Inequality and Uncertainty: The Evolution of Labor Earnings Risk in the U.S. Economy

James Heckman

University of Chicago
American Bar Foundation
University College London

Flavio Cunha

University of Chicago

Invited Lecture for the Far Eastern Meeting of the Econometric Society, Taipei, 11 July 2007

Introduction

- A large literature documents an increase in wage inequality in the American economy over the 1970's and 1980's.

Introduction

- A large literature documents an increase in wage inequality in the American economy over the 1970's and 1980's.
- This has been found in many other countries as well.

Introduction

- A large literature documents an increase in wage inequality in the American economy over the 1970's and 1980's.
- This has been found in many other countries as well.
- This increase in wage inequality has occurred both within and between education-experience groups.

- However, increased variability in wages across people over time is not the same as increased uncertainty in wages.

- However, increased variability in wages across people over time is not the same as increased uncertainty in wages.

- We estimate how much of the recent increase in wage inequality is due to an increase in **heterogeneity** that is predictable by the agents at the age they make their college attendance decisions but is not known to the observing economist, and how much is due to uncertainty at the agent level.

- However, increased variability in wages across people over time is not the same as increased uncertainty in wages.

- We estimate how much of the recent increase in wage inequality is due to an increase in **heterogeneity** that is predictable by the agents at the age they make their college attendance decisions but is not known to the observing economist, and how much is due to uncertainty at the agent level.

- We demonstrate that an increase in microeconomic uncertainty plays an important role in explaining the recent increase in wage inequality.

- Our findings are consistent with the analysis of Gottschalk and Moffitt (1994), who document an increase in "earnings instability" (the $\varepsilon_{s,t}$), demonstrating that the variance of transitory components rose considerably from the period 1970–1978 to the period 1979–1987.

- Our findings are consistent with the analysis of Gottschalk and Moffitt (1994), who document an increase in "earnings instability" (the $\varepsilon_{s,t}$), demonstrating that the variance of transitory components rose considerably from the period 1970–1978 to the period 1979–1987.
- However, their framework cannot distinguish uncertainty from variability.

- Our findings are consistent with the analysis of Gottschalk and Moffitt (1994), who document an increase in “earnings instability” (the $\varepsilon_{s,t}$), demonstrating that the variance of transitory components rose considerably from the period 1970–1978 to the period 1979–1987.
- However, their framework cannot distinguish uncertainty from variability.
- Transitory components as measured by a statistical decomposition of earnings may be perfectly predictable by agents or totally unpredictable.

- Our framework improves on Keane-Wolpin (1997) by estimating sequential updating of serially correlated information sets.

- Our framework improves on Keane-Wolpin (1997) by estimating sequential updating of serially correlated information sets.
- They assume that information shocks are serially independent and identically distributed.

- We use schooling choices to estimate the information sets of agents at the age college enrollment decisions are made.

- We use schooling choices to estimate the information sets of agents at the age college enrollment decisions are made.
- We show that unforecastable components in labor income have increased across cohorts.

- We use schooling choices to estimate the information sets of agents at the age college enrollment decisions are made.
- We show that unforecastable components in labor income have increased across cohorts.
- Earnings instability, or turbulence, has increased substantially.

- We use schooling choices to estimate the information sets of agents at the age college enrollment decisions are made.
- We show that unforecastable components in labor income have increased across cohorts.
- Earnings instability, or turbulence, has increased substantially.
- We model schooling and earnings equations jointly.

A Brief Review of the Previous Literature

- In order to motivate our analysis, it is useful to briefly review previous work

A Brief Review of the Previous Literature

- In order to motivate our analysis, it is useful to briefly review previous work
- The literature on the returns to schooling attempts to estimate the *ex post* rate of return.

A Brief Review of the Previous Literature

- In order to motivate our analysis, it is useful to briefly review previous work
- The literature on the returns to schooling attempts to estimate the *ex post* rate of return.
- *Ex post* returns are interesting historical facts that describe how economies actually reward schooling.

A Brief Review of the Previous Literature

- In order to motivate our analysis, it is useful to briefly review previous work
- The literature on the returns to schooling attempts to estimate the *ex post* rate of return.
- *Ex post* returns are interesting historical facts that describe how economies actually reward schooling.
- *Ex ante* returns are, however, what agents act on.

A Brief Review of the Previous Literature

- In order to motivate our analysis, it is useful to briefly review previous work
- The literature on the returns to schooling attempts to estimate the *ex post* rate of return.
- *Ex post* returns are interesting historical facts that describe how economies actually reward schooling.
- *Ex ante* returns are, however, what agents act on.
- This presentation exposits new methods to estimate *ex ante* returns to schooling.

- The literature on panel data earnings dynamics is not designed to estimate what is in agent information sets.

- The literature on panel data earnings dynamics is not designed to estimate what is in agent information sets.

- See, e.g., Lillard and Willis (1978) and MaCurdy (1982).

- The literature on panel data earnings dynamics is not designed to estimate what is in agent information sets.

- See, e.g., Lillard and Willis (1978) and MaCurdy (1982).

- They estimate earnings equations of the type

$$Y_{i,t} = X_{i,t}\beta + S_i\omega + U_{i,t}, \quad (1)$$

where $Y_{i,t}, X_{i,t}, S_i, U_{i,t}$ denote (for person i at time t) the realized earnings, observable characteristics, educational attainment, and unobservable characteristics, respectively, from the point of view of the observing economist.

- The error term $U_{i,t}$ is often decomposed into two or more components.

- The error term $U_{i,t}$ is often decomposed into two or more components.

- For example, it is common to specify that

$$U_{i,t} = \phi_i + \delta_{i,t}. \tag{2}$$

- The error term $U_{i,t}$ is often decomposed into two or more components.

- For example, it is common to specify that

$$U_{i,t} = \phi_i + \delta_{i,t}. \tag{2}$$

- The term ϕ_i is a person-specific effect.

- The error term $U_{i,t}$ is often decomposed into two or more components.
- For example, it is common to specify that

$$U_{i,t} = \phi_i + \delta_{i,t}. \tag{2}$$

- The term ϕ_i is a person-specific effect.
- The error term $\delta_{i,t}$ is often assumed to follow an ARMA (p, q) process (see Hause, 1980; MaCurdy, 1982), such as $\delta_{i,t} = \rho\delta_{i,t-1} + m_{i,t}$, where $m_{i,t}$ is a mean zero innovation independent of $X_{i,t}$ and the other error components.

- However, the literature is silent about the difference between heterogeneity or variability among persons from the point of view of the observer economist and uncertainty, the unforecastable part of earnings as of a given age.

- However, the literature is silent about the difference between heterogeneity or variability among persons from the point of view of the observer economist and uncertainty, the unforecastable part of earnings as of a given age.

- An alternative specification of the error process postulates a factor structure for earnings,

$$U_{i,t} = \theta_i \alpha_t + \varepsilon_{i,t}, \tag{3}$$

  where θ_i is a vector of skills (e.g., ability, initial human capital, motivation, and the like), α_t is a vector of skill prices, and the $\varepsilon_{i,t}$ are mutually independent mean zero shocks independent of θ_i.

- The predictable components of $U_{i,t}$ will have different effects on choices and economic welfare than the unpredictable components, if people are risk averse and cannot fully insure against uncertainty.

- The predictable components of $U_{i,t}$ will have different effects on choices and economic welfare than the unpredictable components, if people are risk averse and cannot fully insure against uncertainty.
- Statistical decompositions based on (1), (2), and (3) or versions of them describe *ex post* variability but tell us nothing about which components of (1) or (3) are forecastable by agents *ex ante*.

- The predictable components of $U_{i,t}$ will have different effects on choices and economic welfare than the unpredictable components, if people are risk averse and cannot fully insure against uncertainty.
- Statistical decompositions based on (1), (2), and (3) or versions of them describe *ex post* variability but tell us nothing about which components of (1) or (3) are forecastable by agents *ex ante*.
- Is ϕ_i unknown to the agent? $\delta_{i,t}$? Or $\phi_i + \delta_{i,t}$? Or $m_{i,t}$?

- The predictable components of $U_{i,t}$ will have different effects on choices and economic welfare than the unpredictable components, if people are risk averse and cannot fully insure against uncertainty.
- Statistical decompositions based on (1), (2), and (3) or versions of them describe *ex post* variability but tell us nothing about which components of (1) or (3) are forecastable by agents *ex ante*.
- Is ϕ_i unknown to the agent? $\delta_{i,t}$? Or $\phi_i + \delta_{i,t}$? Or $m_{i,t}$?
- In representation (3), the entire vector θ_i, components of the θ_i, the $\varepsilon_{i,t}$, or all of these may or may not be known to the agent at the time schooling choices are made.

Our Approach

- The method is based on the following simple idea.

Our Approach

- The method is based on the following simple idea.
- Suppose that we have data on decisions about a choice variable S.

Our Approach

- The method is based on the following simple idea.
- Suppose that we have data on decisions about a choice variable S.
- The choice variable is assumed to depend, in part, on current and future income, $Y_1, Y_2, \ldots, Y_T$, where T is the horizon for agent decision making, through its present value: $PV = \sum_{t=1}^{T} \left(Y_t / (1+\rho)^{t-1} \right)$, where ρ is the discount rate.

Our Approach

- In the first period, agents only imperfectly predict their future earnings using information $\mathcal{I}$.

Our Approach

- In the first period, agents only imperfectly predict their future earnings using information $\mathcal{I}$.
- S depends on future income, $Y_1, \ldots, Y_T$, through $E\left(PV \mid \mathcal{I}\right)$, where "$E$" denotes expectation.

Our Approach

- In the first period, agents only imperfectly predict their future earnings using information $\mathcal{I}$.
- S depends on future income, $Y_1, \ldots, Y_T$, through $E\left(PV \mid \mathcal{I}\right)$, where "$E$" denotes expectation.
- If, after the choice is made, we actually observe $Y_1, \ldots, Y_T$, we can construct PV *ex post*.

Our Approach

- In the first period, agents only imperfectly predict their future earnings using information $\mathcal{I}$.
- S depends on future income, $Y_1, \ldots, Y_T$, through $E\,(PV \mid \mathcal{I})$, where "E" denotes expectation.
- If, after the choice is made, we actually observe $Y_1, \ldots, Y_T$, we can construct PV *ex post*.
- If the information set is properly specified, the residual corresponding to the component of PV that is not forecastable in the first period, $V = PV - E\,(PV \mid \mathcal{I})$, should not predict S.

Our Approach

- $E(PV \mid \mathcal{I})$ is predictable heterogeneity, allowing for information heterogeneity among agents. V is a measure of uncertainty.

Our Approach

- $E(PV \mid \mathcal{I})$ is predictable heterogeneity, allowing for information heterogeneity among agents. V is a measure of uncertainty.
- This paper develops and applies a method for inferring $\mathcal{I}$ from panel data where the choice is college going.

Our Approach

- $E\,(PV \mid \mathcal{I})$ is predictable heterogeneity, allowing for information heterogeneity among agents. V is a measure of uncertainty.
- This paper develops and applies a method for inferring $\mathcal{I}$ from panel data where the choice is college going.
- Our approach is similar in spirit to a Sims test.

Our Approach

- $E(PV \mid \mathcal{I})$ is predictable heterogeneity, allowing for information heterogeneity among agents. V is a measure of uncertainty.
- This paper develops and applies a method for inferring $\mathcal{I}$ from panel data where the choice is college going.
- Our approach is similar in spirit to a Sims test.
- What components of future earnings cause schooling?

Our Approach

- $E(PV \mid \mathcal{I})$ is predictable heterogeneity, allowing for information heterogeneity among agents. V is a measure of uncertainty.
- This paper develops and applies a method for inferring $\mathcal{I}$ from panel data where the choice is college going.
- Our approach is similar in spirit to a Sims test.
- What components of future earnings cause schooling?
- We go beyond the Sims test by quantifying agent uncertainty.

- Agents have two potential income streams corresponding to the earnings associated with going to college and the earnings associated with not going to college.
- Because we observe the earnings streams of individuals in only one of two possible states (college/no college), it is necessary to account for the missing counterfactual earnings of each person in order to measure unpredictable components.
- This is why we worry about self selection problems in this paper.

The Model

We estimate the information sets of the agents.

Earnings Equations

- Earnings equations for $t = 1, \ldots, T$, which are life cycle outcomes over horizon T.

Earnings Equations

- Earnings equations for $t = 1, \ldots, T$, which are life cycle outcomes over horizon T.
- $(Y_{0,t}, Y_{1,t})$, $t = 1, \ldots, T$, have finite means and can be expressed in terms of conditioning variables X

$$Y_{0,t} = X\beta_{0,t} + U_{0,t}, \quad E(U_{0,t}) < \infty, \tag{4}$$
$$Y_{1,t} = X\beta_{1,t} + U_{1,t}, \quad E(U_{1,t}) < \infty, \tag{5}$$
$$t = 1, \ldots, T.$$

Earnings Equations

- Earnings equations for $t = 1, \ldots, T$, which are life cycle outcomes over horizon T.
- $(Y_{0,t}, Y_{1,t})$, $t = 1, \ldots, T$, have finite means and can be expressed in terms of conditioning variables X

$$Y_{0,t} = X\beta_{0,t} + U_{0,t}, \quad E(U_{0,t}) < \infty, \tag{4}$$
$$Y_{1,t} = X\beta_{1,t} + U_{1,t}, \quad E(U_{1,t}) < \infty, \tag{5}$$
$$t = 1, \ldots, T.$$

- Linearity in parameters plays no essential role in our analysis.

Earnings Equations

- Earnings equations for $t = 1, \ldots, T$, which are life cycle outcomes over horizon T.
- $(Y_{0,t}, Y_{1,t})$, $t = 1, \ldots, T$, have finite means and can be expressed in terms of conditioning variables X

$$\begin{aligned} Y_{0,t} &= X\beta_{0,t} + U_{0,t}, \quad E(U_{0,t}) < \infty, & (4) \\ Y_{1,t} &= X\beta_{1,t} + U_{1,t}, \quad E(U_{1,t}) < \infty, & (5) \end{aligned}$$

$$t = 1, \ldots, T.$$

- Linearity in parameters plays no essential role in our analysis.
- The error terms $U_{s,t}$ are assumed to satisfy $E\left(U_{s,t} \mid X\right) = 0$, $s = 0, 1$.

Choice Equations

- Agents make schooling choices based on expected present value income maximization given information set $\mathcal{I}$.

Choice Equations

- Agents make schooling choices based on expected present value income maximization given information set $\mathcal{I}$.
- Write the index I of present values as

$$I = E\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(Y_{1,t} - Y_{0,t}) - C \,\middle|\, \mathcal{I}\right], \qquad (6)$$

where C is the cost of attending college.

Choice Equations

- Agents make schooling choices based on expected present value income maximization given information set $\mathcal{I}$.

- Write the index I of present values as

$$I = E\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(Y_{1,t} - Y_{0,t}) - C \middle| \mathcal{I}\right], \qquad (6)$$

where C is the cost of attending college.

- We denote by Z and U_C the observable and unobservable determinants of costs, respectively.

- Costs can be written as

$$C = Z\gamma + U_C. \tag{7}$$

- Costs can be written as

$$C = Z\gamma + U_C. \tag{7}$$

- Define

$$\mu_I(X,Z) = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^{t-1} X\left(\beta_{1,t} - \beta_{0,t}\right) - Z\gamma$$

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

$$I = E\left[\mu_I(X,Z) + U_I \middle| \mathcal{I}\right]. \tag{8}$$

- Costs can be written as

$$C = Z\gamma + U_C. \tag{7}$$

- Define

$$\mu_I(X,Z) = \sum_{t=1}^{T} \left(\frac{1}{1+\rho}\right)^{t-1} X\left(\beta_{1,t} - \beta_{0,t}\right) - Z\gamma$$

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

$$I = E\left[\mu_I(X,Z) + U_I \middle| \mathcal{I}\right]. \tag{8}$$

- U_I is the error term in the choice equation and it may or may not include $U_{1,t}$, $U_{0,t}$, or U_C.

Choice Equations

$$S = \mathbf{1}\left[I \geq 0\right]. \tag{9}$$

Test Score Equations

- We have data on a set of cognitive test score equations.

Test Score Equations

- We have data on a set of cognitive test score equations.
- Let M_k denote the agent's score on the k^{th} test (can be any indicator).

Test Score Equations

- We have data on a set of cognitive test score equations.
- Let M_k denote the agent's score on the k^{th} test (can be any indicator).
- M_k have finite means and can be expressed in terms of conditioning variables X^M.

$$M_k = X^M \beta_k^M + U_k^M , \qquad k = 1, 2, \ldots, K. \tag{10}$$

Test Score Equations

- We have data on a set of cognitive test score equations.

- Let M_k denote the agent's score on the k^{th} test (can be any indicator).

- M_k have finite means and can be expressed in terms of conditioning variables X^M.

$$M_k = X^M \beta_k^M + U_k^M, \qquad k = 1, 2, \ldots, K. \tag{10}$$

- We do not need test scores; other measures or longitudinal data on earnings of sufficient length will also work.

Heterogeneity and Uncertainty

- Assume that $X \in \mathcal{I}$.

$$Y_{s,t} = X\beta_{s,t} + E\left(U_{s,t} \mid \mathcal{I}\right) + \left[U_{s,t} - E\left(U_{s,t} \mid \mathcal{I}\right)\right].$$

Heterogeneity and Uncertainty

- Assume that $X \in \mathcal{I}$.

$$Y_{s,t} = X\beta_{s,t} + E\left(U_{s,t} \mid \mathcal{I}\right) + \left[U_{s,t} - E\left(U_{s,t} \mid \mathcal{I}\right)\right].$$

- The component $E\left(U_{s,t} \mid \mathcal{I}\right)$ is available to the agent to help make schooling choices.

Heterogeneity and Uncertainty

- Assume that $X \in \mathcal{I}$.

$$Y_{s,t} = X\beta_{s,t} + E\left(U_{s,t} \mid \mathcal{I}\right) + \left[U_{s,t} - E\left(U_{s,t} \mid \mathcal{I}\right)\right].$$

- The component $E\left(U_{s,t} \mid \mathcal{I}\right)$ is available to the agent to help make schooling choices.
- The component $U_{s,t} - E\left(U_{s,t} \mid \mathcal{I}\right)$ does not enter the schooling equation because it is unknown at the time schooling decisions are made.

- We need to determine which specification of the information set $\mathcal{I}$ best characterizes the dependence between schooling choices and future earnings.

- We need to determine which specification of the information set $\mathcal{I}$ best characterizes the dependence between schooling choices and future earnings.

- We use factor models to represent both $E\left[U_{s,t}\middle|\mathcal{I}\right]$ and $\left(U_{s,t} - E\left[U_{s,t}\middle|\mathcal{I}\right]\right)$.

Factor Models

- Break the error term U_k^M in the test score equations into two components.

$$M_k = X^M \beta_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M. \quad (11)$$

Factor Models

- Break the error term U_k^M in the test score equations into two components.
    - First component is a factor, θ_1

$$M_k = X^M \beta_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M. \qquad (11)$$

Factor Models

- Break the error term U_k^M in the test score equations into two components.
  - First component is a factor, θ_1
  - Second component is unique to test score equation k, ε_k^M.

$$M_k = X^M \beta_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M. \tag{11}$$

Factor Models

- Break the error term U_k^M in the test score equations into two components.
  - First component is a factor, θ_1
  - Second component is unique to test score equation k, ε_k^M.

$$M_k = X^M \beta_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M. \quad (11)$$

- The factor θ_1 is a latent cognitive ability.

Factor Models

- Break the error term U_k^M in the test score equations into two components.
  - First component is a factor, θ_1
  - Second component is unique to test score equation k, ε_k^M.

$$M_k = X^M \beta_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M. \tag{11}$$

- The factor θ_1 is a latent cognitive ability.
- θ_1 is assumed to be independent of X^M and ε_k^M.

Factor Models

- Break the error term U_k^M in the test score equations into two components.
  - First component is a factor, θ_1
  - Second component is unique to test score equation k, ε_k^M.

$$M_k = X^M \beta_k^M + \alpha_k^M \theta_1 + \varepsilon_k^M. \tag{11}$$

- The factor θ_1 is a latent cognitive ability.
- θ_1 is assumed to be independent of X^M and ε_k^M.
- The ε_k^M are mutually independent and independent of θ_1.

Earnings and Choice Equations

- We assume that $U_{0,t}$ and $U_{1,t}$ can be written in factor-structure form

$$U_{i,t} = \alpha_{1,i,t}\theta_1 + \alpha_{2,i,t}\theta_2 + \varepsilon_{i,t}\ , \qquad i = 0, 1.$$

Earnings and Choice Equations

- We assume that $U_{0,t}$ and $U_{1,t}$ can be written in factor-structure form

$$U_{i,t} = \alpha_{1,i,t}\theta_1 + \alpha_{2,i,t}\theta_2 + \varepsilon_{i,t}\,, \qquad i = 0, 1.$$

- Thus,

$$Y_{0,t} = X\beta_{0,t} + \alpha_{1,0,t}\theta_1 + \alpha_{2,0,t}\theta_2 + \varepsilon_{0,t} \quad (12)$$

$$Y_{1,t} = X\beta_{1,t} + \alpha_{1,1,t}\theta_1 + \alpha_{2,1,t}\theta_2 + \varepsilon_{1,t}. \quad (13)$$

Earnings and Choice Equations

- We assume that $U_{0,t}$ and $U_{1,t}$ can be written in factor-structure form

$$U_{i,t} = \alpha_{1,i,t}\theta_1 + \alpha_{2,i,t}\theta_2 + \varepsilon_{i,t}\,, \qquad i = 0, 1.$$

- Thus,

$$Y_{0,t} = X\beta_{0,t} + \alpha_{1,0,t}\theta_1 + \alpha_{2,0,t}\theta_2 + \varepsilon_{0,t} \tag{12}$$

$$Y_{1,t} = X\beta_{1,t} + \alpha_{1,1,t}\theta_1 + \alpha_{2,1,t}\theta_2 + \varepsilon_{1,t}. \tag{13}$$

- We assume that factor θ_j is independent from X, $\varepsilon_{s,t}$, and θ_l for $l \neq j$ and for all s, t.

Earnings and Choice Equations

- We assume that $U_{0,t}$ and $U_{1,t}$ can be written in factor-structure form

$$U_{i,t} = \alpha_{1,i,t}\theta_1 + \alpha_{2,i,t}\theta_2 + \varepsilon_{i,t}\,, \qquad i = 0, 1.$$

- Thus,

$$Y_{0,t} = X\beta_{0,t} + \alpha_{1,0,t}\theta_1 + \alpha_{2,0,t}\theta_2 + \varepsilon_{0,t} \qquad (12)$$

$$Y_{1,t} = X\beta_{1,t} + \alpha_{1,1,t}\theta_1 + \alpha_{2,1,t}\theta_2 + \varepsilon_{1,t}. \qquad (13)$$

- We assume that factor θ_j is independent from X, $\varepsilon_{s,t}$, and θ_l for $l \neq j$ and for all s, t.

- The $\varepsilon_{\ell,t}$, $\ell = 0, 1$ and $t = 1, \ldots, T$, are mutually independent.

$$C = Z\gamma + \alpha_{1,C}\theta_1 + \alpha_{2,C}\theta_2 + \varepsilon_C. \qquad (14)$$

Schooling choice equation:

$$I = E\left[\begin{array}{l} \sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1} X\left(\beta_{1,t}-\beta_{0,t}\right)-Z\gamma \\ +\theta_1\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(\alpha_{1,1,t}-\alpha_{1,0,t}\right)-\alpha_{1,C}\right] \\ +\theta_2\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(\alpha_{2,1,t}-\alpha_{2,0,t}\right)-\alpha_{2,C}\right] \\ +\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(\varepsilon_{1,t}-\varepsilon_{0,t}\right)-\varepsilon_C \end{array}\middle|\ \mathcal{I}\right]. \quad (15)$$

The Estimation of the Information Set

- Need to determine the information set $\mathcal{I}$ of the agent at the age schooling choices are made.

The Estimation of the Information Set

- Need to determine the information set $\mathcal{I}$ of the agent at the age schooling choices are made.

- Define

$$\alpha_{k,I} = \sum_{t=1}^{T} \left(\frac{1}{1+\rho} \right)^{t-1} (\alpha_{k,1,t} - \alpha_{k,0,t}) - \alpha_{k,C} \text{ for } k = 1, 2. \quad (16)$$

- Suppose that $\{\theta_1, \theta_2\} \subset \mathcal{I}$, but $\varepsilon_{s,t} \notin \mathcal{I}$.

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 + \varepsilon_C. \tag{17}$$

- Suppose that $\{\theta_1, \theta_2\} \subset \mathcal{I}$, but $\varepsilon_{s,t} \notin \mathcal{I}$.

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 + \varepsilon_C. \tag{17}$$

- Suppose for the sake of argument that we know $\mu_I(X, Z)$ and $\beta_{s,t}$ for all s and t.

- Suppose that $\{\theta_1, \theta_2\} \subset \mathcal{I}$, but $\varepsilon_{s,t} \notin \mathcal{I}$.

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 + \varepsilon_C. \qquad (17)$$

- Suppose for the sake of argument that we know $\mu_I(X, Z)$ and $\beta_{s,t}$ for all s and t.
- From discrete choice analysis it is well established that under standard conditions, we can proceed as if we know I up to scale.

- Given data $Y_{1,1}$ on X and Z, we can identify the covariance between the terms $I - \mu_I(X, Z)$ and $Y_{1,1} - X\beta_{1,1}$.

- Given data $Y_{1,1}$ on X and Z, we can identify the covariance between the terms $I - \mu_I(X,Z)$ and $Y_{1,1} - X\beta_{1,1}$.
- Under the hypothesis $\{\theta_1, \theta_2\} \subset \mathcal{I}$, this covariance is

$$\text{Cov}\left(I - \mu_I(X,Z), Y_{1,1} - X\beta_{1,1}\right) = \alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1} + \alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2}. \quad (18)$$

- Given data $Y_{1,1}$ on X and Z, we can identify the covariance between the terms $I - \mu_I(X, Z)$ and $Y_{1,1} - X\beta_{1,1}$.
- Under the hypothesis $\{\theta_1, \theta_2\} \subset \mathcal{I}$, this covariance is

$$\text{Cov}\left(I - \mu_I(X, Z), Y_{1,1} - X\beta_{1,1}\right) = \alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1} + \alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2}. \tag{18}$$

- Can test the hypothesis $\{\theta_1, \theta_2\} \subset \mathcal{I}$ against many different alternative hypotheses.

- Given data $Y_{1,1}$ on X and Z, we can identify the covariance between the terms $I - \mu_I(X, Z)$ and $Y_{1,1} - X\beta_{1,1}$.
- Under the hypothesis $\{\theta_1, \theta_2\} \subset \mathcal{I}$, this covariance is

$$\text{Cov}\left(I - \mu_I(X, Z), Y_{1,1} - X\beta_{1,1}\right) = \alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1} + \alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2}. \quad (18)$$

- Can test the hypothesis $\{\theta_1, \theta_2\} \subset \mathcal{I}$ against many different alternative hypotheses.
- Consider the alternative hypothesis that proposes that $\theta_1 \in \mathcal{I}$, but $\theta_2 \notin \mathcal{I}$ and that $E\left[\theta_2 \mid \mathcal{I}\right] = 0$.

- If the alternative is valid, the expected present value of the gain from schooling (15) can be written as

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \varepsilon_C. \tag{19}$$

- If the alternative is valid, the expected present value of the gain from schooling (15) can be written as

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \varepsilon_C. \qquad (19)$$

- In this case, the covariance between the terms $I - \mu_I(X, Z)$ and $Y_{1,1} - X\beta_{1,1}$ is

$$\text{Cov}\left(I - \mu_I(X, Z), Y_{1,1} - X\beta_{1,1}\right) = \alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1}. \qquad (20)$$

- If the alternative is valid, the expected present value of the gain from schooling (15) can be written as

$$I = \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \varepsilon_C. \tag{19}$$

- In this case, the covariance between the terms $I - \mu_I(X, Z)$ and $Y_{1,1} - X\beta_{1,1}$ is

$$\text{Cov}\left(I - \mu_I(X, Z), Y_{1,1} - X\beta_{1,1}\right) = \alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1}. \tag{20}$$

- Difference between (18) and (20) is the term $\alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2}$.

- Can characterize a variety of tests of alternative information structures.

- Can characterize a variety of tests of alternative information structures.
- Define parameters Δ_{θ_1} and Δ_{θ_2} such that

$$\begin{aligned}\operatorname{Cov}\left(I - \mu_I(X, Z), Y_{1,1} - \mu_1(X)\right) \\ - \Delta_{\theta_1}\alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1} - \Delta_{\theta_2}\alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2} = 0.\end{aligned}$$

- Can characterize a variety of tests of alternative information structures.
- Define parameters Δ_{θ_1} and Δ_{θ_2} such that

$$\mathrm{Cov}\left(I - \mu_I(X,Z), Y_{1,1} - \mu_1(X)\right) - \Delta_{\theta_1}\alpha_{1,I}\alpha_{1,1,1}\sigma^2_{\theta_1} - \Delta_{\theta_2}\alpha_{2,I}\alpha_{2,1,1}\sigma^2_{\theta_2} = 0.$$

- Agents know and act on the information contained in factors 1 and 2, so that $\{\theta_1, \theta_2\} \subset \mathcal{I}$, if we reject the hypothesis that both $\Delta_{\theta_1} = 0$ and $\Delta_{\theta_2} = 0$.

- We can actually identify all of the parameters of the model, the function $\mu_I(X, Z)$, the parameters β and α in the test and earnings equations.

- We can actually identify all of the parameters of the model, the function $\mu_I(X, Z)$, the parameters β and α in the test and earnings equations.

- Carneiro, Hansen, and Heckman (2003) present formal proofs of semi-parametric identification of this model.

- We can actually identify all of the parameters of the model, the function $\mu_I(X, Z)$, the parameters β and α in the test and earnings equations.
- Carneiro, Hansen, and Heckman (2003) present formal proofs of semi-parametric identification of this model.
- Normality is *not* required to secure identification, and our estimates are *not* based on normality assumptions.

Empirical Results

- We analyze and compare two distinct samples.

Empirical Results

- We analyze and compare two distinct samples.
- The first sample consists of white males born between 1957 and 1964.

Empirical Results

- We analyze and compare two distinct samples.
- The first sample consists of white males born between 1957 and 1964.
- We obtain information on them from National Longitudinal Survey of Youth (NLSY/1979) data pooled from their birth cohort counterparts from the Panel Survey of Income Dynamics (PSID) data.

- The second sample consists of white males born between 1941 and 1952 who are surveyed in the National Longitudinal Survey (NLS/1966) combined with their birth cohort counterparts from the PSID data.

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

- We consider only two schooling choices: high school and college graduation.
- Both data sets have measures of cognitive test scores.

- We consider only two schooling choices: high school and college graduation.
- Both data sets have measures of cognitive test scores.
- One of the advantages of using factor models instead of the test score itself is that factor models allow for test scores to be noisy measures of cognitive skills.

- For the NLSY/1979, a six factor model fits the data best:

$$\begin{aligned} Y_{s,t} &= X\beta_{s,t} + \theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t} \\ &\quad + \theta_4\alpha_{4,s,t} + \theta_5\alpha_{5,s,t} + \theta_6\alpha_{6,s,t} + \varepsilon_{s,t}, \\ t &= 1,\ldots,T^*,\ s = 0,1, \end{aligned} \tag{21}$$

where $t = 1$ corresponds to age 22 and T^* is age 41. For the NLS/1966, only a five factor model is required to fit the data.

- The cost function C for the 1979 sample is

$$C = Z\gamma + \theta_1\alpha_{1,C} + \theta_2\alpha_{2,C} + \theta_3\alpha_{3,C} + \theta_4\alpha_{4,C} + \theta_5\alpha_{5,C} + \theta_6\alpha_{6,C} + \varepsilon_C. \quad (22)$$

- The cost function C for the 1979 sample is

$$C = Z\gamma + \theta_1\alpha_{1,C} + \theta_2\alpha_{2,C} + \theta_3\alpha_{3,C} + \theta_4\alpha_{4,C} + \theta_5\alpha_{5,C} + \theta_6\alpha_{6,C} + \varepsilon_C. \tag{22}$$

- Each factor θ_k is assumed to be generated by a mixture of J_k normal distributions,

$$\theta_k \sim \sum_{j=1}^{J_k} p_{k,j}\phi\left(\theta_k \mid \mu_{k,j}, \lambda_{k,j}\right),$$

where $\phi\left(\eta \mid \mu_j, \lambda_j\right)$ is a normal density for η with mean μ_j and variance λ_j and $\sum_{j=1}^{J_k} p_{k,j} = 1$, and $p_{k,j} > 0$.

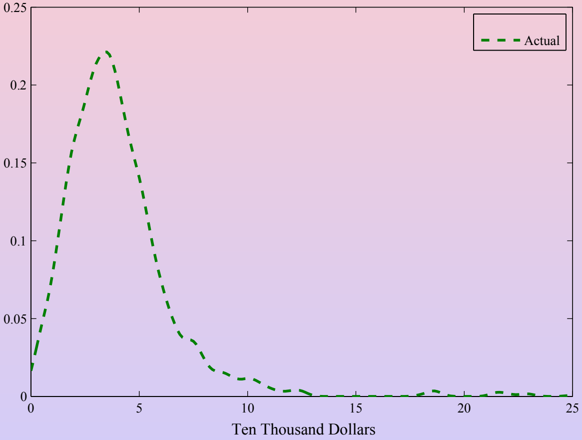
- Ferguson (1983) shows that mixtures of normals with a large number of components approximate any distribution of θ_k arbitrarily well in the ℓ^1 norm.

- Ferguson (1983) shows that mixtures of normals with a large number of components approximate any distribution of θ_k arbitrarily well in the ℓ^1 norm.

- For all factors, a three-component model ($J_k = 3$, $k = 1, \ldots, 6$) is adequate. For all $\varepsilon_{s,t}$ we use a four-component model.

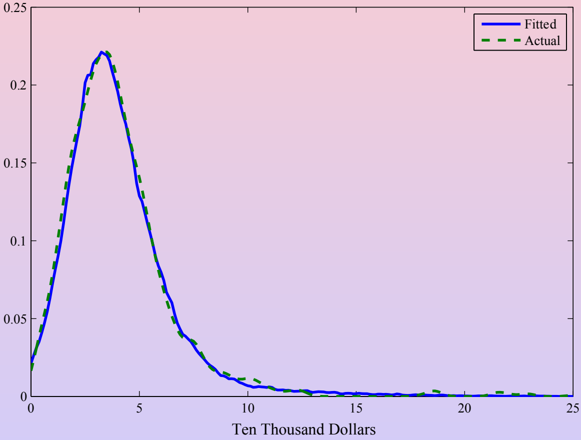
How the model fits the data

Figure 1: Densities of earnings at age 31 (overall sample NLSY/1979)



Let Y denote earnings at age 31 in the overall sample. Here we plot the density functions f(y) generated from the data (the solid curve), against that predicted by the model (the dashed line).

Figure 1: Densities of earnings at age 31 (overall sample NLSY/1979)



Let Y denote earnings at age 31 in the overall sample. Here we plot the density functions f(y) generated from the data (the solid curve), against that predicted by the model (the dashed line).

Empirical Results

- We perform χ^2 goodness-of-fit tests for the earnings correlation matrices.

How the model fits the data

Table 1: 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

Empirical Results

- We relax this assumption and identify the joint distribution of counterfactuals without imposing this condition or other strong assumptions used in the literature.

Empirical Results

- We relax this assumption and identify the joint distribution of counterfactuals without imposing this condition or other strong assumptions used in the literature.

- We identify both *ex ante* and *ex post* joint distributions. Let $E\left(Y_s | \mathcal{I}\right)$ denote the *ex ante* present value of lifetime earnings at schooling level s.

- For a three factor case, the *ex ante* mean present value of earnings is

$$E\left(Y_s \middle| \mathcal{I}\right) = \sum_{t=1}^{T^*} \frac{X\beta_{s,t} + \theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t}}{(1+\rho)^{t-1}}$$

where T^* is the maximum age at which we observe earnings.

- Conditional on covariates X, the covariance between $E(Y_1|\mathcal{I})$ and $E(Y_0|\mathcal{I})$ is

$$\begin{aligned}\operatorname{Cov}(E(Y_1|\mathcal{I}), E(Y_0|\mathcal{I})) \\ = \operatorname{Var}(\theta_1)\left(\sum_{t=1}^{T^*}\frac{\alpha_{1,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*}\frac{\alpha_{1,0,t}}{(1+\rho)^{t-1}}\right) \\ + \cdots + \operatorname{Var}(\theta_3)\left(\sum_{t=1}^{T^*}\frac{\alpha_{3,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*}\frac{\alpha_{3,0,t}}{(1+\rho)^{t-1}}\right).\end{aligned}$$

Table 2A: *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 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 the information set is $\{\theta_1, \theta_2, \theta_3\}$ and $\text{Correlation}(Y_C, Y_H) = 0.1666$.

The Evolution of Joint Distributions and Returns to College

High School \ College	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

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

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

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

High School \ College	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

		College									
		1	2	3	4	5	6	7	8	9	10
High School	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

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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 2B: *Ex Ante* Conditional Distributions for the NLS/1966 (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 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 the information set is $\{\theta_1, \theta_2, \theta_3\}$ and $\text{Correlation}(Y_C, Y_H) = 0.9174$.

High School \ College	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

The Evolution of Joint Distributions and Returns to College

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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

High School \ College	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

- We can also compute the covariance between the present value of *ex post* college and high-school earnings conditional on X. For the NLSY/1979 sample, this is

$$
\begin{aligned}
&\text{Cov}\left(Y_1, Y_0 \mid X\right) \\
&\quad = \text{Var}\left(\theta_1\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{1,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{1,0,t}}{(1+\rho)^{t-1}}\right) \\
&\qquad + \cdots + \text{Var}\left(\theta_6\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{6,1,t}}{(1+\rho)^{t-1}}\right)\left(\sum_{t=1}^{T^*} \frac{\alpha_{6,0,t}}{(1+\rho)^{t-1}}\right).
\end{aligned}
$$

Table 3A: *Ex Post* 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 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 the information set is $\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$ and $\text{Correlation}(Y_C, Y_H) = 0.2842$.

The Evolution of Joint Distributions and Returns to College

High School \ College	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

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School	College 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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

High School \ College	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

High School	College 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

The Evolution of Joint Distributions and Returns to College

High School	College 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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

High School \ College	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 3B: *Ex Post* Conditional Distributions for the NLS/1966 (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 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 the information set is $\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ and $\text{Correlation}(Y_C, Y_H) = 0.6226$.

High School \ College	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

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

High School \ College	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

High School	College 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

High School \ College	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

High School \ College	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

High School	College 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

The Evolution of Joint Distributions and Returns to College

High School \ College	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

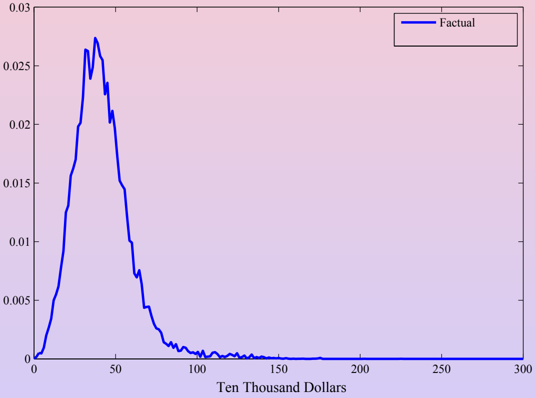
The Evolution of Joint Distributions and Returns to College

High School \ College	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

The Evolution of Joint Distributions and Returns to College

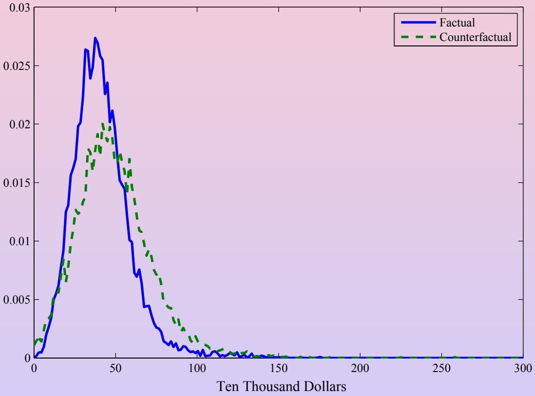
High School \ College	1	2	3	4	5	6	7	8	9	10
1	0.4001	0.1813	0.1023	0.0717	0.0611	0.0406	0.0422	0.0306	0.0337	0.0364
2	0.2144	0.2239	0.1663	0.1207	0.0862	0.0676	0.0486	0.0261	0.0256	0.0205
3	0.1286	0.1716	0.1591	0.1496	0.1181	0.0960	0.0695	0.0515	0.0340	0.0220
4	0.0870	0.1426	0.1551	0.1576	0.1386	0.1131	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

Figure 2A: Density of present value of earnings (high school sample NLSY/1979)



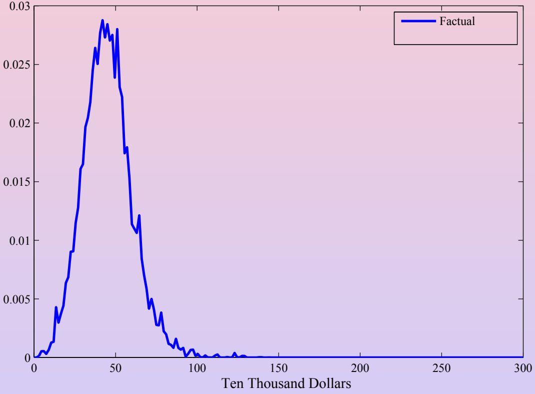
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_0|S=0)$ (the solid curve) against the counterfactual density function $f(y_1|S=0)$ (the dashed curve). We use a discount rate of 5%.

Figure 2A: Density of present value of earnings (high school sample NLSY/1979)



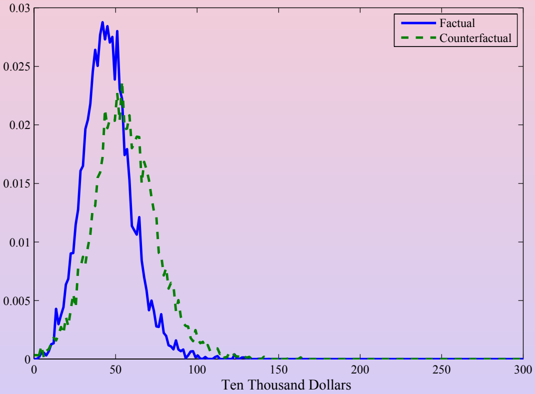
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_0|S=0)$ (the solid curve) against the counterfactual density function $f(y_1|S=0)$ (the dashed curve). We use a discount rate of 5%.

Figure 2B: Densities of present value of earnings (high school sample NLS/1966)



Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector ($S = 0$). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector ($S = 1$). Here we plot the factual density function $f(y_0|S=0)$ (the solid curve) against the counterfactual density function $f(y_1|S=0)$ (the dashed curve). We use a discount rate of 5%.

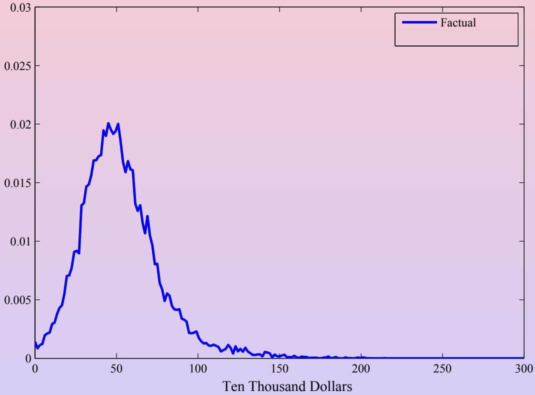
Figure 2B: Densities of present value of earnings (high school sample NLS/1966)



Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_0|S=0)$ (the solid curve) against the counterfactual density function $f(y_1|S=0)$ (the dashed curve). We use a discount rate of 5%.

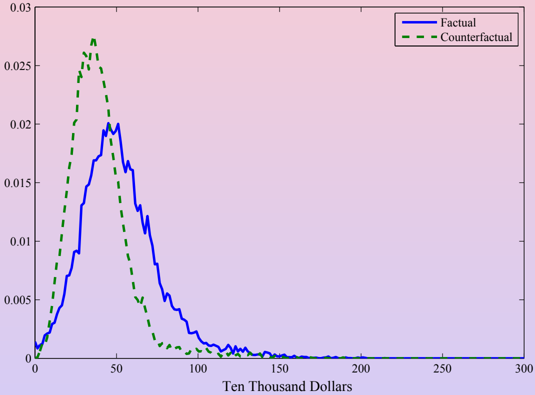
- We compare the actual density of present value of earnings in the college sector with that in the high-school sector.

Figure 3A: Densities of present value of earnings (college sample NLSY/1979)



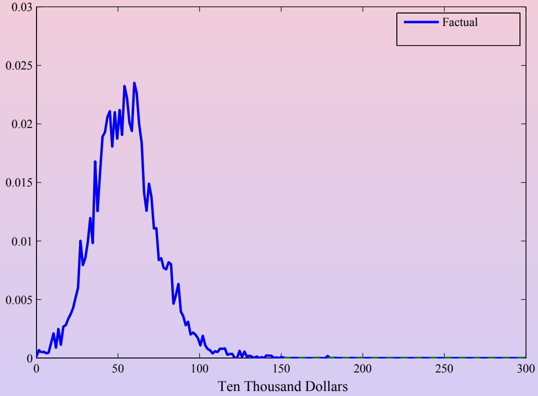
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_1|S=1)$ (the solid curve) against the counterfactual density function $f(y_0|S=1)$ (the dashed curve). We use a discount rate of 5%.

Figure 3A: Densities of present value of earnings (college sample NLSY/1979)



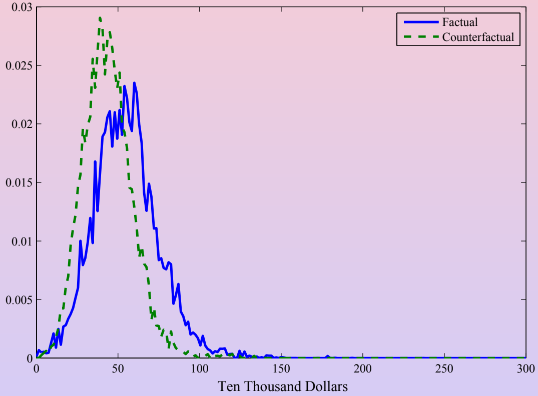
Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector ($S = 0$). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector ($S = 1$). Here we plot the factual density function $f(y_1|S=1)$ (the solid curve) against the counterfactual density function $f(y_0|S=1)$ (the dashed curve). We use a discount rate of 5%.

Figure 3B: Densities of present value of earnings (college sample NLS/1966)



Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_1|S=1)$ (the solid curve) against the counterfactual density function $f(y_0|S=1)$ (the dashed curve). We use a discount rate of 5%.

Figure 3B: Densities of present value of earnings (college sample NLS/1966)



Let Y_0 denote the present value of earnings from age 22 to 41 in the High School sector (S = 0). Let Y_1 denote the present value of earnings from age 22 to 41 in the college sector (S = 1). Here we plot the factual density function $f(y_1|S=1)$ (the solid curve) against the counterfactual density function $f(y_0|S=1)$ (the dashed curve). We use a discount rate of 5%.

- We can compute the percentage of individuals who regret their schooling choice.

- We can compute the percentage of individuals who regret their schooling choice.
- This is reported in Table 5.

Table 5

Percentage that Regret Schooling Choices

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

Table 5

Percentage that Regret 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

- The valuation or net utility function for schooling is

$$I = E\left(\sum_{t=1}^{T^*} \frac{Y_{1,t} - Y_{0,t}}{(1+\rho)^{t-1}} \middle| \mathcal{I} \right) - E\left(C | \mathcal{I} \right).$$

- The valuation or net utility function for schooling is

$$I = E\left(\sum_{t=1}^{T^*} \frac{Y_{1,t} - Y_{0,t}}{(1+\rho)^{t-1}} \middle| \mathcal{I} \right) - E\left(C | \mathcal{I} \right).$$

- In the NLSY/1979, we test, and do not reject, the hypothesis that, at the time they make college going decisions, individuals know their Z and the factors θ_1, θ_2, and θ_3.

- The valuation or net utility function for schooling is

$$I = E\left(\sum_{t=1}^{T^*} \frac{Y_{1,t} - Y_{0,t}}{(1+\rho)^{t-1}} \middle| \mathcal{I} \right) - E\left(C | \mathcal{I} \right).$$

- In the NLSY/1979, we test, and do not reject, the hypothesis that, at the time they make college going decisions, individuals know their Z and the factors θ_1, θ_2, and θ_3.
- They do not know the cohort dummies in X and the factors θ_4, θ_5, θ_6, or $\varepsilon_{s,t}$, $s = 0, 1$, $t = 1, \ldots, T^*$, at the time they make their educational choices.

- Realized earnings in school level s can be written as

$$Y_s = \sum_{t=1}^{T^*} \frac{Y_{s,t}}{(1+\rho)^{t-1}}$$

$$= \sum_{t=1}^{T^*} \frac{\overbrace{X\beta_{s,t} + \theta_1\alpha_{1,s,t} + \theta_2\alpha_{2,s,t} + \theta_3\alpha_{3,s,t}}^{\text{Known to agent}} + \overbrace{\theta_4\alpha_{4,s,t} + \theta_5\alpha_{5,s,t} + \theta_6\alpha_{6,s,t} + \varepsilon_{s,t}}^{\text{Unknown to agent}}}{(1+\rho)^{t-1}}.$$

- We define the residual in the realized present value of earnings as the sum of the unobserved (by the econometrician) components,

$$Q_s = \sum_{t=1}^{T^*} \frac{\begin{pmatrix} \theta_1 \alpha_{1,s,t} + \theta_2 \alpha_{2,s,t} + \theta_3 \alpha_{3,s,t} + \theta_4 \alpha_{4,s,t} \\ + \theta_5 \alpha_{5,s,t} + \theta_6 \alpha_{6,s,t} + \varepsilon_{s,t} \end{pmatrix}}{(1+\rho)^{t-1}}. \tag{23}$$

- For the NLSY/1979, the unforecastable component is

$$P_s = \sum_{t=1}^{T^*} \frac{\theta_4 \alpha_{4,s,t} + \theta_5 \alpha_{5,s,t} + \theta_6 \alpha_{6,s,t} + \varepsilon_{s,t}}{(1+\rho)^{t-1}}. \qquad (24)$$

- We perform a similar analysis for the gross returns to college:

$$R = \sum_{t=1}^{T^*} \frac{Y_{1,t} - Y_{0,t}}{(1+\rho)^{t-1}}.$$

- We perform a similar analysis for the gross returns to college:

$$R = \sum_{t=1}^{T^*} \frac{Y_{1,t} - Y_{0,t}}{(1+\rho)^{t-1}}.$$

- The total residual in the gross returns to college can be defined as $\Delta Q = Q_1 - Q_0$,

$$\Delta Q = \sum_{t=1}^{T^*} \frac{\theta_1 \Delta\alpha_{1,t} + \theta_2 \Delta\alpha_{2,t} + \theta_3 \Delta\alpha_{3,t} + \theta_4 \Delta\alpha_{4,t} + \theta_5 \Delta\alpha_{5,t} + \theta_6 \Delta\alpha_{6,t} + \Delta\varepsilon_t}{(1+\rho)^{t-1}},$$

and the unforecastable component in the gross returns to college is defined as $\Delta P = P_1 - P_0$,

$$\Delta P = \sum_{t=1}^{T^*} \frac{\theta_4 \Delta\alpha_{4,t} + \theta_5 \Delta\alpha_{5,t} + \theta_6 \Delta\alpha_{6,t} + \Delta\varepsilon_t}{(1+\rho)^{t-1}}.$$

Table 6A Unforecastable Components

Evolution of Uncertainty
Panel A: NLS/1966

	College	High School	Returns
Total Residual Variance	460.63	284.81	351.40
Variance of Unforecastable Components	181.37	128.43	327.35

Table 6A Unforecastable Components

Evolution of Uncertainty

Panel A: NLS/1966

	College	High School	Returns
Total Residual Variance	460.63	284.81	351.40
Variance of Unforecastable Components	181.37	128.43	327.35

Panel B: NLSY/1979

	College	High School	Returns
Total Residual Variance	709.75	507.29	906.01
Variance of Unforecastable Components	372.35	272.36	432.87

Table 6A Unforecastable Components

Evolution of Uncertainty

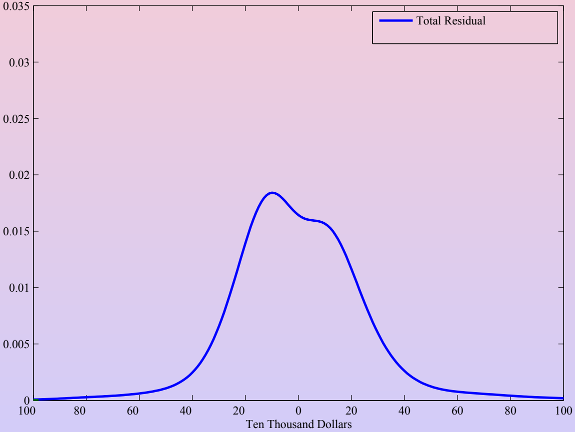
Panel A: NLS/1966			
	College	High School	Returns
Total Residual Variance	460.63	284.81	351.40
Variance of Unforecastable Components	181.37	128.43	327.35
Panel B: NLSY/1979			
	College	High School	Returns
Total Residual Variance	709.75	507.29	906.01
Variance of Unforecastable Components	372.35	272.36	432.87
Panel C: Percentage Increase			
	College	High School	Returns
Percentage Increase in Total Residual Variance	54.08%	78.12%	157.83%
Percentage Increase in Variance of Unforecastable Components	105.30%	112.07%	32.24%

Table 6A Unforecastable Components

Evolution of Uncertainty

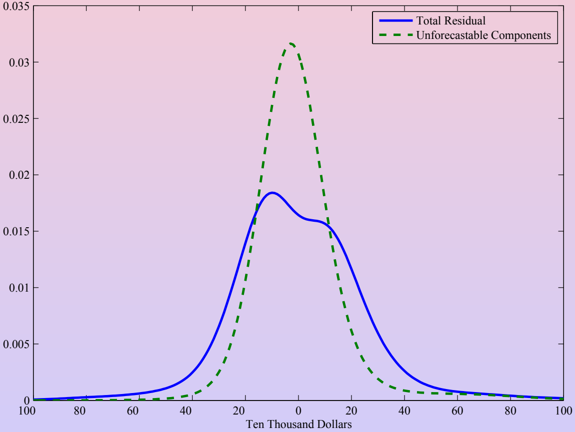
	College	High School	Returns
Panel A: NLS/1966			
Total Residual Variance	460.63	284.81	351.40
Variance of Unforecastable Components	181.37	128.43	327.35
Panel B: NLSY/1979			
	College	High School	Returns
Total Residual Variance	709.75	507.29	906.01
Variance of Unforecastable Components	372.35	272.36	432.87
Panel C: Percentage Increase			
	College	High School	Returns
Percentage Increase in Total Residual Variance	54.08%	78.12%	157.83%
Percentage Increase in Variance of Unforecastable Components	105.30%	112.07%	32.24%
Panel D: Percentage Increase in Total Variance due to Increase in Variance of Uncertainty			
	76.66%	64.69%	19.03%

Figure 5A: Densities of total residual v. unforecastable components in present value of high school earnings (NLSY/1979 sample)



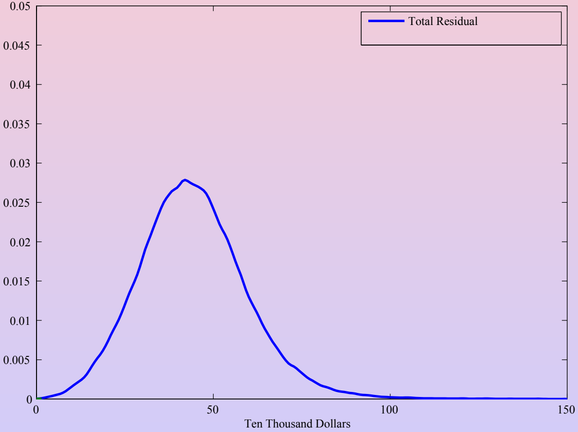
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a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

Figure 5A: Densities of total residual v. unforecastable components in present value of high school earnings (NLSY/1979 sample)



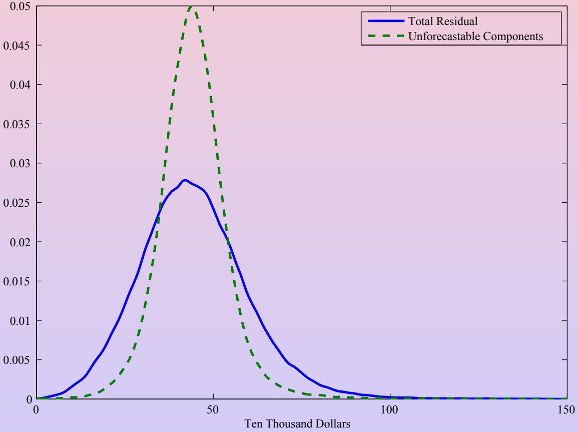
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a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

Figure 5B: Densities of total residual v. unforecastable components in present value of high school earnings (NLS/1966 sample)



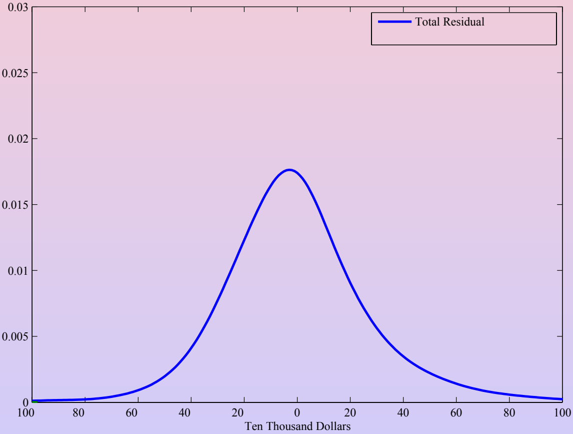
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5%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

Figure 5B: Densities of total residual v. unforecastable components in present value of high school earnings (NLS/1966 sample)



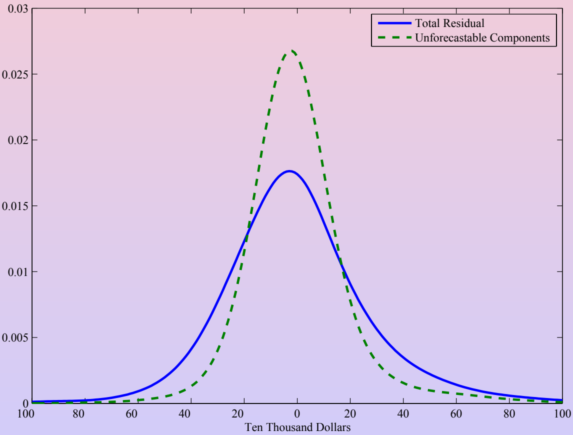
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high-school earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5%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

Figure 6A: Densities of total residual v. unforecastable components in present value of college earnings (NLSY/1979 sample)



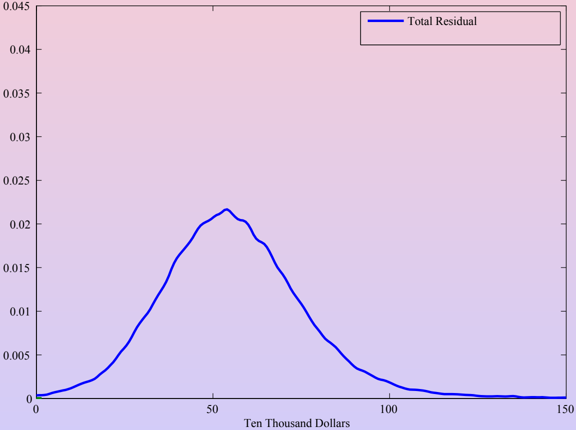
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of college earnings from ages 22 to 41 for the NLSY/1979 sample of white males. The present value of earnings is calculated using a 5% interest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

Figure 6A: Densities of total residual v. unforecastable components in present value of college earnings (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.

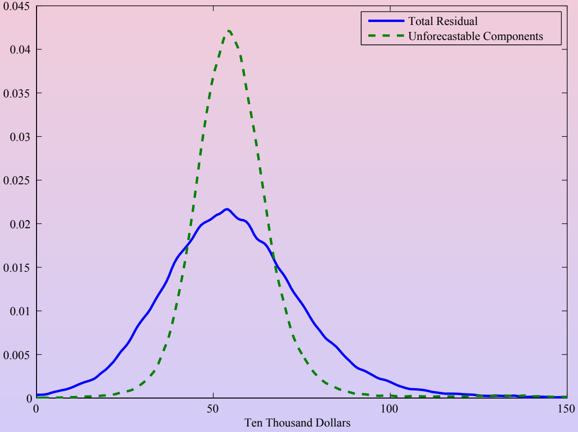
Figure 6B: Densities of total residual v. unforecastable components in present value of college earnings (NLS/1966 sample)



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.

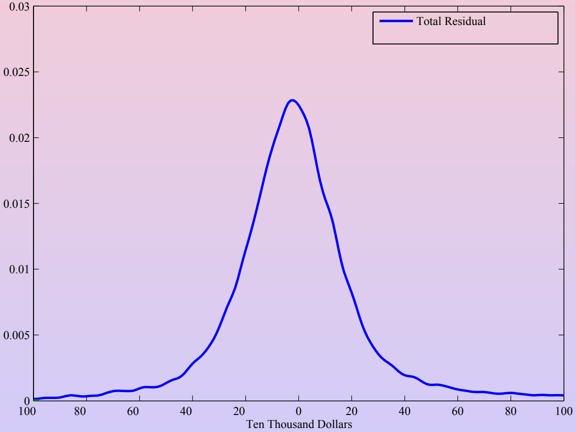
The Evolution of Uncertainty and Heterogeneity

Figure 6B: Densities of total residual v. unforecastable components in present value of college earnings (NLS/1966 sample)



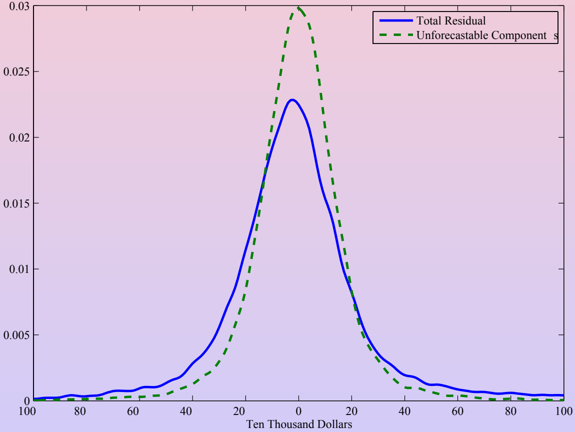
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of college earnings from ages 22 to 41 for the NLS/1966 sample of white males. The present value of earnings is calculated using a 5% interest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

Figure 7A: Densities of total residual v. unforecastable components in returns to college v. high school (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.

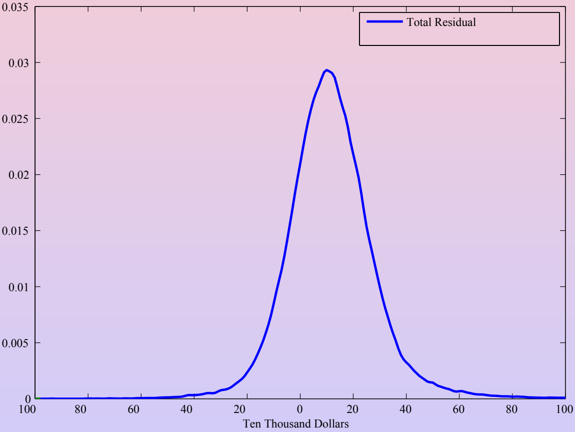
Figure 7A: Densities of total residual v. unforecastable components in returns to college v. high school (NLSY/1979 sample)



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.

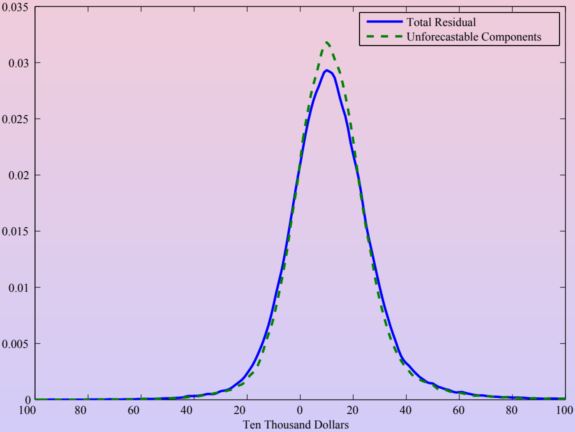
The Evolution of Uncertainty and Heterogeneity

Figure 7B: Densities of total residual v. unforecastable components in returns to college v. high school (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.

Figure 7B: Densities of total residual v. unforecastable components in returns to college v. high school (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.

Table 6B Forecastable Components

Evolution of Heterogeneity

Panel A: NLS/1966

	College	High School	Returns
Total Residual Variance	460.63	284.81	351.40
Variance of Forecastable Components (Heterogeneity)	279.25	156.38	24.05

Table 6B Forecastable Components

Evolution of Heterogeneity

Panel A: NLS/1966			
	College	High School	Returns
Total Residual Variance	460.63	284.81	351.40
Variance of Forecastable Components (Heterogeneity)	279.25	156.38	24.05
Panel B: NLSY/1979			
	College	High School	Returns
Total Residual Variance	709.75	507.29	906.01
Variance of Forecastable Components (Heterogeneity)	337.40	234.93	473.13

Table 6B Forecastable Components

Evolution of Heterogeneity

	College	High School	Returns
Panel A: NLS/1966			
Total Residual Variance	460.63	284.81	351.40
Variance of Forecastable Components (Heterogeneity)	279.25	156.38	24.05
Panel B: NLSY/1979	College	High School	Returns
Total Residual Variance	709.75	507.29	906.01
Variance of Forecastable Components (Heterogeneity)	337.40	234.93	473.13
Panel C: Percentage Increase	College	High School	Returns
Percentage Increase in Total Residual Variance	54.08%	78.12%	157.83%
Percentage Increase in Variance of Forecastable Components	20.82%	50.23%	1866.91%

- About 75% of the increase in the variability in college wages, 65% of the increase in the variability in high school wages, and about 20% of the increase in the variability of returns to college is due to an increase in uncertainty in the American labor market.

The Variance of the Unforecastable Component by Age

- The increase in uncertainty is not uniform across age.

The Variance of the Unforecastable Component by Age

- The increase in uncertainty is not uniform across age.
- The unforecastable components for ages 22 through 25 for the 1979 cohort are given by

$$P_{s,t} = \frac{\varepsilon_{s,t}}{(1-\rho)^{t-1}} \text{ for } t = 1, \ldots, 4, \tag{25}$$

and for ages 26 through 41 by

$$P_{s,t} = \frac{\theta_4 \alpha_{4,s,t} + \theta_5 \alpha_{5,s,t} + \theta_6 \alpha_{6,s,t} + \varepsilon_{s,t}}{(1+\rho)^{t-1}} \text{ for } t = 5, \ldots, T^*. \tag{26}$$

The Variance of the Unforecastable Component by Age

- The increase in uncertainty is not uniform across age.
- The unforecastable components for ages 22 through 25 for the 1979 cohort are given by

$$P_{s,t} = \frac{\varepsilon_{s,t}}{(1-\rho)^{t-1}} \text{ for } t = 1, \ldots, 4, \tag{25}$$

and for ages 26 through 41 by

$$P_{s,t} = \frac{\theta_4 \alpha_{4,s,t} + \theta_5 \alpha_{5,s,t} + \theta_6 \alpha_{6,s,t} + \varepsilon_{s,t}}{(1+\rho)^{t-1}} \text{ for } t = 5, \ldots, T^*. \tag{26}$$

- Figure 8 plots the variance of unforecastable components in high school earnings in NLS/1966 and NLSY/1979.

Figure 8

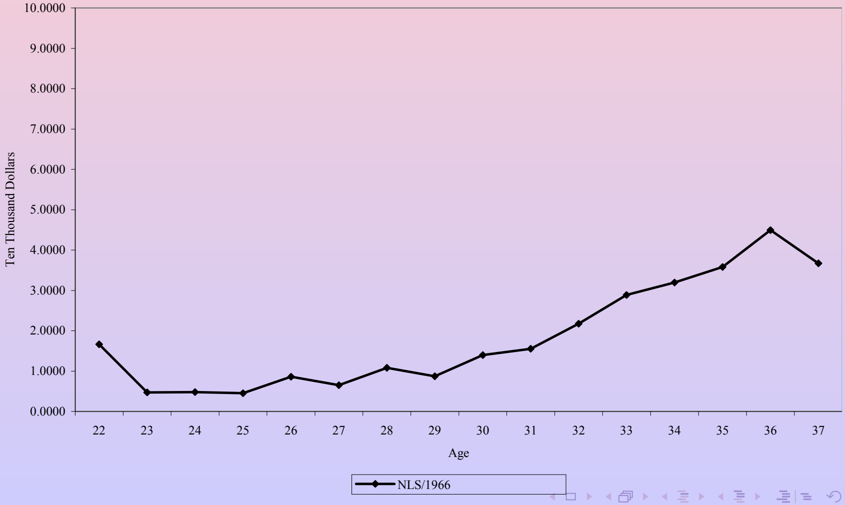

10.0000
9.0000
8.0000
7.0000
6.0000
5.0000
4.0000
3.0000
2.0000
1.0000
0.0000
Ten Thousand Dollars
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
Age
NLS/1966

Figure 8

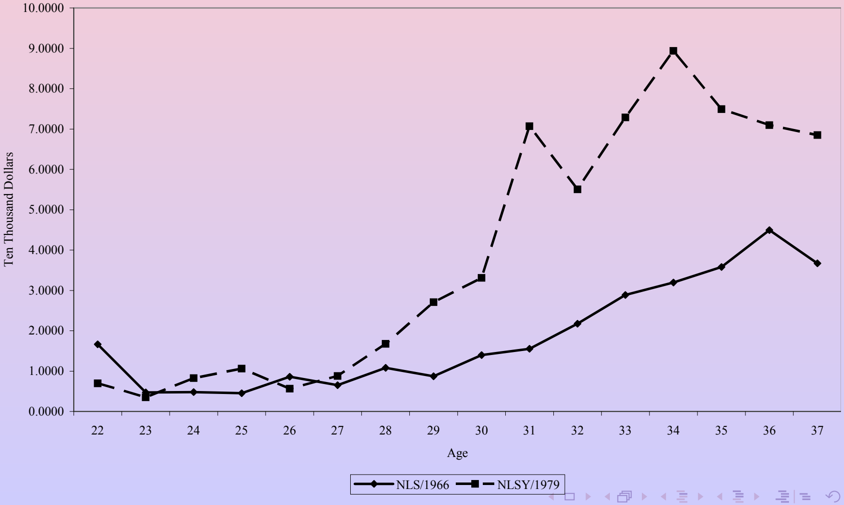

10.0000
9.0000
8.0000
7.0000
6.0000
5.0000
4.0000
3.0000
2.0000
1.0000
0.0000
Ten Thousand Dollars
22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37
Age
NLS/1966
NLSY/1979

Figure 8 note

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 10, 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.

- A similar pattern appears in the variances of the unforecastable components in college earnings.

Figure 9

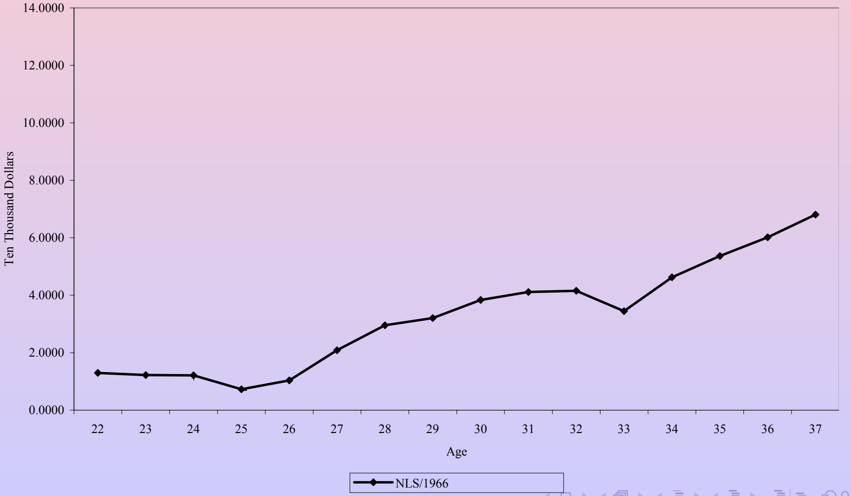

14.0000
12.0000
10.0000
8.0000
6.0000
4.0000
2.0000
0.0000
Ten Thousand Dollars
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
Age
NLS/1966

Figure 9

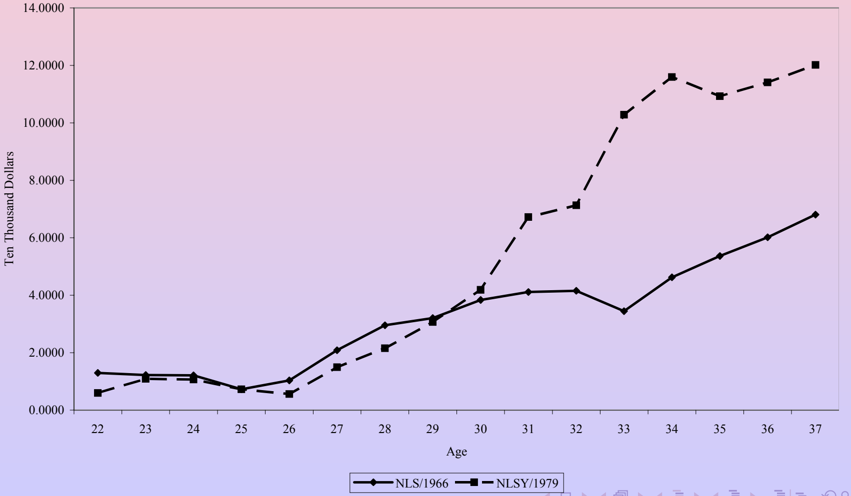

14.0000
12.0000
10.0000
8.0000
6.0000
4.0000
2.0000
0.0000
Ten Thousand Dollars
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
Age
NLS/1966
NLSY/1979

Figure 9 note

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 11, 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.

Accounting for Macro Uncertainty

- Macro uncertainty decreased for later cohorts by 90%.

Accounting for Macro Uncertainty

- Macro uncertainty decreased for later cohorts by 90%.
- These estimates are consistent with the evidence that US business cycle volatility has decreased in recent years.

Accounting for Macro Uncertainty

- Macro uncertainty decreased for later cohorts by 90%.
- These estimates are consistent with the evidence that US business cycle volatility has decreased in recent years.
- At the same time, macro uncertainty is a tiny fraction of total uncertainty for both cohorts (5% for 1966; 1% for 1979).

Table 7

Share of Variance of Business Cycle in Total Variance of Unforecastable Components

	NLS/1966		NLSY/1979	
	Point Estimate	Standard Error	Point Estimate	Standard Error
High School	0.0586	0.0060	0.0069	0.0009
College	0.1193	0.0126	0.0158	0.0021

Table 7 note

Let $Y_{s,t}$ denote the labor income in schooling sector s at age t. Let d_k denote the cohort dummy that takes the value one if the agent was born in year k and zero otherwise. Let X denote the vector of variables containing a dummy indicating whether the agent lived in the South Region at age 14 and a constant term. Let θ_j denote the factor j and $\alpha_{s,t,j}$ denote its factor loading at schooling sector s and age t. Let $\varepsilon_{s,t}$ denote the uniqueness. The model is:

$$Y_{s,t} = X\beta_{s,t} + \sum_{k=\tau_0}^{\tau_1} \gamma_{k,s,t} d_k + \theta_1\alpha_{s,t,1} + \theta_2\alpha_{s,t,2} + \theta_3\alpha_{s,t,3} + \theta_4\alpha_{s,t,4} + \theta_5\alpha_{s,t,5} + \theta_6\alpha_{s,t,6} + \varepsilon_{s,t}.$$

The cohort dummies can capture aggregate shocks. Under this interpretation, we test and reject the hypothesis that the agents know the aggregate shocks at the time of the schooling choice. We test and reject the hypothesis that the agent knows the uniqueness $\varepsilon_{s,t}$ and factors θ_4, θ_5, and θ_6 at the time of the schooling choice. Consequently, the total unforecastable component (aggregate and idiosyncratic components) is given by:

$$\tilde{P}_{s,t} = \sum_{k=\tau_0}^{\tau_1} \gamma_{k,s,t} d_k + \theta_4\alpha_{s,t,4} + \theta_5\alpha_{s,t,5} + \theta_6\alpha_{s,t,6} + \varepsilon_{s,t}.$$

Sequential Revelation of Information, More General Preferences and Market Settings

- We have analyzed a one-shot model of schooling choices.

Sequential Revelation of Information, More General Preferences and Market Settings

- We have analyzed a one-shot model of schooling choices.
- We also assume risk neutrality.

Sequential Revelation of Information, More General Preferences and Market Settings

- We have analyzed a one-shot model of schooling choices.
- We also assume risk neutrality.
- This allows us to use expected present value income maximization as our schooling choice criterion.

- A basic question is “What can be identified in more general environments?”

- A basic question is “What can be identified in more general environments?”
- In the absence of perfect certainty or perfect risk sharing, preferences and market environments also determine schooling choices.

- A basic question is “What can be identified in more general environments?”
- In the absence of perfect certainty or perfect risk sharing, preferences and market environments also determine schooling choices.
- The separation theorem used in this paper that allows consumption and schooling decisions to be analyzed in isolation of each other breaks down.

- If we postulate information arrival processes *a priori*, and assume that preferences are known up to some unknown parameters as in Flavin (1981), Blundell and Preston (1998) and Blundell, Pistaferri, and Preston (2004), we can identify departures from specified market structures.

- If we postulate information arrival processes *a priori*, and assume that preferences are known up to some unknown parameters as in Flavin (1981), Blundell and Preston (1998) and Blundell, Pistaferri, and Preston (2004), we can identify departures from specified market structures.

- An open question, not yet resolved in the literature, is how far one can go in nonparametrically jointly identifying preferences, market structures and agent information sets.

- One can add consumption data to the schooling choice and earnings data to secure identification of risk preference parameters (within a parametric family) and information sets, and to test among alternative models for market environments.

- One can add consumption data to the schooling choice and earnings data to secure identification of risk preference parameters (within a parametric family) and information sets, and to test among alternative models for market environments.

- The lack of full insurance interpretation given to the empirical analysis by Flavin (1981) and Blundell, Pistaferri, and Preston (2004), may instead be a consequence of their misspecification of the generating processes of agent information sets.

Summary and Conclusion

- Increasing wage inequality arises from both increasing micro uncertainty and increasing heterogeneity predictable by agents.

Summary and Conclusion

- Increasing wage inequality arises from both increasing micro uncertainty and increasing heterogeneity predictable by agents.
- About 75% of the increase in the variability in college wages, 65% of the increase in the variability in high school wages, and about 20% of the increase in the variability of returns to college is due to an increase in uncertainty in the American labor market.

Summary and Conclusion

- Increasing wage inequality arises from both increasing micro uncertainty and increasing heterogeneity predictable by agents.
- About 75% of the increase in the variability in college wages, 65% of the increase in the variability in high school wages, and about 20% of the increase in the variability of returns to college is due to an increase in uncertainty in the American labor market.
- About 50% of the variance in wages is due to uncertainty in both years.

- Most agent heterogeneity is microeconomic and this has become even more microeconomic in recent years.

- Most agent heterogeneity is microeconomic and this has become even more microeconomic in recent years.
- Macro forecasting equations understate the extent of true heterogeneity in the economy.

APPENDIX

Identification of the Model

- First consider identification of the test score equations.

APPENDIX

Identification of the Model

- First consider identification of the test score equations.
- Compute the covariances:

$$\text{Cov}\left(M_1 - X^M\beta_1^M, M_2 - X^M\beta_2^M\right) = \alpha_1^M\alpha_2^M\sigma_{\theta_1}^2, \quad (27)$$

$$\text{Cov}\left(M_1 - X^M\beta_1^M, M_3 - X^M\beta_3^M\right) = \alpha_1^M\alpha_3^M\sigma_{\theta_1}^2, \quad (28)$$

$$\text{Cov}\left(M_2 - X^M\beta_2^M, M_3 - X^M\beta_3^M\right) = \alpha_2^M\alpha_3^M\sigma_{\theta_1}^2. \quad (29)$$

- Because the factor θ_1 and uniquenesses ε_k are independently normally distributed random variables, we have identified their distribution.

Earnings and Choice Equations

- We rely on four key assumptions to secure identification.

Earnings and Choice Equations

- We rely on four key assumptions to secure identification.
- First, all of the observable explanatory variables X and Z are independent of the unobservable factors, θ_1 and θ_2, as well as uniquenesses $\varepsilon_{s,t}$ for all s, t.

Earnings and Choice Equations

- We rely on four key assumptions to secure identification.
- First, all of the observable explanatory variables X and Z are independent of the unobservable factors, θ_1 and θ_2, as well as uniquenesses $\varepsilon_{s,t}$ for all s, t.
- Second, θ_1 is independent of θ_2.

Earnings and Choice Equations

- We rely on four key assumptions to secure identification.
- First, all of the observable explanatory variables X and Z are independent of the unobservable factors, θ_1 and θ_2, as well as uniquenesses $\varepsilon_{s,t}$ for all s, t.
- Second, θ_1 is independent of θ_2.
- Third, both θ_1 and θ_2 are independent of ε_C and $\varepsilon_{s,t}$ for all s, t.

- $$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma^2_{\theta_1} & 0 \\ 0 & \sigma^2_{\theta_2} \end{bmatrix}\right).$$

$$\begin{bmatrix} Y_{0,t} \\ Y_{1,t} \end{bmatrix} \mid X \sim N\left(\begin{bmatrix} X\beta_{0,t} \\ X\beta_{1,t} \end{bmatrix}, \begin{bmatrix} \alpha^2_{1,0,t}\sigma^2_{\theta_1} + \alpha^2_{2,0,t}\sigma^2_{\theta_2} + \sigma^2_{\varepsilon_{0,t}} & \alpha_{1,0,t}\alpha_{1,1,t}\sigma^2_{\theta_1} + \alpha_{2,0,t}\alpha_{2,1,t}\sigma^2_{\theta_2} \\ \alpha_{1,0,t}\alpha_{1,1,t}\sigma^2_{\theta_1} + \alpha_{2,0,t}\alpha_{2,1,t}\sigma^2_{\theta_2} & \alpha^2_{1,1,t}\sigma^2_{\theta_1} + \alpha^2_{2,1,t}\sigma^2_{\theta_2} + \sigma^2_{\varepsilon_{1,t}} \end{bmatrix}\right). \quad (30)$$

- $$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma^2_{\theta_1} & 0 \\ 0 & \sigma^2_{\theta_2} \end{bmatrix}\right).$$

- The joint distribution of the labor earnings $Y_{0,t}$, $Y_{1,t}$ conditional on X is

$$\begin{bmatrix} Y_{0,t} \\ Y_{1,t} \end{bmatrix} | X \tag{30}$$
$$\sim N\left(\begin{bmatrix} X\beta_{0,t} \\ X\beta_{1,t} \end{bmatrix}, \begin{bmatrix} \alpha^2_{1,0,t}\sigma^2_{\theta_1} + \alpha^2_{2,0,t}\sigma^2_{\theta_2} + \sigma^2_{\varepsilon_{0,t}} & \alpha_{1,0,t}\alpha_{1,1,t}\sigma^2_{\theta_1} + \alpha_{2,0,t}\alpha_{2,1,t}\sigma^2_{\theta_2} \\ \alpha_{1,0,t}\alpha_{1,1,t}\sigma^2_{\theta_1} + \alpha_{2,0,t}\alpha_{2,1,t}\sigma^2_{\theta_2} & \alpha^2_{1,1,t}\sigma^2_{\theta_1} + \alpha^2_{2,1,t}\sigma^2_{\theta_2} + \sigma^2_{\varepsilon_{1,t}} \end{bmatrix}\right).$$

- From the observed data and the factor structure assumption it follows that

$$E\left(Y_{1,t}\middle|X,S=1\right)=X\beta_{1,t}+\alpha_{1,1,t}E\left[\theta_1|X,S=1\right] \quad (31)$$
$$+\alpha_{2,1,t}E\left[\theta_2|X,S=1\right]+E\left[\varepsilon_{1,t}\middle|X,S=1\right].$$

- From the observed data and the factor structure assumption it follows that

$$E\left(Y_{1,t}\middle|X,S=1\right)=X\beta_{1,t}+\alpha_{1,1,t}E\left[\theta_1|X,S=1\right] \tag{31}$$
$$+\alpha_{2,1,t}E\left[\theta_2|X,S=1\right]+E\left[\varepsilon_{1,t}\middle|X,S=1\right].$$

- The event $S=1$ corresponds to the event $I=E\left(\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(Y_{1,t}-Y_{0,t}\right)-C\middle|\mathcal{I}\right)\geq 0.$

- Assuming that $\varepsilon_{s,t}$ does not enter agent information sets, for the case $\{\theta_1, \theta_2\} \subset \mathcal{I}$ we obtain

$$E\left(\sum_{t=1}^{T} \left(\frac{1}{1+\rho} \right)^{t-1} (Y_{1,t} - Y_{0,t}) - C \middle| \mathcal{I} \right)$$
$$= \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 - \varepsilon_C.$$

- Assuming that $\varepsilon_{s,t}$ does not enter agent information sets, for the case $\{\theta_1, \theta_2\} \subset \mathcal{I}$ we obtain

$$E\left(\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(Y_{1,t} - Y_{0,t}) - C \middle| \mathcal{I}\right)$$
$$= \mu_I(X, Z) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 - \varepsilon_C.$$

- Let η be the linear combination of three independent normal random variables: $\eta = \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 - \varepsilon_C$.

- Assuming that $\varepsilon_{s,t}$ does not enter agent information sets, for the case $\{\theta_1, \theta_2\} \subset \mathcal{I}$ we obtain

$$E\left(\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(Y_{1,t}-Y_{0,t})-C\middle|\mathcal{I}\right)$$
$$=\mu_I(X,Z)+\alpha_{1,I}\theta_1+\alpha_{2,I}\theta_2-\varepsilon_C.$$

- Let η be the linear combination of three independent normal random variables: $\eta=\alpha_{1,I}\theta_1+\alpha_{2,I}\theta_2-\varepsilon_C$.

- Then, $\eta \sim N\left(0,\sigma_\eta^2\right)$, with $\sigma_\eta^2=\alpha_{1,I}^2\sigma_{\theta_1}^2+\alpha_{2,I}^2\sigma_{\theta_2}^2+\sigma_{\varepsilon_c}^2$ and

$$S=1 \Leftrightarrow \eta > -\mu_I(X,Z). \tag{32}$$

- If we replace (32) in (31) and use the fact that $\varepsilon_{s,t}$ is independent of X, Z, and θ,

$$E\left(Y_{1,t} \middle| X, S=1\right) = X\beta_1 + \alpha_{1,1,t} E\left[\theta_1 | X, \eta > -\mu_I(X,Z)\right] \quad (33)$$
$$+ \alpha_{2,1,t} E\left[\theta_2 | X, \eta > -\mu_I(X,Z)\right].$$

Because θ_1, θ_2 and η are normal random variables,

$$\theta_j = \frac{\text{Cov}\left(\theta_j, \eta\right)}{\text{Var}\left(\eta\right)}\eta + \rho_j \text{ for } j = 1, 2, \quad (34)$$

where ρ_j is a mean zero, normal random variable independent from η.

- Because $\operatorname{Cov}(\theta_1, \eta) = \sigma^2_{\theta_1}\alpha_{1,I}$ and $\operatorname{Cov}(\theta_2, \eta) = \sigma^2_{\theta_2}\alpha_{2,I}$ it follows that

$$E[\theta_1 | X, \eta > -\mu_I(X,Z)] = \frac{\sigma^2_{\theta_1}\alpha_{1,I}}{\sigma^2_\eta} E\left[\eta \middle| \eta > -\mu_I(X,Z)\right]$$

and

$$E[\theta_2 | X, \eta > -\mu_I(X,Z)] = \frac{\sigma^2_{\theta_2}\alpha_{2,I}}{\sigma^2_\eta} E\left[\eta \middle| \eta > -\mu_I(X,Z)\right].$$

- Because $\text{Cov}(\theta_1, \eta) = \sigma^2_{\theta_1}\alpha_{1,I}$ and $\text{Cov}(\theta_2, \eta) = \sigma^2_{\theta_2}\alpha_{2,I}$ it follows that

$$E[\theta_1 | X, \eta > -\mu_I(X, Z)] = \frac{\sigma^2_{\theta_1}\alpha_{1,I}}{\sigma^2_\eta} E\left[\eta \middle| \eta > -\mu_I(X, Z)\right]$$

and

$$E[\theta_2 | X, \eta > -\mu_I(X, Z)] = \frac{\sigma^2_{\theta_2}\alpha_{2,I}}{\sigma^2_\eta} E\left[\eta \middle| \eta > -\mu_I(X, Z)\right].$$

- We can rewrite (31) as

$$E\left(Y_{1,t} \middle| \eta \leq -\mu_I(X, Z)\right) = X\beta_{1,t} + \pi_{1,t} \frac{\phi\left(\frac{\mu_I(X,Z)}{\sigma_\eta}\right)}{\Phi\left(\frac{\mu_I(X,Z)}{\sigma_\eta}\right)}. \quad (35)$$

- We can derive a similar expression for mean observed earnings in sector "0":

$$E\left(Y_{0,t}\middle|\,\eta > -\mu_I(X,Z)\right) = X\beta_{0,t} - \pi_{0,t}\frac{\phi\left(\frac{\mu_I(X,Z)}{\sigma_\eta}\right)}{\Phi\left(\frac{\mu_I(X,Z)}{\sigma_\eta}\right)}. \quad (36)$$

- We can derive a similar expression for mean observed earnings in sector "0":

$$E\left(Y_{0,t}\middle|\,\eta > -\mu_I(X,Z)\right) = X\beta_{0,t} - \pi_{0,t}\frac{\phi\left(\frac{\mu_I(X,Z)}{\sigma_\eta}\right)}{\Phi\left(\frac{\mu_I(X,Z)}{\sigma_\eta}\right)}. \quad (36)$$

- Given identification of $\beta_{s,t}$ for all s and t, we can construct the differences $Y_{s,t} - X\beta_{s,t}$ and compute the covariances

$$\mathrm{Cov}\left(M_1 - X^M\beta_1^M, Y_{0,t} - X\beta_{0,t}\right) = \alpha_{1,0,t}\sigma^2_{\theta_1} \quad (37)$$

and

$$\mathrm{Cov}\left(M_1 - X^M\beta_1^M, Y_{1,t} - X\beta_{1,t}\right) = \alpha_{1,1,t}\sigma^2_{\theta_1}. \quad (38)$$

- Note that we can also identify $\alpha_{1,C}/\sigma_\eta$ by computing the covariance

$$\mathrm{Cov}\left(M_1 - X\beta_1^M, \frac{I - \mu_I(X,Z)}{\sigma_\eta}\right) \tag{39}$$
$$= \frac{\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(\alpha_{1,1,t} - \alpha_{1,0,t}) - \alpha_{1,C}}{\sigma_\eta}\sigma_{\theta_1}^2.$$

- Note that we can also identify $\alpha_{1,C}/\sigma_\eta$ by computing the covariance

$$\mathrm{Cov}\left(M_1 - X\beta_1^M, \frac{I - \mu_I(X,Z)}{\sigma_\eta}\right) \tag{39}$$
$$= \frac{\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(\alpha_{1,1,t} - \alpha_{1,0,t}) - \alpha_{1,C}}{\sigma_\eta}\sigma_{\theta_1}^2.$$

- Using (37) and (38), we can identify $\alpha_{1,1,t}$ and $\alpha_{1,0,t}$ for all t.

- Note that we can also identify $\alpha_{1,C}/\sigma_\eta$ by computing the covariance

$$\text{Cov}\left(M_1 - X\beta_1^M, \frac{I - \mu_I(X,Z)}{\sigma_\eta}\right) \tag{39}$$
$$= \frac{\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}(\alpha_{1,1,t} - \alpha_{1,0,t}) - \alpha_{1,C}}{\sigma_\eta}\sigma_{\theta_1}^2.$$

- Using (37) and (38), we can identify $\alpha_{1,1,t}$ and $\alpha_{1,0,t}$ for all t.
- Note that if $T \geq 2$, we can also identify the parameters related to factor θ_2, such as $\alpha_{2,s,t}$ and $\sigma_{\theta_2}^2$.

- To see this, first normalize $\alpha_{2,0,1} = 1$ and compute the covariances:

$$\text{Cov}\left(Y_{0,1} - X\beta_{0,1}, Y_{0,2} - X\beta_{0,2}\right) - \alpha_{1,0,1}\alpha_{1,0,2}\sigma^2_{\theta_1} = \alpha_{2,0,2}\sigma^2_{\theta_2}, \quad (40)$$

$$\text{Cov}\left(Y_{0,1} - X\beta_{0,1}, \frac{I - \mu_I(X,Z)}{\sigma_\eta}\right) - \frac{\alpha_{1,0,1}\sigma^2_{\theta_1}\sum_{t=1}^{T}\left(\alpha_{1,1,t} - \alpha_{1,0,t} - \alpha_{1,C}\right)}{\sigma_\eta}$$
$$= \frac{\sigma^2_{\theta_2}\sum_{t=1}^{T}\left(\alpha_{2,1,t} - \alpha_{2,0,t} - \alpha_{2,C}\right)}{\sigma_\eta}, \quad (41)$$

$$\text{Cov}\left(Y_{0,2} - X\beta_{0,2}, \frac{I - \mu_I(X,Z)}{\sigma_\eta}\right) - \frac{\alpha_{1,0,2}\sigma^2_{\theta_1} \sum_{t=1}^{T} (\alpha_{1,1,t} - \alpha_{1,0,t} - \alpha_{1,C})}{\sigma_\eta}$$
$$= \frac{\alpha_{2,0,2}\sigma^2_{\theta_2} \sum_{t=1}^{T} (\alpha_{2,1,t} - \alpha_{2,0,t} - \alpha_{2,C})}{\sigma_\eta}. \quad (42)$$

- From (40) we can recover $\sigma^2_{\theta_2}$.

- We now add in the information on the covariances from the college earnings equation:

$$\mathrm{Cov}\left(Y_{1,1} - X\beta_{1,1}, Y_{1,2} - X\beta_{1,2}\right) - \alpha_{1,1,1}\alpha_{1,1,2}\sigma^2_{\theta_1} = \alpha_{2,1,1}\alpha_{2,1,2}\sigma^2_{\theta_2}, \tag{43}$$

$$\mathrm{Cov}\left(Y_{1,1} - X\beta_{1,1}, \frac{I - \mu_I(X,Z)}{\sigma_\eta}\right) - \frac{\alpha_{1,1}\sigma^2_{\theta_1}\sum_{t=1}^{T}\left(\alpha_{1,1,t} - \alpha_{1,0,t} - \alpha_{1,C}\right)}{\sigma_\eta} = \frac{\alpha_{2,1,1}\sigma^2_{\theta_2}\sum_{t=1}^{T}\left(\alpha_{2,1,t} - \alpha_{2,0,t} - \alpha_{2,C}\right)}{\sigma_\eta}, \tag{44}$$

$$\text{Cov}\left(Y_{1,2} - X\beta_{1,2}, \frac{I - \mu_I(X, Z)}{\sigma_\eta}\right) - \frac{\alpha_{1,1,2}\sigma^2_{\theta_1} \sum_{t=1}^{T} (\alpha_{1,1,t} - \alpha_{1,0,t} - \alpha_{1,C})}{\sigma_\eta}$$
$$= \frac{\alpha_{2,1,2}\sigma^2_{\theta_2} \sum_{t=1}^{T} (\alpha_{2,1,t} - \alpha_{2,0,t} - \alpha_{2,C})}{\sigma_\eta}. \quad (45)$$

Appendix

Motivating our approach in a traditional model of the returns to schooling

Identifying Information Sets in the Mincer Model of Schooling

- Consider decomposing the “returns” coefficient on schooling in an earnings equation into components that are known at the time schooling choices are made and components that are not known.

Identifying Information Sets in the Mincer Model of Schooling

- Consider decomposing the "returns" coefficient on schooling in an earnings equation into components that are known at the time schooling choices are made and components that are not known.
- Write discounted lifetime earnings of person *i* as

$$Y_i = \alpha + \rho_i S_i + U_i, \tag{46}$$

where ρ_i is the person-specific *ex post* return, S_i is years of schooling, and U_i is a mean zero unobservable.

- We seek to decompose ρ_i into two components $\rho_i = \eta_i + \nu_i$, where η_i is a component known to the agent when he/she makes schooling decisions and ν_i is revealed after the choice is made.

- We seek to decompose ρ_i into two components $\rho_i = \eta_i + \nu_i$, where η_i is a component known to the agent when he/she makes schooling decisions and ν_i is revealed after the choice is made.

- Schooling choices are assumed to depend on what is known to the agent at the time decisions are made, $S_i = \lambda(\eta_i, Z_i, \tau_i)$, where the Z_i are other observed determinants of schooling known to the agent and τ_i represents additional factors unobserved by the analyst but known to the agent.

- We seek to decompose ρ_i into two components $\rho_i = \eta_i + \nu_i$, where η_i is a component known to the agent when he/she makes schooling decisions and ν_i is revealed after the choice is made.

- Schooling choices are assumed to depend on what is known to the agent at the time decisions are made, $S_i = \lambda(\eta_i, Z_i, \tau_i)$, where the Z_i are other observed determinants of schooling known to the agent and τ_i represents additional factors unobserved by the analyst but known to the agent.

- If η_i is known to the agent and acted on, it enters the schooling choice equation. Even if it is known, it may not be acted on.

- If we correctly specify the variables that enter the outcome equation (X) and the variables in the choice equation (Z) that are known to the agent at the time schooling choices are made, local instrumental variable estimates (Heckman and Vytlacil, 2005) identify *ex ante* gross returns.

- If we correctly specify the variables that enter the outcome equation (X) and the variables in the choice equation (Z) that are known to the agent at the time schooling choices are made, local instrumental variable estimates (Heckman and Vytlacil, 2005) identify *ex ante* gross returns.

- The question is how to pick the information set.

- If we correctly specify the variables that enter the outcome equation (X) and the variables in the choice equation (Z) that are known to the agent at the time schooling choices are made, local instrumental variable estimates (Heckman and Vytlacil, 2005) identify *ex ante* gross returns.

- The question is how to pick the information set.

- We consider this problem in the context of the Card model, which, as previously noted, was designed only to estimate *ex post* returns.

The Card Model

- Card presents a version of the Mincer (1974) model, which writes log earnings for person i with schooling level S_i as

$$\ln y_i = \alpha_i + \rho_i S_i, \tag{47}$$

where the “rate of return” ρ_i varies among persons as does the intercept, α_i.

The Card Model

- Card presents a version of the Mincer (1974) model, which writes log earnings for person *i* with schooling level S_i as
$$\ln y_i = \alpha_i + \rho_i S_i, \tag{47}$$
where the "rate of return" ρ_i varies among persons as does the intercept, α_i.

- For the purposes of this discussion think of y_i as an annualized flow of lifetime earnings.

- Let $\alpha_i = \bar{\alpha} + \varepsilon_{\alpha_i}$ and $\rho_i = \bar{\rho} + \varepsilon_{\rho_i}$ where $\bar{\alpha}$ and $\bar{\rho}$ are the means of α_i and ρ_i.

- Let $\alpha_i = \bar{\alpha} + \varepsilon_{\alpha_i}$ and $\rho_i = \bar{\rho} + \varepsilon_{\rho_i}$ where $\bar{\alpha}$ and $\bar{\rho}$ are the means of α_i and ρ_i.

- Thus the means of ε_{α_i} and ε_{ρ_i} are zero.

- Let $\alpha_i = \bar{\alpha} + \varepsilon_{\alpha_i}$ and $\rho_i = \bar{\rho} + \varepsilon_{\rho_i}$ where $\bar{\alpha}$ and $\bar{\rho}$ are the means of α_i and ρ_i.
- Thus the means of ε_{α_i} and ε_{ρ_i} are zero.
- Earnings equation (47) can be written as

$$\ln y_i = \bar{\alpha} + \bar{\rho} S_i + \{\varepsilon_{\alpha_i} + \varepsilon_{\rho_i} S_i\}. \qquad (48)$$

- Let $\alpha_i = \bar{\alpha} + \varepsilon_{\alpha_i}$ and $\rho_i = \bar{\rho} + \varepsilon_{\rho_i}$ where $\bar{\alpha}$ and $\bar{\rho}$ are the means of α_i and ρ_i.
- Thus the means of ε_{α_i} and ε_{ρ_i} are zero.
- Earnings equation (47) can be written as

$$\ln y_i = \bar{\alpha} + \bar{\rho} S_i + \{\varepsilon_{\alpha_i} + \varepsilon_{\rho_i} S_i\}. \tag{48}$$

- Card's model generalizes Rosen's (1977) model to allow for psychic costs of schooling.

- Assuming a person-specific interest rate r_i, we obtain optimal schooling as

$$S_i = \frac{(\rho_i - r_i)}{k}, \tag{49}$$

where k is related to the curvature of psychic costs in schooling.

- Assuming a person-specific interest rate r_i, we obtain optimal schooling as

$$S_i = \frac{(\rho_i - r_i)}{k}, \tag{49}$$

where k is related to the curvature of psychic costs in schooling.

- In the Card model, ρ_i would be a rate of return if there were no direct costs of schooling and everyone faces a constant borrowing rate.

- Assuming a person-specific interest rate r_i, we obtain optimal schooling as

$$S_i = \frac{(\rho_i - r_i)}{k}, \tag{49}$$

where k is related to the curvature of psychic costs in schooling.

- In the Card model, ρ_i would be a rate of return if there were no direct costs of schooling and everyone faces a constant borrowing rate.
- Least squares will not estimate the mean growth rate of earnings with schooling $E(\rho_i)$ unless $\text{Cov}(\rho_i, \rho_i - r_i) = 0$.

- By definition, $\rho_i = \eta_i + \nu_i$.

- By definition, $\rho_i = \eta_i + \nu_i$.
- The cost is r_i.

- By definition, $\rho_i = \eta_i + \nu_i$.

- The cost is r_i.
- Suppose r_i depends on observables (Z_i) and unobservables (ε_i) such that

$$r_i = \gamma_0 + \gamma_1 Z + \varepsilon_i,$$

where ε_i has mean zero and is assumed to be independent of Z_i.

- If we are uncertain about which components of ρ_i enter the schooling equation, we may rewrite (49) as

$$S_i = \lambda_0 + \lambda_1 \eta_i + \lambda_2 \nu_i + \lambda_3 Z_i + \tau_i, \tag{50}$$

where $\lambda_0 = -\frac{\gamma_0}{k}$, $\lambda_1 = \frac{1}{k}$, $\lambda_2 = \frac{1}{k}$ if ν_i is in the information set at the time schooling choices are taken and $\lambda_2 = 0$ otherwise.

- If we are uncertain about which components of ρ_i enter the schooling equation, we may rewrite (49) as

$$S_i = \lambda_0 + \lambda_1 \eta_i + \lambda_2 \nu_i + \lambda_3 Z_i + \tau_i, \tag{50}$$

  where $\lambda_0 = -\frac{\gamma_0}{k}$, $\lambda_1 = \frac{1}{k}$, $\lambda_2 = \frac{1}{k}$ if ν_i is in the information set at the time schooling choices are taken and $\lambda_2 = 0$ otherwise.

- The remaining coefficients are $\lambda_3 = -\frac{\gamma_1}{k}$ and $\tau_i = -\frac{\varepsilon_i}{k}$.

- Suppose that we observe the cost of funds, r_i, and assume that r_i is independent of (ρ_i, α_i).

- Suppose that we observe the cost of funds, r_i, and assume that r_i is independent of (ρ_i, α_i).

- This assumes that the costs of schooling are independent of the "return" ρ_i and the payment to raw ability, α_i.

- Suppose that we observe the cost of funds, r_i, and assume that r_i is independent of (ρ_i, α_i).

- This assumes that the costs of schooling are independent of the "return" ρ_i and the payment to raw ability, α_i.

- Suppose that agents do not know ρ_i at the time they make their schooling decisions but instead know $E(\rho_i) = \bar{\rho}$.

- If agents act on the expected return to schooling, decisions are given by

$$S_i = \frac{\bar{\rho} - r_i}{k}$$

and *ex post* earnings observed after schooling are

$$\ln Y_i = \bar{\alpha} + \bar{\rho} S_i + \{(\alpha_i - \bar{\alpha}) + (\rho_i - \bar{\rho}) S_i\}.$$

- If agents act on the expected return to schooling, decisions are given by

$$S_i = \frac{\bar{\rho} - r_i}{k}$$

and *ex post* earnings observed after schooling are

$$\ln Y_i = \bar{\alpha} + \bar{\rho} S_i + \{(\alpha_i - \bar{\alpha}) + (\rho_i - \bar{\rho}) S_i\}.$$

- In the notation introduced in the Card model, $\eta_i = \bar{\rho}$ and $\nu_i = \rho_i - \bar{\rho}$.

- If agents act on the expected return to schooling, decisions are given by

$$S_i = \frac{\bar{\rho} - r_i}{k}$$

and *ex post* earnings observed after schooling are

$$\ln Y_i = \bar{\alpha} + \bar{\rho} S_i + \{(\alpha_i - \bar{\alpha}) + (\rho_i - \bar{\rho})\, S_i\}\,.$$

- In the notation introduced in the Card model, $\eta_i = \bar{\rho}$ and $\nu_i = \rho_i - \bar{\rho}$.

- $\lambda_2 = 0$ in equation (50) and $\lambda_1 = \frac{1}{k}$.

- In this case,

$$\text{Cov}(Y,S) = \bar{\rho}\,\text{Var}(S),$$

because $(\rho_i - \bar{\rho})$ is independent of S_i.

- In this case,

$$\text{Cov}(Y, S) = \bar{\rho}\,\text{Var}(S),$$

because $(\rho_i - \bar{\rho})$ is independent of S_i.

- Note that, under this information assumption, $(\bar{\alpha}, \bar{\rho})$ can be identified by least squares because $S_i \perp\!\!\!\perp [(\alpha_i - \bar{\alpha}), (\rho_i - \bar{\rho}) S_i]$, where "$\perp\!\!\!\perp$" denotes independence.

- If, on the other hand, agents know ρ_i at the time they make their schooling decisions, OLS breaks down for identifying $\bar{\rho}$ because ρ_i is correlated with S_i.

- If, on the other hand, agents know ρ_i at the time they make their schooling decisions, OLS breaks down for identifying $\bar{\rho}$ because ρ_i is correlated with S_i.

- We can identify $\bar{\rho}$ and the distribution of ρ_i using the method of instrumental variables.

- If, on the other hand, agents know ρ_i at the time they make their schooling decisions, OLS breaks down for identifying $\bar{\rho}$ because ρ_i is correlated with S_i.
- We can identify $\bar{\rho}$ and the distribution of ρ_i using the method of instrumental variables.
- Under our assumptions, r_i is a valid instrument for S_i.

- In this case,

$$\mathrm{Cov}\left(\ln Y_i, S\right) = \bar{\rho}\,\mathrm{Var}\left(S\right) + \mathrm{Cov}\left(S, \left(\rho - \bar{\rho}\right) S\right).$$

- In this case,

$$\text{Cov}\left(\ln Y_i, S\right) = \bar{\rho}\,\text{Var}\left(S\right) + \text{Cov}\left(S, \left(\rho - \bar{\rho}\right) S\right).$$

- Since we observe S and r_i, we can identify $\bar{\rho}$ and can construct the value of $(\rho - \bar{\rho})$ associated with each S, we can form both terms on the right hand side.

- In this case,

$$\mathrm{Cov}\,(\ln Y_i, S) = \bar{\rho}\,\mathrm{Var}\,(S) + \mathrm{Cov}\,(S, (\rho - \bar{\rho})\,S)\,.$$

- Since we observe S and r_i, we can identify $\bar{\rho}$ and can construct the value of $(\rho - \bar{\rho})$ associated with each S, we can form both terms on the right hand side.
- Under the assumption that agents do not know ρ but forecast it by $\bar{\rho}$, ρ is independent of S, so we can test for independence directly.

- In this case,

$$\text{Cov}(\ln Y_i, S) = \bar{\rho}\,\text{Var}(S) + \text{Cov}(S, (\rho - \bar{\rho})\,S).$$

- Since we observe S and r_i, we can identify $\bar{\rho}$ and can construct the value of $(\rho - \bar{\rho})$ associated with each S, we can form both terms on the right hand side.
- Under the assumption that agents do not know ρ but forecast it by $\bar{\rho}$, ρ is independent of S, so we can test for independence directly.
- In this case, the second term on the right hand side is zero and does not contribute to the explanation of $\text{Cov}(\ln Y, S)$.

- Note further that the Durbin (1954) – Wu (1973) – Hausman (1978) test can be used to compare the OLS and IV estimates, which should be the same under the model that assumes that ρ_i is not known at the time schooling decisions are made and that agents base their choice of schooling on $E(\rho_i) = \bar{\rho}$.

- Note further that the Durbin (1954) – Wu (1973) – Hausman (1978) test can be used to compare the OLS and IV estimates, which should be the same under the model that assumes that ρ_i is not known at the time schooling decisions are made and that agents base their choice of schooling on $E(\rho_i) = \bar{\rho}$.

- If the economist does not observe r_i, but instead observes determinants of r_i that are exogenous, it is still possible to conduct a Durbin-Wu-Hausman test to discriminate between the two hypotheses, but one cannot form $\text{Cov}(\rho, S)$ directly.

- This analysis shows that, provided one has a good instrument, it is possible to test for the information in the agent's information set.

- This analysis shows that, provided one has a good instrument, it is possible to test for the information in the agent's information set.
- However, the method is somewhat fragile.

- This analysis shows that, provided one has a good instrument, it is possible to test for the information in the agent's information set.
- However, the method is somewhat fragile.
- If we add selection bias to the Card model (so $E(\alpha \mid S)$ depends on S, something ruled out up to this point), we can identify $\bar{\rho}$ by IV (Heckman and Vytlacil, 1998)

- This analysis shows that, provided one has a good instrument, it is possible to test for the information in the agent's information set.
- However, the method is somewhat fragile.
- If we add selection bias to the Card model (so $E(\alpha \mid S)$ depends on S, something ruled out up to this point), we can identify $\bar{\rho}$ by IV (Heckman and Vytlacil, 1998)
- However, OLS is no longer consistent for $\bar{\rho}$ even if, in making their schooling decisions, agents forecast ρ_i using $\bar{\rho}$.

- Even if there is no selection bias, so that $E(\alpha \mid S)$ does not depend on S, in the case where r_i is not observed, the proposed testing approach based on the Durbin-Wu-Hausman test breaks down if we misspecify the information set.

- Even if there is no selection bias, so that $E(\alpha \mid S)$ does not depend on S, in the case where r_i is not observed, the proposed testing approach based on the Durbin-Wu-Hausman test breaks down if we misspecify the information set.

- Thus if we include the predictors of r_i that predict *ex post* gains $(\rho_i - \bar{\rho})$ and are correlated with S_i, we do not identify $\bar{\rho}$.

- Even if there is no selection bias, so that $E(\alpha \mid S)$ does not depend on S, in the case where r_i is not observed, the proposed testing approach based on the Durbin-Wu-Hausman test breaks down if we misspecify the information set.

- Thus if we include the predictors of r_i that predict *ex post* gains $(\rho_i - \bar{\rho})$ and are correlated with S_i, we do not identify $\bar{\rho}$.

- In general the Durbin-Wu-Hausman test is not informative on the stated question.

- For a familiar example, if local labor market variables proxy the opportunity cost of school (the r_i), and also predict the realization of *ex post* earnings $(\rho_i - \bar{\rho})$, they are invalid instruments.

- For a familiar example, if local labor market variables proxy the opportunity cost of school (the r_i), and also predict the realization of *ex post* earnings $(\rho_i - \bar{\rho})$, they are invalid instruments.

- The question of determining the appropriate information set is front and center and unfortunately cannot, in general, be inferred using IV methods and standard model specification tests.

- The method developed by Cunha, Heckman, and Navarro (2004, 2005) and Cunha and Heckman (2006a) exploits the covariance between S and the realized $\ln(y)$ to determine which components of $\ln(y)$ are known at the time schooling decisions are made.

- The method developed by Cunha, Heckman, and Navarro (2004, 2005) and Cunha and Heckman (2006a) exploits the covariance between S and the realized $\ln(y)$ to determine which components of $\ln(y)$ are known at the time schooling decisions are made.
- Their approach explicitly models selection bias and allows for measurement error in earnings.

- The method developed by Cunha, Heckman, and Navarro (2004, 2005) and Cunha and Heckman (2006a) exploits the covariance between S and the realized $\ln(y)$ to determine which components of $\ln(y)$ are known at the time schooling decisions are made.
- Their approach explicitly models selection bias and allows for measurement error in earnings.
- It does not rely on linearity of the schooling relationship in terms of $\rho - r$.

- The method developed by Cunha, Heckman, and Navarro (2004, 2005) and Cunha and Heckman (2006a) exploits the covariance between S and the realized $\ln(y)$ to determine which components of $\ln(y)$ are known at the time schooling decisions are made.
- Their approach explicitly models selection bias and allows for measurement error in earnings.
- It does not rely on linearity of the schooling relationship in terms of $\rho - r$.
- Their method recognizes the discrete nature of the schooling decision.