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 II focused on IV and showed the relationship between IV and structural selection models in environments with essential heterogeneity: $(Y_1 - Y_0) \not\perp\!\!\!\perp D \mid X$.
- It showed how to obtain marginal returns not just average returns.
- It focused on means.
- Methods for identifying means can also identify marginal distributions.

$$E\left[\mathbf{1}(Y_1 \leq y_1) \mid X\right] = F\left(y_1 \mid X\right)$$
$$E\left[\mathbf{1}(Y_0 \leq y_0) \mid X\right] = F\left(y_0 \mid X\right)$$

- 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:
    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.
- 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 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
$$Y^p = D^p Y_1^p + (1 - D^p)\, Y_0^p$$
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 $\left(Y_0^p \text{ and } Y_1^p;\ Y_0^{p'} \text{ and } Y_1^{p'}\right)$.
- It only compares $F(y^p \mid X)$ and $F(y^{p'} \mid X)$ 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.

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

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

  where Υ 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.

- For the Roy model,

$$D = \mathbf{1}\left[\Upsilon(Y_1) \geq \Upsilon(Y_0)\right].$$

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

$$\Pr(D = 1 \mid \mathcal{I}_E) = \Pr(\Upsilon(Y_1) \geq \Upsilon(Y_0) \mid \mathcal{I}_E).$$

- If the agent's information set is the same as the econometrician's and uses the choice rule $D = \mathbf{1}\left[\Upsilon\left(Y_1\right) \geq \Upsilon\left(Y_0\right)\right]$, then observed choice proportions identify $\Pr\left(D = 1 \mid \mathcal{I}_E\right) = \Pr\left(\Upsilon\left(Y_1\right) \geq \Upsilon\left(Y_0\right) \mid \mathcal{I}_E\right)$.
- But analyses of objective evaluations often condition on information sets other than $\mathcal{I}_E$.
- Need the full joint distribution to compute e.g., $\Pr\left(Y_1 > Y_0\right)$ (the fraction who benefit *ex post*).

- The inequalities of Hoeffding (1940) and Fréchet (1951) state that

$$\begin{aligned} \max\left[F_0\left(y_0 \mid X\right) + F_1\left(y_1 \mid X\right) - 1, 0\right] \\ \leq F\left(y_0, y_1 \mid X\right) \\ \leq \min\left[F_0\left(y_0 \mid X\right), F_1\left(y_1 \mid X\right)\right]. \end{aligned}$$

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

- Example from Heckman, Smith, and Clements (1997).
- 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:

Untreated

Treated		E	N	
	E	P_{EE}	P_{EN}	$P_{E\cdot}$
	N	P_{NE}	P_{NN}	$P_{N\cdot}$
		$P_{\cdot E}$	$P_{\cdot N}$	

- $P_{E\cdot}$, $P_{N\cdot}$, $P_{\cdot E}$, $P_{\cdot N}$ obtained from experiment.

- Estimates of the marginals of the table parameters:

$$P_{E\cdot} = P_{EE} + P_{EN}$$

(employment proportion among the treated)

$$P_{\cdot E} = P_{EE} + P_{NE}$$

(employment proportion among the untreated)

- The treatment effect is usually defined as

$$\begin{aligned} \Delta &= P_{EN} - P_{NE} \qquad \text{net effect} && (1) \\ &= P_{E\cdot} - P_{\cdot E} && (2) \end{aligned}$$

Fréchet-Hoeffding Bounds

$$\max\left[P_{E\cdot} + P_{\cdot E} - 1, 0\right] \leq P_{EE} \leq \min\left[P_{E\cdot}, P_{\cdot E}\right]$$

$$\max\left[P_{E\cdot} - P_{\cdot E}, 0\right] \leq P_{EN} \leq \min\left[P_{E\cdot}, 1 - P_{\cdot E}\right]$$

$$\max\left[-P_{E\cdot} + P_{\cdot E}, 0\right] \leq P_{NE} \leq \min\left[1 - P_{E\cdot}, P_{\cdot E}\right]$$

$$\max\left[1 - P_{E\cdot} - P_{\cdot E}, 0\right] \leq P_{NN} \leq \min\left[1 - P_{E\cdot}, 1 - P_{\cdot E}\right]$$

Fraction Employed in the 16th, 17th or 18th Month after Random Assignment and Fréchet-Hoeffding Bounds on the Probabilities P_{NE} and P_{EN}

National JTPA Study 18 Month Impact Sample Adult Females	
Parameter	Estimate
Fraction employed in the treatment group	0.64
	(0.01)
Fraction employed in the control group	0.61
	(0.01)
Bounds on P_{EN}	[0.03,0.39]
	(0.01),(0.01)
Bounds on P_{NE}	[0.00,0.36]
	(0.00),(0.01)

- 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

$$
\begin{aligned}
&F(y_0, y_1 \mid X) \\
&\quad = \int F_0(y_0 \mid X, q)\, F_1(y_1 \mid X, q)\; d\mu(q \mid X).
\end{aligned}
$$

- Conditional on X, Y_0 and Y_1 are assumed to be deterministically related:

$$Y_1 - Y_0 = \Delta \tag{3}$$

where Δ 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:

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

$$F_0(y_0 \mid X) = \int \tilde{M}(y_0, y_1 \mid X) dF_1(y_1 \mid X).$$

$$
\begin{aligned}
Y_1 &= \mu_1(X) + U_1 \qquad E(U_1 \mid X) = 0 \\
Y_0 &= \mu_0(X) + U_0 \qquad E(U_0 \mid X) = 0
\end{aligned}
$$

$$
\begin{aligned}
Y &= \mu_0(X) + \underbrace{(\mu_1(X) - \mu_0(X) + U_1 - U_0)}_{\beta(X)} D + U_0 \qquad (4) \\
&= \mu_0(X) + (\mu_1(X) - \mu_0(X))\, D + (U_1 - U_0)\, D + U_0 \\
&= \mu_0(X) + \bar{\beta}(X) D + D\eta + U_0
\end{aligned}
$$

$$
\begin{aligned}
\beta(X) &= \bar{\beta}(X) + \eta \\
\eta &= U_1 - U_0
\end{aligned}
$$

$$D = \mathbf{1}\left[Y_1 \geq Y_0\right] \tag{5}$$

$$\begin{aligned} Y_1 &= \mu_1(X) + U_1 \qquad & E(U_1 \mid X) = 0 \\ Y_0 &= \mu_0(X) + U_0 \qquad & E(U_0 \mid X) = 0 \end{aligned}$$

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

- See Aakvik, Heckman, and Vytlacil (2005), Carneiro, Hansen, and Heckman (2001,2003), Cunha, Heckman, and Navarro (2005,2006), and Cunha and Heckman (2006a,b).
- Assume separability:

$$\begin{aligned} Y_1 &= \mu_1(X) + U_1 \\ Y_0 &= \mu_0(X) + U_0. \end{aligned}$$

- Let I denote the latent variable generating schooling choices:

$$\begin{aligned} I &= \mu_I(Z) + U_I \\ D &= \mathbf{1}\left[I \geq 0\right]. \end{aligned}$$

- Normality assumptions make it easy to understand how the method works and can be relaxed.

- 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 θ is the factor that generates dependence across the unobservables:

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

$$\theta \perp\!\!\!\perp (\varepsilon_0, \varepsilon_1, \varepsilon_{U_I}), \quad \varepsilon_0 \perp\!\!\!\perp \varepsilon_1 \perp\!\!\!\perp \varepsilon_{U_I}.$$

- To set the scale of the unobserved factor, we normalize one "loading" (coefficient on θ) to 1.

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

$$\begin{aligned} \mathrm{Cov}\left(U_0, \frac{U_I}{\sigma_{U_I}}\right) &= \frac{\alpha_0 \alpha_{U_I}}{\sigma_{U_I}} \sigma_\theta^2 \\ \mathrm{Cov}\left(U_1, \frac{U_I}{\sigma_{U_I}}\right) &= \frac{\alpha_1 \alpha_{U_I}}{\sigma_{U_I}} \sigma_\theta^2 \end{aligned}$$

where $\sigma_{U_I}^2 = \mathrm{Var}(\varepsilon_{U_I})$.

- We obtain the sign of the dependence between U_0, U_1 because
$$\mathrm{Cov}(U_0, U_1) = \alpha_0\alpha_1\sigma_\theta^2.$$
- 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 θ)

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 ε_M is independent of ε_0, ε_1, ε_{U_I},and θ, as well as (X, Z).

$$\begin{aligned} \mathrm{Cov}(Y_1, M) &= \alpha_1 \alpha_M \sigma_\theta^2 \\ \mathrm{Cov}(Y_0, M) &= \alpha_0 \alpha_M \sigma_\theta^2 \\ \mathrm{Cov}(I, M) &= \frac{\alpha_{U_I}}{\sigma_{U_I}} \alpha_M \sigma_\theta^2 \end{aligned}$$

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

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

$$\frac{\mathrm{Cov}\,(Y_1, I)}{\mathrm{Cov}\,(I, M)} = \frac{\alpha_1 \alpha_{U_I} \sigma_\theta^2}{\alpha_{U_I} \alpha_M \sigma_\theta^2} = \alpha_1$$

because $\alpha_M = 1$.

$$\frac{\mathrm{Cov}\,(Y_1, I)}{\mathrm{Cov}\,(Y_0, I)} = \frac{\alpha_1 \alpha_{U_I} \sigma_\theta^2}{\alpha_0 \alpha_{U_I} \sigma_\theta^2} = \frac{\alpha_1}{\alpha_0}$$

- We obtain σ_θ^2 from $\mathrm{Cov}\,(Y_1, M)$ or $\mathrm{Cov}\,(Y_0, M)$.
- We obtain α_{U_I} (up to scale σ_{U_I}) from $\mathrm{Cov}\,(I, M) = \alpha_{U_I} \alpha_M \sigma_\theta^2$ since we know α_M $(= 1)$ and σ_θ^2.

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 ε_C is independent of θ.

- $U_I = U_1 - U_0 - U_C$, and

$$\begin{aligned} \alpha_{U_I} &= \alpha_1 - \alpha_0 - \alpha_C \\ \varepsilon_{U_I} &= \varepsilon_1 - \varepsilon_0 - \varepsilon_C \\ \text{Var}\left(\varepsilon_{U_I}\right) &= \text{Var}\left(\varepsilon_1\right) + \text{Var}\left(\varepsilon_0\right) + \text{Var}\left(\varepsilon_C\right). \end{aligned}$$

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

- The scale σ_{U_I} is identified if there are variables in X but not in Z.
$$\mathrm{Var}(M) - \alpha_M^2 \sigma_\theta^2 = \sigma_{\varepsilon_M}^2.$$
- We can thus construct the joint distribution of (Y_0, Y_1, C) and hence the joint distribution of (Y_0, Y_1).
- 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.

Example 2. identification without choice data

- Let I be any indicator that depends on θ 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.
- We can identify the joint distribution of (Y_0, Y_1).
- We can identify σ_{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.
- 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

$$\begin{aligned} Y_{1t} &= \mu_{1t}(X) + \alpha_{1t}\theta + \varepsilon_{1t} & t = 1, 2 \\ Y_{0t} &= \mu_{0t}(X) + \alpha_{0t}\theta + \varepsilon_{0t} & t = 1, 2. \end{aligned}$$

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

$$I = (Y_{12} + Y_{11}) - (Y_{02} + Y_{01}) - C$$
$$D = \mathbf{1}[I \geq 0],$$

$$\begin{aligned} I = \mu_{11}(X) + \mu_{12}(X) - \mu_{01}(X) - \mu_{02}(X) - \mu_C(Z) \\ + U_{11} + U_{12} - U_{01} - U_{02} - U_C. \end{aligned}$$

- 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 $(I, Y_{11}, Y_{12}, Y_{01}, Y_{02})$.

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:

$$
\begin{aligned}
\mathrm{Cov}(U_I, U_{12}) &= (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{12}\sigma_\theta^2 \\
\mathrm{Cov}(U_I, U_{11}) &= (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{11}\sigma_\theta^2 \\
\mathrm{Cov}(U_I, U_{01}) &= (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{01}\sigma_\theta^2 \\
\mathrm{Cov}(U_I, U_{02}) &= (\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C)\alpha_{02}\sigma_\theta^2 \\
\mathrm{Cov}(U_{11}, U_{12}) &= \alpha_{11}\alpha_{12}\sigma_\theta^2 \\
\mathrm{Cov}(U_{01}, U_{02}) &= \alpha_{01}\alpha_{02}\sigma_\theta^2 .
\end{aligned}
$$

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

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

$$\frac{\mathrm{Cov}(U_I, U_{12})}{\mathrm{Cov}(U_I, U_{01})} = \alpha_{12}, \qquad \frac{\mathrm{Cov}(U_I, U_{11})}{\mathrm{Cov}(U_I, U_{01})} = \alpha_{11},$$

$$\frac{\mathrm{Cov}(U_I, U_{02})}{\mathrm{Cov}(U_I, U_{01})} = \alpha_{02}.$$

$$\frac{\mathrm{Cov}(U_{11}, U_{12})}{\alpha_{11}\alpha_{12}} = \sigma_\theta^2$$

and

$$\frac{\mathrm{Cov}(U_{01}, U_{02})}{\alpha_{01}\alpha_{02}} = \sigma_\theta^2.$$

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

- We can recover σ_θ^2 (since we know $\alpha_{11}\alpha_{12}$ and $\alpha_{01}\alpha_{02}$) from $\mathrm{Cov}(U_{11}, U_{12})$ and $\mathrm{Cov}(U_{01}, U_{02})$.
- We can also recover α_C since we know σ_θ^2, $\alpha_{12} + \alpha_{11} - \alpha_{02} - \alpha_{01} - \alpha_C$, and α_{11}, α_{12}, α_{01}, α_{02}.
- We can form (conditional on X) $\mathrm{Cov}(Y_{11}, Y_{01}) = \alpha_{11}\alpha_{01}\sigma_\theta^2$; $\mathrm{Cov}(Y_{12}, Y_{01}) = \alpha_{12}\alpha_{01}\sigma_\theta^2$; $\mathrm{Cov}(Y_{11}, Y_{02}) = \alpha_{11}\alpha_{02}\sigma_\theta^2$ and $\mathrm{Cov}(Y_{12}, Y_{02}) = \alpha_{12}\alpha_{02}\sigma_\theta^2$.
- 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 $\mu_{01}(X)$, $\mu_{02}(X)$, $\mu_{11}(X)$, and $\mu_{12}(X)$.

- If the analyst knows θ 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 θ and identify the distribution of θ 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.

- We write

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

$$\begin{aligned} M_1 &= \theta + \varepsilon_1 \\ \frac{M_2}{\lambda_2} &= \theta + \varepsilon_2^* \end{aligned}$$

  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 $g_2(\varepsilon_2^*)$.

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?
  - 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, θ represent the endowments of observable and unobservable skills, respectively. The vectors β, α 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).
- 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 θ and the variance of ε has increased when one compares the period 1970-1978 with the period 1978-1987.
- They call the increase in the variance of temporary shocks ε 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).
- 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.
- 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 θ are known by the agent and which elements of θ 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.
- 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.
- 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 in the following manner:

$$Y_{0,t} = \mu_{0,t} + U_{0,t},\ \ E\left(U_{0,t}\right) = 0 \tag{6}$$

$$Y_{1,t} = \mu_{1,t} + U_{1,t},\ \ E\left(U_{1,t}\right) = 0 \tag{7}$$

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

$$C = \mu_C + U_C \tag{8}$$

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

$$M_k = \mu_k^M + U_k^M,\ \ k = 1, 2, \ldots, K.$$

- The schooling equation is based on

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

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

$$I = E\left\{ \begin{array}{l} \sum_t \left(\frac{1}{1+\rho}\right)^t (\mu_{1,t} - \mu_{0,t}) - \mu_C \\ + \sum_t \left(\frac{1}{1+\rho}\right)^t (U_{1t} - U_{0t}) - U_C \end{array} \middle| \mathcal{I} \right\}$$

- 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.,

$$E \left\{ \begin{array}{l} \sum_t \left(\frac{1}{1+\rho}\right)^t (\mu_{1,t} - \mu_{0,t}) - \mu_C \\ + \sum_t \left(\frac{1}{1+\rho}\right)^t (U_{1,t} - U_{0,t}) - U_C \end{array} \middle| \mathcal{I} \right\} \geq 0$$

- Assume that $U_C, X, Z \in \mathcal{I}$. The event $S = 1$ corresponds to the event

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

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

or, in more compact notation:

$$E\left(\bar{U}\middle|\mathcal{I}\right) - U_C \geq -\mu_I.$$

- Consequently, from data, we can compute:

$$\begin{aligned} & E\left[\, Y_{1,t} \middle| E\left(\bar{U}\middle|\mathcal{I}\right) - U_C \geq -\mu_I\right] \\ = \;\; & \mu_{1,t} + E\left[U_{1,t}\middle| E\left(U\middle|\mathcal{I}\right) - U_C \geq -\mu_I\right] \end{aligned}$$

We want to separate out two **unobservable** components:

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

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

- The component that is unknown by the agent (uncertainty):

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

- How do we formally do it?

Factor models

- Remember that the test score equations were specified as:

$$M_k = \mu_k^M + U_k^M, \;\; k = 1, 2, \ldots, K$$

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

$$M_k = \mu_k^M + \underbrace{\alpha_k^M \theta_1}_{\text{factor}} + \underbrace{\varepsilon_k^M}_{\text{uniqueness}}, \;\; k = 1, 2, \ldots, K.$$

We assume:

- $\theta_1 \sim N\left(0, \sigma^2_{\theta_1}\right)$ (normality is not necessary) and independent from ε_k^M.

- $\varepsilon_k^M \sim N\left(0, \sigma^2_{\varepsilon_k^M}\right)$ (normality is not necessary) and independent from ε_ℓ^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_ℓ^M is captured only by θ_1 for $k \neq \ell$.

- Remember that we proposed the following model for earnings equations:

$$Y_{0,t} = \mu_{0,t} + U_{0,t}$$

$$Y_{1,t} = \mu_{1,t} + U_{1,t}$$

- Now we break unobservables $U_{s,t}$ into three different components to obtain

$$Y_{0,t} = \mu_{0,t} + \underbrace{\alpha_{0,t}\theta_1 + \delta_{0,t}\theta_2}_{\text{factors}} + \underbrace{\varepsilon_{0,t}}_{\text{uniqueness}} \tag{10}$$

$$Y_{1,t} = \mu_{1,t} + \underbrace{\alpha_{1,t}\theta_1 + \delta_{1,t}\theta_2}_{\text{factors}} + \underbrace{\varepsilon_{1,t}}_{\text{uniqueness}} \tag{11}$$

We assume:

- $\theta_2 \sim N\left(0, \sigma^2_{\theta_2}\right)$ (normality is not necessary) and independent from θ_1, and $\{\varepsilon_{0,t}\}$
- $\varepsilon_{s,t} \sim N\left(0, \sigma^2_{s,t}\right)$ (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 θ_1 and θ_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

$$C = \mu_C + \alpha_C \theta_1 + \delta_C \theta_2 + \varepsilon_C, \ \varepsilon_C \sim N\left(0, \sigma_C^2\right). \quad (12)$$

- The schooling equation is generated by

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

- Given (10), (11), and (12) the schooling equation becomes:

$$
\begin{aligned}
I \;=\; & \mu_I + E\left[\theta_1\left(\sum_t\left(\frac{1}{1+\rho}\right)^t(\alpha_{1,t}-\alpha_{0,t})-\alpha_C\right)\middle|\,\mathcal{I}\right]+ \\
& +E\left[\theta_2\left(\sum_t\left(\frac{1}{1+\rho}\right)^t(\delta_{1,t}-\delta_{0,t})-\delta_C\right)\middle|\,\mathcal{I}\right]+ \\
& +E\left[\sum_t\left(\frac{1}{1+\rho}\right)^t(\varepsilon_{1,t}-\varepsilon_{0,t})-\varepsilon_C\middle|\,\mathcal{I}\right]+
\end{aligned}
$$

We assume:

- $\varepsilon_C \in \mathcal{I}$
- $\varepsilon_{s,t} \notin \mathcal{I}$ and $E\left(\varepsilon_{s,t} | \mathcal{I}\right) = 0.$
- 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 *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 $\theta_2 \notin \mathcal{I}$.
- Under (1)-(3) the schooling equation can be written as:

$$\begin{aligned} I &= \mu_I + \sum_t \left(\frac{1}{1+\rho}\right)^t \left[\left(\alpha_{1,t} - \alpha_{0,t}\right)\theta_1 + \left(\delta_{1,t} - \delta_{0,t}\right)\theta_2\right] \\ &\quad -\alpha_C\theta_1 - \delta_C\theta_2 - \varepsilon_C \end{aligned}$$

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:

$$
\begin{aligned}
& \mathrm{Cov}\left(I - \mu_I, Y_{1,1} - \mu_{1,1}\right) \\
= \; & \alpha_{1,1}\left(\sum_t \left(\frac{1}{1+\rho}\right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C\right)\sigma^2_{\theta_1} + \\
& + \delta_{1,1}\left(\sum_t \left(\frac{1}{1+\rho}\right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C\right)\sigma^2_{\theta_2}
\end{aligned}
$$

- Under H_1 :

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

- Another way to state the test is:

$$
\begin{aligned}
& \mathrm{Cov}\left(I - \mu_I, Y_{1,1} - \mu_{1,1}\right) \\
= \; & \alpha_{1,1}\left(\sum_t \left(\frac{1}{1+\rho}\right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C\right)\sigma^2_{\theta_1} + \\
& + \Delta_2 \delta_{1,1}\left(\sum_t \left(\frac{1}{1+\rho}\right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C\right)\sigma^2_{\theta_2}
\end{aligned}
$$

- $H_0 : \Delta_2 \neq 0$ versus $H_1 : \Delta_2 = 0$.
- We iterate among the alternative specifications of $\mathcal{I}$ and produce a model which fits the data best.

Identification

- How can the tests be implemented?
- Take the test score equations:

$$M_k = \mu_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M, \;\; k = 1, 2, \ldots, K. \tag{13}$$

  If $K \geq 3$ we can identify α_k^M and $f(\theta_1)$ up to a normalization (say $\alpha_1^M = 1$).
- Because of independence we can identify μ_k^M from a simple OLS regression in (13).

- We can construct the covariances:

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

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

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

- Consequently, can recover α_2^M, α_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

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

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

$$E\left(Y_{1,t} \middle| S = 1\right) = \mu_{1,t} + E\left(\alpha_{1,t}\theta_1 + \delta_{1,t}\theta_2 + \varepsilon_{1,t} \middle| S = 1\right)$$

- Assuming that the unobservables are all normal, it follows that:

$$E\left(Y_{1,t} \middle| S=1\right) = \mu_{1,t} + \pi_{1,t}\lambda \underbrace{\left(\frac{\sum_t \left(\frac{1}{1+\rho}\right)^t (\mu_{1,t} - \mu_{0,t}) - \mu_C}{\sigma_U} \right)}_{\text{selection correction}} \quad (17)$$

- 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 $\sigma^2_{\theta_1}$ from the test score equations.
- We use the covariance of earnings over time to identify the parameters associated to θ_2 :

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

- Under the normalization $\delta_{1,1} = 1$ we repeat the argument used in test scores and can recover $\delta_{s,t}$ and $\sigma^2_{\theta_2}$.
- It is interesting to note that we can then recover joint distributions:

$$\mathrm{Cov}\left(Y_{1,t}, Y_{0,\tau}\right) = \alpha_{0,\tau}\alpha_{1,t}\sigma^2_{\theta_1} + \delta_{0,\tau}\delta_{1,t}\sigma^2_{\theta_2}$$

- To identify α_C we use:

$$\begin{aligned} & \text{Cov}(M_1 - \mu_1^M, I - \mu_I) \\ = & \left(\sum_t \left(\frac{1}{1+\rho} \right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C \right) \sigma_{\theta_1}^2 \end{aligned}$$

- To identify δ_C we use:

$$
\begin{aligned}
& \mathrm{Cov}\left(Y_{1,1} - X\beta_{1,1}, I - \mu_I\right) \\
= \; & \alpha_{1,1}\left(\sum_t \left(\frac{1}{1+\rho}\right)^t (\alpha_{1,t} - \alpha_{0,t}) - \alpha_C\right)\sigma^2_{\theta_1} + \\
& + \left(\sum_{t=1} \left(\frac{1}{1+\rho}\right)^t (\delta_{1,t} - \delta_{0,t}) - \delta_C\right)\sigma^2_{\theta_2}
\end{aligned}
$$

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.

Data Description

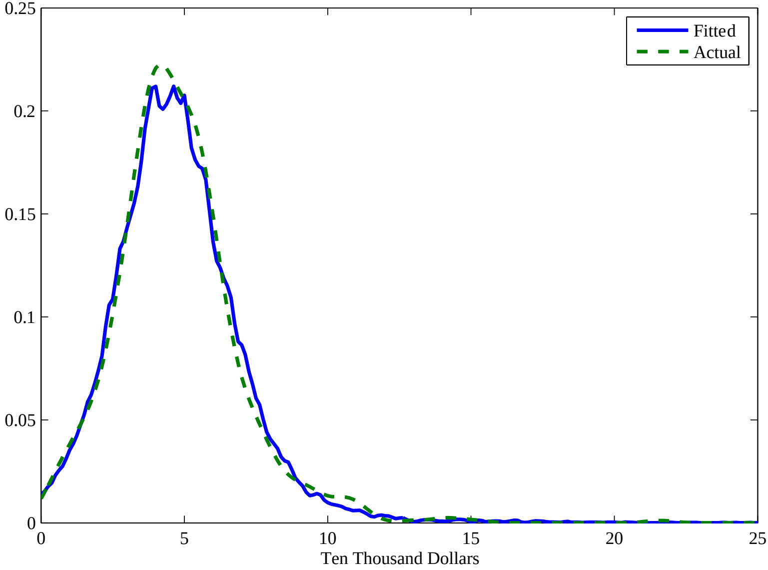
- To study the evolution of labor earnings risk in the U.S. economy we compare two different samples:
  1. NLSY/1979–The first sample consists of white males born between 1957 and 1964 and we obtain their information from NLSY/1979 data pooled their counterparts from the PSID data.
  2. NLS/1966 - The 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.

In both data sets we observe cognitive test scores:

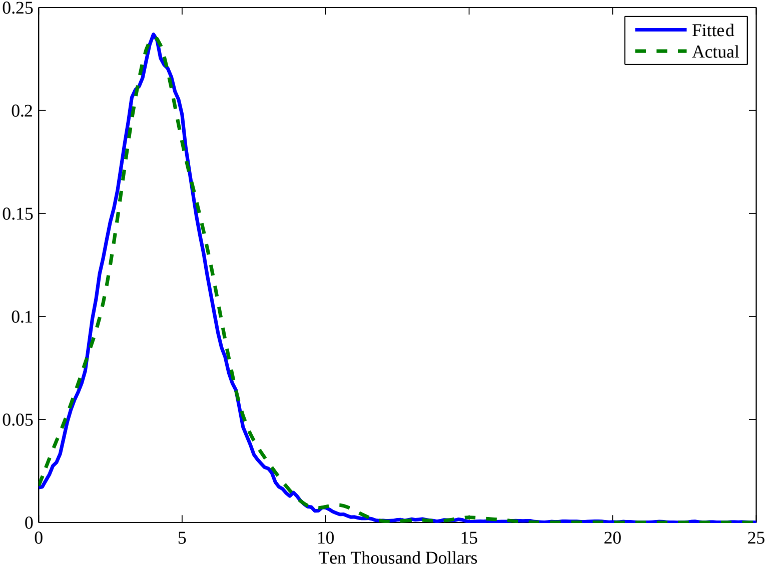
- For the NLSY/1979 we use five components of the ASVAB test battery: arithmetic reasoning, word knowledge, paragraph comprehension, math knowledge and coding speed.
- In the NLS/1966 there are many different achievement tests, but we use the two most commonly reported ones: the OTIS/BETA/GAMMA and the 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."
- 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.

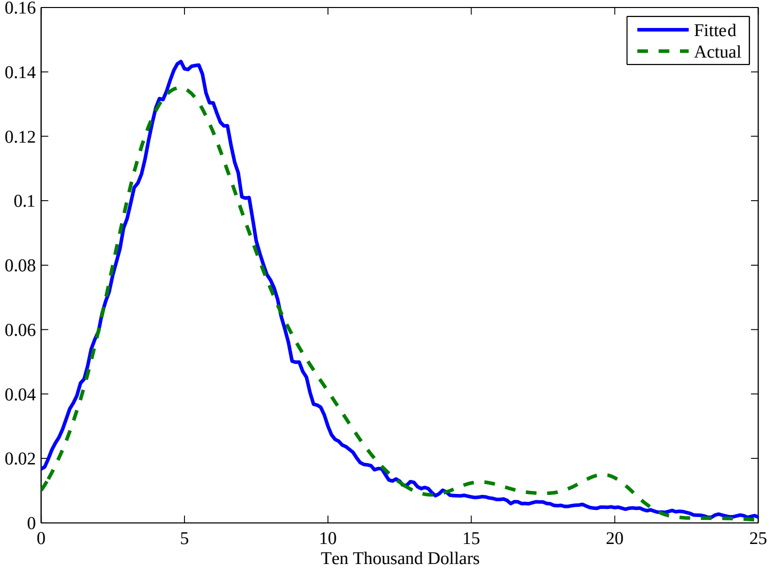
Densities of earnings at age 30 (college sample NLS/1966)



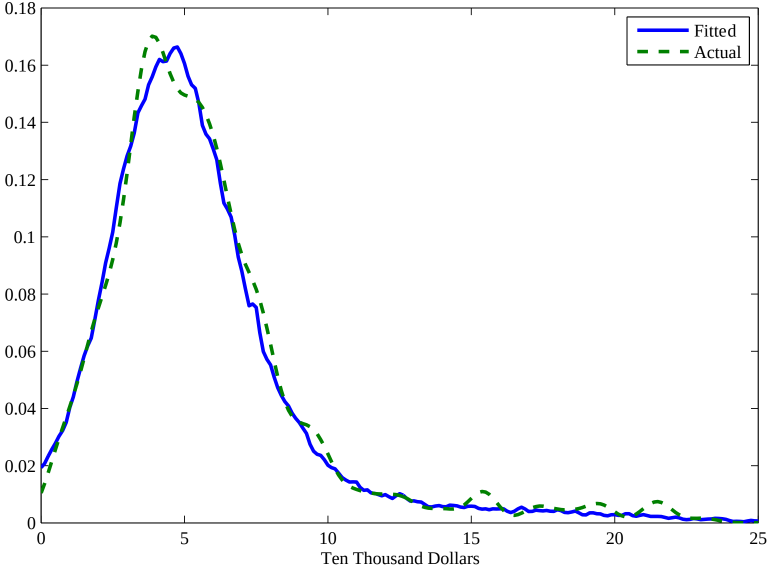
Densities of earnings at age 28 (college sample NLS/1966)



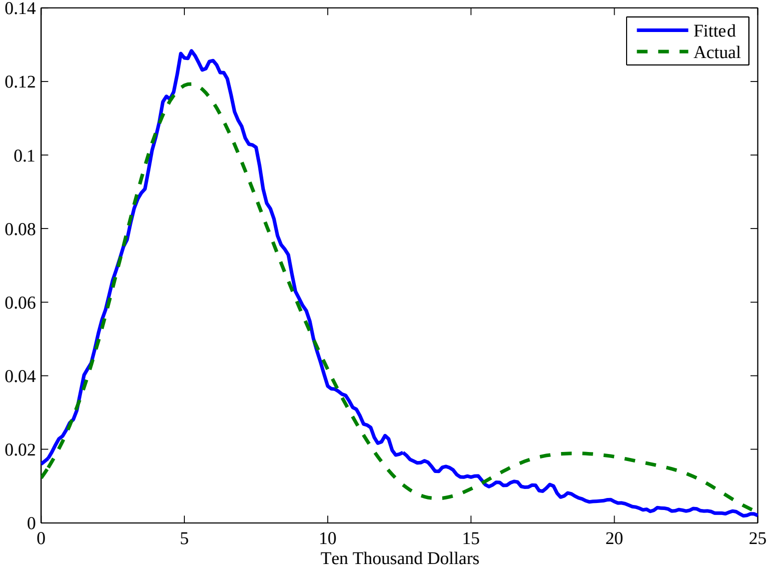
Densities of earnings at age 37 (college sample NLSY/1979)



Densities of earnings at age 24 (college sample NLSY/1979)



Densities of earnings at age 40 (college sample NLS/1979)

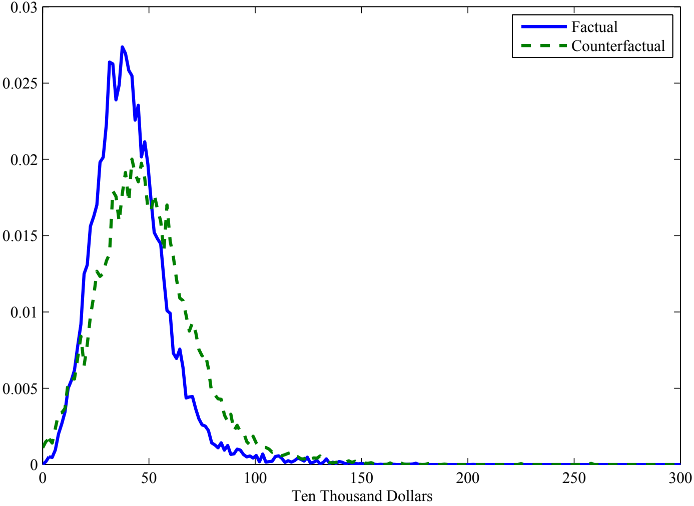


Mean Rates of Return to College by Schooling Group

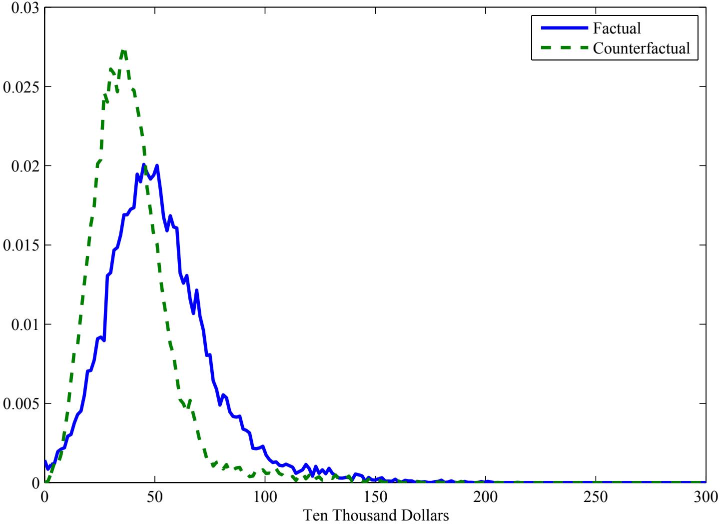
	NLS/1966		NLSY/1979	
Schooling Group	Mean Returns	Standard Error	Mean Returns	Standard Error
High School Graduates	0.2937	0.0083	0.3095	0.0113
College Graduates	0.3307	0.0114	0.3994	0.0129
Individuals at the Margin	0.3081	0.0446	0.3511	0.0535

Note: Under linearity, *ex ante* mean = *ex post* mean.

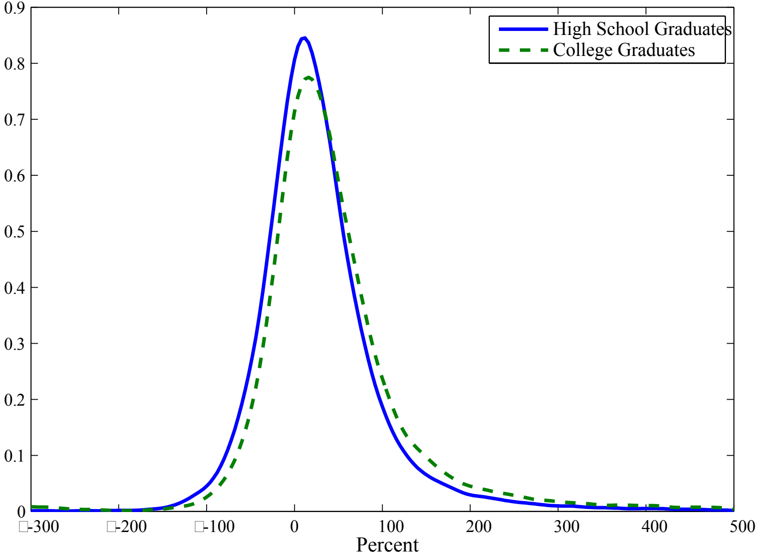
Densities of present value of *ex post* earnings–high school sample NLSY/1979



Densities of present value of *ex post* earnings–college sample NLSY/1979



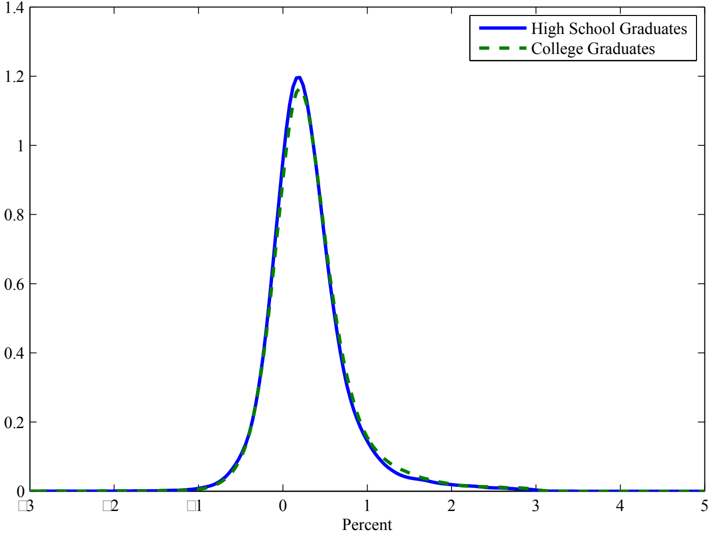
Densities of Returns to College NLSY/1979 Sample



Densities of Returns to College NLSY/1979 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 colege 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, $f(r|S=1)$.

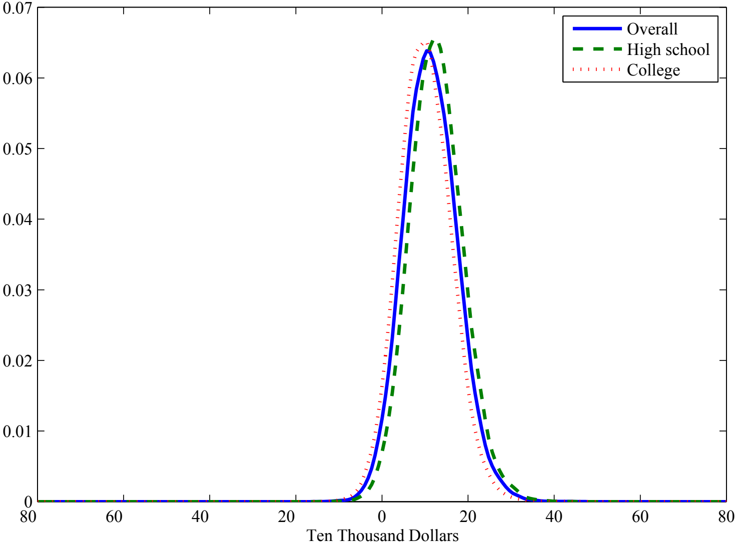
Densities of Returns to College NLS/1966 Sample



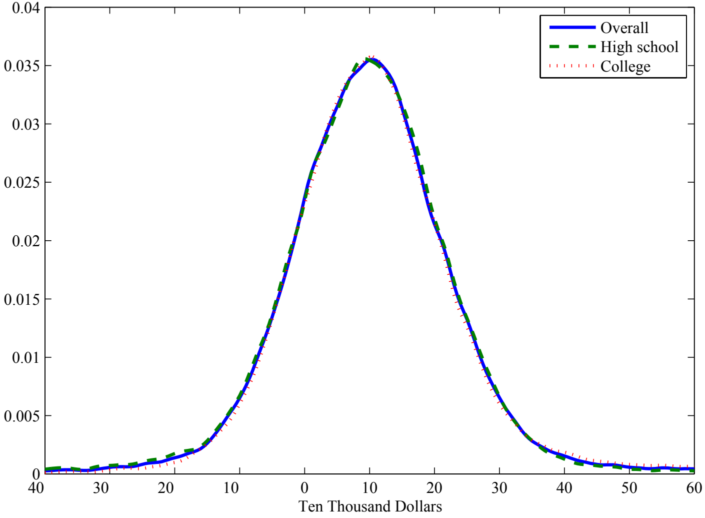
Densities of Returns to College NLS/1966 Sample

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 colege 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 monetary value of psychic cost NLS/1966



Densities of monetary value of psychic cost NLSY/1979



Evolution of Uncertainty

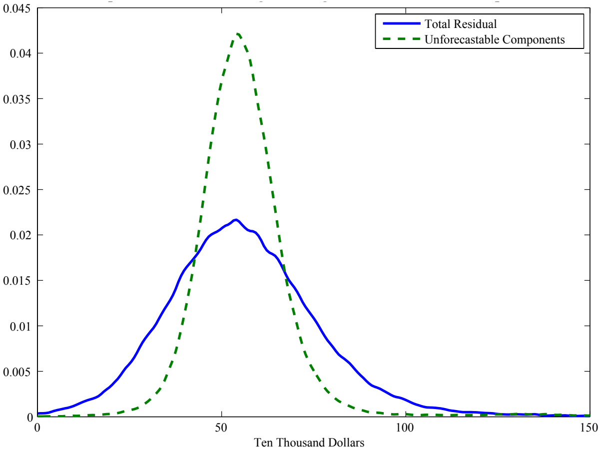
Panel A: NLS/1966			
	College	High School	Returns
Total Residual Variance	460.6260	284.8089	351.4026
Variance of Unforecastable Components	181.3712	128.4315	327.3480
Panel B: NLSY/1979			
	College	High School	Returns
Total Residual Variance	709.7487	507.2910	906.0066
Variance of Unforecastable Components	372.3509	272.3596	432.8733
Panel C: Percentage Increase			
	College	High School	Returns
Percentage Increase in Total Residual Variance	54.083%	78.116%	157.826%
Percentage Increase in Variance of Unforecastable Components	105.298%	112.066%	32.236%

Evolution of Heterogeneity (Diversity)

	College	High School	Returns
Panel A: NLS/1966			
Total Residual Variance	460.6260	284.8089	351.4026
Variance of Forecastable Components (Heterogeneity)	279.2549	156.3774	24.0546
Panel B: NLSY/1979			
	College	High School	Returns
Total Residual Variance	709.7487	507.2910	906.0066
Variance of Forecastable Components (Heterogeneity)	337.3978	234.9314	473.1333
Panel C: Percentage Increase			
	College	High School	Returns
Percentage Increase in Total Residual Variance	54.083%	78.116%	157.826%
Percentage Increase in Variance of Forecastable Components	20.821%	50.234%	1866.914%

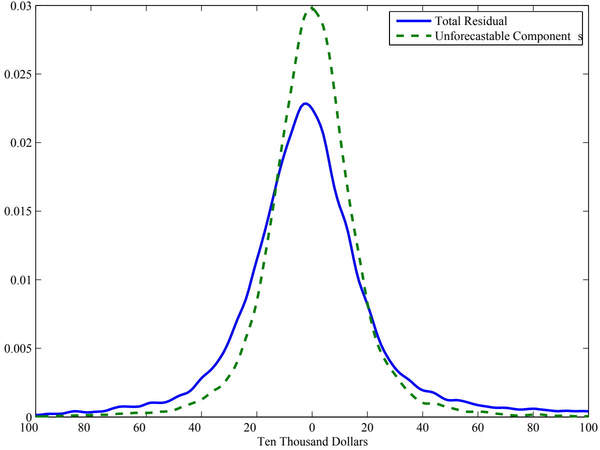
- Greater variance of returns in 1979.
- Greater predictability of returns in 1979 as a fraction of the variance.

The densities of total residual vs unforecastable components in present value of college earnings for the NLS/1966 sample



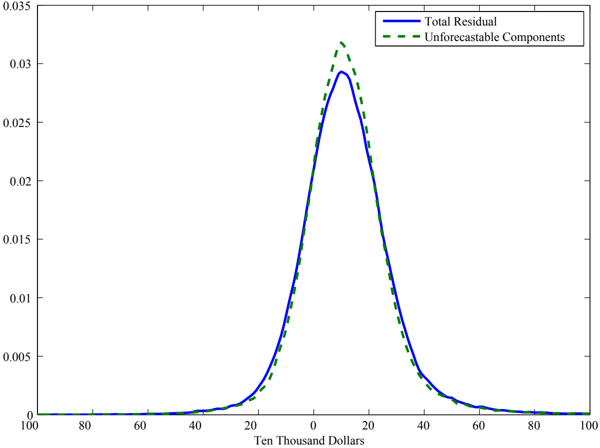
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of college earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5% interest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

Densities of total residual vs unforecastable components returns college vs high school for the NLSY/1979 sample



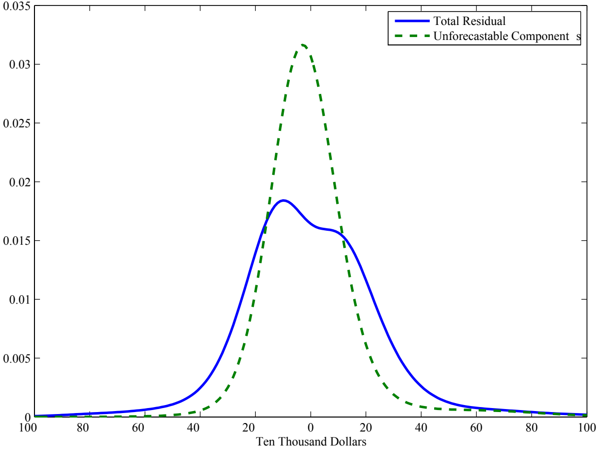
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of earnings differences (or returns to college) for the white males sample of the NLSY/1979 from ages 22 to 41. The present value of returns 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



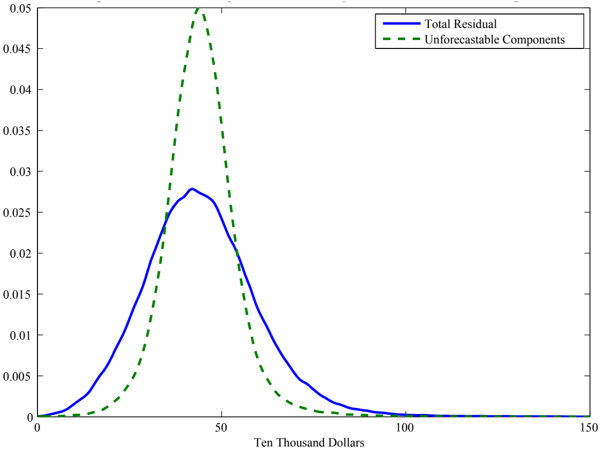
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of earnings differences (or returns to college) for the white males sample of the NLSY/1979 from ages 22 to 41. The present value of returns 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



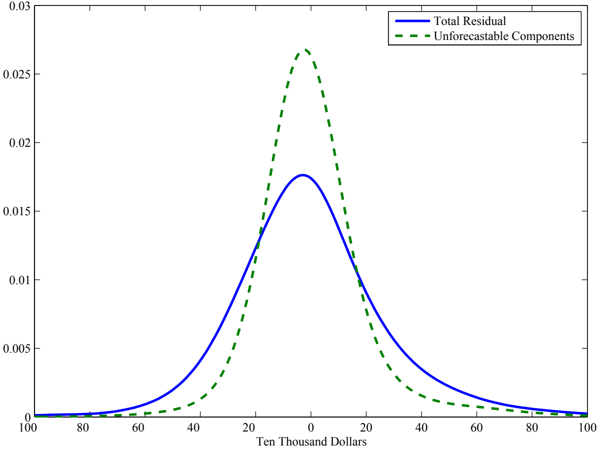
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a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

Schooling Group	NLS/1966	NLSY/1979
Percentage of High School Graduates who Regret Not Graduating from College	0.0966	0.0749
Percentage of College Graduates who Regret Graduating from College	0.0337	0.0311

Table 6: Ex-Ante Conditional Distributions for the NLSY/1979 (College Earnings Conditional on High School Earnings)
$Pr(d_i<Yc<d_i+1 \mid d_j<Yh<d_j+1)$ 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
Individual fixes unknown θ at their means, so Information Set=$\{\theta_1,\theta_2,\theta_3\}$
Corrrelation(Y_C,Y_H) = 0.1666

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

Table 7: Ex-Post Conditional Distributions for the NLSY/1979 (College Earnings Conditional on High School Earnings)
$Pr(d_i<Yc<d_i+1 \mid d_j<Yh<d_j+1)$ 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=$\{\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6\}$
Corrrelation(Y_C,Y_H) = 0.2842

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

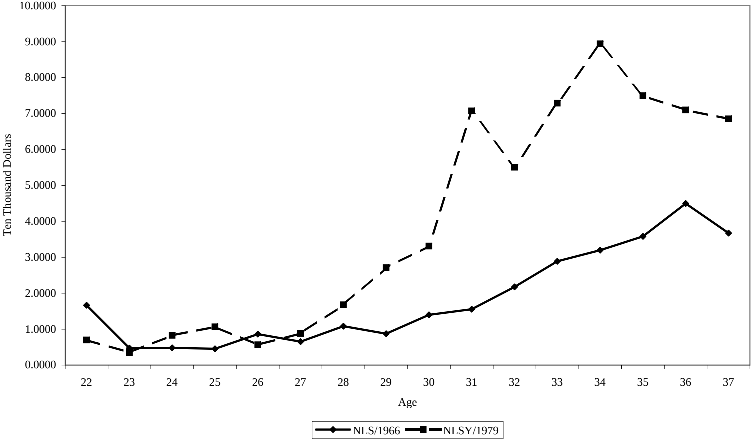
Table 8: Ex-Ante Conditional Distributions for the NLS/1966 (College Earnings Conditional on High School Earnings)
$Pr(d_i<Yc<d_i+1 \mid d_j<Yh<d_j+1)$ where d_i is the ith decile of the College Lifetime Ex-Ante Earnings Distribution and d_j is the jth decile of the High School Ex-Ante Lifetime Earnings Distribution
Individual fixes unknown θ at their means, so Information Set=$\{\theta_1,\theta_2,\theta_3\}$
Corrrelation(Y_C,Y_H) = 0.9174

	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.7036	0.2155	0.0622	0.0137	0.0035	0.0015	0.0000	0.0000	0.0000	0.0000
2	0.2225	0.3780	0.2475	0.1085	0.0285	0.0110	0.0035	0.0000	0.0005	0.0000
3	0.0500	0.2505	0.2960	0.2320	0.1090	0.0455	0.0120	0.0035	0.0015	0.0000
4	0.0145	0.1005	0.2250	0.2585	0.2150	0.1135	0.0545	0.0135	0.0045	0.0005
5	0.0045	0.0435	0.1055	0.1945	0.2545	0.2135	0.1265	0.0460	0.0105	0.0010
6	0.0010	0.0115	0.0435	0.1190	0.2035	0.2455	0.2100	0.1335	0.0295	0.0030
7	0.0000	0.0030	0.0150	0.0500	0.1190	0.2185	0.2705	0.2095	0.1040	0.0105
8	0.0005	0.0000	0.0055	0.0200	0.0555	0.1085	0.2080	0.3125	0.2460	0.0435
9	0.0000	0.0000	0.0005	0.0035	0.0105	0.0380	0.1045	0.2390	0.3920	0.2120
10	0.0000	0.0000	0.0000	0.0005	0.0010	0.0045	0.0105	0.0425	0.2115	0.7295

Table 9: Ex-Post Conditional Distributions for the NLS/1966 (College Earnings Conditional on High School Earnings)
$Pr(d_i<Yc<d_i+1 \mid d_j<Yh<d_j+1)$ 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=$\{\theta_1,\theta_2,\theta_3,\theta_4,\theta_5\}$
Corrrelation(Y_C,Y_H) = 0.6226

	College									
High School	1	2	3	4	5	6	7	8	9	10
1	0.4001	0.1813	0.1023	0.0717	0.0611	0.0406	0.0422	0.0306	0.0337	0.0364
2	0.2144	0.2239	0.1663	0.1207	0.0862	0.0676	0.0486	0.0261	0.0256	0.0205
3	0.1286	0.1716	0.1591	0.1496	0.1181	0.0960	0.0695	0.0515	0.0340	0.0220
4	0.0870	0.1426	0.1551	0.1576	0.1386	0.1131	0.0810	0.0650	0.0365	0.0235
5	0.0450	0.0905	0.1390	0.1400	0.1405	0.1395	0.1165	0.0960	0.0625	0.0305
6	0.0350	0.0720	0.1126	0.1196	0.1456	0.1416	0.1306	0.1211	0.0900	0.0320
7	0.0210	0.0600	0.0710	0.1046	0.1201	0.1521	0.1466	0.1531	0.1126	0.0590
8	0.0205	0.0320	0.0455	0.0816	0.0951	0.1261	0.1562	0.1797	0.1667	0.0966
9	0.0180	0.0205	0.0305	0.0430	0.0755	0.0830	0.1476	0.1741	0.2316	0.1761
10	0.0125	0.0115	0.0235	0.0135	0.0225	0.0415	0.0611	0.1041	0.2077	0.5020

Evolution of variance of unforecastable components–high school sector



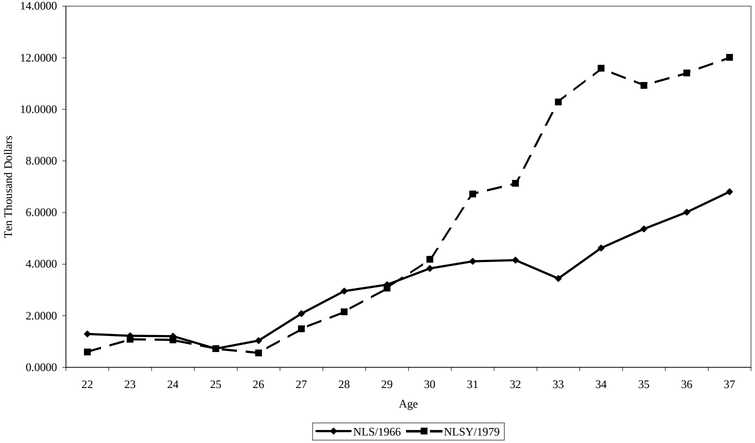
Evolution of variance of unforecastable components–high school sector

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 θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_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 θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_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



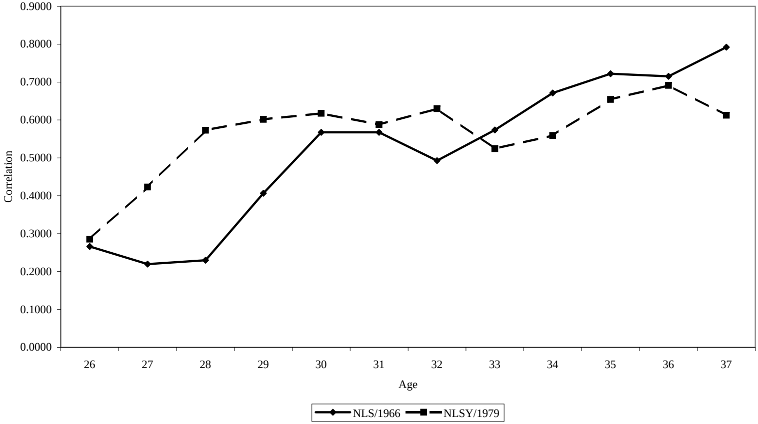
Evolution of variance of unforecastable components–college sector

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 θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_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 θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_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 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



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:

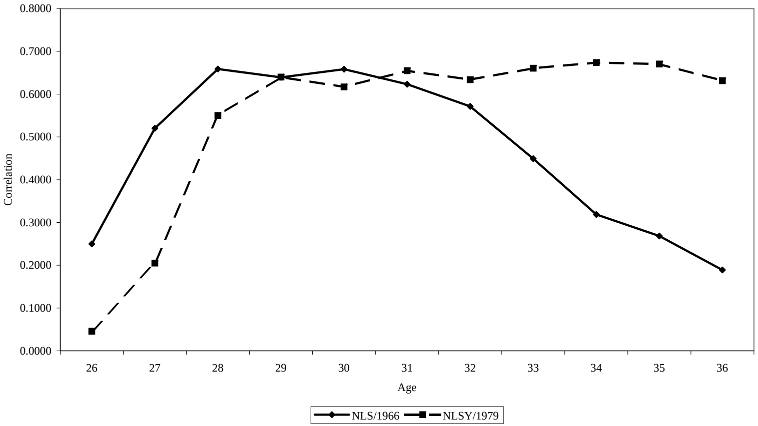
$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_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 θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_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 $P_{s,t+1}$:

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

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



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:

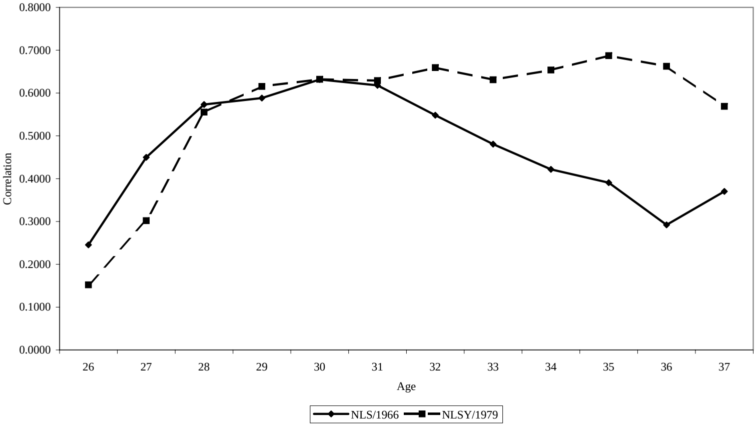
$$Y_{s,t} = X\beta_{s,t} + \theta\alpha_{s,t} + \varepsilon_{s,t}$$

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_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 θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_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 $P_{s,t+1}$:

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

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



One period correlation in unforecastable component of earnings–overall sample

For each schooling level s, at each age t, we model earnings $Y_{s,t}$ according to:

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

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

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

For the NLS/1966 data set, the vector θ contains 5 elements. We test and cannot reject that the agents know the factors θ_1, θ_2, and θ_3 but they don't know factors θ_4, θ_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 θ contains 6 elements. We test and cannot reject that the NLSY/1979 respondents know the factors θ_1, θ_2, and θ_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:

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

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

Counterfactual Simulations

	Total Residual Variance in NLSY/1979			Percentage Change		
	College	**High School**	**Returns**	**College**	**High School**	**Returns**
Estimated (NLSY/1979)	709.7487	507.2910	906.0066	-	-	-
Counterfactual Economy 1	567.9719	333.8619	815.5019	-0.1998	-0.3419	-0.0999
Counterfactual Economy 2	701.6375	471.7632	862.3676	-0.0114	-0.0700	-0.0482
Counterfactual Economy 3	474.8947	294.3037	377.0409	-0.3309	-0.4199	-0.5838
	Variance of Unforecastable Components in NLSY/1979			**Percentage Change**		
	College	**High School**	**Returns**	**College**	**High School**	**Returns**
Estimated (NLSY/1979)	372.3509	272.3596	432.8733	-	-	-
Counterfactual Economy 1	349.4675	205.6334	637.9935	-0.0615	-0.2450	0.4739
Counterfactual Economy 2	364.2397	236.8318	389.2343	-0.0218	-0.1304	-0.1008
Counterfactual Economy 3	192.8493	137.3770	350.4976	-0.4821	-0.4956	-0.1903

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

$$Y_{s,t}^{h} = X\beta_{s,t}^{h} + \theta^{h}\alpha_{s,t} + \varepsilon_{s,t}^{h} \text{ where } h = NLSY/1979,\ NLS/1966.$$

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:

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

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

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

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:

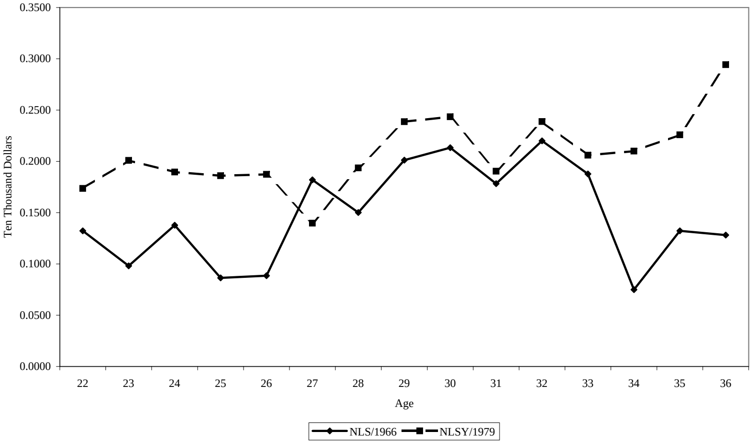
$$P_s^{79} = \sum_{k=4}^{6} Var\left(\theta_k^{79}\right) \left(\sum_{t=22}^{41} \frac{\alpha_{k,s,t}^{79}}{(1+\rho)^{t-22}}\right)^2 + \sum_{t=22}^{41} \frac{Var\left(\varepsilon_{s,t}^{79}\right)}{\left((1+\rho)^{t-22}\right)^2} \tag{2}$$

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 $Var\left(\theta_k^{79}\right)$ with $Var\left(\theta_k^{66}\right)$, for $k = 1,2,...,6$ and fixing $Var\left(\theta_6^{66}\right) = 0$.

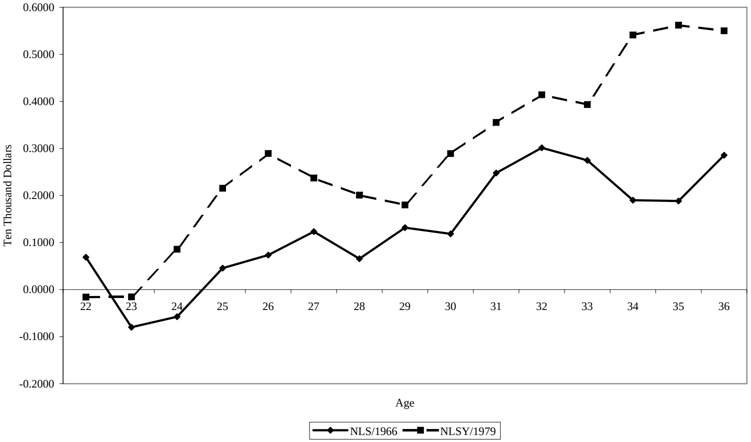
The counterfactual economy 2 is the economy where the distribution of the shocks $\varepsilon_{s,t}^{79}$ are the same as $\varepsilon_{s,t}^{66}$. In this case, we would compute Q_s^{79} and P_s^{79} exactly as above but replacing $Var\left(\varepsilon_{s,t}^{79}\right)$ with $Var\left(\varepsilon_{s,t}^{66}\right)$.

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 $\alpha_{k,s,t}^{66}$.

Evolution of cognitive skill prices – high school sector



Evolution of cognitive skill prices – college sector



Summary

- We discussed the estimation of the distribution of treatment effects (*ex post* or under perfect cerainty).
- 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.
- 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 ε 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.
- See Cunha, Heckman and Schennach (2006a,b).

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 ω^t is denoted $\pi_t\left(\omega^t\right)$.
- 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 $Y_{st}\left(\omega^t\right)$ denote the productivity of agent with schooling level s given history ω^t.

An economic model

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

Consumption Allocation Problem

- Given schooling choice s, the consumption allocation problem of the agent for preference function is

$$V(s) = Max\ E\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t} u\left(c_{s,t}\right)\middle|\mathcal{I}\right] \qquad (18)$$

subject to:

$$\sum_{t=1}^{T}\sum_{\omega^t} q_t\left(\omega^t\right) c_{s,t}^{i}\left(\omega^t\right) = \sum_{t=1}^{T}\sum_{\omega^t} q_t\left(\omega^t\right) Y_{s,t}\left(\omega^t\right) \qquad (19)$$

Lagrange multiplier is λ_s.

Consumption Allocation Problem

- The first-order condition is:

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

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

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

- It is easy to show that

$$c_s = A\left(\rho\right) E\left(\sum_t \left(\frac{1}{1+\rho}\right)^t Y_{s,t}\middle| \mathcal{I}\right) \qquad (20)$$

Consumption Allocation Problem

- We can use (20) in (18) to calculate lifetime utility of schooling level s:

$$V(s) = \frac{1}{A(\rho)} u \left[A(\rho) E \left(\sum_t \left(\frac{1}{1+\rho} \right)^t Y_{s,t} \middle| \mathcal{I} \right) \right]$$

Schooling Decision Problem

- Let C denote the psychic costs associated with schooling choices.
- Let I denote the utility of going to college:

$$I = E\left\{ V(1) - V(0) - \tilde{C} \middle| \mathcal{I} \right\}$$

$$= \frac{1}{A(\rho)} E \left\{ \begin{array}{c} u\left[E\left(A(\rho) \sum_t \left(\frac{1}{1+\rho}\right)^t Y_{1,t} \middle| \mathcal{I} \right) \right] \\ -u\left[E\left(A(\rho) \sum_t \left(\frac{1}{1+\rho}\right)^t Y_{0,t} \middle| \mathcal{I} \right) \right] - C \end{array} \middle| \mathcal{I} \right\}$$

Summary Statistics - NLSY/1979 and PSID[1]

Variable	High School Sample			College Sample		
	Observations	Mean	Standard Error	Observations	Mean	Standard Error
Mother's Education	1045	3.8852	1.1762	843	5.1127	1.5117
Father's Education	1045	3.8699	1.4747	843	5.6987	1.8396
Number of Siblings	1045	3.1129	2.0192	843	2.5101	1.5689
Urban Residence at age 14	1045	0.7378	0.4400	843	0.8470	0.3602
Local Tuition at 4-year college[2]	1045	0.2154	0.0756	843	0.2056	0.0728
Year of Birth is 1958	1045	0.1120	0.3155	843	0.1186	0.3235
Year of Birth is 1959	1045	0.1321	0.3387	843	0.1127	0.3164
Year of Birth is 1960	1045	0.1416	0.3488	843	0.1234	0.3291
Year of Birth is 1961	1045	0.1254	0.3313	843	0.1281	0.3344
Year of Birth is 1962	1045	0.1397	0.3469	843	0.1352	0.3422
Year of Birth is 1963	1045	0.1091	0.3119	843	0.1293	0.3357
Year of Birth is 1964	1045	0.1120	0.3155	843	0.1257	0.3318
Enrolled at School at ASVAB Test Date	538	0.4628	0.4991	465	0.9054	0.2930
Age at ASVAB Test Date	538	19.3457	2.1994	465	19.3462	2.2367
Highest Grade Completed at ASVAB Test Date	538	11.0074	1.3075	465	11.9807	2.0838
ASVAB - Arithmetic Reasoning[3]	538	-0.4783	0.9454	465	0.5687	0.7194
ASVAB - Word Knowledge[3]	538	-0.4310	1.0522	465	0.5222	0.5715
ASVAB - Paragraph Composition[3]	538	-0.4413	1.0869	465	0.5070	0.5463
ASVAB - Coding Speed[3]	538	-0.3277	0.9855	465	0.4096	0.8758
ASVAB - Math Knowledge[3]	538	-0.6211	0.8044	465	0.7035	0.6823

[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[1]

Variable	High School			College		
	Observations	Mean	Standard Error	Observations	Mean	Standard Error
Mother's Education	1117	3.3160	1.1607	1215	4.4568	1.5406
Father's Education	1117	2.9910	1.2798	1215	4.4897	1.9110
Number of Siblings	1117	3.1791	2.2042	1215	2.3358	1.7389
Urban Residence at age 14	1117	0.6634	0.4728	1215	0.8140	0.3893
Local Tuition at 4-year college[2]	1117	0.1577	0.0215	1215	0.1543	0.0214
Year of Birth is 1942	1117	0.0645	0.2457	1215	0.0593	0.2362
Year of Birth is 1943	1117	0.0770	0.2667	1215	0.0601	0.2377
Year of Birth is 1944	1117	0.0609	0.2392	1215	0.0700	0.2552
Year of Birth is 1945	1117	0.0430	0.2029	1215	0.0757	0.2647
Year of Birth is 1946	1117	0.0546	0.2273	1215	0.0840	0.2774
Year of Birth is 1947	1117	0.0985	0.2981	1215	0.1325	0.3392
Year of Birth is 1948	1117	0.1038	0.3052	1215	0.1119	0.3154
Year of Birth is 1949	1117	0.1182	0.3230	1215	0.1119	0.3154
Year of Birth is 1950	1117	0.1334	0.3402	1215	0.0930	0.2906
Year of Birth is 1951	1117	0.1343	0.3411	1215	0.1037	0.3050
Year of Birth is 1952	1117	0.0618	0.2408	1215	0.0494	0.2168
Otis/Beta/Gamma Test[3]	194	-0.5667	0.8026	170	0.6600	0.7971
California Test of Mental Maturity[3]	123	-0.4120	0.8964	95	0.5535	0.8267
Work Knowledge, Occupations[3]	769	0.0854	0.8333	785	0.7025	0.7619
Work Knowledge, Education[3]	772	0.1044	0.8557	787	0.6659	0.7527
Work Knowledge, Earnings Comparison[3]	779	-0.0320	0.9310	791	0.4057	0.9334

[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.

Summary Statistics for Earnings[1] per Schooling Group and Age - NLSY/1979 and PSID[2]

Age	High School Sample			College Sample		
	Observations	Mean	Standard Error	Observations	Mean	Standard Error
22	715	2.0630	1.1722	515	0.8555	0.7747
23	759	2.2700	1.2146	547	1.2363	1.0917
24	795	2.4175	1.4047	588	1.9040	1.3824
25	803	2.5949	1.5033	614	2.5918	1.5434
26	811	2.7158	1.3402	650	3.0224	1.6840
27	829	2.8010	1.4065	668	3.4665	1.9002
28	835	3.0294	1.5780	683	3.8150	1.9740
29	831	3.1140	1.7353	700	4.1289	2.2189
30	814	3.1827	1.5748	706	4.4645	2.3916
31	763	3.4152	2.6028	643	4.8759	3.1723
32	755	3.3891	2.1521	625	5.1221	3.4502
33	658	3.6220	2.5321	572	5.5908	3.8108
34	652	3.6570	3.0182	516	5.8269	3.9961
35	533	3.7389	2.6047	489	6.0398	4.2698
36	530	3.8459	2.4273	445	6.4529	4.2849
37	418	3.8590	2.4618	389	6.9560	4.7483
38	405	3.9803	2.9138	367	7.2928	5.1383
39	341	4.1229	2.9622	315	7.8962	5.8720
40	268	4.2145	2.7394	263	8.0256	5.6185
41	233	4.3307	3.0308	192	8.1747	5.8771

[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

Summary Statistics for Earnings[1] per Schooling Group and Age - NLSY/1966 and PSID[2]

Age	High School Sample			College Sample		
	Observations	Mean	Standard Error	Observations	Mean	Standard Error
22	435	2.7109	1.4637	330	1.3977	1.1195
23	530	2.9363	1.1182	429	2.1406	1.2643
24	601	3.0219	1.2747	580	2.7195	1.4354
25	674	3.2760	1.2396	640	3.2254	1.4483
26	646	3.3748	1.3430	683	3.6320	1.6867
27	653	3.4915	1.3516	698	3.9137	1.8280
28	622	3.5399	1.4160	702	4.3453	2.0456
29	605	3.6311	1.4138	736	4.3880	1.9873
30	593	3.6730	1.5225	699	4.7632	2.1862
31	516	3.5821	1.4622	689	4.9689	2.4219
32	491	3.7744	1.6636	648	5.1731	2.4606
33	432	3.8722	1.7439	618	5.3938	2.5211
34	382	3.8831	1.7366	574	5.6228	3.0574
35	327	4.0330	1.6510	504	5.8789	3.5867
36	280	3.9929	1.7909	476	6.0263	3.7899
37	301	4.0008	1.6679	454	6.3164	4.0824
38	273	3.8033	1.5534	413	6.4203	4.2034
39	247	3.8198	1.7512	407	6.8212	5.1467
40	207	3.7984	1.6144	360	6.7957	4.6554
41	189	3.8534	1.6041	339	6.9379	5.0353

[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[1]

	Otis/Beta/Gamma Test	California Test of Mental Maturity	Work Knowledge, Occupations	Work Knowledge, Education and Occupation	Work Knowledge, Earnings Comparison
Otis/Beta/Gamma Test	1.0000	N/A[2]	0.4289	0.4457	0.0765
California Test of Mental Maturity	N/A[2]	1.0000	0.2036	0.1714	0.1112
Work Knowledge, Occupations	0.4289	0.2036	1.0000	0.9374	0.4464
Work Knowledge, Education	0.4457	0.1714	0.9374	1.0000	0.1068
Work Knowledge, Earnings Comparison	0.0765	0.1112	0.4464	0.1068	1.0000

[1]We control for mother's and fatber'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[1,2]

	High School Earnings Equations						College Earnings Equations					
Age	Loading on Factor 1	Loading on Factor 2	Loading on Factor 3	Loading on Factor 4	Loading on Factor 5	Loading on Factor 6	Loading on Factor 1	Loading on Factor 2	Loading on Factor 3	Loading on Factor 4	Loading on Factor 5	Loading on Factor 6
22		1.00	0.00	0.00	0.00	0.00			0.00	0.00	0.00	0.00
23			0.00	0.00	0.00	0.00			0.00	0.00	0.00	0.00
24			1.00	0.00	0.00	0.00				0.00	0.00	0.00
25				0.00	0.00	0.00				0.00	0.00	0.00
26				1.00	0.00	0.00					0.00	0.00
27					0.00	0.00					0.00	0.00
28					1.00	0.00						0.00
29						0.00						0.00
30						1.00						
31												
32												
33												
34												
35												
36												
37												
38												
39												
40												
41												

[1]The empty cells correspond to factor loadings that are estimated, not normalized.

Normalizations on Factor Loadings: NLSY/1966[1]

	High School Earnings Equations					College Earnings Equations				
Age	Loading on Factor 1	Loading on Factor 2	Loading on Factor 3	Loading on Factor 4	Loading on Factor 5	Loading on Factor 1	Loading on Factor 2	Loading on Factor 3	Loading on Factor 4	Loading on Factor 5
22		1.00	0.00	0.00	0.00			0.00	0.00	0.00
23			0.00	0.00	0.00			0.00	0.00	0.00
24			1.00	0.00	0.00				0.00	0.00
25				0.00	0.00				0.00	0.00
26				1.00	0.00					0.00
27					0.00					0.00
28					1.00					
29										
30										
31										
32										
33										
34										
35										
36										
37										
38										
39										
40										
41										

[1]The empty cells correspond to factor loadings that are estimated, not normalized.

χ^2 Goodness of Fit Test*
NLS/1966 - White Males

	High School		College		Overall	
Age	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**
22	15.9253	20.7143	20.4226	24.9958	40.3457	44.9853
23	34.2904	35.1725	26.0629	31.4104	58.3545	56.9424
24	36.6979	40.1133	78.7201	41.3371	169.9024	74.4683
25	29.5390	43.7730	48.0949	48.6024	127.8453	82.5287
26	40.5503	42.5570	36.2933	44.9853	129.4631	80.2321
27	47.3649	50.9985	49.0855	44.9853	134.1814	83.6753
28	39.4947	47.3999	34.6742	46.1943	109.4980	85.9649
29	32.8477	42.5570	23.7741	49.8018	73.8654	88.2502
30	31.7484	40.1133	28.8962	47.3999	115.9924	81.3810
31	29.0581	36.4150	38.1067	47.3999	66.6677	67.5048
32	29.9466	35.1725	53.5870	47.3999	96.7551	72.1532
33	28.8073	32.6706	33.5289	42.5570	97.1011	67.5048
34	27.0961	30.1435	43.5183	42.5570	92.3537	59.3035
35	29.6717	26.2962	33.9107	33.9244	88.6647	52.1923
36	18.9902	22.3620	26.2794	32.6706	53.1827	47.3999
37	21.6758	22.3620	31.1112	32.6706	76.8646	48.6024
38	14.4640	21.0261	22.3595	31.4104	44.2595	48.6024
39	18.4237	21.0261	23.7976	31.4104	40.6077	41.3371
40	17.4722	19.6751	25.3994	28.8693	66.4910	36.4150
41	13.6884	14.0671	18.1718	26.2962	29.3257	31.4104

* 95% Confidence, equiprobable bins with aprox. 20 people per bin. A χ^2 statistic lower than the critical value indicates a "good" fit.

χ^2 Goodness of Fit Test*
NLSY/1979 - White Males

	High School		College		Overall	
Period	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**	**χ^2 statistic**	**Critical Value**
22	57.3492	47.3999	39.7989	35.1725	167.6037	77.9305
23	42.6356	50.9985	49.0880	36.4150	152.2525	81.3810
24	52.0194	52.1923	197.1642	42.5570	259.3508	84.8206
25	42.4359	46.1943	94.7095	43.7730	175.4185	84.8206
26	62.4019	53.3835	70.2619	43.7730	157.0157	89.3912
27	43.6728	52.1923	42.3386	44.9853	180.5021	91.6702
28	54.9250	48.6024	40.9173	46.1943	174.9548	91.6702
29	49.1212	48.6024	36.1557	46.1943	128.6085	89.3912
30	50.4962	50.9985	41.6969	47.3999	135.4479	89.3912
31	49.6975	48.6024	30.6494	43.7730	119.7788	84.8206
32	44.5459	50.9985	34.5965	42.5570	142.0935	83.6753
33	35.4077	43.7730	28.7575	38.8851	148.9012	77.9305
34	40.7768	42.5570	28.4552	38.8851	95.9926	68.6693
35	41.8859	36.4150	30.6125	33.9244	68.3855	62.8296
36	36.2069	38.8851	39.1018	30.1435	91.1547	62.8296
37	34.6365	31.4104	21.0079	27.5871	73.0057	52.1923
38	24.2197	28.8693	20.5837	27.5871	40.6111	49.8018
39	29.6366	27.5871	29.2055	28.8693	60.9063	42.5570
40	14.3437	21.0261	11.5051	19.6751	29.0430	33.9244
41	17.9075	19.6751	16.6693	16.9190	39.8928	33.9244

* 95% Confidence, equiprobable bins with aprox. 20 people per bin. A χ^2 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

	High School	College	Overall
NLS/1966 - 5 Factors	15.6968	210.4133	114.8754
NLS/1979 - 6 Factors	70.6451	156.5446	187.5425
NLS/1979 - 5 Factors	64.2682	309.2815	226.2401
Critical Value*	222.0741	222.0741	222.0741

* 95% Confidence