Revisiting U.S. Wage Inequality at the Bottom 50%

Oren Danieli
Tel Aviv University

NBER Labor Studies 2023

Inequality Trends at the Bottom 50%

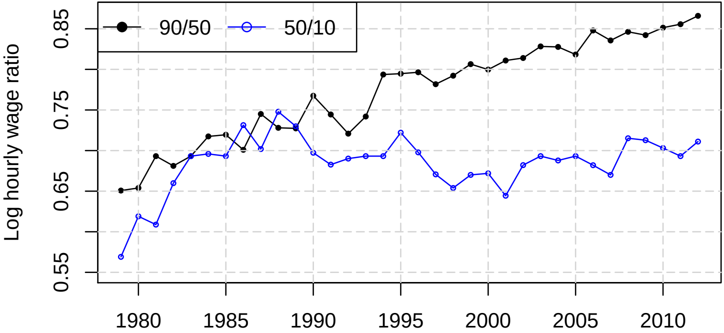


Figure: 90/50 and 50/10 Log Hourly Wage Ratio

Quantiles are calculated for all workers with positive earnings at the hours level, using sample weights multiplied by hours worked. Source: CPS Outgoing Rotation Groups

Leading Hypotheses

In the early 1980s, inequality is rising in both parts of the distribution

- Skill-Biased Technological Change (Katz & Murphy, 1992)

In late 1980s - 1990s inequality decreases at the bottom

- "Wage Polarization" - decline in middle wages ▸ Figure
- Routine-Biased Technological Change (Autor, Katz & Kearney, 2006; Acemoglu & Autor, 2011)
- Decrease in demand for workers performing routine tasks
- Key support: job/employment polarization (Goos et al., 2014)

Key Challenges to RBTC

1. **Why should middle wages relatively decline?**
   - Routine workers are dispersed almost equally at bottom 50%
   - ▸ Figure
2. **Why did middle wages stopped declining around 2000?**
   - Employment polarization continues long after
3. **Why does the market adjusts almost entirely through quantities?**
   - Price changes (wages) is too small to generate trend in wages
     - Autor, Katz & Kearney (2005) and Firpo, Fortin & Lemieux (2013)

Key Challenges to RBTC

1. **Why should middle wages relatively decline?**
   - Routine workers are dispersed almost equally at bottom 50%
   - ▸ Figure

2. **Why did middle wages stopped declining around 2000?**
   - Employment polarization continues long after

3. **Why does the market adjusts almost entirely through quantities?**
   - Price changes (wages) is too small to generate trend in wages
     - Autor, Katz & Kearney (2005) and Firpo, Fortin & Lemieux (2013)

This Paper: A new theory for the trends in the bottom 50% of the income distribution that addresses these challenges

This Paper

1. Theory
    - Small (but important) modification to RBTC
    - Skill-Replacing RBTC
    - Tech does not (directly) replace workers it replaces their skill

This Paper

1. Theory
   - Small (but important) modification to RBTC
   - Skill-Replacing RBTC
   - Tech does not (directly) replace <u>workers</u> it replaces their <u>skill</u>

2. New Empirical Facts
   - Decline in return to skill in routine occupations
   - Reallocation of low-skill workers into routine occupations
   - Interactive-Fixed-Effect-Model

This Paper

1. Theory
   - Small (but important) modification to RBTC
   - Skill-Replacing RBTC
   - Tech does not (directly) replace workers it replaces their skill
2. New Empirical Facts
   - Decline in return to skill in routine occupations
   - Reallocation of low-skill workers into routine occupations
   - Interactive-Fixed-Effect-Model
3. Decomposition
   - 93% of wage polarization can be attributed to SR-RBTC
   - Skewness Decomposition

Theoretical Framework

Assumptions

Building on Jung and Mercenier (2014) and Cortes (2016)

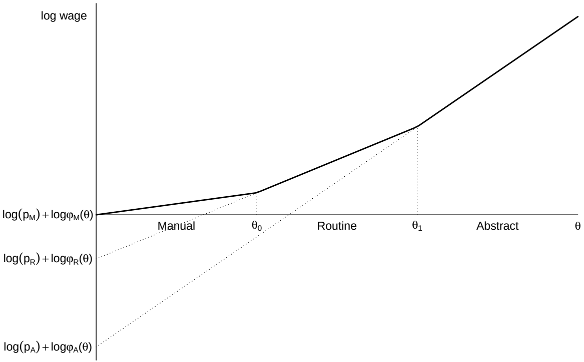
- Workers have one-dimensional skill θ_i
  - Most results hold for multi-dimensional skill
- Three occupations: Manual, Routine, Abstract
- **Key Assumption:** Comparative advantage

$$\forall\theta : \frac{\partial \log \varphi_M(\theta)}{\partial\theta} < \frac{\partial \log \varphi_R(\theta)}{\partial\theta} < \frac{\partial \log \varphi_A(\theta)}{\partial\theta}$$

Theorem (JM): Under these assumptions, there exist two thresholds θ_0, θ_1 such that $\theta < \theta_0$ sort into M, $\theta_0 < \theta < \theta_1$ sort into R and $\theta_1 < \theta$ sort into A.

General Equilibrium

Jung & Mercenier Sorting



RBTC

Focus only on effect on the routine occupation. The production function in the routine occupation is:

$$\varphi_R(\theta_i; \tau) = \left(\theta_i^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

where τ is the technology that is common across all workers. RBTC is $\tau \uparrow$

- $\sigma = 1$ skill neutral similar to Acemoglu & Autor (2011)
- $\sigma < 1$ skill enhancing
- $\sigma > 1$ **skill replacing**

Skill Replacing Technology

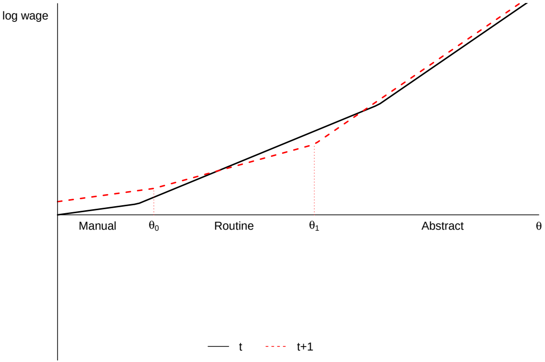
I will focus on the case of Skill-Replacing RBTC

- Increase in τ when $\sigma > 1$

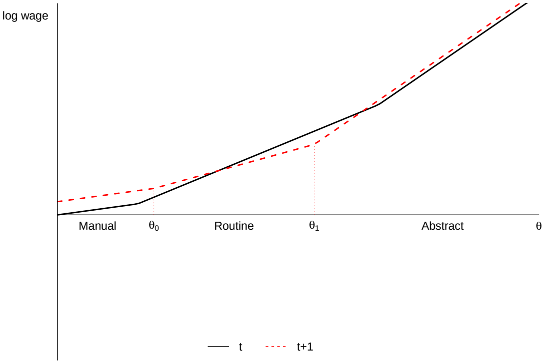
Examples:

- Arithmetic skills are replaced with calculators
- Memory skills are replaced with computers
- Physical strength is replaced with machinery

First Stage: Wage Polarization



First Stage: Wage Polarization

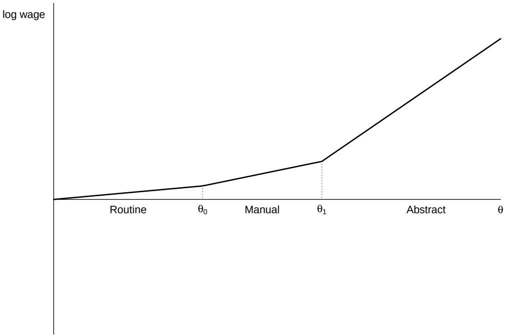


1. **Why should middle wages relatively decline?**

A: Because these are the highest skill routine workers

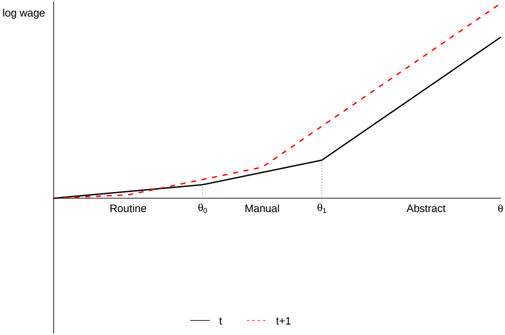
Second Stage: Bottom 50% Inequality Rises

Large SR-RBTC: comp. advantage flips $\frac{\partial \log \varphi_R(\theta;\tau)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta)}{\partial \theta}$



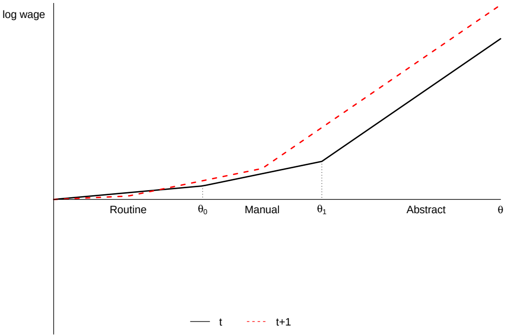
Second Stage: Bottom 50% Inequality Rises

Large SR-RBTC: comp. advantage flips $\frac{\partial \log \varphi_R(\theta;\tau)}{\partial\theta} < \frac{\partial \log \varphi_M(\theta)}{\partial\theta}$



Second Stage: Bottom 50% Inequality Rises

Large SR-RBTC: comp. advantage flips $\frac{\partial \log \varphi_R(\theta;\tau)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta)}{\partial \theta}$



2. **Why did middle wages stopped declining around 2000?**

A: Middle-wage workers are no longer in the routine occupation

- bottom 50% inequality could increase

List of New Predictions

1. Decline in return to skill in routine occupations

List of New Predictions

1. Decline in return to skill in routine occupations
2. Routine workers gradually become less skilled
   - Eventually, routine workers have less skill than manual

List of New Predictions

1. Decline in return to skill in routine occupations
2. Routine workers gradually become less skilled
    - Eventually, routine workers have less skill than manual
3. Routine workers become more concentrated at lower wages

Empirical Results

IFEM

Skill is not directly observed

- I use panel data, assume that skill is constant over time

Use Interactive Fixed Effect Model (**IFEM**) ▸ Why?

$$\log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \alpha_{jt}\theta_i + \varepsilon_{ijt}$$

i - worker, j - 3 occupation categories, t - year and X_{it} experience^2.

We are interested in:

1. How $\alpha_{routine,t}$ changes with time
2. How average routine skill $\frac{1}{N_R}\sum_{i \in R} \widehat{\theta}_i$ change

▸ Estimation

IFEM

Skill is not directly observed

- I use panel data, assume that skill is constant over time

Use Interactive Fixed Effect Model (**IFEM**) ▸ Why?

$$\log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \quad \theta_i + \varepsilon_{ijt}$$

i - worker, j - 3 occupation categories, t - year and X_{it} experience^2.

We are interested in:

1. How $\alpha_{routine,t}$ changes with time
2. How average routine skill $\frac{1}{N_R} \sum_{i \in R} \widehat{\theta_i}$ change

▸ Estimation

IFEM

Skill is not directly observed

- I use panel data, assume that skill is constant over time

Use Interactive Fixed Effect Model (**IFEM**) ▸ Why?

$$\log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \alpha_{jt} \theta_i + \varepsilon_{ijt}$$

i - worker, j - 3 occupation categories, t - year and X_{it} experience^2.

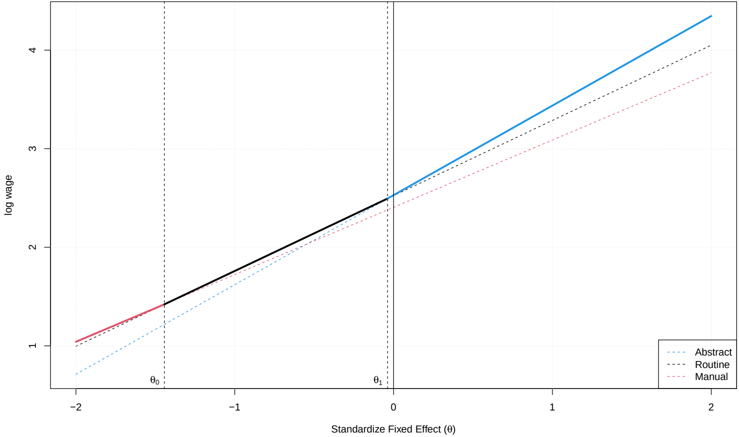
We are interested in:

1. How $\alpha_{routine,t}$ changes with time
2. How average routine skill $\frac{1}{N_R} \sum_{i \in R} \widehat{\theta}_i$ change

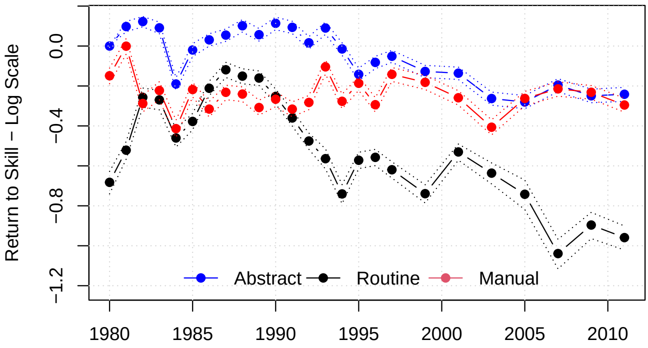
▸ Estimation

Results for 1-Year: 1987

Predicted log wage in each occupation as a function of skill θ

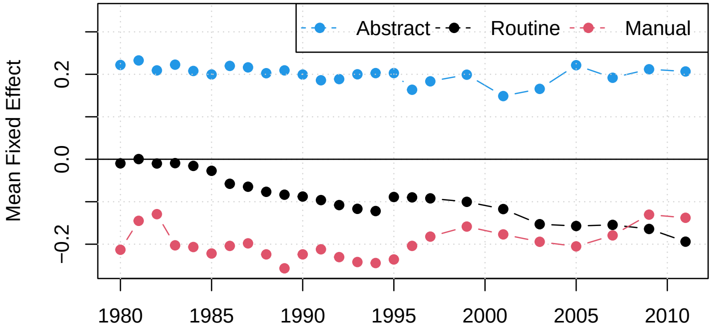


Long Term Trend of α_{jt}



▸ 1-Digit Occupational Category

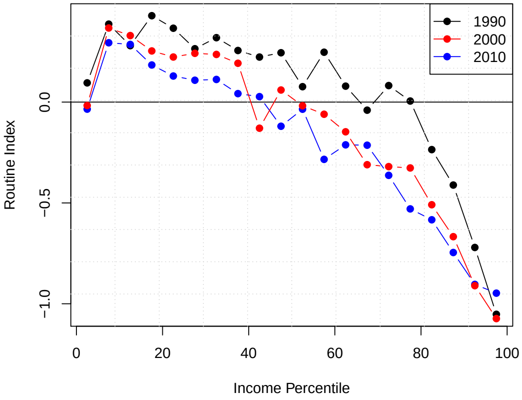
Decline in Skill in Routine Occupations



▸ 1-Digit Occupational Category

Routine by Income Percentile

Routine task intensity measured by occupation with O*NET



Quantifying the Role of SR-RBTC Using Skewness Decomposition

Why Decompose?

SR-RBTC is consistent with the data

- But is it large enough to explain the full wage trend?
- Or maybe other explanations also play a role

This is the motivation for decomposition exercise

- Which share of the overall trend can be attributed to different hypotheses
- Focus in the period of "wage polarization"
- Inequality at the bottom is relatively stable afterwards

Skewness Decomposition

Can measure wage polarization with the third-moment: Skewness

▸ Skewness Over Time ▸ Influence Function

$$\mu_3(Y) = E\left[\left(\frac{Y-\mu}{\sigma}\right)^3\right]$$

Skewness Decomposition

Can measure wage polarization with the third-moment: Skewness

▸ Skewness Over Time ▸ Influence Function

$$\mu_3(Y) = E\left[\left(\frac{Y-\mu}{\sigma}\right)^3\right]$$

Similar to variance, skewness has a simple decomposition

$$\mu_3(Y) = \underbrace{E\left[\mu_3(Y|X)\right]}_{Within} + \underbrace{\mu_3\left(E\left[Y|X\right]\right)}_{Between} + \underbrace{3COV\left(E\left[Y|X\right], V\left[Y|X\right]\right)}_{Correlation}$$

Interpretation

$$\mu_3(Y) = \underbrace{E[\mu_3(Y|X)]}_{Within} + \underbrace{\mu_3(E[Y|X])}_{Between} + \underbrace{3COV(E[Y|X], V[Y|X])}_{Correlation}$$

Set X to be occupation

- **Within component** - non-occupation explanations (residual)
- **Between component** - skill-neutral RBTC: decrease in routine wages
  - Should be main change in Acemoglu & Autor (2011) ($p_R \downarrow$)
- **Correlation component** - higher if:
  - High paying occupations have higher inequality.
  - Low paying occupations have lower inequality.
  - SR-RBTC: decrease in inequality within (low-paid) routine occupations
  - Captures violation of ignorability

Skewness Decomposition by Occupation

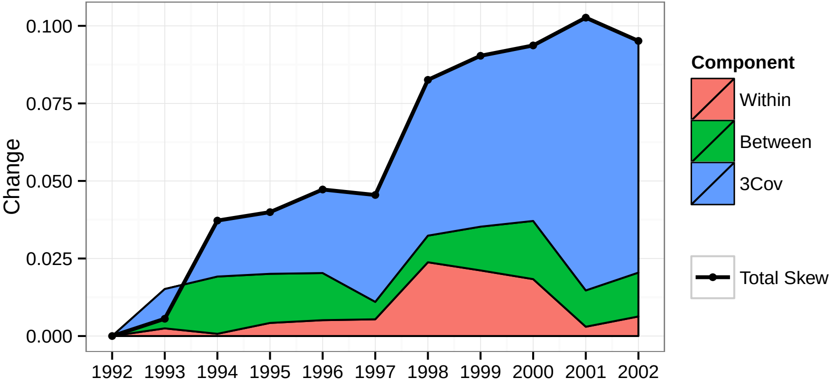


Figure: Skewness Decomposition Changes 1992-2002

Data resource: CPS-ORG

3 digit Industry | Years of School | 1980-2010

Changes in Variance

- Increase in the covariance component is driven by within-occupation inequality ▸ Details
- Inequality is increases at high-paying and decreases at low-paying occupations ▸ Details
- The decrease in inequality in low paying occupations is unique for the 1990s ▸ Details
- This decrease is concentrated in routine occupations ▸ Details

Changes in Variance

- Increase in the covariance component is driven by within-occupation inequality ▸ Details
- Inequality is increases at high-paying and decreases at low-paying occupations ▸ Details
- The decrease in inequality in low paying occupations is unique for the 1990s ▸ Details
- This decrease is concentrated in routine occupations ▸ Details

3. **Why does the market adjust through quantities?**
   A: Significant wage changes within routine occupations

Conclusion

Key Takeaways

1. SR-RBTC model can explain the puzzles with RBTC
   - Why middle wage decline in 1990s
   - Why inequality at the bottom fluctuates
   - Why previous decomposition methods did not work
2. Predictions of the model are verified in the data
3. Skewness Decomposition shows this explains most of the trend
   - R-package available at CRAN

Thank You!

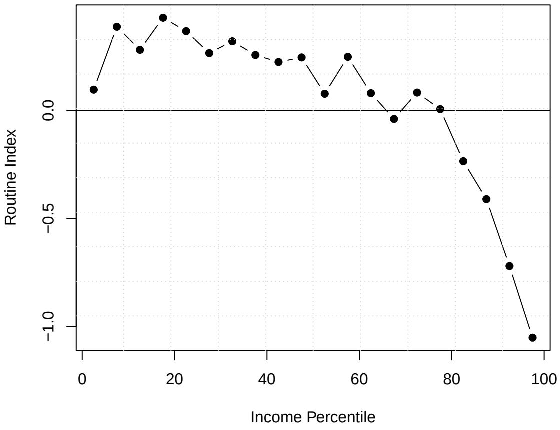
Appendix

Wage Growth by 5% Bins

Return

Routine Level by Income Percentile

Replication of Figure in Autor & Dorn (2013, Fig 4)



Routine index is defined using O*NET data ▸ Details ▸ Return

Routine Index O*NET

Following Acemoglu-Autor (2011) use O*NET to take the average of

- Pace determined by speed of equipment
- Controlling machines and processes
- Spend time making repetitive motions.
- Importance of repeating the same tasks
- Importance of being exact or accurate
- Structured v. Unstructured work (reverse)

Proposition 1

Proposition: Let $w_a < w_b$ denote wages of two routine workers. The effect of RBTC ($\tau \uparrow$) on the wage ratio $\frac{w_b}{w_a}$ depends on

$$sign\left(\frac{\partial \frac{w_b}{w_a}}{\partial \tau}\right) = sign\left(1 - \sigma\right)$$

Return

RBTC

Focus only on effect on the routine occupation. RBTC is $\tau \uparrow$

$$\varphi_R(\theta_i;\tau) = \left(\theta_i^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

- $\sigma = 1$ skill neutral similar to Acemoglu & Autor (2011)
- $\sigma < 1$ skill enhancing
- $\sigma > 1$ **skill replacing**
- Return

General Equilibrium ▸ Return

Total amount produced from each intermediate good

$$M = \int_{\theta_{min}}^{\theta_0} \varphi_M(\theta)\,\mathrm{d}\theta \quad R = \int_{\theta_0}^{\theta_1} \varphi_R(\theta)\,\mathrm{d}\theta \quad A = \int_{\theta_1}^{\theta_{max}} \varphi_A(\theta)\,\mathrm{d}\theta$$

General Equilibrium ▸ Return

Total amount produced from each intermediate good

$$M = \int_{\theta_{min}}^{\theta_0} \varphi_M(\theta)\, \mathrm{d}\theta \quad R = \int_{\theta_0}^{\theta_1} \varphi_R(\theta)\, \mathrm{d}\theta \quad A = \int_{\theta_1}^{\theta_{max}} \varphi_A(\theta)\, \mathrm{d}\theta$$

The final good is the output of a CES function with $\rho < 0$

$$Y = (M^\rho + R^\rho + A^\rho)^{\frac{1}{\rho}}$$

Manual and abstract workers become more productive through complementarities

- ▸ Theorem

SR-RBTC

I will focus on the case of Skill-Replacing RBTC

- Increase in τ when $\sigma > 1$

As technology advances ($\tau \uparrow$) the routine occupation see a decline in:

- Price of routine goods (p_R)
- Employment

SR-RBTC

I will focus on the case of Skill-Replacing RBTC

- Increase in τ when $\sigma > 1$

As technology advances ($\tau \uparrow$) the routine occupation see a decline in:

- Price of routine goods (p_R)
- Employment

- Mean skill level ($E[\theta_i|R]$)
- Inequality within the routine occupation

SR-RBTC: First Stage

Impact on bottom 50% inequality changes with time

- Divide it into two stages

In the first stage, τ is still "small"

- Comparative advantage still holds
- Returns to skill are higher in R then M

During the first stage, overall wage trend would be U-Shaped

- ▸ Theorems

GE Theorem

Theorem

Assume $\rho < 0$, so $\tau \uparrow$ implies decrease in p_R and the income share of routine workers

- Does not depend on σ
- Empirically shown by Cortes (2016), Eden & Gaggl (2018)

▸ Return

Weaker Assumptions

Theorem

Assuming a skill replacing technology ($\sigma > 1$). An RBTC (increase in τ) would generate:

1. *A decline in gaps between routine workers who do not switch occupations*
2. *The most skilled routine workers would leave the routine occupation ($\frac{\partial \theta_1}{\partial \tau} < 0$)*
3. *Wages for the highest skill routine worker (θ_1) would fall relative to any other worker.*

Stronger Assumptions

Assume $0 < \frac{d\theta_0}{d\tau} < \left|\frac{d\theta_1}{d\tau}\right|$ as seen in the data.

Theorem

SR-RBTC generates

1. *Decline in: employment, within occupation inequality and mean skill level in the routine occupation.*
2. *Overall wage trend would be U-shaped ("wage polarization")*

▸ Return

Theorem

Theorem

There exists $\widetilde{\tau}$, such that for every $\tau \geq \widetilde{\tau}$

$$\frac{\partial \log \varphi_R(\theta;\tau)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta)}{\partial \theta}$$

and routine workers would earn the lowest wages.
Any additional SR-RBTC ($\tau \uparrow$) would (still)

- *Decrease employment in the routine occupation ($\frac{d\theta_0}{d\tau}$<0)*
- *Decrease gaps between routine workers who do not switch occupation*

▸ Return

Testing Decline in Return to Skill

The key prediction of the model is that inequality is declining within routine occupations

- But this is only for "stayers" - those who do not switch occupations
- Overall inequality in routine occupations is affected by compositional changes

Testing Decline in Return to Skill

The key prediction of the model is that inequality is declining within routine occupations

- But this is only for "stayers" - those who do not switch occupations
- Overall inequality in routine occupations is affected by compositional changes

There are several challenges in measuring inequality for stayers

1. Regression to mean
2. Selected sample (especially over long time periods)
3. Can be confused with income volatility

▸ Return

Estimation

θ_i is a nuisance parameter. Can only get some estimate of it $\widehat{\theta}_i$ based on a small number of observations. ▸ Details

Problem: $\widehat{\theta}_i$ is noisy, so least squares will suffer from attenuation bias because

$$E\left[\widehat{\theta}_i \varepsilon_{ijt}\right] \neq 0$$

Therefore we need additional moments.

- Holtz-Eakin et al. (1988), Ahn et al. (2001)
- ▸ Details

Years of Schooling ‣ Return

$$\log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \alpha_{jt} \theta_i + \varepsilon_{ijt}$$

I use years of schooling S_i as an instrument

- For every occupation j and year t

$$E\left[S_i \varepsilon_{ijt} | j, t\right] = 0$$
$$E\left[X_{ijt} \varepsilon_{ijt} | j, t\right] = 0$$
$$E\left[\varepsilon_{ijt} | j, t\right] = 0$$

  where ε_{ijt} is a function of the parameters $(\alpha_{jt}, \beta_{jt}, \lambda_{jt})$.
- Estimate using GMM.

Years of Schooling ▸ Return

$$\log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \alpha_{jt}\theta_i + \varepsilon_{ijt}$$

I use years of schooling S_i as an instrument

- For every occupation j and year t

$$E[S_i \varepsilon_{ijt} | j, t] = 0$$
$$E[X_{ijt} \varepsilon_{ijt} | j, t] = 0$$
$$E[\varepsilon_{ijt} | j, t] = 0$$

  where ε_{ijt} is a function of the parameters $(\alpha_{jt}, \beta_{jt}, \lambda_{jt})$.

- Estimate using GMM.

Biased when school affects wages not through "main skill"

- Example: bonus for useless degrees
- Results with three skills ▸ Details

IFEM-Literature

- Holtz-Eakin et al. (1988) use lagged variables
  - Violated (for instance) is ε are serially correlated
- Ahn et al. (2001) add assumption on covariance structure of $V(\varepsilon_{ijt})$
  - For instance - constant variance for ε
  - Rules out changes in volatility

Return

$\widehat{\theta}$

Define

$$\nu_{ijt} = \frac{1}{\alpha_{jt}}\left(y_{ijt} - \beta_{jt}X_{ijt} - \lambda_{jt}\right) = \theta_i + \frac{\varepsilon_{ijt}}{\alpha_{jt}}$$

For every $\sum_t w_{ijt} = 1$ can define

$$\widehat{\theta}\left(y_i, X_i, \alpha, \beta, \lambda\right) = \sum_t w_{ijt}\overline{\nu_{ijt}} = \theta_i + \widetilde{\varepsilon}_i \tag{1}$$

such that

$$y_{ijt} - \beta_{jt}X_{it} - \lambda_{jt} - \alpha_{jt}\widehat{\theta}_i = \varepsilon_{ijt} - \alpha_{jt}\widetilde{\varepsilon}_i = \epsilon_{ijt}$$

I choose $w_{ijt} = \frac{\alpha_{jt}^2}{\sum_{j't'} \alpha_{j't'}^2}$ which minimizes the mean squared error $\overline{\epsilon_{ijt}^2}$.

▸ Return

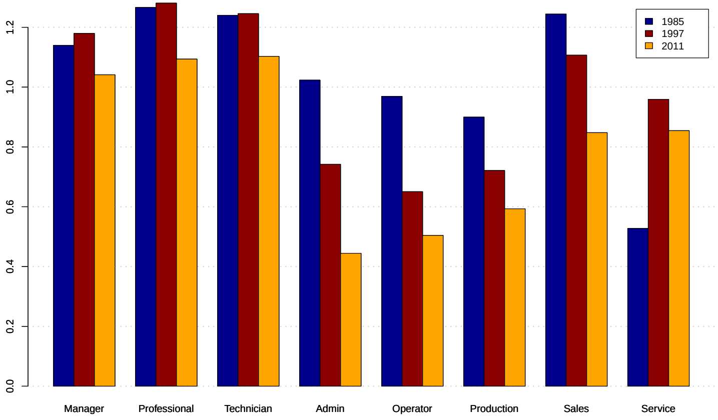
Three Skills

Estimate IFEM with

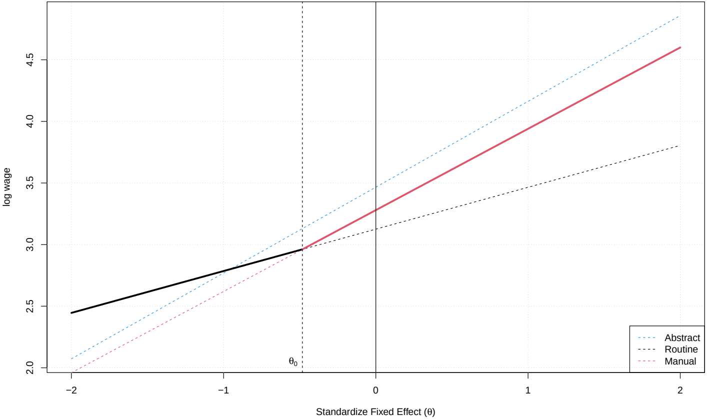
$$\log w_{ijt} = \beta_{ijt} X_{it} + \lambda_{jt} + \alpha_{jt}\theta_{ij} + \varepsilon_{ijt}$$

	Abstract	Routine	Manual
Abstract	1		
Routine	.74	1	
Manual	.83	.69	1

α_{jt} by 1-Digit ▸ Return

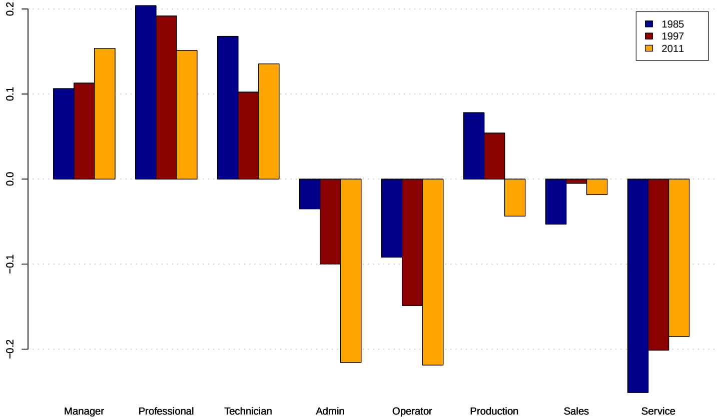


IFEM 2011

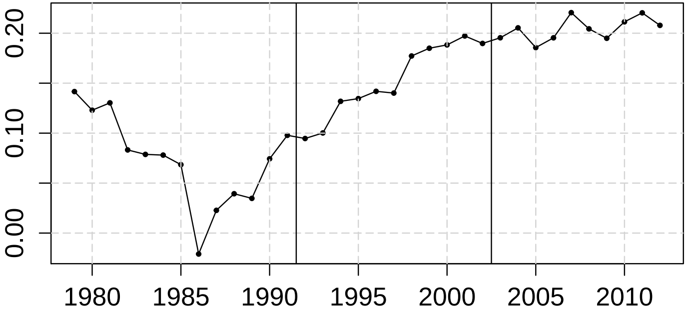


Return

Decline in Skill in Routine: 1-Digit ▸ Return



Skewness Trend ▸ Back



The vertical lines are where changes in occupational coding took part. Source: CPS Outgoing Rotation Groups

Robustness ▸ Back

Looking by other categories yields large residual component

- ▸ 3 digit Industry
- ▸ Years of School

Decomposing jointly shows occupations explain the large increase

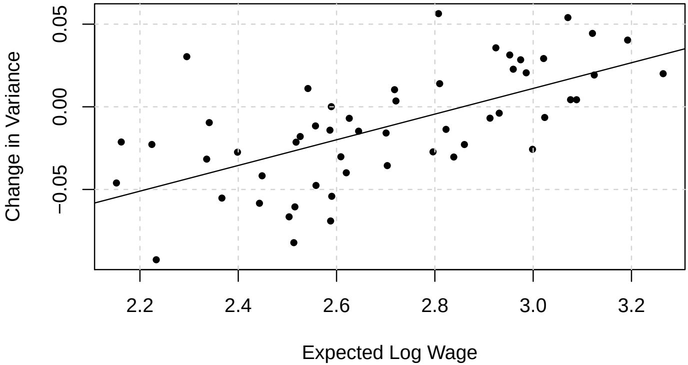
- ▸ Details

- ▸ 3 digit Industry
- ▸ Years of School

Longer time period ▸ Details

Using imputed wages ▸ Details

Changes in Variance 1992-2002 ▸ Return



Documented before by Firpo et al. (2013)

- Explains full increase in covariance component ▸ Decompose

Variance Trends in Other Decades ▸ Return

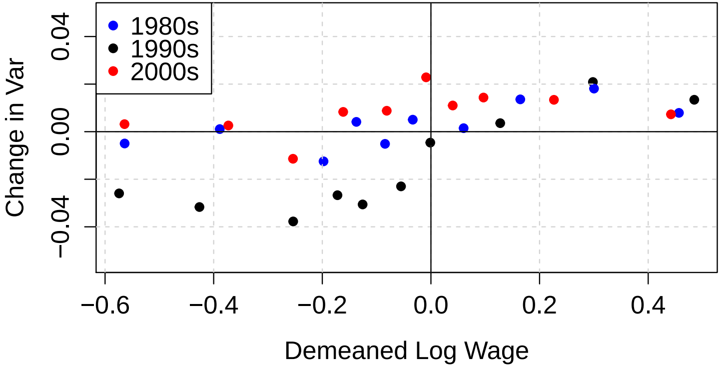


Figure: Change in $V[\ln w|occ]$ by $E[\ln w|occ]$ - Binned Scatter Plot

Data resource: CPS-ORG

Variance Trend in Routine/Non-routine Occupations ‣ Return

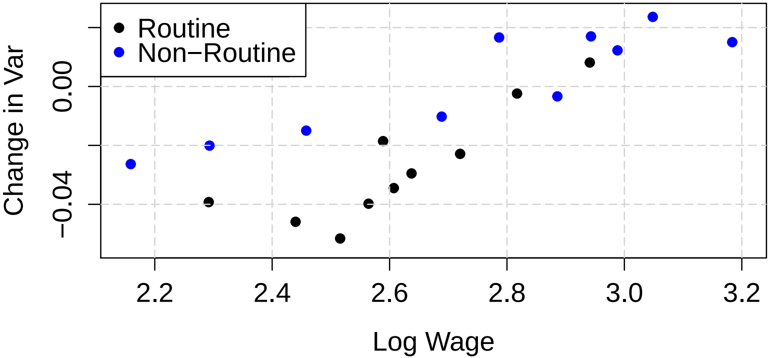


Figure: Change in $V[\ln w|occ]$ by $E[\ln w|occ]$ 1992-2002

Data resource: CPS-ORG. Routine occupations are administrators, producers and operators. Categories are divided same as in Acemoglu & Autor (2011)

Counterfactual Covariance ▸ Return

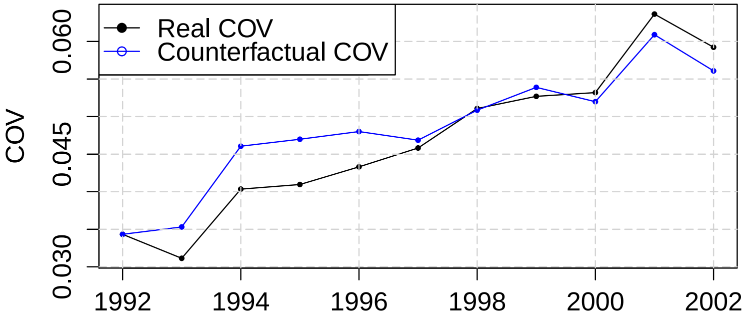
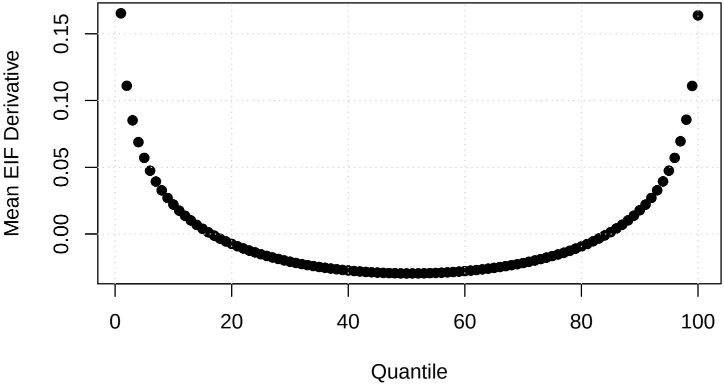


Figure: Covariance of Expectation and Variance of Log-Wage

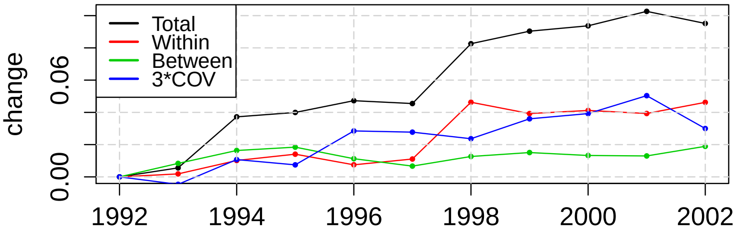
Data resource: CPS-ORG ▸ return

Influence Function



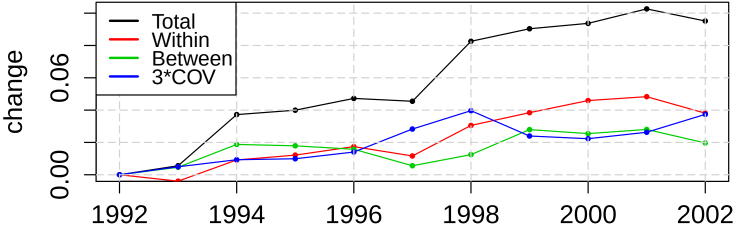
▸ Return

Decomposing by Industry



Return

Decomposing by Education and Experience



Return

Linear Skewness Decomposition

If $Y = \sum_i X_i$ can write

$$\mu_3(Y) = \sum_i \mu_3(X_i) + \sum_i \sum_{j \neq i} COV\left(X_i^2, X_j\right) + \sum_i \sum_{j \neq i} \sum_{k \neq i,j} E\left[X_i X_j X_k\right] \tag{2}$$

and decompose into several components. The simple skins decomposition is for $Y = E[Y|X] + \varepsilon$

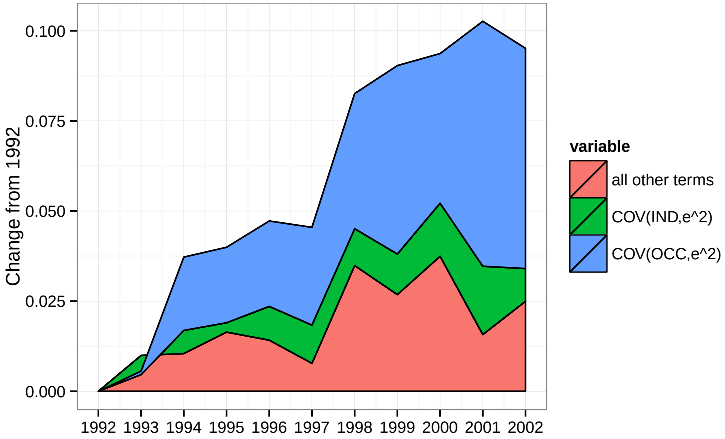
Can first run a regression such as

$$\ln w_i = occ_i + ind_i + \varepsilon_i$$

and decompose by each component.

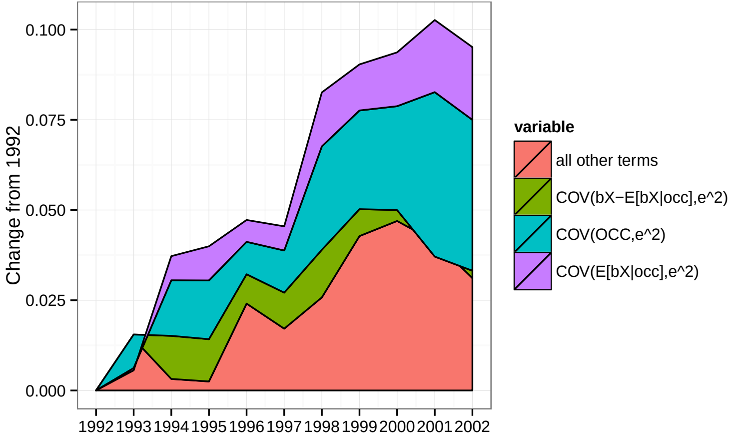
Return

Joint Occupation-Industry Decomposition



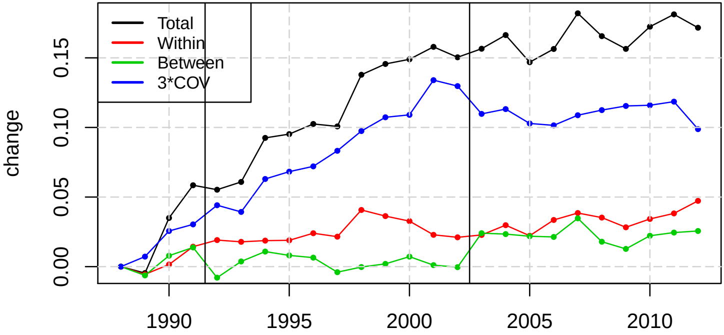
▸ Return

Joint Occupation-School-Experience Decomposition



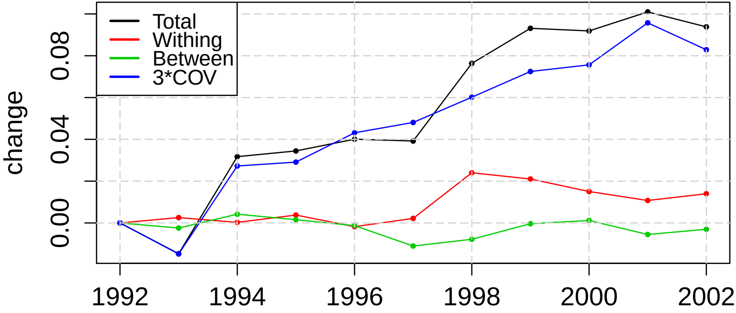
Return

Decomposition with Imputed Wages



Return

Decomposition with Imputed Wages



▸ Return