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

James Heckman
University of Chicago
University College Dublin

Sergio Urzúa
Northwestern University

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', $s < s'$:

$$\begin{aligned} Y(s,t) &= \text{ earnings at schooling } s \text{ at age } t \\ Y(s',t) &= \text{ earnings at schooling } s' \text{ at age } t \end{aligned}$$

- 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 $s \longrightarrow s'$ 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.
- IRR $\gtrless r$ 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.
- Builds on previous work on dynamic selection into schooling (Altonji, 1993; Keane and Wolpin, 1997, 2001; Eckstein and Wolpin, 1999; Arcidiacono, 2004; Cameron and Heckman, 1998, 2001).
- 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 rocognize 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.

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

Simple Model

- 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, ϵ_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

$$Y_s = \left[\sum_{x=0}^{T} (1+r)^{-x} Y(s,x) \right] \epsilon_s.$$

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- 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.

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

$$\bar{Y}_s = E(Y_s \mid \mathcal{I}_{s-1}).$$

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- 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

$$Y_s \leq \frac{E(V_{s+1} \mid \mathcal{I}_s)}{1+r},$$

so the value of schooling level s, V_s, is

$$V_s = \max\left\{ Y_s, \frac{E(V_{s+1} \mid \mathcal{I}_s)}{1+r} \right\}$$

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

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- 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

$$Y_s \leq \frac{E(V_{s+1} \mid \mathcal{I}_s)}{1+r},$$

so the value of schooling level s, V_s, is

$$V_s = \max \left\{ Y_s, \frac{E(V_{s+1} \mid \mathcal{I}_s)}{1+r} \right\}$$

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

- At $\bar{S}$, after all information is revealed, $V_{\bar{S}} = Y_{\bar{S}} = \bar{Y}_{\bar{S}} \epsilon_{\bar{S}}$.

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

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

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

where $E(V_{s+1} \mid \mathcal{I}_s)$ may depend on ϵ_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 $s+1$:

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

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

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

$$\bar{Y}_s E\left[\left(\epsilon_s \mid \epsilon_s > \frac{E(V_{s+1} \mid \mathcal{I}_s)}{(1+r)\bar{Y}_s}\right)\right]. \tag{1}$$

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- The expected value of schooling level s as perceived at current schooling $s-1$ is:

$$
\begin{aligned}
& E(V_s \mid \mathcal{I}_{s-1}) \\
& = \left(1 - E\left(p_{s+1,s} \mid \mathcal{I}_{s-1}\right)\right) \bar{Y}_s E\left[\left(\epsilon_s \mid \epsilon_s > \frac{E_s(V_{s+1} \mid \mathcal{I}_s)}{(1+r)\bar{Y}_s}\right) \middle| \mathcal{I}_{s-1}\right] \\
& \quad + E\left(p_{s+1,s} \mid \mathcal{I}_{s-1}\right)\left(\frac{E(V_{s+1} \mid \mathcal{I}_{s-1})}{1+r}\right).
\end{aligned}
$$

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- The expected value of schooling level s as perceived at current schooling $s-1$ is:

$$
\begin{aligned}
& E(V_s \mid \mathcal{I}_{s-1}) \\
& = (1 - E\left(p_{s+1,s} \mid \mathcal{I}_{s-1}\right)) \bar{Y}_s E\left[\left(\epsilon_s \mid \epsilon_s > \frac{E_s(V_{s+1} \mid \mathcal{I}_s)}{(1+r)\bar{Y}_s}\right) \middle| \mathcal{I}_{s-1}\right] \\
& \quad + E\left(p_{s+1,s} \mid \mathcal{I}_{s-1}\right)\left(\frac{E(V_{s+1} \mid \mathcal{I}_{s-1})}{1+r}\right).
\end{aligned}
$$

- 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 **option value** of schooling s, as perceived after completing $s-1$ levels of schooling is

$$O_{s,s-1} = E\left([V_s - Y_s] \mid \mathcal{I}_{s-1}\right).$$

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

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

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

$$R_{s,s-1} = \frac{E(V_s \mid \mathcal{I}_{s-1}) - Y_{s-1}}{Y_{s-1}}. \quad (2)$$

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

$$R_{s,s-1} = \frac{E(V_s \mid \mathcal{I}_{s-1}) - Y_{s-1}}{Y_{s-1}}. \tag{2}$$

- 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

$$\widetilde{R}_{s,s-1} = \frac{E(V_s \mid \mathcal{I}_{s-1}) - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}}.$$

- 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

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

$$Y_{s-1} + C_{s-1} \leq \frac{E(V_s \mid \mathcal{I}_{s-1})}{1+r}, \tag{3}$$

i.e.,

$$r \leq \frac{E(V_s \mid \mathcal{I}_{s-1}) - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}} = \widetilde{R}_{s,s-1},$$

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

- $\widetilde{R}_{s,s-1}$ is an appropriate *ex ante* rate of return concept because if
$$Y_{s-1} + C_{s-1} \leq \frac{E(V_s \mid \mathcal{I}_{s-1})}{1+r}, \tag{3}$$
i.e.,
$$r \leq \frac{E(V_s \mid \mathcal{I}_{s-1}) - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}} = \widetilde{R}_{s,s-1},$$
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
$$\frac{V_s - (Y_{s-1} + C_{s-1})}{Y_{s-1} + C_{s-1}}.$$

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- 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).

Rate of Return to Schooling with Uncertainty, Learning About State-specific Shocks

- 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.

Example

- 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 ϵ_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 ϵ_s is independent and identically distributed log-normal: $\log(\epsilon_s) \sim N(0, \sigma)$ 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 (s)	Transition Probability ($p_{s,s-1}$)	Proportional Increase in $\bar{Y}$	Proportional Increase in Observed Earnings	Option/ Total Value $\frac{O_{s,s-1}}{E(V_s \mid \mathcal{I}_{s-1})}$	Avg. Return $E(R_{s,s-1} \mid \mathcal{I}_s)$	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: $\bar{Y}_{s+1} = (1 + \rho_{s+1})\bar{Y}_s$ with $\rho_2 = 0.1,\ \rho_3 = 0.3,\ \rho_4 = 0.1,\ \rho_5 = 0.2$)

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.

Notes

The simulated model assumes lifetime earnings for someone with s years of school equal $\bar{Y}_s \epsilon_s$ where ϵ_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

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

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

$$\bar{Y}_s E\left[\left(\epsilon_s \mid \epsilon_s > \frac{E_s(V_{s+1} \mid \mathcal{I}_s)}{(1+r)\bar{Y}_s}\right)\middle| \mathcal{I}_{s-1}\right],$$

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 $R_{s,s-1} = \frac{E(V_s|\mathcal{I}_{s-1}) - Y_{s-1}}{Y_{s-1}}$.

Notes (cont.)

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

$$E\left[R_{s,s-1} \mid \epsilon_{s-1} \leq \frac{E_{s-1}(V_s \mid \mathcal{I}_s)}{(1+r)\bar{Y}_{s-1}} \middle| \mathcal{I}_{s-1}\right].$$

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

$$E\left[R_{s,s-1} \mid \epsilon_{s-1} > \frac{E_{s-1}(V_s \mid \mathcal{I}_s)}{(1+r)\bar{Y}_{s-1}} \middle| \mathcal{I}_{s-1}\right].$$

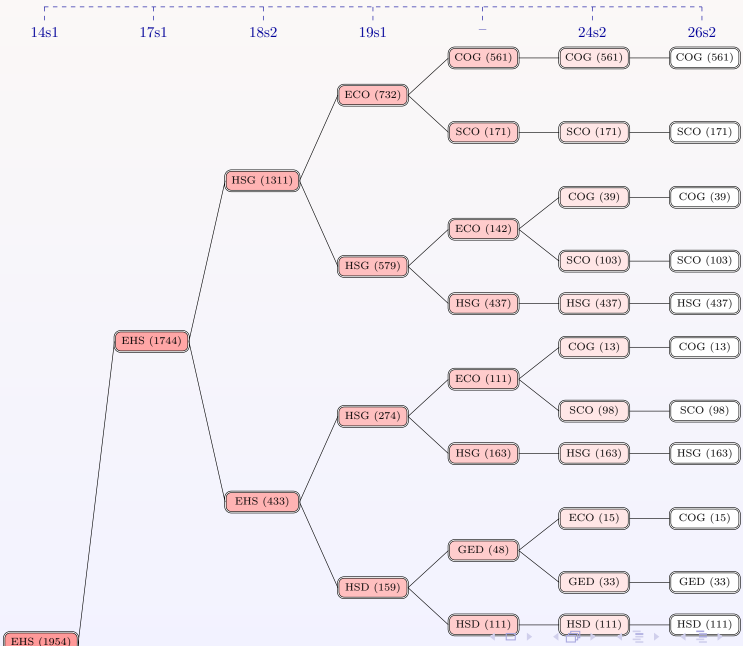
The marginal treatment effect equals $r = 0.10$. OLS (Mincer) estimate is the coefficient on schooling in a log earnings regression (the Mincer return).

A More General Model with Delay, Dropout and Return

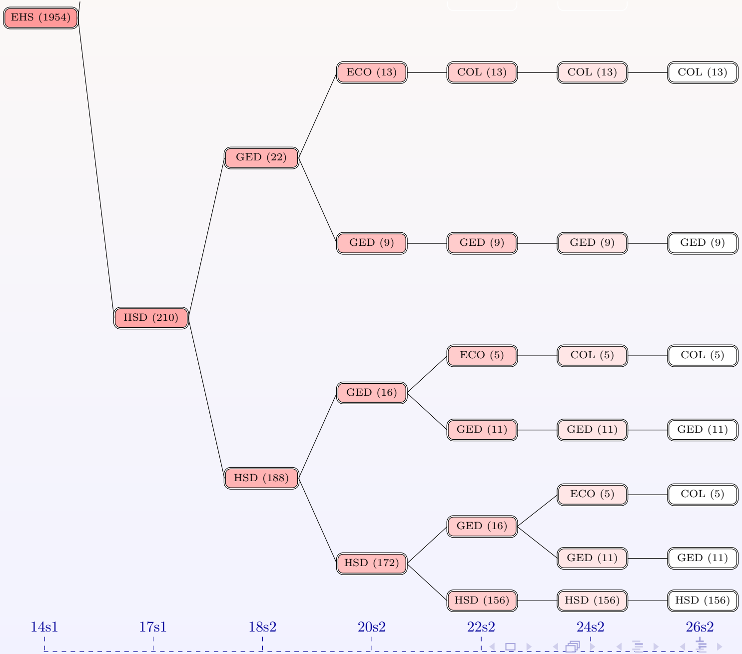
Our Data: The Decision Tree Agents Face

- 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

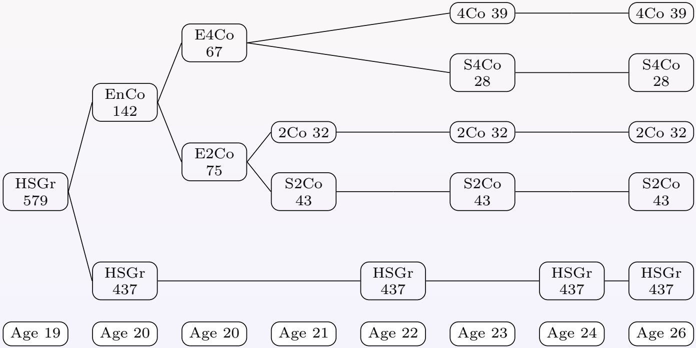


Given High School Enrollment at Age 19 — Bottom of Tree



Evolution of Schooling Attainment
Given High School Enrollment at Age 19 — Top of Tree

NLSY79 — White Males



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.

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

Assume exponential discounting for this paper.

This is relaxed in work underway.

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

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

Full state at t is $(s(t), t, \mathcal{I}(s,t))$.

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

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

$$
\begin{aligned}
&V(s(t), t \mid \mathcal{I}(s(t), t)) \\
&= \max_{s(t) \in \kappa(s,t)} \left\{ R(s(t), t) + \left(\frac{1}{1+r}\right) \max_{\{s(t+1)\} \in \kappa(s(t), t+1)} E(V(\mathcal{I}(s(t+1), t+1) \mid \mathcal{I}(s,t), t)) \right\}
\end{aligned}
$$

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

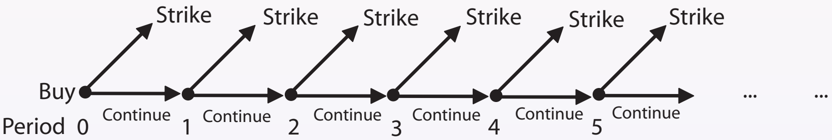
$$R(s, t) = Y(s, t) - C(s, t)$$

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

$$E\left[PV(s(t), t) \mid \mathcal{I}(s(t), t)\right] = E\left[\sum_{\ell=0}^{T} \frac{R(s(t), \ell)}{(1+r)^{\ell - t}} \middle| \mathcal{I}(s(t), t)\right]$$

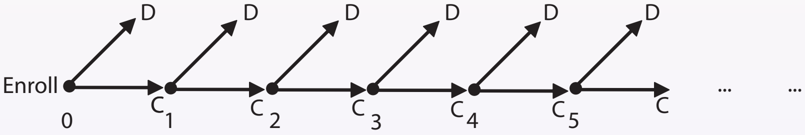
- There are a variety of possible option values.
- Consider a simple schooling model that illustrates our main points.

Figure: American Option



- 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.

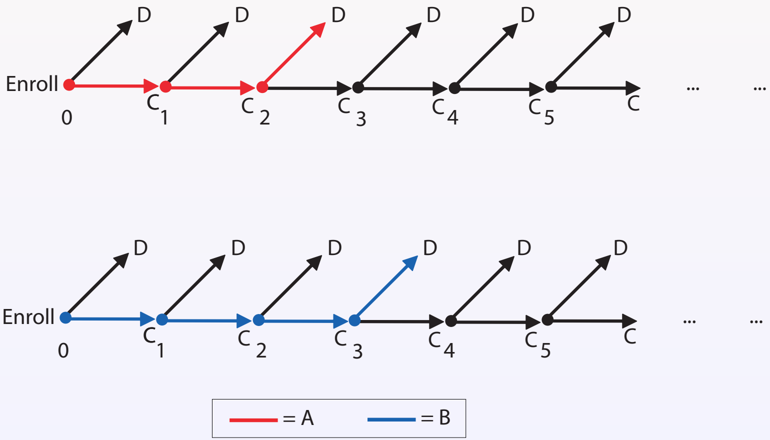
Figure: Standard Schooling Model with Irreversible Choices (Once you stop, you cannot return)



D = Dropout C = Continue

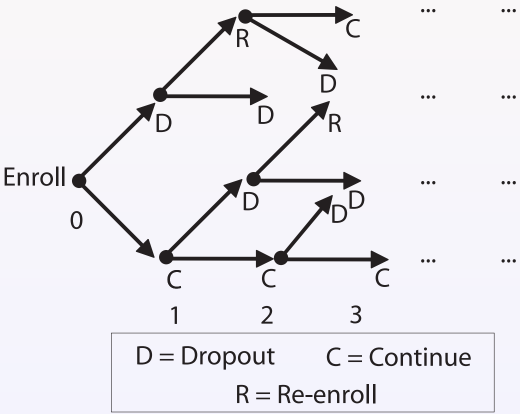
- Each stage is one period.
- This is the “standard” framework (consistent with Mincer).
- By analogy with the American options literature, one definition of the option value is the value of the program associated with enrolling and being able to drop out.
- Unlike that case, you can get a flow if you 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)

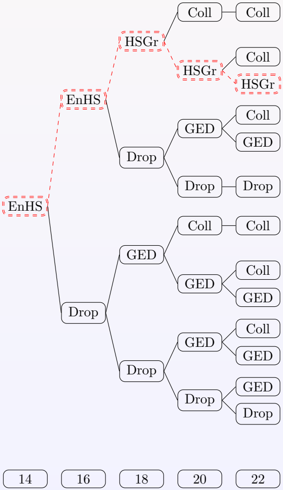
Figure: More General



- 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

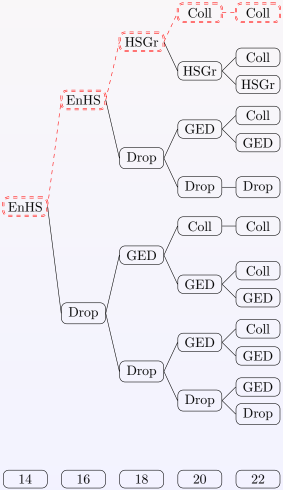
Option Value of College Calculation Schematic

Traditional High School Path



Option Value of College Calculation Schematic

Traditional College Path

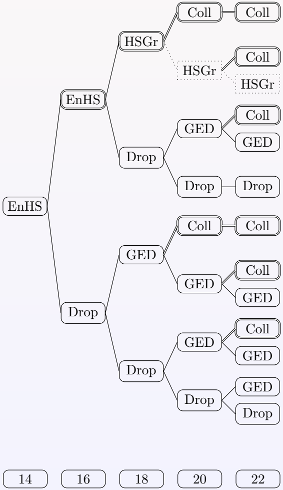


- Distinguish the option value of delaying one period.
- Option value of being a high school grad at time $t+1$ as perceived at time t:

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

- 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, $s = 1, \ldots, \overline{S}$ 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

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

- 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

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

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

$$\begin{aligned} R(s, t^*) &+ \frac{E[R(s, t^*+1) \mid \mathcal{I}(s, t^*)]}{1+r} \\ &+ \frac{1}{(1+r)^2} E[V(s', t^*+2) \mid \mathcal{I}(s, t^*)]. \end{aligned}$$

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 $t-1$:

$$\frac{V\left[s(t), t \mid \mathcal{I}\left(s(t-1), t-1\right)\right] - E\left[PV\left(s(t-1), t \mid \mathcal{I}(s(t-1), t-1)\right)\right]}{E\left[PV(s(t-1), t-1 \mid \mathcal{I}(s(t-1), t-1))\right]}$$

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):

$$\frac{V\left(s(t), t \mid \mathcal{I}(s(t-1), t-1\right) - V\left(s(t-1), t \mid \mathcal{I}(s(t-1), t-1\right)}{V\left(s(t-1), t \mid \mathcal{I}(s(t-1), t-1\right)}$$

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.
  - (i) Ever making the choice
  - (ii) 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 Θ be set of factors.
- Agents update information using the factor structure.
- Occupying a state can reveal a factor.

Model's Ingredients

- Earnings:

$$Y_i(s,t,\theta) \quad = \quad \alpha_t(s) + \beta_t(s)\theta_i + \epsilon^Y_{t,i}(s) \text{ for } s \in \mathcal{S}$$

where $\epsilon^Y_{t,i}(s) \perp\!\!\!\perp \theta_i, \forall t$, and $\epsilon^Y_{t,i}(s) \perp\!\!\!\perp \epsilon^Y_{t',i'}(s) \forall t, t', i, i'$

Model's Ingredients

- Earnings:

$$Y_i(s,t,\theta) \quad = \quad \alpha_t(s) + \beta_t(s)\theta_i + \epsilon^Y_{t,i}(s) \text{ for } s \in \mathcal{S}$$

where $\epsilon^Y_{t,i}(s) \perp\!\!\!\perp \theta_i, \forall t$, and $\epsilon^Y_{t,i}(s) \perp\!\!\!\perp \epsilon^Y_{t',i'}(s) \forall t, t', i, i'$

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

$$\begin{aligned} C_i(s(t), s(t+1), t, \theta) = \lambda_t(s(t), s(t+1)) + \lambda^\theta_t(s(t), s(t+1))\theta_i \\ + \varepsilon^C_{t,i}(s(t), s(t+1)) \end{aligned}$$

where $\varepsilon^C_{t,i}(s(t), s(t+1)) \perp\!\!\!\perp \theta_i$,
$\varepsilon^C_{t,i}(s(t), s(t+1)) \perp\!\!\!\perp \varepsilon^C_{t',i'}(s(t), s(t+1))$ for any t, t' ($t \neq t'$ for individual i) and individuals i, i'.

Model Ingredients

- 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.

Model's Ingredients

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

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

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

Model's Ingredients

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

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

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.

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 θ which arrive when states are experienced (*ex ante* expected θ are zero).

Solution: Backward Induction

There are terminal states and solve by backward induction.

The Likelihood Function

- Let $\{\mathbf{X}_i, \mathbf{T}_i, \mathbf{Y}_i, \mathbf{I}_i\}$ denote the observed data of agent i.

$$\begin{aligned} \mathbf{X}_i &: \text{ regressors} \\ \mathbf{T}_i &: \text{ test scores} \\ \mathbf{Y}_i &: \text{ earnings} \\ \mathbf{I}_i &: \text{ indicator variables for choice of state at time } t \end{aligned}$$

- Let $\mathbf{\Theta}_i$ denote the set of unobserved (by economist) factors that contributes to the experience of agent i.
- Let $\{\beta, \alpha, \mu, \sigma^2, \mathbf{p}, \nu, \tau^2, \pi\}$ denote the set of all parameters in all equations.

$$\begin{aligned} & f\left(\mathbf{X}_i, \mathbf{T}_i, \mathbf{Y}_i, \mathbf{I}_i, \mathbf{\Theta}_i \mid \mathbf{X}_i; \beta, \alpha, \mu, \sigma^2, \mathbf{p}, \nu, \tau^2, \pi\right) \\ &= f\left(\mathbf{X}_i, \mathbf{T}_i, \mathbf{Y}_i, \mathbf{I}_i | \mathbf{\Theta}_i, \mathbf{X}_i; \beta, \alpha, \mu, \sigma^2, \mathbf{p}\right) \prod_{\theta_t \in \mathbf{\Theta}_i} f_{\theta_t}\left(\theta_t; \nu, \tau^2, \pi\right) \end{aligned}$$

- (log) likelihood function

$$l\left(\beta,\alpha,\mu,\sigma^2,\mathbf{p},\nu,\tau^2,\pi \middle| \left\{\mathbf{X}_i,\mathbf{T}_i,\mathbf{Y}_i,\mathbf{I}_i\right\}_{i=1}^{N}\right)$$
$$=\sum_{i=1}^{N}\log\left\{\int\cdots\int_{\theta\in\mathbf{\Theta}_i}\left[\prod_{T_i\in\mathcal{T}_i} f_T\left(T_i|\theta_i'\right)\right]\left[\prod_{S\in\mathcal{S}_i} f_{Y_S}\left(Y_{S,i}|\theta_i'\right)\right]\right.$$
$$\left.\left[\prod_{S\in\mathcal{S}_i} g_S\left(1|\theta_i'\right)\right]\prod_{\theta_t\in\mathbf{\Theta}_i} f_{\theta_r}\left(\theta_r|\nu,\tau^2,\pi\right)d\theta_r\right\}$$

- We estimate using mixture of normals for $\underset{\sim}{\theta}$ 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.
- *Ex ante vs. ex post* rate of return.

Goodness of Fit

The Data

- NLSY
- Some background statistics on the evolution of schooling.

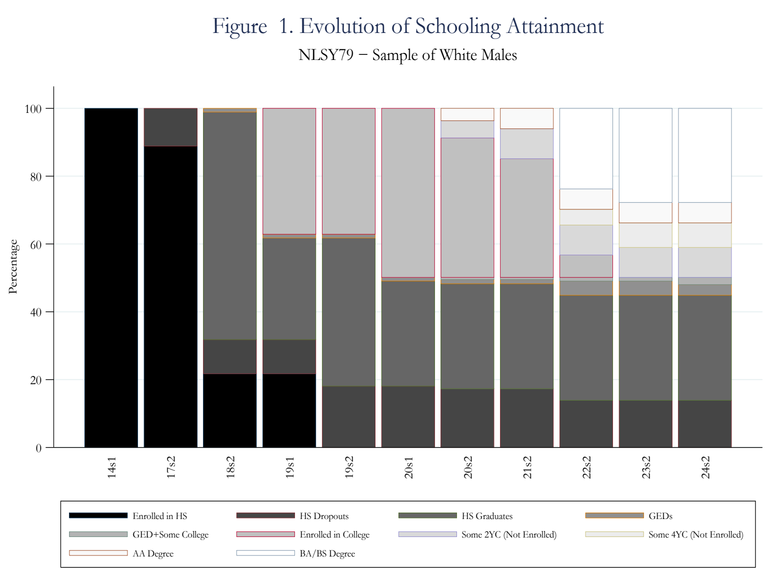
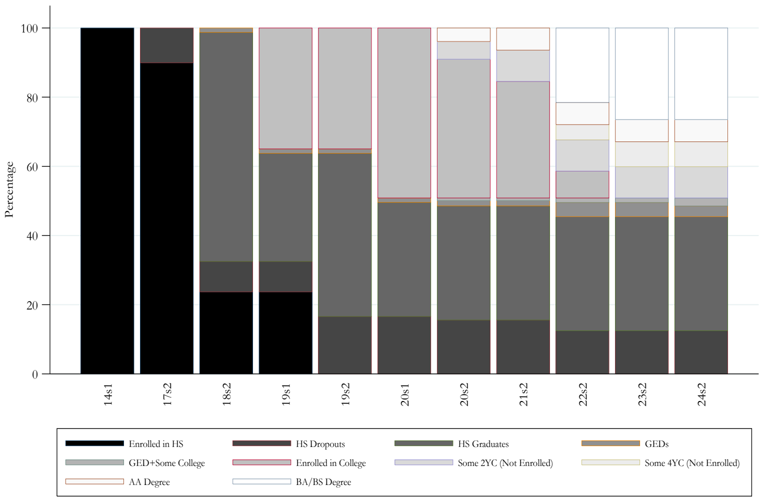


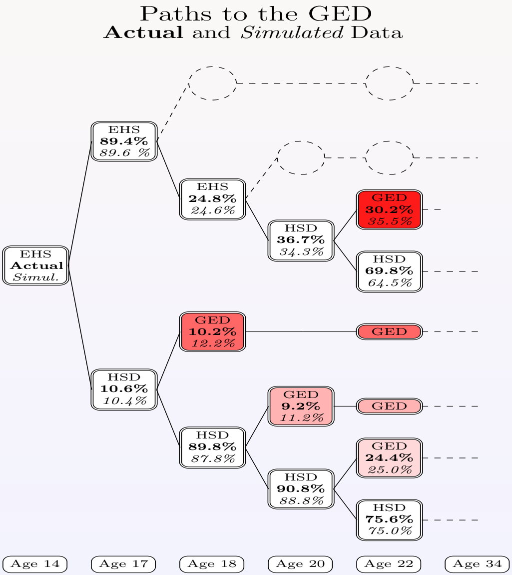
Figure 1. Evolution of Schooling Attainment
NLSY79 – Sample of White Males
Percentage
0
20
40
60
80
100
14s1
17s2
18s2
19s1
19s2
20s1
20s2
21s2
22s2
23s2
24s2
Enrolled in HS
HS Dropouts
HS Graduates
GEDs
GED+Some College
Enrolled in College
Some 2YC (Not Enrolled)
Some 4YC (Not Enrolled)
AA Degree
BA/BS Degree

Figure 2. Evolution of Schooling Attainment

NLSY79 – Simulated Sample

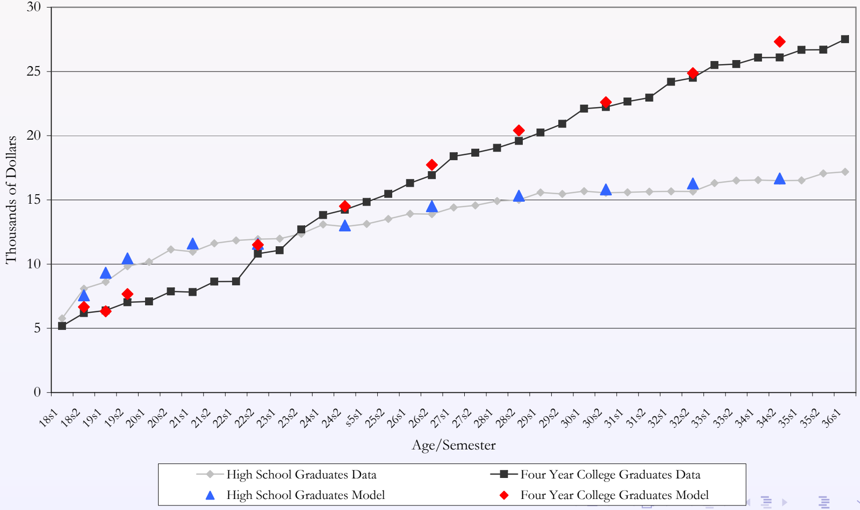


GED


Paths to the GED
Actual and Simulated Data
EHS
Actual
Simul.
EHS
89.4%
89.6 %
HSD
10.6%
10.4%
EHS
24.8%
24.6%
GED
10.2%
12.2%
HSD
89.8%
87.8%
HSD
36.7%
34.3%
GED
9.2%
11.2%
HSD
90.8%
88.8%
GED
30.2%
35.5%
HSD
69.8%
64.5%
GED
GED
GED
24.4%
25.0%
HSD
75.6%
75.0%
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

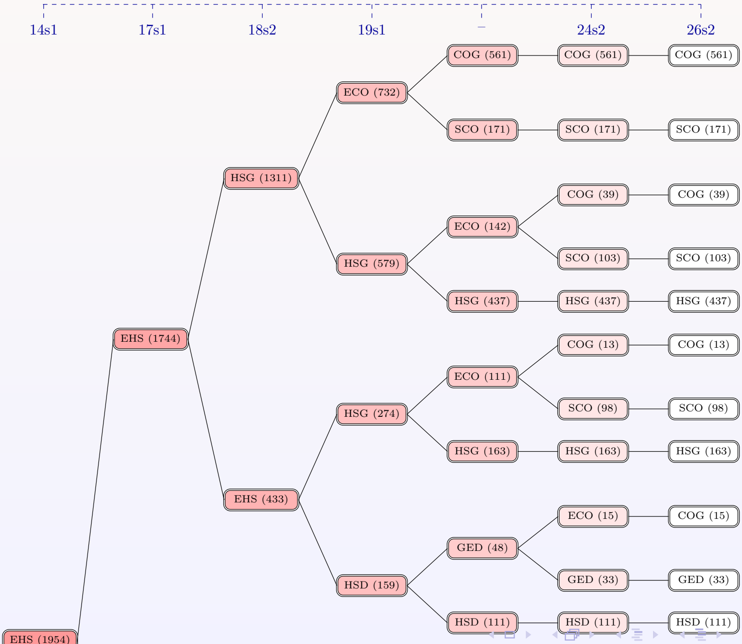


Present Discounted Value of Earnings, Rate of Returns and Internal Rate of Returns High School Graduates versus Four Year College Graduates

	Data		Model	
	High School Graduates	Four Year College Grad	High School Graduates	Four Year College Grad
Present Value of Earnings [a]	379.994	459.014	377.188	450.795
Rate of Return	20.8%		19.5%	
IRR	9%		9%	
Mincer Coefficient (Age 30)	9.45%		8.9%	

Note: (a) We assume an annual discount rate of 3%.

Given High School Enrollment at Age 18 — Top of Tree

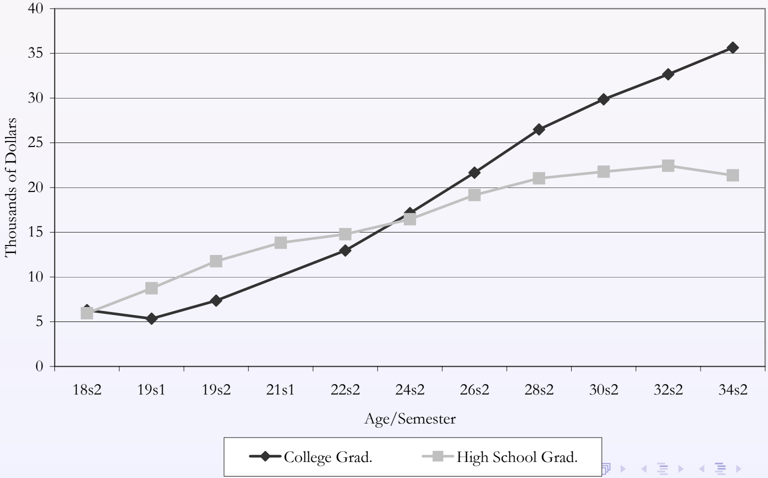


Present Discounted Value of Earnings, Rate of Returns and Internal Rate of Returns High School Graduates versus Four Year College Graduates

	Ever Four Year College Grad.		Four Year College Grad. On Time		Ever Four Year College Grad. - Top Cognitive Ability	
	High School Graduates	Four Year College Grad	High School Graduates	Four Year College Grad	High School Graduates	Four Year College Grad
Present Value of Earnings [(a)]	377.188	450.795	394.413	509.991	482.037	551.627
Rate of Return	19.5%		29.3%		14.4%	
IRR	9%		11%		6%	

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

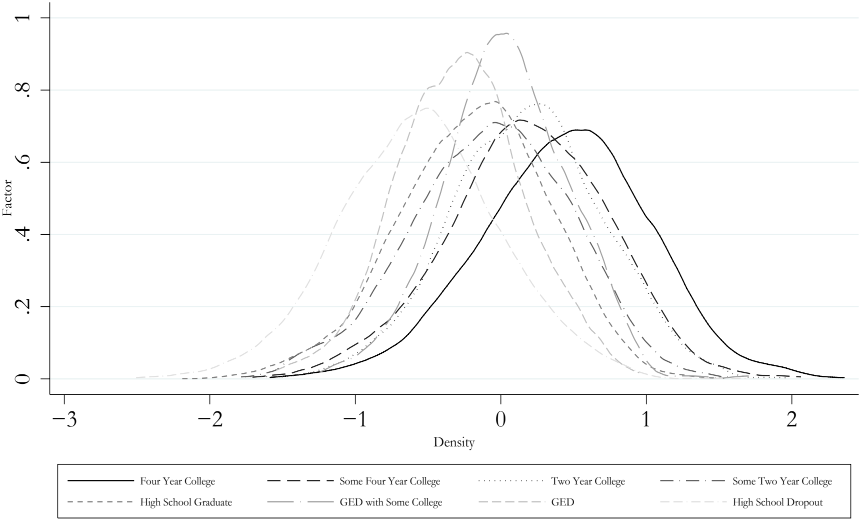
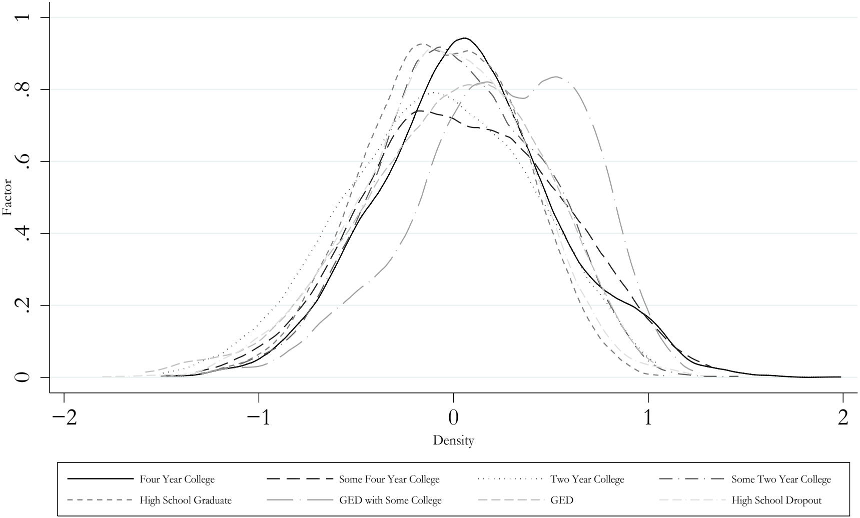
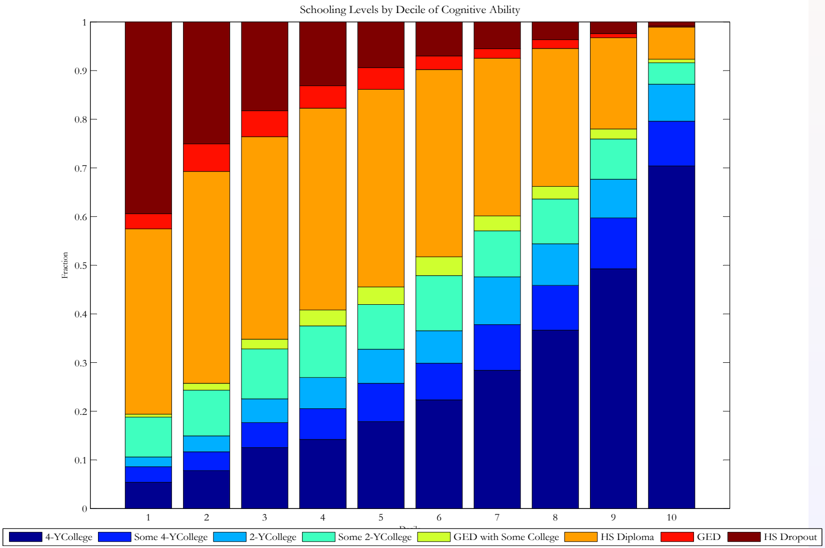
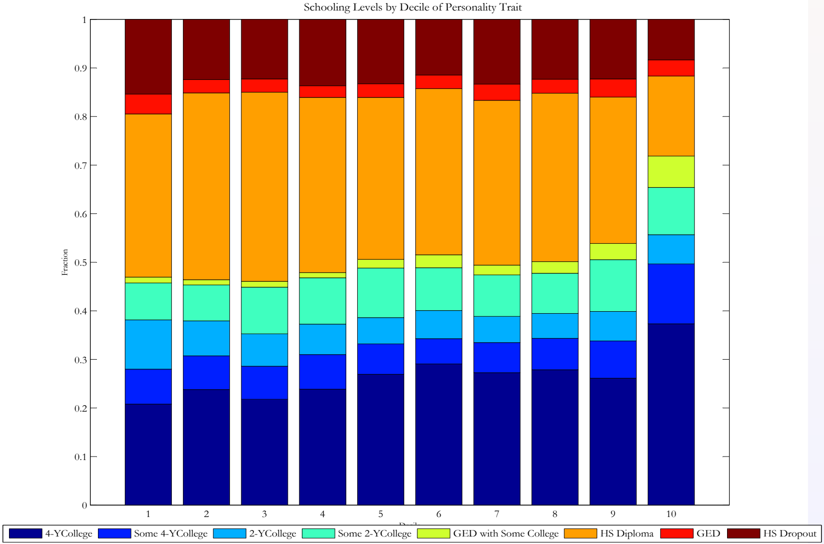


Figure 2. Distribution of Non−Cognitive Factor
By Final Schooling Level (overall)

NLSY79 − Sample of White Males




Schooling Levels by Decile of Cognitive Ability
Fraction
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
2
3
4
5
6
7
8
9
10
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
Fraction
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1 2 3 4 5 6 7 8 9 10
4-YCollege
Some 4-YCollege
2-YCollege
Some 2-YCollege
GED with Some College
HS Diploma
GED
HS Dropout

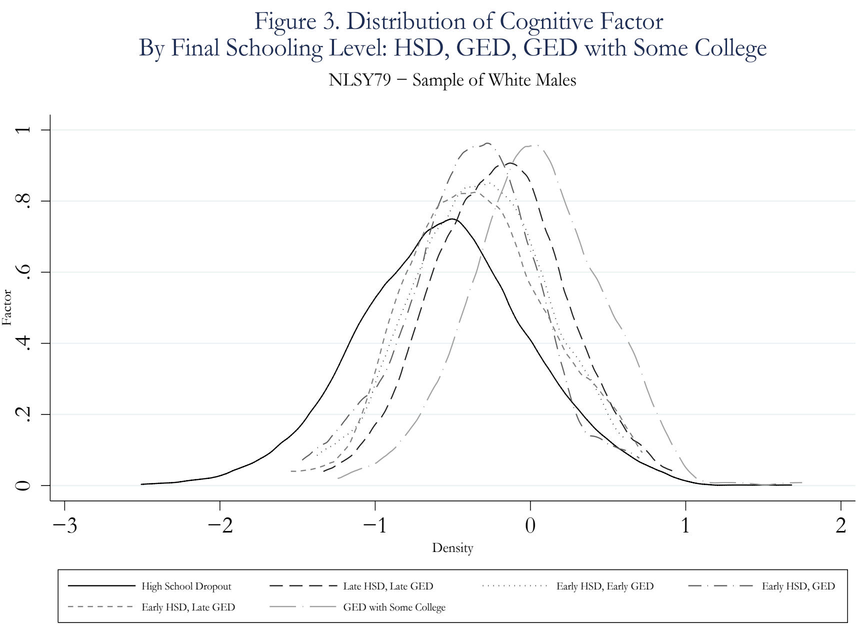


Figure 3. Distribution of Cognitive Factor
By Final Schooling Level: HSD, GED, GED with Some College
NLSY79 − Sample of White Males
Factor
0
.2
.4
.6
.8
1
−3
−2
−1
0
1
2
Density
High School Dropout
Late HSD, Late GED
Early HSD, Early GED
Early HSD, GED
Early HSD, Late GED
GED with Some College

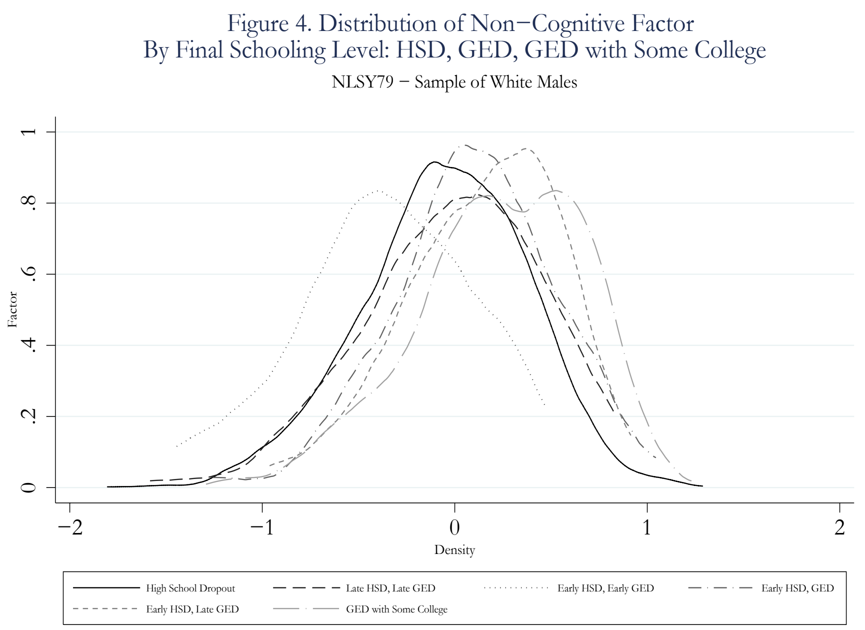
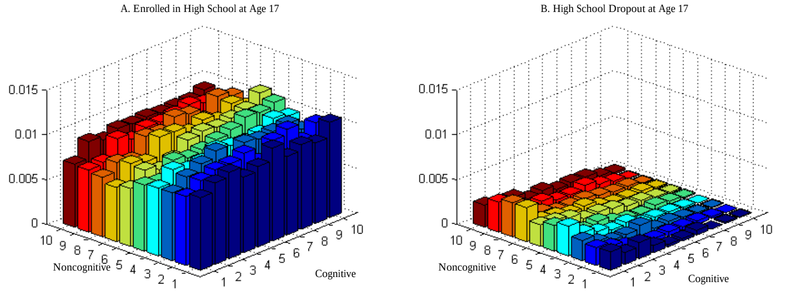


Figure 4. Distribution of Non−Cognitive Factor
By Final Schooling Level: HSD, GED, GED with Some College
NLSY79 − Sample of White Males
Factor
0
.2
.4
.6
.8
1
−2
−1
0
1
2
Density
High School Dropout
Late HSD, Late GED
Early HSD, Early GED
Early HSD, GED
Early HSD, Late GED
GED with Some College

Sorting of People into Schools Based on Cognitive and Noncognitive Abilities

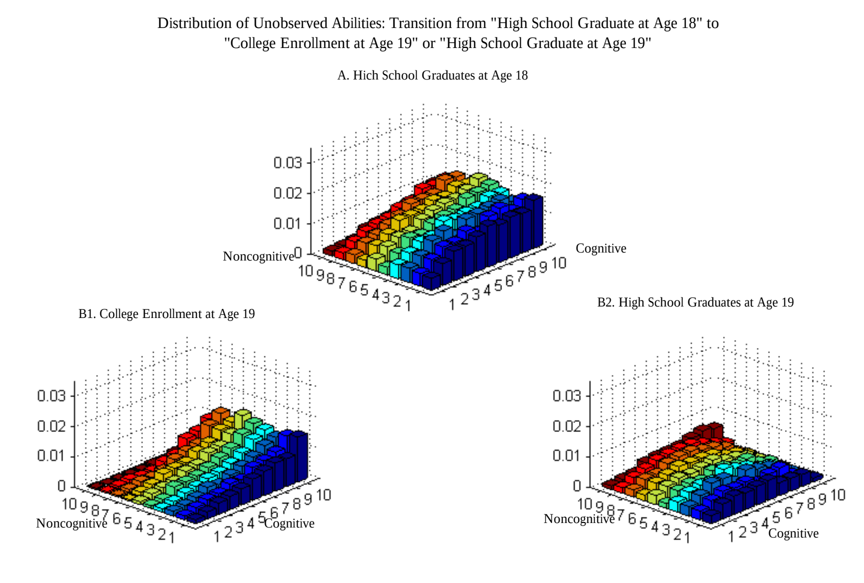
Schooling

Distribution of Unobserved Abilities by Schooling Level at Age 17: Transition from "Enrolled in HS" (Age 14) to "Enrolled in HS" or "HS Dropout"



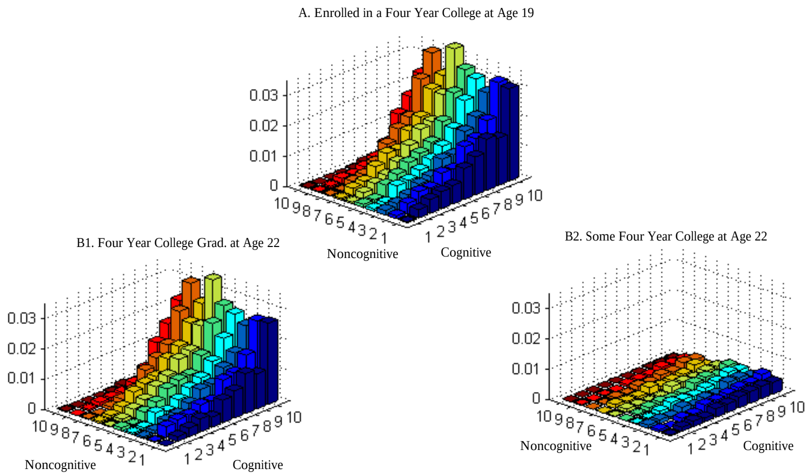
Note: We use the convention that decile 1 enters the lowest ability levels, whereas decile 10 contains the highest ability levels. The levels are computed using the overall distribution.

Schooling

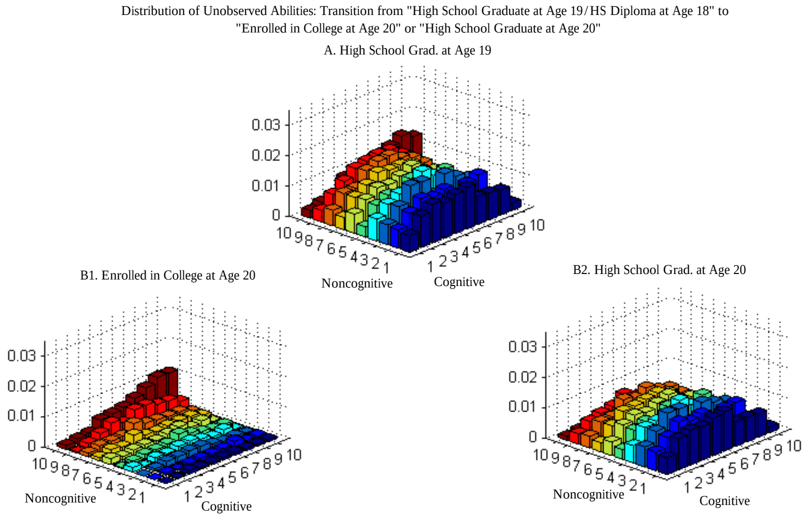

Distribution of Unobserved Abilities: Transition from "High School Graduate at Age 18" to "College Enrollment at Age 19" or "High School Graduate at Age 19"
A. Hich School Graduates at Age 18
0.03
0.02
0.01
0
Noncognitive
Cognitive
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10
B1. College Enrollment at Age 19
0.03
0.02
0.01
0
Noncognitive
Cognitive
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10
B2. High School Graduates at Age 19
0.03
0.02
0.01
0
Noncognitive
Cognitive
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10

Schooling

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"

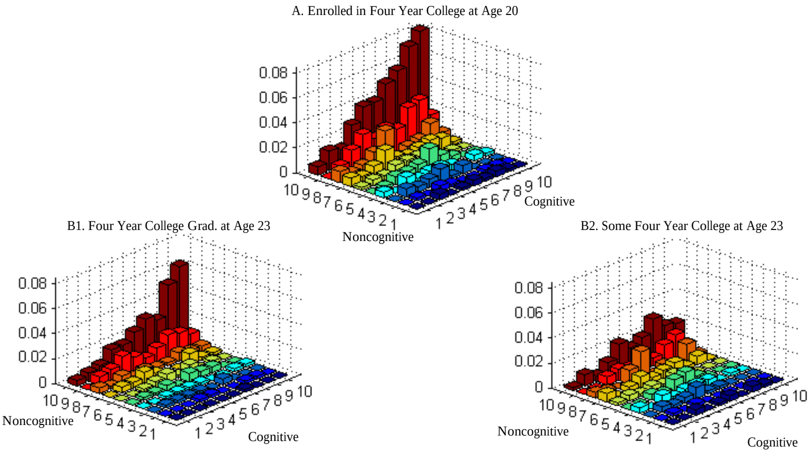


Schooling

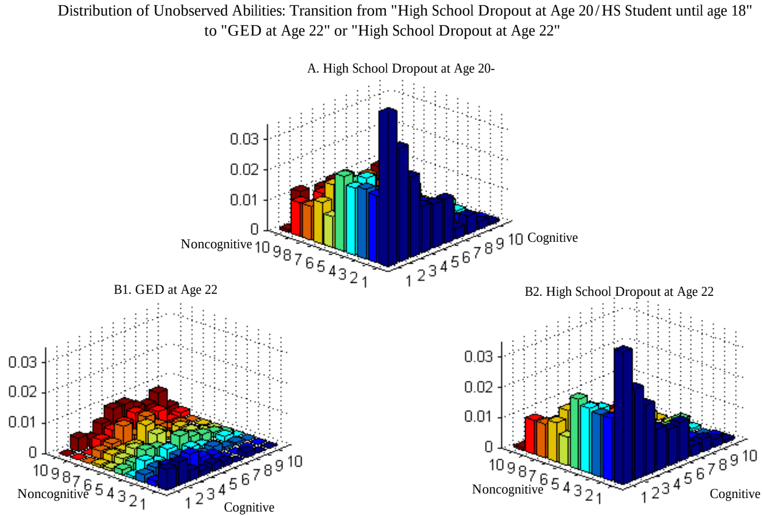

Distribution of Unobserved Abilities: Transition from "High School Graduate at Age 19/HS Diploma at Age 18" to "Enrolled in College at Age 20" or "High School Graduate at Age 20"
A. High School Grad. at Age 19
B1. Enrolled in College at Age 20
B2. High School Grad. at Age 20
0.03
0.02
0.01
0
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10
Noncognitive
Cognitive

Schooling

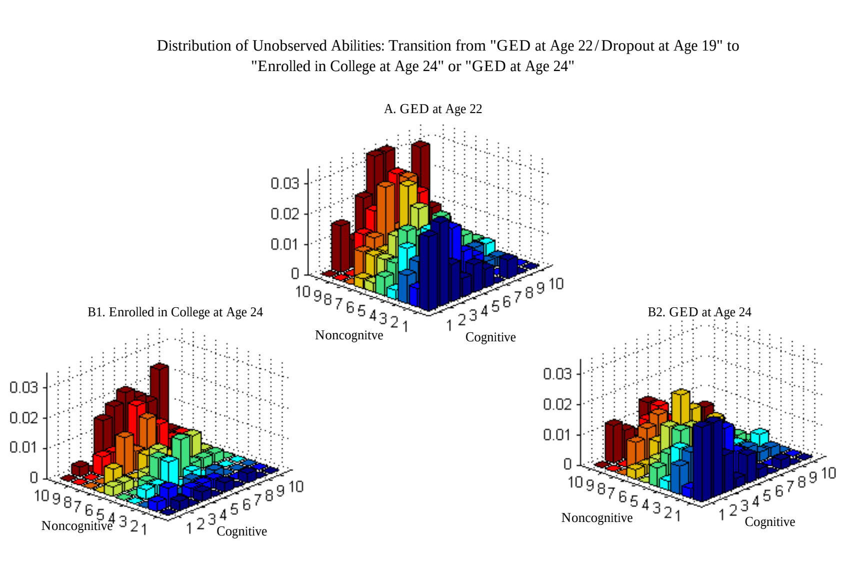
Distribution of Unobserved Abilities: Transition from "Enrolled in Four Year College at Age 20/2nd Sem. / HS Diploma at Age 18" to "Four Year College Grad. at Age 23" or "Some Four Year College at Age 23"



Schooling


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
0.03
0.02
0.01
0
Noncognitive
Cognitive
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10

Schooling

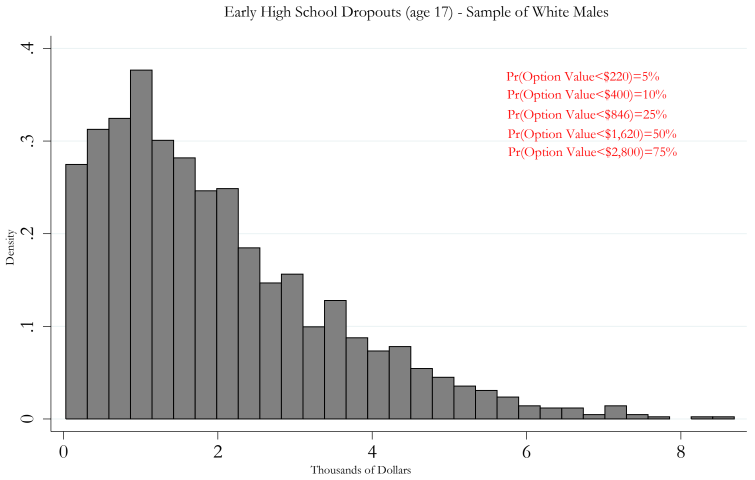

Distribution of Unobserved Abilities: Transition from "GED at Age 22/Dropout at Age 19" to "Enrolled in College at Age 24" or "GED at Age 24"
A. GED at Age 22
0.03
0.02
0.01
0
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10
Noncognitve
Cognitive
B1. Enrolled in College at Age 24
0.03
0.02
0.01
0
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10
Noncognitive
Cognitive
B2. GED at Age 24
0.03
0.02
0.01
0
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10
Noncognitive
Cognitive

Option Values of Various Educational States

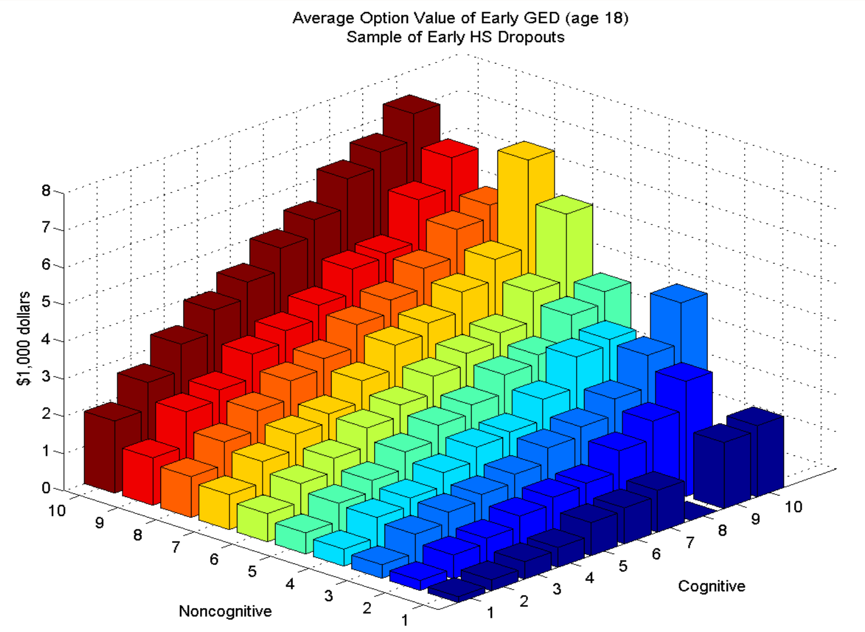
- Consider estimating the option value of the GED as part of a general project to estimate the option value of different types of schooling and training.
- A stochastic dynamic programming model with information updating.
- A few people benefit. Most do not.

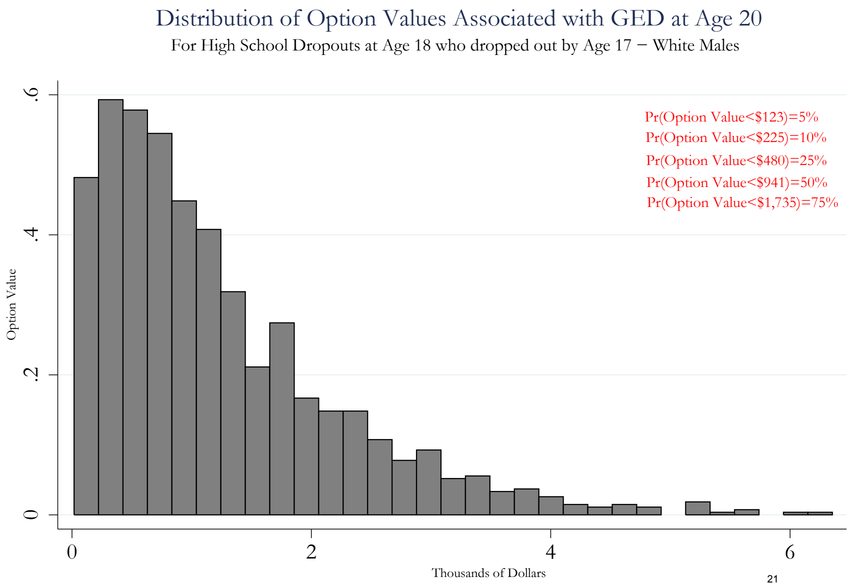
- First consider option values for enrollment in a state *at one age* compared to another state *at that age*.
- We later consider the value of having the option *whenever* it is used.

Distribution of Option Values - Early GED

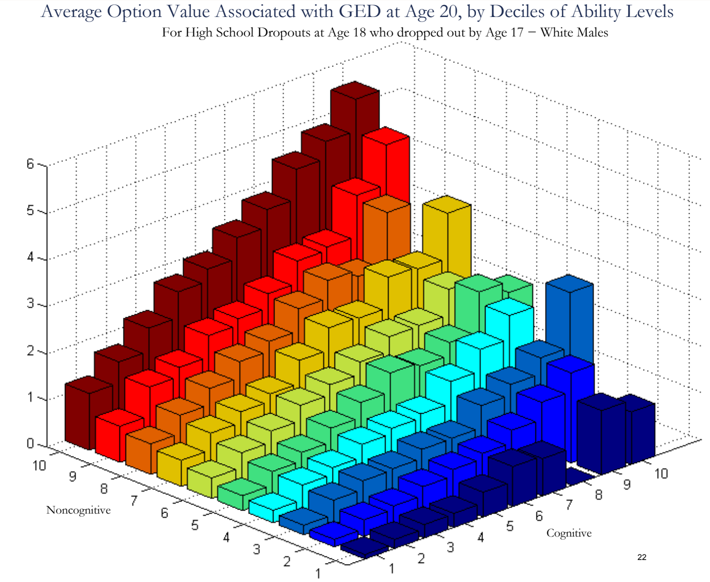


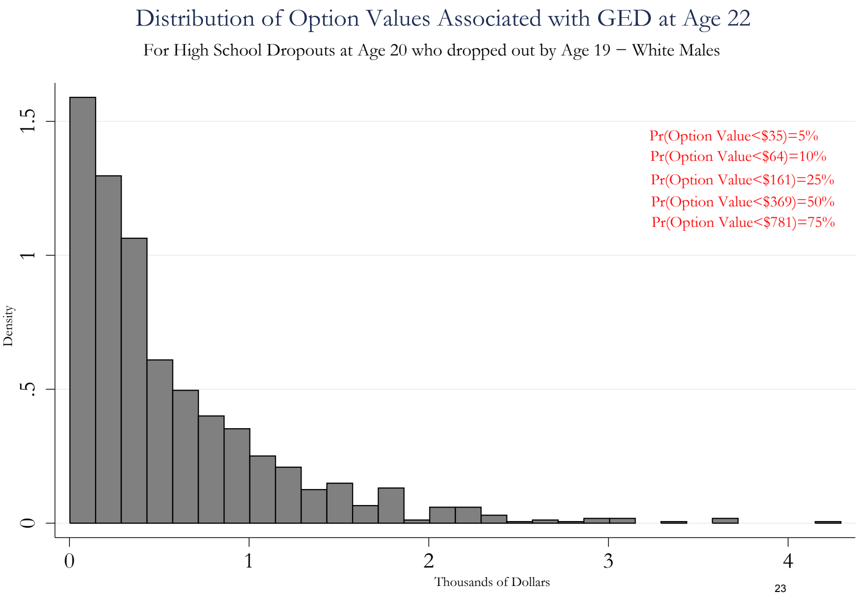
Note: Source Heckman and Urzua (2008).


Average Option Value of Early GED (age 18)
Sample of Early HS Dropouts
$1,000 dollars
0
1
2
3
4
5
6
7
8
Noncognitive
1
2
3
4
5
6
7
8
9
10
Cognitive
1
2
3
4
5
6
7
8
9
10

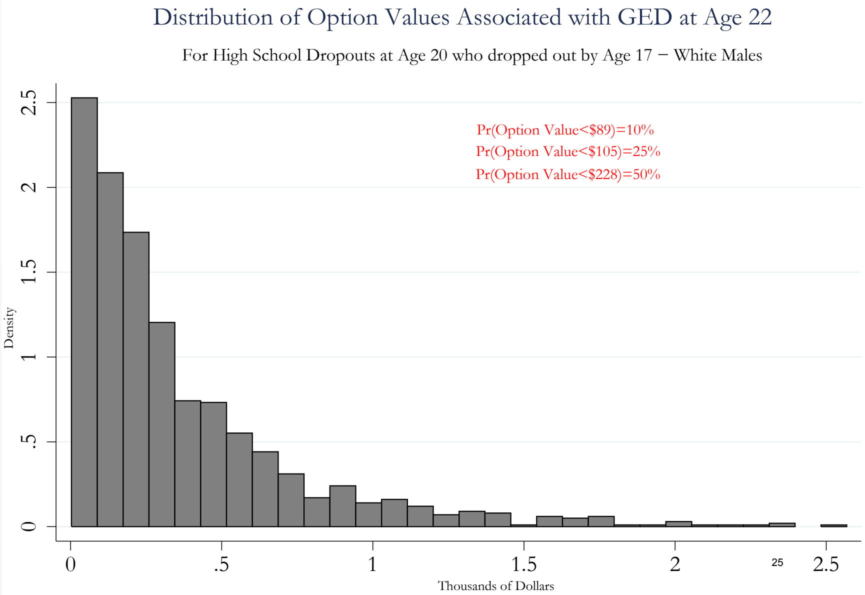


Note: Source Heckman and Urzua (2008).


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
0
1
2
3
4
5
6
Noncognitive
1
2
3
4
5
6
7
8
9
10
Cognitive
1
2
3
4
5
6
7
8
9
10
22



Note: Source Heckman and Urzua (2008).



Note: Source Heckman and Urzua (2008).

Simulation Exercise: The Effects of Eliminating the GED[a]

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

Schooling Level	Simulated (1)	No GED (2)	Change in Rate (2)-(1)	% Change ((2)/(1) - 1)%
Four Year College	26.4%	28.0%	1.6%	6.1%
Some Four Year College	7.0%	7.8%	0.8%	11.4%
Two Year College	5.8%	6.3%	0.5%	8.5%
Some Two Year College	9.3%	9.8%	0.5%	5.0%
Some College GED	2.9%	-	-	-
High School Graduates	32.8%	35.0%	2.1%	6.5%
GEDs	3.6%	-	-	-
High School Dropouts	12.1%	13.1%	1.0%	8.4%

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 ($\mathcal{I}=\{f_C, f_N, \theta\}$)

Option Value and Its Decomposition	Unconditional	Cognitive Ability [c] $f_C<d_1^C$	Cognitive Ability [c] $f_C>d_{10}^C$	Noncognitive Ability [c] $f_N<d_1^N$	Noncognitive Ability [c] $f_N>d_{10}^N$	Both Abilities [c] $f_C<d_1^C$, $f_N<d_1^N$	Both Abilities [c] $f_C>d_{10}^C$, $f_N>d_{10}^N$
Overall Option Value Associated with College Enrollment=(1)+(2)+(3)+(4)	27,474	6,574	61,179	12,903	90,045	1,297	171,680
Decomposition:							
(1) College Enrollment at Age 19 (on time) after Graduating from HS at Age 18 (on time) [b]	16,604	395	50,036	11,683	27,151	659	77,486
(2) College Enrollment at Age 20 (delayed) after Graduating from HS at Age 18 (on time)	5,390	1,490	7,244	960	24,185	271	55,900
(3) College Enrollment at Age 20 (delayed) after Graduating from HS at Age 19 (delayed)	5,452	4,685	3,869	223	38,696	353	38,264
(4) College Enrollement at Age 24 after dropping out from HS at Age 19 and obtaining GED at age 22	28	4	30	37	13	14	30

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_j^k 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 [a]
Dynamic Schooling Model with Learning vs. without Learning

Option Value and Its Decomposition	Learning $I=\{f_C,f_N,\theta\}$	No Learning $I=\{f_C,f_N\}$	*Difference*
Overall Option Value Associated with College Enrollment=(1)+(2)+(3)+(4)	27,474	14,408	*13,066*
Decomposition:			
(1) College Enrollment at Age 19 (on time) after Graduating from HS at Age 18 (on time) [b]	16,604	8,861	*7,743*
(2) College Enrollment at Age 20 (delayed) after Graduating from HS at Age 18 (on time)	5,390	1,745	*3,645*
(3) College Enrollment at Age 20 (delayed) after Graduating from HS at Age 19 (delayed)	5,452	3,784	*1,668*
(4) College Enrollement at Age 24 after dropping out from HS at Age 19 and obtaining GED at age 22	28	18	*10*

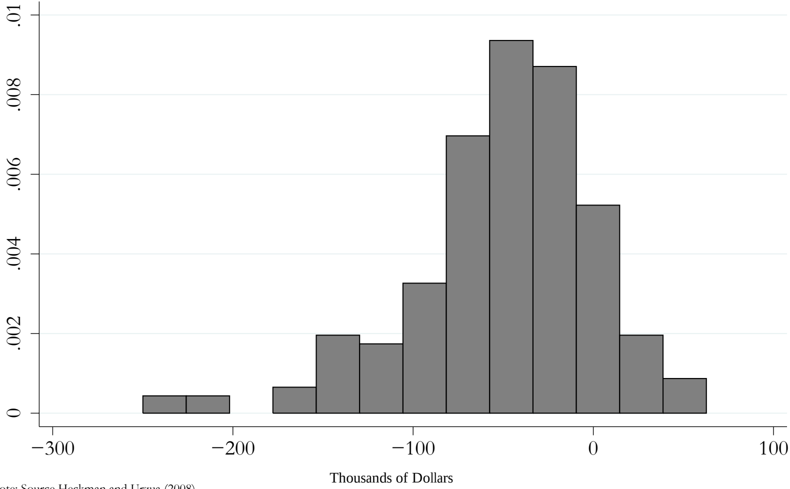
Notes: (a) All the numbers are in thousands of dollars at age 17. The sample is unchanged across simulations, that is, the numbers are computed for those agents enrolled in high school at age 17 under the three factor model.; (b) In this case the option value is generated imposing that the agent cannot 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

Initial State	Average Treatment Effect	Treament on the Treated	Treatment on the Untreated
Unconditional			
(1) High School Grad. At Age 18			
Ever	4.80%	18.27%	-12.20%
Once for All	12.60%	20.25%	2.90%
(2) High School Grad. At Age 19	-51.30%	160%	-127%
Grad. High School at Age 18			
(3) High School Grad. At Age 19	45.05%	487%	-175%
Enrolled in HS at Age 18			
Low Ability Individuals ($f_C<d_5^C$ and $f_N<d_5^N$)			
(1) High School Grad. At Age 18			
Ever	-2.92%	9.06%	-9.02%
Once for All	-2.72%	9.64%	-9.01%
(2) High School Grad. At Age 19	-111.90%	38%	-125.20%
Grad. High School at Age 18			
(3) High School Grad. At Age 19	-118.80%	36.40%	-136.20%
Enrolled in HS at Age 18			
High Ability Individuals ($f_C>d_4^C$ and $f_N>d_4^N$)			
(1) High School Grad. At Age 18			
Ever	7.14%	20.52%	-13.57%
Once for All	20.12%	23.32%	15.17%
(2) High School Grad. At Age 19	-6.65%	183.20%	-117.50%
Grad. High School at Age 18			
(3) High School Grad. At Age 19	221.10%	566.50%	-138.40%
Enrolled in HS at Age 18			

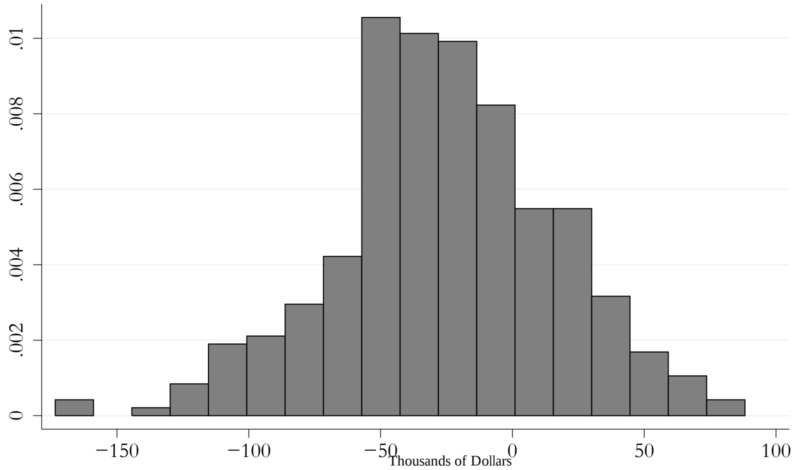
Costs

Distribution of Costs: Transition from "High School Dropout at Age 17" to "GED at Age 18"

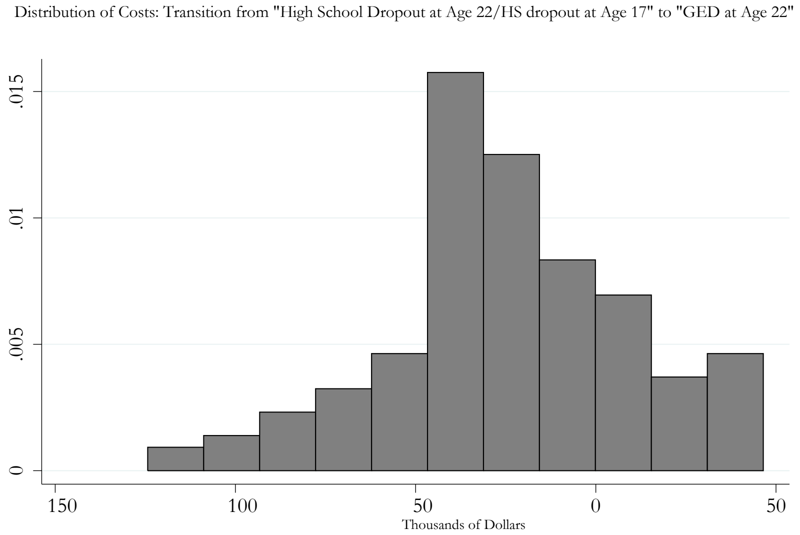


Note: Source Heckman and Urzua (2008).

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).

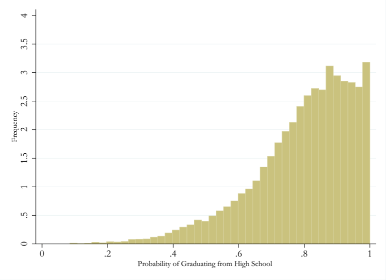


Note: Source Heckman and Urzua (2008).

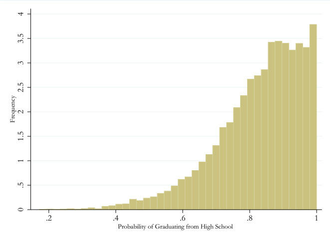
Support Conditions Satisfied?

Figure. Support Conditions for the Analysis of High School Graduation

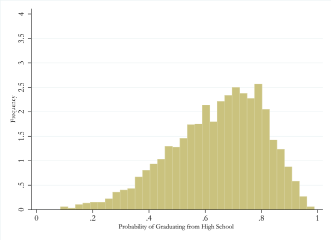
A. Overall Sample



B. Sample of High School Graduates



C. Sample of High School Dropouts



Summary

- 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.
- Builds on previous work on dynamic selection into schooling (Altonji, 1993; Keane and Wolpin, 1997, 2001; Eckstein and Wolpin, 1999; Arcidiacono, 2004; Cameron and Heckman, 1998, 2001).
- 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.