Wage Dynamics and Returns to Unobserved Skill

Lance Lochner
University of Western Ontario

Youngmin Park
Bank of Canada

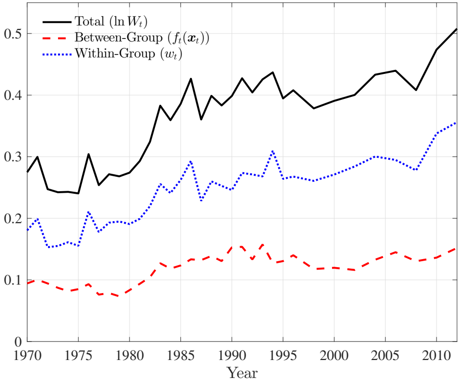
Youngki Shin
McMaster University

Ninth BIS CCA Research Conference

June 5, 2018

The views expressed in this presentation are the authors' and do not necessarily represent those of the Bank of Canada

Rising Wage Inequality in the US



- Substantial increase within (education, race, and experience) groups $\Rightarrow$ rising 'residual' wage inequality
- Source: variance of log hourly wages for American men from the Panel Study of Income Dynamics (PSID)

Rising Returns to Unobserved Skill?

- Since Juhn, Murphy & Pierce (1993), many economists have equated rising residual inequality with an increase in the returns to unobserved ability/skill:
  - Let $w_{i,t}$ be the log wage residual of individual i in year t:

$$w_{i,t} = \underbrace{\mu_t}_{\text{return}} \times \underbrace{\theta_i}_{\text{unobserved skill}}$$

- Motivated an influential literature on skill-biased technical change
  - Many studies specifically aimed to explain rising returns to unobserved skill/ability (e.g., Acemoglu 1998, Galor & Moav 2000)
- Still the dominant interpretation within labor economics (e.g., Autor, Katz & Kearney 2008)

Other Interpretations

- Lemieux (2006) argues that some of the increase in residual inequality is explained by an increase in the variance of unobserved skills
  - $\text{Var}(\theta_t)$ increased over time due to an increase in highly educated and older workers who have larger within-group skill dispersion

$$\underbrace{\text{Var}(\theta_t)}_{\text{total var}} = \sum_j \underbrace{p_{j,t}}_{\text{fraction of } j} \underbrace{\text{Var}(\theta_j)}_{\text{var of } j}$$

- Short-term fluctuations in earnings have also increased (Gottschalk & Moffitt 1994)
  - Reflects measurement errors or transitory shocks unrelated to skills:

$$w_{i,t} = \underbrace{\theta_i}_{\text{permanent}} + \underbrace{\varepsilon_{i,t}}_{\text{transitory}}$$

  - Evidence based on panel data (PSID) that $\text{Var}(\varepsilon_t)$ increased over time
  - Motivated a literature on trends in consumption inequality (e.g., Krueger & Perri 2006, Blundell, Pistaferri & Preston 2008)

Identifying the Causes of Rising Residual Inequality

- We show how panel data on wages can be used to separately identify the evolution of:
  - Returns to unobserved skills
  - Distributions of unobserved skills
  - Volatility of transitory wage shocks
- Key idea: 'long' autocovariances of wages identify the evolution of skill returns
  - Transitory shocks do not drive long-term differences in wages across individuals
  - Motives a new Instrumental Variable (IV) estimation strategy for estimating skill returns
    - Requires no assumptions on experience or time effects on variances of skills or shocks

Main Empirical Findings (from PSID)

- Returns to unobserved skill *fell* substantially from mid-1980s through mid-1990s
  - Decline was more dramatic for those who didn't attend college
  - Consistent with falling return to measured cognitive ability (AFQT score) since 1980s (Castex & Dechter 2014)
  - Fundamentally different from returns estimated under assumption of time-invariant skill distributions (e.g., Juhn, Murphy, & Pierce 1993, Moffitt & Gottschalk 2012)
- Rising residual inequality is driven by an increase in the variance of unobserved skill
  - Variance of lifecycle skill growth, not variance of initial skills
  - Not accounted for in composition effects of Lemieux (2006)

Explaining the Falling Returns

- We develop a new quantitative framework to decompose the changes in returns to unobserved skill into changes in demand and supply
  - Based on a job assignment model of Sattinger (1979)
  - Our estimates can be combined with the restrictions implied by the model to recover changes in skill demand and supply
- Decomposition suggests that both supply and demand factors played important roles
  - Decline in skill demand was more important than supply shifts for non-college workers

Modeling Log Wage Residuals

We observe log wage residuals of a large number of individuals for periods $t = \underline{t}, \underline{t}+1, \ldots, \overline{t}$

$$\begin{aligned} w_{i,t} &= \mu_t \theta_{i,t} + \varepsilon_{i,t} \\ \theta_{i,t} &= \theta_{i,t-1} + \nu_{i,t} \end{aligned}$$

- $\theta_{i,t}$ reflects time-varying unobserved skills
- μ_t reflects the labor market returns to unobserved skills
  - Shifts relative wages (log wage differential) between high and low skill
- $\varepsilon_{i,t}$ reflects transitory components of wage
  - Year to year wage fluctuations unrelated to skills (e.g., measurement error, wage dynamics induced by labor market frictions)

Normalizations & Assumptions

$$\begin{aligned} w_{i,t} &= \mu_t \theta_{i,t} + \varepsilon_{i,t} \\ \theta_{i,t} &= \theta_{i,t-1} + \nu_{i,t} \end{aligned}$$

- All components are mean zero:
  $\mathrm{E}[\theta_t] = \mathrm{E}[\varepsilon_t] = \mathrm{E}[\nu_t] = 0$
- Transitory components are uncorrelated with skills:
  $\mathrm{Cov}(\varepsilon_t, \theta_{t'}) = \mathrm{Cov}(\varepsilon_t, \nu_{t'}) = 0, \ \forall (t, t')$
- Serial correlation in transitory component dies out after k periods:
  $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t'}) = 0$ for $|t' - t| \geq k$
- Skill changes are uncorrelated with past skills:
  $\mathrm{Cov}(\nu_t, \theta_{t'}) = 0$ for $t' < t$

Identifying Skill Returns Over Time

Substituting in for $\theta_{i,t-1} = (w_{i,t-1} - \varepsilon_{i,t-1})/\mu_{t-1}$ yields

$$\begin{aligned} w_{i,t} =& \mu_t(\theta_{i,t-1} + \nu_{i,t}) + \varepsilon_{i,t} \\ =& \mu_t \left(\frac{w_{i,t-1} - \varepsilon_{i,t-1}}{\mu_{t-1}} + \nu_{i,t} \right) + \varepsilon_{i,t} \\ =& \frac{\mu_t}{\mu_{t-1}} w_{i,t-1} + \left(\varepsilon_{i,t} - \frac{\mu_t}{\mu_{t-1}} \varepsilon_{i,t-1} + \nu_{i,t} \right) \end{aligned}$$

- OLS is inconsistent because $w_{i,t-1}$ is correlated with $\varepsilon_{i,t-1}$
- *Past* residuals are valid instruments for $w_{i,t-1}$ since $\nu_{i,t}$ is uncorrelated with past skills and serial correlation in $\varepsilon_{i,t}$ dies out after k periods
- Probability limit of the IV estimator is

$$\frac{\text{Cov}(w_t, w_{t'})}{\text{Cov}(w_{t-1}, w_{t'})} = \frac{\mu_t \mu_{t'} \text{Var}(\theta_{t'})}{\mu_{t-1} \mu_{t'} \text{Var}(\theta_{t'})} = \frac{\mu_t}{\mu_{t-1}}, \quad \text{for } t' < t - k$$

- Therefore, μ_t / μ_{t-1} is identified for all $t > \underline{t} + k$ ▸ Identifying Early Returns
  - Normalizing $\mu_{t^*} = 1$ for some t^* sets the units for unobserved skill

Identifying the Rest

Once skill returns have been identified, the rest is identified except for the last k periods

- We can identify $\text{Var}(\theta_t|c)$ (and then $\text{Var}(\nu_t|c)$) from

$$\text{Var}(\theta_t|c) = \frac{\text{Cov}(w_t, w_{t'}|c)}{\mu_t \mu_{t'}} \quad \text{for } t' \geq t + k$$

- Next, we can identify

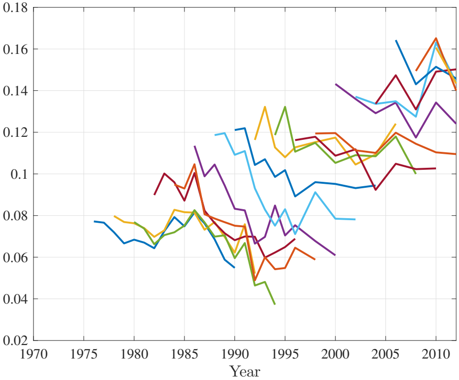
$$\text{Cov}(\varepsilon_t, \varepsilon_{t'}|c) = \text{Cov}(w_t, w_{t'}|c) - \mu_t \mu_{t'} \text{Var}(\theta_t|c) \quad \text{for } t \leq t' \leq t + k$$

PSID Data

- PSID is a longitudinal survey of a representative sample of US individuals and their families
- We use earnings for calendar years 1970-2012
  - Annual up to 1996, biennial thereafter
- Average hourly wages: annual earnings divided by annual hours worked
- Annual earnings: household head's total wages and salaries (excluding farm and business income)
- Select male household heads with ages 16-64 and experiences 1-40
  - Resulting dataset has 3,766 men and 44,547 observations
- To obtain residuals, we run cross-sectional OLS regressions of log wages on regressors including experience, race, and education, separately by year and college attendance status

Long Autocovariances

- $\text{Cov}(w_b, w_t) = \mu_b \mu_t \text{Var}(\theta_b)$ for $t - b \geq k = 6$ plotted



- Negative slopes in late 1980s and 1990s suggest *declining* μ_t
- Upward shifting lines suggest *increasing* $\text{Var}(\theta_b)$ ▸ More

IV Estimation of Changes in Skill Returns

- Recall: for $t' < t - k$,

$$\frac{\text{Cov}(w_t, w_{t'})}{\text{Cov}(w_{t-1}, w_{t'})} = \frac{\mu_t}{\mu_{t-1}}$$

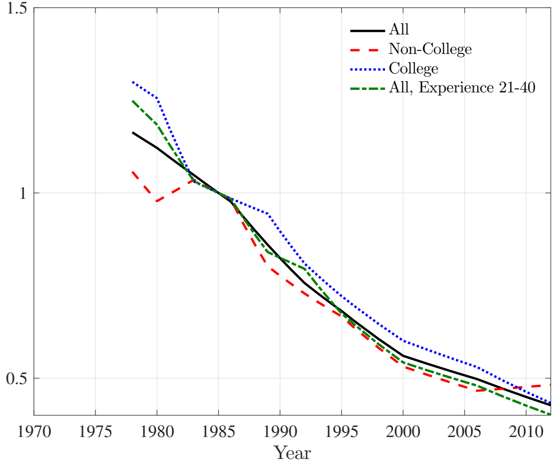
- Assuming $k = 6$, we estimate $(\mu_t - \mu_{t-2})/\mu_{t-2}$ by regressing $w_{i,t} - w_{i,t-2}$ on $w_{i,t-2}$ using the following instruments:
  - $w_{i,t-8}$ and $w_{i,t-9}$ for 1979-1995
  - $w_{i,t-8}$ and $w_{i,t-10}$ for 1996-2012

2SLS Estimates of 2-year Growth Rates of μ_t

	1979-80	1981-83	1984-86	1987-89	1990-92	1993-95	1996-2000	2002-06	2008-12
					A. All men				
	-0.036	-0.044	-0.046	-0.081*	-0.082*	-0.067	-0.075*	-0.039	-0.050
	(0.045)	(0.038)	(0.038)	(0.034)	(0.035)	(0.035)	(0.025)	(0.028)	(0.027)
Obs	1,349	2,077	2,188	2,245	2,189	2,095	2,122	2,129	1,968
F-Stat	163.09	191.61	114.85	209.42	227.13	286.96	369.09	344.25	341.36
					B. All men with 21–40 years of experience (at year t)				
	-0.052	-0.088*	-0.031	-0.100*	-0.036	-0.104*	-0.084*	-0.040	-0.058
	(0.050)	(0.043)	(0.050)	(0.046)	(0.044)	(0.045)	(0.030)	(0.032)	(0.031)
Obs	928	1,323	1,244	1,211	1,244	1,300	1,427	1,591	1,493
F-Stat	117.23	132.19	66.26	130.53	132.83	201.62	295.75	281.91	267.83
					C. Non-college-educated men (all experience levels)				
	-0.075	0.039	-0.035	-0.127*	-0.062	-0.057	-0.087*	-0.043	0.011
	(0.061)	(0.056)	(0.060)	(0.050)	(0.058)	(0.054)	(0.043)	(0.047)	(0.075)
Obs	740	1,080	997	965	897	851	862	826	615
F-Stat	81.85	85.23	39.48	98.34	92.27	91.33	121.44	142.56	104.92
					D. College-educated men (all experience levels)				
	-0.034	-0.123*	-0.030	-0.028	-0.097*	-0.074	-0.070*	-0.041	-0.065*
	(0.061)	(0.048)	(0.049)	(0.047)	(0.047)	(0.046)	(0.031)	(0.034)	(0.029)
Obs	508	884	1,046	1,109	1,107	1,242	1,252	1,293	1,141
F-Stat	100.95	115.03	123.38	97.29	122.42	208.04	260.47	218.64	229.40

Standard errors in parentheses. * Denotes significance at 0.05 level.

μ_t Implied by 2SLS Estimates ($\mu_{1985} = 1$)



- Return to skill declined around 50% from 1985 to 2005

Minimum Distance (MD) Estimation

- We now use all autocovariances of $w_{i,t}$ to estimate all parameters
  - Total 2,824 autocovariances
- MD estimator $\hat{\mathbf{\Lambda}}$ minimizes the distance between the data and model covariances

$$\min_{\mathbf{\Lambda}} \sum_{s,c,t' \leq t} \left\{ \widehat{\text{Cov}}(w_t, w_{t'}|s,c) - \text{Cov}(w_t, w_{t'}|s,c,\mathbf{\Lambda}) \right\}^2$$

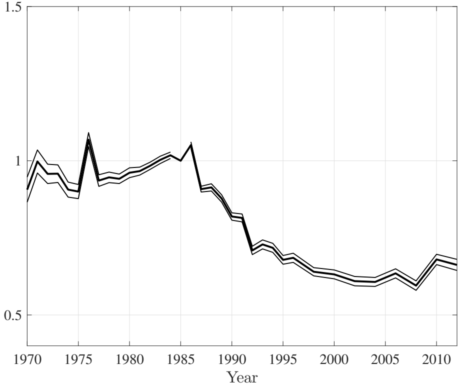
where s is the indicator for college attendance

- For a given $\mathbf{\Lambda}$, we construct model covariances by assuming:

$$\theta_{i,t} = \psi_i + \sum_{j=0}^{t-c_i+1} \nu_{i,t-j}, \quad \varepsilon_{i,t} = \sum_{j=0}^{\min\{5,t-c_i+1\}} \beta_j \xi_{i,t-j}$$

  - We allow for cohort-specific variances $\text{Var}(\psi|c), \text{Var}(\nu_t|c), \text{Var}(\xi_t|c)$

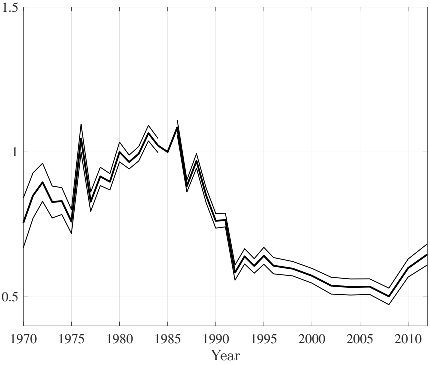
MD Estimates of μ_t: Full Sample ($\mu_{1985} = 1$)



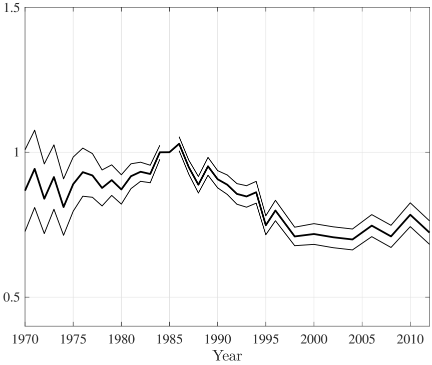
- Substantial decline from 1985 to 2005 consistent with IV estimates

Importance of Time-varying Skill Variance
Accounting for Heterogeneity in Skill Growth

μ_t Estimated Separately by College Attendance



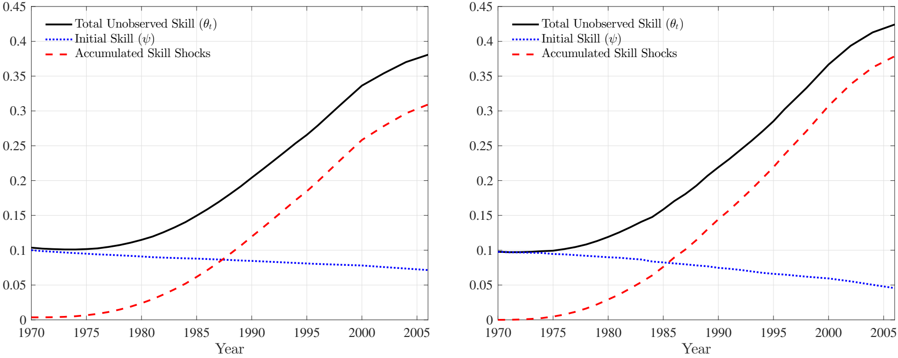
(a) Non-College



(b) College

- Reject the hypothesis of identical returns for two groups
- Stronger decline for non-college workers

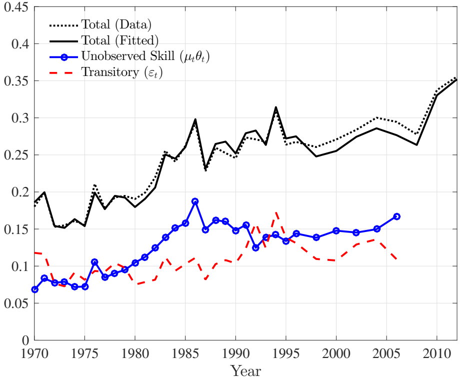
Estimated Variance of Unobserved Skill



(a) Non-College (b) College

- Rising skill variance driven by increasing variance in accumulated skill shocks rather than initial skill

Decomposition of Residual Variance



- Long-run trend in residual variance largely driven by unobserved skill

By College Attendance
Other estimates

Interpreting the Falling Returns

- Why have the returns to skill fallen?
  - Changes in demand or supply?
- We offer a new quantitative framework based on the job assignment model of Sattinger (1979) and Gabaix & Landier (2008)
  - Abstract from transitory wages (i.e., treat as measurement errors)
  - Model gives log wage equation consistent with empirical model
  - Equilibrium conditions can be combined with our estimates to recover changes in demand and supply

Job Assignment Model

- Workers differ by skill $\Theta_t = g_t(\boldsymbol{x}_t) + \theta_t$, normally distributed with mean $\mathrm{E}[\Theta_t]$ and standard deviation $\sigma(\Theta_t)$
- Jobs differ by productivity Z_t, normally distributed with mean $\mathrm{E}[Z_t]$ and standard deviation $\sigma(Z_t)$
- Output is produced through one-to-one matching of workers and jobs

$$\ln Y_t(\Theta_t, Z_t) = \lambda_t \Theta_t + \gamma_t Z_t$$

- Competitive labor market with hedonic wage function $W_t(\Theta_t)$
- Taking $W_t(\Theta_t)$ as given, employer with productivity Z_t solves

$$\max_{\Theta_t} \left\{ Y_t(\Theta_t, Z_t) - W_t(\Theta_t) \right\}$$

  - Let the 'matching function', $\hat{Z}_t(\Theta_t)$, be the inverse of the solution
- Labor market clearing condition:

$$\hat{Z}_t(\Theta_t) = \mathrm{E}[Z_t] + \frac{\sigma(Z_t)}{\sigma(\Theta_t)}(\Theta_t - \mathrm{E}[\Theta_t]),$$

Closed Form Formula for the Return to Skill

Recall: $\ln Y_t(\Theta_t, Z_t) = \lambda_t \Theta_t + \gamma_t Z_t$

$$\frac{d \ln W_t(\Theta_t)}{d\Theta_t} = \underbrace{\lambda_t}_{\text{direct}} + \underbrace{\gamma_t \frac{d\hat{Z}_t(\Theta_t)}{d\Theta_t}}_{\text{sorting}} = \lambda_t + \gamma_t \frac{\sigma(Z_t)}{\sigma(\Theta_t)}$$

- More skilled workers receive higher wages because
  - They produce more at any given job (λ_t)
  - They also work at more productive jobs ($\gamma_t \sigma(Z_t)/\sigma(\Theta_t)$)
- Sorting effect depends on the slope of the matching function $\sigma(Z_t)/\sigma(\Theta_t)$
  - Small if everyone works at the same job ($\sigma(Z_t) \approx 0$)
  - Large if everyone has the same skill ($\sigma(\Theta_t) \approx 0$)

Recovering Demand and Supply Factors

- Equating the estimated return with its theoretical counterpart:

$$\mu_t = \lambda_t + \gamma_t \frac{\sigma(Z_t)}{\sigma(\Theta_t)}$$

- We also derive the formula for the labor share:

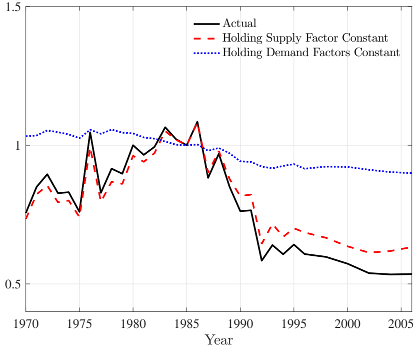
$$\frac{W_t(\Theta_t)}{Y_t(\Theta_t, \hat{Z}_t(\Theta_t))} = \frac{\lambda_t}{\lambda_t + \gamma_t \frac{\sigma(Z_t)}{\sigma(\Theta_t)}} \Rightarrow \lambda_t = \mu_t \times \text{labor share}$$

- Variance of skill:

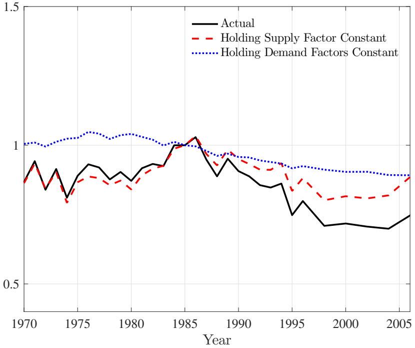
$$\sigma(\Theta_t)^2 = \frac{\text{Var}(\ln W_t) - \text{Var}(\varepsilon_t)}{\mu_t^2}$$

- Finally, $\gamma_t \sigma(Z_t) = (\mu_t - \lambda_t)\sigma(\Theta_t)$

Effects of Demand and Supply Factors on Skill Returns



(a) Non-College



(b) College

- Falling skill returns driven mostly by demand factors for non-college workers
- Both supply and demand forces driving decline for college workers

Implications for the Nature of Technical Changes

Demand-driven fall in skill returns after the mid-1980s

- Challenges the skill-biased technical change hypothesis
  - New technologies complement skilled labor and raise its relative demand
  - Skilled workers are also better at adopting new technologies ('Nelson-Phelps hypothesis')
- Consistent with Schumpeterian growth through creative destruction
  - Innovations involve radically new techniques rather than improvements in existing methods
  - Some skills become obsolete, while new set of skills are required
    - Workers who are equally productive today might become differentiated tomorrow $\Rightarrow$ large skill changes
  - Similar to 'turbulence' by Ljungqvist & Sargent (1998, 2008)

Conclusions

- We use panel data to separately identify changes in skill returns from changes in the distributions of labor market skills
  - Simple IV strategy can identify changes in the return to skill
- Using the PSID, we show that
  - Skill returns have *declined* substantially since the mid-1980s
    - Stronger declines for non-college men
  - Variance of unobserved skills increased markedly
    - Driven by increase in variances of skill growth shocks
- Develop an equilibrium framework to interpret the falling returns
  - Driven by both demand and supply factors
  - Fall in demand is more important for non-college workers

Thank You!

Appendix

Identifying Early Skill Returns

- Using future residuals as instruments would allow us to identify growth in early returns, but these are biased: for $t' \geq t + k$,

$$\frac{\text{Cov}(w_t, w_{t'})}{\text{Cov}(w_{t-1}, w_{t'})} = \frac{\mu_t \mu_{t'} \text{Var}(\theta_t)}{\mu_{t-1}\mu_{t'} \text{Var}(\theta_{t-1})} = \frac{\mu_t}{\mu_{t-1}} \left(\frac{\text{Var}(\theta_{t-1}) + \text{Var}(\nu_t)}{\text{Var}(\theta_{t-1})} \right)$$

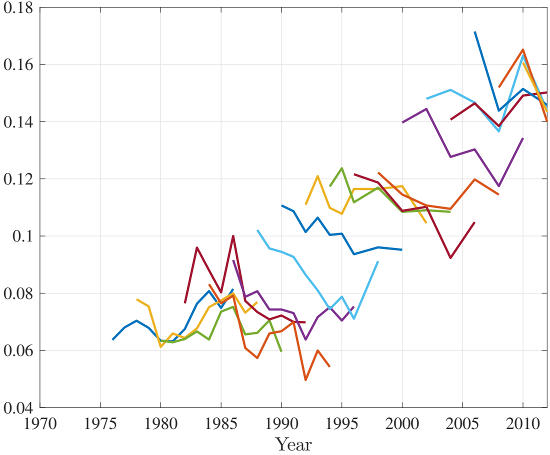
- Can difference this out if we have two cohorts such that $\text{Var}(\nu_t | c) = \text{Var}(\nu_t | c')$
  - Satisfied with U-shaped variance in age/experience (Baker & Solon 2003, Blundell, Graber & Mogstad 2015)
- For $t' \geq t + k$,

$$\frac{\text{Cov}(w_t, w_{t'}|c) - \text{Cov}(w_t, w_{t'}|c')}{\text{Cov}(w_{t-1}, w_{t'}|c) - \text{Cov}(w_{t-1}, w_{t'}|c')} = \frac{\mu_t}{\mu_{t-1}}$$

- So μ_t/μ_{t-1} is identified for all t if $\overline{t} - \underline{t} \geq 2k$

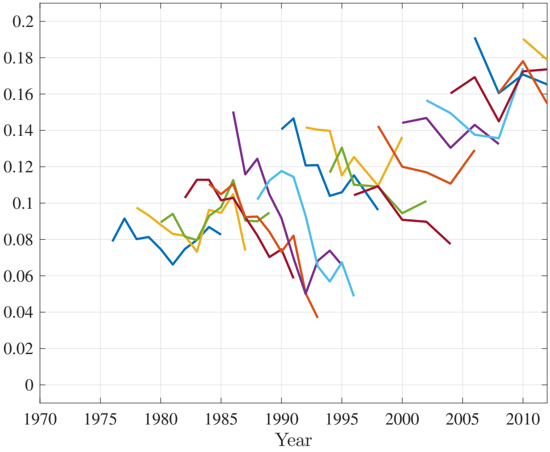
Back

'Balanced' Sample



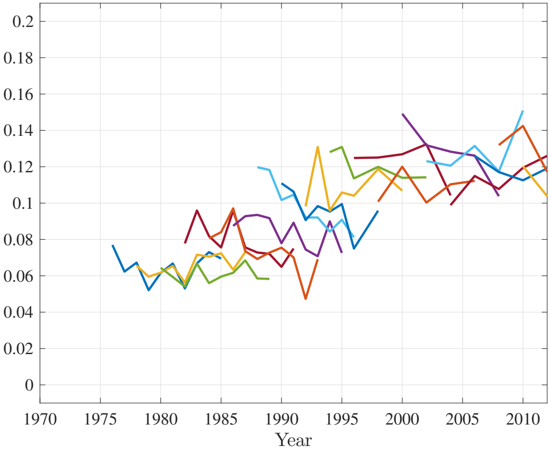
Back

16-30 Years of Experience



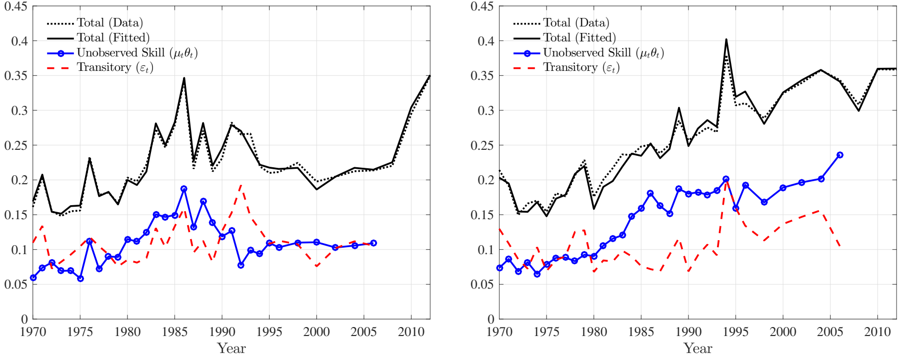
◂ Back

1-15 Years of Experience



◂ Back

Variance Decomposition: Estimated by College Attendance

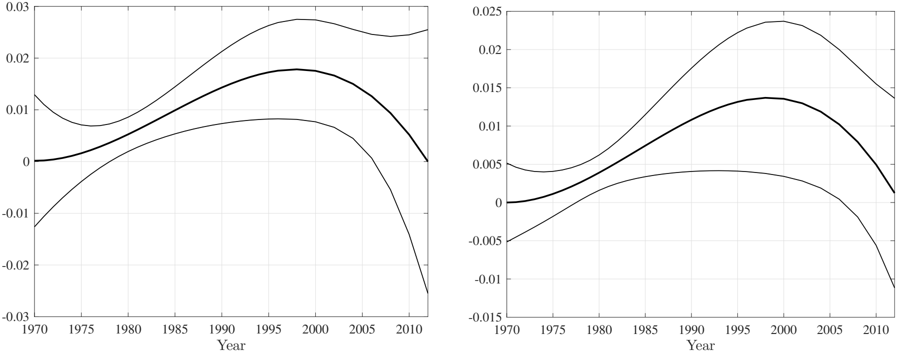


(a) Non-College

(b) College

Back

Time Trends in Skill Shock Variances ($\pi(t)$)

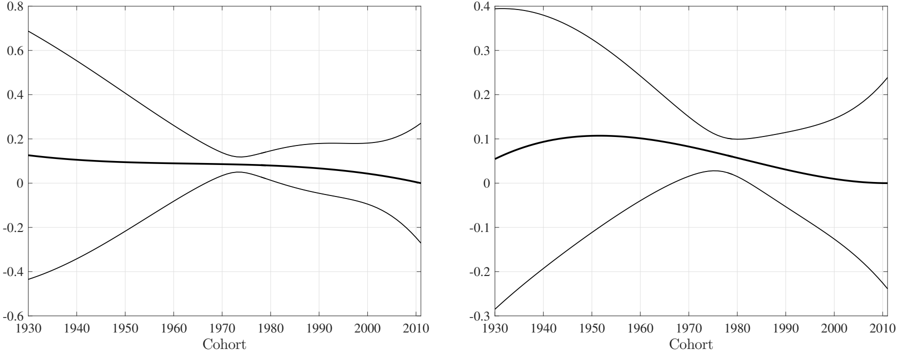


(a) Non-College (b) College

- Rising variance of permanent shocks not purely a composition effect (as in Lemieux (2006))

Back

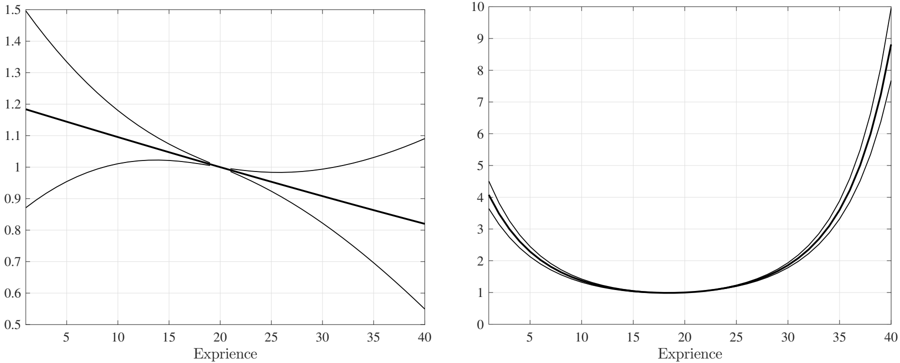
Cohort Trends in Initial Skill Variances



(a) Non-College (b) College

Back

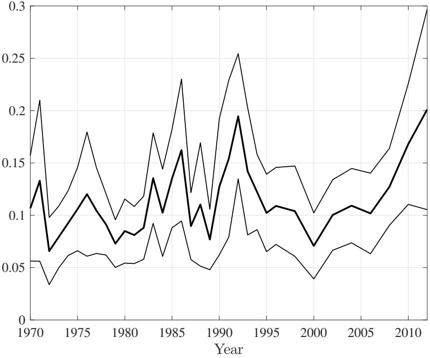
Experience Trends in Skill Shock Variances



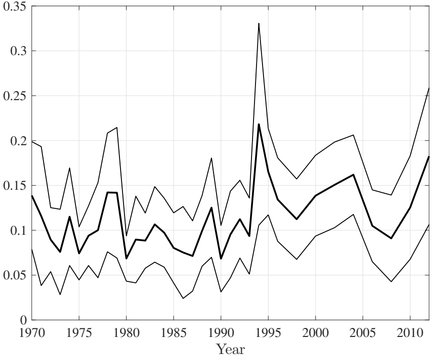
(a) Non-College (b) College

Back

Time Trends in Transitory Shock Variances



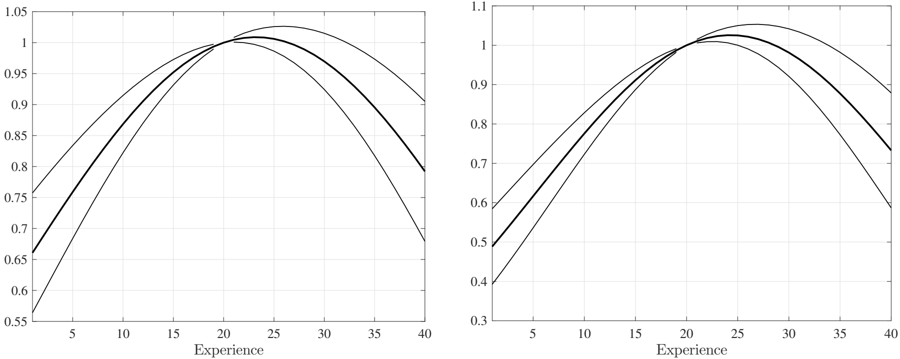
(a) Non-College



(b) College

Back

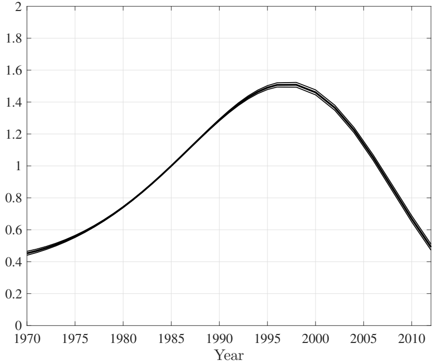
Experience Trends in Transitory Shock Variances



(a) Non-College (b) College

◂ Back

χ_t



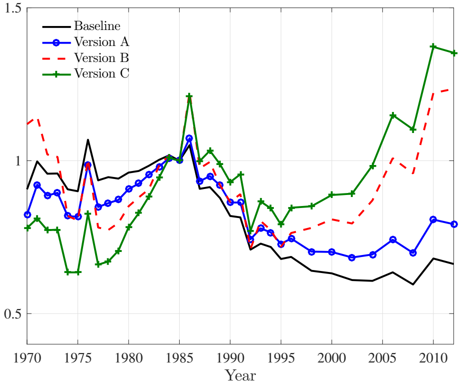
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1970
1975
1980
1985
1990
1995
2000
2005
2010
Year

◂ Back

Importance of Accounting for Time-Varying Skill Variance

- Haider (2001) and Moffitt and Gottschalk (2012) estimated similar models using PSID, but they reached different conclusions about μ_t
- We argue that the difference is due to the time-invariant skill distribution assumed in the previous MD estimates
  - Initial skill variances are identical across cohorts and skill shock variances are constant over time
- Without time-varying skill variances, the model is 'forced' to explain the increase in residual variance via increasing μ_t

Estimated μ_t under Different Restrictions: Full Sample

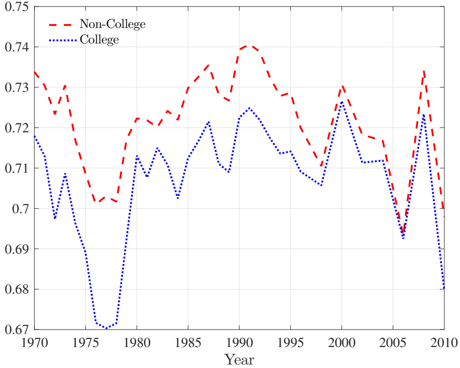


- Version A: ARMA(1,1) transitory component (rather than MA(5))
- Version B: Version A+no time trend in skill shock variance $\pi(t)$
- Version C: Version B+cohort-invariant initial skill variance $\mathrm{Var}(\psi|c)$
    - Very similar to Haider (2001) and Moffitt and Gottschalk (2012)

◂ Back

Labor Share

- Challenge: we need sector-specific labor shares
- We use industry-specific labor shares weighted by fraction of workers by education type
- KLEMS data for US, 1970-2010 (Jorgenson, et al. 2012)



Back

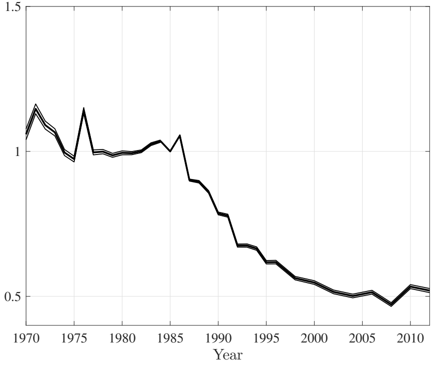
Accounting for Heterogeneous Skill Growth

- Rising skill variance may reflect systematic growth rather than shocks
  - 'Heterogeneous income profile': e.g., Baker (1997), Guvenen (2009)
- Consider general HIP process:

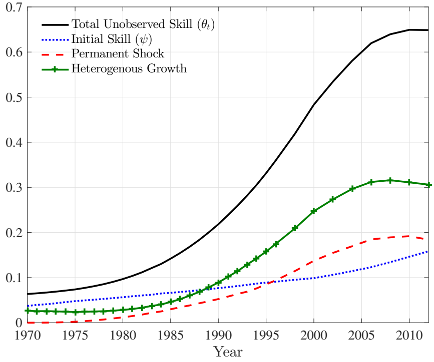
$$\theta_{i,t} = \theta_{i,t-1} + \chi_t \eta(e_{i,t})\delta_i + \nu_{i,t}$$

  - δ_i is a mean zero individual-specific growth rate
    - Uncorrelated with all shocks, but may be correlated with ψ
    - Assume no cohort trend in $\text{Var}(\delta)$ and $\text{Cov}(\delta, \psi)$
  - $\eta(e) = \max\{1 - e/30, 0\}$ accounts for diminishing growth rates
    - No systematic growth after 30 years of experience$\Rightarrow$ helps identify μ_t
  - χ_t allows for time-varying differences in systematic skill growth
    - Assume cubic polynomial in time

Estimates with HIP: Full Sample



(a) μ_t



(b) Skill Variance Decomposition

- Systematic skill growth is important for rising skill variance
- But time pattern of μ_t is robust to the HIP process

▸ χ_t ◂ Back