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
- 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%

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
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

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

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 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
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 $\theta_i$
  - 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 $\theta_0, \theta_1$ 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

\[ \log(\text{wage}) = \log(p_M) + \log(\phi_M(\theta)) + \log(p_R) + \log(\phi_R(\theta)) + \log(p_A) + \log(\phi_A(\theta)) \]
Focus only on effect on the routine occupation. The production function in the routine occupation is:

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

where $\tau$ 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 $\tau$ 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

1. Why should middle wages relatively decline?
   A: Because these are the highest skill routine workers.
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 \( \frac{\partial \log \varphi_R(\theta; \tau)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta)}{\partial \theta} \)

- Why did middle wages stopped declining around 2000?
  - A: Middle-wage workers are no longer in the routine occupation

- Bottom 50% inequality could increase
Second Stage: Bottom 50% Inequality Rises

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

![Graph showing log wage over different occupations and time periods](image_url)
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
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
Skill is not directly observed

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

Use Interactive Fixed Effect Model (IFEM)

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

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

We are interested in:

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

Estimation
Skill is not directly observed

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

Use Interactive Fixed Effect Model (IFEM) Why?

\[
\log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \theta_i + \varepsilon_{ijt}
\]

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

We are interested in:

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

Estimation
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 \( \alpha_{\text{routine},t} \) changes with time
2. How average routine skill \( \frac{1}{N_R} \sum_{i \in R} \hat{\theta}_i \) change

Estimation
Results for 1-Year: 1987

Predicted log wage in each occupation as a function of skill $\theta$
Long Term Trend of $\alpha_{jt}$

![Graph showing the long-term trend of $\alpha_{jt}$ over the years 1980 to 2010. The graph illustrates the change in return to skill for different occupational categories: Abstract, Routine, and Manual. The x-axis represents the years, and the y-axis represents the return to skill on a log scale. The occupational categories are highlighted with different colors and styles: Abstract in blue, Routine in black, and Manual in red.](image-url)

- 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

![Graph showing routine task intensity by income percentile](image-url)
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

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

Can measure wage polarization with the third-moment: Skewness

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

Similar to variance, skewness has a simple decomposition

\[ \mu_3(Y) = E [\mu_3(Y|X)] + \mu_3(E[Y|X]) + 3\text{COV}(E[Y|X], V[Y|X]) \]

Within \hspace{1cm} Between \hspace{1cm} Correlation
Interpretation

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

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.
- Inequality is increases at high-paying and decreases at low-paying occupations.
- The decrease in inequality in low paying occupations is unique for the 1990s.
- This decrease is concentrated in routine occupations.
Changes in Variance

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

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

![Graph showing wage growth by 5% bins](image-url)
Routine Level by Income Percentile

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

Routine index is defined using O*NET data
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

$$\text{sign} \left( \frac{\partial \frac{w_b}{w_a}}{\partial \tau} \right) = \text{sign} (1 - \sigma)$$
Focus only on effect on the routine occupation. RBTC is $\tau \uparrow$

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

- $\sigma = 1$ skill neutral similar to Acemoglu & Autor (2011)
- $\sigma < 1$ skill enhancing
- $\sigma > 1$ skill replacing
Total amount produced from each intermediate good

\[
M = \int_{\theta_{\text{min}}}^{\theta_0} \varphi_M(\theta) \, d\theta \quad R = \int_{\theta_0}^{\theta_1} \varphi_R(\theta) \, d\theta \quad A = \int_{\theta_1}^{\theta_{\text{max}}} \varphi_A(\theta) \, d\theta
\]
Total amount produced from each intermediate good

\[ M = \int_{\theta_{\text{min}}}^{\theta_{0}} \varphi_M (\theta) \, d\theta \quad R = \int_{\theta_0}^{\theta_1} \varphi_R (\theta) \, d\theta \quad A = \int_{\theta_1}^{\theta_{\text{max}}} \varphi_A (\theta) \, d\theta \]

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
I will focus on the case of Skill-Replacing RBTC

- Increase in $\tau$ 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 $\tau$ 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, $\tau$ 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
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 $\sigma$
- Empirically shown by Cortes (2016), Eden & Gaggl (2018)
Weaker Assumptions

Theorem

Assuming a skill replacing technology ($\sigma > 1$). An RBTC (increase in $\tau$) 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 ($\theta_1$) 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")**
Theorem

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

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

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
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.
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
Estimation

$\theta_i$ is a nuisance parameter. Can only get some estimate of it $\hat{\theta}_i$ based on a small number of observations.

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

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

Therefore we need additional moments.

- Holtz-Eakin et al. (1988), Ahn et al. (2001)
\[ \log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \alpha_{jt} \theta_i + \epsilon_{ijt} \]

I use years of schooling \( S_i \) as an instrument

- For every occupation \( j \) and year \( t \)

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

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

- Estimate using GMM.
Years of Schooling

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

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 $\varepsilon_{ijt}$ 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
Holtz-Eakin et al. (1988) use lagged variables
- Violated (for instance) is $\varepsilon$ are serially correlated

Ahn et al. (2001) add assumption on covariance structure of $V(\varepsilon_{ijt})$
- For instance - constant variance for $\varepsilon$
- Rules out changes in volatility
Define

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

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

\[ \hat{\theta} (y_i, X_i, \alpha, \beta, \lambda) = \sum_t w_{ijt} \nu_{ijt} = \theta_i + \tilde{\varepsilon}_i \] \hspace{1cm} (1)

such that

\[ y_{ijt} - \beta_{jt} X_{it} - \lambda_{jt} - \alpha_{jt} \hat{\theta}_i = \varepsilon_{ijt} - \alpha_{jt} \tilde{\varepsilon}_i = \epsilon_{ijt} \]

I choose \( w_{ijt} = \frac{\alpha_{jt}^2}{\sum_{j't'} \alpha_{j't'}^2} \) which minimizes the mean squared error \( \varepsilon_{ijt}^2 \).
Estimate IFEM with

\[ \log w_{ijt} = \beta_{ijt} X_{it} + \lambda_{jt} + \alpha_{jt} \theta_{ij} + \varepsilon_{ijt} \]
<table>
<thead>
<tr>
<th></th>
<th>Abstract</th>
<th>Routine</th>
<th>Manual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Routine</td>
<td>.74</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Manual</td>
<td>.83</td>
<td>.69</td>
<td>1</td>
</tr>
</tbody>
</table>
$\alpha_{jt}$ by 1-Digit
Abstract
Routine
Manual

Return
Decline in Skill in Routine: 1-Digit

Manager Professional Technician Admin Operator Production Sales Service
−0.2 −0.1 0.0 0.1 0.2

1985       1997
2011

1985 1997 2011

Manager Professional Technician Admin Operator Production Sales Service
−0.2 −0.1 0.0 0.1 0.2
The vertical lines are where changes in occupational coding took part. Source: CPS Outgoing Rotation Groups
Robustness

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

Documented before by Firpo et al. (2013)

- Explains full increase in covariance component
Variance Trends in Other Decades

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

**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)
Figure: Covariance of Expectation and Variance of Log-Wage

Data resource: CPS-ORG
Influence Function

![Graph showing the relationship between quantile and mean EIF derivative.]
Decomposing by Industry
Decomposing by Education and Experience
Linear Skewness Decomposition

If \( Y = \sum_i X_i \) can write

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

(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.
Joint Occupation-Industry Decomposition
Joint Occupation-School-Experience Decomposition
Decomposition with Imputed Wages

The graph illustrates changes in various economic indicators from 1990 to 2010. The y-axis represents the change in percentage, ranging from 0.00 to 0.15. The x-axis represents the years: 1990, 1995, 2000, 2005, and 2010.

Key indicators include:
- **Total**
- **Within**
- **Between**
- **3*COV**

Each category is represented by a different color and line style. The graph shows trends and changes over the specified period.
Decomposition with Imputed Wages

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Within</th>
<th>Between</th>
<th>3*COV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>