Separating Heterogeneity from Uncertainty
Decomposing Trends in Inequality in Earnings into Forecastable and Uncertain Components Extract (*JOLE*, 2016)

Flavio Cunha and James Heckman

Econ 350, Winter 2019

I. Introduction

Basic idea

- Decision variable C_1 (say consumption of an agent in the first period of life), depends on incomes $Y_1, \ldots, Y_T$ over horizon T that are realized after the consumption choice is taken.
- Permanent income hypothesis the correlation between C_1 and future Y_t is a measure of how much of future Y_t is known and acted on when agents make their consumption decisions.
- See, e.g., Flavin (1981).

Basic idea, Cont'd

- At issue is whether agents *act* on information that they know.
- They may not because:
  - **Credit constraints:** reduce the ability of agents to transfer known future income to the present.
  - They may be irrational.
- All statistical decompositions of earnings processes vulnerable to these criticisms

$$Y_t = Y_t^{\text{permanent}} + U_t^{\text{transitory}}$$

- Agents only imperfectly predict their future earnings using information set $\mathcal{I}_1$.
- Suppose that C_1 depends on future Y_t through expected present value, $E\left(PV_1 \middle| \mathcal{I}_1\right)$
- $PV_1 = \sum_{t=1}^{T} \frac{Y_t}{(1+\rho)^{t-1}}$, and ρ is the discount rate.
- Assumes asset market in which agents can lend or borrow against verifiable future income.
- After the choice of C_1 is made, we actually observe $Y_1, \ldots, Y_T$.
- 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, $V_1 = PV_1 - E\left(PV \middle| \mathcal{I}_1\right)$, should not predict C_1; $E\left(PV_1 \middle| \mathcal{I}_1\right)$ is predictable.
- V_1 arises from uncertainty.

- The variance in PV_1 that is unpredictable using $\mathcal{I}_1$ is a measure of uncertainty as of the first period.
- The Sims (1972) test for noncausality is based on a related idea in a linear prediction framework.
- Whereas Sims tests whether future Y_t predict current C_1, we measure what fraction of future Y_t predict current C_1 and use a more general prediction process.
- Use college attendance choices as its decision variable to estimate uncertainty.
- Accordingly, we measure uncertainty at only one stage of the life cycle.
- Can use consumption, labor supply, etc. (see Navarro and Zhou, 2017).

II. Model

A. Earnings Equations

- Roy model (1951)
- Two lifetime potential earnings streams, $(Y_{0,t}, Y_{1,t})$, $t = 1, \dots, T$, for schooling levels "0" and "1."

$$Y_{0,t} = \boldsymbol{X}\beta_{\mathbf{0,t}} + U_{0,t} \tag{1}$$

$$Y_{1,t} = \boldsymbol{X}\beta_{\mathbf{1,t}} + U_{1,t}, \qquad t = 1, \ldots, T, \tag{2}$$

- $E\left(U_{s,t} \mid \boldsymbol{X}\right) = 0,\ s = 0, 1,\ t = 1, \cdots, T.$

B. Choice Equations

$$I = E\left[\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(Y_{1,t} - Y_{0,t}\right) - C \,\middle|\, \mathcal{I}_1\right], \tag{3}$$

$$C = \boldsymbol{Z}\gamma + U_C. \tag{4}$$

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

and

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

- Substituting in (1), (2), and (4) into decision rule (3):

$$I = E\left[\mu_I(\boldsymbol{X}, \boldsymbol{Z}) + U_I \middle| \mathcal{I}_1\right]. \tag{5}$$

- $E\left(U_I|\mathcal{I}_1\right)$
- $U_{1,t}$, $U_{0,t}$, or U_C, depending on what is in the agent's information set.

$$S = \mathbf{1}\left[I \geq 0\right]. \tag{6}$$

C. Cognitive Ability

- Let M_k denote an agent's score on the k^{th} test.
- M_k have finite means and can be expressed in terms of conditioning variables $\boldsymbol{X}^M$.

D. Heterogeneity and Uncertainty

$$Y_{s,t} = E\left(Y_{s,t} \mid \mathcal{I}_1\right) + V_{s,t}, \qquad s = 0, 1, \quad t = 1, \ldots, T.$$

E. Factor Models

- Let factors and factor loadings
- $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$ and $\boldsymbol{\alpha_{s,t}} = (\alpha_{1,s,t}, \ldots, \alpha_{K,s,t})$.
- $\varepsilon_{s,t}, s \in \{0, 1\}, t \in \{1, \cdots, T\}$
- $\boldsymbol{\theta}, \boldsymbol{X}$ and $\boldsymbol{Z}$.
- $\theta \perp\!\!\!\perp (X, Z)$

$$U_{s,t} = \boldsymbol{\theta}\boldsymbol{\alpha}_{\boldsymbol{s},\boldsymbol{t}} + \varepsilon_{s,t}, \quad s = 0, 1, \quad t = 1, \ldots, T. \tag{7}$$

- Equation for psychic and pecuniary cost is decomposed in a fashion similar to the earnings equations

$$C = \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\theta}\boldsymbol{\alpha}_{\boldsymbol{C}} + \varepsilon_C, \tag{8}$$

$$I = \tag{9}$$

$$E\left[\ \sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\boldsymbol{X}\left(\boldsymbol{\beta_{1,t}}-\boldsymbol{\beta_{0,t}}\right)-\boldsymbol{Z}\gamma+\boldsymbol{\theta\alpha_{I}}+\sum_{t=1}^{T}\left(\frac{1}{1+\rho}\right)^{t-1}\left(\varepsilon_{1,t}-\varepsilon_{0,t}\right)-\varepsilon_{C}\ \middle|\mathcal{I}_1\right]$$

Define:

$$\boldsymbol{\alpha_I} = \sum_{t=1}^{T} \left(\frac{1}{1+\rho} \right)^{t-1} (\boldsymbol{\alpha_{1,t}} - \boldsymbol{\alpha_{0,t}}) - \boldsymbol{\alpha_C}.$$

F. Test Score Equations

(Measures of θ)

$$M_k = \boldsymbol{X}^M \boldsymbol{\beta_k^M} + \theta_1 \alpha_k^M + \varepsilon_k^M, k = 1, \ldots, K \tag{10}$$

G. The Estimation of Predictable Components of Future Earnings

Example

$$I = \mu_I(\boldsymbol{X}, \boldsymbol{Z}) + \alpha_{1,I}\theta_1 + \alpha_{2,I}\theta_2 + \varepsilon_C. \tag{11}$$

- Using standard results in the theory of discrete choice, we can proceed as if we observe I in equations (6) and (11) up to an unknown positive scale.
- Thus from the discrete choice on schooling we observe the index generating the choices up to scale.
- From the correlation between S and realized incomes, we can form (up to scale) the covariance between I and $Y_{s,t}$, $t = 1, \ldots, T$ for $s = 0$ or 1.

Conditional on $\boldsymbol{X}, \boldsymbol{Z}$ this covariance is

$$\begin{aligned} \text{Cov}\left(I, Y_{s,t} | \boldsymbol{X}, \boldsymbol{Z}\right) = \alpha_{1,I}\alpha_{1,s,t}\sigma^2_{\theta_1} + \alpha_{2,I}\alpha_{2,s,t}\sigma^2_{\theta_2}, \\ s = 0, 1, t = 1, \cdots, T. \end{aligned} \quad (12)$$

- Suppose next that θ_2 is not known, or is known and not acted on by the agent when schooling choices are made.
- In this case, $\alpha_{2,I} = 0$.
- If neither θ_2 nor θ_1 is known, or acted on by the agent, $\alpha_{1,I} = \alpha_{2,I} = 0$.
- For panels of earnings histories of length 3 or more ($T \geq 3$) and with three or more measures of cognition ($K \geq 3$), we can use the system of covariances in (12) joined with the information from the covariances between M_k and I and M_k and $Y_{s,t}$ to identify the model and infer the number of factors.

Applications

- Carneiro et al. (2005)
- Cunha et al. (2005)
- Heckman et al. (2006)
- Abbring and Heckman (2007)
- Cunha and Heckman (2008)

- The cited papers establish conditions for identifying $\sigma^2_{\theta_1}, \sigma^2_{\theta_2}, \alpha_{1,s,t}$ and $\alpha_{2,s,t}$, $s = 0, 1$, $t = 1, \ldots, T$. (See Cunha et al., 2005.)
- If component (factor) θ_1 appears in the period t earnings equation ($\alpha_{1,s,t} \neq 0$) is correlated with I and is acted on by the agent in making schooling choices (so $\alpha_{1,I} \neq 0$), then θ_1 is predictable (in $\mathcal{I}_1$) at the time schooling decisions are being made.
- If earnings component θ_2 is uncorrelated with I, then $\alpha_{2,I} = 0$ and θ_2 is not acted on by the agent in making schooling choices and we say that it is unpredictable at the time schooling choices are made.

III. Empirical Results

Table 1: Schooling Choice and Rates of Return per Year of College: Comparison Across Cohorts

	NLS/66	**NLSY/79**
High School Graduates	58.17%	64.19%
College Graduates	41.83%	35.81%
Mincer Returns to College[1]	9.01%	11.96%
Mincer Returns to College[2]	10.17%	12.41%
Mincer Returns to College[3]	8.17%	11.00%

[1] Pooled OLS Regression, controlling only for Mincer Experience and Mincer Experience Squared

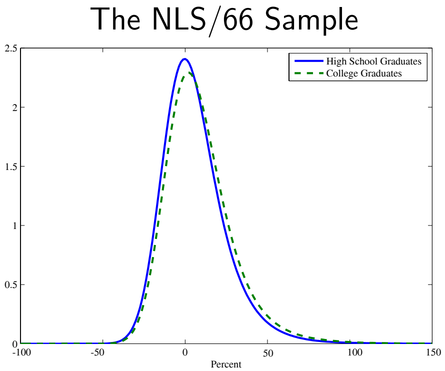
[2] Pooled OLS Regression, controlling for Mincer Experience, Mincer Experience Squared, and Year Dummies

[3] Pooled OLS Regression, controlling for Mincer Experience, Mincer Experience Squared, Cognitive Skills, Urban and South Residence at Age 14, and Year Dummies (Dependent Variable: Log Earnings).

Table 2: Mean Rates of Return per Year of College by Schooling Group

	NLS/66		NLSY/79	
Schooling Group	Mean Returns	Standard Error	Mean Returns	Standard Error
High School Graduates	0.0592	0.0046	0.0955	0.0063
College Graduates	0.0877	0.0070	0.1355	0.0080
Individuals at the Margin	0.0750	0.0178	0.1184	0.0216

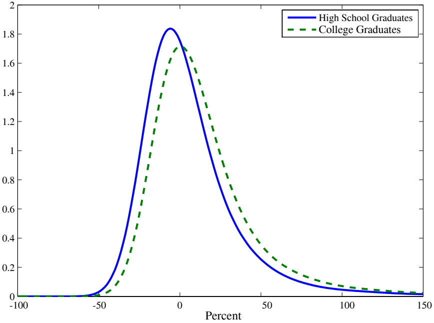
Figure 1: Densities of Returns to College



Let Y_0, Y_1 denote the present value of earnings from age 22 to age 36 in the high school and college sectors, respectively. Define ex post returns to college as the ratio $R = (Y_1 - Y_0)/Y_0$. Let $f(r)$ denote the density function of the random variable R. The solid line is the density of ex post returns to college for high school graduates, that is $f(r|S = 0)$. The dashed line is the density of ex post returns to college for college graduates, that is, $f(r|S = 1)$. This assumes that the agent chooses schooling without knowing θ_3 and the innovations $\varepsilon_{s,t}$ for s = high school, college and $t = 22, \cdots, 36$.

Figure 1: Densities of Returns to College

The NLSY/79 Sample



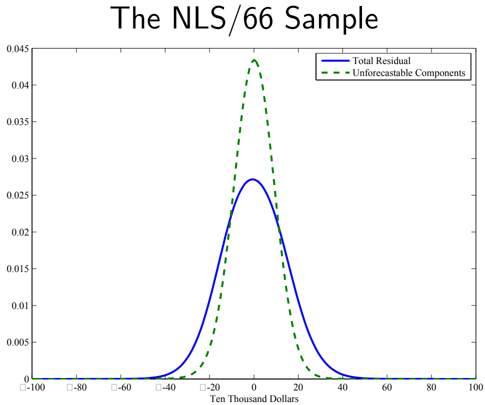
Let Y_0, Y_1 denote the present value of earnings from age 22 to age 36 in the high school and college sectors, respectively. Define ex post returns to college as the ratio $R = (Y_1 - Y_0)/Y_0$. Let $f(r)$ denote the density function of the random variable R. The solid line is the density of ex post returns to college for high school graduates, that is $f(r|S=0)$. The dashed line is the density of ex post returns to college for college graduates, that is, $f(r|S=1)$. This assumes that the agent chooses schooling without knowing θ_3 and the innovations $\varepsilon_{s,t}$ for $s =$ high school, college and $t = 22, \cdots, 36$.

Table 3: Evolution of Uncertainty

	College	High School	Returns
NLS/1966			
	College	High School	Returns
Total Variance	195.882	136.965	611.245
Variance of Unforecastable Components	76.332	31.615	167.187
Variance of Forecastable Components	119.550	105.350	444.058
NLS/1979			
	College	High School	Returns
Total Variance	292.368	165.350	823.200
Variance of Unforecastable Components	84.464	48.137	221.976
Variance of Forecastable Components	207.904	117.214	601.223
Evolution			
Percentage Increase in Total Variance	49.26%	20.72%	34.68%
Percentage Increase in Variance of Unforecastable Components	10.65%	52.26%	32.77%
Percentage Increase in Variance of Forecastable Components	73.90%	11.26%	35.39%
Percentage Increase in Total Variance by Source			
	College	High School	Returns
Percentage Increase in Total Variance due to Unforecastable Components	8.43%	58.20%	25.85%
Percentage Increase in Total Variance due to Forecastable Components	91.57%	41.80%	74.15%

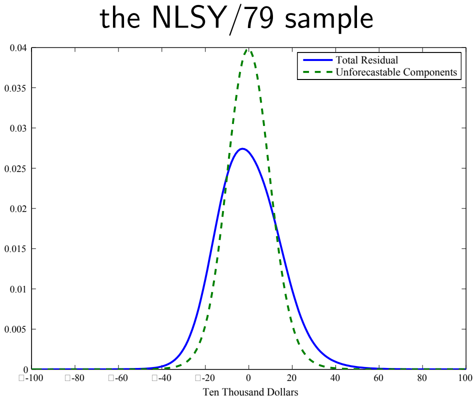
THE UNIVERSITY OF CHICAGO

Figure 2: The Densities of Total Residual vs. Unforecastable Components in Present Value of High School Earnings



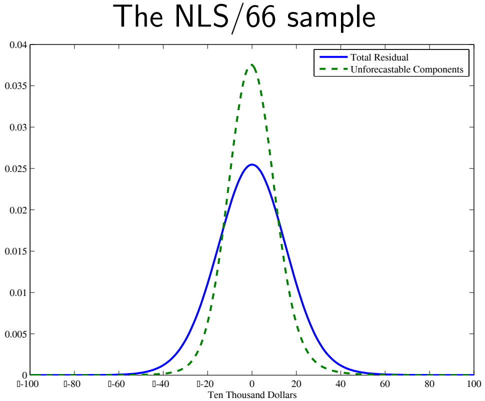
In this figure we plot the density of total residual (the solid curve) against the density of the unforecastable components (the dashed curve) for the present value of high school earnings from ages 22 to 36 for the NLS/66 and NLSY/79 samples of white males. The present value of earnings is calculated using a 5% interest rate.

Figure 2: The Densities of Total Residual vs. Unforecastable Components in Present Value of High School Earnings



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 36 for the NLS/66 and NLSY/79 samples of white males. The present value of earnings is calculated using a 5% interest rate.

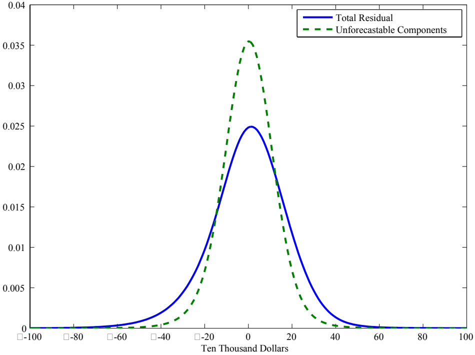
Figure 3: The Densities of Total Residual vs. Unforecastable Components in Present Value of College Earnings



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 36 for the NLS/66 and NLSY/79 samples of white males. The present value of earnings is calculated using a 5% interest rate.

Figure 3: The Densities of Total Residual vs. Unforecastable Components in Present Value of College Earnings

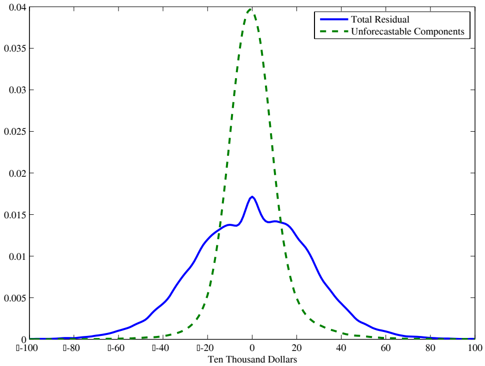
The NLSY/79 sample



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 36 for the NLS/66 and NLSY/79 samples of white males. The present value of earnings is calculated using a 5% interest rate.

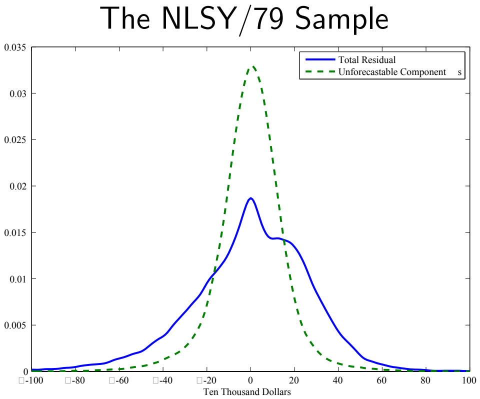
Figure 4: The Densities of Total Residual vs. Forecastable Components Returns College vs. High School

The NLS/66 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) from ages 22 to 36 for the NLS/66 and NLSY/79 samples of white males. The present value of earnings is calculated using a 5% interest rate.

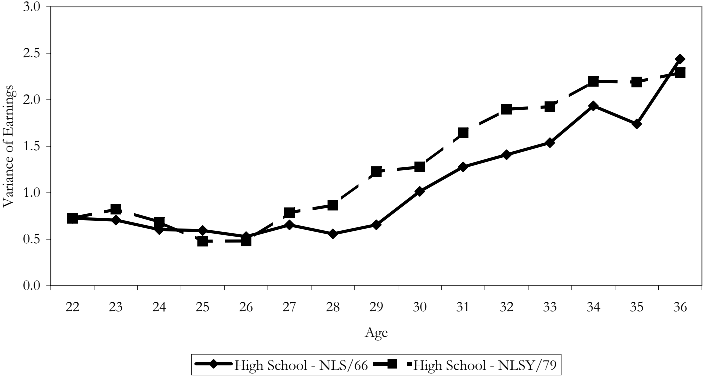
Figure 4: The Densities of Total Residual vs. Forecastable Components Returns College vs. High School



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) from ages 22 to 36 for the NLS/66 and NLSY/79 samples of white males. The present value of earnings is calculated using a 5% interest rate.

Figure 5: Profile of Variance of Uncertainty

High School Sample, NLS/66 vs NLSY/79



THE UNIVERSITY OF CHICAGO

Figure 5: Profile of Variance of Uncertainty

College Sample, NLS/66 vs NLSY/79

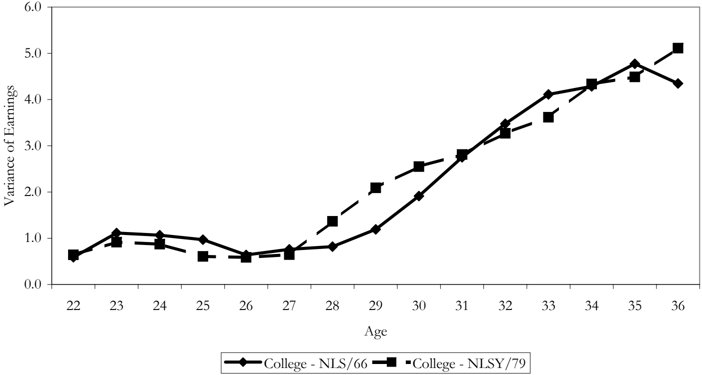


Figure 6: Profile of Variance of Heterogeneity

High School Sample, NLS/66 vs NLSY/79

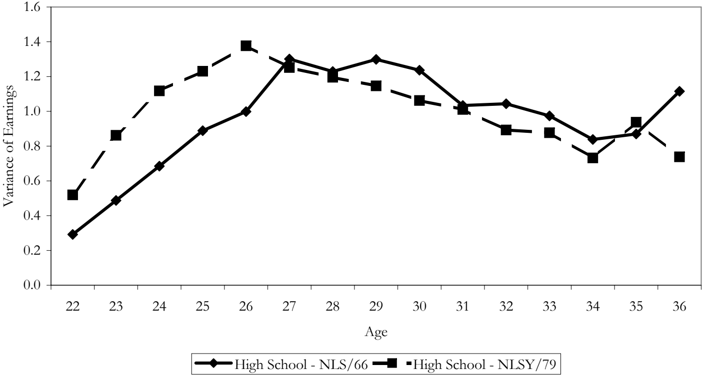
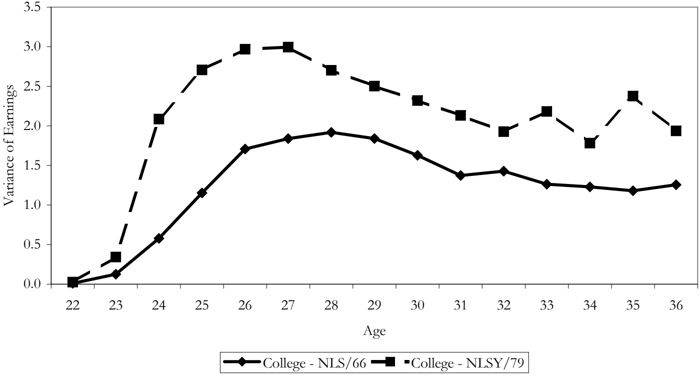


Figure 6: Profile of Variance of Heterogeneity

College Sample, NLS/66 vs NLSY/79



THE UNIVERSITY OF CHICAGO

Table 4: 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	0.1111	0.0147	0.0156	0.0020
College	0.0452	0.0077	0.0392	0.0052
Overall	0.0679	0.0107	0.0328	0.0042

Table 5: Predictable Heterogeneity

A. Gini Decomposition			
	NLS/66	NLSY/79	% Growth
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.1803	0.2088	15.85%
Counterfactual: Predictable Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.1591	0.1825	14.73%

B. The Theil Entropy Index T (Overall)			
	NLS/66	NLSY/79	% Growth
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.0502	0.0693	37.98%
Counterfactual: Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.0390	0.0522	33.76%
Within Schooling Groups			
	NLS/66	NLSY/79	% Change
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.0491	0.0631	28.53%
Counterfactual: Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.0378	0.0465	22.85%
Between Schooling Groups			
	NLS/66	NLSY/79	% Change
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.0011	0.0062	447.37%
Counterfactual: Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.0011	0.0057	394.22%

- [1] Let $Y_{k,s,t,i}$ denote the earnings of an agent i, $i = 1, ..., n_k$, at age t, $t = 1, ..., T$, in schooling level s, $s =$ high school, college, and cohort k, $k = NLS/1966, NLSY/1979$.
- We model earnings $Y_{k,s,t,i}$ as:

$$Y_{k,s,t,i} = \mu_{s,k}\left(\boldsymbol{X}_{k,i}\right) + \theta_{1,k,i}\alpha_{1,k,s,t} + \theta_{2,k,i}\alpha_{2,k,s,t} + \theta_{3,k,i}\alpha_{3,k,s,t} + \varepsilon_{k,s,t,i}. \tag{1}$$

- Present value of earnings in schooling level s, $Y_{k,s,i}$, is $Y_{k,s,i} = \sum_{t=1}^{T^*} \frac{Y_{k,s,t,i}}{(1+\rho)^{t-1}}$.
- Observed truncated present value of earnings is $Y_{k,i} = S_{k,i} Y_{k,1,i} + (1 - S_{k,i}) Y_{k,0,i}$.
- Let $C_{k,i}$ denote the direct costs for individual i in cohort k.
- The schooling choice is:

$$S_{k,i} = 1 \Leftrightarrow E\left(Y_{k,1,i} - Y_{k,0,i} - C_{k,i} \middle| \mathcal{I}_k \right) \geq 0. \tag{2}$$

- This is the factual economy. We then compute the average present value of earnings across all individuals in cohort k, $\mu_k = \frac{1}{n} \sum_{i=1}^{n_k} Y_{k,i}$.
- For a given inequality aversion parameter ϵ, we compute the level of permanent income $\bar{Y}_k(\epsilon)$ that generates the same welfare as the social welfare of the actual distribution in cohort k:

$$\frac{\left[\bar{Y}_k(\epsilon)\right]^{1-\epsilon} - 1}{1-\epsilon} = \frac{1}{n_k} \sum_{i=1}^{n_k} \frac{(Y_{k,i})^{1-\epsilon} - 1}{1-\epsilon}$$

- For each value of ϵ, the Atkinson Index is $A(\epsilon) = 1 - \frac{\bar{Y}_k(\epsilon)}{\mu_k}$.
- In this row, we show the Atkinson Index for the observed present value of earnings $Y_{k,i}$ for different values of ϵ.
- [2] We simulate the economy by replacing (1) with:

$$Y^h_{k,s,t,i} = \mu_{s,k}(\boldsymbol{X}_{k,i}) + \theta_{1,k,i}\alpha_{1,k,s,t} + \theta_{2,k,i}\alpha_{2,k,s,t},$$

where $Y^h_{k,s,t,i}$ are the individual earnings when idiosyncratic uncertainty is completely shut down.

- The present value of earnings when only predictable heterogeneity is accounted for is constructed in a similar manner: $Y^h_{k,s,i} = \sum_{t=1}^{T^*} \frac{Y^h_{k,s,t,i}}{(1+\rho)^{t-1}}$.
- The schooling choices are as determined in (2).
- In this row, we show the Atkinson Index for the observed present value of earnings $Y^h_{k,i}$ for different values of ϵ when we constrain schooling choices, $S_{k,i}$, to be observed in the factual economy.

Table 6: Atkinson Index

$$A(\varepsilon) = 1 - \frac{\overline{Y}_k(\varepsilon)}{\mu_k}$$

		$\varepsilon = 0.5$	
	NLS/66	NLSY/79	%Change
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.0276	0.0389	0.4111
Counterfactual: Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.0213	0.0286	0.3437
		$\varepsilon = 1.5$	
	NLS/66	NLSY/79	%Change
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.0968	0.1467	0.5147
Counterfactual: Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.0716	0.0980	0.3687

Table 6: Atkinson Index, Cont.

		$\varepsilon = 1.0$	
	NLS/66	NLSY/79	%Change
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.0586	0.0847	0.4446
Counterfactual: Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.0447	0.0604	0.3503
		$\varepsilon = 2.0$	
	NLS/66	NLSY/79	%Change
Factual Economy: Predictable Heterogeneity and Uncertainty[1]	0.1627	0.2627	0.6149
Counterfactual: Fixing Schooling Choices as in Factual Economy			
Predictable Heterogeneity Only[2]	0.1060	0.1506	0.4205

Figure 7: Mean Earnings Profile NLSY/66, Comparison Across Schooling Within Cohorts

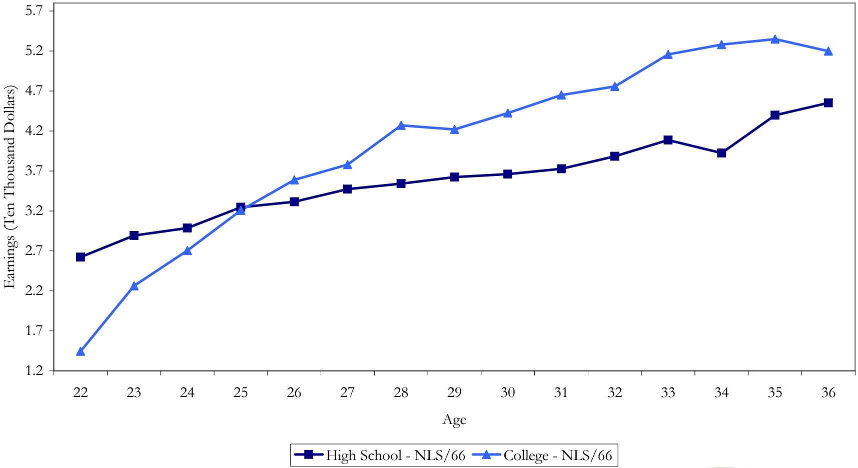


Figure 8: Mean Earnings Profile NLSY/79, Comparison Across Schooling Within Cohorts

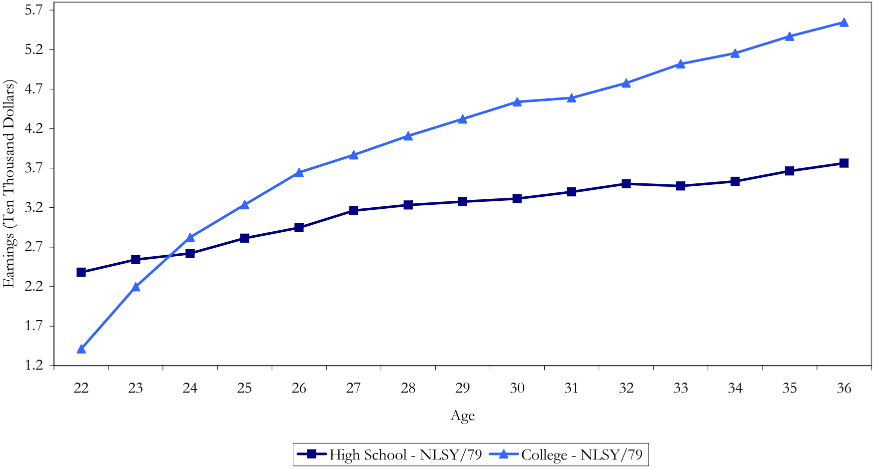


Figure 9: Standard Deviation of Earnings, High School Sample, Comparison Within Schooling Groups Across Cohorts

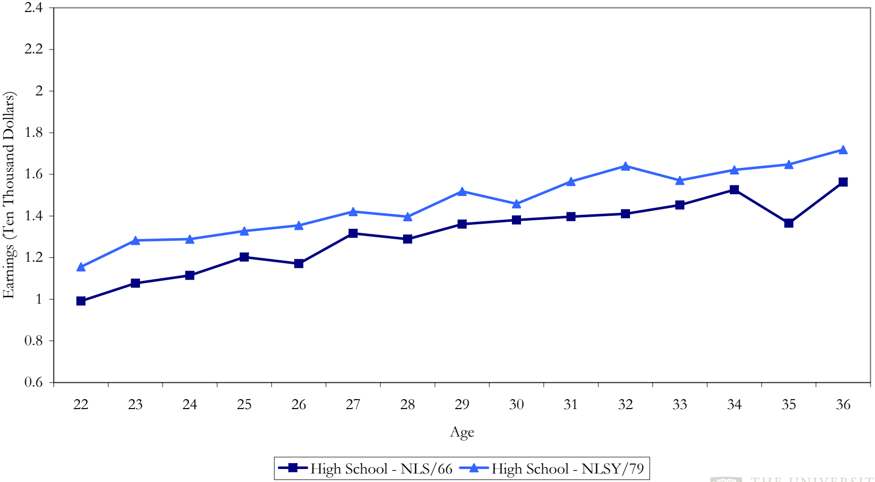


Figure 10: Standard Deviation of Earnings, College Sample, Comparison Within Schooling Groups Across Cohorts

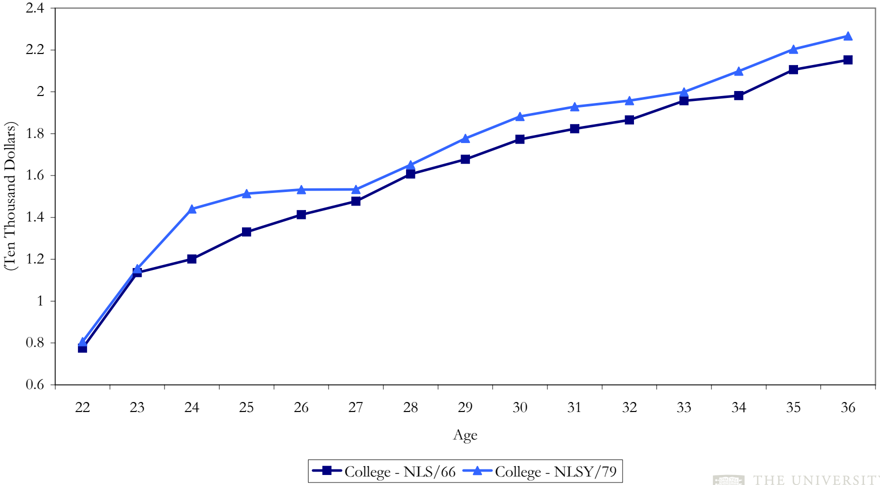
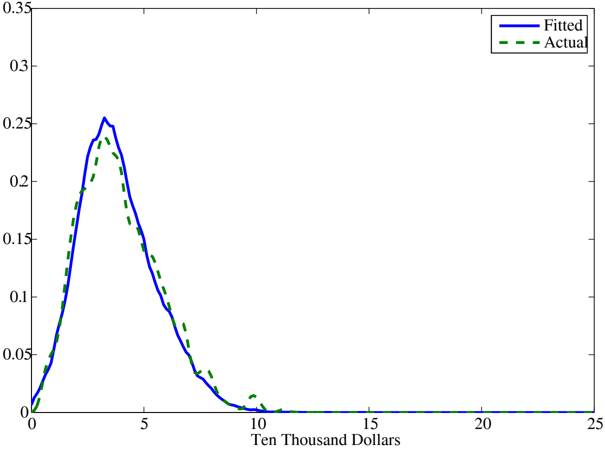
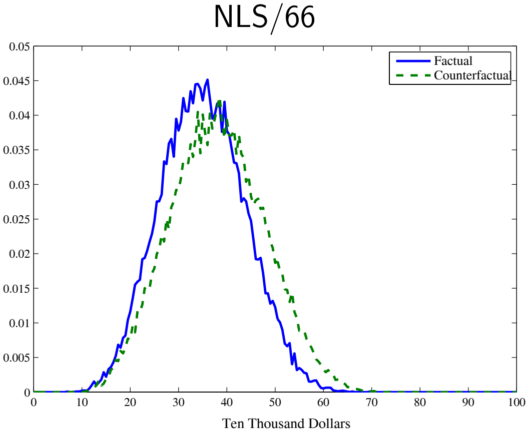


Figure 11: Densities of Earnings at Age 33, Overall Sample NLSY/79



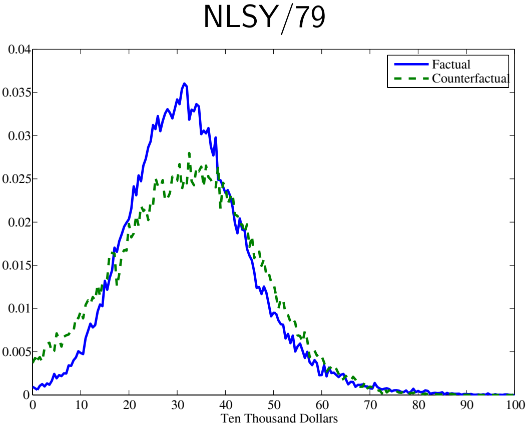
Let Y denote earnings at age 33 in the overall sample. Here we plot the density functions $f(y)$ generated from the data (the solid curve), against that predicted by the model (the dashed line).

Figure 12: Densities of Present Value Earnings, High School Sample



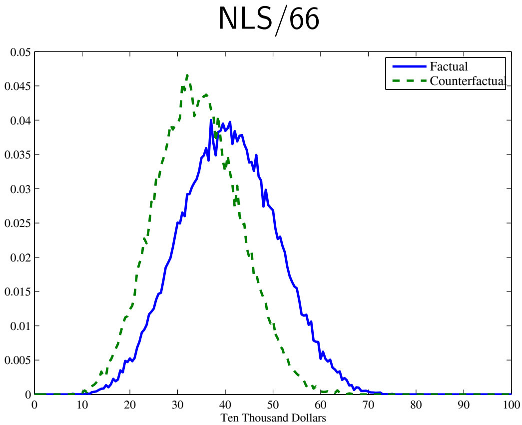
Present value of earnings from age 22 to 36 for High School Graduates discounted using an interest rate of 5%. Here we plot the factual density function $f(y_0|S=0)$ (the solid curve), against the counterfactual density function $f(y_1|S=0)$ (the dashed line). We use kernel density estimation to smooth these functions.

Figure 12: Densities of Present Value Earnings, High School Sample



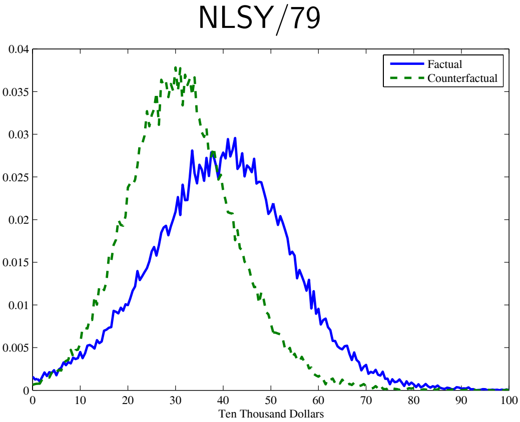
Present value of earnings from age 22 to 36 for High School Graduates discounted using an interest rate of 5%. Here we plot the factual density function $f(y_0|S=0)$ (the solid curve), against the counterfactual density function $f(y_1|S=0)$ (the dashed line). We use kernel density estimation to smooth these functions.

Figure 13: Densities of Present Value of Earnings, College Sample



Present value of earnings from age 22 to 36 for College Graduates discounted using an interest rate of 5%. 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 line). We use kernel density estimation to smooth these functions.

Figure 13: Densities of Present Value of Earnings, College Sample



Present value of earnings from age 22 to 36 for College Graduates discounted using an interest rate of 5%. 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 line). We use kernel density estimation to smooth these functions.