A very old topic - data goes back to 1905

**TABLE I. — MEAN DAILY WAGES OF ITALIAN WOMEN ACCORDING TO THEIR AGES AND THE SIZES OF THE ESTABLISHMENTS IN WHICH THEY WERE AT WORK**

<table>
<thead>
<tr>
<th>Age of Employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 20 Employees</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>15-20</td>
</tr>
<tr>
<td>20-35</td>
</tr>
<tr>
<td>35-55</td>
</tr>
<tr>
<td>Above 55</td>
</tr>
</tbody>
</table>

"Il progresso dell' Economia politica dipenderà pel futuro in gran parte dalla ricerca di leggi empiriche, ricavate dalla statistica, e che si paragoneranno poi colle leggi teoriche note, o che ne faranno conoscere di nuove." Pareto.
Many papers over the decades also similar findings of a large firm pay premium – e.g.

Slichter (1950)
Lester (1967)
Brown and Medoff (1989)
Oi and Idson (1999)
Our Large Firm Wage Premium (LFWP) Paper

Use two massive datasets – SSA data on all W2 pay slips since 1978 and Census data on all firms since 1976, finding:

1. LFWP falls by about 50%, mainly due to falling large firm AKM Fixed-Effect (not due to less worker sorting)

2. Appears to particularly impact lower paid/educated workers

3. Associated with two industry factors in particular:
   • Shrinking manufacturing (which has a high LFWP)
   • Growth of low paying service sector (e.g. big-box retail)
Outline

1) Data (SSA and Census Data)

2) SSA results

3) Census Data

4) Implications for inequality
Social Security Administration (SSA) data is the Master Earnings File (MEF)

Universe of all W-2s from 1978 to 2013 (about 100m per year)

For each job: SSN, EIN and total compensation:

“Total compensation includes: wages, salaries, tips, restricted stock grants, exercised stock options, severance payments, & all other types of income considered remuneration for labor services by the IRS.”
## Example W2

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Wages, tips, other compensation</td>
</tr>
<tr>
<td>2</td>
<td>Federal income tax withheld</td>
</tr>
<tr>
<td>3</td>
<td>Social security wages</td>
</tr>
<tr>
<td>4</td>
<td>Social security tax withheld</td>
</tr>
<tr>
<td>5</td>
<td>Medicare wages and tips</td>
</tr>
<tr>
<td>6</td>
<td>Medicare tax withheld</td>
</tr>
<tr>
<td>7</td>
<td>Social security tips</td>
</tr>
<tr>
<td>8</td>
<td>Allocated tips</td>
</tr>
<tr>
<td>9</td>
<td>Control number</td>
</tr>
<tr>
<td>10</td>
<td>Dependent care benefits</td>
</tr>
<tr>
<td>11</td>
<td>Nonqualified plans</td>
</tr>
<tr>
<td>12a</td>
<td>See instructions for box 12</td>
</tr>
<tr>
<td>12b</td>
<td></td>
</tr>
<tr>
<td>12c</td>
<td></td>
</tr>
<tr>
<td>12d</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Statutory employee retirement plan, Third-party sick pay</td>
</tr>
<tr>
<td>14</td>
<td>Other</td>
</tr>
<tr>
<td>15</td>
<td>State, Employee’s state ID number</td>
</tr>
<tr>
<td>16</td>
<td>State wages, tips, etc.</td>
</tr>
<tr>
<td>17</td>
<td>State income tax</td>
</tr>
<tr>
<td>18</td>
<td>Local wages, tips, etc.</td>
</tr>
<tr>
<td>19</td>
<td>Local income tax</td>
</tr>
<tr>
<td>20</td>
<td>Locality name</td>
</tr>
</tbody>
</table>

**Employee’s social security number:** 000-00-0000
**Employer identification number (EIN):** 999-99-9999
**Wages, tips, other compensation:** 1,000,000

**W-2 Wage and Tax Statement:** 2014

---

*Copy A For Social Security Administration — Send this entire page with Form W-3 to the Social Security Administration; photocopies are not acceptable.*
What is an **EIN** (Employer Identification Number)?

Any firm with an employee (so issues a W-2) must have an EIN.

Bureau of Labor Statistics uses the EIN as its definition of a firm.

Many organizations have one (e.g. Facebook, Walmart Stores).

Others have many, e.g.

- Stanford has 4 EINs (1 for the university, 1 for each hospital and 1 for the bookstore).
- The 6165 public companies in D&B have 19,969 EINs.
Individual earnings percentiles
(10%=$10k, 50%=$40k, 90%=$100k, 99%=$350k)
Firm size percentiles: unweighted & emp weighted

Unweighted, Median = 3 employees

Employee weighted, Median = 1000 employees
Census data is Longitudinal Business Database

Contains all establishments from 1976 to 2015

Census groups into firms based on ownership and control

Industry defined by largest employment across establishments

Earnings data from the IRS, so similar W2 definition as SSA
Outline

1) SSA Data

2) SSA Results

3) Census Results

4) Implications for inequality
LFWP measured by yearly regression coefficient of log(earnings) on log(size) - falls by ≈50% since 1970s
Increase in earnings from 100 to 10,000 person firm

- 1980-1984: 47%
- 1985-1989: 44%
- 1990-1994: 41%
- 1995-1999: 31%
- 2000-2004: 27%
- 2005-2009: 22%
- 2010-2013: 20%
Analysis with the Abowd, Kramarz and Margolis (1999) and Card, Henning and Kline (2013) Model

Statistical Model for Individual Log Annual Earnings

\[ y_{ijt} = \alpha_i + \psi_j + X_{ijt} \beta + \varepsilon_{ijt} \]

- Fixed worker component \( \alpha \) (e.g. education, innate ability, etc.)
- Fixed firm component \( \Psi \) (e.g. rent sharing, efficiency wages, etc.)
- Time varying worker characteristics \( X \) (here age and age squared)

Estimate Separately in 7-Year Intervals from 1980 to 2013

- 1980-1986 (first): 5.2m firms, 65m workers, 332m worker years
- 2007-2013 (last): 5.2m firms, 81m workers, 414m worker-year

Details in Song et al. (forthcoming 2019)
AKM regression equation: \[ y_{it} = \alpha_i + \psi_j(i,t) + x_{it}'\beta + r_{it} \]
AKM regression equation: \( y_{it} = \alpha_i + \psi_{j(i,t)} + x'_{it} \beta + r_{it} \)
Almost 90% of the decline in the large firm wage premium comes from the fall in the firm effect.

Notes: Firm size groups: 1=1-10, 2=10-50, 3=50-250, 4=250-1K, 5=1-2.5K, 6=2.5-10K, 7=10-15K, 8=15K+. Age/year effects and the residual term are omitted.
Fall in LFWP more for lower end workers: earnings

Figure 3. Relationship Between Firm Size and AKM Match Component by WFE Quartile
Fall in LFWP more for lower end workers: **education**

**Figure:** LFWP for Two Education Groups, Relative to Firm Size 100 or Less

Outline

1) SSA Data

2) SSA result

3) Census results

4) Implications for inequality
Census data similar 44% drop in LFWP

Notes: Obtained from firm-level data in the US Census Longitudinal Business Database. Results from annual employment-weighted regressions. The Y-axis represents the increase in log(firm mean wage) associated with a given increase in log(firm employment). Regressions run for each year with firm-employment weights.
Composition accounts for $\approx \frac{1}{4}$ of the LFWP drop

Notes: Obtained from firm-level data in the US Census Longitudinal Business Database. The solid line shows estimated from annual employment-weighted regressions. The dotted line shows estimates from regressions weighted by employment* (industry employment in 1976/industry employment in year t).
Composition change reflects mainly the shrinkage of manufacturing, which has a high LFWP.

Notes: Obtained from firm-level data in the US Census Longitudinal Business Database. Results from employment-weighed regressions using data from 1980. LFWP defined as regression coefficient of log(mean wages) on log(firm employment) by industry and year.
Industry characteristic changes also matter: low pay industries (e.g. retail & admin) now have larger firms.
Controlling for industry composition and characteristics accounts $\approx \frac{2}{3}$ of the fall in LF WP.

Notes: From US Census Longitudinal Business Database. The black line shows the benchmark regression estimates from annual employment-weighted regressions. The red line adds industry fixed effects. The Y-axis represents the coefficient from regressing log(average wage) on log(firm employment) by year.
Remaining 1/3 due to drops of LFWP within individual industries – in particular Retail

Notes: From US Census Longitudinal Business Database. Results from regressions of log(average wages) on log(firm employment) by industry and year. Industry line thickness scaled to average employment share from 1978 to 2015.
Summary

1. LFWP falls by about 50%, mainly due to falling large firm AKM Fixed-Effect (not due to less worker sorting)

2. Appears to particularly impact lower paid/educated workers

3. Associated with two industry factors in particular:
   - Shrinking manufacturing (which has a high LFWP)
   - Expansion of low paying service sector (e.g. big box retail)
Outline

1) SSA Data

2) SSA result

3) Census results

4) Implications for inequality
Two offsetting impacts on inequality

(1) Large firms pay more, so reducing the LFWP reduces between firm inequality

(2) LFWP falling faster for lower-end workers, increasing within firm inequality
Firms 100 ≤ employees < 1k, percentiles since 1981

Source: “Firming up inequality” (2019), Song, Price, Guvenen, Bloom and von Wachter
Firms 10k ≤ employees, percentiles since 1981

Source: “Firming up inequality” (2019), Song, Price, Guvenen, Bloom and von Wachter
Find that the net impact of decline in the LFWP probably relatively small

<table>
<thead>
<tr>
<th>Panel A: Between-/Within-Firm Size Class Variance Decomposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval 1</td>
</tr>
<tr>
<td>0.791</td>
</tr>
<tr>
<td>Between Variance</td>
</tr>
<tr>
<td>Within Variance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: AKM Components of Between-Firm Size Class Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var Worker Effect</td>
</tr>
<tr>
<td>Var Firm Effect</td>
</tr>
<tr>
<td>Cov Worker-Firm Effect</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: AKM Components of Within-Firm Size Class Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var Worker Effect</td>
</tr>
<tr>
<td>Var Firm Effect</td>
</tr>
<tr>
<td>Cov Worker-Firm Effect</td>
</tr>
<tr>
<td>N (millions)</td>
</tr>
</tbody>
</table>

Notes: Firms are groups into 5 classes based on the size of their workforce: 1 to 20, 21 to 100, 101 to 1000, 1001 to 10000, and over 10000.

The fall in LFWP reduces between firm size class inequality

This fall in between firm size class inequality from the falling large firm AKM firm-effect

Within firm size class inequality rising from more sorting & segregation (Song et al. 2019)
Conclusions

1. US large-firm wage premium (LFWP) has been falling for over 30 years, and now about half of its value in 1980

2. Appears to due to large firms are cutting their pay premium

3. Fall in LFWP particularly for lower education employees

4. Declining manufacturing and expanding services (e.g. big-box retail) appears to account for much of this
Back Up
The disappearing large-firm wage premium seems to come from a falling large-firm AKM fixed-effect—in words, large-firms no longer pay “extra.”

AKM Component Decomp of Corr(Log Earnings, Log Firm Size) by Interval

Correlation Coefficient


Interval

AKM Component
log(y)
WFE
FFE
XB
R
In numbers, almost 90% of the drop is from the decline in the firm effect

<table>
<thead>
<tr>
<th></th>
<th>Log Earnings (1)</th>
<th>Worker Effect (2)</th>
<th>Firm Effect (3)</th>
<th>Age Effect (4)</th>
<th>AKM Residual (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interval 1: 1980-86</td>
<td>0.080</td>
<td>0.016</td>
<td>0.057</td>
<td>0.007</td>
<td>0.001</td>
</tr>
<tr>
<td>Interval 5: 2007-13</td>
<td>0.039</td>
<td>0.019</td>
<td>0.021</td>
<td>-0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>Change</td>
<td>-0.041</td>
<td>0.003</td>
<td>-0.036</td>
<td>-0.008</td>
<td>0.000</td>
</tr>
<tr>
<td>Share (Percent)</td>
<td>-</td>
<td>(-7.5)</td>
<td>(86.8)</td>
<td>(20.2)</td>
<td>(0.5)</td>
</tr>
</tbody>
</table>

Table 1—Change in LFWP Regression Coefficients by AKM Components
The firm size and earnings correlation: 1978-2013