

Evolution of Inequality within Firms

SED Abstract

Jae Song* David Price[†] Fatih Guvenen[‡] Nick Bloom[§]

February 7, 2014

Extended Abstract

Since the 1970s, the US economy has witnessed a dramatic rise in wage inequality. An enormous body of theoretical and empirical research has been conducted over the past two decades in an attempt to understand the causes of this trend. While much has been learned from these analyses, several major questions still remain unanswered. An important set of open questions concerns the link between wage inequality on the worker side to trends in the behavior of the firms and industries that employ these workers. A major difficulty with studying questions of this sort has been the lack of a matched employer-employee dataset in the United States.

This paper uses a unique and confidential dataset—the Master Earnings File (MEF)—from the U.S. Social Security Administration (SSA) records, which makes it feasible to construct such a matched dataset. The MEF contains earnings and employment histories for the entire U.S. population from 1978 to (currently) 2012. It also contains unique Employer Identification Numbers (EINs) for each job held by each worker, which makes it feasible to construct a matched employer-employee dataset containing all US firms and all US workers. The matched dataset contains over 4 billion observations (5+ million firms and 100+ million workers per year for 35 years) and is the first of its kind for the United States. It enables us to characterize heterogeneity within each firm by earnings,

*Social Security Administration

[†]Stanford

[‡]University of Minnesota and NBER

[§]Stanford and NBER

gender, age, immigrant status, and other characteristics, thereby allowing us to shed light on questions precisely at the nexus of firms and their employees.

At this point, the matched dataset has been constructed and we have obtained answers to some of the questions listed below. However, the confidential nature of the data does not allow us to share these findings yet. We will be able to do so once we obtain SSA authorization to disseminate our findings (in early Spring).

The project addresses the following sets of questions:

- Firms vs. Workers: To what extent has the rising wage inequality in the U.S. been driven by increasing inequality between firms, and to what extent by changes that occurred within firms? How has this decomposition varied during recessions, as opposed to during periods of growth? The unique employer-employee matched data allows us to calculate this decomposition precisely over the period being considered.
- Does firm type matter?: The dispersion of wages within firms varies widely. In what ways is it related to firm characteristics such as firm size, firm age, and industry? How has this dispersion, and the characteristics associated with it, changed over time? What characteristics of firms are associated with increased productivity, as measured by change in wage income? How is current compensation related to future wage growth? By matching firm characteristics to wage outcomes, we can understand how they are related. And in addition to yearly snapshots, we can use EINs to track firms over time, understanding how each adjusts to new economic conditions.
- The role of gender and age: How much of the gender wage gap and age premium is explained by differences within and between firms? How have these firm behaviors changed over time, and how are they related to firm characteristics? Because we know the characteristics of each worker within a firm, we can compare the compensation of employees with different characteristics within each firm.
- The impact of recessions: How do different firms adjust their employment practices to the state of the economy? For example, in recessions, do firms trim from the bottom of the wage distribution within the firm? Who is more likely to leave during recessions: men or women, young or old workers, high or low paid workers within

each firm? If firms' total wage bill decreases in recessions, to what extent is this because firms reduced the number of workers they employ, as opposed to lowering average wages? Because our dataset extends back for thirty-five years, from 1978 to 2012, we are able to study how firms have reacted to four different recessions and three full expansions.

- Wage variance and higher moments: How have the higher order moments of the wage distribution changed within and between firms over time? For example, is the economy-wide wage polarization we see based on polarization between firms, or within firms? How is the wage earned by the highest-paid employee related to the wages earned by the lower-income employees in the same firm? Analyzing a data that includes every wage record in every firm allows us to understand firms' wage structure in much more detail than has been done in the past.

This dataset presents an opportunity to look inside every firm in the economy, understanding the details of its compensation structure. This new understanding can help us to understand the decisions they make, and can raise new and interesting questions that are not possible to address with the aggregated data typically studied.