MIT 14.662 Spring 2026: Lecture 2 — Educational Production and Wage Structure

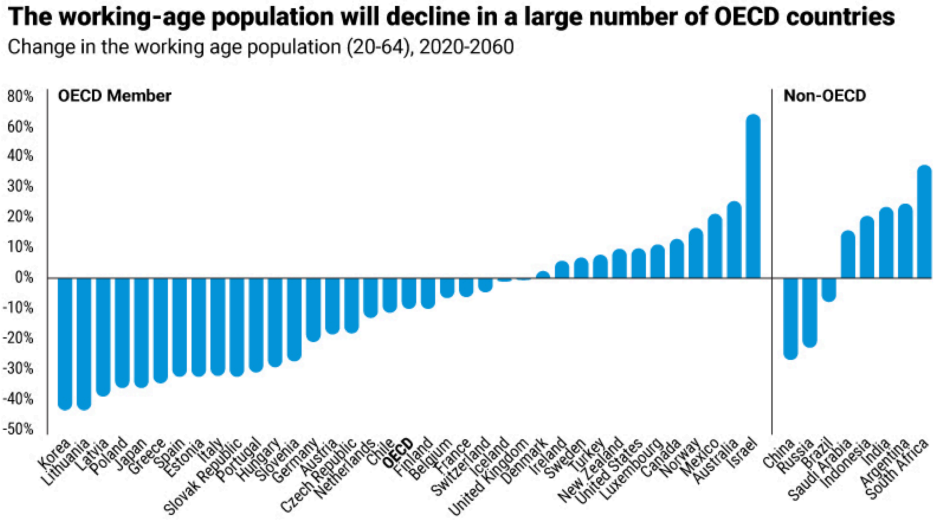
David Autor, MIT and NBER

February 4, 2026 (rev 2026/02/04)

Agenda

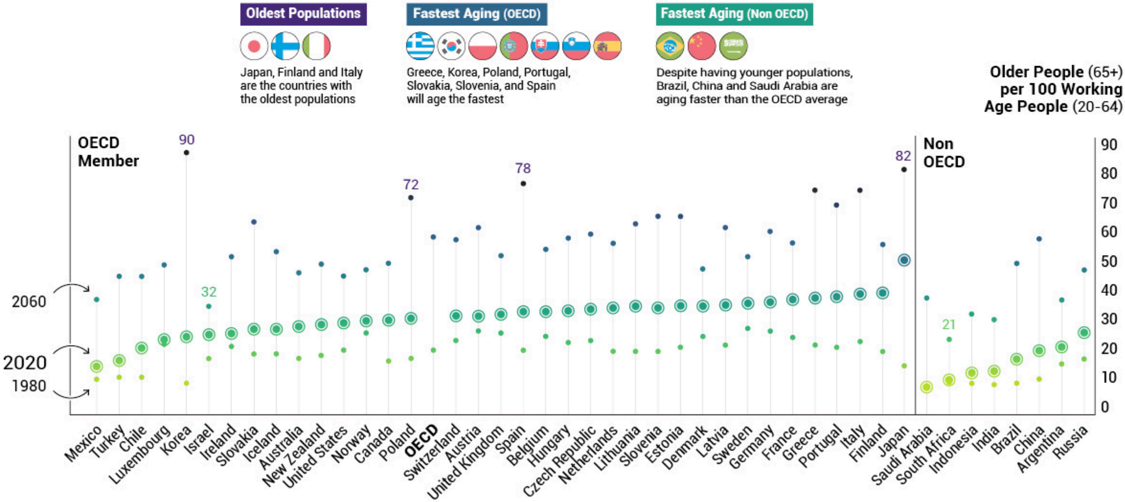
1. **Some motivating fact patterns**
2. **Supply shifts and cohort effects**
   - The puzzling age-profile of the college premium
   - Formalizing a model with imperfect age-group substitutability
   - Cohort effects in a model without cohort quality differences
   - Estimation of full model
3. **Has there been a decline in the quality of college graduates?**
   - Motivation
   - Estimating the Carneiro-Lee model in light of Card-Lemieux

Shrinking working-age populations in OECD



OECD 2020

Rising dependency ratios worldwide



The Visual Capitalist 2020

Why demographics matter for wage structure

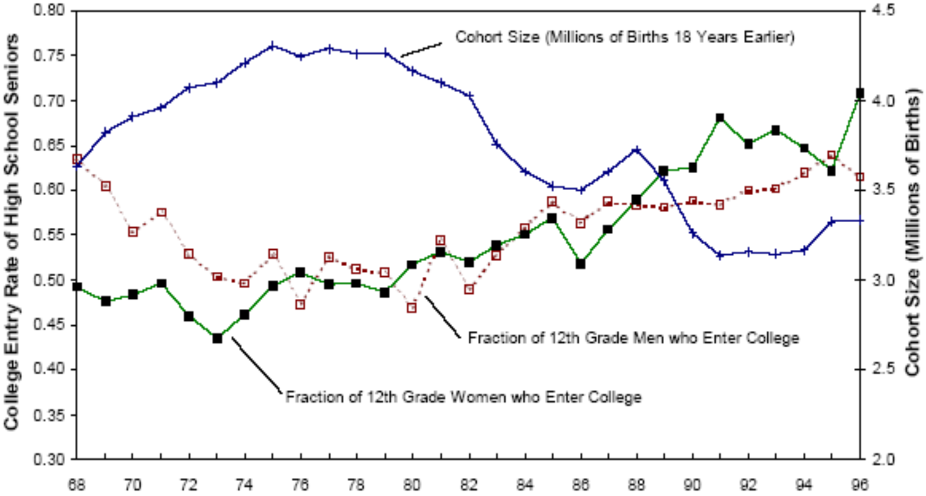
Cohort size and educational attainment jointly determine labor supply

- Smaller birth cohorts $\rightarrow$ fewer workers entering the labor market
- Changes in college attainment rates amplify or offset these effects
- Together, these determine the relative supply of college versus non-college labor
- Relative supply shifts $\rightarrow$ changes in the college wage premium

This lecture

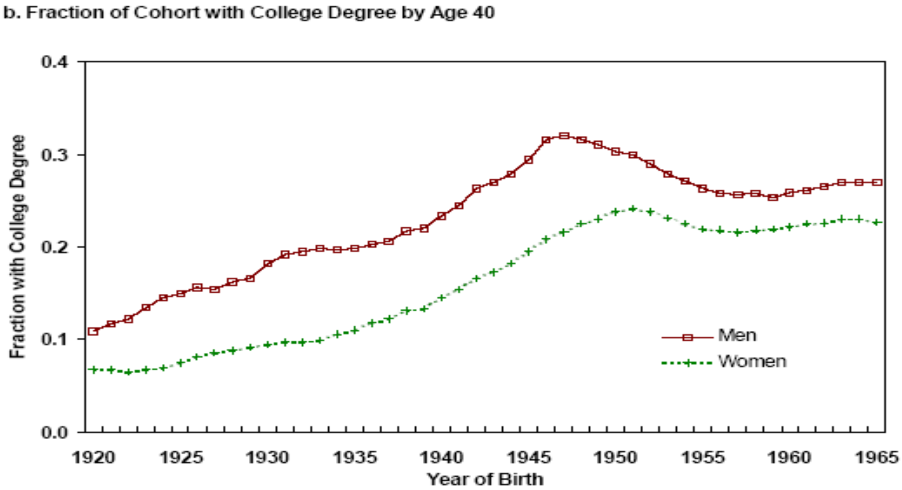
- How do supply shifts affect the wage structure?
- Why do cohort effects emerge even without quality differences?
- Has the quality of college graduates declined as enrollment expanded?

Birth cohort size and college completion



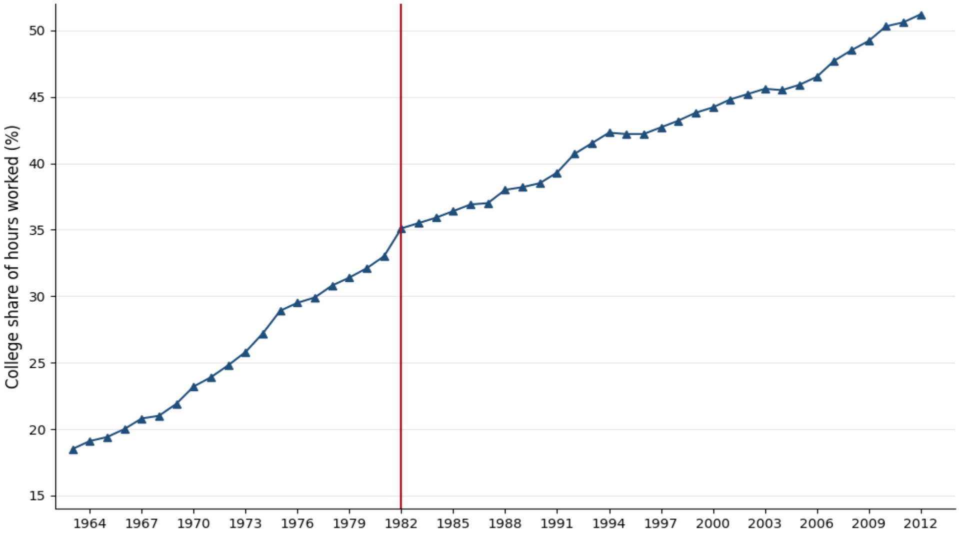
Card and Lemieux, 2002

The plateauing of college attainment in the US

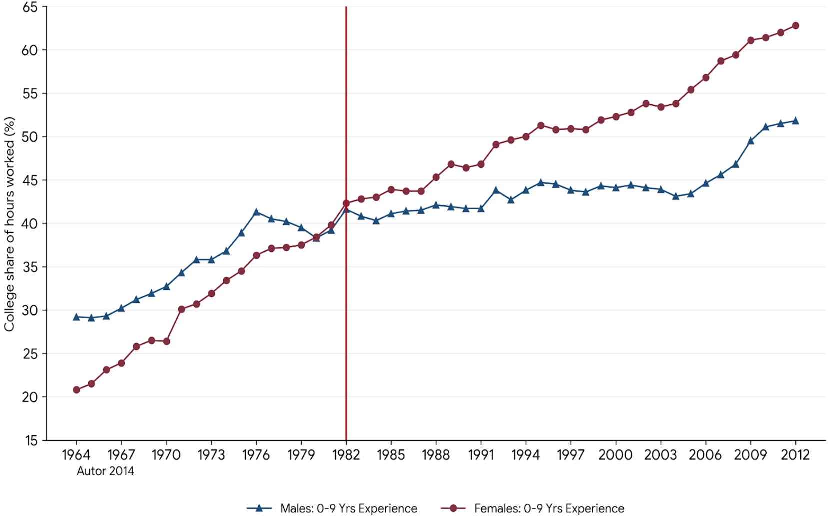


Card and Lemieux, 2002

College share of US hours worked (%), 1963 – 2012



College share of US hours worked, 1963 – 2012: Men and women with <10 years of potential experience



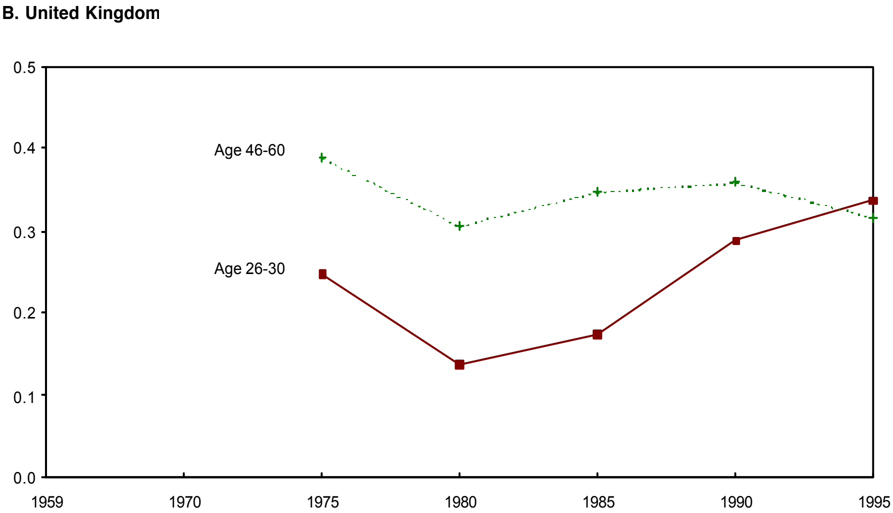
Agenda

Agenda

Card-Lemieux '01: "Can Falling Supply Explain the Rising Return to College for Younger Men?" College/high-school wage diffs among US men, 1959–1995

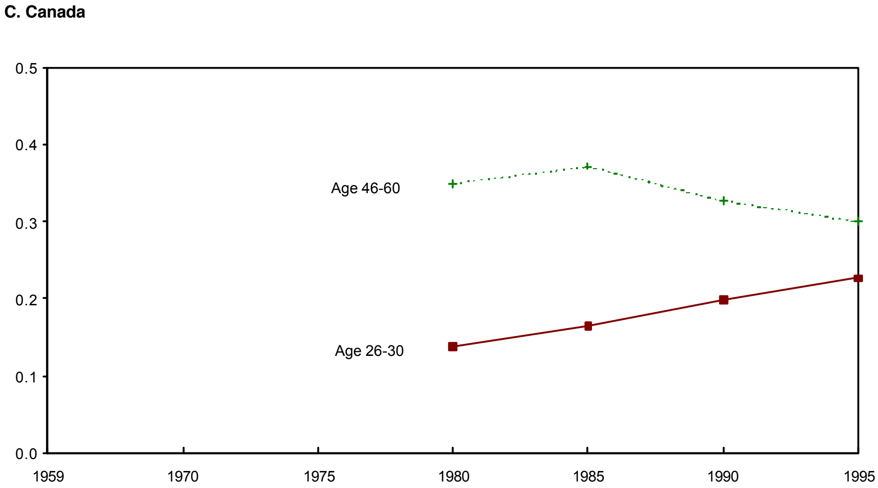
Card and Lemieux 2001, QJE

College/high-school wage diffs for younger and older men in the U.K.



Card and Lemieux 2001, QJE

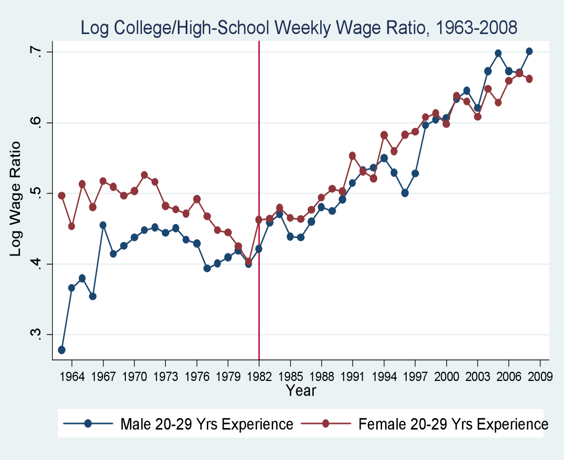
College/high-school wage diffs for younger and older men in Canada

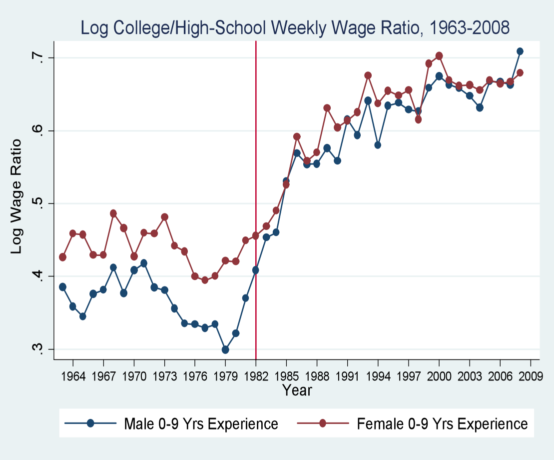


Card and Lemieux 2001, QJE

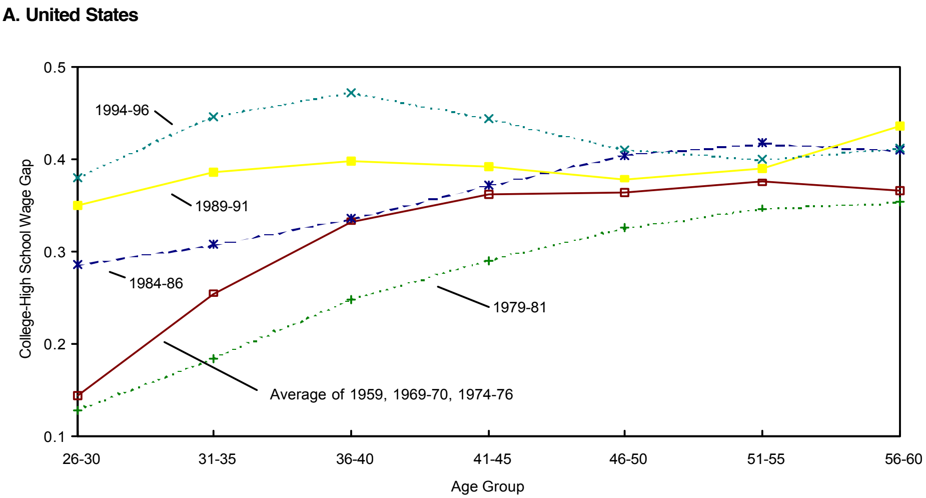
College premium: Young v. experienced workers in US

College Premium among Experienced Workers (Left Panel)
versus Recent Entrants (Right Panel)



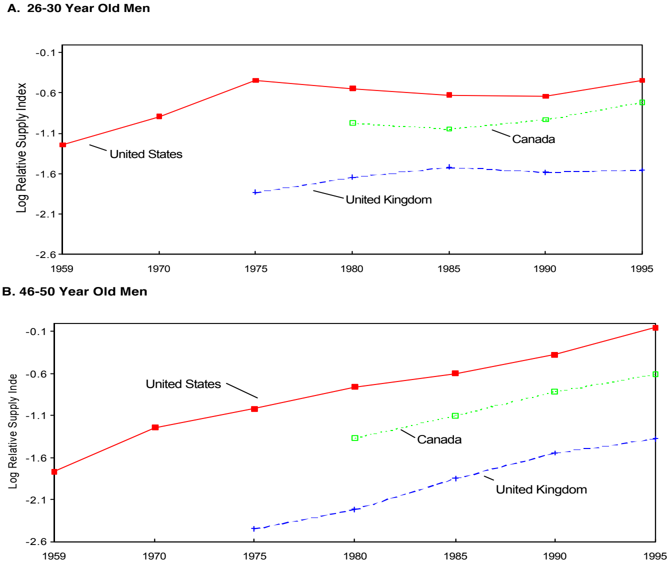


Age profile of college/high-school premium: 1958-1996



Card and Lemieux 2001, QJE

Age-group specific relative supplies



Card and Lemieux 2001, QJE

Agenda

1. **Some motivating fact patterns**
2. **Supply shifts and cohort effects**
   - The puzzling age-profile of the college premium
   - Formalizing a model with imperfect age-group substitutability
   - Cohort effects in a model without cohort quality differences
   - Estimation of full model
3. **Has there been a decline in the quality of college graduates?**
   - Motivation
   - Estimating the Carneiro-Lee model in light of Card-Lemieux

Card-Lemieux 2001: Formalization

Nested, two-level CES model

- Upper level: Identical to simple Katz-Murphy model
  - Output a f'n of H_t, L_t, A_{Ht}, A_{Lt}
- Lower level: Supplies of each education group are themselves CES aggregates of the labor supply of different age groups
  - Aggregate education supplies H_t, L_t depend on age-group specific supplies H_{jt}, L_{jt}

Card-Lemieux 2001: Formalization

Education supplies (lower level of CES model)

$$H_t = \left(\sum_j \alpha_j H_{jt}^{\eta}\right)^{1/\eta} \text{ and } L_t = \left(\sum_j \beta_j L_{jt}^{\eta}\right)^{1/\eta}$$

- $\sigma_A = 1/(1-\eta)$ is the elasticity of subst across age groups j
- α_j and β_j are efficiency parameters, assumed fixed by age group (do not vary across cohorts or over time)
- H_{jt}, L_{jt} are age-group specific supplies of H and L workers in period t
- Note: if $\eta = 1 \to \sigma_A = \infty$, age groups are perfect substitutes—though they may have different efficiencies, given by α_j, β_j

Aggregate output

- Determined by *total* H and L supply and time-varying efficiency parameters A_{Ht}, A_{Lt}

$$Y_t = f(H_t, L_t, A_{Ht}, A_{Lt})$$

- Assume $Y(\cdot)$ is also CES

$$Y_t = (A_{Ht} H_t^{\rho} + A_{Lt} L_t^{\rho})^{1/\rho}$$

with $\sigma = 1/(1-\rho)$, where σ is the aggregate elasticity of subst btwn H and L workers

Card-Lemieux 2001: Wages

Competitive wage setting — Wages equal marginal products

- Economy operates on the demand curve

$$w_{jt} = \frac{\partial Y_t}{\partial L_{jt}} = \frac{\partial Y_t}{\partial L_t} \times \frac{\partial L_t}{\partial L_{jt}}$$
$$= A_{Lt} L_t^{\rho-\eta} \pi_t \times \beta_j L_{jt}^{\eta-1}$$

- where

$$\pi_t = (A_{Lt} L_t^{\rho} + A_{Ht} H_t^{\rho})^{1/\rho-1},$$

  and similarly for the wages of college graduates

- Provided that $\eta < 1$, the age-specific wage (by education) is declining in age-specific supply
- CES nesting dolls

Card-Lemieux 2001: Relative wages

Relative wages of H versus L workers in same cohort

$$\ln\left(\frac{w_{jt}^{H}}{w_{jt}^{L}}\right) \equiv r_{jt} = \ln(A_{Ht}/A_{Lt}) + \ln(\alpha_j/\beta_j) - \frac{1}{\sigma}\ln\left(H_t/L_t\right)$$
$$- \frac{1}{\sigma_A}\left[\ln(H_{jt}/L_{jt}) - \ln(H_t/L_t)\right] + e_{jt}$$

The relative H/L wage ratio for cohort j depends on four factors

1. The skill factor-augmenting parameters A_{Ht}/A_{Lt}
2. The age-specific efficiency parameters α_j/β_j
3. The aggregate supply of H_t/L_t
4. Gap btwn the relative supply of H_{jt}/L_{jt} (in a cohort) and aggregate overall supply $\ln(H_t/L_t)$, written as $\ln(H_{jt}/L_{jt}) - \ln(H_t/L_t)$

Agenda

1. **Some motivating fact patterns**

2. **Supply shifts and cohort effects**
   - The puzzling age-profile of the college premium
   - Formalizing a model with imperfect age-group substitutability
   - Cohort effects in a model without cohort quality differences
   - Estimation of full model

3. **Has there been a decline in the quality of college graduates?**
   - Motivation
   - Estimating the Carneiro-Lee model in light of Card-Lemieux

There are no cohort quality differences in model—yet, cohort effects emerge

How could supply shifts generate 'cohorts effects'?

- Suppose that the log relative supply (*not wage ratio*) of workers of age j in year t is

$$\ln\left(H_{jt}/L_{jt}\right) = \lambda_{t-j} + \phi_j$$

(1) A cohort effect for the group λ_{t-j} ($t-j$ is cohort's birth year)

(2) An age effect $\phi_j = \ln(\alpha_j/\beta_j)$ that is common across cohorts

- λ_{t-j} is cohort specific relative supply of H versus L labor, and ϕ_j is an age effect
- Age effects ϕ_j allows relative supply term to vary with age, but age profile constant across cohorts
- Assumption: λ_{t-j} is fixed for a cohort—cohorts do not obtain (much) additional education after labor market entry

Card-Lemieux 2001: Estimation

Rewrite equation above in estimable form as

$$\ln\left(w^H_{jt}/w^L_{jt}\right) \equiv r_{jt} = \ln(A_{Ht}/A_{Lt}) - \frac{1}{\sigma_A}\phi_j + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma}\right)\ln\left(H_t/L_t\right) - \frac{1}{\sigma_A}\lambda_{t-j} + e_{jt}$$

1. Year-specific factors, common across age groups:

$$\ln(A_{Ht}/A_{Lt}) + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma}\right)\ln\left(H_t/L_t\right)$$

2. Age-group specific factors, common across years

$$-\frac{1}{\sigma_A}\phi_j = -\frac{1}{\sigma_A}\ln(\alpha_j/\beta_j)$$

3. Cohort-specific constants: $-\frac{1}{\sigma_A}\lambda_{t-j}$
4. And a residual: e_{jt}

Special cases when cohort effects will not be evident

- Recall that

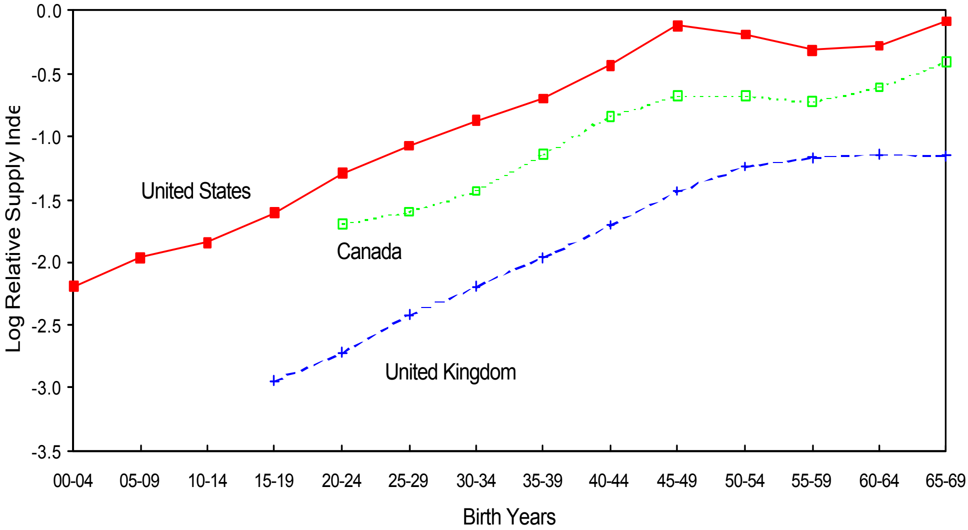
$$\ln\left(H_{jt}/L_{jt}\right) = \lambda_{t-j} + \phi_j$$

and $r_{jt} \equiv \ln\left(w^H_{jt}/w^L_{jt}\right) =$

$$= \ln(A_{Ht}/A_{Lt}) - \frac{1}{\sigma_A}\phi_j + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma}\right)\ln\left(H_t/L_t\right) - \frac{1}{\sigma_A}\lambda_{t-j} + e_{jt}$$
$$= \ln(A_{Ht}/A_{Lt}) - \frac{1}{\sigma_A}\phi_j - \frac{1}{\sigma}\ln\left(H_t/L_t\right) - \frac{1}{\sigma_A}\left[\ln(H_{jt}/L_{jt}) - \ln(H_t/L_t)\right] + e_{jt}$$

- **When will cohort effects not be evident?**
  1. When $1/\sigma_A \approx 0$, ($\sigma_A \rightarrow \infty$): Age groups perfect substitutes
  2. When $\frac{1}{\sigma_A}\left[\ln(H_{jt}/L_{jt}) - \ln(H_t/L_t)\right] \approx$ constant: cohort effects present but not identifiable
  3. **BUT**, if $\ln(H_{jt}/L_{jt}) - \ln(H_t/L_t)$ varies over time, r_{jt} will exhibit 'cohort effects'

Special case may be relevant for 1900-1945 birth cohorts



Card and Lemieux 2001, QJE

Existence of cohort effects?

C-L want to estimate

$$r_{jt} = b_j + C_{t-j} + d_t + e_{jt}.$$

What's the problem?

- *Not estimable*—cohort effects are a linear combination of the $b_{j's}$ and the $d_{t's}$
- Even if we believe that age, cohort and time effects exist, cannot identify them

Workaround

- Restrict the cohort effects to be the same for the 10 oldest cohorts
- Allows identification of the age effects along with the cohort effects for younger ages
- Identification by assumption, but it may be reasonable.

Estimating age, time and cohort effects—Restricting oldest 10 cohorts to same cohort effect: United States

	United States		
	No cohort effects	10 oldest cohorts only	10 oldest coh. eff. same
Year effects			
1970	0.026	0.020	0.020
	(0.021)	(0.010)	(0.009)
1975	−0.020	−0.024	−0.024
	(0.021)	(0.010)	(0.009)
1980	−0.049	−0.060	−0.062
	(0.019)	(0.011)	(0.009)
1985	0.058	0.017	0.015
	(0.020)	(0.013)	(0.010)
1990	0.099	0.022	0.022
	(0.020)	(0.015)	(0.011)
1995	0.141	0.024	0.034
	(0.021)	(0.019)	(0.014)
Cohort effects:			
1950–1954	—	—	0.040
			(0.011)
1955–1959	—	—	0.124
			(0.013)
1960–1964	—	—	0.178
			(0.016)
1964–1969	—	—	0.175
			(0.024)
Degrees of freedom	36	26	32
χ^2 (p-value)	295.01	48.84	51.09
	(0.00)	(0.00)	(0.02)
R^2	0.87	0.97	0.98

Card & Lemieux *QJE* 2001

Estimating age, time and cohort effects—Restricting oldest 10 cohorts to same cohort effect: US, UK, and Canada

	United States			United Kingdom			Canada		
	No cohort effects	10 oldest cohorts only	10 oldest coh. eff. same	No cohort effects	7 oldest cohorts only	7 oldest coh. eff. same	No cohort effects	6 oldest cohorts only	6 oldest coh. eff. same
Year effects									
1970	0.026	0.020	0.020	—	—	—	—	—	—
	(0.021)	(0.010)	(0.009)						
1975	−0.020	−0.024	−0.024	0.000	0.000	0.000	—	—	—
	(0.021)	(0.010)	(0.009)						
1980	−0.049	−0.060	−0.062	−0.077	−0.086	−0.076	0.000	0.000	0.000
	(0.019)	(0.011)	(0.009)	(0.026)	(0.021)	(0.018)			
1985	0.058	0.017	0.015	−0.045	−0.057	−0.069	0.020	0.007	−0.004
	(0.020)	(0.013)	(0.010)	(0.027)	(0.025)	(0.021)	(0.019)	(0.017)	(0.013)
1990	0.099	0.022	0.022	0.021	−0.041	−0.037	0.031	−0.011	−0.025
	(0.020)	(0.015)	(0.011)	(0.028)	(0.028)	(0.025)	(0.017)	(0.018)	(0.016)
1995	0.141	0.024	0.034	0.051	−0.060	−0.039	0.043	−0.038	−0.039
	(0.021)	(0.019)	(0.014)	(0.030)	(0.038)	(0.031)	(0.017)	(0.021)	(0.021)
Cohort effects:									
1950–1954	—	—	0.040	—	—	−0.009	—	—	0.028
			(0.011)			(0.019)			(0.015)
1955–1959	—	—	0.124	—	—	0.075	—	—	0.076
			(0.013)			(0.025)			(0.021)
1960–1964	—	—	0.178	—	—	0.134	—	—	0.133
			(0.016)			(0.032)			(0.027)
1964–1969	—	—	0.175	—	—	0.162	—	—	0.142
			(0.024)			(0.046)			(0.036)
Degrees of freedom	36	26	32	24	14	20	18	9	14
χ^2 (p-value)	295.01	48.84	51.09	48.31	10.76	15.33	66.06	12.00	20.51
	(0.00)	(0.00)	(0.02)	(0.00)	(0.70)	(0.76)	(0.00)	(0.21)	(0.11)
R^2	0.87	0.97	0.98	0.77	0.85	0.92	0.89	0.90	0.97

Card & Lemieux *QJE* 2001

Estimating substitutability among age groups

Taking parametric model to the data

- Estimate age-group specific elasticities of substitution with

$$\begin{aligned} r_{jt} &= b_j + d_t - (1/\sigma_A)\ln(H_{jt}/L_{jt}) + e_{jt} \\ &= b_j + d_t + \gamma \ln(H_{jt}/L_{jt}) + e_{jt} \end{aligned}$$

- Notes
  - $\widehat{b_j} = \ln(\alpha_j/\beta_j)$
  - $\widehat{d_t} = \ln(A_{Ht}/A_{Lt}) - (\frac{1}{\sigma} - \frac{1}{\sigma_A})\ln(H_t/L_t)$
  - $\hat{\gamma} = -1/\sigma_A$
  - Structure of the CES model allows us to estimate σ_A while absorbing the main effects of σ and H/L.

Estimating σ_A

	United States		United Kingdom		Canada	
	(1)	(2)	(3)	(4)	(5)	(6)
Age-group specific relative supply	−0.203	−0.265	−0.233	−0.261	−0.165	−0.161
	(0.019)	(0.026)	(0.058)	(0.071)	(0.042)	(0.040)
Trend	—	0.012	—	0.013	—	0.006
		(0.001)		(0.003)		(0.001)
Year effects:						
1970	0.104	—	—	—	—	—
	(0.012)					
1975	0.124	—	0.0	—	—	—
	(0.017)					
1980	0.129	—	−0.032	—	0.0	—
	(0.019)		(0.023)			
1985	0.255	—	0.060	—	0.029	—
	(0.020)		(0.034)		(0.014)	
1990	0.301	—	0.149	—	0.054	—
	(0.021)		(0.039)		(0.014)	
1995	0.365	—	0.199	—	0.089	—
	(0.023)		(0.044)		(0.017)	
Degrees of freedom	35	40	23	26	17	19
χ^2 (p-value)	66.62	209.34	25.78	52.03	35.00	35.68
	(0.00)	(0.00)	(0.31)	(0.00)	(0.01)	(0.01)
R^2	0.97	0.91	0.86	0.72	0.94	0.94

Standard errors are in parentheses. Models are fit by weighted least squares to the age-group by year college-high school wage gaps shown in Table I. Weights are inverse sampling variances of the estimated wage gaps. All models include age effects. For the United States and the United Kingdom, the years indicated when reporting the estimated year effects are the midpoints of the year intervals shown in Table I.

Card and Lemieux 2001, QJE

Agenda

1. **Some motivating fact patterns**

2. **Supply shifts and cohort effects**
   - The puzzling age-profile of the college premium
   - Formalizing a model with imperfect age-group substitutability
   - Cohort effects in a model without cohort quality differences
   - Estimation of full model

3. **Has there been a decline in the quality of college graduates?**
   - Motivation
   - Estimating the Carneiro-Lee model in light of Card-Lemieux

Estimating the full model: What's missing?

Now, construct a measure of aggregate and cohort specific supply

- One ingredient still missing...
- Estimates of age-specific efficiency parameters α_j, β_j

C-L estimate by fitting

$$\ln(w_{jt}^{L}) + \frac{1}{\widehat{\sigma}_A} \ln L_{jt} = \ln(A_{Lt} L_t^{\rho-\eta} \pi_t) + \ln \beta_j + e_j$$

$$\ln(w_{jt}^{H}) + \frac{1}{\widehat{\sigma}_A} \ln H_{jt} = \ln(A_{Ht} H_t^{\rho-\eta} \pi_t) + \ln \alpha_j + e'_j$$

Estimated for each skill group, pooling across all time periods t

- $\ln(A_{Ht} H_t^{\rho-\eta} \pi_t)$ is absorbed by a set of year dummies
- Efficiency parameters α_j and β_j are estimated with age dummies

Second stage estimation

We've estimated σ_A and the α_j and β_j. But we don't yet know

1. σ, the overall elasticity of substitution between H and L
2. Evolution of skill demands $\Delta \ln A_{Ht}/A_{Lt}$
3. To what degree 'cohort effects' are explained by acceleration/deceleration of supply

Second stage estimation

Given estimates of $\alpha_j, \beta_j, \sigma_A$ we can do the grand estimation

$$r_{jt} = \ln(A_{Ht}/A_{Lt}) + \ln(\alpha_j/\beta_j) - \frac{1}{\sigma}\ln\left(H_t/L_t\right) - \frac{1}{\sigma_A}\left[\ln(H_{jt}/L_{jt}) - \ln(H_t/L_t)\right] + e_{jt}$$

Notice: C-L include two supply measures as regressors

1. Aggregate supply measure $\ln\left(H_t/L_t\right)$. Coefficient provides an estimate of $1/\sigma$
2. Deviation of the cohort supply measure from the aggregate measure, $\ln(H_{jt}/L_{jt}) - \ln\left(H_t/L_t\right)$. Coefficient provides a *second* estimate of $1/\sigma_A$
3. C-L also estimate 2nd stage models where constructed supply index, $\ln(\hat{H}_t/\hat{L}_t)$, incorporates estimates of σ_A and $\alpha_j's, \beta_j's$

C-L main estimates for US

	United States		
	(1)	(2)	(3)
Age-group specific relative supply	−0.202 (0.026)	−0.209 (0.025)	−0.208 (0.019)
Trend	0.017 (0.002)	0.020 (0.002)	0.015 (0.002)
1980 dummy	—	—	−0.057 (0.011)
Katz-Murphy aggr. supply index	−0.414 (0.047)	—	—
Aggr. supply index for men with imperfect substitution across age groups	—	−0.483 (0.053)	−0.327 (0.051)
Degrees of freedom	39	39	38
χ^2 (p-value)	143.05 (0.00)	138.02 (0.00)	81.01 (0.00)
R^2	0.94	0.95	0.96

Card and Lemieux 2001, QJE

C-L main estimates for UK and Canada

	United Kingdom			Canada		
	(4)	(5)	(6)	(7)	(8)	(9)
Age-group specific relative supply	−0.233	−0.233	−0.233	−0.166	−0.165	−0.165
	(0.078)	(0.078)	(0.059)	(0.041)	(0.041)	(0.042)
Trend	0.021	0.018	0.020	−0.001	−0.002	−0.006
	(0.007)	(0.006)	(0.005)	(0.007)	(0.015)	(0.024)
1980 dummy	—	—	−0.073	—	—	−0.006
			(0.016)			(0.026)
Katz-Murphy aggr. supply index	−0.466	—	—	0.069	—	—
	(0.156)			(0.247)		
Aggr. supply index for men with imperfect substitution across age groups	—	−0.340	−0.416	—	0.134	0.275
		(0.114)	(0.087)		(0.547)	(0.826)
Degrees of freedom	25	25	24	18	18	17
χ^2 (p-value)	50.39	50.47	27.34	35.08	35.12	35.00
	(0.00)	(0.00)	(0.29)	(0.01)	(0.01)	(0.01)
R^2	0.73	0.73	0.86	0.94	0.94	0.94

Robustness: Def'n of college workers, hours and hourly earnings, using experience rather than age cohorts

TABLE V
ROBUSTNESS OF THE RESULTS TO ALTERNATIVE MEASURES OF THE COLLEGE-HIGH SCHOOL WAGE GAP, UNITED STATES

	Wage gap by age group				
	1959–1995		1975–1995		Wage gap by Experience Group, 1959–1995
	Coll + vs. HS	Coll "labor" vs. HS "labor"	AWE of FT wkrs	Av. hourly earnings	
	(1)	(2)	(3)	(4)	(5)
Age-group rel. supply	−0.157	−0.125	−0.237	−0.218	−0.107
	(0.023)	(0.020)	(0.033)	(0.035)	(0.048)
Aggregate supply index	−0.562	−0.426	−0.355	−0.400	−0.618
	(0.056)	(0.042)	(0.135)	(0.146)	(0.103)
Trend	0.026	0.020	0.017	0.017	0.024
	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Degrees of freedom	39	39	25	25	39
R^2	0.96	0.95	0.93	0.91	0.73

Card and Lemieux 2001, QJE

Including women and men: makes a big difference

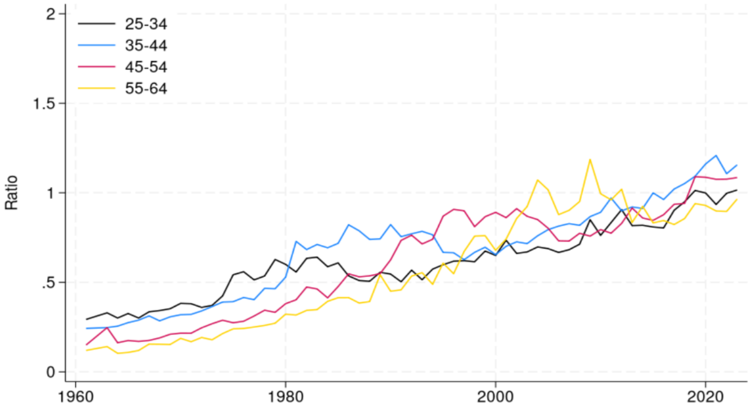
	(1)	(2)	(3)	(4)	(5)
Age-group specific relative supply	—	—	−0.221	−0.230	−0.223
			(0.020)	(0.031)	(0.022)
Aggregate supply index (men and women)	—	—	—	−0.865	−0.628
				(0.091)	(0.074)
Time trend	—	—	—	0.035	0.027
				(0.003)	0.003
Year effects:					
1970	0.037	0.033	0.034	—	—
	(0.019)	(0.009)	(0.009)		
1975	−0.009	−0.010	−0.001	—	—
	(0.019)	(0.008)	(0.009)		
1980	−0.035	−0.045	−0.028	—	−0.057
	(0.017)	(0.008)	(0.008)		(0.009)
1985	0.061	0.025	0.058	—	—
	(0.017)	(0.009)	(0.008)		
1990	0.124	0.058	0.112	—	—
	(0.017)	(0.009)	(0.008)		
1995	0.174	0.087	0.161	—	—
	(0.018)	(0.011)	(0.009)		
Cohort effects:					
1950–1954	—	0.033	—	—	—
		(0.009)			
1955–1959	—	0.106	—	—	—
		(0.011)			
1960–1964	—	0.145	—	—	—
		(0.013)			
1965–1969	—	0.133	—	—	—
		(0.019)			
Degrees of freedom	36	32	35	39	38
χ^2 (p-value)	331.80	56.31	73.91	194.42	93.46
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
R^2	0.89	0.98	0.98	0.94	0.97

Card and Lemieux 2001, QJE

Unknown: How well this exercise would work with 25 more years of data (I)

Figure 7: Relative supply of college educated to high school educated labor, by gender and broad age group

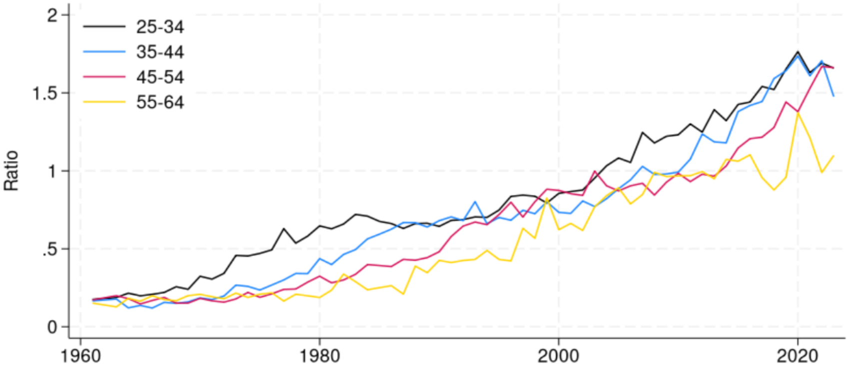
Panel A: Men



Bengali, Valletta, Zhao 2025

Unknown: How well this exercise would work with 25 more years of data (II)

Panel B: Women



Note: Years are survey reference years. Relative supply is the ratio of the college educated labor supply to the high school educated labor supply, as defined in the text; however, relative to the text description, the series in this figure are calculated for broader, 10-year age groups. The sample used to compute supply includes all classes of part- and full-time workers of all education levels, aged 25–64. Labor supplied by individuals with less than a high school degree, some college, or above a college degree is allocated to 'college educated labor' and 'high school educated labor' as described in the text.
Source: Authors' calculations from CPS ASEC microdata.

Bengali, Valletta, Zhao 2025

Conclusions

1. **Substance**
   - Appears to be something fundamentally relevant (or at least useful) about this model
   - C&L get an amazing amount of mileage out of simple two-factor skill premium
     - Over time
     - Across countries
     - Between skill groups
     - Among age-groups within a skill group
2. **Craft and style**
   - Well chosen problem and model
   - Connects to frontier of literature (at that time)
   - Extends, reinterprets, or (in some cases) overturns existing knowledge

From quantity to quality

Card-Lemieux explain cohort effects via quantity of skilled labor

- Cohorts with more college graduates have lower college wage premiums
- Driven by imperfect substitutability across age groups
- No role for differences in *quality* across cohorts

But could changes in cohort quality also matter?

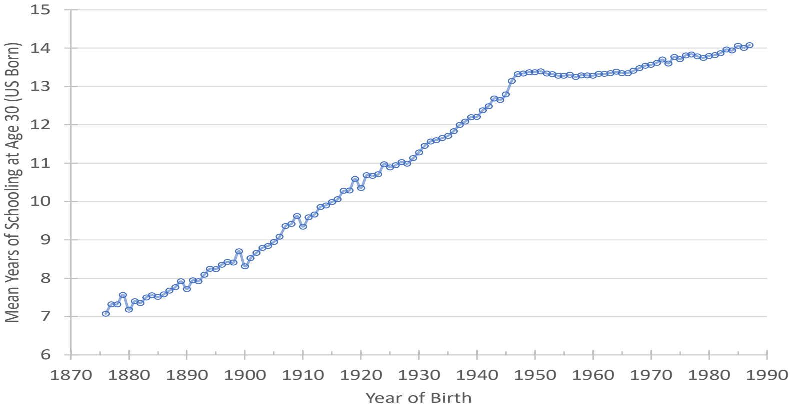
- As college enrollment expands, are ‘marginal’ students lower quality?
- Does education quality decline when classrooms are more crowded?
- If so, how does this affect measured wage premiums?

Challenge: Quality and quantity stories make similar predictions—both suggest cohort wage premiums fall when college enrollment rises

Agenda

Agenda

So much more schooling – Has quality been diluted?



Sources and Notes: US Census IPUMS data from 1940 to 2000 and CPS MORG data from 2005 to 2018. The figure updates Goldin and Katz (2007, figure 7). See the on-line appendix for details.

Carneiro-Lee 2011: Want to estimate 'quality-adjusted' changes in college/HS wage premium, 1960-2000

Secular Δ's in college-going create shifts in cohort supply

- Might also create differences in the cohort *quality* of education
  - 'Lower-quality' individuals go on to college
  - Quality of education deteriorates when there is a large influx of students (or both)
- How this might affect the wage structure—specifically, the measured college/high-school gap?
  - Clear for *wage level*
  - Non-obvious for *wage gap*

Cross-state scatters: SAT scores vs. SAT test taking rates

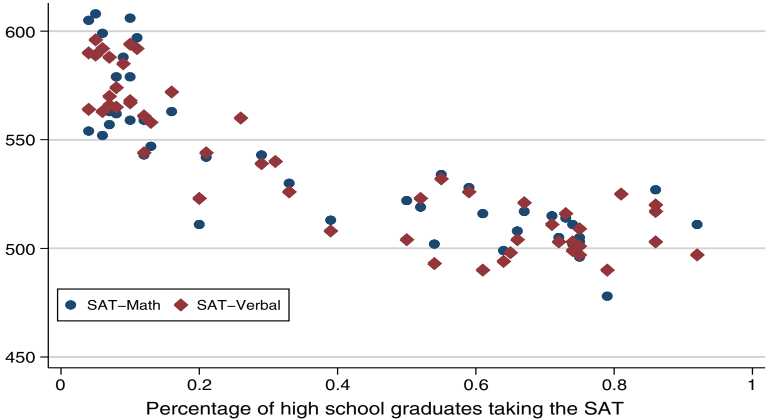


FIGURE 3. SAT SCORES AND PARTICIPATION RATES

Note: The figure displays average verbal and math SAT scores by state (in 2004) against the percentage of high school graduates who take the SAT in each state.

Carneiro and Lee, 1011

Cross-state regressions: SAT scores vs. SAT test taking rates

Table 7: Regression of SAT Scores on the Percentage of High School Graduates Taking SAT, with Year and State Dummies - 1993/2004 (except 1995 and 1998)

	(1)	(2)
	SAT Math	SAT Verbal
Percentage of High School Graduates Taking SAT	-16.503 [7.275]**	-26.268 [6.629]***
Observations	510	510
R-squared	0.99	0.99
Included Dummy Variables		
(Year)	√	√
(State)	√	√

Notes: Robust standard errors in brackets, clustered on the region of residence-schooling-year cell. * significant at 10%; ** significant at 5%; *** significant at 1%.

Carneiro and Lee, 2011

Evidence from IALS: cohorts with more college grads have lower quantitative literacy among college grads

Variable	(1) College	(2) High School	(3) All
Proportion in College	-1.248	0.604	-0.112
	[0.581]**	[0.533]	[0.384]
College			0.977
			[0.047]***
Age	0.101	0.001	0.037
	[0.024]***	[0.019]	[0.013]***
Age Squared	-0.001	0.000	0.000
	[0.000]***	[0.000]	[0.000]***
Constant	-0.794	-0.651	-1.050
	[0.416]*	[0.444]	[0.295]***
Observations	962	1503	2465
R-squared	0.03	0.00	0.22

Notes: The dependent variable is the standardized quantitative literacy score of individuals aged 25-60 in the US sample of the International Adult Literacy Survey. Proportion in college is the percentage of individuals with some college or more and is computed for each cohort. The variable "College" is a dummy variable for individuals with some college or more. Robust standard errors in brackets, clustered at the cohort level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Carneiro and Lee, 2011

Agenda

Carneiro-Lee 2011: Want to estimate ‘quality-adjusted’ changes in college/HS wage premium, 1960-2000

Challenge for Carneiro-Lee: Their hypothesis is almost isomorphic to Card-Lemieux ’01

- Both predict that education 'returns' rise differentially for young college grads when the supply of young college grads decelerates
- In the Card-Lemieux view, it stems from a slowdown in the supply of young college adults
- In the Carneiro-Lee view, it could alternatively stem from an *improvement* in the quality of college relative to non-college adults

Carneiro-Lee 2011: Estimation approach

Carneiro-Lee draw on technique from Card and Krueger 1992, JPE

- To identify the effect of cohort quality on earnings...
- Regress wages of workers *working outside of their home region* (9 Census regions) on educational composition of birth cohort
- Assumes cross-region moves are exogenous to wages—though there are fixes

Carneiro-Lee 2011: Estimation approach

Write the wage of college worker i **as** $\ln w^c_{iatrb}$:

$$\ln w^c_{iatrb} = \gamma_{atr} + \gamma_{ab} + \gamma_{tb} + \phi\left(\tilde{P}_{t-a,b}\right) + e^c_{iatrb}.$$

- a is current age
- t is year
- r is the region of work
- b is the region of birth
- superscript c indicates that i is a college worker

Coefficient of interest is ϕ

- Odds of proportion of cohort that attended college

$$\tilde{P}_{t-a,b} = P_{t-a,b} / \left(1 - P_{t-a,b}\right)$$

Carneiro-Lee 2011: Estimation approach

Estimating equation for log college wage

$$\ln w^{c}_{iatrb} = \gamma_{atr} + \gamma_{ab} + \gamma_{tb} + \phi\left(\tilde{P}_{t-a,b}\right) + e^{c}_{iatrb}$$

- γ_{atr} is a full set of interactions between age, year and region of work. These absorb average wage levels of all college workers by age in year t in each region
- γ_{ab} takes out average wage levels of workers by each region-of-birth by age group
- γ_{tb} takes out average wages of workers by each region of birth by year

Carneiro-Lee 2011: Estimation approach

Available variation

- Wages vary across schooling × year × age × residence-region × birth-region cells
- Weeks worked (for labor supply models) vary across schooling × year × age × residence-region cells
- Composition varies across schooling × year × age × birth-region

$$\ln w^c_{iatrb} = \gamma_{atr} + \gamma_{ab} + \gamma_{tb} + \phi\left(\tilde{P}_{t-a,b}\right) + e^c_{iatrb}$$

Identifying variation (what's leftover)

- Cohort by birth-region variation in wage levels, ID'd by individuals born in regions b but working in regions $r \neq b$
- Paper also attempts to address the selective migration issue with parametric corrections, none perfect

College premium in region of residence (by cohort) vs. college proportion in region of birth (by cohort)

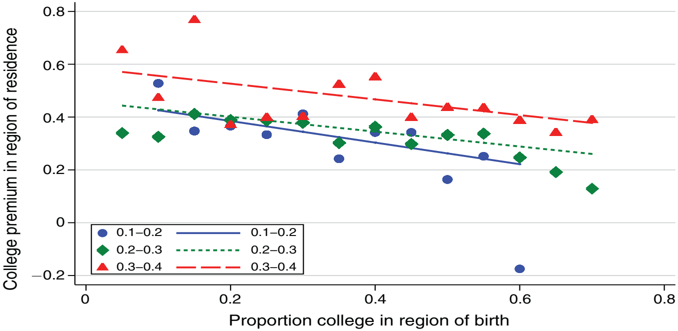


Figure 1—College Premium in Region of Residence as a Function of Proportion College in Region of Birth

Notes: Using US Census data between 1960 and 2000, the figure graphs the college premium in each (year) × (age) × (region of birth) × (region of residence) cell against the proportion in college by cohort and region of birth, after grouping individuals into three sets of regional labor markets: those with a high share of college educated workers (30–40 percent), those with a medium share (20–30 percent), and those with a low share (10–20 percent).

Carneiro and Lee, 2011

Carneiro-Lee reduced form and structural estimates for college earnings as a function of supply and 'quality'

	(1) Reduced-Form Model (Controlling for Quality)	(2) Structural Model (Controlling for Quality)	(3) Structural Model (Without Controlling for Quality)
Panel A - College			
Odds of Proportion in College	-0.086	-0.092	
	[0.036]**	[0.032]***	
Quality-Adjusted Log Weeks		-0.110	
		[0.022]***	
Log Weeks			-0.188
			[0.019]***
Observations	2598	2598	2598

Included Explanatory Variables			
(Year)×(Age)×(Region of Residence)	√		
(Year)×(Region of Residence)×(Region of Birth)	√	√	√
(Age)×(Common Linear Time Trends)		√	√
(Year)×(Region of Birth)	√	√	√
(Age)×(Region of Birth)	√	√	√
(Observed Migration Probability)	√	√	√
(Observed Migration Probability)^2	√	√	√
(Staying Probability)	√	√	√
(Staying Probability)^2	√	√	√
(Observed Migr. Prob.)×(Staying Prob.)	√	√	√

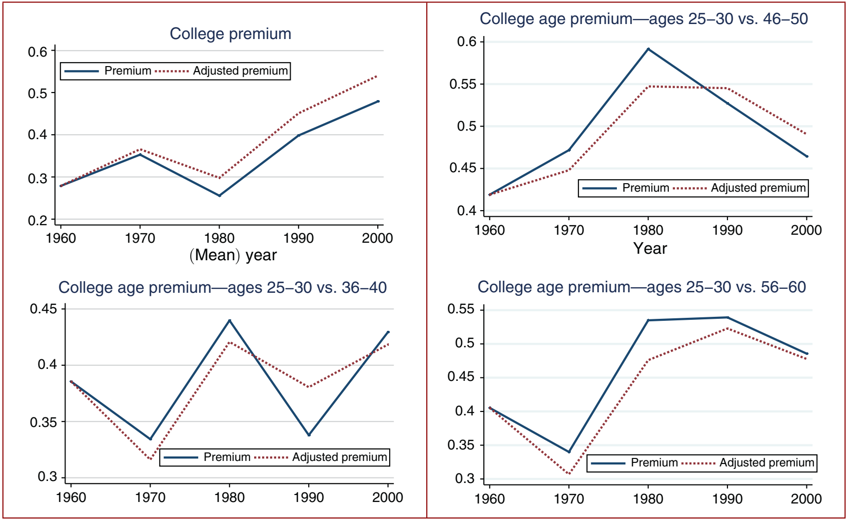
Carneiro and Lee, 2011

Carneiro-Lee reduced form and structural estimates for high-school earnings as a function of supply and 'quality'

	(1) Reduced-Form Model (Controlling for Quality)	(2) Structural Model (Controlling for Quality)	(3) Structural Model (Without Controlling for Quality)
Panel B - High School			
Odds of Proportion in College	-0.032	-0.023	
	[0.022]	[0.024]	
Quality-Adjusted Log Weeks		-0.091	
		[0.012]***	
Log Weeks			-0.090
			[0.012]***
Observations	2692	2692	2692
Included Explanatory Variables			
(Year)×(Age)×(Region of Residence)	√		
(Year)×(Region of Residence)×(Region of Birth)	√	√	√
(Age)×(Common Linear Time Trends)		√	√
(Year)×(Region of Birth)	√	√	√
(Age)×(Region of Birth)	√	√	√
(Observed Migration Probability)	√	√	√
(Observed Migration Probability)^2	√	√	√
(Staying Probability)	√	√	√
(Staying Probability)^2	√	√	√
(Observed Migr. Prob.)×(Staying Prob.)	√	√	√

Carneiro and Lee, 2011

‘Quality-adjusted’ college wage premiums (Carneiro-Lee ’11)



Carneiro-Lee 2011: Interpretation

Interpreting the coefficient of -0.086 (Panel A of Table 1) on $\tilde{P}$

- If college enrollment rises from 50% to 60%, $\tilde{P}$ increases from 1 to 1.5
  - Implies a fall in college wages of 0.0425

If wage declines due to fall in quality of marginal college-goers

- Let w_1 be the average log wage of inframarginal students
- If marginal students are $17\% = 10/60$ of college goers, then the observed college wage is:

$$\Delta \bar{w} = -(0.086 \times 0.5) = 0.83 w_1 + 0.17 w_2 - w_1$$
$$-(0.086 \times 0.5) = 0.17(w_2 - w_1)$$
$$w_1 - w_2 = 0.253$$

- Implies that the marginal college-goers are 25% less skilled than the inframarginal college-goers