The Evolution of Labor Earnings Risk in the U.S. Economy

James Heckman
University College Dublin and University of Chicago

Flavio Cunha
University of Chicago

Econometric Policy Evaluation, Lecture III
Koopmans Memorial Lectures
Cowles Foundation
Yale University
September 29, 2006
• Lecture I presented a general framework for policy evaluation.
Lecture I presented a general framework for policy evaluation.

Lecture II focused on IV and showed the relationship between IV and structural selection models in environments with essential heterogeneity:

\[(Y_1 - Y_0) \perp \not \perp D \mid X.\]
Lecture I presented a general framework for policy evaluation.

Lecture II focused on IV and showed the relationship between IV and structural selection models in environments with essential heterogeneity:

\[(Y_1 - Y_0) \perp \!
\!
\!
\!
\!
\!
\!
\perp D \mid X.\]

It showed how to obtain marginal returns not just average returns.
Lecture I presented a general framework for policy evaluation.

Lecture II focused on IV and showed the relationship between IV and structural selection models in environments with essential heterogeneity:

\[(Y_1 - Y_0) \not\perp\not\perp D \mid X.\]

It showed how to obtain marginal returns not just average returns.

It focused on means.
• Lecture I presented a general framework for policy evaluation.

• Lecture II focused on IV and showed the relationship between IV and structural selection models in environments with essential heterogeneity:

\[(Y_1 - Y_0) \perp \!\!\!\!\perp D \mid X.\]

• It showed how to obtain marginal returns not just average returns.

• It focused on means.

• Methods for identifying means can also identify marginal distributions.

\[E [\mathbf{1}(Y_1 \leq y_1) \mid X] = F (y_1 \mid X)\]
\[E [\mathbf{1}(Y_0 \leq y_0) \mid X] = F (y_0 \mid X)\]
These lectures focused primarily on ex post analyses and did not account for uncertainty.

“Ex post calculations of capital accumulation have their place in economic and statistical history; they are useful measures for economic progress; but they are of no use to theoretical economists who are trying to find out how the system works, because they have no significance for conduct.
These lectures focused primarily on ex post analyses and did not account for uncertainty.

Today I want to consider the analysis of distributions of outcomes ex ante and ex post.

“Ex post calculations of capital accumulation have their place in economic and statistical history; they are useful measures for economic progress; but they are of no use to theoretical economists who are trying to find out how the system works, because they have no significance for conduct.
These lectures focused primarily on *ex post* analyses and did not account for uncertainty.

Today I want to consider the analysis of distributions of outcomes *ex ante* and *ex post*.

As Hicks (1946, p. 179) puts it,

“*Ex post calculations of capital accumulation have their place in economic and statistical history; they are useful measures for economic progress; but they are of no use to theoretical economists who are trying to find out how the system works, because they have no significance for conduct.*”
Two distinct problems that we consider:
Two distinct problems that we consider:

1. Identifying distributions of treatment effects.
Two distinct problems that we consider:

1. Identifying distributions of treatment effects.
2. Explicitly introducing uncertainty.
Two distinct problems that we consider:

1. Identifying distributions of treatment effects.
2. Explicitly introducing uncertainty.

First consider recovering *ex post* joint distributions.
We do not usually observe \((Y_0, Y_1)\) as a pair.
We do not usually observe \((Y_0, Y_1)\) as a pair.

The problem of recovering joint distributions from cross section data has two aspects.
The problem

- We do not usually observe \((Y_0, Y_1)\) as a pair.
- The problem of recovering joint distributions from cross section data has two aspects.
- The first is the **selection problem**.
The problem

- We do not usually observe \((Y_0, Y_1)\) as a pair.

- The problem of recovering joint distributions from cross section data has two aspects.

- The first is the **selection problem**.

- From data on outcomes, \(F_1(y_1 \mid D = 1, X)\), \(F_0(y_0 \mid D = 0, X)\), under what conditions can one recover \(F_1(y_1 \mid X)\) and \(F_0(y_0 \mid X)\), respectively?
The problem

- We do not usually observe \((Y_0, Y_1)\) as a pair.

- The problem of recovering joint distributions from cross section data has two aspects.

  - The first is the selection problem.

  - From data on outcomes, \(F_1(y_1 \mid D = 1, X)\), \(F_0(y_0 \mid D = 0, X)\), under what conditions can one recover \(F_1(y_1 \mid X)\) and \(F_0(y_0 \mid X)\), respectively?

- The second problem is the evaluation problem: how to construct the joint distribution of \(F(y_0, y_1 \mid X)\) from the two marginal distributions.
Why bother identifying joint distributions?

- Why not settle for the marginals $F(y_0 \mid X)$ and $F(y_1 \mid X)$?
Why bother identifying joint distributions?

- Why not settle for the marginals $F(y_0 \mid X)$ and $F(y_1 \mid X)$?
- These can be identified in a straightforward way using standard tools.
Why bother identifying joint distributions?

- Why not settle for the marginals $F(y_0 \mid X)$ and $F(y_1 \mid X)$?

- These can be identified in a straightforward way using standard tools.

- The literature on the measurement of economic inequality as surveyed by Foster and Sen (1997) focuses on marginal distributions across different policy states.
Why bother identifying joint distributions?

- Why not settle for the marginals $F(y_0 \mid X)$ and $F(y_1 \mid X)$?
- These can be identified in a straightforward way using standard tools.
- The literature on the measurement of economic inequality as surveyed by Foster and Sen (1997) focuses on marginal distributions across different policy states.
- Invoking the anonymity postulate, it does not keep track of individual fortunes across different policy states.
Why not settle for the marginals $F(y_0 \mid X)$ and $F(y_1 \mid X)$?

These can be identified in a straightforward way using standard tools.

The literature on the measurement of economic inequality as surveyed by Foster and Sen (1997) focuses on marginal distributions across different policy states.

Invoking the anonymity postulate, it does not keep track of individual fortunes across different policy states.

It does not consider mechanisms of assignment of treatment.
Thus, in comparing policies $p$ and $p'$, it compares the marginal distributions of

$$Y^p = D^p Y^p_1 + (1 - D^p) Y^p_0$$

and

$$Y^{p'} = D^{p'} Y^{p'}_1 + (1 - D^{p'}) Y^{p'}_0,$$

where $D^p$ and $D^{p'}$ are the treatment choice indicators under policies $p$ and $p'$, respectively.
Thus, in comparing policies $p$ and $p'$, it compares the marginal distributions of

$$Y^p = D^p Y_1^p + (1 - D^p) Y_0^p$$

and

$$Y^{p'} = D^{p'} Y_{1}^{p'} + \left(1 - D^{p'}\right) Y_{0}^{p'},$$

where $D^p$ and $D^{p'}$ are the treatment choice indicators under policies $p$ and $p'$, respectively.

It does not seek information on the subjective valuations of the policy change or the components of the treatment distributions under each policy

$$\left(Y_0^p \text{ and } Y_1^p; Y_0^{p'} \text{ and } Y_1^{p'}\right).$$
Thus, in comparing policies $p$ and $p'$, it compares the marginal distributions of

$$Y^p = D^p Y^p_1 + (1 - D^p) Y^p_0$$

and

$$Y^{p'} = D^{p'} Y^{p'}_1 + \left(1 - D^{p'}\right) Y^{p'}_0,$$

where $D^p$ and $D^{p'}$ are the treatment choice indicators under policies $p$ and $p'$, respectively.

It does not seek information on the subjective valuations of the policy change or the components of the treatment distributions under each policy ($Y^p_0$ and $Y^p_1$; $Y^{p'}_0$ and $Y^{p'}_1$).

It only compares $F(y^p | X)$ and $F(y^{p'} | X)$ in making comparisons of welfare.
Some economists appeal to classical welfare economics and classical decision theory to argue that marginal distributions of treatment outcomes are all that is required to evaluate policy.
Why bother identifying joint distributions?

- Some economists appeal to classical welfare economics and classical decision theory to argue that marginal distributions of treatment outcomes are all that is required to evaluate policy.

- The argument is that under expected utility maximization with information set $\mathcal{I}$, the agent should be assigned to (choose) treatment 1 if

$$E (\Upsilon (Y_1) - \Upsilon (Y_0) \mid \mathcal{I}) > 0,$$

where $\Upsilon$ is the preference function and $\mathcal{I}$ is the appropriate information set.
Some economists appeal to classical welfare economics and classical decision theory to argue that marginal distributions of treatment outcomes are all that is required to evaluate policy.

The argument is that under expected utility maximization with information set $\mathcal{I}$, the agent should be assigned to (choose) treatment 1 if

$$E (\Upsilon (Y_1) - \Upsilon (Y_0) \mid \mathcal{I}) > 0,$$

where $\Upsilon$ is the preference function and $\mathcal{I}$ is the appropriate information set.

For other criteria used in classical decision theory, marginal distributions are all that is required.
If one seeks to know the proportion of people who benefit from the program in gross terms ($\Pr( Y_1 > Y_0)$), one needs to know the joint distribution of $(Y_0, Y_1)$ given the appropriate information set.
If one seeks to know the proportion of people who benefit from the program in gross terms ($\Pr (Y_1 > Y_0)$), one needs to know the joint distribution of $(Y_0, Y_1)$ given the appropriate information set.

For the Roy model,

$$D = 1 [\gamma (Y_1) \geq \gamma (Y_0)].$$
Why bother identifying joint distributions?

- If one seeks to know the proportion of people who benefit from the program in gross terms ($\Pr(Y_1 > Y_0)$), one needs to know the joint distribution of $(Y_0, Y_1)$ given the appropriate information set.

- For the Roy model,

$$D = 1 \left[ \gamma(Y_1) \geq \gamma(Y_0) \right].$$

- In this case, the probability of selecting treatment given the econometrician’s information set $\mathcal{I}_E$ is

$$\Pr(D = 1 \mid \mathcal{I}_E) = \Pr(\gamma(Y_1) \geq \gamma(Y_0) \mid \mathcal{I}_E).$$
If the agent’s information set is the same as the econometrician’s and uses the choice rule
\( D = 1 [\Upsilon (Y_1) \geq \Upsilon (Y_0)] \), then observed choice proportions identify
\[ \Pr (D = 1 | I_E) = \Pr (\Upsilon (Y_1) \geq \Upsilon (Y_0) | I_E). \]
Why bother identifying joint distributions?

- If the agent’s information set is the same as the econometrician’s and uses the choice rule
  \[ D = 1 \left[ \Upsilon(Y_1) \geq \Upsilon(Y_0) \right] \], then observed choice proportions identify
  \[ \Pr(D = 1 | \mathcal{I}_E) = \Pr(\Upsilon(Y_1) \geq \Upsilon(Y_0) | \mathcal{I}_E). \]

- But analyses of objective evaluations often condition on information sets other than \( \mathcal{I}_E \).
Why bother identifying joint distributions?

- If the agent’s information set is the same as the econometrician’s and uses the choice rule $D = 1 [\Upsilon (Y_1) \geq \Upsilon (Y_0)]$, then observed choice proportions identify
  $$\Pr (D = 1 | I_E) = \Pr (\Upsilon (Y_1) \geq \Upsilon (Y_0) | I_E).$$
  
- But analyses of objective evaluations often condition on information sets other than $I_E$.

- Need the full joint distribution to compute e.g.,
  $$\Pr (Y_1 > Y_0)$$
  (the fraction who benefit ex post).
The inequalities of Hoeffding (1940) and Fréchet (1951) state that

\[
\max [F_0(y_0 | X) + F_1(y_1 | X) - 1, 0] \\
\leq F(y_0, y_1 | X) \\
\leq \min [F_0(y_0 | X), F_1(y_1 | X)].
\]
The inequalities of Hoeffding (1940) and Fréchet (1951) state that

\[
\max \left[ F_0(y_0 \mid X) + F_1(y_1 \mid X) - 1, 0 \right] \\
\leq F(y_0, y_1 \mid X) \\
\leq \min \left[ F_0(y_0 \mid X), F_1(y_1 \mid X) \right].
\]

Assume \( F_1, F_0 \) are strictly monotonic.
The inequalities of Hoeffding (1940) and Fréchet (1951) state that

\[
\max \ [F_0 (y_0 \mid X) + F_1 (y_1 \mid X) - 1, 0] \\
\leq F (y_0, y_1 \mid X) \\
\leq \min \ [F_0 (y_0 \mid X), F_1 (y_1 \mid X)].
\]

- Assume \( F_1, F_0 \) are strictly monotonic.
- Upper bound \( Y_1 = F_1^{-1} (F_0 (Y_0)) \).
The inequalities of Hoeffding (1940) and Fréchet (1951) state that

\[
\max \left[ F_0(y_0 | X) + F_1(y_1 | X) - 1, 0 \right] \\
\leq F(y_0, y_1 | X) \\
\leq \min \left[ F_0(y_0 | X), F_1(y_1 | X) \right].
\]

- Assume \( F_1, F_0 \) are strictly monotonic.
- Upper bound \( Y_1 = F_1^{-1}(F_0(Y_0)) \).
- Lower bound \( Y_1 = F_1^{-1}(1 - F_0(Y_0)) \).

A discrete outcome example.

A discrete outcome example.

Let $(E, E)$ denote the event “employed with treatment” and “employed without treatment” and let $(E, N)$ be the event “employed with treatment, not employed without treatment”.

This model for outcomes can be written in the form of a contingency table:

<table>
<thead>
<tr>
<th></th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E$</td>
<td>$N$</td>
</tr>
<tr>
<td>$E$</td>
<td>$P_{EE}$</td>
<td>$P_{EN}$</td>
</tr>
<tr>
<td>$N$</td>
<td>$P_{NE}$</td>
<td>$P_{NN}$</td>
</tr>
<tr>
<td></td>
<td>$P_{.E}$</td>
<td>$P_{.N}$</td>
</tr>
</tbody>
</table>
This model for outcomes can be written in the form of a contingency table:

<table>
<thead>
<tr>
<th></th>
<th>Untreated</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E$</td>
<td>$N$</td>
</tr>
<tr>
<td></td>
<td>$E$</td>
<td>$P_{EE}$</td>
</tr>
<tr>
<td>$N$</td>
<td>$P_{NE}$</td>
<td>$P_{NN}$</td>
</tr>
<tr>
<td></td>
<td>$P_E.$</td>
<td>$P_N.$</td>
</tr>
</tbody>
</table>

$P_E.$, $P_N.$, $P_E.$, $P_N.$ obtained from experiment.
Bounds from classical probability inequalities

- Estimates of the marginals of the table parameters:

\[ P_E = P_{EE} + P_{EN} \]

(employment proportion among the treated)

\[ P_E = P_{EE} + P_{NE} \]

(employment proportion among the untreated)
Estimates of the marginals of the table parameters:

\[ P_{E.} = P_{EE} + P_{EN} \]

(employment proportion among the treated)

\[ P_{.E} = P_{EE} + P_{NE} \]

(employment proportion among the untreated)

The treatment effect is usually defined as

\[ \Delta = P_{EN} - P_{NE} \quad \text{net effect} \quad (1) \]

\[ = P_{E.} - P_{.E} \quad (2) \]
Fréchet-Hoeffding Bounds

\[
\max [P_{E.} + P_{.E} - 1, 0] \leq P_{EE} \leq \min [P_{E.}, P_{.E}]
\]

\[
\max [P_{E.} - P_{.E}, 0] \leq P_{EN} \leq \min [P_{E.}, 1 - P_{.E}]
\]

\[
\max [-P_{E.} + P_{.E}, 0] \leq P_{NE} \leq \min [1 - P_{E.}, P_{.E}]
\]

\[
\max [1 - P_{E.} - P_{.E}, 0] \leq P_{NN} \leq \min [1 - P_{E.}, 1 - P_{.E}]
\]
Fraction Employed in the 16th, 17th or 18th Month after Random Assignment and Fréchet-Hoeffding Bounds on the Probabilities $P_{NE}$ and $P_{EN}$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction employed in the treatment group</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>Fraction employed in the control group</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>Bounds on $P_{EN}$</td>
<td>[0.03, 0.39]</td>
</tr>
<tr>
<td></td>
<td>(0.01), (0.01)</td>
</tr>
<tr>
<td>Bounds on $P_{NE}$</td>
<td>[0.00, 0.36]</td>
</tr>
<tr>
<td></td>
<td>(0.00), (0.01)</td>
</tr>
</tbody>
</table>

Notes: 1. Employment percentages are based on self-reported employment in months 16, 17 and 18 after random assignment. A person is coded as employed if the sum of their self-reported earnings over these three months is positive. 2. $P_{ij}$ is the probability of having employment status $i$ in the treated state and employment status $j$ in the untreated state, where $i$ and $j$ take on the values $E$ for employed and $N$ for not employed. The Fréchet-Hoeffding bounds are given in the text. 3. Standard errors are discussed in Heckman, Smith and Clements (1997). Source: Heckman, Smith and Clements (1997).
Requires access to variables $Q$ that have the property that conditional on $Q$, $F(y_0 \mid D = 0, X, Q) = F(y_0 \mid X, Q)$ and $F(y_1 \mid D = 1, X, Q) = F(y_1 \mid X, Q)$. 

Matching assumes that conditional on observed variables, $Q$, there is no selection problem. (In linear equation, OLS is matching) 

$Y_0 \perp \perp D \mid Q$ and $Y_1 \perp \perp D \mid Q$.

Identify the joint distribution $F(y_0, y_1 \mid X) = \int F_0(y_0 \mid X, q) F_1(y_1 \mid X, q) d\mu(q \mid X)$. 

Bounds from classical probability inequalities
Bounds from classical probability inequalities

- Requires access to variables $Q$ that have the property that conditional on $Q$, $F(y_0 \mid D = 0, X, Q) = F(y_0 \mid X, Q)$ and $F(y_1 \mid D = 1, X, Q) = F(y_1 \mid X, Q)$.

- Matching assumes that conditional on observed variables, $Q$, there is no selection problem. (In linear equation, OLS is matching)
Bounds from classical probability inequalities

- Requires access to variables $Q$ that have the property that conditional on $Q$, $F(y_0 \mid D = 0, X, Q) = F(y_0 \mid X, Q)$ and $F(y_1 \mid D = 1, X, Q) = F(y_1 \mid X, Q)$.

- Matching assumes that conditional on observed variables, $Q$, there is no selection problem. (In linear equation, OLS is matching)

- $(Y_0 \perp\!\!\!\!\perp D \mid Q)$ and $(Y_1 \perp\!\!\!\!\perp D \mid Q)$.
Requires access to variables $Q$ that have the property that conditional on $Q$, $F(y_0 \mid D = 0, X, Q) = F(y_0 \mid X, Q)$ and $F(y_1 \mid D = 1, X, Q) = F(y_1 \mid X, Q)$.

Matching assumes that conditional on observed variables, $Q$, there is no selection problem. (In linear equation, OLS is matching)

$(Y_0 \perp \perp D \mid Q)$ and $(Y_1 \perp \perp D \mid Q)$.

Identify the joint distribution

$$F (y_0, y_1 \mid X) = \int F_0 (y_0 \mid X, q) F_1 (y_1 \mid X, q) \, d\mu (q \mid X).$$
Conditional on $X$, $Y_0$ and $Y_1$ are assumed to be deterministically related:

$$Y_1 - Y_0 = \Delta$$  \hspace{1cm} (3)

where $\Delta$ is a constant given $X$. 

Bounds from classical probability inequalities
Conditional on \( X \), \( Y_0 \) and \( Y_1 \) are assumed to be deterministically related:

\[
Y_1 - Y_0 = \Delta
\]  

where \( \Delta \) is a constant given \( X \).

\[
F_1(Y_1) = F_0(Y_0 + \Delta).
\]
Conditional on $X$, $Y_0$ and $Y_1$ are assumed to be deterministically related:

$$Y_1 - Y_0 = \Delta$$

where $\Delta$ is a constant given $X$.

$F_1(Y_1) = F_0(Y_0 + \Delta)$.

Can generalize using Fréchet upper and lower bounds.
Conditional on $X$, $Y_0$ and $Y_1$ are assumed to be deterministically related:

$$Y_1 - Y_0 = \Delta$$

(3)

where $\Delta$ is a constant given $X$.

- $F_1(Y_1) = F_0(Y_0 + \Delta)$.

- Can generalize using Fréchet upper and lower bounds.

- Enforce perfect ranking but not equality of differences across all quantiles.
Markov kernels $M(y_1, y_0 \mid X)$ and $\tilde{M}(y_0, y_1 \mid X)$ that map marginals into marginals:

$$F_1(y_1 \mid X) = \int M(y_1, y_0 \mid X) dF_0(y_0 \mid X),$$

$$F_0(y_0 \mid X) = \int \tilde{M}(y_0, y_1 \mid X) dF_1(y_1 \mid X).$$
Bounds from classical probability inequalities

\[ Y_1 = \mu_1(X) + U_1 \quad E(U_1 \mid X) = 0 \]
\[ Y_0 = \mu_0(X) + U_0 \quad E(U_0 \mid X) = 0 \]

\[ Y = \mu_0(X) + (\mu_1(X) - \mu_0(X) + U_1 - U_0) D + U_0 \quad (4) \]

\[ \beta(X) = \bar{\beta}(X) + \eta \]
\[ \eta = U_1 - U_0 \]
$D = 1 [Y_1 \geq Y_0]$  \hspace{1cm} (5)

$Y_1 = \mu_1(X) + U_1 \hspace{1cm} E(U_1 \mid X) = 0$

$Y_0 = \mu_0(X) + U_0 \hspace{1cm} E(U_0 \mid X) = 0$

- Can identify $F(y_0, y_1)$ under Roy assumption and some variation in the $X$. (Heckman and Honoré, 1990)
Using additional information


- Assume separability:

\[
Y_1 = \mu_1(X) + U_1 \\
Y_0 = \mu_0(X) + U_0.
\]
Using additional information


- Assume separability:

  \[
  Y_1 = \mu_1(X) + U_1 \\
  Y_0 = \mu_0(X) + U_0.
  \]

- Let \( I \) denote the latent variable generating schooling choices:

  \[
  I = \mu_I(Z) + U_I \\
  D = 1[I \geq 0].
  \]

Assume separability:

\[ Y_1 = \mu_1(X) + U_1 \]
\[ Y_0 = \mu_0(X) + U_0. \]

Let \( I \) denote the latent variable generating schooling choices:

\[ I = \mu_I(Z) + U_I \]
\[ D = 1[I \geq 0]. \]

Normality assumptions make it easy to understand how the method works and can be relaxed.
Restrict the dimension of the unobservables.
Using additional information

- Restrict the dimension of the unobservables.

- If we have many measurements relative to the dimensionality of the latent variables, we get identification of the joint distribution.
Restrict the dimension of the unobservables.

If we have many measurements relative to the dimensionality of the latent variables, we get identification of the joint distribution.

Assume a one factor model where $\theta$ is the factor that generates dependence across the unobservables:

\begin{align*}
U_0 &= \alpha_0 \theta + \varepsilon_0 \\
U_1 &= \alpha_1 \theta + \varepsilon_1 \\
U_I &= \alpha_{U_I} \theta + \varepsilon_{U_I}
\end{align*}

\[ \theta \perp \perp (\varepsilon_0, \varepsilon_1, \varepsilon_{U_I}), \quad \varepsilon_0 \perp \perp \varepsilon_1 \perp \perp \varepsilon_{U_I}. \]
• Restrict the dimension of the unobservables.

• If we have many measurements relative to the dimensionality of the latent variables, we get identification of the joint distribution.

• Assume a one factor model where $\theta$ is the factor that generates dependence across the unobservables:

\[
\begin{align*}
U_0 &= \alpha_0 \theta + \varepsilon_0 \\
U_1 &= \alpha_1 \theta + \varepsilon_1 \\
U_I &= \alpha_{U_I} \theta + \varepsilon_{U_I}
\end{align*}
\]

\[\theta \perp \perp (\varepsilon_0, \varepsilon_1, \varepsilon_{U_I}), \quad \varepsilon_0 \perp \perp \varepsilon_1 \perp \perp \varepsilon_{U_I}.\]

• To set the scale of the unobserved factor, we normalize one “loading” (coefficient on $\theta$) to 1.
Assume that $E(U_0) = 0$, $E(U_1) = 0$ and $E(U_I) = 0$ and $E(\theta) = 0$. 
Using additional information

- Assume that $E(U_0) = 0$, $E(U_1) = 0$ and $E(U_I) = 0$ and $E(\theta) = 0$.

- From standard analysis of censored models, we can recover the distribution of $\left(U_0, \frac{U_I}{\sigma_{U_I}}\right)$ and $\left(U_1, \frac{U_I}{\sigma_{U_I}}\right)$ (Heckman, 1990)
Using additional information

- Assume that $E(U_0) = 0$, $E(U_1) = 0$ and $E(U_I) = 0$ and $E(\theta) = 0$.

- From standard analysis of censored models, we can recover the distribution of $(U_0, \frac{U_I}{\sigma_U})$ and $(U_1, \frac{U_I}{\sigma_U})$ (Heckman, 1990)

- From the joint distributions of $(U_0, \frac{U_I}{\sigma_U})$ and $(U_1, \frac{U_I}{\sigma_U})$ we can identify

  $$\text{Cov} \left( U_0, \frac{U_I}{\sigma_U} \right) = \frac{\alpha_0 \alpha U_I}{\sigma_U} \sigma_{\theta}^2$$

  $$\text{Cov} \left( U_1, \frac{U_I}{\sigma_U} \right) = \frac{\alpha_1 \alpha U_I}{\sigma_U} \sigma_{\theta}^2$$

where $\sigma_{U_I}^2 = \text{Var}(\varepsilon_{U_I})$. 
We obtain the sign of the dependence between $U_0$, $U_1$ because
\[
\text{Cov}(U_0, U_1) = \alpha_0 \alpha_1 \sigma^2. 
\]
We obtain the sign of the dependence between $U_0, U_1$ because

$$\text{Cov}(U_0, U_1) = \alpha_0 \alpha_1 \sigma_\theta^2.$$ 

Can’t identify other parameters without further assumptions.
Using additional information

- We obtain the sign of the dependence between $U_0, U_1$ because

  $$\text{Cov}(U_0, U_1) = \alpha_0 \alpha_1 \sigma^2_\theta.$$ 

- Can’t identify other parameters without further assumptions.

- With additional information, we can identify the full joint distribution of $(U_0, U_1, U_I)$. 
Suppose we have additional “measurements” (e.g., a test score; labor supply; outcomes generated by $\theta$)
Example 1. Access to a single proxy measure (e.g., a test score)

\[ M = \mu_M(X) + U_M \]

where

\[ U_M = \alpha_M \theta + \varepsilon_M \]

so

\[ M = \mu_M(X) + \alpha_M \theta + \varepsilon_M \]

where \( \varepsilon_M \) is independent of \( \varepsilon_0, \varepsilon_1, \varepsilon_U, \) and \( \theta, \) as well as \( (X, Z) \).

\[
\begin{align*}
\text{Cov} (Y_1, M) &= \alpha_1 \alpha_M \sigma_\theta^2 \\
\text{Cov} (Y_0, M) &= \alpha_0 \alpha_M \sigma_\theta^2 \\
\text{Cov} (I, M) &= \frac{\alpha_U I}{\sigma_U} \alpha_M \sigma_\theta^2
\end{align*}
\]
Example 1. Access to a single proxy measure (e.g., a test score)

- Normalize the loading on the proxy (or test score) to one ($\alpha_M = 1$).

$$\frac{\text{Cov} \left( Y_1, I \right)}{\text{Cov} \left( I, M \right)} = \frac{\alpha_1 \alpha_U \sigma^2_{\theta}}{\alpha_U \alpha_M \sigma^2_{\theta}} = \alpha_1$$

because $\alpha_M = 1$.

$$\frac{\text{Cov} \left( Y_1, I \right)}{\text{Cov} \left( Y_0, I \right)} = \frac{\alpha_1 \alpha_U \sigma^2_{\theta}}{\alpha_0 \alpha_U \sigma^2_{\theta}} = \frac{\alpha_1}{\alpha_0}$$
Using additional information

Example 1. Access to a single proxy measure (e.g., a test score)

- Normalize the loading on the proxy (or test score) to one ($\alpha_M = 1$).

$$\frac{\text{Cov} (Y_1, I)}{\text{Cov} (I, M)} = \frac{\alpha_1 \alpha_u \sigma_\theta^2}{\alpha_u \alpha_M \sigma_\theta^2} = \alpha_1$$

because $\alpha_M = 1$.

$$\frac{\text{Cov} (Y_1, I)}{\text{Cov} (Y_0, I)} = \frac{\alpha_1 \alpha_u \sigma_\theta^2}{\alpha_0 \alpha_u \sigma_\theta^2} = \frac{\alpha_1}{\alpha_0}$$

- We obtain $\sigma_\theta^2$ from $\text{Cov} (Y_1, M)$ or $\text{Cov} (Y_0, M)$. 
Using additional information

Example 1. Access to a single proxy measure (e.g., a test score)

- Normalize the loading on the proxy (or test score) to one ($\alpha_M = 1$).

\[
\frac{\text{Cov} (Y_1, I)}{\text{Cov} (I, M)} = \frac{\alpha_1 \alpha_U \sigma^2_\theta}{\alpha_U \alpha_M \sigma^2_\theta} = \alpha_1
\]

because $\alpha_M = 1$.

\[
\frac{\text{Cov} (Y_1, I)}{\text{Cov} (Y_0, I)} = \frac{\alpha_1 \alpha_U \sigma^2_\theta}{\alpha_0 \alpha_U \sigma^2_\theta} = \frac{\alpha_1}{\alpha_0}
\]

- We obtain $\sigma^2_\theta$ from Cov ($Y_1, M$) or Cov ($Y_0, M$).

- We obtain $\alpha_U$ (up to scale $\sigma_U$) from

\[
\text{Cov} (I, M) = \alpha_U \alpha_M \sigma^2_\theta
\]

since we know $\alpha_M (= 1)$ and $\sigma^2_\theta$. 
Example 1. Access to a single proxy measure (e.g., a test score)

- The model is overidentified.
Using additional information

Example 1. Access to a single proxy measure (e.g., a test score)

- The model is overidentified.
- We write out the decision rule in terms of costs, we can characterize the latent variable determining choices as:

\[ I = Y_1 - Y_0 - C \]

where \( C = \mu_C(Z) + U_C \) and \( U_C = \alpha_C \theta + \varepsilon_C \), and \( \varepsilon_C \) is independent of \( \theta \).
Example 1. Access to a single proxy measure (e.g., a test score)

- The model is overidentified.
- We write out the decision rule in terms of costs, we can characterize the latent variable determining choices as:

\[ I = Y_1 - Y_0 - C \]

where \( C = \mu_C(Z) + U_C \) and \( U_C = \alpha_C \theta + \varepsilon_C \), and \( \varepsilon_C \) is independent of \( \theta \).

- \( U_I = U_1 - U_0 - U_C \), and

\[
\begin{align*}
\alpha_{U_I} &= \alpha_1 - \alpha_0 - \alpha_C \\
\varepsilon_{U_I} &= \varepsilon_1 - \varepsilon_0 - \varepsilon_C \\
\text{Var} \left( \varepsilon_{U_I} \right) &= \text{Var} \left( \varepsilon_1 \right) + \text{Var} \left( \varepsilon_0 \right) + \text{Var} \left( \varepsilon_C \right).
\end{align*}
\]
Using additional information

Example 1. Access to a single proxy measure (e.g., a test score)

- The scale $\sigma_{U_l}$ is identified if there are variables in $X$ but not in $Z$.

$$\text{Var}(M) - \alpha^2_M \sigma^2_{\theta} = \sigma^2_{\varepsilon_M}.$$
Example 1. Access to a single proxy measure (e.g., a test score)

- The scale $\sigma_{U_l}$ is identified if there are variables in $X$ but not in $Z$.

$$\text{Var}(M) - \alpha_M^2 \sigma_\theta^2 = \sigma_{\varepsilon_M}^2.$$ 

- We can thus construct the joint distribution of $(Y_0, Y_1, C)$ and hence the joint distribution of $(Y_0, Y_1)$. 
Example 1. Access to a single proxy measure (e.g., a test score)

- The scale $\sigma_{U_i}$ is identified if there are variables in $X$ but not in $Z$.

$$\text{Var}(M) - \alpha_M^2 \theta = \sigma_{\varepsilon M}^2.$$ 

- We can thus construct the joint distribution of $(Y_0, Y_1, C)$ and hence the joint distribution of $(Y_0, Y_1)$.

- We have assumed normality because it is convenient to do so. Carneiro, Hansen, and Heckman (2003); Cunha and Heckman (2006b); Cunha, Heckman, and Navarro (2005,2006); and Cunha, Heckman, and Schennach (2006a,b) relax this assumption.
Using additional information

Example 2. identification without choice data

- Let $I$ be any indicator that depends on $\theta$ and assume that it is observed.
Using additional information

Example 2. identification without choice data

- Let \( I \) be any indicator that depends on \( \theta \) and assume that it is observed.

- By limit operations (\( P(X, Z) \rightarrow 0 \) or \( P(X, Z) \rightarrow 1 \) along certain sequences in its support) or some randomization we observe triplets \((Y_0, M, I), (Y_1, M, I)\).
Example 2. identification without choice data

- Let $I$ be any indicator that depends on $\theta$ and assume that it is observed.
- By limit operations ($P(X, Z) \to 0$ or $P(X, Z) \to 1$ along certain sequences in its support) or some randomization we observe triplets $(Y_0, M, I)$, $(Y_1, M, I)$.
- Not $Y_0$ and $Y_1$ together.
Example 2. identification without choice data

- Let \( I \) be any indicator that depends on \( \theta \) and assume that it is observed.

- By limit operations (\( P(X, Z) \rightarrow 0 \) or \( P(X, Z) \rightarrow 1 \) along certain sequences in its support) or some randomization we observe triplets \((Y_0, M, I), (Y_1, M, I)\).

- Not \( Y_0 \) and \( Y_1 \) together.

- We can identify all of the variances and covariances of the factor model as well as the factor loadings up to one normalization.
Example 2. identification without choice data

- Let $I$ be any indicator that depends on $\theta$ and assume that it is observed.
- By limit operations ($P(X, Z) \to 0$ or $P(X, Z) \to 1$ along certain sequences in its support) or some randomization we observe triplets $(Y_0, M, I)$, $(Y_1, M, I)$.
- Not $Y_0$ and $Y_1$ together.
- We can identify all of the variances and covariances of the factor model as well as the factor loadings up to one normalization.
- We can identify the joint distribution of $(Y_0, Y_1)$. 

\[ \text{\[ \begin{align*} \sigma_{UI} & \text{ rather than normalizing it to one.} \end{align*} \] \] \]
Using additional information

Example 2. identification without choice data

Let $I$ be any indicator that depends on $\theta$ and assume that it is observed.

By limit operations ($P(X, Z) \to 0$ or $P(X, Z) \to 1$ along certain sequences in its support) or some randomization we observe triplets $(Y_0, M, I), (Y_1, M, I)$.

Not $Y_0$ and $Y_1$ together.

We can identify all of the variances and covariances of the factor model as well as the factor loadings up to one normalization.

We can identify the joint distribution of $(Y_0, Y_1)$.

We can identify $\sigma_{U_I}$ rather than normalizing it to one.
Example 3. Two (or more) periods of panel data on outcomes

For each person we have two periods of outcome data in one counterfactual state or the other.
Using additional information

Example 3. Two (or more) periods of panel data on outcomes

- For each person we have two periods of outcome data in one counterfactual state or the other.
- We observe choices and associated outcomes
  \[ Y_{jt} = \mu_{jt}(X) + U_{jt}, \ j = 0, 1, \ t = 1, 2. \]
Example 3. Two (or more) periods of panel data on outcomes

- For each person we have two periods of outcome data in one counterfactual state or the other.
- We observe choices and associated outcomes
  \[ Y_{jt} = \mu_{jt}(X) + U_{jt}, \ j = 0, 1, \ t = 1, 2. \]
- We write
  \[ U_{1t} = \alpha_{1t} \theta + \varepsilon_{1t} \quad \text{and} \quad U_{0t} = \alpha_{0t} \theta + \varepsilon_{0t} \]
  to obtain
  \[ Y_{1t} = \mu_{1t}(X) + \alpha_{1t} \theta + \varepsilon_{1t} \quad t = 1, 2 \]
  \[ Y_{0t} = \mu_{0t}(X) + \alpha_{0t} \theta + \varepsilon_{0t} \quad t = 1, 2. \]
Example 3. Two (or more) periods of panel data on outcomes

\[ I = (Y_{12} + Y_{11}) - (Y_{02} + Y_{01}) - C \]

\[ D = 1[I \geq 0], \]

\[ I = \mu_{11}(X) + \mu_{12}(X) - \mu_{01}(X) - \mu_{02}(X) - \mu_C(Z) \]

\[ + U_{11} + U_{12} - U_{01} - U_{02} - U_C. \]

\[ \]

- From normality, we can recover the joint distributions of \((I, Y_{11}, Y_{12})\) and \((I, Y_{01}, Y_{02})\) but not directly the joint distribution of \((I, Y_{11}, Y_{12}, Y_{01}, Y_{02})\).
Using additional information

Example 3. Two (or more) periods of panel data on outcomes

- Thus, conditioning on $X$ and $Z$ we can recover the joint distribution of $(U_I, U_{01}, U_{02})$ and $(U_I, U_{11}, U_{12})$ but apparently not that of $(U_I, U_{01}, U_{02}, U_{11}, U_{12})$. 
Using additional information

Example 3. Two (or more) periods of panel data on outcomes

- Thus, conditioning on $X$ and $Z$ we can recover the joint distribution of $(U_I, U_{01}, U_{02})$ and $(U_I, U_{11}, U_{12})$ but apparently not that of $(U_I, U_{01}, U_{02}, U_{11}, U_{12})$.

- From the available data, we can identify the following covariances:

\[
\text{Cov}(U_I, U_{12}) = (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{12}\sigma_\theta^2
\]
\[
\text{Cov}(U_I, U_{11}) = (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{11}\sigma_\theta^2
\]
\[
\text{Cov}(U_I, U_{01}) = (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{01}\sigma_\theta^2
\]
\[
\text{Cov}(U_I, U_{02}) = (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{02}\sigma_\theta^2
\]
\[
\text{Cov}(U_{11}, U_{12}) = \alpha_{11}\alpha_{12}\sigma_\theta^2
\]
\[
\text{Cov}(U_{01}, U_{02}) = \alpha_{01}\alpha_{02}\sigma_\theta^2.
\]
Example 3. Two (or more) periods of panel data on outcomes

- **Normalize** $\alpha_{01} = 1$. Then,

  \[
  \frac{\text{Cov}(U_I, U_{12})}{\text{Cov}(U_I, U_{01})} = \alpha_{12}, \quad \frac{\text{Cov}(U_I, U_{11})}{\text{Cov}(U_I, U_{01})} = \alpha_{11},
  \]

  \[
  \frac{\text{Cov}(U_I, U_{02})}{\text{Cov}(U_I, U_{01})} = \alpha_{02}.
  \]

  \[
  \frac{\alpha_{11} \alpha_{12}}{\alpha_{11} \alpha_{12}} = \sigma^2_{\theta}
  \]

  and

  \[
  \frac{\text{Cov}(U_{01}, U_{02})}{\alpha_{01} \alpha_{02}} = \sigma^2_{\theta}.
  \]
Example 3. Two (or more) periods of panel data on outcomes

- We can recover $\sigma_\theta^2$ (since we know $\alpha_{11}\alpha_{12}$ and $\alpha_{01}\alpha_{02}$) from $\text{Cov}(U_{11}, U_{12})$ and $\text{Cov}(U_{01}, U_{02})$. 

Using additional information
Example 3. Two (or more) periods of panel data on outcomes

- We can recover $\sigma_\theta^2$ (since we know $\alpha_{11}\alpha_{12}$ and $\alpha_{01}\alpha_{02}$) from $\text{Cov}(U_{11}, U_{12})$ and $\text{Cov}(U_{01}, U_{02})$.

- We can also recover $\alpha_C$ since we know $\sigma_\theta^2$, $\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C$, and $\alpha_{11}, \alpha_{12}, \alpha_{01}, \alpha_{02}$. 

Using additional information
Example 3. Two (or more) periods of panel data on outcomes

- We can recover $\sigma_\theta^2$ (since we know $\alpha_{11}\alpha_{12}$ and $\alpha_{01}\alpha_{02}$) from $\text{Cov}(U_{11}, U_{12})$ and $\text{Cov}(U_{01}, U_{02})$.

- We can also recover $\alpha_C$ since we know $\sigma_\theta^2$, $\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C$, and $\alpha_{11}, \alpha_{12}, \alpha_{01}, \alpha_{02}$.

- We can form (conditional on $X$)
  \[
  \text{Cov}(Y_{11}, Y_{01}) = \alpha_{11}\alpha_{01}\sigma_\theta^2; \quad \text{Cov}(Y_{12}, Y_{01}) = \alpha_{12}\alpha_{01}\sigma_\theta^2; \\
  \text{Cov}(Y_{11}, Y_{02}) = \alpha_{11}\alpha_{02}\sigma_\theta^2 \text{ and } \text{Cov}(Y_{12}, Y_{02}) = \alpha_{12}\alpha_{02}\sigma_\theta^2.
  \]
Example 3. Two (or more) periods of panel data on outcomes

- We can recover $\sigma^2_\theta$ (since we know $\alpha_{11}\alpha_{12}$ and $\alpha_{01}\alpha_{02}$) from $\text{Cov}(U_{11}, U_{12})$ and $\text{Cov}(U_{01}, U_{02})$.

- We can also recover $\alpha_C$ since we know $\sigma^2_\theta$, 
  $\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C$, and $\alpha_{11}, \alpha_{12}, \alpha_{01}, \alpha_{02}$.

- We can form (conditional on $X$) 
  $\text{Cov}(Y_{11}, Y_{01}) = \alpha_{11}\alpha_{01}\sigma^2_\theta$; 
  $\text{Cov}(Y_{12}, Y_{01}) = \alpha_{12}\alpha_{01}\sigma^2_\theta$; 
  $\text{Cov}(Y_{11}, Y_{02}) = \alpha_{11}\alpha_{02}\sigma^2_\theta$ and 
  $\text{Cov}(Y_{12}, Y_{02}) = \alpha_{12}\alpha_{02}\sigma^2_\theta$.

- Thus we can identify the joint distribution of 
  $(Y_{01}, Y_{02}, Y_{11}, Y_{12}, C)$ since we can identify $\mu_C(Z)$ from 
  the schooling choice equation since we know
  $\mu_{01}(X), \mu_{02}(X), \mu_{11}(X), \text{ and } \mu_{12}(X)$.
If the analyst knows $\theta$ and can condition on it, we obtain the conditional independence assumption of matching, $(M-1)$:

$$(Y_0, Y_1) \independent D \mid X, Z, \theta.$$
If the analyst knows $\theta$ and can condition on it, we obtain the conditional independence assumption of matching, (M-1):

$$(Y_0, Y_1) \perp \perp D \mid X, Z, \theta.$$ 

Aakvik, Heckman, and Vytlacil (2005) proxy for $\theta$ and identify the distribution of $\theta$ under the following assumption:

$$\theta \perp \perp X, Z.$$ 

Thus the factor approach is a version of matching on unobservables.
If the analyst knows $\theta$ and can condition on it, we obtain the conditional independence assumption of matching, (M-1):

$$(Y_0, Y_1) \perp \perp D \mid X, Z, \theta.$$ 

Aakvik, Heckman, and Vytlacil (2005) proxy for $\theta$ and identify the distribution of $\theta$ under the following assumption:

$$\theta \perp \perp X, Z.$$ 

Thus the factor approach is a version of matching on unobservables.

We do not need normality (Kotlarski’s Theorem).
Theorem

\( T_1 = \theta + \varepsilon_1 \)

and

\( T_2 = \theta + \varepsilon_2 \)

and \( \theta \perp \perp \varepsilon_1 \perp \perp \varepsilon_2 \), the means of all three generating random variables are finite and are normalized to \( E(\varepsilon_1) = E(\varepsilon_2) = 0 \), and the conditions of Fubini’s theorem are satisfied for each random variable, and the random variables possess nonvanishing (a.e.) characteristic functions, then the densities of \((\theta, \varepsilon_1, \varepsilon_2), g(\theta), g_1(\varepsilon_1), g_2(\varepsilon_2), \) respectively, are identified.
Applied to our context, consider the first two equations of a vector of indicators $M$. 

\[ M_1 = \lambda_1 \theta + \epsilon_1, \]
\[ M_2 = \lambda_2 \theta + \epsilon_2, \]

where $\lambda_2 \neq 0$. 

Applying Kotlarski’s Theorem, we can nonparametrically identify the densities 

\[ g_{\theta}(\theta), \quad g_{\epsilon_1}(\epsilon_1), \quad g_{\epsilon^*_2}(\epsilon^*_2). \]
Applied to our context, consider the first two equations of a vector of indicators \( M \).

We write

\[
\begin{align*}
M_1 & = \lambda_1 \theta + \varepsilon_1, \text{ where } \lambda_1 = 1 \\
M_2 & = \lambda_2 \theta + \varepsilon_2, \text{ where } \lambda_2 \neq 0
\end{align*}
\]

\[
\begin{align*}
\frac{M_1}{\lambda_2} & = \theta + \varepsilon^*_2 \\
\frac{M_2}{\lambda_2} & = \theta + \varepsilon^*_2
\end{align*}
\]

where \( \varepsilon^*_2 = \varepsilon_2 / \lambda_2 \).
Nonparametric extensions

- Applied to our context, consider the first two equations of a vector of indicators $M$.

- We write

$$M_1 = \lambda_1 \theta + \varepsilon_1, \text{ where } \lambda_1 = 1$$
$$M_2 = \lambda_2 \theta + \varepsilon_2, \text{ where } \lambda_2 \neq 0$$

$$M_1 = \theta + \varepsilon_1$$
$$\frac{M_2}{\lambda_2} = \theta + \varepsilon^*_2$$

where $\varepsilon^*_2 = \varepsilon_2 / \lambda_2$.

- Applying Kotlarski’s Theorem, we can nonparametrically identify the densities $g_\theta(\theta)$, $g_1(\varepsilon_1)$, and $g_2(\varepsilon^*_2)$. 
Thus far we have ignored uncertainty which is an essential feature of a modern economy.
Accounting for uncertainty

- Thus far we have ignored uncertainty which is an essential feature of a modern economy.
- For the rest of this talk, we focus on a specific problem.

According to Levy and Murnane (1992):

- Earnings inequality was stable in the 1970s but increased rapidly over the 1980s.
- Inequality between age-education groups was stable in the 1970s and rose sharply in the 1980s.
- Inequality within age-education groups has grown steadily since the 1970s.
- This trend stopped in the mid 1990s.
Accounting for uncertainty

- Thus far we have ignored uncertainty which is an essential feature of a modern economy.
- For the rest of this talk, we focus on a specific problem.
- To understand the evolution of inequality and uncertainty in labor earnings for the U.S. economy.

According to Levy and Murnane (1992):
- Earnings inequality was stable in the 1970s but increased rapidly over the 1980s.
- Inequality between age-education groups was stable in the 1970s and rose sharply in the 1980s.
- Inequality within age-education groups has grown steadily since the 1970s.
- This trend stopped in the mid 1990s.
Accounting for uncertainty

- Thus far we have ignored uncertainty which is an essential feature of a modern economy.

- For the rest of this talk, we focus on a specific problem.

- To understand the evolution of inequality and uncertainty in labor earnings for the U.S. economy.

- According to Levy and Murnane (1992):
Accounting for uncertainty

- Thus far we have ignored uncertainty which is an essential feature of a modern economy.
- For the rest of this talk, we focus on a specific problem.
- To understand the evolution of inequality and uncertainty in labor earnings for the U.S. economy.
- According to Levy and Murnane (1992):
  - Earnings inequality was stable in the 1970s but increased rapidly over the 1980s.
Accounting for uncertainty

- Thus far we have ignored uncertainty which is an essential feature of a modern economy.
- For the rest of this talk, we focus on a specific problem.
- To understand the evolution of inequality and uncertainty in labor earnings for the U.S. economy.
- According to Levy and Murnane (1992):
  - Earnings inequality was stable in the 1970s but increased rapidly over the 1980s.
  - Inequality between age-education groups was stable in the 1970s and rose sharply in the 1980s.
Accounting for uncertainty

- Thus far we have ignored uncertainty which is an essential feature of a modern economy.
- For the rest of this talk, we focus on a specific problem.
- To understand the evolution of inequality and uncertainty in labor earnings for the U.S. economy.
- According to Levy and Murnane (1992):
  - Earnings inequality was stable in the 1970s but increased rapidly over the 1980s.
  - Inequality between age-education groups was stable in the 1970s and rose sharply in the 1980s.
  - Inequality within age-education groups has grown steadily since the 1970s.
Accounting for uncertainty

- Thus far we have ignored uncertainty which is an essential feature of a modern economy.
- For the rest of this talk, we focus on a specific problem.
- To understand the evolution of inequality and uncertainty in labor earnings for the U.S. economy.
- According to Levy and Murnane (1992):
  - Earnings inequality was stable in the 1970s but increased rapidly over the 1980s.
  - Inequality between age-education groups was stable in the 1970s and rose sharply in the 1980s.
  - Inequality within age-education groups has grown steadily since the 1970s.
  - This trend stopped in the mid 1990s.
Accounting for uncertainty

- How should we interpret this increase?

\[ Y_{i,s,t} = X_{i,s,t} \beta_s \alpha_t + \theta_i \alpha_{s,t} + \epsilon_{i,s,t} \]

\( Y_{i,s,t} \) are earnings of person \( i \) at time \( t \) in sector \( s \), \( i = 1, \ldots, I \), \( t = 1, \ldots, T \), \( s = 1, \ldots, S \). The vectors \( X, \theta \) represent the endowments of observable and unobservable skills, respectively. The vectors \( \beta, \alpha \) are prices of the skills.
Accounting for uncertainty

- How should we interpret this increase?
  - More uncertainty?
Accounting for uncertainty

- How should we interpret this increase?
  - More uncertainty?
  - More heterogeneity? (Or diversity among agents?)
Accounting for uncertainty

- How should we interpret this increase?
  - More uncertainty?
  - More heterogeneity? (Or diversity among agents?)

- One way to think about these issues is using the Gorman-Lancaster characteristics model of earnings (see, for example, Heckman and Scheinkman, 1987):

\[ Y_{i,s,t} = X_{i,s,t}\beta_{s,t} + \theta_i\alpha_{s,t} + \varepsilon_{i,s,t} \]
Accounting for uncertainty

- How should we interpret this increase?
  - More uncertainty?
  - More heterogeneity? (Or diversity among agents?)

- One way to think about these issues is using the Gorman-Lancaster characteristics model of earnings (see, for example, Heckman and Scheinkman, 1987):

  \[ Y_{i,s,t} = X_{i,s,t} \beta_{s,t} + \theta_i \alpha_{s,t} + \varepsilon_{i,s,t} \]

- \( Y_{i,s,t} \) are earnings of person \( i \) at time \( t \) in sector \( s \), \( i = 1, \ldots, I \), \( t = 1, \ldots, T \), \( s = 1, \ldots, S \).
Accounting for uncertainty

- How should we interpret this increase?
  - More uncertainty?
  - More heterogeneity? (Or diversity among agents?)

- One way to think about these issues is using the Gorman-Lancaster characteristics model of earnings (see, for example, Heckman and Scheinkman, 1987):

\[ Y_{i,s,t} = X_{i,s,t} \beta_{s,t} + \theta_i \alpha_{s,t} + \varepsilon_{i,s,t} \]

- \( Y_{i,s,t} \) are earnings of person \( i \) at time \( t \) in sector \( s \), \( i = 1, \ldots, I, \ t = 1, \ldots, T, \ s = 1, \ldots, S \).

- The vectors \( X, \theta \) represent the endowments of observable and unobservable skills, respectively. The vectors \( \beta, \alpha \) are prices of the skills.
Accounting for uncertainty

- $\varepsilon_{i,s,t}$ could represent unmeasured (but known by the individual) factors that affect outcomes $Y_{i,s,t}$ or productivity shocks (not known by the individual).
Accounting for uncertainty

- $\varepsilon_{i,s,t}$ could represent unmeasured (but known by the individual) factors that affect outcomes $Y_{i,s,t}$ or productivity shocks (not known by the individual).

- Gottschalk and Moffitt (1994) separate permanent from transitory shocks by considering a version of the model:

$$\log Y_{i,t} = X_{i,t} \beta_t + \theta_i + \varepsilon_{i,t}$$
Accounting for uncertainty

- $\varepsilon_{i,s,t}$ could represent unmeasured (but known by the individual) factors that affect outcomes $Y_{i,s,t}$ or productivity shocks (not known by the individual).

- Gottschalk and Moffitt (1994) separate permanent from transitory shocks by considering a version of the model:

  $$\log Y_{i,t} = X_{i,t}\beta_t + \theta_i + \varepsilon_{i,t}$$

- They show that both the variance of $\theta$ and the variance of $\varepsilon$ has increased when one compares the period 1970-1978 with the period 1978-1987.
Accounting for uncertainty

- $\varepsilon_{i,s,t}$ could represent unmeasured (but known by the individual) factors that affect outcomes $Y_{i,s,t}$ or productivity shocks (not known by the individual).

- Gottschalk and Moffitt (1994) separate permanent from transitory shocks by considering a version of the model:

\[
\log Y_{i,t} = X_{i,t} \beta_t + \theta_i + \varepsilon_{i,t}
\]

- They show that both the variance of $\theta$ and the variance of $\varepsilon$ has increased when one compares the period 1970-1978 with the period 1978-1987.

- They call the increase in the variance of temporary shocks $\varepsilon$ an increase in earnings instability.
Accounting for uncertainty

Today we focus our attention on how much of the increase in inequality is forecastable by agents early in life (i.e., around ages 17-18).
Today we focus our attention on how much of the increase in inequality is forecastable by agents early in life (i.e., around ages 17-18).

We call heterogeneity the part of lifetime inequality that is forecastable at ages 17-18.
Accounting for uncertainty

- Today we focus our attention on how much of the increase in inequality is forecastable by agents early in life (i.e., around ages 17-18).

- We call heterogeneity the part of lifetime inequality that is forecastable at ages 17-18.

- We call uncertainty the part of lifetime inequality that is not forecastable at ages 17-18.
Accounting for uncertainty

- Today we focus our attention on how much of the increase in inequality is forecastable by agents early in life (i.e., around ages 17-18).

- We call heterogeneity the part of lifetime inequality that is forecastable at ages 17-18.

- We call uncertainty the part of lifetime inequality that is not forecastable at ages 17-18.

- We build on Gottschalk and Moffitt (1994) and analyze the dynamics of heterogeneity and uncertainty in the U.S. economy.
Accounting for uncertainty

- We build on Carneiro, Hansen and Heckman (2003) and Cunha, Heckman and Navarro (2005) to estimate uncertainty facing new cohorts of labor market entrants and how uncertainty evolves over cohorts.
Accounting for uncertainty

- We build on Carneiro, Hansen and Heckman (2003) and Cunha, Heckman and Navarro (2005) to estimate uncertainty facing new cohorts of labor market entrants and how uncertainty evolves over cohorts.

- We estimate the information set that agents act on. We do not impose it.
Accounting for uncertainty

- We build on Carneiro, Hansen and Heckman (2003) and Cunha, Heckman and Navarro (2005) to estimate uncertainty facing new cohorts of labor market entrants and how uncertainty evolves over cohorts.

- We estimate the information set that agents act on. We do not impose it.

- We show how heterogeneity and uncertainty change over time by analyzing two distinct cohorts: The NLS/1966 versus the NLSY/1979.
Accounting for uncertainty

- We build on Carneiro, Hansen and Heckman (2003) and Cunha, Heckman and Navarro (2005) to estimate uncertainty facing new cohorts of labor market entrants and how uncertainty evolves over cohorts.

- We estimate the information set that agents act on. We do not impose it.

- We show how heterogeneity and uncertainty change over time by analyzing two distinct cohorts: The NLS/1966 versus the NLSY/1979.

- By exploring schooling choices together with realized earnings, we are able to distinguish which elements of \( \theta \) are known by the agent and which elements of \( \theta \) are not known at the time of schooling choice (i.e., among the early cohorts).
Accounting for uncertainty

- We find that lifetime earnings inequality has a substantial predictable component for the agent by age 17–18.
Accounting for uncertainty

- We find that lifetime earnings inequality has a substantial predictable component for the agent by age 17–18.
- Forecastability at age 17-18 for the NLS/1966 was both relatively and absolutely larger than for the NLSY/1979 cohort.
The model

This model builds on the framework developed for *ex post* models.
The model

- This model builds on the framework developed for *ex post* models.
- Schooling choices $S$
The model

- This model builds on the framework developed for *ex post* models.
- Schooling choices $S$
- Use $S$ instead of $D$ because we are considering schooling.
The model

- This model builds on the framework developed for *ex post* models.
- Schooling choices $S$
- Use $S$ instead of $D$ because we are considering schooling.
- Realized Earnings $Y_t = SY_{1,t} + (1 - S) Y_{0,t}$ for $t = 1, 2$. 
The model

- This model builds on the framework developed for *ex post* models.
- Schooling choices $S$
- Use $S$ instead of $D$ because we are considering schooling.
- Realized Earnings $Y_t = SY_{1,t} + (1 - S) Y_{0,t}$ for $t = 1, 2$.
- Explanatory variables in earnings equations $X$. 

Determinants of cost $Z$. A set of $K$ test scores $M_1, M_2, \ldots, M_K$ for each individual.
The model

- This model builds on the framework developed for *ex post* models.
- Schooling choices $S$
- Use $S$ instead of $D$ because we are considering schooling.
- Realized Earnings $Y_t = SY_{1,t} + (1 - S) Y_{0,t}$ for $t = 1, 2$.
- Explanatory variables in earnings equations $X$.
- Determinants of cost $Z$. 
This model builds on the framework developed for ex post models.

Schooling choices $S$

Use $S$ instead of $D$ because we are considering schooling.

Realized Earnings $Y_t = SY_{1,t} + (1 - S) Y_{0,t}$ for $t = 1, 2$.

Explanatory variables in earnings equations $X$.

Determinants of cost $Z$.

A set of $K$ test scores $M_1, M_2, \ldots, M_K$ for each individual.
This model builds on the framework developed for *ex post* models.

- Schooling choices $S$
- Use $S$ instead of $D$ because we are considering schooling.
- Realized Earnings $Y_t = SY_{1,t} + (1 - S) Y_{0,t}$ for $t = 1, 2$.
- Explanatory variables in earnings equations $X$.
- Determinants of cost $Z$.
- A set of $K$ test scores $M_1, M_2, \ldots, M_K$ for each individual.
- Explanatory variables in test score equations $X^M$. 
We assume that $Y_{s,t}$ for $s = 0, 1$, $t = 1, \ldots, T$ can be decomposed in the following manner:

$$Y_{0,t} = \mu_{0,t} + U_{0,t}, \quad E(U_{0,t}) = 0$$  \hspace{1cm} (6)$$

$$Y_{1,t} = \mu_{1,t} + U_{1,t}, \quad E(U_{1,t}) = 0$$  \hspace{1cm} (7)$$
The model

- We assume that $Y_{s,t}$ for $s = 0, 1$, $t = 1, \ldots, T$ can be decomposed in the following manner:

\[
Y_{0,t} = \mu_{0,t} + U_{0,t}, \quad E(U_{0,t}) = 0
\]  
\[
Y_{1,t} = \mu_{1,t} + U_{1,t}, \quad E(U_{1,t}) = 0
\]  

- The psychic costs $C$ are decomposed in observable $Z$ and unobservable $U_C$ determinants in the following manner:

\[
C = \mu_C + U_C
\]
The model

- We assume that $Y_{s,t}$ for $s = 0, 1$, $t = 1, \ldots, T$ can be decomposed in the following manner:

\[
Y_{0,t} = \mu_{0,t} + U_{0,t}, \quad E(U_{0,t}) = 0 \quad (6)
\]

\[
Y_{1,t} = \mu_{1,t} + U_{1,t}, \quad E(U_{1,t}) = 0 \quad (7)
\]

- The psychic costs $C$ are decomposed in observable $Z$ and unobservable $U_C$ determinants in the following manner

\[
C = \mu_C + U_C \quad (8)
\]

- The test score $M_k$ follows a linear in parameters model where $X^M$ are test score predictors:

\[
M_k = \mu_k^M + U_k^M, \quad k = 1, 2, \ldots, K.
\]
The schooling equation is based on

\[
I = E \left[ \sum_t \left( \frac{1}{1 + \rho} \right)^t (Y_{1t} - Y_{0t}) - C \right] \ (9)
\]
The schooling equation is based on

$$I = E \left[ \sum_t \left( \frac{1}{1 + \rho} \right)^t (Y_{1t} - Y_{0t}) - C \right] I$$ (9)

If we replace (6), (7), and (8) into (9) we get:

$$I = E \left\{ \sum_t \left( \frac{1}{1 + \rho} \right)^t (\mu_{1,t} - \mu_{0,t}) - \mu_C \right\} I$$

$$+ \sum_t \left( \frac{1}{1 + \rho} \right)^t (U_{1t} - U_{0t}) - U_C \right\} I$$
The model

We observe college earnings $Y_{1,t}$ only for the individuals who choose $S = 1$. 
The model

- We observe college earnings $Y_{1,t}$ only for the individuals who choose $S = 1$.
- $S = 1$ if, and only if $I > 0$, i.e.,

$$E \left\{ \sum_t \left( \frac{1}{1+\rho} \right)^t (\mu_{1,t} - \mu_{0,t}) - \mu_C + \sum_t \left( \frac{1}{1+\rho} \right)^t (U_{1,t} - U_{0,t}) - U_C | \mathcal{I} \right\} \geq 0$$
Assume that $U_C, X, Z \in \mathcal{I}$. The event $S = 1$ corresponds to the event

$$E \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (U_{1,t} - U_{0,t}) \middle| \mathcal{I} \right) - U_C$$

$$\geq \mu_C - \sum_t \left( \frac{1}{1 + \rho} \right)^t (\mu_{1,t} - \mu_{0,t}).$$

or, in more compact notation:

$$E (\bar{U} \middle| \mathcal{I}) - U_C \geq -\mu_I.$$
Consequently, from data, we can compute:

\[
E \left[ Y_{1,t} \mid E (\bar{U} \mid I) - U_C \geq -\mu_I \right] \\
= \mu_{1,t} + E \left[ U_{1,t} \mid E (U \mid I) - U_C \geq -\mu_I \right]
\]

We want to separate out two \textbf{unobservable} components:
Consequently, from data, we can compute:

\[
E \left[ Y_{1,t} \mid E (\bar{U} \mid I) - U_C \geq -\mu_I \right] = \mu_{1,t} + E \left[ U_{1,t} \mid E (U \mid I) - U_C \geq -\mu_I \right]
\]

We want to separate out two unobservable components:

- The component that is known and acted on by the agent (heterogeneity):

\[
E \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (U_{1,t} - U_{0,t}) \right| I \right).
\]
The component that is unknown by the agent (uncertainty):

\[
\left(\sum_t \left( \frac{1}{1 + \rho} \right)^t (U_{1,t} - U_{0,t}) \right) - E \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (U_{1,t} - U_{0,t}) \bigg| I \right).
\]
The component that is unknown by the agent (uncertainty):

$$\left( \sum_{t} \left( \frac{1}{1 + \rho} \right)^{t} (U_{1,t} - U_{0,t}) \right)$$

$$- E \left( \sum_{t} \left( \frac{1}{1 + \rho} \right)^{t} (U_{1,t} - U_{0,t}) \mid I \right).$$

How do we formally do it?
Factor models

Remember that the test score equations were specified as:

\[ M_k = \mu_k^M + U_k^M, \quad k = 1, 2, \ldots, K \]

We assume:
Remember that the test score equations were specified as:

\[ M_k = \mu_k^M + U_k^M, \quad k = 1, 2, \ldots, K \]

Now we break unobservable \( U_k^M \) into factors and uniquenesses to obtain

\[ M_k = \mu_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M, \quad k = 1, 2, \ldots, K. \]

We assume:
Factor models

- Remember that the test score equations were specified as:

\[ M_k = \mu_k^M + U_k^M, \quad k = 1, 2, \ldots, K \]

- Now we break unobservable \( U_k^M \) into factors and uniquenesses to obtain

\[ M_k = \mu_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M, \quad k = 1, 2, \ldots, K. \]

We assume:

- \( \theta_1 \sim N(0, \sigma_{\theta_1}^2) \) (normality is not necessary) and independent from \( \varepsilon_k^M \).
\[ \varepsilon_k^M \sim N \left(0, \sigma^2_{\varepsilon_k^M} \right) \text{ (normality is not necessary) and independent from } \varepsilon_{\ell}^M \text{ for } \ell \neq k. \]
Factor models

- \( \varepsilon_k^M \sim N \left( 0, \sigma^2_{\varepsilon_k^M} \right) \) (normality is not necessary) and independent from \( \varepsilon_{\ell}^M \) for \( \ell \neq k \).
- \( \alpha_1^M = 1 \) (recall that in factor analysis one such normalization is always necessary because scales are arbitrary).
Factor models

\[ \varepsilon_k^M \sim \mathcal{N} \left( 0, \sigma_{\varepsilon_k}^2 \right) \] (normality is not necessary) and independent from \( \varepsilon_{\ell}^M \) for \( \ell \neq k \).

\[ \alpha_1^M = 1 \] (recall that in factor analysis one such normalization is always necessary because scales are arbitrary).

Note, in particular, that the covariance between \( U_k^M \) and \( U_\ell^M \) is captured only by \( \theta_1 \) for \( k \neq \ell \).
Remember that we proposed the following model for earnings equations:

\[ Y_{0,t} = \mu_{0,t} + U_{0,t} \]

\[ Y_{1,t} = \mu_{1,t} + U_{1,t} \]
Remember that we proposed the following model for earnings equations:

\[ Y_{0,t} = \mu_{0,t} + U_{0,t} \]

\[ Y_{1,t} = \mu_{1,t} + U_{1,t} \]

Now we break unobservables \( U_{s,t} \) into three different components to obtain

\[ Y_{0,t} = \mu_{0,t} + \alpha_{0,t}\theta_1 + \delta_{0,t}\theta_2 + \varepsilon_{0,t} \]  
\( \text{(10)} \)

\[ Y_{1,t} = \mu_{1,t} + \alpha_{1,t}\theta_1 + \delta_{1,t}\theta_2 + \varepsilon_{1,t} \]  
\( \text{(11)} \)
We assume:

- $\theta_2 \sim N \left(0, \sigma_{\theta_2}^2\right)$ (normality is not necessary) and independent from $\theta_1$, and $\{ \varepsilon_0, t \}$
We assume:

- $\theta_2 \sim N\left(0, \sigma_{\theta_2}^2\right)$ (normality is not necessary) and independent from $\theta_1$, and $\{\varepsilon_{0,t}\}$

- $\varepsilon_{s,t} \sim N\left(0, \sigma_{s,t}^2\right)$ (normality is not necessary) and independent from $\varepsilon_{s',\tau}$ for $\tau \neq t$. 

Note, again, that the dependence between $Y_{s,t}$ and $Y_{s',t'}$ is captured only by $\theta_1$ and $\theta_2$. 

Remember that we proposed the following model for costs $C$:

$$C = \mu_C + U_C$$
We assume:

- $\theta_2 \sim N(0, \sigma^2_{\theta_2})$ (normality is not necessary) and independent from $\theta_1$, and $\{\varepsilon_{0,t}\}$
- $\varepsilon_{s,t} \sim N(0, \sigma^2_{s,t})$ (normality is not necessary) and independent from $\varepsilon_{s',\tau}$ for $\tau \neq t$.
- $\delta_{1,1} = 1$. 

Note, again, that the dependence between $Y_s, t$ and $Y_{s'}, t'$ is captured only by $\theta_1$ and $\theta_2$. Remember that we proposed the following model for costs $C$:

$$C = \mu_C + U_C$$
We assume:

- \( \theta_2 \sim N(0, \sigma_{\theta_2}^2) \) (normality is not necessary) and independent from \( \theta_1 \), and \( \{\varepsilon_{0,t}\} \)

- \( \varepsilon_{s,t} \sim N(0, \sigma_{s,t}^2) \) (normality is not necessary) and independent from \( \varepsilon_{s',\tau} \) for \( \tau \neq t \).

- \( \delta_{1,1} = 1 \).

- Note, again, that the dependence between \( Y_{s,t} \) and \( Y_{s',t'} \) is captured only by \( \theta_1 \) and \( \theta_2 \).
We assume:

- $\theta_2 \sim N \left(0, \sigma_{\theta_2}^2 \right)$ (normality is not necessary) and independent from $\theta_1$, and $\{\varepsilon_{0,t}\}$
- $\varepsilon_{s,t} \sim N \left(0, \sigma_{s,t}^2 \right)$ (normality is not necessary) and independent from $\varepsilon_{s',\tau}$ for $\tau \neq t$.
- $\delta_{1,1} = 1$.

Note, again, that the dependence between $Y_{s,t}$ and $Y_{s',t'}$ is captured only by $\theta_1$ and $\theta_2$.

Remember that we proposed the following model for costs $C$:

$$C = \mu_C + U_C$$
Now we decompose the residuals $U_C$ in three different components to obtain

$$C = \mu_C + \alpha_C \theta_1 + \delta_C \theta_2 + \varepsilon_C, \quad \varepsilon_C \sim N(0, \sigma_C^2).$$

(12)
Now we decompose the residuals $U_C$ in three different components to obtain

$$C = \mu_C + \alpha_C \theta_1 + \delta_C \theta_2 + \varepsilon_C, \quad \varepsilon_C \sim \mathcal{N}(0, \sigma^2_C). \quad (12)$$

The schooling equation is generated by

$$I = E \left\{ \mu_I + \sum_t \left( \frac{1}{1 + \rho} \right)^t (U_{1t} - U_{0t}) - U_C \mid \mathcal{I} \right\}$$
Given (10), (11), and (12) the schooling equation becomes:

\[
I = \mu_I + E \left[ \theta_1 \left( \sum_t \left( \frac{1}{1+\rho} \right)^t \left( \alpha_{1,t} - \alpha_{0,t} \right) \right) \bigg| \mathcal{I} \right] + \\
+ E \left[ \theta_2 \left( \sum_t \left( \frac{1}{1+\rho} \right)^t \left( \delta_{1,t} - \delta_{0,t} \right) \right) \bigg| \mathcal{I} \right] + \\
+ E \left[ \sum_t \left( \frac{1}{1+\rho} \right)^t \left( \varepsilon_{1,t} - \varepsilon_{0,t} \right) \bigg| \mathcal{I} \right]
\]
We assume:

\[ \varepsilon_C \in \mathcal{I} \]
We assume:

- $\varepsilon_C \in I$
- $\varepsilon_{s,t} \notin I$ and $E(\varepsilon_{s,t}|I) = 0.$
We assume:

- $\varepsilon_C \in \mathcal{I}$
- $\varepsilon_{s,t} \notin \mathcal{I}$ and $E(\varepsilon_{s,t}|\mathcal{I}) = 0$.
- We postulate $H_0 : \{\theta_1, \theta_2\} \subset \mathcal{I}$. We test among alternative specifications of $\mathcal{I}$.
We assume:

- $\varepsilon_C \in \mathcal{I}$
- $\varepsilon_{s,t} \notin \mathcal{I}$ and $E(\varepsilon_{s,t}|\mathcal{I}) = 0$.
- We postulate $H_0 : \{\theta_1, \theta_2\} \subset \mathcal{I}$. We test among alternative specifications of $\mathcal{I}$.
- We don’t want to impose \textit{a priori} that certain factors are in the information set of the agents.
We want to determine whether $\theta_1 \in \mathcal{I}$ or $\theta_1 \notin \mathcal{I}$ and whether $\theta_2 \in \mathcal{I}$ or $\theta_2 \notin \mathcal{I}$.
We want to determine whether $\theta_1 \in I$ or $\theta_1 \notin I$ and whether $\theta_2 \in I$ or $\theta_2 \notin I$.

Under (1)-(3) the schooling equation can be written as:

$$l = \mu_I + \sum_t \left( \frac{1}{1 + \rho} \right)^t \left[ (\alpha_{1,t} - \alpha_{0,t}) \theta_1 + (\delta_{1,t} - \delta_{0,t}) \theta_2 \right] - \alpha_C \theta_1 - \delta_C \theta_2 - \varepsilon_C$$
How to identify the information set of the agent

Postulate $H_0 : \theta_1, \theta_2 \in \mathcal{I}$ against $H_1 : \theta_1 \in \mathcal{I}$ but $\theta_2 \notin \mathcal{I}$. 
How to identify the information set of the agent

- Postulate $H_0 : \theta_1, \theta_2 \in \mathcal{I}$ against $H_1 : \theta_1 \in \mathcal{I}$ but $\theta_2 \notin \mathcal{I}$.

- How do we test it? Under $H_0$:

$$\text{Cov} (I - \mu_I, Y_{1,1} - \mu_{1,1}) = \alpha_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \right) \sigma_{\theta_1}^2 +$$

$$+ \delta_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C \right) \sigma_{\theta_2}^2$$
How to identify the information set of the agent

Under $H_1$:

\[
\text{Cov}(I - \mu_1, Y_{1,1} - \mu_{1,1}) = \alpha_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\alpha_{1,t} - \alpha_{0,t} - \alpha_C) \right) \sigma_{\theta_1}^2
\]
Another way to state the test is:

\[
\text{Cov} (I - \mu_I, Y_{1,1} - \mu_{1,1})
\]

\[
= \alpha_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \right) \sigma_{\theta_1}^2 +
\]

\[
+ \Delta_2 \delta_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C \right) \sigma_{\theta_2}^2
\]

\(H_0: \Delta_2 \neq 0\) versus \(H_1: \Delta_2 = 0\). We iterate among the alternative specifications of \(I\) and produce a model which fits the data best.
Another way to state the test is:

\[
\text{Cov}(I - \mu_I, Y_{1,1} - \mu_{1,1}) = \alpha_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \right) \sigma_{\theta_1}^2 + \\
+ \Delta_2 \delta_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C \right) \sigma_{\theta_2}^2
\]

\[H_0 : \Delta_2 \neq 0 \text{ versus } H_1 : \Delta_2 = 0.\]
Another way to state the test is:

\[
\text{Cov} (I - \mu_I, Y_{1,1} - \mu_{1,1}) \\
= \alpha_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \right) \sigma_{\theta_1}^2 + \\
+ \Delta_2 \delta_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C \right) \sigma_{\theta_2}^2
\]

- \( H_0 : \Delta_2 \neq 0 \) versus \( H_1 : \Delta_2 = 0 \).
- We iterate among the alternative specifications of \( \mathcal{I} \) and produce a model which fits the data best.
How can the tests be implemented?
How can the tests be implemented?

Take the test score equations:

\[ M_k = \mu_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M, \quad k = 1, 2, \ldots, K. \]  \hspace{1cm} (13)

If \( K \geq 3 \) we can identify \( \alpha_k^M \) and \( f(\theta_1) \) up to a normalization (say \( \alpha_1^M = 1 \)).
How can the tests be implemented?

Take the test score equations:

\[ M_k = \mu_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M, \quad k = 1, 2, \ldots, K. \]  

(13)

If \( K \geq 3 \) we can identify \( \alpha_k^M \) and \( f(\theta_1) \) up to a normalization (say \( \alpha_1^M = 1 \)).

Because of independence we can identify \( \mu_k^M \) from a simple OLS regression in (13).
We can construct the covariances:

\[ \text{Cov} \left( M_1 - \mu_1^M, M_2 - \mu_2^M \right) = \alpha_2^M \sigma_{\theta_1}^2 \]  \hspace{1cm} (14)

\[ \text{Cov} \left( M_1 - \mu_1^M, M_3 - \mu_3^M \right) = \alpha_3^M \sigma_{\theta_1}^2 \]  \hspace{1cm} (15)

\[ \text{Cov} \left( M_3 - \mu_3^M, M_2 - \mu_2^M \right) = \alpha_3^M \alpha_2^M \sigma_{\theta_1}^2 \]  \hspace{1cm} (16)
We can construct the covariances:

\[
\text{Cov} \left( M_1 - \mu_1^M, M_2 - \mu_2^M \right) = \alpha_2^M \sigma_{\theta_1}^2 \tag{14}
\]

\[
\text{Cov} \left( M_1 - \mu_1^M, M_3 - \mu_3^M \right) = \alpha_3^M \sigma_{\theta_1}^2 \tag{15}
\]

\[
\text{Cov} \left( M_3 - \mu_3^M, M_2 - \mu_2^M \right) = \alpha_3^M \alpha_2^M \sigma_{\theta_1}^2 \tag{16}
\]

Consequently, can recover \( \alpha_2^M, \alpha_3^M, \text{ and } \sigma_{\theta_1}^2 \).
We can construct the covariances:

\[
\text{Cov} \left( M_1 - \mu_1^M, M_2 - \mu_2^M \right) = \alpha_2^M \sigma_{\theta_1}^2 \tag{14}
\]

\[
\text{Cov} \left( M_1 - \mu_1^M, M_3 - \mu_3^M \right) = \alpha_3^M \sigma_{\theta_1}^2 \tag{15}
\]

\[
\text{Cov} \left( M_3 - \mu_3^M, M_2 - \mu_2^M \right) = \alpha_3^M \alpha_2^M \sigma_{\theta_1}^2 \tag{16}
\]

Consequently, can recover \( \alpha_2^M \), \( \alpha_3^M \), and \( \sigma_{\theta_1}^2 \).

Recent work by Schennach (2004) allows these to be identified under much more general conditions than independence.
Consider now the college earnings equation

\[ Y_{1,t} = \mu_{1,t} + \alpha_{1,t} \theta_1 + \delta_{1,t} \theta_2 + \varepsilon_{1,t} \]
Consider now the college earnings equation

\[ Y_{1,t} = \mu_{1,t} + \alpha_{1,t} \theta_1 + \delta_{1,t} \theta_2 + \varepsilon_{1,t} \]

We cannot use OLS regression anymore because of the selection problem:

\[ E(Y_{1,t} \mid S = 1) = \mu_{1,t} + E(\alpha_{1,t} \theta_1 + \delta_{1,t} \theta_2 + \varepsilon_{1,t} \mid S = 1) \]
Assuming that the unobservables are all normal, it follows that:

\[ E(Y_{1,t} \mid S = 1) = \mu_{1,t} + \pi_{1,t} \lambda \left( \sum_t \left( \frac{1}{1+\rho} \right)^t (\mu_{1,t} - \mu_{0,t}) - \mu_C \right) \]

\[ \sigma_U \]

\text{selection correction}
Under normality we can use standard selection estimators and recover $\mu_{1,t}$. 
Under normality we can use standard selection estimators and recover $\mu_{1,t}$.

Normality is not required, just easy to understand. It motivates how it is possible to identify these parameters.
Under normality we can use standard selection estimators and recover $\mu_{1,t}$.

Normality is not required, just easy to understand. It motivates how it is possible to identify these parameters.

We can do the same analysis for high school earnings equations and recover $\mu_{0,t}$. 
Under normality we can use standard selection estimators and recover $\mu_{1,t}$.

Normality is not required, just easy to understand. It motivates how it is possible to identify these parameters.

We can do the same analysis for high school earnings equations and recover $\mu_{0,t}$.

$\pi_{s,t}$ is a coefficient that depends on $\rho, \sigma_{\theta_1}^2, \sigma_{\theta_2}^2, \sigma_{\epsilon_C}^2, \alpha_{s,t}$, and $\delta_{s,t}$ for $s = 0, 1$ and $t = 1, 2, \ldots, T$. 
Under normality we can use standard selection estimators and recover $\mu_{1,t}$.

Normality is not required, just easy to understand. It motivates how it is possible to identify these parameters.

We can do the same analysis for high school earnings equations and recover $\mu_{0,t}$

$\pi_{s,t}$ is a coefficient that depends on $\rho, \sigma^2_{\theta_1}, \sigma^2_{\theta_2}, \sigma^2_{\varepsilon_C}, \alpha_{s,t}$, and $\delta_{s,t}$ for $s = 0, 1$ and $t = 1, 2, \ldots, T$.

Once we recover $\beta_{s,t}$ we can compute the covariances:

$$\text{Cov} \left( M_1 - \mu_1^M, Y_{s,t} - \mu_{s,t} \right) = \alpha_{s,t} \sigma^2_{\theta_1}$$
And it is easy to see that we can identify $\alpha_{s,t}$ for all $s, t$ because we have already determined $\sigma_{\theta_1}^2$ from the test score equations.
Identification

- And it is easy to see that we can identify $\alpha_{s,t}$ for all $s, t$ because we have already determined $\sigma_{\theta_1}^2$ from the test score equations.

- We use the covariance of earnings over time to identify the parameters associated to $\theta_2$:

$$\text{Cov} \left( Y_{s,\tau} - \mu_{s,\tau}, Y_{s,t} - \mu_{s,t} \right) = \alpha_{s,\tau} \alpha_{s,t} \sigma_{\theta_1}^2 + \delta_{s,\tau} \delta_{s,t} \sigma_{\theta_2}^2$$
And it is easy to see that we can identify $\alpha_{s,t}$ for all $s, t$ because we have already determined $\sigma_{\theta_1}^2$ from the test score equations.

We use the covariance of earnings over time to identify the parameters associated to $\theta_2$:

$$\text{Cov} \left( Y_{s,\tau} - \mu_{s,\tau}, Y_{s,t} - \mu_{s,t} \right) = \alpha_{s,\tau} \alpha_{s,t} \sigma_{\theta_1}^2 + \delta_{s,\tau} \delta_{s,t} \sigma_{\theta_2}^2$$

Under the normalization $\delta_{1,1} = 1$ we repeat the argument used in test scores and can recover $\delta_{s,t}$ and $\sigma_{\theta_2}^2$. 
And it is easy to see that we can identify $\alpha_{s,t}$ for all $s, t$ because we have already determined $\sigma_{\theta_1}^2$ from the test score equations.

We use the covariance of earnings over time to identify the parameters associated to $\theta_2$:

$$\text{Cov} (Y_{s,\tau} - \mu_{s,\tau}, Y_{s,t} - \mu_{s,t}) = \alpha_{s,\tau} \alpha_{s,t} \sigma_{\theta_1}^2 + \delta_{s,\tau} \delta_{s,t} \sigma_{\theta_2}^2$$

Under the normalization $\delta_{1,1} = 1$ we repeat the argument used in test scores and can recover $\delta_{s,t}$ and $\sigma_{\theta_2}^2$.

It is interesting to note that we can then recover joint distributions:

$$\text{Cov} (Y_{1,t}, Y_{0,\tau}) = \alpha_{0,\tau} \alpha_{1,t} \sigma_{\theta_1}^2 + \delta_{0,\tau} \delta_{1,t} \sigma_{\theta_2}^2$$
To identify $\alpha_C$ we use:

$$\text{Cov}(M_1 - \mu_1^M, I - \mu_I) = \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \right) \sigma_{\theta_1}^2$$
To identify $\delta_C$ we use:

\[
\text{Cov} \left( Y_{1,1} - X_1 \beta_{1,1}, I - \mu_1 \right) = \alpha_{1,1} \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \right) \sigma^2_{\theta_1} + \\
+ \left( \sum_{t=1} \left( \frac{1}{1 + \rho} \right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C \right) \sigma^2_{\theta_2}
\]
Summary of Empirical Results

- Returns to college have increased in past 20 years.
Summary of Empirical Results

- Returns to college have increased in past 20 years.
- Predictable components are more than half of 1966 variance for college and about half for high school in 1966.
Summary of Empirical Results

- Returns to college have increased in past 20 years.
- Predictable components are more than half of 1966 variance for college and about half for high school in 1966.
- Variance of residual earnings increases across cohorts.
Summary of Empirical Results

- Returns to college have increased in past 20 years.
- Predictable components are more than half of 1966 variance for college and about half for high school in 1966.
- Variance of residual earnings increases across cohorts.
- Earnings variances less predictable for 1979 than 1966.
Summary of Empirical Results

- Returns to college have increased in past 20 years.
- Predictable components are more than half of 1966 variance for college and about half for high school in 1966.
- Variance of residual earnings increases across cohorts.
- Earnings variances less predictable for 1979 than 1966.
- Increase in unforecastability happens after age 30.
Summary of Empirical Results

- Returns to college have increased in past 20 years.
- Predictable components are more than half of 1966 variance for college and about half for high school in 1966.
- Variance of residual earnings increases across cohorts.
- Earnings variances less predictable for 1979 than 1966.
- Increase in unforecastability happens after age 30.
- Persistence of shocks has increased over time.
Summary of Empirical Results

- Returns to college have increased in past 20 years.
- Predictable components are more than half of 1966 variance for college and about half for high school in 1966.
- Variance of residual earnings increases across cohorts.
- Earnings variances less predictable for 1979 than 1966.
- Increase in unforecastability happens after age 30.
- Persistence of shocks has increased over time.
- Cognitive skill prices as in Gorman-Lancaster model have gone up.
To study the evolution of labor earnings risk in the U.S. economy we compare two different samples:
Data Description

To study the evolution of labor earnings risk in the U.S. economy we compare two different samples:

1. NLSY/1979–The first sample consists of white males born between 1957 and 1964 and we obtain their information from NLSY/1979 data pooled their counterparts from the PSID data.

2. NLS/1966 - The second sample consists of the white males born between 1941 and 1952 and are surveyed from the NLS/1966 combined with their counterparts from the PSID data.

We consider only two schooling choices: high school and college graduation.

We consider labor income from ages 22 to 41. Concepts of labor income are the same in both years.
Data Description

To study the evolution of labor earnings risk in the U.S. economy we compare two different samples:

1. NLSY/1979–The first sample consists of white males born between 1957 and 1964 and we obtain their information from NLSY/1979 data pooled their counterparts from the PSID data.

2. NLS/1966 - The second sample consists of the white males born between 1941 and 1952 and are surveyed from the NLS/1966 combined with their counterparts from the PSID data.
To study the evolution of labor earnings risk in the U.S. economy we compare two different samples:

1. NLSY/1979–The first sample consists of white males born between 1957 and 1964 and we obtain their information from NLSY/1979 data pooled their counterparts from the PSID data.

2. NLS/1966 - The second sample consists of the white males born between 1941 and 1952 and are surveyed from the NLS/1966 combined with their counterparts from the PSID data.

We consider only two schooling choices: high school and college graduation.
To study the evolution of labor earnings risk in the U.S. economy we compare two different samples:

1. NLSY/1979—The first sample consists of white males born between 1957 and 1964 and we obtain their information from NLSY/1979 data pooled their counterparts from the PSID data.

2. NLS/1966 - The second sample consists of the white males born between 1941 and 1952 and are surveyed from the NLS/1966 combined with their counterparts from the PSID data.

- We consider only two schooling choices: high school and college graduation.

- We consider labor income from ages 22 to 41.
To study the evolution of labor earnings risk in the U.S. economy we compare two different samples:

1. **NLSY/1979**—The first sample consists of white males born between 1957 and 1964 and we obtain their information from NLSY/1979 data pooled their counterparts from the PSID data.

2. **NLS/1966** — The second sample consists of the white males born between 1941 and 1952 and are surveyed from the NLS/1966 combined with their counterparts from the PSID data.

- We consider only two schooling choices: high school and college graduation.
- We consider labor income from ages 22 to 41.
- Concepts of labor income are the same in both years.
In both data sets we observe cognitive test scores:

- For the NLSY/1979 we use five components of the ASVAB test battery: arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge and coding speed.
In both data sets we observe cognitive test scores:

- For the NLSY/1979 we use five components of the ASVAB test battery: arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge and coding speed.

- In the NLS/1966 there are many different achievement tests, but we use the two most commonly reported ones: the OTIS/BETA/GAMMA and the California Test of Mental Maturity (CTMM).
In both data sets we observe cognitive test scores:

- For the NLSY/1979 we use five components of the ASVAB test battery: arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge and coding speed.

- In the NLS/1966 there are many different achievement tests, but we use the two most commonly reported ones: the OTIS/BETA/GAMMA and the California Test of Mental Maturity (CTMM).

- One problem in the NLS/1966 sample is that there are no respondents for whom we observe scores from two distinct tests.
We complement the information from these test scores by considering other proxies for cognitive achievement. These are the tests on “knowledge of the world of work.”
We complement the information from these test scores by considering other proxies for cognitive achievement. These are the tests on “knowledge of the world of work.”

Even after controlling for parental education, number of siblings, urban residence at age 14, and dummies for year of birth, the “knowledge of the world of work” test scores are correlated with the cognitive test scores. The correlation with OTIS/BETA/GAMMA and CTMM is stronger for the occupation and education tests than for the earnings-comparison test.
Let $Y_1$ denote earnings at age 30. Here we plot the density functions $f(y_1|S=1)$ generated from the data (the solid curve), against that predicted by the model (the dashed line).
Let $Y_1$ denote earnings at age 28. Here we plot the density functions $f(y_1|S=1)$ generated from the data (the solid curve), against that predicted by the model (the dashed line).
Densities of earnings at age 37 (college sample NLSY/1979)

Let $Y_1$ denote earnings at age 37. Here we plot the density functions $f(y_1|S=1)$ generated from the data (the solid curve), against that predicted by the model (the dashed line).
Let $Y_1$ denote earnings at age 34. Here we plot the density functions $f(y_1|S=1)$ generated from the data (the solid curve), against that predicted by the model (the dashed line).
Let $Y_1$ denote earnings at age 40. Here we plot the density functions $f(y_1|S=1)$ generated from the data (the solid curve), against that predicted by the model (the dashed line).
### Mean Rates of Return to College by Schooling Group

<table>
<thead>
<tr>
<th>Schooling Group</th>
<th>NLS/1966</th>
<th></th>
<th>NLSY/1979</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Returns</td>
<td>Standard Error</td>
<td>Mean Returns</td>
<td>Standard Error</td>
</tr>
<tr>
<td>High School Graduates</td>
<td>0.2937</td>
<td>0.0083</td>
<td>0.3095</td>
<td>0.0113</td>
</tr>
<tr>
<td>College Graduates</td>
<td>0.3307</td>
<td>0.0114</td>
<td>0.3994</td>
<td>0.0129</td>
</tr>
<tr>
<td>Individuals at the Margin</td>
<td>0.3081</td>
<td>0.0446</td>
<td>0.3511</td>
<td>0.0535</td>
</tr>
</tbody>
</table>

Note: Under linearity, *ex ante* mean = *ex post* mean.
Let $Y_0$ and $Y_1$ denote the present value of high-school and college earnings, respectively. Here we plot the factual density function of high-school earnings for high-school graduates, $f(y_0|S=0)$ (the solid curve), against the counterfactual of college earnings for high-school graduates, $f(y_1|S=0)$ (the dashed line). We use kernel density estimation to smooth these functions. The present value of earnings are calculated using an interest rate of 5%.
Let $Y_0$ and $Y_1$ denote the present value of high-school and college earnings, respectively. Here we plot the factual density function of college earnings for college graduates, $f(y_1|S=1)$ (the solid curve), against the counterfactual of high-school earnings for college graduates, $f(y_0|S=1)$ (the dashed line). We use kernel density estimation to smooth these functions. The present value of earnings are calculated using an interest rate of 5%.
Let \( Y_0, Y_1 \) denote the present value of earnings from age 22 to age 41 in the high school and college sectors, respectively. Define ex post returns to college as the ratio \( R = (Y_1 - Y_0)/Y_0 \). Let \( f(r) \) denote the density function of the ex post returns to college \( R \).

The solid line is the density of ex post returns to college for high school graduates, that is, \( f(r|S=0) \). The dashed line is the density of ex post returns to college for college graduates, that is, \( f(r|S=1) \).
Let $Y_0$, $Y_1$ denote the present value of earnings from age 22 to age 41 in the high school and college sectors, respectively. Define ex post returns to college as the ratio $R=(Y_1-Y_0)/Y_0$. Let $f(r)$ denote the density function of the ex post returns to college $R$. The solid line is the density of ex post returns to college for high school graduates, that is, $f(r|S=0)$. The dashed line is the density of ex post returns to college for college graduates, that is, $f(r|S=1)$. 
Let $Y_0, Y_1$ denote the present value of earnings from age 22 to age 41 in the high school and college sectors, respectively. Define ex post returns to college as the ratio $R = (Y_1 - Y_0)/Y_0$. Let $f(r)$ denote the density function of the ex post returns to college $R$. The solid line is the density of ex post returns to college for high school graduates, that is, $f(r|S=0)$. The dashed line is the density of ex post returns to college for college graduates, that is, $f(r|S=1)$. 

Web Appendix Figure 4.6

Densities of Returns to College NLS/1966 Sample
Let $Y_0, Y_1$ denote the present value of earnings from age 22 to age 41 in the high school and college sectors, respectively. Define ex post returns to college as the ratio $R=(Y_1-Y_0)/Y_0$. Let $f(r)$ denote the density function of the ex post returns to college $R$. The solid line is the density of ex post returns to college for high school graduates, that is, $f(r|S=0)$. The dashed line is the density of ex post returns to college for college graduates, that is, $f(r|S=1)$. 

Densities of Returns to College NLS/1966 Sample
Densities of monetary value of psychic cost NLS/1966

Figure 8

Densities of monetary value of psychic cost
NLS/1966

Overall
High school
College
Densities of monetary value of psychic cost NLSY/1979

Figure 8: Densities of monetary value of psychic cost NLSY/1979

Overall
High school
College
### Evolution of Uncertainty

**Panel A: NLS/1966**

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Residual Variance</td>
<td>460.6260</td>
<td>284.8089</td>
<td>351.4026</td>
</tr>
<tr>
<td>Variance of Unforecastable Components</td>
<td>181.3712</td>
<td>128.4315</td>
<td>327.3480</td>
</tr>
</tbody>
</table>

**Panel B: NLSY/1979**

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Residual Variance</td>
<td>709.7487</td>
<td>507.2910</td>
<td>906.0066</td>
</tr>
<tr>
<td>Variance of Unforecastable Components</td>
<td>372.3509</td>
<td>272.3596</td>
<td>432.8733</td>
</tr>
</tbody>
</table>

**Panel C: Percentage Increase**

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage Increase in Total Residual Variance</td>
<td>54.083%</td>
<td>78.116%</td>
<td>157.826%</td>
</tr>
<tr>
<td>Percentage Increase in Variance of Unforecastable Components</td>
<td>105.298%</td>
<td>112.066%</td>
<td>32.236%</td>
</tr>
</tbody>
</table>
Evolution of Heterogeneity (Diversity)

Panel A: NLS/1966

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Residual Variance</td>
<td>460.6260</td>
<td>284.8089</td>
<td>351.4026</td>
</tr>
<tr>
<td>Variance of Forecastable Components (Heterogeneity)</td>
<td>279.2549</td>
<td>156.3774</td>
<td>24.0546</td>
</tr>
</tbody>
</table>

Panel B: NLSY/1979

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Residual Variance</td>
<td>709.7487</td>
<td>507.2910</td>
<td>906.0066</td>
</tr>
<tr>
<td>Variance of Forecastable Components (Heterogeneity)</td>
<td>337.3978</td>
<td>234.9314</td>
<td>473.1333</td>
</tr>
</tbody>
</table>

Panel C: Percentage Increase

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage Increase in Total Residual Variance</td>
<td>54.083%</td>
<td>78.116%</td>
<td>157.826%</td>
</tr>
<tr>
<td>Percentage Increase in Variance of Forecastable Components</td>
<td>20.821%</td>
<td>50.234%</td>
<td>1866.914%</td>
</tr>
</tbody>
</table>
Greater variance of returns in 1979.
- Greater variance of returns in 1979.
- Greater predictability of returns in 1979 as a fraction of the variance.
The densities of total residual vs unforecastable components in present value of college earnings for the NLS/1966 sample

In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of college earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5% interest rate.
Densities of total residual vs unforecastable components returns college vs high school for the NLSY/1979 sample

In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of earnings differences (or returns to college) for the white males sample of the NLSY/1979 from ages 22 to 41. The present value of returns to college is calculated using a 5% interest rate.
Densities of total residual vs unforecastable components returns college vs high school for the NLS/1966 sample

In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of earnings differences (or returns to college) for the white males sample of the NLSY/1979 from ages 22 to 41. The present value of returns to college is calculated using a 5% interest rate.
The densities of total residual vs unforecastable components in present value of high school earnings for the NLSY/1979 sample

In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a 5% interest rate.
The densities of total residual vs unforecastable components in present value of high school earnings for the NLS/1966 sample

In this figure, we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5% interest rate.
The densities of total residual vs unforecastable components in present value of college earnings for the NLSY/1979 sample

In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of college earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a 5% interest rate.
## Percentage that Regret Their Schooling Choices

<table>
<thead>
<tr>
<th>Schooling Group</th>
<th>NLS/1966</th>
<th>NLSY/1979</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of High School Graduates who Regret Not Graduating from College</td>
<td>0.0966</td>
<td>0.0749</td>
</tr>
<tr>
<td>Percentage of College Graduates who Regret Graduating from College</td>
<td>0.0337</td>
<td>0.0311</td>
</tr>
</tbody>
</table>
Table 6: Ex-Ante Conditional Distributions for the NLSY/1979 (College Earnings Conditional on High School Earnings)

\[ \Pr(d_i < Y_c < d_i+1 \mid d_j < Y_h < d_j+1) \]

where \( d_i \) is the \( i \)th decile of the College Lifetime Ex-Ante Earnings Distribution and \( d_j \) is the \( j \)th decile of the High School Ex-Ante Lifetime Earnings Distribution.

Individual fixes unknown \( \theta \) at their means, so Information Set=\{\( \theta_1, \theta_2, \theta_3 \}\)

\[ \text{Correlation}(Y_C, Y_H) = 0.1666 \]

<table>
<thead>
<tr>
<th>High School</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2995</td>
<td>0.1685</td>
<td>0.1114</td>
<td>0.0789</td>
<td>0.0570</td>
<td>0.0413</td>
<td>0.0393</td>
<td>0.0431</td>
<td>0.0471</td>
<td>0.1137</td>
</tr>
<tr>
<td>2</td>
<td>0.2273</td>
<td>0.2119</td>
<td>0.1597</td>
<td>0.1271</td>
<td>0.0907</td>
<td>0.0678</td>
<td>0.0450</td>
<td>0.0288</td>
<td>0.0180</td>
<td>0.0236</td>
</tr>
<tr>
<td>3</td>
<td>0.1532</td>
<td>0.1840</td>
<td>0.1656</td>
<td>0.1472</td>
<td>0.1146</td>
<td>0.0914</td>
<td>0.0642</td>
<td>0.0434</td>
<td>0.0230</td>
<td>0.0132</td>
</tr>
<tr>
<td>4</td>
<td>0.1110</td>
<td>0.1368</td>
<td>0.1492</td>
<td>0.1474</td>
<td>0.1418</td>
<td>0.1184</td>
<td>0.0882</td>
<td>0.0588</td>
<td>0.0334</td>
<td>0.0148</td>
</tr>
<tr>
<td>5</td>
<td>0.0748</td>
<td>0.1100</td>
<td>0.1244</td>
<td>0.1413</td>
<td>0.1459</td>
<td>0.1403</td>
<td>0.1172</td>
<td>0.0836</td>
<td>0.0462</td>
<td>0.0162</td>
</tr>
<tr>
<td>6</td>
<td>0.0494</td>
<td>0.0866</td>
<td>0.1146</td>
<td>0.1204</td>
<td>0.1371</td>
<td>0.1399</td>
<td>0.1283</td>
<td>0.1242</td>
<td>0.0736</td>
<td>0.0258</td>
</tr>
<tr>
<td>7</td>
<td>0.0306</td>
<td>0.0582</td>
<td>0.0904</td>
<td>0.1094</td>
<td>0.1264</td>
<td>0.1436</td>
<td>0.1506</td>
<td>0.1430</td>
<td>0.1064</td>
<td>0.0414</td>
</tr>
<tr>
<td>8</td>
<td>0.0236</td>
<td>0.0348</td>
<td>0.0531</td>
<td>0.0769</td>
<td>0.0989</td>
<td>0.1252</td>
<td>0.1638</td>
<td>0.1799</td>
<td>0.1676</td>
<td>0.0761</td>
</tr>
<tr>
<td>9</td>
<td>0.0264</td>
<td>0.0262</td>
<td>0.0316</td>
<td>0.0459</td>
<td>0.0651</td>
<td>0.0929</td>
<td>0.1308</td>
<td>0.1784</td>
<td>0.2431</td>
<td>0.1594</td>
</tr>
<tr>
<td>10</td>
<td>0.0457</td>
<td>0.0182</td>
<td>0.0214</td>
<td>0.0216</td>
<td>0.0321</td>
<td>0.0446</td>
<td>0.0772</td>
<td>0.1176</td>
<td>0.2291</td>
<td>0.3925</td>
</tr>
</tbody>
</table>
Table 7: Ex-Post Conditional Distributions for the NLSY/1979 (College Earnings Conditional on High School Earnings)

\[
\Pr(d_i < Y_c < d_i + 1 | d_j < Y_h < d_j + 1) \text{ where } d_i \text{ is the } i \text{th decile of the College Lifetime Ex-Ante Earnings Distribution and } d_j \text{ is the } j \text{th decile of the High School Ex-Ante Lifetime Earnings Distribution}
\]

Information Set=\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}

Correlation(Y_C,Y_H) = 0.2842

<table>
<thead>
<tr>
<th>High School</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2118</td>
<td>0.1614</td>
<td>0.1188</td>
<td>0.0932</td>
<td>0.0782</td>
<td>0.0654</td>
<td>0.0532</td>
<td>0.0554</td>
<td>0.0651</td>
<td>0.0974</td>
</tr>
<tr>
<td>2</td>
<td>0.1684</td>
<td>0.1777</td>
<td>0.1557</td>
<td>0.1213</td>
<td>0.1038</td>
<td>0.0862</td>
<td>0.0640</td>
<td>0.0516</td>
<td>0.0417</td>
<td>0.0296</td>
</tr>
<tr>
<td>3</td>
<td>0.1374</td>
<td>0.1676</td>
<td>0.1464</td>
<td>0.1390</td>
<td>0.1244</td>
<td>0.0954</td>
<td>0.0754</td>
<td>0.0577</td>
<td>0.0333</td>
<td>0.0234</td>
</tr>
<tr>
<td>4</td>
<td>0.1080</td>
<td>0.1336</td>
<td>0.1433</td>
<td>0.1378</td>
<td>0.1213</td>
<td>0.1115</td>
<td>0.0980</td>
<td>0.0746</td>
<td>0.0475</td>
<td>0.0243</td>
</tr>
<tr>
<td>5</td>
<td>0.0787</td>
<td>0.1105</td>
<td>0.1232</td>
<td>0.1335</td>
<td>0.1345</td>
<td>0.1291</td>
<td>0.1144</td>
<td>0.0862</td>
<td>0.0614</td>
<td>0.0286</td>
</tr>
<tr>
<td>6</td>
<td>0.0656</td>
<td>0.1028</td>
<td>0.1149</td>
<td>0.1201</td>
<td>0.1276</td>
<td>0.1330</td>
<td>0.1250</td>
<td>0.0998</td>
<td>0.0823</td>
<td>0.0288</td>
</tr>
<tr>
<td>7</td>
<td>0.0548</td>
<td>0.0779</td>
<td>0.0842</td>
<td>0.1097</td>
<td>0.1196</td>
<td>0.1224</td>
<td>0.1410</td>
<td>0.1331</td>
<td>0.1132</td>
<td>0.0441</td>
</tr>
<tr>
<td>8</td>
<td>0.0428</td>
<td>0.0507</td>
<td>0.0741</td>
<td>0.0880</td>
<td>0.0994</td>
<td>0.1224</td>
<td>0.1410</td>
<td>0.1585</td>
<td>0.1539</td>
<td>0.0693</td>
</tr>
<tr>
<td>9</td>
<td>0.0416</td>
<td>0.0436</td>
<td>0.0474</td>
<td>0.0577</td>
<td>0.0803</td>
<td>0.1001</td>
<td>0.1277</td>
<td>0.1728</td>
<td>0.1939</td>
<td>0.1348</td>
</tr>
<tr>
<td>10</td>
<td>0.0386</td>
<td>0.0204</td>
<td>0.0269</td>
<td>0.0292</td>
<td>0.0339</td>
<td>0.0520</td>
<td>0.0704</td>
<td>0.1155</td>
<td>0.1945</td>
<td>0.4186</td>
</tr>
</tbody>
</table>
Table 8: Ex-Ante Conditional Distributions for the NLS/1966 (College Earnings Conditional on High School Earnings)

Pr(d_i<Yc<d_i+1 | d_j<Yh<d_j+1) where d_i is the i-th decile of the College Lifetime Ex-Ante Earnings Distribution and d_j is the j-th decile of the High School Ex-Ante Lifetime Earnings Distribution.

Individual fixes unknown θ at their means, so Information Set={θ_1,θ_2,θ_3}

Correlation(Y_C,Y_H) = 0.9174

<table>
<thead>
<tr>
<th>High School</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7036</td>
<td>0.2155</td>
<td>0.0622</td>
<td>0.0137</td>
<td>0.0035</td>
<td>0.0015</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>0.2225</td>
<td>0.3780</td>
<td>0.2475</td>
<td>0.1085</td>
<td>0.0285</td>
<td>0.0110</td>
<td>0.0035</td>
<td>0.0000</td>
<td>0.0005</td>
<td>0.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.0500</td>
<td>0.2505</td>
<td>0.2960</td>
<td>0.2320</td>
<td>0.1090</td>
<td>0.0455</td>
<td>0.0120</td>
<td>0.0035</td>
<td>0.0015</td>
<td>0.0000</td>
</tr>
<tr>
<td>4</td>
<td>0.0145</td>
<td>0.1005</td>
<td>0.2250</td>
<td>0.2585</td>
<td>0.2150</td>
<td>0.1135</td>
<td>0.0545</td>
<td>0.0135</td>
<td>0.0045</td>
<td>0.0005</td>
</tr>
<tr>
<td>5</td>
<td>0.0045</td>
<td>0.0435</td>
<td>0.1055</td>
<td>0.1945</td>
<td>0.2545</td>
<td>0.2135</td>
<td>0.1265</td>
<td>0.0460</td>
<td>0.0105</td>
<td>0.0010</td>
</tr>
<tr>
<td>6</td>
<td>0.0010</td>
<td>0.0115</td>
<td>0.0435</td>
<td>0.1190</td>
<td>0.2035</td>
<td>0.2455</td>
<td>0.2100</td>
<td>0.1335</td>
<td>0.0295</td>
<td>0.0030</td>
</tr>
<tr>
<td>7</td>
<td>0.0000</td>
<td>0.0030</td>
<td>0.0150</td>
<td>0.0500</td>
<td>0.1190</td>
<td>0.2185</td>
<td>0.2705</td>
<td>0.2095</td>
<td>0.1040</td>
<td>0.0105</td>
</tr>
<tr>
<td>8</td>
<td>0.0005</td>
<td>0.0000</td>
<td>0.0055</td>
<td>0.0200</td>
<td>0.0555</td>
<td>0.1085</td>
<td>0.2080</td>
<td>0.3125</td>
<td>0.2460</td>
<td>0.0435</td>
</tr>
<tr>
<td>9</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0005</td>
<td>0.0035</td>
<td>0.0105</td>
<td>0.0380</td>
<td>0.1045</td>
<td>0.2390</td>
<td>0.3920</td>
<td>0.2120</td>
</tr>
<tr>
<td>10</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0005</td>
<td>0.0010</td>
<td>0.0045</td>
<td>0.0105</td>
<td>0.0425</td>
<td>0.2115</td>
<td>0.7295</td>
</tr>
</tbody>
</table>
Table 9: Ex-Post Conditional Distributions for the NLS/1966 (College Earnings Conditional on High School Earnings)

Pr(d_i<Yc<d_i+1 | d_j<Yh<d_j+1) where d_i is the ith decile of the College Lifetime Ex-Ante Earnings Distribution and d_j is the jth decile of the High School Ex-Ante Lifetime Earnings Distribution

Information Set={\theta_1,\theta_2,\theta_3,\theta_4,\theta_5}

Correlation(Y_C,Y_H) = 0.6226

<table>
<thead>
<tr>
<th>High School</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4001</td>
<td>0.1813</td>
<td>0.1023</td>
<td>0.0717</td>
<td>0.0611</td>
<td>0.0406</td>
<td>0.0422</td>
<td>0.0306</td>
<td>0.0337</td>
<td>0.0364</td>
</tr>
<tr>
<td>2</td>
<td>0.2144</td>
<td>0.2239</td>
<td>0.1663</td>
<td>0.1207</td>
<td>0.0862</td>
<td>0.0676</td>
<td>0.0486</td>
<td>0.0261</td>
<td>0.0256</td>
<td>0.0205</td>
</tr>
<tr>
<td>3</td>
<td>0.1286</td>
<td>0.1716</td>
<td>0.1591</td>
<td>0.1496</td>
<td>0.1181</td>
<td>0.0960</td>
<td>0.0695</td>
<td>0.0515</td>
<td>0.0340</td>
<td>0.0220</td>
</tr>
<tr>
<td>4</td>
<td>0.0870</td>
<td>0.1426</td>
<td>0.1551</td>
<td>0.1576</td>
<td>0.1386</td>
<td>0.1131</td>
<td>0.0810</td>
<td>0.0650</td>
<td>0.0365</td>
<td>0.0235</td>
</tr>
<tr>
<td>5</td>
<td>0.0450</td>
<td>0.0905</td>
<td>0.1390</td>
<td>0.1400</td>
<td>0.1395</td>
<td>0.1165</td>
<td>0.0960</td>
<td>0.0625</td>
<td>0.0305</td>
<td>0.0305</td>
</tr>
<tr>
<td>6</td>
<td>0.0350</td>
<td>0.0720</td>
<td>0.1126</td>
<td>0.1196</td>
<td>0.1456</td>
<td>0.1416</td>
<td>0.1306</td>
<td>0.1211</td>
<td>0.0900</td>
<td>0.0320</td>
</tr>
<tr>
<td>7</td>
<td>0.0210</td>
<td>0.0600</td>
<td>0.0710</td>
<td>0.1046</td>
<td>0.1201</td>
<td>0.1521</td>
<td>0.1466</td>
<td>0.1531</td>
<td>0.1126</td>
<td>0.0590</td>
</tr>
<tr>
<td>8</td>
<td>0.0205</td>
<td>0.0320</td>
<td>0.0455</td>
<td>0.0816</td>
<td>0.0951</td>
<td>0.1261</td>
<td>0.1562</td>
<td>0.1797</td>
<td>0.1667</td>
<td>0.0966</td>
</tr>
<tr>
<td>9</td>
<td>0.0180</td>
<td>0.0205</td>
<td>0.0305</td>
<td>0.0430</td>
<td>0.0755</td>
<td>0.0830</td>
<td>0.1476</td>
<td>0.1741</td>
<td>0.2316</td>
<td>0.1761</td>
</tr>
<tr>
<td>10</td>
<td>0.0125</td>
<td>0.0115</td>
<td>0.0235</td>
<td>0.0135</td>
<td>0.0225</td>
<td>0.0415</td>
<td>0.0611</td>
<td>0.1041</td>
<td>0.2077</td>
<td>0.5020</td>
</tr>
</tbody>
</table>
Evolution of variance of unforecastable components–high school sector

![Graph showing the evolution of variance of unforecastable components for high school sector. The x-axis represents age, and the y-axis represents Ten Thousand Dollars. Two lines are plotted: one for NLS/1966 and another for NLSY/1979.](image_url)
For each schooling level $s$, at each age $t$, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X \beta_{s,t} + \theta \alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector $\theta$ contains 5 elements. We test and cannot reject that the agents know the factors $\theta_1, \theta_2$, and $\theta_3$ but they don’t know factors $\theta_4, \theta_5$, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector $\theta$ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors $\theta_1, \theta_2$, and $\theta_3$ but they don’t know factors $\theta_4, \theta_5, \theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t} \theta_4 + \alpha_{5,s,t} \theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t} \theta_4 + \alpha_{5,s,t} \theta_5 + \alpha_{6,s,t} \theta_6 + \varepsilon_{s,t}$. In Figure 1, we compare the variance of $P_{s,t}$ from NLS/1966 (the solid curve) with the one from NLSY/1979 (the dashed curve) at different ages of the individuals who are high-school graduates. We see that until age 27, the estimated variance of $P_{s,t}$ from NLS/1966 and NLSY/1979 are very similar, but from age 28 on, the variance of $P_{s,t}$ from NLSY/1979 is much larger than the counterpart from NLS/1966.
Evolution of variance of unforecastable components–college sector

![Graph showing the evolution of variance of unforecastable components in the college sector over age. The graph plots age on the x-axis and ten thousand dollars on the y-axis. Two lines represent NLS/1966 and NLSY/1979 datasets, illustrating variance changes over time.](image-url)
For each schooling level $s$, at each age $t$, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector $\theta$ contains 5 elements. We test and cannot reject that the agents know the factors $\theta_1$, $\theta_2$, and $\theta_3$ but they don’t know factors $\theta_4$, $\theta_5$, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, \ldots, 41$. For the NLSY/1979 data set, the vector $\theta$ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors $\theta_1$, $\theta_2$, and $\theta_3$ but they don’t know factors $\theta_4$, $\theta_5$, $\theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, \ldots, 41$. Let $P_{s,t}$ denote the unforecastable components at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. In Figure 2, we compare the variance of $P_{s,t}$ from NLS/1966 (the solid curve) with the one from NLSY/1979 (the dashed curve) at different ages of the individuals who are college graduates. We see that until age 30, the estimated variance of $P_{s,t}$ from NLS/1966 and NLSY/1979 are very similar, but from age 31 on, the variance of $P_{s,t}$ from NLSY/1979 is much larger than the counterpart from NLS/1966.
One period correlation in unforecastable component of earnings—high school sample
For each schooling level \( s \), at each age \( t \), we model earnings \( Y_{s,t} \) according to:

\[
Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}
\]

For the NLS/1966 data set, the vector \( \theta \) contains 5 elements. We test and cannot reject that the agents know the factors \( \theta_1, \theta_2, \) and \( \theta_3 \) but they don’t know factors \( \theta_4, \theta_5, \) and \( \varepsilon_{s,t} \) at the time of their schooling choice, for \( s = 0, 1 \) and \( t = 22, \ldots, 41 \). For the NLSY/1979 data set, the vector \( \theta \) contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors \( \theta_1, \theta_2, \) and \( \theta_3 \) but they don’t know factors \( \theta_4, \theta_5, \theta_6 \) and \( \varepsilon_{s,t} \) at the time of their schooling choice, for \( s = 0, 1 \) and \( t = 22, \ldots, 41 \). Let \( P_{s,t} \) denote the unforecastable components at the time of the schooling choice. For the NLS/1966, \( P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t} \). For the NLSY/1979, \( P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t} \). Let \( \phi(s,t) \) denote the correlation between \( P_{s,t} \) and \( P_{s,t+1} \):

\[
\phi(s,t) = \text{Corr}(P_{s,t}, P_{s,t+1})
\]

In Figure 3, we plot \( \phi(s,t) \) from NLS/1966 (the solid curve) with the one from NLSY/1979 (the dashed curve) at different ages of the individuals who are high-school graduates. We see that for both NLS/1966 and NLSY/1979, \( \phi(s,t) \) tend to increase at earlier ages (from age 26 to age 30).
One period correlation in unforecastable component of earnings—college sample
For each schooling level $s$, at each age $t$, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector $\theta$ contains 5 elements. We test and cannot reject that the agents know the factors $\theta_1, \theta_2,$ and $\theta_3$ but they don’t know factors $\theta_4, \theta_5,$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector $\theta$ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors $\theta_1, \theta_2,$ and $\theta_3$ but they don’t know factors $\theta_4, \theta_5, \theta_6$ and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. Let $\phi(s,t)$ denote the correlation between $P_{s,t}$ and $P_{s,t+1}$:

$$\phi(s,t) = \text{Corr}(P_{s,t}, P_{s,t+1})$$

In Figure 4, we plot $\phi(s,t)$ from NLS/1966 (the solid curve) with the one from NLSY/1979 (the dashed curve) at different ages of the individuals who are college graduates. We see that for NLS/1966 follows a hump-shaped profile, but not so much the one-period correlation for the NLSY/1979.
One period correlation in unforecastable component of earnings—overall sample

![Graph showing the correlation of earnings over age with NLS/1966 and NLSY/1979 data.]
For each schooling level $s$, at each age $t$, we model earnings $Y_{s,t}$ according to:

$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

The overall earnings at age $t$, $Y_t$, is defined as:

$$Y_t = S (X\beta_{1,t} + \theta\alpha_{1,t} + \varepsilon_{1,t}) + (1 - S) (X\beta_{0,t} + \theta\alpha_{0,t} + \varepsilon_{0,t})$$

For the NLS/1966 data set, the vector $\theta$ contains 5 elements. We test and cannot reject that the agents know the factors $\theta_1, \theta_2$, and $\theta_3$ but they don’t know factors $\theta_4, \theta_5$, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. For the NLSY/1979 data set, the vector $\theta$ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors $\theta_1, \theta_2$, and $\theta_3$ but they don’t know factors $\theta_4, \theta_5, \theta_6$, and $\varepsilon_{s,t}$ at the time of their schooling choice, for $s = 0, 1$ and $t = 22, ..., 41$. Let $P_{s,t}$ denote the unforecastable components at the time of the schooling choice. For the NLS/1966, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \varepsilon_{s,t}$. For the NLSY/1979, $P_{s,t} = \alpha_{4,s,t}\theta_4 + \alpha_{5,s,t}\theta_5 + \alpha_{6,s,t}\theta_6 + \varepsilon_{s,t}$. Define $P_t$ as the unforecastable component for overall earnings at the time of the schooling choice:

$$P_t = S (\alpha_{4,1,t}\theta_4 + \alpha_{5,1,t}\theta_5 + \varepsilon_{1,t}) + (1 - S) (\alpha_{4,0,t}\theta_4 + \alpha_{5,0,t}\theta_5 + \varepsilon_{0,t})$$

Let $\phi(t)$ denote the correlation between $P_t$ and $P_{t+1}$:
### Counterfactual Simulations

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated (NLSY/1979)</td>
<td>709.7487</td>
<td>507.2910</td>
<td>906.0066</td>
</tr>
<tr>
<td>Counterfactual Economy 1</td>
<td>567.9719</td>
<td>333.8619</td>
<td>815.5019</td>
</tr>
<tr>
<td>Counterfactual Economy 2</td>
<td>701.6375</td>
<td>471.7632</td>
<td>862.3676</td>
</tr>
<tr>
<td>Counterfactual Economy 3</td>
<td>474.8947</td>
<td>294.3037</td>
<td>377.0409</td>
</tr>
</tbody>
</table>

### Percentage Change

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated (NLSY/1979)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Counterfactual Economy 1</td>
<td>-0.1998</td>
<td>-0.3419</td>
<td>-0.0999</td>
</tr>
<tr>
<td>Counterfactual Economy 2</td>
<td>-0.0114</td>
<td>-0.0700</td>
<td>-0.0482</td>
</tr>
<tr>
<td>Counterfactual Economy 3</td>
<td>-0.3309</td>
<td>-0.4199</td>
<td>-0.5838</td>
</tr>
</tbody>
</table>

### Variance of Unforecastable Components in NLSY/1979

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated (NLSY/1979)</td>
<td>372.3509</td>
<td>272.3596</td>
<td>432.8733</td>
</tr>
<tr>
<td>Counterfactual Economy 1</td>
<td>349.4675</td>
<td>205.6334</td>
<td>637.9935</td>
</tr>
<tr>
<td>Counterfactual Economy 2</td>
<td>364.2397</td>
<td>236.8318</td>
<td>389.2343</td>
</tr>
<tr>
<td>Counterfactual Economy 3</td>
<td>192.8493</td>
<td>137.3770</td>
<td>350.4976</td>
</tr>
</tbody>
</table>

### Percentage Change

<table>
<thead>
<tr>
<th></th>
<th>College</th>
<th>High School</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated (NLSY/1979)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Counterfactual Economy 1</td>
<td>-0.0615</td>
<td>-0.2450</td>
<td>0.4739</td>
</tr>
<tr>
<td>Counterfactual Economy 2</td>
<td>-0.0218</td>
<td>-0.1304</td>
<td>-0.1008</td>
</tr>
<tr>
<td>Counterfactual Economy 3</td>
<td>-0.4821</td>
<td>-0.4956</td>
<td>-0.1903</td>
</tr>
</tbody>
</table>
For each schooling level $s$, at each age $t$, we model earnings as:

$$Y_{s,t}^h = X \beta_{s,t}^h + \theta^h \alpha_{s,t} + \varepsilon_{s,t}^h$$

where we introduce the superscript $h$ to make our explanation of Table 5 clearer. For the NLSY/1979, we fit a six-factor model, while for the NLS/1966, we fit a five-factor model. For survey $h$, the present value of earnings from ages 22 through 41 in schooling level $s$ is:

$$Y_s^h = \sum_{t=22}^{41} \left( X \beta_{s,t}^h + \sum_{k=1}^{K_h} \theta^h_k \alpha_{k,s,t}^h + \varepsilon_{s,t}^h \right) \frac{1}{(1 + \rho)^{t-22}}, \ s = 0, 1; h = 66, 79; K_h = 5 \text{ if } h = 66, \ K_h = 6 \text{ if } h = 79.$$

The total residual variance at schooling level $s$ in NLSY/1979 is $Q_{s}^{79}$ as:

$$Q_{s}^{79} = \sum_{k=1}^{6} \text{Var} \left( \theta_k^{79} \right) \left( \sum_{t=22}^{41} \frac{\alpha_{k,s,t}^{79}}{(1 + \rho)^{t-22}} \right)^2 + \sum_{t=22}^{41} \frac{\text{Var} \left( \varepsilon_{s,t}^{79} \right)}{(1 + \rho)^{t-22}}$$

(1)

Given our estimated information set, the variance of unforecastable components at the time of the schooling choice of an individual in the NLSY/1979 sample, $P_{s}^{79}$, is:

$$P_{s}^{79} = \sum_{k=4}^{6} \text{Var} \left( \theta_k^{79} \right) \left( \sum_{t=22}^{41} \frac{\alpha_{k,s,t}^{79}}{(1 + \rho)^{t-22}} \right)^2 + \sum_{t=22}^{41} \frac{\text{Var} \left( \varepsilon_{s,t}^{79} \right)}{(1 + \rho)^{t-22}}$$

(2)

The counterfactual economy 1 is simulated as the economic environment where the distribution of the factors in NLSY/1979 were exactly the same as in NLS/1966. In this counterfactual economy we would compute $Q_{s}^{79}$ and $P_{s}^{79}$ exactly as above, except that we would replace $\text{Var} \left( \theta_k^{79} \right)$ with $\text{Var} \left( \theta_k^{66} \right)$, for $k = 1, 2, ..., 6$ and fixing $\text{Var} \left( \theta_6^{66} \right) = 0$.

The counterfactual economy 2 is the economy where the distribution of the shocks $\varepsilon_{s,t}^{79}$ are the same as $\varepsilon_{s,t}^{66}$. In this case, we would compute $Q_{s}^{79}$ and $P_{s}^{79}$ exactly as above but replacing $\text{Var} \left( \varepsilon_{s,t}^{79} \right)$ with $\text{Var} \left( \varepsilon_{s,t}^{66} \right)$.

Finally, counterfactual economy 3 is the economy where the factor loadings $\alpha_{k,s,t}^{79}$ are the same as $\alpha_{k,s,t}^{66}$. We can obtain $Q_{s}^{79}$ and $P_{s}^{79}$ after replacing $\alpha_{k,s,t}^{79}$ with $\alpha_{k,s,t}^{66}$. 

---

**Counterfactual Simulations**

<table>
<thead>
<tr>
<th>Counterfactual Economy</th>
<th>Total Residual Variance</th>
<th>Variance of Unforecastable Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counterfactual Economy 1</td>
<td>349.4675</td>
<td>205.6334</td>
</tr>
<tr>
<td>Counterfactual Economy 2</td>
<td>364.2397</td>
<td>236.8318</td>
</tr>
<tr>
<td>Counterfactual Economy 3</td>
<td>474.8947</td>
<td>294.3037</td>
</tr>
<tr>
<td>Estimated (NLSY/1979)</td>
<td>372.3509</td>
<td>272.3596</td>
</tr>
</tbody>
</table>

---

For each schooling level $s$, at each age $t$, we model earnings as:

$$Y_{s,t}^h = X \beta_{s,t}^h + \theta^h \alpha_{s,t} + \varepsilon_{s,t}^h$$

where we introduce the superscript $h$ to make our explanation of Table 5 clearer. For the NLSY/1979, we fit a six-factor model, while for the NLS/1966, we fit a five-factor model. For survey $h$, the present value of earnings from ages 22 through 41 in schooling level $s$ is:

$$Y_s^h = \sum_{t=22}^{41} \left( X \beta_{s,t}^h + \sum_{k=1}^{K_h} \theta^h_k \alpha_{k,s,t}^h + \varepsilon_{s,t}^h \right) \frac{1}{(1 + \rho)^{t-22}}, \ s = 0, 1; h = 66, 79; K_h = 5 \text{ if } h = 66, \ K_h = 6 \text{ if } h = 79.$$

The total residual variance at schooling level $s$ in NLSY/1979 is $Q_{s}^{79}$ as:

$$Q_{s}^{79} = \sum_{k=1}^{6} \text{Var} \left( \theta_k^{79} \right) \left( \sum_{t=22}^{41} \frac{\alpha_{k,s,t}^{79}}{(1 + \rho)^{t-22}} \right)^2 + \sum_{t=22}^{41} \frac{\text{Var} \left( \varepsilon_{s,t}^{79} \right)}{(1 + \rho)^{t-22}}$$

(1)

Given our estimated information set, the variance of unforecastable components at the time of the schooling choice of an individual in the NLSY/1979 sample, $P_{s}^{79}$, is:

$$P_{s}^{79} = \sum_{k=4}^{6} \text{Var} \left( \theta_k^{79} \right) \left( \sum_{t=22}^{41} \frac{\alpha_{k,s,t}^{79}}{(1 + \rho)^{t-22}} \right)^2 + \sum_{t=22}^{41} \frac{\text{Var} \left( \varepsilon_{s,t}^{79} \right)}{(1 + \rho)^{t-22}}$$

(2)

The counterfactual economy 1 is simulated as the economic environment where the distribution of the factors in NLSY/1979 were exactly the same as in NLS/1966. In this counterfactual economy we would compute $Q_{s}^{79}$ and $P_{s}^{79}$ exactly as above, except that we would replace $\text{Var} \left( \theta_k^{79} \right)$ with $\text{Var} \left( \theta_k^{66} \right)$, for $k = 1, 2, ..., 6$ and fixing $\text{Var} \left( \theta_6^{66} \right) = 0$.

The counterfactual economy 2 is the economy where the distribution of the shocks $\varepsilon_{s,t}^{79}$ are the same as $\varepsilon_{s,t}^{66}$. In this case, we would compute $Q_{s}^{79}$ and $P_{s}^{79}$ exactly as above but replacing $\text{Var} \left( \varepsilon_{s,t}^{79} \right)$ with $\text{Var} \left( \varepsilon_{s,t}^{66} \right)$.

Finally, counterfactual economy 3 is the economy where the factor loadings $\alpha_{k,s,t}^{79}$ are the same as $\alpha_{k,s,t}^{66}$. We can obtain $Q_{s}^{79}$ and $P_{s}^{79}$ after replacing $\alpha_{k,s,t}^{79}$ with $\alpha_{k,s,t}^{66}$. 

---

**Counterfactual Simulations**

<table>
<thead>
<tr>
<th>Counterfactual Economy</th>
<th>Total Residual Variance</th>
<th>Variance of Unforecastable Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counterfactual Economy 1</td>
<td>349.4675</td>
<td>205.6334</td>
</tr>
<tr>
<td>Counterfactual Economy 2</td>
<td>364.2397</td>
<td>236.8318</td>
</tr>
<tr>
<td>Counterfactual Economy 3</td>
<td>474.8947</td>
<td>294.3037</td>
</tr>
<tr>
<td>Estimated (NLSY/1979)</td>
<td>372.3509</td>
<td>272.3596</td>
</tr>
</tbody>
</table>
Evolution of cognitive skill prices – high school sector
Evolution of cognitive skill prices – college sector
We discussed the estimation of the distribution of treatment effects (*ex post* or under perfect certainty).
We discussed the estimation of the distribution of treatment effects (*ex post* or under perfect certainty).

Show how to extract uncertainty facing agents.
Summary

- We discussed the estimation of the distribution of treatment effects (*ex post* or under perfect certainty).

- Show how to extract uncertainty facing agents.

- We use schooling choices to infer the agent information sets at the time of the schooling choice.
Summary

- We discussed the estimation of the distribution of treatment effects (ex post or under perfect certainty).

- Show how to extract uncertainty facing agents.

- We use schooling choices to infer the agent information sets at the time of the schooling choice.

- A number of papers has used this strategy to separate heterogeneity from uncertainty: Carneiro, Hansen and Heckman (2003), Cunha, Heckman, and Navarro (2005), Navarro (2005), Cunha and Heckman (2006a).
The main idea: choices agents make are source of information about what they know and act on.
The main idea: choices agents make are source of information about what they know and act on.

Using more choices allows us to make less strict econometric assumptions. For example, Cunha and Heckman (2006b) show that we can model “uncertainty” better by looking at different risks people face. In particular, we can break the assumption that \( \varepsilon \) is independent over time (important for quantitative results of incomplete markets as in Aiyagari, 1994).
Recent work by Schennach (2004) allows us to break new ground. Her work does not require the strong independence assumptions as Carneiro, Hansen, and Heckman (2003), so we can study aggregate shocks.
Recent work by Schennach (2004) allows us to break new ground. Her work does not require the strong independence assumptions as Carneiro, Hansen, and Heckman (2003), so we can study aggregate shocks.

See Cunha, Heckman and Schennach (2006a,b).
An economic model

- Agents live for $T$ periods.
An economic model

- Agents live for \( T \) periods.
- In each period there is a realization of a stochastic event \( \omega_t \in \Omega \).
An economic model

- Agents live for $T$ periods.
- In each period there is a realization of a stochastic event $\omega_t \in \Omega$.
- Let the histories of events up to and until time $t$ be denoted $\omega^t = \{\omega_1, \omega_2, \ldots, \omega_t\}$.
An economic model

- Agents live for $T$ periods.
- In each period there is a realization of a stochastic event $\omega_t \in \Omega$.
- Let the histories of events up to and until time $t$ be denoted $\omega^t = \{\omega_1, \omega_2, \ldots, \omega_t\}$.
- The unconditional probability of a particular sequence of events $\omega^t$ is denoted $\pi_t(\omega^t)$. 
An economic model

- Agents live for $T$ periods.
- In each period there is a realization of a stochastic event $\omega_t \in \Omega$.
- Let the histories of events up to and until time $t$ be denoted $\omega^t = \{\omega_1, \omega_2, \ldots, \omega_t\}$.
- The unconditional probability of a particular sequence of events $\omega^t$ is denoted $\pi_t(\omega^t)$.
- In the first period, before any stochastic event is realized, agents choose schooling level $S$ and how to allocate consumption across states of nature and over time.
An economic model

- Agents live for $T$ periods.
- In each period there is a realization of a stochastic event $\omega_t \in \Omega$.
- Let the histories of events up to and until time $t$ be denoted $\omega^t = \{\omega_1, \omega_2, \ldots, \omega_t\}$.
- The unconditional probability of a particular sequence of events $\omega^t$ is denoted $\pi_t(\omega^t)$.
- In the first period, before any stochastic event is realized, agents choose schooling level $S$ and how to allocate consumption across states of nature and over time.
- Let $Y_{st}(\omega^t)$ denote the productivity of agent with schooling level $s$ given history $\omega^t$. 
An economic model

- $c_{st}(\omega^t)$ is consumption of an agent with schooling level $s$ at period $t$ and history $\omega^t$. 
An economic model

- $c_{st}(\omega^t)$ is consumption of an agent with schooling level $s$ at period $t$ and history $\omega^t$.
- $q_t(\omega^t)$ is price of an AD security that delivers one unit of period-$t$ consumption good if the history $\omega^t$ is realized and zero otherwise.
An economic model

- $c_{st}(\omega^t)$ is consumption of an agent with schooling level $s$ at period $t$ and history $\omega^t$.
- $q_t(\omega^t)$ is price of an AD security that delivers one unit of period-$t$ consumption good if the history $\omega^t$ is realized and zero otherwise.
- The productivity $Y_{st}(\omega^t)$ has a stochastic component.
An economic model

- $c_{st}(ω^t)$ is consumption of an agent with schooling level $s$ at period $t$ and history $ω^t$.
- $q_t(ω^t)$ is price of an AD security that delivers one unit of period-$t$ consumption good if the history $ω^t$ is realized and zero otherwise.
- The productivity $Y_{st}(ω^t)$ has a stochastic component.
- There is no aggregate uncertainty. All uncertainty is idiosyncratic.
An economic model

- $c_{st}(\omega^t)$ is consumption of an agent with schooling level $s$ at period $t$ and history $\omega^t$.
- $q_t(\omega^t)$ is price of an AD security that delivers one unit of period-$t$ consumption good if the history $\omega^t$ is realized and zero otherwise.
- The productivity $Y_{st}(\omega^t)$ has a stochastic component.
- There is no aggregate uncertainty. All uncertainty is idiosyncratic.
- Consumption goods can be produced according to a constant returns to scale technology that depends only on labor.
Given schooling choice $s$, the consumption allocation problem of the agent for preference function is

$$V(s) = \text{Max } E \left[ \sum_{t=1}^{T} \left( \frac{1}{1 + \rho} \right)^t u(c_{s,t}) \right]$$

subject to:

$$\sum_{t=1}^{T} \sum_{\omega^t} q_t(\omega^t) c_{s,t}^{i}(\omega^t) = \sum_{t=1}^{T} \sum_{\omega^t} q_t(\omega^t) Y_{s,t}(\omega^t)$$

Lagrange multiplier is $\lambda_s$. 
The first-order condition is:

\[ \lambda_s q_t (\omega^t) = \left( \frac{1}{1 + \rho} \right)^t \pi_t (\omega^t) u' \left[ c_{s,t}^i (\omega^t) \right] \]
The first-order condition is:

\[ \lambda_s q_t (\omega^t) = \left( \frac{1}{1 + \rho} \right)^t \pi_t (\omega^t) \ u' \left[ c_{s,t} (\omega^t) \right] \]

No aggregate uncertainty implies the equilibrium consumption allocation must be such that:

\[ c_{s,t} (\omega^t) = c_s \]
The first-order condition is:

$$\lambda_s q_t (\omega^t) = \left( \frac{1}{1 + \rho} \right)^t \pi_t (\omega^t) u' \left[ c_{s,t} (\omega^t) \right]$$

No aggregate uncertainty implies the equilibrium consumption allocation must be such that:

$$c_{s,t} (\omega^t) = c_s$$

It is easy to show that

$$c_s = A(\rho) E \left( \sum_t \left( \frac{1}{1 + \rho} \right)^t Y_{s,t} \right) \quad (20)$$
We can use (20) in (18) to calculate lifetime utility of schooling level $s$:

$$V(s) = \frac{1}{A(\rho)} u \left[ A(\rho) E \left( \sum_{t} \left( \frac{1}{1 + \rho} \right)^t Y_{s,t} \right | I \right]$$
Let $C$ denote the psychic costs associated with schooling choices.
Let $C$ denote the psychic costs associated with schooling choices.

Let $I$ denote the utility of going to college:

$$I = E \left\{ V(1) - V(0) - \bar{C} \mid \mathcal{I} \right\}$$

$$= \frac{1}{A(\rho)} E \left\{ u \left[ E \left( A(\rho) \sum_t \left( \frac{1}{1+\rho} \right)^t Y_{1,t} \bigg\vert \mathcal{I} \right) \right] - u \left[ E \left( A(\rho) \sum_t \left( \frac{1}{1+\rho} \right)^t Y_{0,t} \bigg\vert \mathcal{I} \right) \right] - C \right\}$$
### Summary Statistics - NLSY/1979 and PSID

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations High School Sample</th>
<th>Observations College Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observations</td>
<td>Mean</td>
</tr>
<tr>
<td>Mother's Education</td>
<td>1045</td>
<td>3.8852</td>
</tr>
<tr>
<td>Father's Education</td>
<td>1045</td>
<td>3.8699</td>
</tr>
<tr>
<td>Number of Siblings</td>
<td>1045</td>
<td>3.1129</td>
</tr>
<tr>
<td>Urban Residence at age 14</td>
<td>1045</td>
<td>0.7378</td>
</tr>
<tr>
<td>Local Tuition at 4-year college</td>
<td>1045</td>
<td>0.2154</td>
</tr>
<tr>
<td>Year of Birth is 1958</td>
<td>1045</td>
<td>0.1120</td>
</tr>
<tr>
<td>Year of Birth is 1959</td>
<td>1045</td>
<td>0.1321</td>
</tr>
<tr>
<td>Year of Birth is 1960</td>
<td>1045</td>
<td>0.1416</td>
</tr>
<tr>
<td>Year of Birth is 1961</td>
<td>1045</td>
<td>0.1254</td>
</tr>
<tr>
<td>Year of Birth is 1962</td>
<td>1045</td>
<td>0.1397</td>
</tr>
<tr>
<td>Year of Birth is 1963</td>
<td>1045</td>
<td>0.1091</td>
</tr>
<tr>
<td>Year of Birth is 1964</td>
<td>1045</td>
<td>0.1120</td>
</tr>
<tr>
<td>Enrolled at School at ASVAB Test Date</td>
<td>538</td>
<td>0.4628</td>
</tr>
<tr>
<td>Age at ASVAB Test Date</td>
<td>538</td>
<td>19.3457</td>
</tr>
<tr>
<td>Highest Grade Completed at ASVAB Test Date</td>
<td>538</td>
<td>11.0074</td>
</tr>
<tr>
<td>ASVAB - Arithmetic Reasoning</td>
<td>538</td>
<td>-0.4783</td>
</tr>
<tr>
<td>ASVAB - Word Knowledge</td>
<td>538</td>
<td>-0.4310</td>
</tr>
<tr>
<td>ASVAB - Paragraph Composition</td>
<td>538</td>
<td>-0.4413</td>
</tr>
<tr>
<td>ASVAB - Coding Speed</td>
<td>538</td>
<td>-0.3277</td>
</tr>
<tr>
<td>ASVAB - Math Knowledge</td>
<td>538</td>
<td>-0.6211</td>
</tr>
</tbody>
</table>

1. The sample consists of white males born between 1957 and 1964 who are high school or college graduates.
2. In ten thousand dollars. The tuition figures are inflation-adjusted using the CPI. The base year is 2000.
3. Not available for PSID respondents.
### Summary Statistics - NLS/1966 and PSID

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>High School</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Mean)</td>
<td>(Std. Error)</td>
<td>Mean (Mean)</td>
</tr>
<tr>
<td>Mother's Education</td>
<td>1117</td>
<td>3.3160</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>1.1607</td>
<td>1.5406</td>
<td></td>
</tr>
<tr>
<td>Father's Education</td>
<td>1117</td>
<td>2.9910</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>1.2798</td>
<td>1.9110</td>
<td></td>
</tr>
<tr>
<td>Number of Siblings</td>
<td>1117</td>
<td>3.1791</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>2.2042</td>
<td>1.7389</td>
<td></td>
</tr>
<tr>
<td>Urban Residence at age 14</td>
<td>1117</td>
<td>0.6634</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.4728</td>
<td>0.3893</td>
<td></td>
</tr>
<tr>
<td>Local Tuition at 4-year college²</td>
<td>1117</td>
<td>0.1577</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.0215</td>
<td>0.0214</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1942</td>
<td>1117</td>
<td>0.0645</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.2457</td>
<td>0.2362</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1943</td>
<td>1117</td>
<td>0.0770</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.2667</td>
<td>0.2377</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1944</td>
<td>1117</td>
<td>0.0609</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.2392</td>
<td>0.2552</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1945</td>
<td>1117</td>
<td>0.0430</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.2029</td>
<td>0.2647</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1946</td>
<td>1117</td>
<td>0.0546</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.2273</td>
<td>0.2774</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1947</td>
<td>1117</td>
<td>0.0985</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.2981</td>
<td>0.3392</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1948</td>
<td>1117</td>
<td>0.1038</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.3052</td>
<td>0.3154</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1949</td>
<td>1117</td>
<td>0.1182</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.3230</td>
<td>0.3154</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1950</td>
<td>1117</td>
<td>0.1334</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.3402</td>
<td>0.2906</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1951</td>
<td>1117</td>
<td>0.1343</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.3411</td>
<td>0.3050</td>
<td></td>
</tr>
<tr>
<td>Year of Birth is 1952</td>
<td>1117</td>
<td>0.0618</td>
<td>1215</td>
</tr>
<tr>
<td></td>
<td>0.2408</td>
<td>0.2168</td>
<td></td>
</tr>
<tr>
<td>Otis/Beta/Gamma Test³</td>
<td>194</td>
<td>-0.5667</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>0.8026</td>
<td>0.7971</td>
<td></td>
</tr>
<tr>
<td>California Test of Mental Maturity³</td>
<td>123</td>
<td>-0.4120</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>0.8964</td>
<td>0.8267</td>
<td></td>
</tr>
<tr>
<td>Work Knowledge, Occupations³</td>
<td>769</td>
<td>0.0854</td>
<td>785</td>
</tr>
<tr>
<td></td>
<td>0.8333</td>
<td>0.7619</td>
<td></td>
</tr>
<tr>
<td>Work Knowledge, Education³</td>
<td>772</td>
<td>0.1044</td>
<td>787</td>
</tr>
<tr>
<td></td>
<td>0.8557</td>
<td>0.7527</td>
<td></td>
</tr>
<tr>
<td>Work Knowledge, Earnings Comparison³</td>
<td>779</td>
<td>-0.0320</td>
<td>791</td>
</tr>
<tr>
<td></td>
<td>0.9310</td>
<td>0.9334</td>
<td></td>
</tr>
</tbody>
</table>

¹ The sample consists of white males born between 1941 and 1952 who are high school or college graduates
² In ten thousand dollars. The tuition figures are inflation-adjusted using the CPI. The base year is 2000.
³ Not available for PSID respondents.
<table>
<thead>
<tr>
<th>Age</th>
<th>High School Sample</th>
<th>College Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observations</td>
<td>Mean</td>
</tr>
<tr>
<td>22</td>
<td>715</td>
<td>2.0630</td>
</tr>
<tr>
<td>23</td>
<td>759</td>
<td>2.2700</td>
</tr>
<tr>
<td>24</td>
<td>795</td>
<td>2.4175</td>
</tr>
<tr>
<td>25</td>
<td>803</td>
<td>2.5949</td>
</tr>
<tr>
<td>26</td>
<td>811</td>
<td>2.7158</td>
</tr>
<tr>
<td>27</td>
<td>829</td>
<td>2.8010</td>
</tr>
<tr>
<td>28</td>
<td>835</td>
<td>3.0294</td>
</tr>
<tr>
<td>29</td>
<td>831</td>
<td>3.1140</td>
</tr>
<tr>
<td>30</td>
<td>814</td>
<td>3.1827</td>
</tr>
<tr>
<td>31</td>
<td>763</td>
<td>3.4152</td>
</tr>
<tr>
<td>32</td>
<td>755</td>
<td>3.3891</td>
</tr>
<tr>
<td>33</td>
<td>658</td>
<td>3.6220</td>
</tr>
<tr>
<td>34</td>
<td>652</td>
<td>3.6570</td>
</tr>
<tr>
<td>35</td>
<td>533</td>
<td>3.7389</td>
</tr>
<tr>
<td>36</td>
<td>530</td>
<td>3.8459</td>
</tr>
<tr>
<td>37</td>
<td>418</td>
<td>3.8590</td>
</tr>
<tr>
<td>38</td>
<td>405</td>
<td>3.9803</td>
</tr>
<tr>
<td>39</td>
<td>341</td>
<td>4.1229</td>
</tr>
<tr>
<td>40</td>
<td>268</td>
<td>4.2145</td>
</tr>
<tr>
<td>41</td>
<td>233</td>
<td>4.3307</td>
</tr>
</tbody>
</table>

1 In ten thousand dollars. The earnings figures are inflation-adjusted using the CPI. The base year is 2000.
2 The sample consists of white males born between 1957 and 1964 who are high school or college graduates.
<table>
<thead>
<tr>
<th>Age</th>
<th>High School Sample</th>
<th>College Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observations</td>
<td>Mean</td>
</tr>
<tr>
<td>22</td>
<td>435</td>
<td>2.7109</td>
</tr>
<tr>
<td>23</td>
<td>530</td>
<td>2.9363</td>
</tr>
<tr>
<td>24</td>
<td>601</td>
<td>3.0219</td>
</tr>
<tr>
<td>25</td>
<td>674</td>
<td>3.2760</td>
</tr>
<tr>
<td>26</td>
<td>646</td>
<td>3.3748</td>
</tr>
<tr>
<td>27</td>
<td>653</td>
<td>3.4915</td>
</tr>
<tr>
<td>28</td>
<td>622</td>
<td>3.5399</td>
</tr>
<tr>
<td>29</td>
<td>605</td>
<td>3.6311</td>
</tr>
<tr>
<td>30</td>
<td>593</td>
<td>3.6730</td>
</tr>
<tr>
<td>31</td>
<td>516</td>
<td>3.5821</td>
</tr>
<tr>
<td>32</td>
<td>491</td>
<td>3.7744</td>
</tr>
<tr>
<td>33</td>
<td>432</td>
<td>3.8722</td>
</tr>
<tr>
<td>34</td>
<td>382</td>
<td>3.8831</td>
</tr>
<tr>
<td>35</td>
<td>327</td>
<td>4.0330</td>
</tr>
<tr>
<td>36</td>
<td>280</td>
<td>3.9929</td>
</tr>
<tr>
<td>37</td>
<td>301</td>
<td>4.0008</td>
</tr>
<tr>
<td>38</td>
<td>273</td>
<td>3.8033</td>
</tr>
<tr>
<td>39</td>
<td>247</td>
<td>3.8198</td>
</tr>
<tr>
<td>40</td>
<td>207</td>
<td>3.7984</td>
</tr>
<tr>
<td>41</td>
<td>189</td>
<td>3.8534</td>
</tr>
</tbody>
</table>

1 In ten thousand dollars. The earnings figures are inflation-adjusted using the CPI. The base year is 2000.

2 The sample consists of white males born between 1941 and 1952 who are high school or college graduates.
### Raw Correlation of Test Scores from NLS/1966

<table>
<thead>
<tr>
<th></th>
<th>Otis/Beta/Gamma Test</th>
<th>California Test of Mental Maturity</th>
<th>Work Knowledge, Occupations</th>
<th>Work Knowledge, Education and Occupation</th>
<th>Work Knowledge, Earnings Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otis/Beta/Gamma Test</td>
<td>1.0000</td>
<td>N/A²</td>
<td>0.4289</td>
<td>0.4457</td>
<td>0.0765</td>
</tr>
<tr>
<td>California Test of Mental Maturity</td>
<td>N/A²</td>
<td>1.0000</td>
<td>0.2036</td>
<td>0.1714</td>
<td>0.1112</td>
</tr>
<tr>
<td>Work Knowledge, Occupations</td>
<td>0.4289</td>
<td>0.2036</td>
<td>1.0000</td>
<td>0.9374</td>
<td>0.4464</td>
</tr>
<tr>
<td>Work Knowledge, Education</td>
<td>0.4457</td>
<td>0.1714</td>
<td>0.9374</td>
<td>1.0000</td>
<td>0.1068</td>
</tr>
<tr>
<td>Work Knowledge, Earnings Comparison</td>
<td>0.0765</td>
<td>0.1112</td>
<td>0.4464</td>
<td>0.1068</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

1. We control for mother's and father's education, urban residency at age 14, and year of birth.

2. Individuals report either Otis/Beta/Gamma or the California Test of Mental Maturity, but not both.
### Normalizations on Factor Loadings: NLSY/1979,2

#### High School Earnings Equations

<table>
<thead>
<tr>
<th>Age</th>
<th>Loading on Factor 1</th>
<th>Loading on Factor 2</th>
<th>Loading on Factor 3</th>
<th>Loading on Factor 4</th>
<th>Loading on Factor 5</th>
<th>Loading on Factor 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>23</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>24</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>26</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>28</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>29</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>30</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>31</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>32</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>34</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>35</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>36</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>37</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>38</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>39</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>40</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>41</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loading on Factor 1</th>
<th>Loading on Factor 2</th>
<th>Loading on Factor 3</th>
<th>Loading on Factor 4</th>
<th>Loading on Factor 5</th>
<th>Loading on Factor 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### College Earnings Equations

The empty cells correspond to factor loadings that are estimated, not normalized.
Normalizations on Factor Loadings: NLSY/1966\(^1\)

<table>
<thead>
<tr>
<th>Age</th>
<th>High School Earnings Equations</th>
<th>College Earnings Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Loading on Factor 1</td>
<td>Loading on Factor 2</td>
</tr>
<tr>
<td>22</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>23</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>24</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>25</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>26</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>27</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>28</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>29</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>30</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>31</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>32</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>33</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>34</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>35</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>36</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>37</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>38</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>39</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>40</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>41</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

\(^1\) The empty cells correspond to factor loadings that are estimated, not normalized.
### $\chi^2$ Goodness of Fit Test*

#### NLS/1966 - White Males

<table>
<thead>
<tr>
<th>Age</th>
<th>High School</th>
<th>College</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$ statistic</td>
<td>Critical Value</td>
<td>$\chi^2$ statistic</td>
</tr>
<tr>
<td>22</td>
<td>15.9253</td>
<td>20.7143</td>
<td>20.4226</td>
</tr>
<tr>
<td>23</td>
<td>34.2904</td>
<td>35.1725</td>
<td>26.0629</td>
</tr>
<tr>
<td>24</td>
<td>36.6979</td>
<td>40.1133</td>
<td>78.7201</td>
</tr>
<tr>
<td>25</td>
<td>29.5390</td>
<td>43.7730</td>
<td>48.0949</td>
</tr>
<tr>
<td>26</td>
<td>40.5503</td>
<td>42.5570</td>
<td>36.2933</td>
</tr>
<tr>
<td>27</td>
<td>47.3649</td>
<td>50.9985</td>
<td>49.0855</td>
</tr>
<tr>
<td>28</td>
<td>39.4947</td>
<td>47.3999</td>
<td>34.6742</td>
</tr>
<tr>
<td>29</td>
<td>32.8477</td>
<td>42.5570</td>
<td>23.7741</td>
</tr>
<tr>
<td>30</td>
<td>31.7484</td>
<td>40.1133</td>
<td>28.8962</td>
</tr>
<tr>
<td>31</td>
<td>29.0581</td>
<td>36.4150</td>
<td>38.1067</td>
</tr>
<tr>
<td>32</td>
<td>29.9466</td>
<td>35.1725</td>
<td>53.5870</td>
</tr>
<tr>
<td>33</td>
<td>28.8073</td>
<td>32.6706</td>
<td>33.5289</td>
</tr>
<tr>
<td>34</td>
<td>27.0961</td>
<td>30.1435</td>
<td>43.5183</td>
</tr>
<tr>
<td>35</td>
<td>29.6717</td>
<td>26.2962</td>
<td>33.9107</td>
</tr>
<tr>
<td>36</td>
<td>18.9902</td>
<td>22.3620</td>
<td>26.2794</td>
</tr>
<tr>
<td>37</td>
<td>21.6758</td>
<td>22.3620</td>
<td>31.1112</td>
</tr>
<tr>
<td>38</td>
<td>14.4640</td>
<td>21.0261</td>
<td>22.3595</td>
</tr>
<tr>
<td>39</td>
<td>18.4237</td>
<td>21.0261</td>
<td>23.7976</td>
</tr>
<tr>
<td>40</td>
<td>17.4722</td>
<td>19.6751</td>
<td>25.3994</td>
</tr>
</tbody>
</table>

* 95% Confidence, equiprobable bins with approx. 20 people per bin. A $\chi^2$ statistic lower than the critical value indicates a "good" fit.
### χ² Goodness of Fit Test*

**NLSY/1979 - White Males**

<table>
<thead>
<tr>
<th>Period</th>
<th>χ² statistic</th>
<th>Critical Value</th>
<th>χ² statistic</th>
<th>Critical Value</th>
<th>χ² statistic</th>
<th>Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>57.3492</td>
<td>47.3999</td>
<td>39.7989</td>
<td>35.1725</td>
<td>167.6037</td>
<td>77.9305</td>
</tr>
<tr>
<td>23</td>
<td>42.6356</td>
<td>50.9985</td>
<td>49.0880</td>
<td>36.4150</td>
<td>152.2525</td>
<td>81.3810</td>
</tr>
<tr>
<td>24</td>
<td>52.0194</td>
<td>52.1923</td>
<td>197.1642</td>
<td>42.5570</td>
<td>259.3508</td>
<td>84.8206</td>
</tr>
<tr>
<td>25</td>
<td>42.4359</td>
<td>46.1943</td>
<td>94.7095</td>
<td>43.7730</td>
<td>175.4185</td>
<td>84.8206</td>
</tr>
<tr>
<td>26</td>
<td>62.4019</td>
<td>53.3835</td>
<td>70.2619</td>
<td>43.7730</td>
<td>157.0157</td>
<td>89.3912</td>
</tr>
<tr>
<td>27</td>
<td>43.6728</td>
<td>52.1923</td>
<td>42.3386</td>
<td>44.9853</td>
<td>180.5021</td>
<td>91.6702</td>
</tr>
<tr>
<td>28</td>
<td>54.9250</td>
<td>48.6024</td>
<td>40.9173</td>
<td>46.1943</td>
<td>174.9548</td>
<td>91.6702</td>
</tr>
<tr>
<td>29</td>
<td>49.1212</td>
<td>48.6024</td>
<td>36.1557</td>
<td>46.1943</td>
<td>128.6085</td>
<td>89.3912</td>
</tr>
<tr>
<td>30</td>
<td>50.4962</td>
<td>50.9985</td>
<td>41.6969</td>
<td>47.3999</td>
<td>135.4479</td>
<td>89.3912</td>
</tr>
<tr>
<td>31</td>
<td>49.6975</td>
<td>48.6024</td>
<td>30.6494</td>
<td>43.7730</td>
<td>119.7788</td>
<td>84.8206</td>
</tr>
<tr>
<td>32</td>
<td>44.5459</td>
<td>50.9985</td>
<td>34.5965</td>
<td>42.5570</td>
<td>142.0935</td>
<td>83.6753</td>
</tr>
<tr>
<td>33</td>
<td>35.4077</td>
<td>43.7730</td>
<td>28.7575</td>
<td>38.8851</td>
<td>148.9012</td>
<td>77.9305</td>
</tr>
<tr>
<td>34</td>
<td>40.7768</td>
<td>42.5570</td>
<td>28.4552</td>
<td>38.8851</td>
<td>95.9926</td>
<td>68.6693</td>
</tr>
<tr>
<td>35</td>
<td>41.8859</td>
<td>36.4150</td>
<td>30.6125</td>
<td>33.9244</td>
<td>68.3855</td>
<td>62.8296</td>
</tr>
<tr>
<td>36</td>
<td>36.2069</td>
<td>38.8851</td>
<td>39.1018</td>
<td>30.1435</td>
<td>91.1547</td>
<td>62.8296</td>
</tr>
<tr>
<td>37</td>
<td>34.6365</td>
<td>31.4104</td>
<td>21.0079</td>
<td>27.5871</td>
<td>73.0057</td>
<td>52.1923</td>
</tr>
<tr>
<td>38</td>
<td>24.2197</td>
<td>28.8693</td>
<td>20.5837</td>
<td>27.5871</td>
<td>40.6111</td>
<td>49.8018</td>
</tr>
<tr>
<td>39</td>
<td>29.6366</td>
<td>27.5871</td>
<td>29.2055</td>
<td>28.8693</td>
<td>60.9063</td>
<td>42.5570</td>
</tr>
</tbody>
</table>

* 95% Confidence, equiprobable bins with approx. 20 people per bin. A χ² statistic lower than the critical value indicates a "good" fit.
### Test of Equality of Predicted versus Actual Correlation

**Matrices of Earnings (from ages 22 to 41)**

**NLSY/1979 and NLS/1966**

<table>
<thead>
<tr>
<th></th>
<th>High School</th>
<th>College</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLS/1966 - 5 Factors</td>
<td>15.6968</td>
<td>210.4133</td>
<td>114.8754</td>
</tr>
<tr>
<td>NLS/1979 - 6 Factors</td>
<td>70.6451</td>
<td>156.5446</td>
<td>187.5425</td>
</tr>
<tr>
<td>NLS/1979 - 5 Factors</td>
<td>64.2682</td>
<td>309.2815</td>
<td>226.2401</td>
</tr>
<tr>
<td>Critical Value*</td>
<td>222.0741</td>
<td>222.0741</td>
<td>222.0741</td>
</tr>
</tbody>
</table>

* *95% Confidence*