: All figures are in 2000 US Dollars.
Evolution of Schooling Attainment
Given High School Enrollment at Age 19 — Top of Tree
NLSY79 — White Males

<table>
<thead>
<tr>
<th>Age 19</th>
<th>Age 20</th>
<th>Age 20</th>
<th>Age 21</th>
<th>Age 22</th>
<th>Age 23</th>
<th>Age 24</th>
<th>Age 26</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSGr 579</td>
<td>EnCo 142</td>
<td>E4Co 67</td>
<td>4Co 39</td>
<td>4Co 39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSGr 437</td>
<td>E2Co 75</td>
<td>2Co 32</td>
<td>2Co 32</td>
<td>2Co 32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSGr 437</td>
<td>S2Co 43</td>
<td>S4Co 28</td>
<td>S4Co 28</td>
<td>S2Co 43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSGr 437</td>
<td>S2Co 43</td>
<td>S2Co 43</td>
<td>S2Co 43</td>
<td>S2Co 43</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HSGr = High School Graduate
EnCo = Enrolled in College
EnNCo = Enrolled in N-Year College
SNCO = Some N-Year College
NCo = N-Year College Completed
State Space

(1) Information may arrive in each state and at each age;

(2) Information can be state specific (e.g. you learn something about yourself by going to college but you do not learn if you do not go to college).

(3) $\mathcal{I}(s, t)$ is the information set.

(4) Specify a (state)$\times$(age) specification $(s, t)$. 
Economic Model

At each age $t$ and state $s$ there is a current period net reward.

$$R(s, t), \quad s \in S, \quad t \in \{1, \ldots, T\}$$

Assume exponential discounting for this paper.

This is relaxed in work underway.

$I(s, t)$ is the information set in state $s$ at time $t$.

$s(t) \in S$ is the choice agents make at age $t$.

Full state at $t$ is $(s(t), t, I(s, t))$.

Assume a non-stochastic discount rate $r$.

$\kappa(s, t)$ is choice set open to a person when they are at schooling $s$ at time $t$. 
At time $t$ in state $s$, agent has a sequence of possible future choices. $\kappa(s, t)$ depends on current information $\mathcal{I}(s, t)$.

The agent’s value function at time $t$ is

$$V(s(t), t \mid \mathcal{I}(s, t)) = \max_{[s(\tau) \in \kappa(s, \tau)]_{\tau \geq t}} E \left[ \sum_{\tau \geq t} \frac{R(s(\tau), \tau)}{(1 + r)^{\tau-t}} \middle| \mathcal{I}(s, t) \right].$$
\[ V(s(t), t \mid I(s(t), t)) \]
\[ = \max_{s(t) \in \kappa(s, t)} \left\{ R(s(t), t) + \left( \frac{1}{1 + r} \right) \max_{\{s(t+1)\} \in \kappa(s(t), t+1)} E(V(I(s(t+1), t+1) \mid I(s, t), t)) \right\} \]

Let \( C(s, t) \) be per-period cost (associated with each schooling level)

\[ R(s, t) = Y(s, t) - C(s, t) \]

Actually, we work with a more general cost function \( C(s', s, t) \).

The stopping (for ever) value at \( s(t) \) at time \( t \) is

\[ E \left[ PV(s(t), t) \mid I(s(t), t) \right] = E \left[ \sum_{\ell=0}^{T} \frac{R(s(t), \ell)}{(1 + r)^{\ell-t}} \mid I(s(t), t) \right] \]
There are a variety of possible option values.

Consider a simple schooling model that illustrates our main points.
Option value: What you pay to have the “right to strike” in different periods. (Initial price given initial information.)

If you strike, you get the value of the portfolio in that period.

You get no current flow if you continue.

Option value is price of this stream at date “0”

Continuation values can be computed at each stage.
Figure: Standard Schooling Model with Irreversible Choices (Once you stop, you cannot return)

Each stage is one period.
This is the “standard” framework (consistent with Mincer).
By analogy with the American options literature, one definition of the option value is the value of the program associated with enrolling and being able to drop out.
Unlike that case, you can get a flow if you continue.
The option value can be computed up front or at each stage. (continuation value)
The pairwise internal rates of return compare one of the many branches with the other.
Figure:

Compare A with B (two of many possible choices)
Two dimensions: Level attained and when it is attained
Can define “option value” as return to the more general program
Or value of “options” open up by attaining a level of schooling
Option Value of College Calculation Schematic

Traditional High School Path
Option Value of College Calculation Schematic

Traditional College Path

- EnHS
- GED
- Coll
- Drop
- HSGr
- Coll
- Drop
- EnHS
- GED
- Coll
- Drop
- HSGr
- Coll
- Drop
- HSGr
- Coll
- Drop
- HSGr
- Coll
- Drop

Values: 14, 16, 18, 20, 22
Distinguish the option value of delaying one period.

Option value of being a high school grad at time $t + 1$ as perceived at time $t$:

$$E \max [V(\text{College}(t + 1)), V(\text{HS}(t + 1)) \mid \text{Enrolled}, t, \mathcal{I}(\text{enrolled}, t)] - E [V(\text{HS}(t + 1)) \mid \text{Enrolled}, t, \mathcal{I}(\text{enrolled}, t)]$$

Option value of ever being a high school graduate is computed across all branches (compare values of staying on in high school).

Can be compared to value of remaining as high school dropout.
Option Value of College Calculation Schematic
A Comparison between American Options and the Options Relevant to Education

- We generalize previous work by having people pick schooling levels $s, s = 1, \ldots, S$ at different times $t$.
- The per period reward for a person with schooling $s$ at time $t$ is $R(s, t)$.
- If you stop at $s$ forever at time $t^*$, the present value at time $t^*$ is

$$PV(s, t^*) = \sum_{t=t^*}^{T} \frac{E[R(s, t) | I(s, t^*)]}{(1 + r)^{t-t^*}}$$

- This is like cashing out of an American option.
If you stop at \( s \) at time \( t^* \), and then go on to \( s' \) the next period, the return evaluated at time \( t \) at state \( s' \) is

\[
R(s, t^*) + \frac{1}{1 + r} E[V(s', t + 1) | \mathcal{I}(s, t)].
\]

Similarly, there is the strategy that has the agent stay on at \( s \) for two periods then moves, etc.

\[
R(s, t^*) + \frac{E[R(s, t^* + 1) | \mathcal{I}(s, t^*)]}{1 + r} + \frac{1}{(1 + r)^2} E[V(s', t^* + 2) | \mathcal{I}(s, t^*)].
\]
Ex Ante Rate of Return to Choice $s(t)$ as viewed at $t - 1$

(a) Irreversible case for terminal state: Ex Ante return to advancing to $s(t)$ from $s(t - 1)$ given stopping value at $t - 1$:

$$V[s(t), t \mid I(s(t - 1), t - 1)] - E[PV(s(t - 1), t \mid I(s(t - 1), t - 1))]$$

$$E[PV(s(t - 1), t - 1 \mid I(s(t - 1), t - 1))]$$

If $> r$, continue.

Otherwise stop.
Ex Ante Rate of Return to Choice \( s(t) = \bar{s} \) at \( t \) given choice \( s(t - 1) \) as viewed at \( t - 1 \)

(b) **General case** (not necessarily terminal states):

\[
\frac{V(s(t), t | \mathcal{I}(s(t - 1), t - 1) - V(s(t - 1), t | \mathcal{I}(s(t - 1), t - 1))}{V(s(t - 1), t | \mathcal{I}(s(t - 1), t - 1)}
\]

If \( > r \), continue to \( s(t) \).

Otherwise stop.

- This is the rate of return to getting \( s(t) \) at \( t \). The agent might make the choice at \( t + 1 \) instead.

- *Ex post* returns are computed using different agent information sets.
In Summary, in Our Model

- Two distinct concepts.
  1. Ever making the choice
  2. Making the choice at $t$

- The traditional approach (Becker-Mincer) assumes schooling decisions are made at fixed ages and are irreversible.

- Compares two streams only

- Our evidence shows the opposite is true. A lot of fluidity — delay, dropping in and out.
Econometric Model

- We postulate a factor structure for arrival of information.
- Let $\Theta$ be set of factors.
- Agents update information using the factor structure.
- Occupying a state can reveal a factor.
Model’s Ingredients

- **Earnings:**

  \[ Y_i(s, t, \theta) = \alpha_t(s) + \beta_t(s)\theta_i + \epsilon_{t,i}(s) \text{ for } s \in S \]

where \( \epsilon_{t,i}(s) \perp \theta_i, \forall t \), and \( \epsilon_{t,i}(s) \perp \epsilon_{t',i'}(s) \forall t, t', i, i' \)
Ingredients

Model's Ingredients

- **Earnings:**

  \[ Y_i(s, t, \theta) = \alpha_t(s) + \beta_t(s)\theta_i + \epsilon_{t,i}(s) \text{ for } s \in S \]

  where \( \epsilon_{t,i}(s) \perp \perp \theta_i, \forall t \), and \( \epsilon_{t,i}(s) \perp \perp \epsilon_{t',i'}(s) \forall t, t', i, i' \)

- **Schooling Costs:** Cost of going from \( s(t) \) to \( s(t + 1) \)

  \[ C_i(s(t), s(t + 1), t, \theta) = \lambda_t(s(t), s(t + 1)) + \lambda^\theta_t(s(t), s(t + 1))\theta_i \]

  \[ + \epsilon_{t,i}^C(s(t), s(t + 1)) \]

  where \( \epsilon_{t,i}^C(s(t), s(t + 1)) \perp \perp \theta_i \),

  \( \epsilon_{t,i}^C(s(t), s(t + 1)) \perp \perp \epsilon_{t',i'}^C(s(t), s(t + 1)) \) for any \( t, t' \) (\( t \neq t' \) for individual \( i \)) and individuals \( i, i' \).
In a data set with truncated earnings histories due to upper limits on the length of panels, the estimated “cost” is true cost minus the discounted earnings history after truncation.

Model is identified using the analysis of Abbring and Heckman (2007) and Heckman and Navarro (2007).

Identification and interpretation of the factor structure models are facilitated by test score equations.
Ingredients

Model's Ingredients

- **Measurement System**: For agent $i$ the test score $j$ is:

\[ T_i(j, \theta) = \pi(j) + \pi^T(\theta) \theta_i + e_i(j) \]

where $e_i(j) \perp \theta$, and $e'_i(j') \perp e_i(j)$. Test scores include cognitive and non-cognitive terms. Test Score: Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Mathematical Knowledge, Numerical Operations, Coding Speed, Rotter Locus of Control Scale, Rosenberg Self-Esteem Scale.
Ingredients

Model’s Ingredients

- **Measurement System:** For agent $i$ the test score $j$ is:

$$T_i(j, \theta) = \pi(j) + \pi^T(\theta) \theta_i + e_i(j)$$

where $e_i(j) \perp \perp \theta$, and $e'_i(j') \perp \perp e_i(j)$.

- Test scores include cognitive and non-cognitive terms.

- **Test Score:** Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Mathematical Knowledge, Numerical Operations, Coding Speed, Rotter Locus of Control Scale, Rosenberg Self-Esteem Scale.
Ingredients

Model’s Ingredients

- Relies on “full support” conditions (identification at infinity) which we check.

- Information updating:
  1. Agents know the $X$.
  2. They know the parameters including factor loadings in cost and outcome equations.
  3. They learn about the $e$ (ex ante set to zero).
  4. They learn about components of $\theta$ which arrive when states are experienced (ex ante expected $\theta$ are zero).
Solution: Backward Induction

There are terminal states and solve by backward induction.
The Likelihood Function

- Let \( \{X_i, T_i, Y_i, I_i\} \) denote the observed data of agent \( i \).
  - \( X_i \): regressors
  - \( T_i \): test scores
  - \( Y_i \): earnings
  - \( I_i \): indicator variables for choice of state at time \( t \)

- Let \( \Theta_i \) denote the set of unobserved (by economist) factors that contributes to the experience of agent \( i \).

- Let \( \{\beta, \alpha, \mu, \sigma^2, p, \nu, \tau^2, \pi\} \) denote the set of all parameters in all equations.

\[
f \left( X_i, T_i, Y_i, I_i, \Theta_i \mid X_i; \beta, \alpha, \mu, \sigma^2, p, \nu, \tau^2, \pi \right) \\
= f \left( X_i, T_i, Y_i, I_i \mid \Theta_i, X_i; \beta, \alpha, \mu, \sigma^2, p \right) \prod_{\theta_t \in \Theta_i} f_{\theta_t} \left( \theta_t; \nu, \tau^2, \pi \right)
\]
(log) likelihood function

\[
I \left( \beta, \alpha, \mu, \sigma^2, p, \nu, \tau^2, \pi \mid \{X_i, T_i, Y_i, I_i\}_{i=1}^N \right)
= \sum_{i=1}^N \log \left\{ \int \cdots \int_{\theta \in \Theta_i} \left[ \prod_{T_i \in T_i} f_T (T_i \mid \theta'_i) \right] \left[ \prod_{S \in S_i} f_{Y_S} (Y_{S,i} \mid \theta'_i) \right] \left[ \prod_{S \in S_i} g_S (1 \mid \theta'_i) \right] \prod_{\theta_t \in \Theta_i} f_{\theta_t} (\theta_t \mid \nu, \tau^2, \pi) \right\} \prod_{\theta_t \in \Theta_i} d\theta_r
\]

We estimate using mixture of normals for \( \theta \) and the e.
Empirical Results

- Presentation of goodness of fit for model.
- Presentation of marginal gains at different margins under different information sets.
- Presentations of distributions of gains overall and by margins.
- Distributions of costs by transition.
- Sorting evidence.
- Option values
  1. Overall
  2. Decomposed by transition
  3. For each transition contribution due to learning and contribution due to nonlinearity
- Comparing IRR with correct rate of return.
- \textit{Ex ante vs. ex post} rate of return.
Goodness of Fit
The Data

- NLSY
- Some background statistics on the evolution of schooling.
Figure 1. Evolution of Schooling Attainment

NLSY79 – Sample of White Males

Note: NLSY79 − Sample of White Males

Legend:
- Enrolled in HS
- HS Graduates
- HS Dropouts
- Some College
- GED
- Some 2YC (Not Enrolled)
- Some 4YC (Not Enrolled)
- AA Degree
- BA/BS Degree

1981: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1982: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1983: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1984: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1985: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1986: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1987: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1988: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1989: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1990: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1991: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1992: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1993: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1994: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1995: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1996: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1997: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1998: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
1999: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2000: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2001: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2002: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2003: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2004: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2005: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2006: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2007: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2008: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2009: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2010: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2011: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2012: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2013: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2014: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2015: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2016: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2017: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2018: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2019: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2020: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2021: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2022: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2023: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
2024: Enrolled in HS, HS Graduates, HS Dropouts, Some College, GED, Some 2YC (Not Enrolled)
Figure 2. Evolution of Schooling Attainment

NLSY79 – Simulated Sample
Paths to the GED

**Actual** and **Simulated Data**

- **EHS** = Enrolled in High School
- **HSD** = High School Dropout
- **GED** = GED Recipient

Age 14 | Age 17 | Age 18 | Age 20 | Age 22 | Age 34
Traditional Rates of Return and the Model Estimated Rate of Return (Including Option Value Incentives)
Earnings per Age/Semester Ever High School Graduates versus Ever Four Year College Graduates Model versus Data
## Present Discounted Value of Earnings, Rate of Returns and Internal Rate of Returns High School Graduates versus Four Year College Graduates

<table>
<thead>
<tr>
<th>Present Value of Earnings (a)</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High School Graduates</td>
<td>Four Year College Grad</td>
</tr>
<tr>
<td>Present Value of Earnings</td>
<td>379.994</td>
<td>459.014</td>
</tr>
<tr>
<td>Rate of Return</td>
<td>20.8%</td>
<td>19.5%</td>
</tr>
<tr>
<td>IRR</td>
<td>9%</td>
<td>9%</td>
</tr>
<tr>
<td>Mincer Coefficient (Age 30)</td>
<td>9.45%</td>
<td>8.9%</td>
</tr>
</tbody>
</table>

Note: (a) We assume an annual discount rate of 3%.
Figure 1. Evolution of Education Achieved
NLSY79 - White Males

EHS = Enrolled in High School
HSG = High School Graduate
HSD = High School Dropout
ECO = Enrolled in College
SCO = Some College Completed
GED = GED Completed

Note: All figures are in 2000 US Dollars.
Present Discounted Value of Earnings, Rate of Returns and Internal Rate of Returns High School Graduates versus Four Year College Graduates

<table>
<thead>
<tr>
<th>Present Value of Earnings (a)</th>
<th>Ever Four Year College Grad.</th>
<th>Four Year College Grad. On Time</th>
<th>Ever Four Year College Grad. - Top Cognitive Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School Graduates</td>
<td>377.188</td>
<td>394.413</td>
<td>482.037</td>
</tr>
<tr>
<td>Four Year College Grad</td>
<td>450.795</td>
<td>509.991</td>
<td>551.627</td>
</tr>
<tr>
<td>Rate of Return</td>
<td>19.5%</td>
<td>29.3%</td>
<td>14.4%</td>
</tr>
<tr>
<td>IRR</td>
<td>9%</td>
<td>11%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Note: (a) We assume an annual discount rate of 3%.
Earning Profiles for High School Graduates at Age 18 Ever Four Year College Grad. Versus High School Grad. Top Decile of Cognitive Ability ($f_C > d_{10}^C$).
Estimated Distributions of Abilities

- First, overall distributions.
- Evidence on sorting by ability type.
Figure 1. Distribution of Cognitive Factor By Final Schooling Level (overall)

NLSY79 – Sample of White Males

Legend:
- Solid line: Four Year College
- Dashed line: Some Four Year College
- Dot-dashed line: Two Year College
- Dotted line: Some Two Year College
- Dashed-dotted line: High School Graduate
- Grey line: GED with Some College
- Black line: GED
- Dotted-dashed line: High School Dropout
Figure 2. Distribution of Non-Cognitive Factor  
By Final Schooling Level (overall)  
NLSY79 – Sample of White Males

Note:

NLSY79 – Sample of White Males

Figure 2. Distribution of Non-Cognitive Factor  
By Final Schooling Level (overall)  
NLSY79 – Sample of White Males

Note:
Schooling Levels by Decile of Cognitive Ability

<table>
<thead>
<tr>
<th>Decile</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
</tr>
<tr>
<td>8</td>
<td>0.8</td>
</tr>
<tr>
<td>9</td>
<td>0.9</td>
</tr>
<tr>
<td>10</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Legend:
- 4-YCollege
- Some 4-YCollege
- 2-YCollege
- Some 2-YCollege
- GED with Some College
- HS Diploma
- GED
- HS Dropout
Schooling Levels by Decile of Personality Trait

<table>
<thead>
<tr>
<th>Decile</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
</tr>
<tr>
<td>8</td>
<td>0.8</td>
</tr>
<tr>
<td>9</td>
<td>0.9</td>
</tr>
<tr>
<td>10</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Figure 3. Distribution of Cognitive Factor
By Final Schooling Level: HSD, GED, GED with Some College

NLSY79 – Sample of White Males
Figure 4. Distribution of Non-Cognitive Factor By Final Schooling Level: HSD, GED, GED with Some College

NLSY79 – Sample of White Males
Sorting of People into Schools Based on Cognitive and Noncognitive Abilities
Note: We use the convention that decile 1 enters the lowest ability levels, whereas decile 10 contains the highest ability levels. The levels are computed using the overall distribution.
Distribution of Unobserved Abilities: Transition from "High School Graduate at Age 18" to "College Enrollment at Age 19" or "High School Graduate at Age 19"

A. High School Graduates at Age 18

B. College Enrollment at Age 19

B2. High School Graduates at Age 19

C. Cognitive Noncognitive

Cognitive Noncognitive Noncognitive
Distribution of Unobserved Abilities: Transition from "Enrolled in a Four Year College at Age 19/2st Sem." to "Four Year College Graduate at Age 22" or "Some Four Year College at Age 22"

A. Enrolled in a Four Year College at Age 19

B1. Four Year College Grad. at Age 22

B2. Some Four Year College at Age 22
Distribution of Unobserved Abilities: Transition from "High School Graduate at Age 19/HS Diploma at Age 18" to "Enrolled in College at Age 20" or "High School Graduate at Age 20"

A. High School Grad. at Age 19

B1. Enrolled in College at Age 20

B2. High School Grad. at Age 20
Distribution of Unobserved Abilities: Transition from "Enrolled in Four Year College at Age 20/ 2nd Sem. / HS Diploma at Age 18" to "Four Year College Grad. at Age 23" or "Some Four Year College at Age 23"

A. Enrolled in Four Year College at Age 20

B1. Four Year College Grad. at Age 23
B2. Some Four Year College at Age 23

Cognitive

Noncognitive
Distribution of Unobserved Abilities: Transition from "High School Dropout at Age 20/ HS Student until age 18" to "GED at Age 22" or "High School Dropout at Age 22"

A. High School Dropout at Age 20-

B1. GED at Age 22

B2. High School Dropout at Age 22
Distribution of Unobserved Abilities: Transition from "GED at Age 22/Dropout at Age 19" to "Enrolled in College at Age 24" or "GED at Age 24"
Option Values of Various Educational States

- Consider estimating the option value of the GED as part of a general project to estimate the option value of different types of schooling and training.

- A stochastic dynamic programming model with information updating.

- A few people benefit. Most do not.
First consider option values for enrollment in a state *at one age* compared to another state *at that age*.

We later consider the value of having the option *whenever* it is used.
Early High School Dropouts (age 17) - Sample of White Males

$\Pr(\text{Option Value} < \$220) = 5\%$
$\Pr(\text{Option Value} < \$400) = 10\%$
$\Pr(\text{Option Value} < \$846) = 25\%$
$\Pr(\text{Option Value} < \$1,620) = 50\%$
$\Pr(\text{Option Value} < \$2,800) = 75\%$

Note: Source Heckman and Urzua (2008).
Average Option Value of Early GED (age 18)
Sample of Early HS Dropouts

$1,000 dollars

Noncognitive

Cognitive
Distribution of Option Values Associated with GED at Age 20
For High School Dropouts at Age 18 who dropped out by Age 17 – White Males

Note: Source Heckman and Urzua (2008).
Average Option Value Associated with GED at Age 20, by Deciles of Ability Levels

For High School Dropouts at Age 18 who dropped out by Age 17 – White Males

Cognitive
Noncognitive

22
Distribution of Option Values Associated with GED at Age 22

For High School Dropouts at Age 20 who dropped out by Age 19 – White Males

Note: Source Heckman and Urzua (2008.

Pr(Option Value<$35)=5% 
Pr(Option Value<$64)=10% 
Pr(Option Value<$161)=25% 
Pr(Option Value<$369)=50% 
Pr(Option Value<$781)=75%

Note: Source Heckman and Urzua (2008).
Distribution of Option Values Associated with GED at Age 22

For High School Dropouts at Age 20 who dropped out by Age 17 — White Males

Pr(Option Value<$89)=10%
Pr(Option Value<$105)=25%
Pr(Option Value<$228)=50%

Note: Source Heckman and Urzua (2008).
Simulation Exercise: The Effects of Eliminating the GED

Note: The numbers in columns (1) and (2) are computed as fractions of the overall population.

<table>
<thead>
<tr>
<th>Schooling Level</th>
<th>Simulated (1)</th>
<th>No GED (2)</th>
<th>Change in Rate (2)-(1)</th>
<th>% Change ((2)/(1) - 1)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four Year College</td>
<td>26.4%</td>
<td>28.0%</td>
<td>1.6%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Some Four Year College</td>
<td>7.0%</td>
<td>7.8%</td>
<td>0.8%</td>
<td>11.4%</td>
</tr>
<tr>
<td>Two Year College</td>
<td>5.8%</td>
<td>6.3%</td>
<td>0.5%</td>
<td>8.5%</td>
</tr>
<tr>
<td>Some Two Year College</td>
<td>9.3%</td>
<td>9.8%</td>
<td>0.5%</td>
<td>5.0%</td>
</tr>
<tr>
<td>Some College GED</td>
<td>2.9%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>High School Graduates</td>
<td>32.8%</td>
<td>35.0%</td>
<td>2.1%</td>
<td>6.5%</td>
</tr>
<tr>
<td>GEDs</td>
<td>3.6%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>High School Dropouts</td>
<td>12.1%</td>
<td>13.1%</td>
<td>1.0%</td>
<td>8.4%</td>
</tr>
</tbody>
</table>
Decompose: High School vs. College (enrollment)

(a) Option values by contribution from each transition.

(b) Sources: learning and nonlinearity.

(c) True Rate of return by age and by transition (perceived at different ages and transitions).
The Option Value of College Enrollment
High School Students at Age 17 (a)
Dynamic Schooling Model with Learning ($\mathcal{Z} = \{f_C, f_N, \theta\}$)

<table>
<thead>
<tr>
<th>Option Value and Its Decomposition</th>
<th>Unconditional</th>
<th>Cognitive Ability (c)</th>
<th>Noncognitive Ability (c)</th>
<th>Both Abilities (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Option Value Associated with College Enrollment=(1)+(2)+(3)+(4)</td>
<td>27,474</td>
<td>6,574</td>
<td>61,179</td>
<td>12,903</td>
</tr>
<tr>
<td>Decomposition:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) College Enrollment at Age 19 (on time) after graduating from HS at Age 18 (on time) (b)</td>
<td>16,604</td>
<td>395</td>
<td>50,036</td>
<td>11,683</td>
</tr>
<tr>
<td>(2) College Enrollment at Age 20 (delayed) after graduating from HS at Age 18 (on time)</td>
<td>5,390</td>
<td>1,490</td>
<td>7,244</td>
<td>960</td>
</tr>
<tr>
<td>(3) College Enrollment at Age 20 (delayed) after graduating from HS at Age 19 (delayed)</td>
<td>5,452</td>
<td>4,685</td>
<td>3,869</td>
<td>223</td>
</tr>
<tr>
<td>(4) College Enrollment at Age 24 after dropping out from HS at Age 19 and obtaining GED at age 22</td>
<td>28</td>
<td>4</td>
<td>30</td>
<td>37</td>
</tr>
</tbody>
</table>

Notes: (a) All the numbers are in thousands of dollars at age 17; (b) In this case the option value is generated imposing that the agent cannot go back to college in the future. The option value associated with this possibility is presented in (2); (c) $d_k^j$ denotes the j-th decile associated with factor k. The deciles are computed from the overall distributions of abilities.
### The Contribution of Learning to the Option Value of College Enrollment

#### High School Students at Age 17

#### Dynamic Schooling Model with Learning vs. without Learning

<table>
<thead>
<tr>
<th>Option Value and Its Decomposition</th>
<th>Learning I={f_C,f_N,θ}</th>
<th>No Learning I={f_C,f_N}</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Option Value Associated with College Enrollment= (1)+(2)+(3)+(4)</td>
<td>27,474</td>
<td>14,408</td>
<td>13,066</td>
</tr>
<tr>
<td>Decomposition:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) College Enrollment at Age 19 (on time) after Graduating from HS at Age 18 (on time)</td>
<td>16,604</td>
<td>8,861</td>
<td>7,743</td>
</tr>
<tr>
<td>(2) College Enrollment at Age 20 (delayed) after Graduating from HS at Age 18 (on time)</td>
<td>5,390</td>
<td>1,745</td>
<td>3,645</td>
</tr>
<tr>
<td>(3) College Enrollment at Age 20 (delayed) after Graduating from HS at Age 19 (delayed)</td>
<td>5,452</td>
<td>3,784</td>
<td>1,668</td>
</tr>
<tr>
<td>(4) College Enrollment at Age 24 after dropping out from HS at Age 19 and obtaining GED at age 22</td>
<td>28</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

Notes: (a) All the numbers are in thousands of dollars at age 17. The sample is unchanged across simulations, that is, the numbers are computed for those agents enrolled in high school at age 17 under the three factor model.; (b) In this case the option value is generated imposing that the agent cannot go back to college in the future. The option value associated with this possibility is presented in (2).
# High School Grad. versus. College Enrollment

## True Rate of Return

<table>
<thead>
<tr>
<th>Initial State</th>
<th>Average Treatment Effect</th>
<th>Treatment on the Treated</th>
<th>Treatment on the Untreated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unconditional</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) High School Grad. At Age 18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>4.80%</td>
<td>18.27%</td>
<td>-12.20%</td>
</tr>
<tr>
<td>Once for All</td>
<td>12.60%</td>
<td>20.25%</td>
<td>2.90%</td>
</tr>
<tr>
<td>(2) High School Grad. At Age 19</td>
<td>-51.30%</td>
<td>160%</td>
<td>-127%</td>
</tr>
<tr>
<td><strong>Grad. High School at Age 18</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) High School Grad. At Age 19</td>
<td>45.05%</td>
<td>487%</td>
<td>-175%</td>
</tr>
<tr>
<td><strong>Enrolled in HS at Age 18</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Ability Individuals (f_c&lt;d_5^C and f_N&lt;d_5^N)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) High School Grad. At Age 18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>-2.92%</td>
<td>9.06%</td>
<td>-9.02%</td>
</tr>
<tr>
<td>Once for All</td>
<td>-2.72%</td>
<td>9.64%</td>
<td>-9.01%</td>
</tr>
<tr>
<td>(2) High School Grad. At Age 19</td>
<td>-111.90%</td>
<td>38%</td>
<td>-125.20%</td>
</tr>
<tr>
<td><strong>Grad. High School at Age 18</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) High School Grad. At Age 19</td>
<td>-118.80%</td>
<td>36.40%</td>
<td>-136.20%</td>
</tr>
<tr>
<td><strong>Enrolled in HS at Age 18</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Ability Individuals (f_c&gt;d_4^C and f_N&gt;d_4^N)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) High School Grad. At Age 18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever</td>
<td>7.14%</td>
<td>20.52%</td>
<td>-13.57%</td>
</tr>
<tr>
<td>Once for All</td>
<td>20.12%</td>
<td>23.32%</td>
<td>15.17%</td>
</tr>
<tr>
<td>(2) High School Grad. At Age 19</td>
<td>-6.65%</td>
<td>183.20%</td>
<td>-117.50%</td>
</tr>
<tr>
<td><strong>Grad. High School at Age 18</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) High School Grad. At Age 19</td>
<td>221.10%</td>
<td>566.50%</td>
<td>-138.40%</td>
</tr>
<tr>
<td><strong>Enrolled in HS at Age 18</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Costs
Distribution of Costs: Transition from "High School Dropout at Age 17" to "GED at Age 18"

Note: Source Heckman and Urzua (2008).
Distribution of Costs: Transition from "High School Dropout at Age 20/HS student until Age 19" to "GED at Age 22"

Note: Source Heckman and Urzua (2008).
Note: Source Heckman and Urzua (2008).
Support Conditions Satisfied?
Figure. Support Conditions for the Analysis of High School Graduation

A. Overall Sample

B. Sample of High School Graduates

C. Sample of High School Dropouts

Notes: Each panel presents the distribution of the probability of graduating from high school. This probability is estimated from the structural model, and consequently, it is estimated taking into account the dynamic schooling decisions as well as unobserved heterogeneity.
We develop a model of educational choices with uncertainty, learning about serially correlated shocks, dropout and delay. We consider high school, dropout, GED and college choices jointly. We generalize the rate of return and show the inadequacy of the IRR and rates of return in this more general setting. Option values are computed by stage and due to nonlinearity and uncertainty. *Ex ante*/*ex post* distinctions are substantial. They are substantial and raise the rate of return substantially beyond traditional measures.
Theoretical Contributions of this Paper

- A dynamic sequential model of educational choices among discrete states with option values arising from learning and nonlinearity of reward functions at different stages of the life cycle.

- We build a model of schooling connecting high school dropping out, GED attainment, delay, college choices and returns.

- Define the correct concept of the rate of return to schooling in a dynamic model with uncertainty, nonlinearity and delay.


- Like Arcidiacono (2004), we model learning about persistent shocks (see also Miller, 1984; Pakes, 1986; and others).
• Agents are risk neutral.

• Our model is identified semiparametrically:

(i) non-parametric identification of distributions of unobservables that are *serially persistent*;

(ii) earnings equations parametric (but flexible functional forms).
Empirical Contributions of This Paper

- Estimate true rates of return and compare with IRR.
- Decompose option values by stages (educational choices and times choices are made; account for delay).
- Estimate at each stage the respective contributions of non-linearity and learning to option values and rates of return.
- Estimate contributions of both cognitive and noncognitive skills to returns and costs.
- We analyze jointly high school dropout and GED returns, as well as returns to two year and four year colleges (Eckstein-Wolpin, 1999).
- Schooling states $s$ need not be ordered.