

Earnings Inequality Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data

John M. Abowd, Kevin L. McKinney and Nellie Zhao

Allied Social Sciences Associations
American Economic Association
New Measures of Human Capital and Their Application
January 5, 2018 8:00-10:00am Philadelphia, PA

Full text: <http://www.journals.uchicago.edu/doi/full/10.1086/694104>

Acknowledgements and Disclaimer

- This research was conducted while all 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 (LEHD) Program, which was partially supported by NSF Grants SES-9978093, SES-0339191, and ITR-0427889; NIA Grant AG018854; and grants from the Sloan Foundation. Abowd also acknowledges direct support from NSF Grants SES-0339191, CNS-0627680, SES-0922005, TC-1012593, and SES-1131848

Background

- How has earnings inequality changed over time?
- Complete analysis of LEHD jobs (1990-2013) using worker-based, not job-based, annual earnings
- Examination of the relevant frame for individual earnings inequality
- How do administrative data compare to household surveys?
- What are the trends in cross-sectional and dynamic earnings inequality?

Data: Worker-level Analysis

- Frame: all jobs in the LEHD infrastructure 1990-2013
- Real, annual earnings analyzed for four regimes:
 - 1: 1990-2013 [19%]* (AK CO MD ID IL IN KS LA MO WA WI)
 - 2: 1995-2013 [68%] (+ AZ CA FL GA MN NC NY OR MT PA RI SD TX WY)
 - 3: 1998-2013 [86%] (+ CT HI KY MI ND NJ NM NV SC TN VA WV)
 - 4: 2004-2013 [100%] (+ AL AR DC DE IA MA MS NE NH OH OK UT VT)
 - Done to allow for differential entry dates of states into the Local Employment Dynamics federal/state partnership without using a statistical missing data model (as is done in the establishment analysis)

*Percent of Quarterly Census of Employment and Wages (QCEW) month-one employment in 2012:Q1.

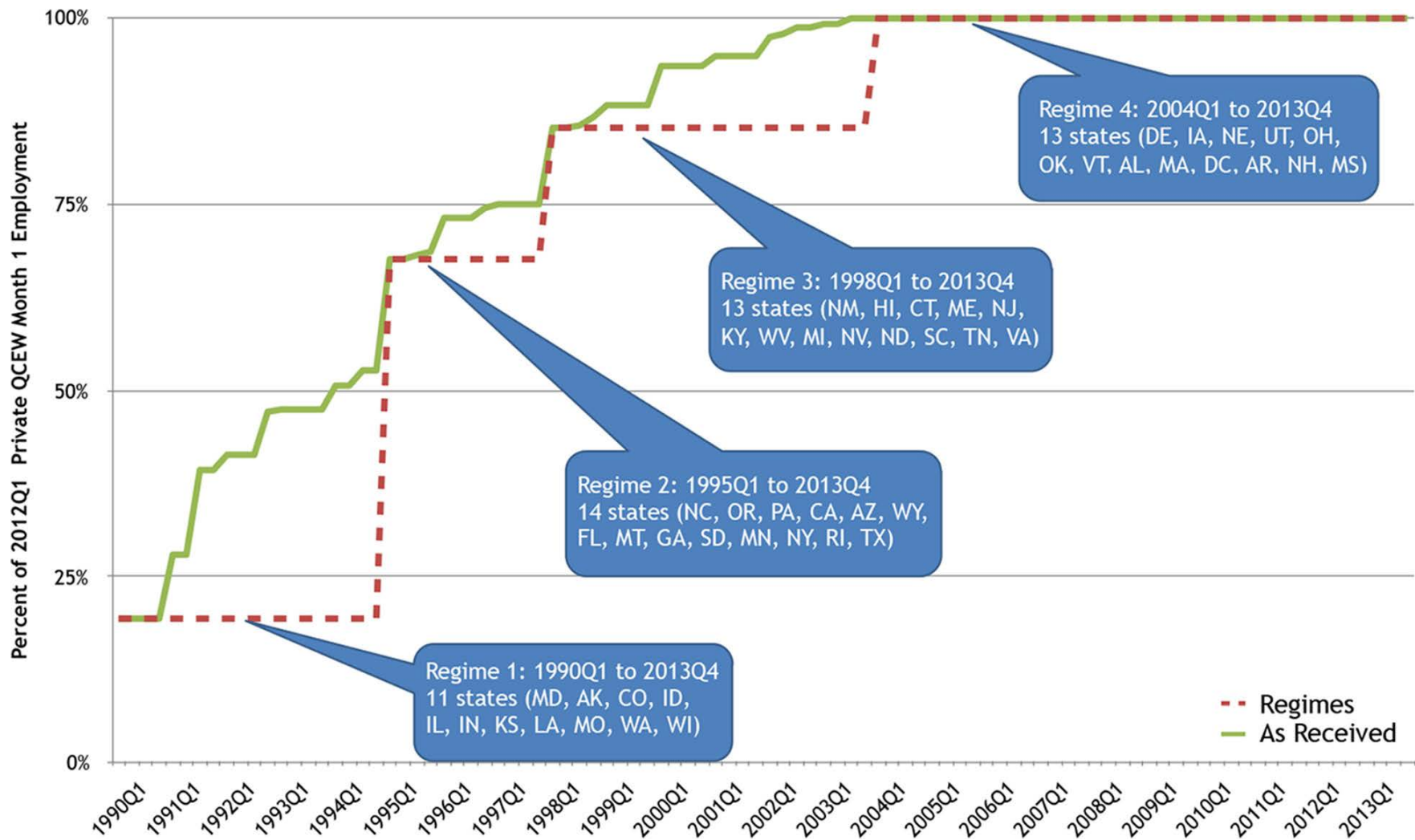


Fig. 1. Longitudinal Employer-Household Dynamics (LEHD) infrastructure data as received and by date regimes. The figure shows the coverage of the LEHD infrastructure data expressed as a percentage of 2012Q1 private Quarterly Census of Employment and Wages (QCEW) employment as received (solid green line) and by date regime (dashed red line). Office of Personnel Management data for federal workers are not shown in this figure but are available beginning in 2000Q1.

Basic Earnings Definitions

- Job earnings: UI-covered wage and salary payments between one statutory employer and one statutory employee
 - Relevant statute is the state Unemployment Insurance system. Covers private firms and state/local government.
 - No self-employment earnings unless the proprietor drew a salary (indistinguishable from other employees in this case)
 - Federal data from the Office of Personnel Management (recast to be comparable to UI data for earnings measurement)
- Worker earnings: Sum of all job earnings for a given person over the time period (year)
- All earnings measures are in year 2000 dollars (CPI-U).

Persons Eligible to Work and Entry/Exit from the Labor Market

- Frame: All persons with a valid SSN (on Census Numident), age 18-70 sometime during 1990-2013
- Year must also be greater than or equal to SSN year of issue and less than or equal to year of death (if available)
- Person is active in the labor market in a given year when positive UI earnings are found
- Removed from analysis any year where worker earnings are based on more than 12 jobs (out of frame)
- Valid SSN, age, and number of jobs restrictions remove about between 6% and 8% of the workers/year

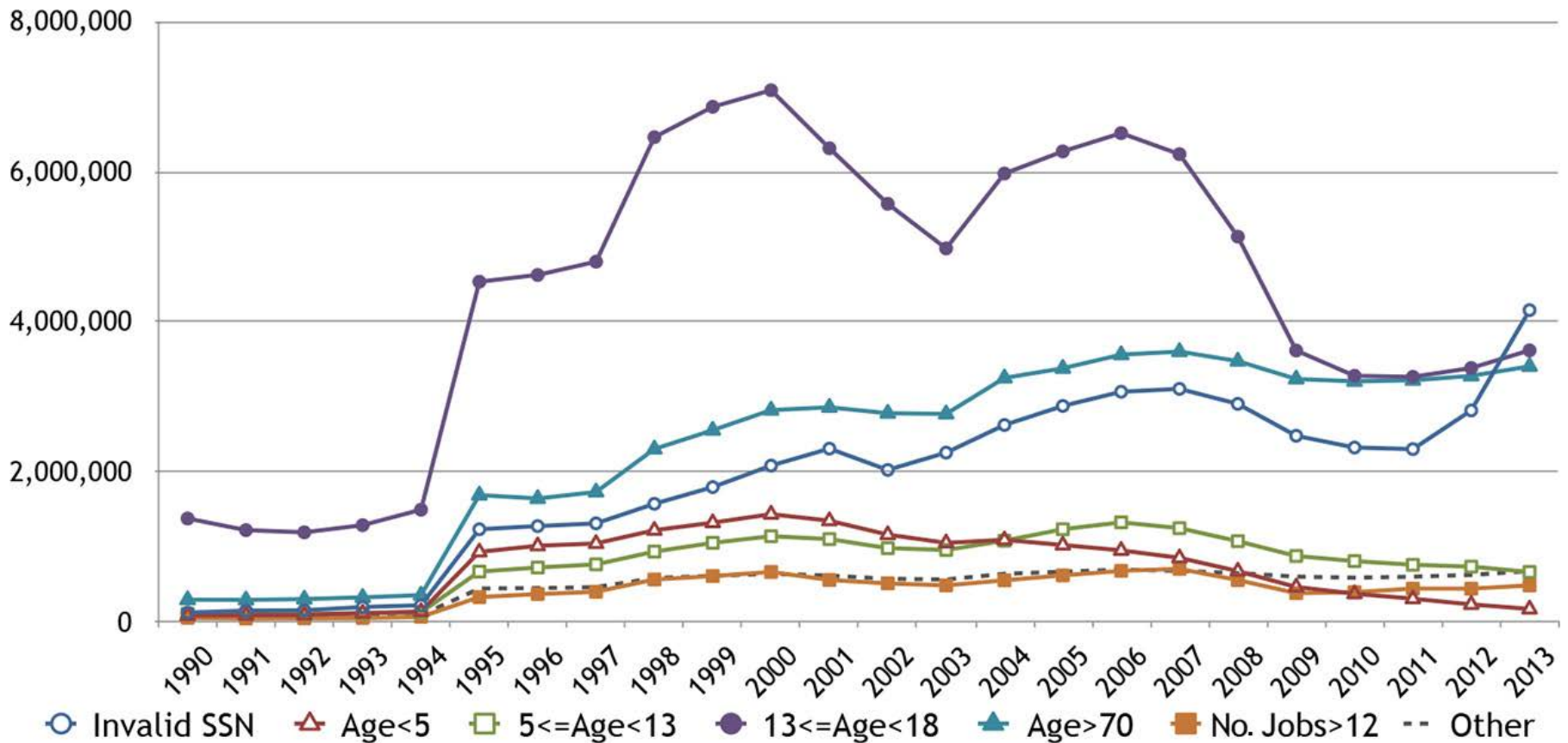


Fig. A1. Immigrant candidates: excluded earnings records. This figure presents the count of earnings records excluded from the eligible-workers frame each year, disaggregated by the different eligibility requirements the record failed to meet: (i) records that are only on the unemployment insurance (“Invalid SSN” [Social Security number]), (ii) records where the SSN is valid but the age of the worker is less than 5 years old (“Age < 5”), (iii) records where the worker is between 5 and 13 years old (“5 ≤ Age < 13”), (iv) records where the worker is between 13 and 18 years old (“13 ≤ Age < 18”), (v) records where the worker is more than 70 years old (“Age > 70”), (vi) records where the worker has more than 12 jobs a year (“No. Jobs > 12”), and (vii) records that fail to meet the other eligibility requirements (“Other”), such as the year being greater than or equal to the SSN year of issue and less than the year of death (when available).

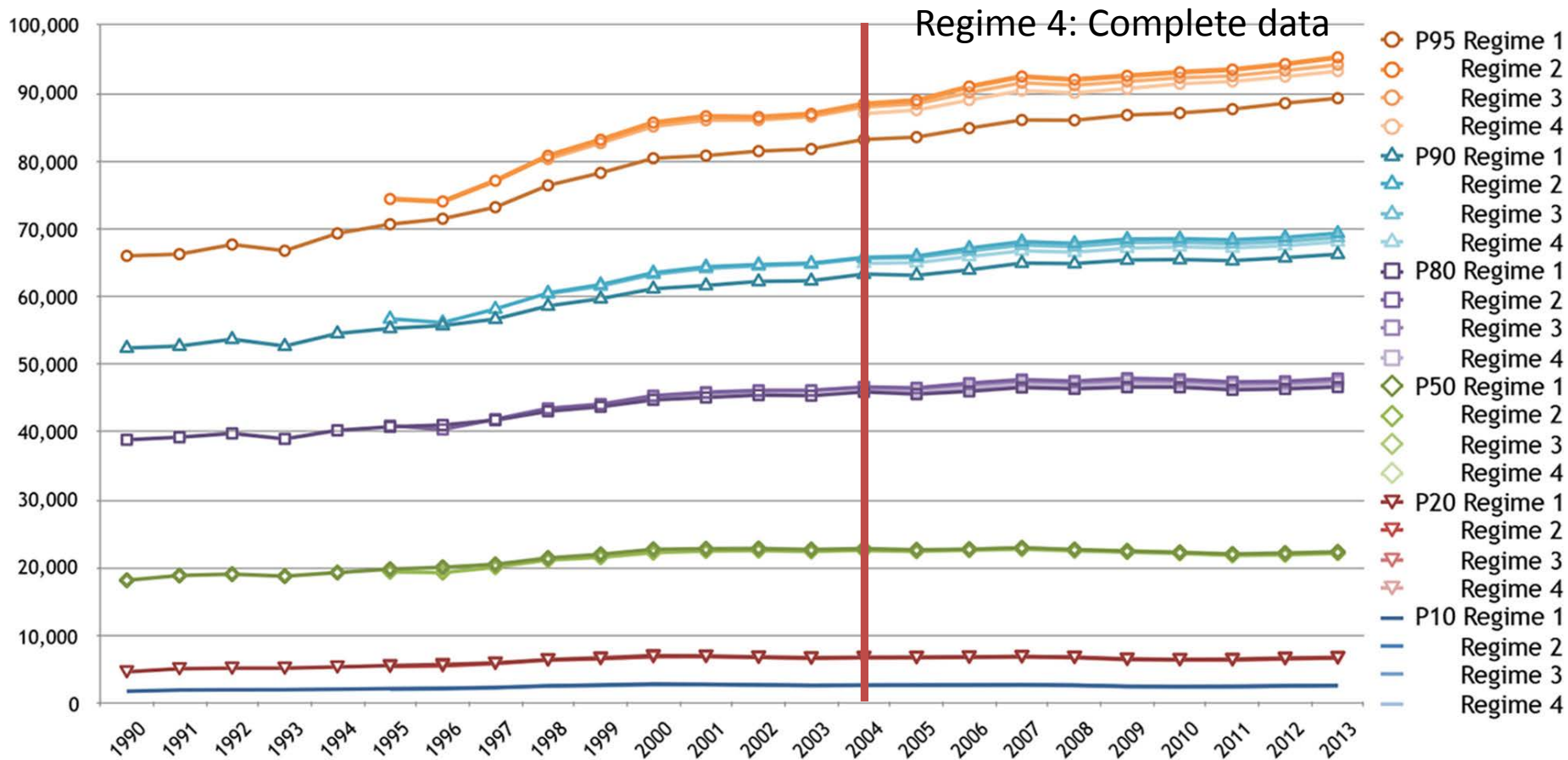


Fig. 2. Percentiles of the earnings distribution for eligible workers by cumulative date regime. The figure plots the 10th, 20th, 50th, 80th, 90th, and 95th percentiles of the earnings distribution of eligible workers by cumulative date regime and year. The cumulative date regime plots the data for all regimes less than or equal to the indicated regime. The figure is not a cumulative distribution, although it looks like one. For example, “P95 Regime 2” indicates the 95th percentile for all states in regimes 1 and 2.

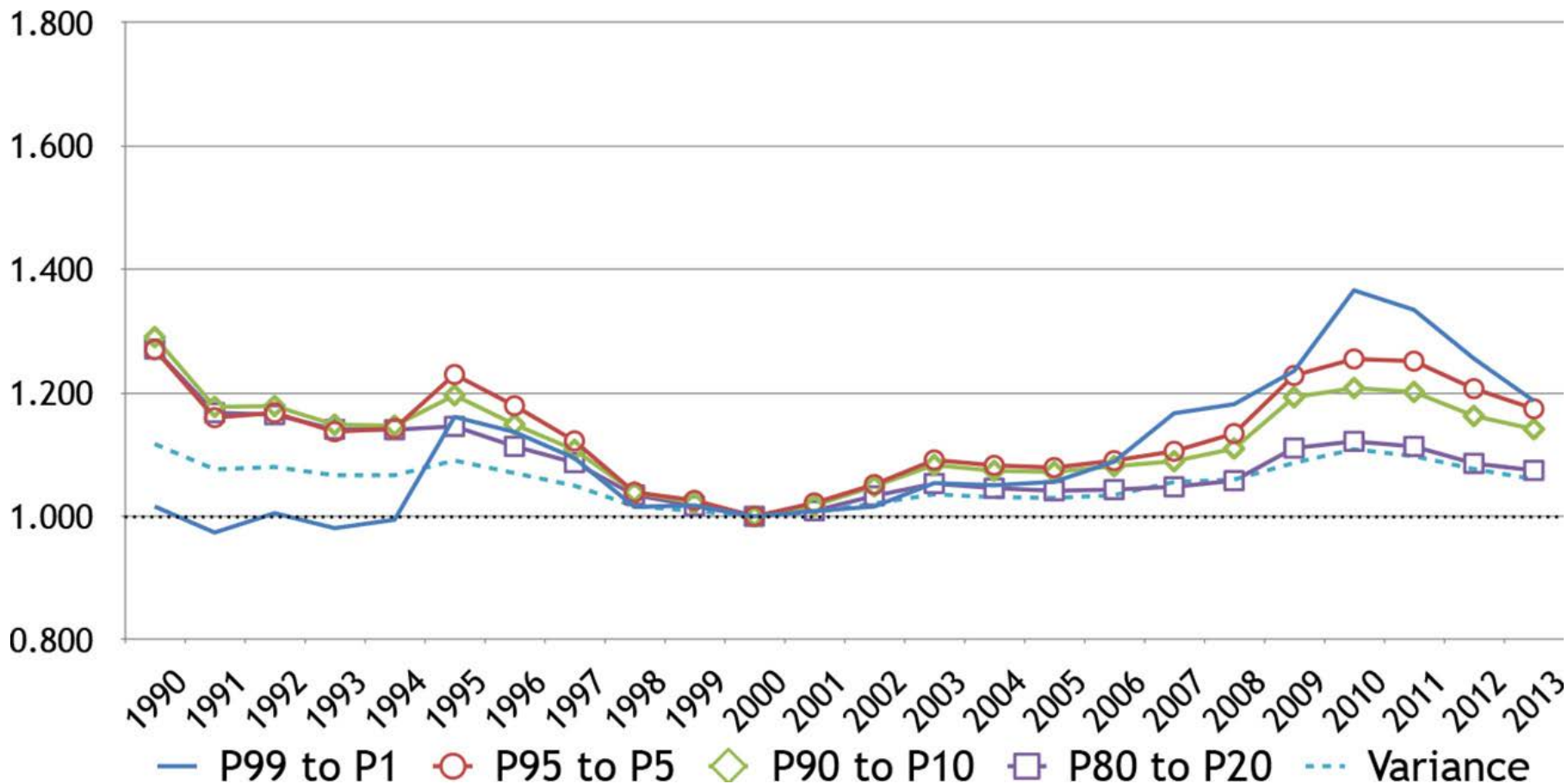


Fig. 3. Selected inequality measures 1990–2013, relative to 2000. The figure presents measures of earnings inequality for eligible workers in all states relative to 2000 from 1990 to 2013. The measures of earnings inequality considered are (i) the ratio of the 99th to the 1st percentile (“P99 to P1”), (ii) the ratio of the 95th to the 5th percentile (“P95 to P5”), (iii) the ratio of the 90th to the 10th percentile (“P90 to P10”), (iv) the ratio of the 80th to the 20th percentile (“P80 to P20”), and (v) the variance of log annual earnings (“Variance”). Results are based on the eligible-workers frame from the Longitudinal Employer-Household Dynamics infrastructure files.

Comparison of LEHD with Linked CPS/ACS

- CPS-ASEC: all persons with positive earnings from 1989-2003
- ACS: all persons with positive earnings from 2000-2013
- CPS/ACS data series are interpolated between 2000-2003
- CPS respondents reported earnings for the previous calendar year
- ACS respondents reported earnings for the previous 12 months relative to the survey date (assigned to the calendar year with the greatest overlap)
- LEHD: Unemployment Insurance (UI) worker earnings (firm-reported administrative data) for the eligible-worker frame

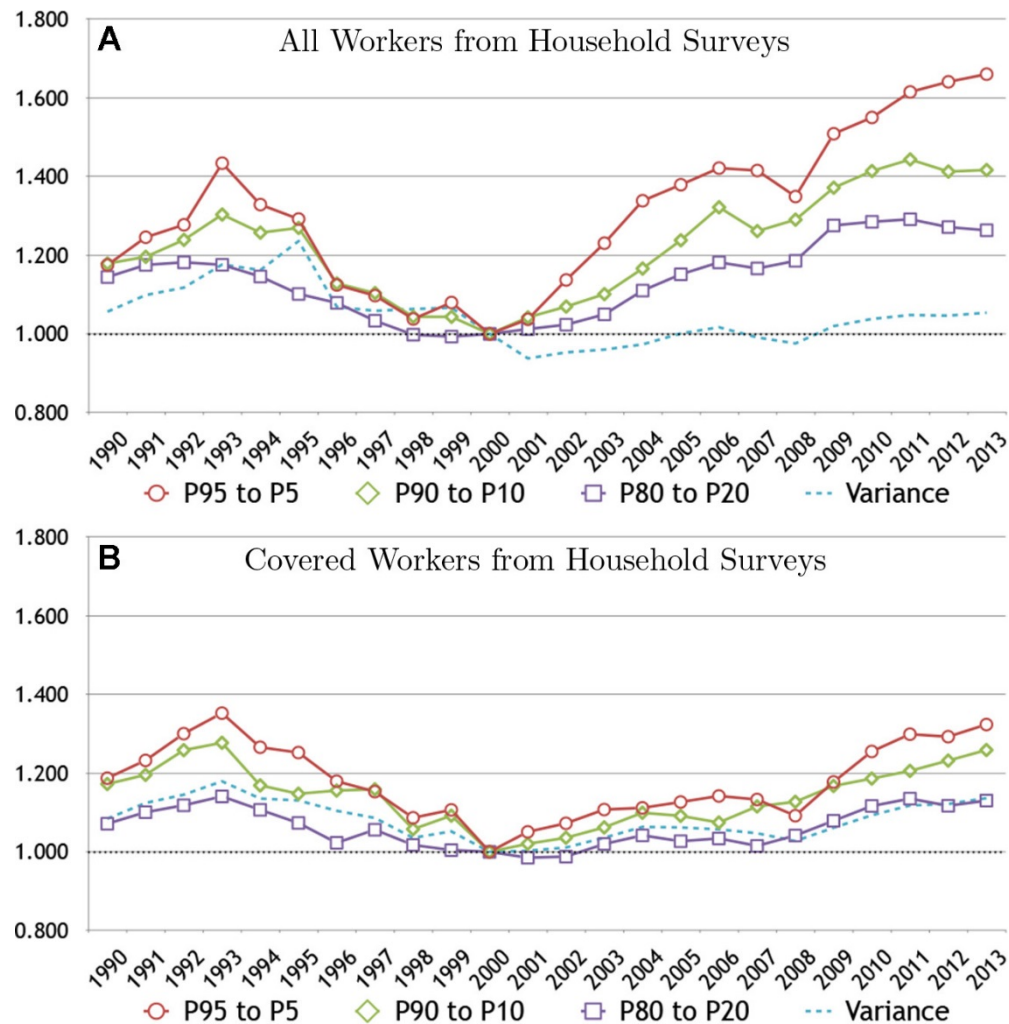


Fig. 5. Selected inequality measures 1990–2013, relative to 2000 (Current Population Survey [CPS]/American Community Survey [ACS]). *A*, All workers from household surveys. *A* presents measures of earnings inequality for all workers in CPS/ACS relative to 2000 from 1990 to 2013. *B*, Covered workers from household surveys. *B* presents measures of earnings inequality for covered workers in CPS/ACS relative to 2000 from 1990 to 2013. The measures of earnings inequality considered are (i) the ratio of the 95th to the 5th percentile (“P95 to P5”), (ii) the ratio of the 90th to the 10th percentile (“P90 to P10”), (iii) the ratio of the 80th to the 20th percentile (“P80 to P20”), and (iv) the variance of log annual earnings (“Variance”).

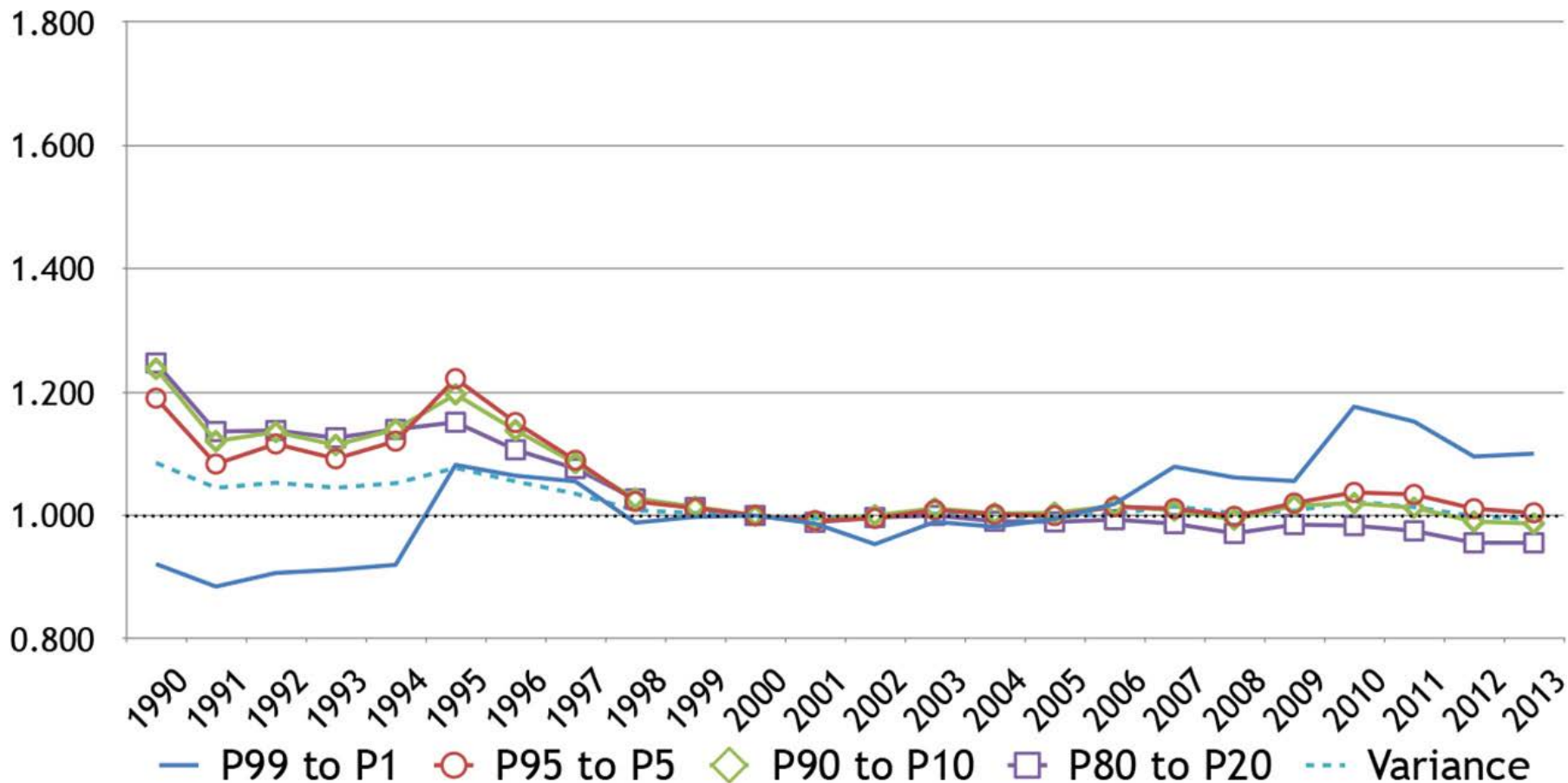


Fig. B3. Selected inequality measures from 1990 to 2013, relative to 2000 (all workers). The measures of earnings inequality considered are (i) the ratio of the 99th to the 1st percentile (“P99 to P1”), (ii) the ratio of the 95th to the 5th percentile (“P95 to P5”), (iii) the ratio of the 90th to the 10th percentile (“P90 to P10”), (iv) the ratio of the 80th to the 20th percentile (“P80 to P20”), and (v) the variance of log annual earnings (“Variance”).

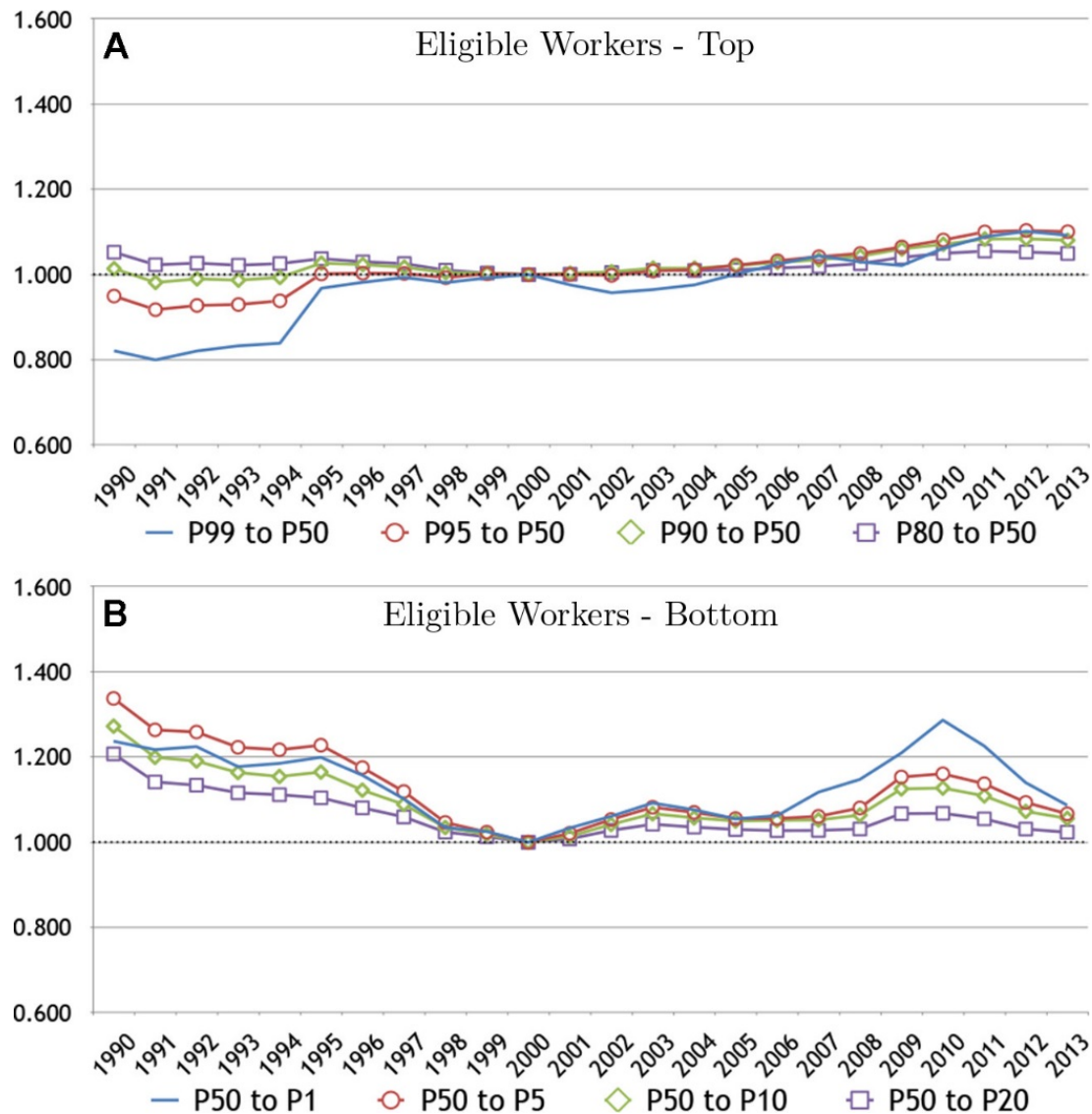


Fig. 4. Selected inequality measures for the top and bottom of the earnings distribution 1990–2013, relative to 2000. *A* and *B* decompose the 99/1 ratio, the 95/5 ratio, the 90/10 ratio, and the 80/20 ratio for eligible workers in all states relative to 2000 from 1990 to 2013 relative to the median. *A* plots the following ratios for the top half of the earnings distribution: (i) the ratio of the 99th to the 50th percentile (“P99 to P50”), (ii) the ratio of the 95th to the 50th percentile (“P95 to P50”), (iii) the ratio of the 90th to the 50th percentile (“P90 to P50”), and (iv) the ratio of the 80th to the 50th percentile (“P80 to P50”). *B* plots the following ratios for the bottom half of the earnings distribution: (i) the ratio of the 50th to the 1st percentile (“P50 to P1”), (ii) the ratio of the 50th to the 5th percentile (“P50 to P5”), (iii) the ratio of the 50th to the 10th percentile (“P50 to P10”), and (iv) the ratio of the 50th to the 20th percentile (“P50 to P20”). The estimates are based on the eligible-workers frame from the Longitudinal Employer-Household Dynamics infrastructure files. See figure B4 for comparable data using the all-worker frame.

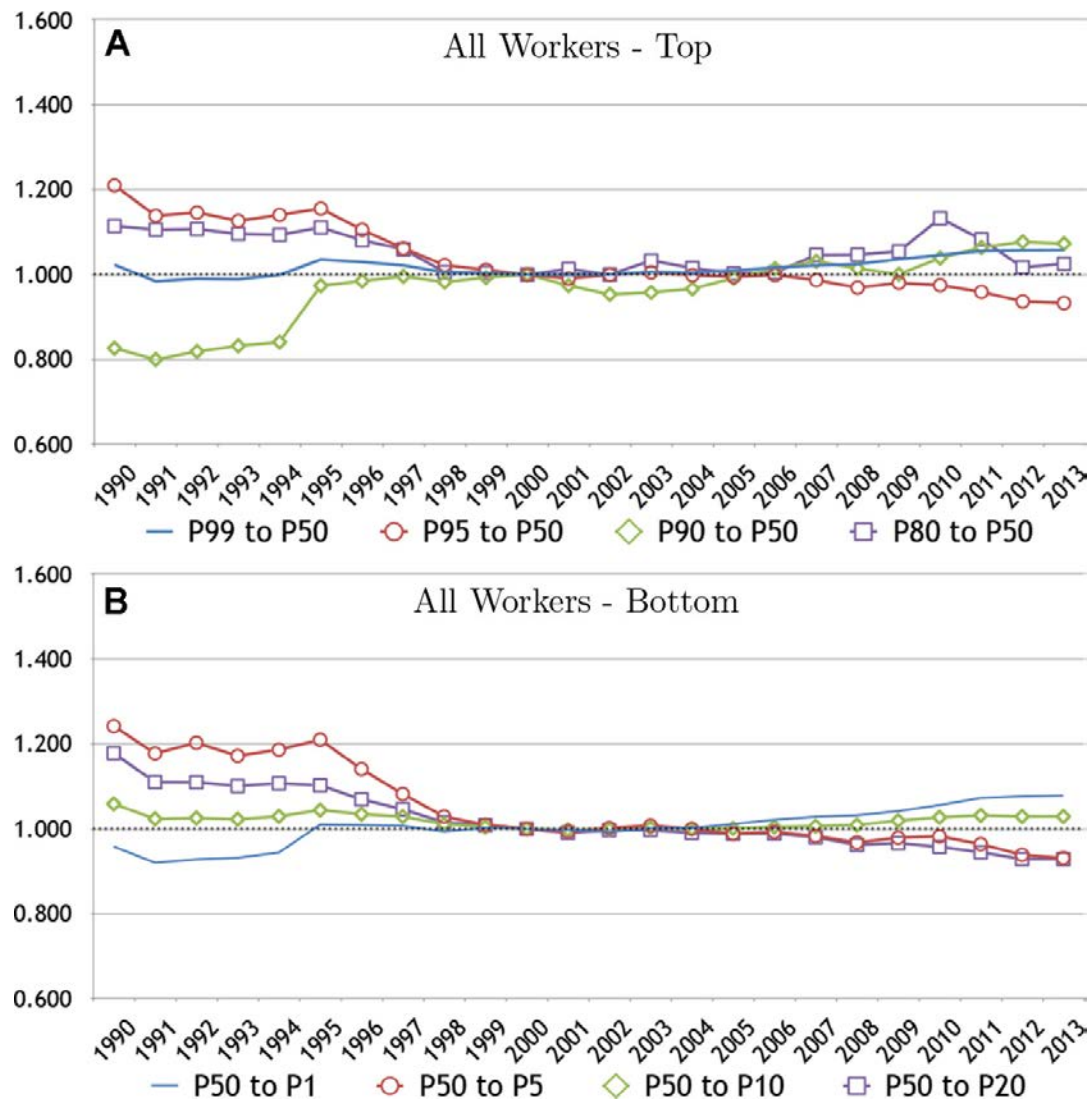


Fig. B4. Selected inequality measures for the top and bottom of the earnings distribution from 1990 to 2013, relative to 2000 (all workers). *A* and *B* decompose the 99/1 ratio, the 95/5 ratio, the 90/10 ratio, and the 80/20 ratio for all workers in all states relative to 2000 from 1990 to 2013 relative to the median. *A* plots the following ratios for the top half of the earnings distribution: (i) the ratio of the 99th to the 50th percentile (“P99 to P50”), (ii) the ratio of the 95th to the 50th percentile (“P95 to P50”), (iii) the ratio of the 90th to the 50th percentile (“P90 to P50”), and (iv) the ratio of the 80th to the 50th percentile (“P80 to P50”). *B* plots the following ratios for the bottom half of the earnings distribution: (i) the ratio of the 50th to the 1st percentile (“P50 to P1”), (ii) the ratio of the 50th to the 5th percentile (“P50 to P5”), (iii) the ratio of the 50th to the 10th percentile (“P50 to P10”), and (iv) the ratio of the 50th to the 20th percentile (“P50 to P20”). The estimates are based on the all-workers frame from the Longitudinal Employer-Household Dynamics infrastructure files.

Properly Measuring Inequality

- LEHD data contain earnings records for workers both with and without a valid SSN
- Inequality trends differ when using jobs for all workers v. jobs for eligible workers
- Earnings inequality increases after 2000 when using jobs for eligible workers, and in the household survey data
- Except for the 99-1 ratio, earnings inequality is relatively stable after 2000 when using all jobs for all workers, unlike in the household survey data
- We think that the eligible-worker frame is the correct one for studying earning inequality

Analyzing Inequality Dynamics

- It is essential to track long periods of labor force inactivity (zero earnings)
- Especially before, during and after the Great Recession, individuals with long histories of positive annual earnings experienced up to multiple consecutive years of zero earnings
- Much of the dynamics is movement between active (positive earnings) and inactive (zero earnings)



Fig. 9. Flows into and out of inactivity. Estimates are based on the authors' calculations using transitions into and out of the eligible-workers frame used to construct the earnings distributions, including inactive workers, and transitions between the earnings categories.

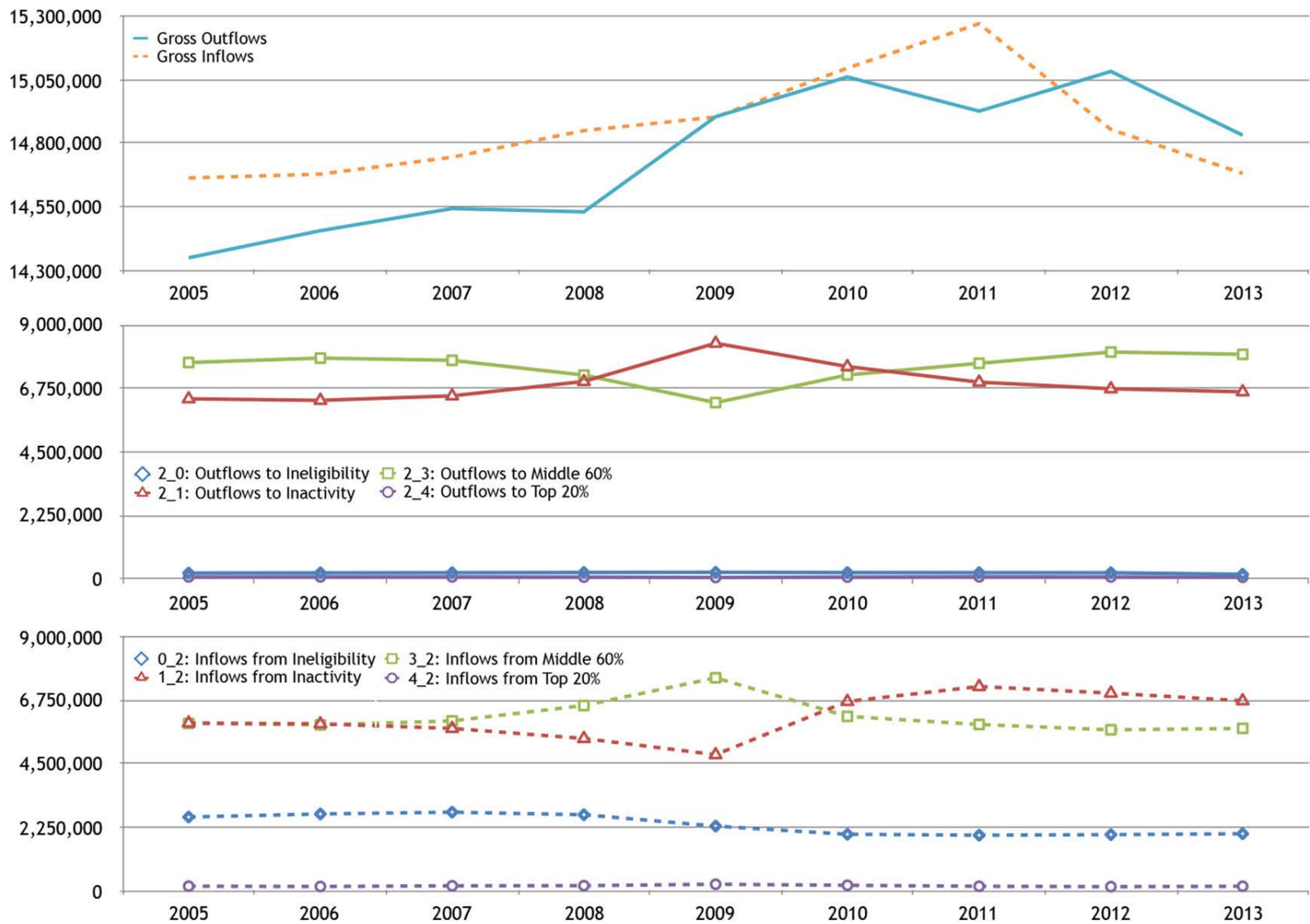


Fig. 10. Flows into and out of the bottom 20% of the earnings distribution. Estimates are based on the authors' calculations using transitions into and out of the eligible-workers frame used to construct the earnings distributions, including inactive workers, and transitions between the earnings categories.

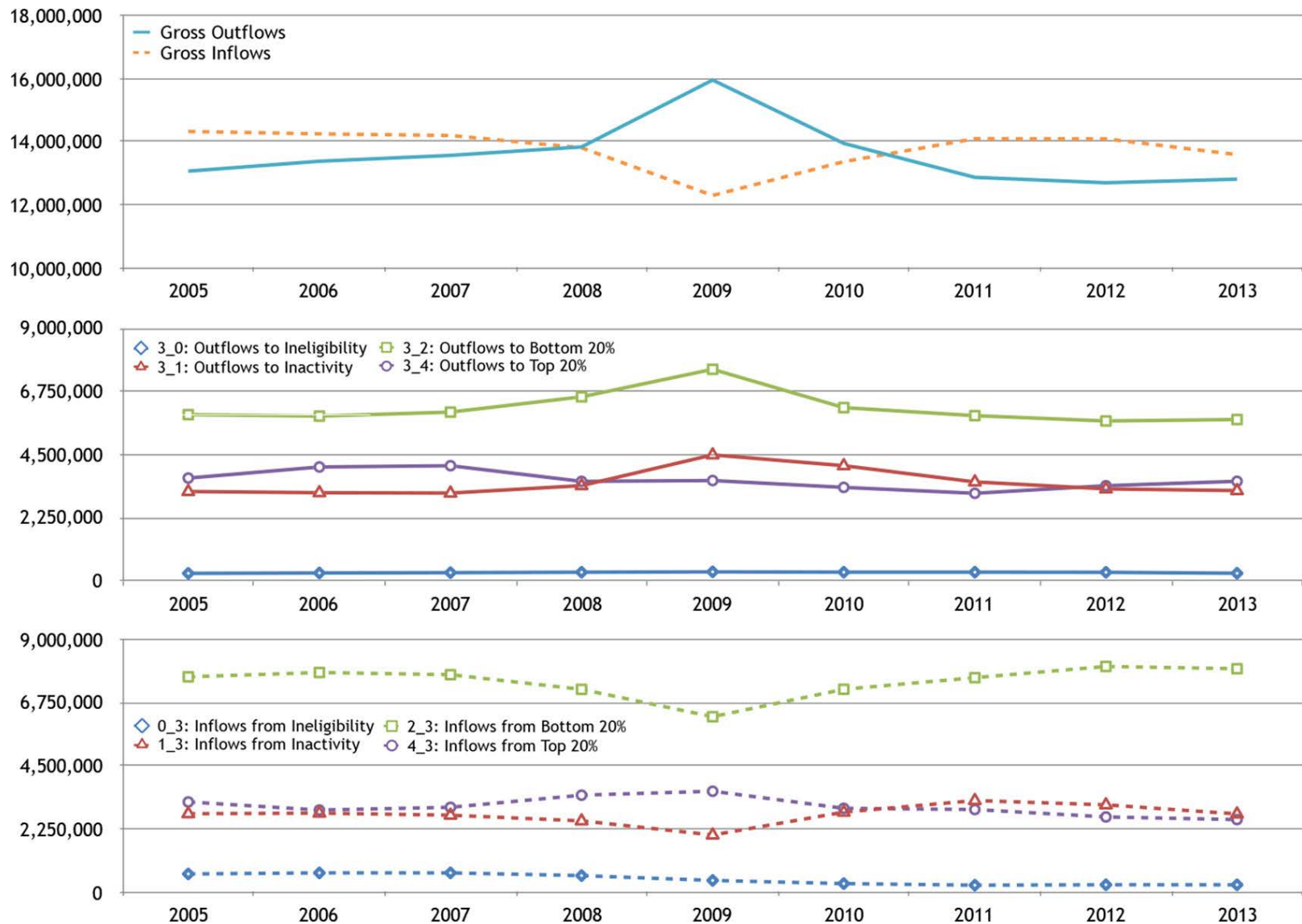


Fig. 11. Flows into and out of the middle 60% of the earnings distribution. Estimates are based on the authors' calculations using transitions into and out of the eligible-workers frame used to construct the earnings distributions, including inactive workers, and transitions between the earnings categories.

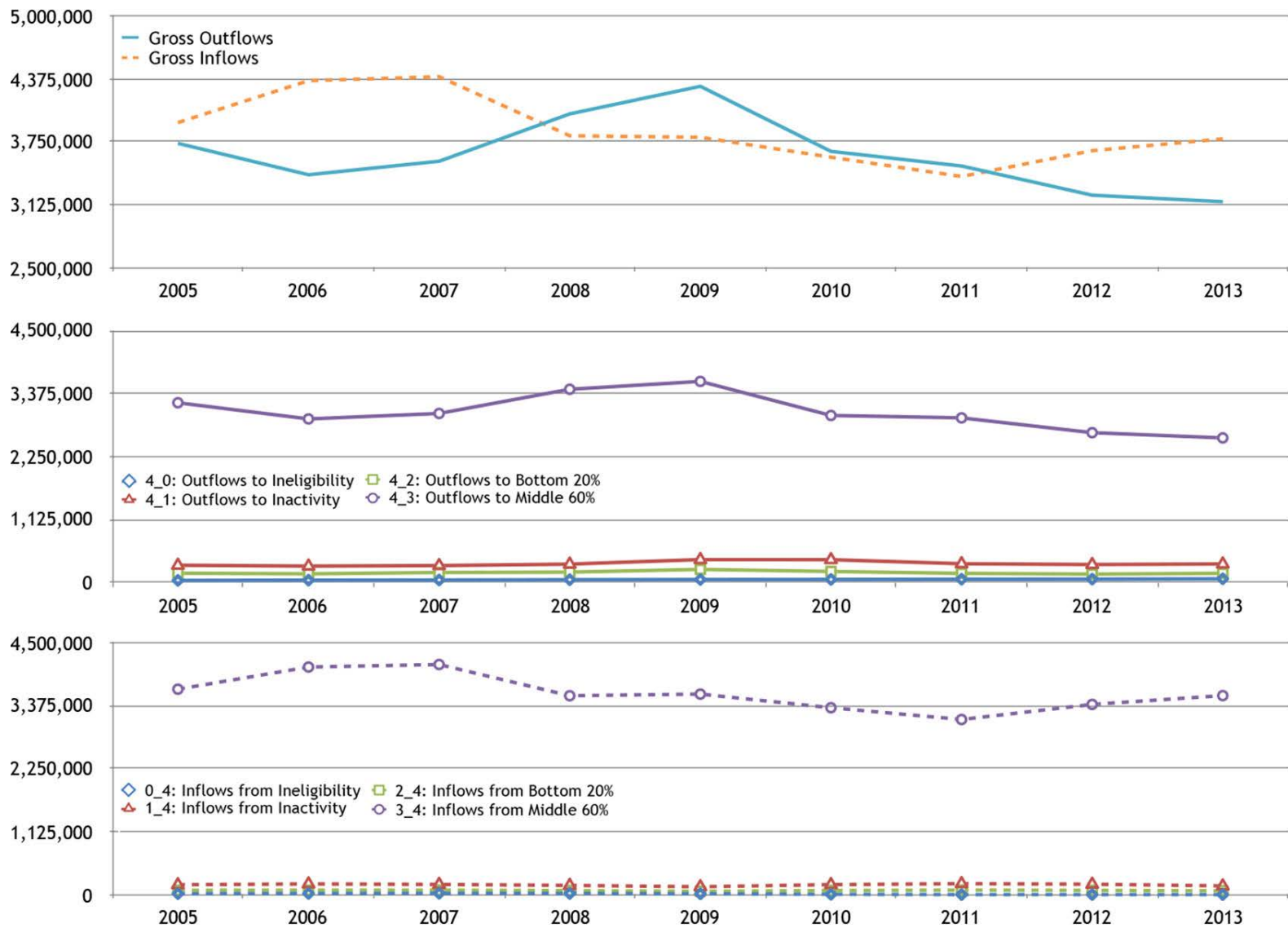


Fig. 12. Flows into and out of the top 20% of the earnings distribution. Estimates are based on the authors' calculations using transitions into and out of the eligible-workers frame used to construct the earnings distributions, including inactive workers, and transitions between the earnings categories.

Conclusions

- State entry year to LEHD does not materially affect the trends in earnings inequality, but LEHD data contain a relatively large number of workers with very low earnings
- Earnings inequality has increased since 2000 in the administrative data using the eligible-workers frame
- Increase in inequality is even greater during the great recession when you include eligible workers with zero earnings