Earnings Inequality and the Role of the Firm

Kevin L McKinney and John M Abowd
U.S. Census Bureau
January 6, 2018
AEA/LERA Annual Meeting
Acknowledgements and Disclaimer

- This research was conducted while the authors were supported by the U.S. Census Bureau. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the Census Bureau or any other research sponsor. All results have been reviewed to ensure that no confidential information is disclosed.

- This research uses data from the Census Bureau’s Longitudinal Employer-Household Dynamics Program, which was partially supported by NSF Grants SES-9978093, SES-0339191, and ITR-0427889; NIA Grant AG018854; and grants from the Sloan Foundation.
Outline

- Background
- Data (Designed vs Found)
- Sample Selection (Eligible Workers)
- Earnings Inequality Trends
- Job Earnings Model Estimation
- Job Earnings Decomposition
- Worker Earnings Mobility Results
- Conclusion
Background

- Increasing earnings inequality since 2000
- What explains the large difference in earnings across workers?
  - Portable worker skill and experience?
  - Where you work?
- Both are important and together explain about 45% of the total variation in earnings across jobs
Data

- U.S. Census Bureau’s Longitudinal Employer Household Dynamics (LEHD) linked employer employee data
- Analysis Variable: Real annual earnings at all jobs
- Available Period: 1990-2013
- Analysis Period: 2004-2013
- Data for all states, DC, and federal workers are available beginning in 2004
- Covers the period before, during, and after the great recession
LEHD Data is “Found”

- LEHD data is not designed to be a reliable national worker frame
- A job in LEHD data is the relation between a statutory employer and a statutory employee
- A job should appear in LEHD data if the firm is covered by the state Unemployment Insurance system, except:
  - Not all firms are covered (about 90% of NIPA W&S data)
  - State entry occurs sporadically over 15 years
  - Earnings are filed using inconsistent/incorrect person identifiers
- For the purpose of measuring individual earnings inequality, jobs must be assigned to a worker
- We create a reliable national worker frame by using only jobs associated with an “eligible worker”
What are Eligible Workers?

- We use the SSA Numident (list of officially issued SSN’s) to create a consistent frame of persons eligible to work every year from 2004-2013
  - Age 18-70, SSN issued, no death report
- Combine the annual list of eligible workers with the same year LEHD jobs data to determine active status
  - Include earnings from all jobs during the year if fewer than 12 jobs are reported, zero otherwise
- Workers (“immigrant candidates”) on the LEHD jobs data that do not match to the SSA Numident or matches with more than 12 jobs per year are removed
1990-1994: 19% of QCEW Employment (11 states)


1998-2003: 86% of QCEW Employment (38 states)

2004-2013: 100% of QCEW Employment (50 states+DC+OPM)
Percent of All Worker Jobs Associated with Immigrant Candidates By Year

- **1990-1994:** 19% of QCEW Employment (11 states)
- **1995-1997:** 68% of QCEW Employment (25 states) Entry of CA, FL, NY, and TX
- **1998-2003:** 86% of QCEW Employment (38 states)
- **2004-2013:** 100% of QCEW Employment (50 states+DC+OPM)
Comparison of Earnings Inequality Trends

- Statistics for the Eligible Workers and the All Workers Samples
  - Ratio of the 99th and the 1st percentiles
  - Ratio of the 95th and the 5th percentiles
  - Ratio of the 90th and the 10th percentiles
  - Ratio of the 80th and the 20th percentiles
  - Variance of Log Annual Earnings
Selected Inequality Measures: Eligible Workers Relative to 2000

- r_99_01
- r_95_05
- r_90_10
- r_80_20
- VarLnEarn

United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov
Selected Inequality Measures: All Workers Relative to 2000

- $r_{99\_01}$
- $r_{95\_05}$
- $r_{90\_10}$
- $r_{80\_20}$
- $\text{VarLnEarn}$
Earnings Decomposition

- Estimate a fixed person ($\theta$) and fixed firm ($\psi$) effects earnings model.
- Dependent variable ($y$): log real (2000 CPI) annual earnings at all eligible jobs
- Covariates ($x$): constant, demographic characteristics interacted with actual labor force experience, labor force attachment variables, and aggregate labor market conditions variables
Model Estimation

\[ \ln y_{ijt} = \alpha + x_{it}\beta + \theta_i + \psi_j + \varepsilon_{ijt} \]

- Observations (person firm year): \(~2\) billion
- Persons \((i)\): \(~201\) million
- Firms \((j)\): \(~14.6\) million
- Jobs \((i \times j)\): \(~826\) million
- Years \((t)\): 2004 ... 2013
Job Level Results

- Model explains about 85% of the job-year variation in log earnings
- Decompose each log job-year earnings observation into the following components
  - Worker skill: $h_{it} = \alpha + x_{it}^{\text{exper}} \beta^{\text{exper}} + \theta_i$
  - Labor force attachment: $x_{it}^{\text{lfa}} \beta^{\text{lfa}}$
  - Psi: $\psi_j$
  - Other: $x_{it}^{\text{other}} \beta^{\text{other}}$
  - Residual: $\varepsilon_{it}$
Model Variance Components (scaled to sum to 1)

- lfa, 0.4832
- psi, 0.1301
- h, psi, 0.0308
- h, lfa, 0.1097
- other diag, 0.0002
- other nodiag, -0.0033
- psi, lfa, 0.1182
- h, lfa, 0.1311
Job Level Results (continued)

- Worker skill \((h_{it})\) and the firm \((\psi_j)\) main effects each explain about 13% of log job earnings variance.

- Worker skill and firm main effects have a positive covariance component (3%).

- Both the worker skill and the firm components have substantial positive covariance with labor force attachment (11% and 12% respectively).

- Labor force attachment is the dominant component (about 48%).
Jobs to Workers

- The job level estimation results are used to decomposes earnings into a person specific portable component, a firm level component, and a residual.
- The goal of this paper is to explore how the person and firm specific components vary by annual worker earnings.
- However, first we need to aggregate the components across jobs for workers with multiple employers during the year.
Creating Worker-Year Earnings Components

- Worker-Year Earnings: Sum the dollar value of earnings across all eligible jobs for each worker-year
- Worker Skill: Log worker skill is the same for all jobs within a worker-year
  - Convert each job skill component to dollars, sum, and then take the log of the sum
- Log firm component varies for each job within a worker-year
  - Estimate the dollar value of the firm and non-firm component of each job $y_{ijt}^{firm} = y_{ijt} - \exp(\ln y_{ijt} - \psi_j)$
  - Sum dollar value firm and non-firm components across jobs
  - Recover the all jobs log firm component by taking the difference between all jobs log earnings and the all jobs log non-firm component
Binning the Earnings Components

- Place each eligible worker-year observation for each measure (annual real earnings ($y_{it}$), worker skill ($h_{it}$), and firm ($\psi_{it}$)) in one of three bins
  - Bin 2: Bottom 20%, Bin 3: Middle 60%, Bin 4: Top 20%
  - Bin boundary values estimated separately for each measure using log values and all observations
- Bin 1 is reserved for eligible workers with no observed earnings in a particular year
  - Eligible workers have a valid SSN, are between the ages of 18-70, SSN issued, and not reported dead
Year-to-Year Earnings Mobility

- **Within:** Earnings change, but the change is such that the earnings bin in the previous and the current year are the same
- **Between:** Earnings change, but the change is such that the earnings bin in the previous year differs from the earnings bin in the current year
- Worker must be employed in the previous year
  - Patterns 1_1, 1_2, 1_3, and 1_4 are excluded
  - 12 possible earnings/inactivity mobility patterns
Earnings Mobility by Previous Year Earnings Bin

1 - Inactive
2 - Bottom 20%
3 - Middle 60%
4 - Top 20%
Putting Everything Together

- The next three charts combine the worker year-to-year earnings mobility results with the worker level earnings decomposition estimates.
- To reduce clutter we show results only for the largest earnings mobility flows (representing 90% of workers).
- Each bubble represents a specific worker, firm, and earnings mobility pattern.
- Previous year earnings is on the horizontal axis and current year earnings is on the vertical axis.
- Results are the average of nine year-to-year earnings mobility pairs (2004-2013).
Top 20%: Average Earnings by Worker, Firm, and Flow Type

Average Current Year Earnings vs. Average Previous Year Earnings

- Top Top 4_4, $143,296
- Top Middle 4_4, $92,507
- Top Bottom 4_4, $80,717
- Middle Top 4_4, $67,457
- Middle Middle 4_4, $61,625
- Middle Bottom 4_4, $56,789

United States Census Bureau
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov
Middle 60%: Earnings by Worker, Firm, and Flow Type

Average Current Year Earnings vs. Average Previous Year Earnings

- Middle Top 3_4, $38,199
- Middle Middle 3_4, $37,532
- Top Middle 3_3, $33,479
- Middle Top 3_3, $31,965
- Top Bottom 3_3, $27,830
- Bottom Top 3_3, $24,984
- Middle Middle 3_3, $24,165
- Bottom Middle 3_3, $17,361
- Middle Bottom 3_3, $15,820
- Middle Bottom 3_2, $15,344
- Middle Middle 3_2, $15,071
- Bottom Bottom 3_3, $12,342
- Bottom Middle 3_2, $11,989
- Middle Bottom 3_2, $10,788
- Bottom Bottom 3_2, $10,000
- Middle Middle 3_1, $10,000

Note: The earnings are listed for various worker, firm, and flow type categories.
Bottom 20%: Earnings by Worker, Firm, and Flow Type

Average Current Year Earnings vs. Average Previous Year Earnings

- Bottom Bottom 2_3, $3,398
- Middle Middle 2_3, $3,706
- Middle Bottom 2_3, $3,540
- Bottom Middle 2_3, $3,726
- Bottom Bottom 2_2, $2,748
- Middle Bottom 2_2, $2,774
- Bottom Bottom 2_2, $2,282
- Middle Middle 2_2, $3,077
- Bottom Middle 2_2, $2,781
- Bottom Bottom 2_1, $1,014
- Middle Bottom 2_1, $1,749
- Bottom Middle 2_1, $1,749
- Middle Middle 2_1, $1,828
- Bottom Bottom 2_1, $1,014
- Middle Middle 2_1, $2,653
Conclusion

- Like all list based frames, administrative data cannot be used without ancillary information to insure the frame is representative of the target population
- Earnings heterogeneity across firms is a substantial component of earnings inequality
  - A top skill worker at a top paying firm earns about $51,000 (55%, top earnings bin) more than a worker in the same skill class at a middle paying firm
  - A middle skill worker at a top paying firm earns about $6,000 (9%, top earnings bin) or $8,000 (24%, middle earnings bin) more than a worker in the same skill class at a middle paying firm
  - A bottom skill worker at a middle paying firm earns about $5,000 (41%, middle earnings bin) or $500 (22%, bottom earnings bin) more than a worker in the same skill class at a bottom paying firm
Conclusion (continued)

- Earnings are substantially higher for top skill workers at top paying firms.
- Middle skill workers at top paying firms benefit substantially less (9% vs 55%).
- Low paying firms tend to be concentrated in the “leisure and hospitality” and the “education and health” sectors.
- High paying firms tend to be concentrated in the “manufacturing” and the “prof/bus services” sectors.