View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Decomposing the Sources of Earnings Inequality

Assessing the Role of Reallocation

October 2006

Fredrik Andersson
Cornell University

Elizabeth E. Davis
University of Minnesota

Matthew L. Freedman
University of Maryland

Julia I. Lane
NORC/University of Chicago

Brian P. McCall
University of Minnesota

L. Kristin Sandusky
U.S. Census Bureau

Thanks to Charlie Brown , David Card, Erica Groshen and the participants of the labor lunch at University of California, Berkeley for their useful comments. The analysis presented here has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grant SES-9978093 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant 5 R01 AG018854-02, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author(s) and do not represent the views of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau is preparing to support external researchers' use of these data; please contact U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA.

Abstract

This paper uses matched employer-employee data from the Longitudinal Employer Household Dynamics database to investigate the contribution of worker and firm reallocation to *within industry* changes in wage inequality between 1992 and 2003. We find that the entry and exit of firms and the sorting of workers and firms based on underlying worker "skills" are important determinants of changes in industry earnings distributions over time. Our results suggest that the underlying dynamics of earnings inequality are complex and are due to factors that cannot be measured in standard cross-sectional data.

1. Introduction

Disentangling the sources of changes in earnings inequality has long been a challenge. The literature has provided both demand and supply side explanations, including, for example, skill-biased technological change, minimum wage adjustments, changes in workforce composition, and declines in unionization. Yet, although wages are determined by the interaction of both firms and workers, most analytical work has been based on the examination of cross-sectional surveys of workers. This means that little is known about the impact on the earnings distribution of changes in the types of firms and the allocation of workers across those firms.

In this paper we use matched administrative data that has longitudinal information on both workers and the firms for which they work to begin to fill this gap in knowledge. Our focus complements earlier worker-based studies that examine changes in *within group* inequality by examining *within industry* inequality.

The analysis in this paper uses the new data to advance knowledge about the sources of changes in earnings inequality in several ways. First, it quantifies the impact of changes in workforce composition, particularly workforce skill and experience, on the earnings distribution by examining the reallocation of workers into and out of the workforce. Second, it tracks the reallocation of workers across jobs to examine the earnings impact of changing firm wage premia, due to changes in unionization, compensating differentials, or rent-sharing. Finally, it examines the impact of firm entry and exit, and the resulting job allocation, on the earnings distribution, providing commensurate insight into the impact of changing production processes. Because our interest is primarily on understanding the impact of changes in firms and the allocation of workers across firms, and rather than on changing industry structure, we examine each industry separately. In addition, the relatively short period of time for which the data are available does not permit us to address the issue of changes in the price of skill.

The paper is structured as follows. After a brief review of the literature and discussion of the data, we present some basic empirical facts about the changes in earnings distributions in each industry sector. We then develop an econometric method for decomposing the sources of change in the earnings distributions when employer-employee matched data are available. The remaining sections of the paper describe the results of performing these decompositions and summarize the implications.

In general, we find that there is no one single "silver bullet" explaining changes in the earnings distribution in each industry. Even when the direction of change is similar across industries, the underlying contributing factors can be very different. Most interestingly, even in industries in which overall inequality is trending in opposite directions, the influence of one set of factors can be consistently in the same direction.

Not surprisingly, given the extensive amounts of worker and firm reallocation that have been documented in the literature, we find that both types of reallocation have large effects on different parts of the earnings distribution. In particular, the entry and exit of firms and sorting of workers and firms based on underlying worker "skills" are important determinants of changes in industry earnings distributions over time. By and large, new firms act to improve the lot of workers at the bottom of the income distribution, but at the same time, existing low-wage firms have expanded their share of employment of low-wage workers. The former result is consistent with the notion that new firms are more productive than old, while the latter is consistent with the fears of policy-makers that there are fewer "high-wage" jobs available to low-wage workers. Although our analysis focuses on the consequences of within-industry reallocation, these results suggest caution when searching for simple answers to questions raised by complex economic phenomena.

2. Background

a) Earnings Inequality

Despite a vast literature that attempts to disentangle the sources of increased inequality that has occurred in recent decades, there is still not complete consensus.[1] A large number of researchers agree that the change was driven by skill biased technical change interacting in complex ways with changes in unionization, management structure, and international trade (see, e.g., Acemoglou 2002). However, some researchers point to changes in the composition of the workforce as an important contributor to earnings inequality changes (Lemieux, 2004). Others point to structural changes. Card and DiNardo (2002) argue that changes in the minimum wages and declines in unionization were the dominant contributors to recent observed trends in inequality. Fortin and Lemieux (1997) also examine the impact of de-regulation in the 1980s, focusing on the transportation, communication, and banking industries.

Recently, Autor, Katz and Kearney (2005) have found that since the late 1980's there has been a divergence in the change in wage inequality between the upper and lower halves of the wage distribution with the lower half, as measured by the 50-10 difference in log wages, either being compressed or not changing and the upper half, as measured by the 90-50 difference, exhibiting a steady rise in inequality. They find that labor force compositional shifts have increased wage inequality with the impact being an offset to countervailing wage compression movements in the lower half of the distribution and a reinforcing effect on residual wage inequality increases in the upper half of the distribution.

Although almost all of the literature is driven by analyses of worker-based surveys, most notably the Current Population Survey (CPS), there is some firm-based evidence that suggests that changes in the distribution of wages may be due in part to changes on the firm side of the labor market. Bernard and Jensen (1998) find that increases in wage inequality across states are highly correlated with shifts in industrial composition, particularly the decline in manufacturing. Burgess, Lane and McKinney (2001) observe sizeable differences in the trends in earnings inequality across industries in Maryland. Other studies have also established the role of firm effects on wages and on wage inequality. Davis and Haltiwanger (1991) find that firm size is an important determinant of wages, and that wage inequality has shifted both among and within

[1] For a relatively recent survey of the wage inequality literature see Katz and Autor (1999).

manufacturing plants. Abowd, Kramarz and Margolis (1999) examine the relationship between "high wage firms," or firms that seem to pay a wage premium or markup, and "high wage workers," who earn a more than we would expect given their observable characteristics, most likely as a return to unobserved skill.

b) Worker and Firm Reallocation

The re-allocation approach in this paper is different from that taken in the previous literature, which has attempted to directly disentangle the relative contribution of different factors on changes in wage inequality using time series of cross-sectional datasets. By employing longitudinal matched employee-employer data in this study we can focus on the impact of types of workers, types of firms, and the match between the two, on changes in the earnings distribution over time. The results can then be compared with the direct evidence for consistency (or inconsistency). In particular, changes in workforce composition that lead to an increasingly skilled workforce would be consistent with demand side explanations such as skill biased technological change. Similarly, firm entry and exit that lead to high premium firms replacing low premium firms can be linked to firm learning and selection of new technologies (Jovanovic (1982), Ericson and Pakes (1995), Haltiwanger, Lane and Spletzer, 2006).[2] Changing sectoral earnings inequality in low-wage and highly unionized industries would be consistent with hypotheses about the impact of changing unionization and real minimum wages.

A priori, it is certainly clear that there is sufficient turbulence in each of these factors – workforce composition, the types of firms, and the reallocation of workers -- to effect change in the earnings distribution. The potential to change even the most stable workforce at the firm level over a decade or more is quite substantial. Burgess, Lane and Stevens (2000) point out that some 42% of workers are still employed by the same employer in non-manufacturing; 32% in manufacturing after 9 years. In addition, there is ample room for changing firm and industry structure to alter the economic landscape. Davis, Haltiwanger, and Schuh (1996) document the large magnitude of job creation and destruction. They also note the dominance of idiosyncratic factors in accounting for the observed fast pace of job reallocation, while later work (Foster et al, 2005) suggests that

---

[2] It is worth noting that Aghion (2001) has argued that this is an important determinant of earnings inequality.

more productive firms replace less productive ones. Firm entry and exit play a role as well. Spletzer (2000) reports that forty percent of new businesses die within three years of their birth, and more than half of all jobs destroyed in a three-year period are due to the death of establishments.

3. The LEHD Data

The approach is made possible by the existence of a new database created by the Longitudinal Employer and Household Dynamics (LEHD) Program at the U.S. Census Bureau[3]. These data enable us to match workers with past and present employers, together with employer and worker characteristics (Abowd, Haltiwanger and Lane, 2004). This database consists of quarterly records of the employment and earnings of almost all individuals from the unemployment insurance systems of a number of US states in the 1990s[4]. These type of data have been extensively described elsewhere (Haltiwanger, Lane and Spletzer, 2006), but it is worth noting that there are several advantages over household based survey data. The data are current, and the dataset is extremely large. Since the scope of the data is almost the universe of employers and workers in the covered private sector[5], it is possible to trace the movements of workers across earnings categories and across employers. The Unemployment Insurance records have also been matched to internal administrative records containing information on date of birth, place of birth, race, and sex for all workers, thus providing limited demographic information.

Of particular importance given the focus of this study is the reasonably accurate reporting of both earnings and industry. A recent paper by Hirsch and Schumacher (2004) points out that as many of 30% of respondents to the Current Population Survey – the major source of information on earnings inequality in the literature - do not respond to income questions, and are consequently imputed. In the LEHD data, the earnings are quite accurately reported, since there are financial penalties for misreporting. In addition,

[3] Much more detail is provided in the appendix.
[4] Because of the sensitivity of these data it is worth noting that the data are anonymized before they are used in any Census Bureau projects. Any research that is engaged in must be for statistical purposes only, and under Title 13 of the U.S. code, any breach of confidentiality can result in prosecution in which violators are subject to a $250,000 fine and/or 5 years in jail.

[5] Stevens (2002) describes coverage issues related to the LEHD database.

there is substantial internal evidence from the LEHD program that workers not only often misreport earnings, but also do not accurately identify their industry at the major industry level (Decressin et al, 2006).

Because almost all jobs held by all workers in the covered private sector workforce are included the LEHD dataset, it is possible to analyze two different facets of the labor market – both jobs and employment. The two obviously differ to the extent that there is multiple job holding, and to the degree in which there is churning of workers through different sets of jobs. When we use workers as the unit of analysis, we will typically describe their employment with their main (or dominant) employer over the year, and characterize that employer's industry, size, and turnover rates. Earnings refer to quarterly earnings, and we have no information on either wage rates or hours and weeks worked.

Another feature of the LEHD data is new measures of human capital. While standard measures of human capital include such variables as education and experience, other measures, such as ability or family background, have rarely been able to be captured. However, work by Juhn, Murphy, and Pierce (1993), for example, demonstrates that a major contribution to increased earnings inequality in the 1980s was an increase in return to "unmeasured" characteristics—for example, interpersonal skills. The LEHD dataset permits the quantification of the value of these measures, although not permitting a decomposition of the source (Abowd, Kramarz, and Margolis, 1999; Abowd, Lengermann, and McKinney, 2003). This is achieved by capturing the portable component of individual earnings[6]—that component that belongs to an individual as she or he moves from job to job in the labor market and that is separate from the type of firm for which she or he works. We use two measures of human capital: the part associated with the person effect—the unobservable individual heterogeneity—and the time varying experience measure. In interpreting the human capital measure, several remarks should be made. First, the human capital measure is not simply a ranking of the wage of the worker, precisely because wages include both person and firm effects. Second, the person effect will reflect the influence of any time-invariant personal characteristics. Thus, it will reflect factors such as the education level of the individual and other observable

[6] The log real annualized full-time, full-year wage rate described in Abowd et al, 2003.

accumulated skill correlates and it will reflect unobserved dimensions of skill. At the same time, it is a measure of human capital that abstracts from firm effects that may be present in measures based upon observable characteristics – but will not reflect either firm-specific human capital or, in a related manner, match effects.

We also analyze the role of the firm effect. The firm effect literally captures the extent to which the firm the worker is attached pays above or below average wages (after controlling for person effects) and may reflect many factors including rent sharing, firm specific human capital, compensating differentials or unionization effects (Abowd, Lengermann, and McKinney, 2003, Andersson, Holzer and Lane, 2005).

In order to analyze the widest possible time interval (1992 – 2003) we limit our data to four states: California, Illinois, Maryland and North Carolina. In 2003 these states accounted for approximately 21% of U.S. employment.[7]

4. Basic Empirical Facts by Sector

a) Changes in inequality

We use the 1992-2003 difference in log (real) wages at different percentiles to illustrate the changes in the earnings distributions across industry sectors.[8] Table 1 shows the 90th, 50th (median) and 10th percentile of earnings in 2003 and the 90-10, 90-50, and 50-10 log wage differences by sector.

An examination of the first three columns of this table demonstrates that there are substantial earnings differences across industries. For example, median earnings are over twice as high in mining as in the agriculture/fishing/forestry sector, and similar differences hold at both the 90th and 10th percentiles. Earnings at the high and low end of the distribution also vary greatly across sectors. The highest 90th percentile earnings are found in the finance/insurance sector ($114,000), while the lowest 10th percentile earnings are in retail and agriculture/fishing/forestry (both under $10,000). The distribution of earnings also varies across sectors, particularly the 90-10 and 90-50 log wage differences. The 2003 90-10 log wage gap is highest in services, finance/insurance,

[7] The fraction was computed using data from the Current Population Survey's Monthly Outgoing Rotation Groups for 2003.

[8] We use the Standard Industrial Classification (SIC) to identify industry sectors. The public sector is omitted.

wholesale, retail, and manufacturing. These same five industries also had the highest 90-50 log wage differences in 2003. In contrast, inequality at the lower end of the earnings distribution does not vary as much across industries, though services had the largest 50-10 log wage difference.

Figure 1 depicts empirical estimates of the cumulative distribution functions of annualized earnings for all sectors and, as expected, shows a rightward shift from 1992 to 2003. When we repeat the same exercise on an industry by industry basis, we find that although the direction of change is the same for each industry, the most marked rightward shifts were in agriculture/fishing/forestry and the finance/insurance sectors. Other sectors, including manufacturing, transportation/communication, and wholesale, had only modest shifts but with a tendency toward increasing inequality, with a (larger) shift right at the top of the distribution. The most remarkable result, however, is the *lack* of volatility in the earnings distributions shown in these figures, which is especially notable given the enormous amount of economic change that firms have faced over this time period.

The overall rightward shift was not by the same amount at all points of the earnings distribution in each industry, as evidenced in Table 1. In fact, earnings inequality as measured by the 90-10 log wage difference declined in four industries: agriculture, mining, construction, and retail trade. By contrast, earnings inequality increased in five industries: manufacturing, transportation/communication, wholesale, finance/insurance, and services.

In order to compare changes in wage inequality in the upper and lower tails of the earnings distribution, Figures 2-4 plot the 90-10, 90-50, and 50-10 wage gaps by industry, showing clearly the variability in trends across sectors. In three of the four industries in which overall inequality (the 90-10 log wage difference) declined, much or all of the decrease was in the lower half of the earnings distribution (as measured by the 50-10 log wage gap). Only in mining was there much of a decline in the upper half (the 90-50 log difference). In contrast, lower-tail earnings inequality did not increase in the five industries in which overall inequality increased. The increase in earnings inequality in manufacturing, transportation/communication, wholesale, finance/insurance, and services occurred in the upper tail of the earnings distribution. The distance between the 90th and 50th percentiles increased in these six industries, while the 50-10 log difference stayed about the same or decreased slightly. These results confirm the findings of Autor et al. (2005), who find using CPS data that economy-wide, the 90-50 wage gap grew through the 1990s while the 50-10 difference leveled off after about 1987. Yet we also see differences in trends in upper and lower tail inequality across the sectors.

To compare the changes observed in our four-state LEHD data with the U.S. labor market as a whole we computed estimates of the log real weekly earnings percentiles for 1992 and 2003 for all U.S. workers using data from the 1992 and 2003 Current Population Survey's Monthly Outgoing Rotation Groups (CPS-MORG). These estimates are provided in Table 2, Panel A. Panel B of Table 2 presents similar statistics from the CPS-MORG when the samples are limited to the same four states used in our LEHD analyses.

As can be seen from the tables the estimated change in the 90-10 difference between 1992 and 2003 was somewhat higher for the full-sample CPS-MORG data (.09) as compared to the LEHD (.06). The estimated change in the 90-50 difference for the LEHD was actually higher than the CPS-MORG (.09 for LEHD versus .06 for CPS-MORG). The estimated change in the 50-10 difference, however, was substantially lower for the LEHD data (-.03 for LEHD versus .03 for CPS-MORG). As is shown in Panel B of Table 2 much of the differences can be accounted for by the fact that the LEHD contains only four states.

b) Changes in types of workers

One possible reason for these changes in the earnings distribution is that workforce characteristics have changed over time. That there is certainly ample potential for such changes to occur is evident from an examination of Table 3. In manufacturing, for example, of the more than 5 million workers who were employed in either 1992, 2003 or both years, more than 40% were only in the industry in 1992, 35% were only in the industry in 2003, and only 21% were there in both years. As one might expect, the change in the workforce is even more marked in the retail food industry: 39% were only in the industry in 1992; almost half were only in the industry in 2003, and fewer than 12% were in the industry in both years.

The evidence presented in Table 4 shows that this mobility did not translate into enormous swings in the age, gender and skill distribution of workers, although there were some dramatic changes in the allocation of workers across sectors. In particular, the mining and manufacturing sectors shrank as services expanded, and they remained predominantly male and skewed toward older workers. By contrast, industries such as finance, insurance, and real estate as well as services continued to have more females and younger workers. Similarly, although the skill level of the workforce increased in all industries (using both the overall measure of human capital, which includes experience, and the individual fixed effect), the swings are not substantial.

c) Changes in types of firms

Another possible reason for changes in earnings inequality is changes in the types of firms that are hiring workers. We examine this possibility in Table 5, which can be read in the same way as Table 3. In manufacturing, for example, of the more than 100,000 firms who employed individuals in either 1992, 2003 or both years, about 36% were only in the industry in 1992, 37% were only in the industry in 2003, and only 27% were there in both years. The rates are even lower in industries with more small firms: in retail trade, 40% were only in the industry in 1992; 41% were only in the industry in 2003, and about 20% were in the industry in both years. While industry differences exist, as both panels of Figure 5 show, all industries had high rates of firm entry and exit over the 1992-2003 period.

d) *Changes in the joint distribution of human-capital/ pay policy pairings*

One source of change in earnings inequality is changes in the joint distribution of employee human capital and firm pay levels. Suppose we estimate a linear panel data model with fixed firm and individual effects such as that described in Abowd, Kramarz, and Margolis (1999),

$$y_{it} = \mathbf{x_{it}}\boldsymbol{\beta} + \theta_i + \psi_{j(i,t)} + \varepsilon_{it}$$

where y_{it} are individual i's earnings at time t, $\mathbf{x_{it}}$ is a vector of observed productivity measures of individual i at time t, θ_i is an individual fixed effect that measures an individual's unobserved productivity or human capital and $\psi_{j(i,t)}$ the fixed effect of the firm that individual i works for at time t and measures a firm's pay policy. Then changes in the distribution of earnings may be due to changes in the joint distribution of θ and ψ.

For example, over time it may be the case that high θ individuals are more likely to work at high ψ firms and low θ individuals are more likely to work at low ψ firms, which would tend to increase earnings inequality. Using the estimated values of θ and ψ this model, Figure 6 plots the joint distributions of θ and ψ for 1992 and 2003.[9] As can be seen from these graphs between 1992 and 2003 the likelihood of a low θ individual being at a high ψ firm has declined. Thus individuals with low skill levels are less likely in 2003 than 2002 to be paired with firms with high pay policies. Figure 7 displays the expected values of θ by decile groups of ψ for 1992 and 2003. This figure clearly shows that there has been a large upward shift between 1992 and 2003 in the expected value of θ for the highest ψ decile group of firms. Whether this finding is a result of entry and exit of different types of firms and workers or is due to a reshuffling of worker-firm matches, however, cannot be determined these figures.

Although these descriptive statistics hint at the potential for worker and job reallocation to affect earnings inequality, they report average effects for all firms and workers in the industry. To further investigate the trends in inequality, in the next section we develop an econometric approach to examine different points of the earnings distribution.

5. Decomposing Earnings Distribution Changes with Employer-Employee Matched Data

In this section, we develop econometric methods for decomposing changes in earnings distributions when employer-employee matched panel data are available. For the sake of describing the earnings decomposition method, we shall initially assume that we have only one continuous exogenous predictor variable x. Moreover, since we are examining earnings data as opposed to wage data, we do not attempt to examine the impact of changes in the real value of the minimum wage over time (see DiNardo, Fortin and Lemieux, 1996; and Lee, 1999).

[9] The estimates included age variables as controls.

Let θ be a variable representing an individual's (unobserved) productivity that is assumed to be constant over time. Further, let ψ represent a firm's (unobserved) pay policy variable that is also assumed to be constant over time. An individual's earnings is assumed to be determined by the function $y = g(\varepsilon, \theta, \psi, x)$ where ε is a random error component which is assumed to be independent of x, θ, and ψ.

For expositional simplicity, we assume that the variables x, θ, ψ, and ε have a continuous joint probability density function f_t for each time period t=1,2.

(1) $\quad \mathrm{d}P_t(\varepsilon, \theta, \psi, x) = f_t(\varepsilon, \theta, \psi, x)\mathrm{dt}$.

One additional facet of the data is the fact that between the two time periods, within an industry firms can be created or destroyed and workers may enter or exit. Thus, for both firms and workers there are stayers (*s*), leavers (*l*) and new entrants (*n*) Now, we can rewrite the joint distribution in (1) at time period 1 as a mixture of these worker-firm types:

(2)
$$\begin{aligned} f_1(\varepsilon, \theta, \psi, x) &= p_1(\text{w}=s,\ \text{f}=s) f_1^{ss}(\varepsilon, \theta, \psi, x) + p_1(\text{w}=l,\ \text{f}=s) f_1^{ls}(\varepsilon, \theta, \psi, x) + \\ & p_1(\text{w}=s,\ \text{f}=l) f_1^{sl}(\varepsilon, \theta, \psi, x) + p_1(\text{w}=l,\ \text{f}=l) f_1^{ll}(\varepsilon, \theta, \psi, x) \end{aligned} .$$

where p_1(w=s,f=*s*) is the fraction of worker-firm matches where both firm and worker remain in the industry until time 2, p_1(w=*l*,f=s) is the fraction of worker-firm matches where the firm remains in the industry until time 2 but the worker leaves, p_1(w=*s*,f=*l*) is the fraction of worker-firm matches where the worker remains in the industry until time 2 but the firm leaves, and p_1(w=*l*,f=*l*) is the fraction of worker-firm matches where both the worker and firm leave by time 2. The distributions, $f_1^{ss}(\varepsilon, \theta, \psi, x)$, $f_1^{ls}(\varepsilon, \theta, \psi, x)$, $f_1^{sl}(\varepsilon, \theta, \psi, x)$, and $f_1^{ll}(\varepsilon, \theta, \psi, x)$ are the analogous conditional distributions. On the other hand, the joint distribution in (1) at time period 2 as can be written as

(3)
$$\begin{aligned} f_2(\varepsilon, \theta, \psi, x) &= p_2(\text{w}=s,\ \text{f}=s) f_2^{ss}(\varepsilon, \theta, \psi, x) + p_2(\text{w}=n,\ \text{f}=s) f_2^{ns}(\varepsilon, \theta, \psi, x) + \\ & p_2(\text{w}=s,\ \text{f}=n) f_2^{sn}(\varepsilon, \theta, \psi, x) + p_2(\text{w}=n,\ \text{f}=n) f_2^{nn}(\varepsilon, \theta, \psi, x) \end{aligned} .$$

where *n* indicates new entrants into the industry between time 1 and time 2.

The order of the sequential decomposition may differ. In this paper, we first analyze the extent to which worker entry and exit has changed the earnings distribution by considering the counterfactual of what if there had been no exit and entry of workers. In that situation, (3) becomes

(4) $$f_2^{w}(\varepsilon,\theta,\psi,x) = \frac{p_2(\text{w=}s\text{, f=s})f_2^{ss}(\varepsilon,\theta,\psi,x) + p_1(\text{w=}l\text{, f=s})f_1^{ls}(\varepsilon,\theta,\psi,x) + p_2(\text{w=s, f=}n)f_2^{sn}(\varepsilon,\theta,\psi,x)}{R}$$

where

$R = p_2(\text{w=}s\text{, f=}s) + p_1(\text{w=}l\text{, f=}s) + p_2(\text{w=}s\text{, f=}n)$.

Here, we have assumed that, had those individuals who left actually stayed, they would have matched with firms in a manner analogous to the distribution of workers who actually left those firms that stayed in the industry.

Next, we consider the impact of the change in the distribution of x. Here, we note that, for example,

$f_2^{ss}(\varepsilon,\theta,\psi,x) \equiv f_2^{ss}(\varepsilon,\theta,\psi \mid x) f_2(x)$

and replace $f_2(x)$ by $f_1(x)$:

(5) $$f_2^{ss,x}(\varepsilon,\theta,\psi,x) \equiv f_2^{ss}(\varepsilon,\theta,\psi \mid x) f_1(x) = f_2^{ss}(\varepsilon,\theta,\psi,x)\left(f_1(x) \big/ f_2(x) \right)$$

The other terms in (4) are modified in a similar fashion. Thus we have

(6) $$f_2^{w,x}(\varepsilon,\theta,\psi,x) = \frac{p_2(\text{w=s, f=s})f_2^{ss,x}(\varepsilon,\theta,\psi,x) + p_1(\text{w=}l\text{, f=s})f_1^{ls,x}(\varepsilon,\theta,\psi,x) + p_2(\text{w=s, f=}n)f_2^{sn,x}(\varepsilon,\theta,\psi,x)}{R}$$

Next, we look at the impact of firm entry and exit by looking at the counterfactual that assumes that the set of firms (as well as workers and x) at time 2 is the same as time 1:

(7) $$f_2^{w,x,e}(\varepsilon,\theta,\psi,x) = p_1(\text{w=s, f=s})f_2^{ss,x}(\varepsilon,\theta,\psi,x) + p_1(\text{w=}l\text{, f=s})f_1^{ls,x}(\varepsilon,\theta,\psi,x) + p_1(\text{w=}l\text{, f=}l)f_1^{ll,x}(\varepsilon,\theta,\psi,x)$$

Finally, after we have restricted the set of firms and workers to be the same as in period 1, it is still possible to exam how the distribution of θ given ψ may have changed between periods 1 and 2 due to a reallocation of workers across firms within the industry. Now,

(8) $$\begin{aligned} f_2^{w,x,e}(\varepsilon,\theta,\psi,x) &\equiv f_2^{w,x,e}(\varepsilon,\theta,\psi,x \mid \theta,\psi)\, f_2^{w,x,e}(\theta,\psi) \\ &= f_2^{w,x,e}(\varepsilon,\theta,\psi,x \mid \theta,\psi)\, f_2^{w,x,e}(\theta \mid \psi)\, f_1(\psi) \end{aligned}$$

so we can define

(9) $$f_2^{w,x,e,a}(\varepsilon,\theta,\psi,x) \equiv f_2^{w,x,e}(\varepsilon,\theta,\psi,x \mid \theta,\psi)\, f_1^{w,x,e}(\theta \mid \psi)\, f_1(\psi)$$

where the superscript refers to holding the allocation of workers to firms constant. From these counterfactual distributions, we can decompose changes in the earnings distribution. Let Y be the range of y and let A be a subset of Y (i.e. $A \subset Y$). Then,

(10) $$P_t(y \in A) = \int_{\{\varepsilon,\theta,\psi,x\}:g(\varepsilon,\theta,\psi,x)\in A} f_t(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx\,.$$

We can then define the counterfactual probabilities by

(11) $$P_2^{w}(y \in A) = \int_{\{\varepsilon,\theta,\psi,x\}:g(\varepsilon,\theta,\psi,x)\in A} f_2^{w}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx\,,$$

(12) $$P_2^{w,x}(y \in A) = \int_{\{\varepsilon,\theta,\psi,x\}:g(\varepsilon,\theta,\psi,x)\in A} f_2^{w,x}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx$$

(13) $$P_2^{w,x,e}(y \in A) = \int_{\{\varepsilon,\theta,\psi,x\}:g(\varepsilon,\theta,\psi,x)\in A} f_2^{w,x,e}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx$$

and

(14) $$P_2^{w,x,e,a}(y \in A) = \int_{\{\varepsilon,\theta,\psi,x\}:g(\varepsilon,\theta,\psi,x)\in A} f_2^{w,x,e,a}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx\,.$$

The "Oaxaca type" decomposition of the change in the probability of the event $y \in A$ can then be written as

$$
(15) \quad
\begin{aligned}
&P_2(y\in A)-P_1(y\in A)=\\
&(P_2(y\in A)-P_2^{w}(y\in A))+(P_2^{w}(y\in A)-P_2^{w,x}(y\in A))\\
&+(P_2^{w,x}(y\in A)-P_2^{w,x,e}(y\in A))+(P_2^{w,x,e}(y\in A)-P_2^{w,x,e,a}(y\in A))\\
&+(P_2^{w,x,e,a}(y\in A)-P_2(y\in A))
\end{aligned}
$$

Suppose, in general, that we wish to decompose the expected value of some function r of earnings, E(r(y)). Then

$$
\begin{aligned}
&E_2(r(y))-E_1(r(y))=\\
&\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx-\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_1(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx=\\
&\left(\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx-\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2^{w}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx\right)+\\
&\left(\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2^{w}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx-\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2^{w,x}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx\right)+\\
&\left(\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2^{w,x}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx-\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2^{w,x,e}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx\right)+\\
&\left(\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,s))\,f_2^{w,x,e}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx-\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_2^{w,x,e,a}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx\right)+\\
&\left(\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,s))\,f_2^{w,x,e,a}(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx-\int\limits_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\,f_1(\varepsilon,\theta,\psi,x)\,d\varepsilon d\theta d\psi dx\right)
\end{aligned}
$$

where D denotes the domain of ($\varepsilon,\theta,\psi,x$). Note that in (15) we have $r(y)=I(y\in A)$ where I is the indicator function.[10]

[10] This decomposition technique can be extended to the case where the earnings function $y=g(\varepsilon,\theta,\psi,x)$ varies across time (i.e. $y=g_t(\varepsilon,\theta,\psi,x)$) by incorporating an additional decomposition step that would measure the impact of this "structural" change on the distribution of earnings.

We apply this general decomposition technique to the LEHD data described above so that we can explore the determinants of wage and earnings distribution changes over time for different industries. To put more structure on the relationship between y and $(\varepsilon, \theta, \psi, x)$ we assume that the relationship takes the form of a linear panel data model with fixed firm and individual effects such as described in Abowd, Kramarz, and Margolis (1999). So, the function g has the following form

(16) $$y_{it} = g(\varepsilon, \theta, \psi, \mathbf{x}) = \mathbf{x_{it}}\boldsymbol{\beta} + \theta_i + \psi_{j(i,t)} + \varepsilon_{it} \ .$$

Thus, we assume that no structural change has occurred over this time period and (16) was used to estimate the determinants of log earnings using the entire LEHD data.

Since our focus is on examining changes in earnings distributions over time within industries, from this estimation we calculate $dP_t(\varepsilon, \theta, \psi, \mathbf{x})$ for each industry. In the LEHD data all exogenous variables (e.g., age, gender) are discrete. Thus, the discussion presented above that analyzed decompositions with a continuous explanatory variable, does not directly apply. With discrete explanatory variables, however, we simply estimate the distribution of $(\varepsilon, \theta, \psi)$ for each distinct category of exogenous variable within each industry-time cell.

To perform this decomposition, we will be required to estimate the continuous distribution of $(\varepsilon, \theta, \psi)$ for several categories of **x** within each of the ten industries. Since the number of observations in the LEHD data is extremely large, we accomplished this task by discretizing these variables. We discretized each variable by breaking the range into 100 mutually exclusive intervals and assigning the midpoint value to each observation that falls within the interval. This method is applied for all intervals except the lowest and highest intervals (which are unbounded). For the highest (lowest) interval we assign a value that equals the average of the lower (higher) boundary value and the highest (lowest) observed value in the (industry) sample. We denote the discretized vales by $(\varepsilon^d, \theta^d, \psi^d)$. Earnings are then recomputed using the discretized values by:

(17) $$y_{it}^d = \mathbf{x_{it}}\boldsymbol{\beta} + \theta_i^d + \psi_{j(i,t)}^d + \varepsilon_{it}^d \, .$$

The decompositions are then performed on y_{it}^d using the discrete analogs of the equations presented above.

6. Decomposition Results

First we decomposed the changes in the earnings distributions into the amount due to worker entry and exit, changes in observable characteristics, firm entry and exit and changes in the distribution of worker unobserved attributes (θ) for a given firm pay policy (ψ) – i.e., sorting[11]. The results of these decompositions are presented in Table 6.[12]

A graphical display of the decomposition of the entire log earnings distribution is presented in Figures 8 and 9. The decomposition is portrayed in the figures in terms of the differences in the cumulative density functions (c.d.f's). Figure 8 presents the difference in the c.d.f.'s for all sectors between 1992 and 2003, while Figure 9 presents the decomposition of this change into the various factors. Tables 6 and 7 provide the results of the decomposition.

The results presented in the first three rows of Table 6 decompose the sources of earnings changes at three different points of the distribution: the 90th, 50th and 10th percentiles, while the last three describe the effect on changes in inequality. Thus, the first column indicates that log earnings for a worker in the 10th percentile stood at 9.379 in 2003; in the 50th percentile at 10.341; and in the 90th percentile at 11.322. As a result, the 90th percentile worker made roughly 194% more than the 10th percentile worker, and 98% more than the median worker, while the median worker made about 96% more than the 10th percentile worker. A comparison of this with the 1992 distribution, presented in the second to last and last columns, shows that inequality as measured by the 90/10 ratio was about 6.4 log points greater in 2003 than in 1992: the 90th percentile worker made 188% as much as the 10th percentile worker. This increase came about entirely from an increase in the 90-50 difference: the gain in earnings of the 90th percentile worker relative to the median worker was almost 10 log points; the 50-10 difference actually decreased.

These gross changes, however, mask considerable flux in the earnings distribution due to changes in the underlying factors, which are spelled out in the intervening columns. An examination of the second column reveals that changes in the sector

---

[11] One additional step needed when investigating the decomposition of earnings changes for all sectors is the change in the employment distribution of sectors over time. This step is performed first.

[12] While the order of the decomposition may affect the results, switching the order of worker entry and exit and firm entry and exit in the decompositions led to similar findings. For the sake of brevity these results are not reported.

distribution of employment led to small increases in both the 90-50 and 50-10 log earnings differences. Worker entry and exit had no impact on the 50-10 log earnings difference and resulted in a slight decrease in the 90-50 log earnings difference. More important were changes in observed worker characteristics, which led to decreases in both the 50-10 and 90-50 log earnings differences. While entry and exit of firms led to large decreases in the 50-10 and 90-50 log earnings differences, these effects were almost entirely offset by the impact of sorting of workers and firms.

The results presented in Table 7 present exactly the same set of sequential decompositions of the cumulative distribution functions: the first panel presents the results for industries in which inequality declined, the second for industries in which inequality increased.

The first striking result upon examining the decompositions is the degree to which each separate factor affects the earnings distribution, even in industries where – in net terms – there are not substantial changes in earnings. In the services sector, for example, the 90-10 log wage gap rose only slightly from 1.92 in 1992 to 1.97 in 2003 (Table 7). Earnings rose similarly at both ends of the distribution: log earnings in the 10th percentile rose from 9.24 in 1992 to 9.35 in 2003, while log earnings in the 90th percentile in 1992 was 11.16 but had risen to 11.33 by 2003. Yet our analysis reveals remarkable differences in the way in which changes in the composition of workers, firms and the match between the two affects earnings in different parts of the distribution – often in offsetting ways.

Looking first at the four industries in which overall inequality (as measured by the 90-10 log wage difference) declined, Table 7 shows the change in three inequality measures due to each of the factors. The second column reveals that, despite the high levels of worker churning, the churning was among workers of the same average skill level (θ), resulting in basically no change in inequality. That this is true in every industry suggests that, by and large, workforce quality within each industry is quite persistent, which is consistent with work by Haltiwanger, Lane, and Spletzer (2006). An analysis of the third column reveals that, holding θ constant, the aging of the workforce (and the associated returns to experience) acted to decrease earnings inequality in three of the four industries, with little impact in the retail sector. Increased experience led to increased earnings at both ends of the distribution, but with a larger impact at the 10th percentile than at the 90th (except in retail), thus decreasing inequality.

The entry and exit of firms clearly has an enormous impact on the earnings distribution, as is evident from a comparison of column three with column four in Table 7. In mining, column four (compared to column three) shows that if no firm entry or exit had occurred between 1992 and 2003, the 90-10 log wage gap would have swung by over 120 log points, most of which occurred between the 50th and 10th percentiles. Notably, firm entry and exit typically acted to increase earnings at the bottom end of the distribution more than at the top, and in some cases, decreased earnings at the top end, resulting in a decline in the 90/10 ratio in each industry.

Finally, the effect of the match between workers and firms is evident in a comparison between the fourth and fifth columns. Sorting of workers across different sets of firms actually had a larger negative impact on earnings for workers in the 10th percentile than in the 90th percentile in the four industries with declining inequality. Sorting acted to raise 90th percentile earnings in three of the four industries. The net result, however, was an increase in earnings inequality in each of the industries.

The second panel in Table 7 reports the decompositions for the five industries in which overall earnings inequality increased (as measured by the 90-l0 log wage gap). As was the case for the declining-inequality industries, worker churning had little effect on the earnings distribution in these sectors. Comparing the second and third column shows that increased experience lowered inequality by raising earnings more at the bottom than the top of the earnings distribution in each of the industries. Interestingly, however, the impact of changing experience on earnings, while largely symmetric across the distribution within industries, is quite different across industries. For example, changes in experience affected the 10th and 50th earnings percentiles in manufacturing by 20 log points, compared with closer to 10 log points in services.

The effect of firm entry and exit was substantial in these five industries, and in general led to a decrease in earnings inequality. In most of the five industries, entry and exit of firms raised earnings at the bottom more than at the top end of the distribution. In wholesale, the 90th percentile of earnings dropped considerably due to firm entry and exit. Finally, comparing the fourth and fifth columns indicates that sorting of workers among firms generally led to an increase in inequality. Earnings at the bottom of the distribution were much lower due to sorting, leading to a rise in all three inequality measures in all industries. This effect was quite large in manufacturing and services.

Overall, the decompositions in Table 7 show that, while trends in overall inequality as measured by the 90-10 log wage difference diverged across industries, similar factors were at work. Worker entry and exit had little effect on the wage inequality measures, despite high levels of worker churning in the economy. Increased experience (as measured by the aging of workers) acted to increase earnings at all levels, but with a larger impact at the lower end of the distribution. Thus, increased experience tended to lessen wage inequality in all industries. Firm entry and exit and sorting of workers were the biggest factors, with the former acting to decrease inequality offset by the latter effect increasing it in most of the industries. Yet despite the similarities in underlying factors, the size of these effects differed considerably across the industries.

The Kullback-Leibler measure of the distance provides a more aggregate summary statistic for comparing these trends.[13] Table 8 presents the decomposition of the Kullback-Leibler measure between the 1992 and 2003 earnings distributions. There is some evidence that worker entry and exit tended to widen the distance between the 1992 and 2003 earnings distributions in mining, manufacturing, and transportation/communication, while shrinking the distance in all other industries. Changes in observable characteristics tended to widen the distance between the two distributions for all but those three industries as well. Firm entry and exit narrowed the distance between the 1992 and 2003 earnings distributions across all industries. In contrast, sorting of workers among firms widened the distance across all industries.

To summarize, earnings distributions changed differently across all industries over the 1992 to 2003 period. However, while worker entry and exit into an industry appeared to have little effect on the industry earnings distributions over that time period, firm entry and exit tended to compress the dispersion of within-industry earnings distributions, while resorting among firms and workers tended to widen the dispersion of within-industry earnings distributions. Changes in the observable characteristics of

---

[13] The Kullback – Leibler measure for two density functions f_1 and f_2 is defined by $\int_0^\infty [f_1(w) - f_2(w)] \ln(f_1(w)/f_2(w))dw$.

workers, which in our data are primarily the aging of workers within an industry, lead to an increase in all the percentiles we examined, to varying extents, in all industries.

When focusing on the observed decrease in the 50-10 percentile difference, overall and across many industries, it appears that some of this decrease is due to changes in observable characteristics while the remaining appears to be due to the fact that entry and exit of firms and changes in employee – employer matches have big but opposite impacts, the negative impact of entry and exit of firms tends to dominate the positive impact of the change in employer-employee matches. The one factor that appears to have played a role in the increase in the 90-50 percentile difference observed both overall in for many industries is the positive impact of changes in employee – employer matches. Both changes in observable characteristics and firm entry and exit tended to lower this difference.

Before concluding it is important to note that the decomposition of changes in the earnings distribution, while accounting for worker entry and exit into an industry, did not account for the source of worker entry and the destination of those who exit. Workers can enter a particular industry either by leaving another industry or by entering the sample over the period. Similarly workers can exit an industry by moving to another industry or leaving the sample.[14] Some data on the observed and unobserved measures of human capital for these industry entrants (leavers) as well as the pay policies of the firms the join (leave) are presented in Table 9.

Among workers entering industries between 1992 and 2003 approximately 81% were individuals not in the sample in 1992. Industry entrants who are new entrants into the sample between 1992 and 2003 have both lower observed and unobserved skill levels and work at firms who pay less than workers who remained in the industry between 1992 and 2003. While workers who were observed to switch industries between 1992 and 2003 also had lower unobserved skills and worked at firms that tended to pay less than those who didn't switch, the magnitude of the difference was considerably smaller than for

[14] Workers entering the sample can be new labor force entrants who resided in our group of states or migrants from other states. Workers leaving the sample can be workers leaving the labor force or workers moving to a state outside our group of states. Unfortunately, we can not distinguish these different groups of sample entrants and leavers.

those who were new entrants into the sample. Moreover, the observed skills of industry switchers were similar to non-switchers.

Among workers exiting industries between 1992 and 2003 roughly three-quarters had also left the sample. While sample leavers had lower unobserved skills than those who switched industries, their observed skill level was slightly higher. Moreover, they tended to leave firms that had pay policies similar to those from which industry switchers left.

One of the interesting findings is the very large and offsetting effects of entry/exit of firms and the change in worker-firm sorting. An examination of the tables reveals that the largest manifestation of this phenomenon occurs in the 10th percentile of the distribution. This result raises the possibility that the sorting adjustment is due to the substantial increase of the minimum wage in California in 2002 to $6.75, while the other states remained at $5.15. In other words, old California firms would still have been at a lower minimum wage in the counterfactual, not in reality, while new entering firms would have entered with a higher minimum. We examine this possibility in Table 10.

In the service industry, which has large numbers of minimum wage workers, if California is included, the counterfactual had there been no entry and exit of firms is that the log wage in services would drop by .993. Without California, the same counterfactual is that the log wage would drop by only .129. In contrast, the 50th percentile drop of .186 with California included has a much smaller change to a .013 drop with California excluded; similarly, the 90th percentile drop is .002 with California included and the change is only .004 with California excluded.

Of course, other factors may also be at play. The other industry with substantial numbers of minimum wage workers is the retail trade industry, and an examination of Table 10 reveals little evidence of the same pattern in that industry. However, an analysis of the Current Population Survey data shows a substantial drop in the proportion of part-time workers in the industry: from 33% in 1992 to 22% in 2003. The observed phenomenon might be due to exiting firms disportionately hiring part-time workers, while firms entering hire disproportionately many full-time workers. This explanation would thus focus more on hours than wages, and rely on the fact that part-time workers more likely to be at the lower end of earnings distribution.

7. Conclusions

In this paper, we used a new linked employer-employee dataset from the Longitudinal Employer Household Dynamics Program to explore changes in the earnings distributions across sectors of the economy. We hoped to advance knowledge on the way in which the reallocation of jobs and workers affect changes in earnings inequality by focusing on within industry changes in earnings inequality. We directly examined the way in which changes in workforce composition, firm entry and exit and job reallocation affect industry-specific earnings distributions. Finally, we directly examined the degree to which changes in the matching of workers and firms affect earnings inequality.

While there were differences across industries in the magnitudes and directions of change in various aspects of the earnings distribution over the 1992 to 2003 time period, our earnings decompositions revealed that most factors had similar qualitative effects across all industries. In particular, even in industries in which there was very little change on the aggregate earnings distribution between 1998 and 2003, there were enormous, albeit offsetting, changes in the factors contributing to earnings change. Similar factors were at work in industries with declining inequality as well as those with increasing inequality. The magnitudes of these effects, however, varied considerably.

In particular, we found that worker entry and exit had very little impact on changes in the earnings distributions over this time period for the industries examined. In other words, despite the ample opportunities for firms to change their workforce composition, industry workforces remained, by and large, very similar, and earnings gains due to experience tended to be higher at the lower end of the distribution. This does not lend credence to the notion that individual firms are changing their production technologies in a way that is biased towards skill.

Changes in observable characteristics, which mainly involved the aging of the workforce within each industry, tended to shift the earnings distributions of all industries to the right. The effect of an increasingly experienced workforce on earnings inequality is to decrease it in three of the four declining inequality and in all of the five increasing inequality industries – in each case primarily by increasing earnings at the bottom of the earnings distribution.

On the other hand, the net impact of firm entry and exit is to reduce the dispersion of earnings for all industries. In almost all industries this effect acted to increase earnings at the bottom end of the distribution more than at the top. Since firm wage premia are likely to primarily reflect rent sharing, unionization and/or efficiency wage payments, it is difficult to reconcile the fact that these premia are disproportionately being paid to workers at the bottom end of the earnings distribution with a declining importance of wage setting institutions for low wage workers. In addition, we do not find the changing sectoral earnings inequality in low-wage and highly unionized industries that would be consistent with hypotheses about the impact of changing unionization and real minimum wages.

Finally, sorting of workers based on the “human capital” measures over time tended to increase the dispersion of industry earnings distributions between 1992 and 2003. This is consistent with the idea that the driving force of economic change is the entry and exit of firms, and can be linked to the selection of new technologies, and the associated workforce, by new firms.

Our findings suggest that even when earnings distributions seem superficially not to change, or to shift in opposite directions, the extensive amounts of worker and firm reallocation that have been documented in the literature do have large effects on different parts of the earnings distribution. In particular, the entry and exit of firms and sorting of workers and firms based on underlying worker “skills” are important determinants of changes in industry earnings distributions over time.

Our paper demonstrates the utility of matched employer – employee panel data in decomposing changes in earnings distributions over time. Our results suggest that the underlying dynamics of earnings inequality are complex, and are due to factors that cannot be measured in standard cross-sectional data.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms," *Econometrica*. 67: 251-333.

Abowd, John, Paul Lengermann and Kevin McKinney "The Measurement of Human Capital in the U.S. Economy," March 2003, U.S Census Bureau working paper.

Abowd, John, John Haltiwanger and Julia Lane "Integrated Longitudinal Employee-Employer Data for the United States" *American Economic Review*, May 2004

Acemoglu, Daron. 2002. "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 40(1), 7-72.

Andersson, Fredrik, Harry Holzer, and Julia Lane. 2005. *Moving up or Moving On: Who Advances in the Low Wage Labor Market*, Russell Sage Foundation.

Bernard, Andrew B. and J. Bradford Jensen. 1998. Understanding Increasing and Decreasing Wage Inequality," Unpublished paper (Yale University).

Burgess, Simon, Julia Lane, and David Stevens. 2000. "Job Flows and Worker Flows in the Life Cycle of the Firm" *Oxford Bulletin of Economics and Statistics*, 62, 885-908.

Card, David and John DiNardo. 2002. "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20(4): 733-783.

Davis, Steven J., and John Haltiwanger. 1991. "Wage Dispersion within and Between Manufacturing Plants," *Brookings Papers on Economic Activity: Microeconomics*, 115-180.

Davis, Steven J., John Haltiwanger and Scott Schuh. 1996. "Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts." *Small Business Economics, 8*(4): 297-315.

Decressin, Anja, Tomeka Hill, and Julia Lane. 2006. "Employer Provided Health Insurance: What can be learned from the Form 5500?" mimeo, Longitudinal Employer and Household Dynamics, Bureau of Census.

DiNardo, John, Nicole Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992," *Econometrica* Vol. 64, 1001-1004.

Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, 62, 53-82.

Fortin, Nicole and Thomas Lemieux. 1997. "Institutional Changes and Rising Wage Inequality: Is There a Linkage?" *Journal of Economic Perspectives* 11 (2, Spring): 75-96.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2005. "Relocation, Firm Turnover and Efficiency: Selection of Productivity or Profitability," N.B.E.R. Working Paper #11555, Cambridge, MA.

Haltiwanger, John, Julia Lane and Jim Spletzer, 2006 "Wage, Productivity and the Dynamic Interaction of Businesses and Workers" *Labour Economics*, in press.

Hirsch, Barry and Edward Schumacher. 2004. "Match Bias in Wage Gap Estimates Due to Earnings Imputation" *Journal of Labor Economics*, 22 (3): 689-722.

Jovanovic, Boyan, 1982. "Selection and the Evolution of Industry," *Econometrica* 50(3): 649-670.

Juhn, C., K. Murphy and B. Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy* 101 (3): 410-442.

Katz, Lawrence and David Autor. 1999. "Changes in the Wage Structure and Earnings Inequality," in *Handbook of Labor Economics Vol. 3*, O Ashenfelter and D. Card Eds. (North Holland: Amsterdam): 1463-1555.

Lee, David S. 1999. "Wage Inequality in the United States During the 1980s: Rising Dispersion or Falling Minimum Wage," *Quarterly Journal of Economics*, 114, 977-1023.

Lemieux, Thomas, 2004 "Residual Wage Inequality: A Reexamination" working paper, Univeristy of British Columbia.

Spletzer, James, R. 2000. "The Contribution of Establishment Births and Deaths to Employment Growth," *Journal of Business and Economic Statistics*, 18(1): 113-126.

Stevens, David W. 2002. "Employment that is not Covered by State Unemployment Insurance Laws," Technical Paper TP-2002-16, Longitudinal Employer Household Dynamics, U.S. Census Bureau.

Appendix A

This appendix provides some additional detail on measures of mobility and earnings that were constructed using the LEHD data.

Measuring Labor Force Participation

The datasets provide summary statistics (pooled across 4 states) of the earnings and human capital distributions for each sector and each year. For each measure of earnings or human capital, the files provide summary statistics of the distributions of several sets of workers. The characteristics used to identify these different sets of workers are summarized below.

Dominant Employer

A worker's dominant employer is the SEIN (state employer identification number – this is the state UI administrative unit) that contributes the most to the worker's earnings in each year. Thus, each worker employed during a year has one (and only one) dominant employer per year.

Full Time Workers

We use data from Current Population Survey in combination with LEHD state data to impute whether or not a worker is employed full time in each year at his "main" job (analogous to "dominant employer" concept used in LEHD state data). We use CPS variables to perform this imputation using a logit model, and the dependent variable was taken from the CPS question of whether or not the respondent was employed full time at the main employer last year.

Three characteristics of the findings suggest that this imputation was quite successful. First, the standard errors on the coefficients were very small. Second, for individuals found in both the CPS and the LEHD state data, the imputation results were very similar to the observed outcomes. Third, for all individuals, the predicted probabilities of working full time were clustered into two groups such that predicted probabilities for all members of one group were extremely high and the predicted outcomes for the second group were extremely low. More discussion of this imputation can be found in Abowd, Lengermann and McKinney (2003) (hereafter ALM).

Work at End of Quarter One

The distribution of workers employed in a sector at a particular point in time may differ substantially from the distribution of all workers working in the sector at any time during the year. To obtain a "snapshot" of the earnings and human capital distribution in each sector, it is necessary to identify those workers most likely working at a certain point in time. Because we have only quarterly employment information for each worker, we are constrained to approximate the point-in- time employment for each sector. The point of time we have chosen is the end of quarter one of each year. This timing is consistent with the timing of the employment count reported by businesses in the Economic Census and other business surveys. To identify those workers who are most likely working at the same employer at the end of quarter one, we use an indicator of employment at an SEIN in quarter one and in quarter two. The reasoning here is that workers employed at the same SEIN for these two continuous quarters are, with high probability, continuously employed during the two quarters and thus working at the SEIN at the end of the first quarter.

Measuring Earnings:

The LEHD data permit the construction of several different earnings measures. We choose to focus on primarily on log real annualized earnings.

Because we do not observe hours worked in the data but instead only observe quarters worked, we have constructed the "annualized" earnings measure which is, for each worker, the full-time full year earnings equivalent. This variable is adjusted for discontinuities in labor market attachment during the year and is used as the dependent variable in the decomposition of the individual's "wage" into person effect, firm effect, and an experience component. First, we define full quarter employment in quarter t as having an employment history with positive earnings for quarters t - 1 , t , and t +1. Continuous employment during quarter t means having an employment history with positive earnings for either t-1 and t or t and t+1. Employment spells that are neither full quarter nor continuous are designated discontinuous. If the individual was full quarter employed for at least one quarter at the dominant employer, the annualized wage is computed as 4 times average full quarter earnings at that employer (total full quarter earnings divided by the number of full quarters worked). This accounts for 84% of the person-year-state observations in our eventual analysis sample. Otherwise, if the individual was continuously employed for at least one quarter at the dominant employer, the annualized wage is average earnings in all continuous quarters of employment at the dominant employer multiplied by 8 (i.e., 4 quarters divided by expected employment duration during the continuous quarters of 0.5). This accounts for 11% of all observations. For the remaining 5%, annualized wages are average earnings in each quarter multiplied by 12 (i.e., 4 quarters divided by an expected employment duration during discontinuous quarters of 0.33). For additional details, please see ALM.

Measuring Human Capital:

The details of the human capital measures are contained in ALM. In the reported statistics, there are three measures reported: overall human capital (h), the person effect (θ) and the experience effect. Note the overall human capital measure is the sum of the person effect, the experience effect and a reference constant (see in particular equation (25) in ALM). Note that by construction the grand mean of the person effect is zero so some workers (groups) have negative person effects. All components are from a log specification so differences across workers (groups) are interpretable in terms of log differences.

When computing the worker and firm fixed effects, only dominant job spells held by workers who are between 18 and 70 years old and who are imputed to work full time at that job are used to compute human capital measures. Thus, only workers who have been imputed to work full time in at least one job will have a valid person effect. However, once calculated, these measures may be applied to any job spell (dominant or other, full time or other) held by the worker.

Defining the Sample of Workers

The data include year and sector specific earnings and human capital statistics for workers who we have identified as having a dominant employer that year (where the dominant employer is defined as the employer contributing the most to the workers annual earnings), who are imputed to work full time that year, who we have identified as likely working at the end of the first quarter that year, and who have earnings (real) of at least $250 in at least one quarter of the year.

Table 1: Earnings Levels, Differences and Changes by Sector, 1992-2003
Longitudinal Employer Household Dynamics

Sector	2003 Earnings Percentiles			90-10 Log Wage Difference		90-50 Log Wage Difference		50-10 Log Wage Difference	
	90th	50th	10th	2003	Change from 1992	2003	Change from 1992	2003	Change from 1992
AgricFishForestry	$44,149	$19,234	$9,126	1.58	-0.14	0.83	-0.02	0.75	-0.12
Mining	$82,705	$45,879	$22,427	1.30	-0.09	0.59	-0.06	0.72	-0.03
Construction	$73,174	$34,181	$14,831	1.60	-0.12	0.76	-0.01	0.83	-0.11
Manufacturing	$90,650	$34,176	$15,183	1.79	0.13	0.98	0.16	0.81	-0.02
TransCommunication	$82,987	$39,597	$15,959	1.65	0.13	0.74	0.12	0.91	0.01
Wholesale	$96,084	$34,852	$15,307	1.84	0.11	1.01	0.12	0.82	0.00
Retail	$51,404	$19,820	$8,512	1.80	-0.05	0.95	0.01	0.85	-0.06
Finance/Insurance	$114,428	$37,083	$16,244	1.95	0.18	1.13	0.15	0.83	0.03
Services	$83,079	$31,346	$11,523	1.98	0.05	0.97	0.08	1.00	-0.03
All Sectors	**$82,207**	**$31,477**	**$11,992**	**1.93**	**0.06**	**0.96**	**0.10**	**0.97**	**-0.03**

Notes: Based on data from CA, IL, MD and NC.

Table 2A: Weekly Earning Percentile Differences and Changes by Sector, 1992-2003 Current Population Surveys Monthly Outgoing Rotation Groups: All States						
	90-10 Log Wage Difference		**90-50 Log Wage Difference**		**50-10 Log Wage Difference**	
SIC Sector	**2003**	**Change from 1992**	**2003**	**Change from 1992**	**2003**	**Change from 1992**
AgricFishForestry	1.46	0.00	0.79	0.07	0.67	-0.06
Mining	1.34	-0.12	0.66	-0.10	0.68	-0.02
Construction	1.42	0.04	0.73	0.06	0.69	-0.02
Manufacturing	1.51	0.04	0.84	0.07	0.67	0.14
TransCommunication	1.57	0.19	0.72	0.13	0.85	0.06
Wholesale	1.59	0.07	0.82	0.01	0.77	0.06
Retail	1.88	0.04	0.91	-0.01	0.97	0.04
Finance/Insurance	1.75	0.14	0.96	0.12	0.78	0.02
Services	1.97	0.05	0.92	0.04	1.05	0.00
All Sectors	**1.86**	**0.09**	**0.89**	**0.06**	**0.97**	**0.03**
Notes: Data weighted using CPS earnings weights.						

Table 2B: Log Real Weekly Earning Percentiles Differences and Changes by Sector, 1992-2003 Current Population Surveys Monthly Outgoing Rotation Groups: CA IL MD NC						
	90-10 Log Wage Difference		**90-50 Log Wage Difference**		**50-10 Log Wage Difference**	
SIC Sector	**2003**	**Change from 1992**	**2003**	**Change from 1992**	**2003**	**Change from 1992**
AgricFishForestry	1.46	0.22	0.97	0.24	0.49	-0.02
Mining	1.38	0.04	0.96	0.31	0.43	-0.27
Construction	1.46	-0.02	0.75	0.10	0.71	-0.11
Manufacturing	1.69	0.13	0.97	0.17	0.72	-0.05
TransCommunication	1.63	0.36	0.74	0.20	0.89	0.16
Wholesale	1.63	0.13	0.87	0.13	0.76	0.01
Retail	1.89	0.03	0.96	0.02	0.93	0.01
Finance/Insurance	1.78	0.08	0.95	0.12	0.83	-0.04
Services	2.01	0.04	0.95	0.03	1.06	0.01
All Sectors	**1.87**	**0.05**	**0.92**	**0.04**	**0.95**	**0.01**

Notes: Data weighted using CPS earnings weights.

Table 3: Worker Mobility In and Out of Industrial Sectors				
		Proportion in Industry Sector		
Sector	Number of Workers in 1992 & 2003	in 1992 and not 2003	in 2003 and not 1992	in 1992 and 2003
AgricFishForestry	578,036	39.13%	47.97%	12.89%
Mining	67,888	56.47%	29.20%	14.33%
Construction	1,511,595	32.21%	53.40%	14.38%
Manufacturing	5,145,894	43.89%	34.75%	21.35%
TransCommunication	1,775,581	36.58%	44.18%	19.24%
Wholesale Trade	2,006,918	41.03%	46.86%	12.10%
Retail Trade	4,214,151	39.03%	49.25%	11.71%
FIRE	2,101,998	35.97%	47.10%	16.93%
Services	10,196,180	31.00%	50.97%	18.02%
Source: LEHD data for CA, IL, MD and NC.				

Table 4: Changes in Workforce Composition Between 1992 and 2003								
		Change in						
			Proportion of workforce that is				Human Capital	Individual Fixed Effects
Sector	Employment in 1992	Employment	Male	14-29	30-49	50+	(log points)	(log points)
AgricFishForestry	300,709	17%	-7%	-6%	1%	5%	0.17	0.05
Mining	48,063	-39%	2%	0%	-11%	10%	0.11	0.08
Construction	704,268	46%	-1%	-5%	1%	4%	0.1	0.03
Manufacturing	3,357,441	-14%	2%	-7%	-1%	7%	0.17	0.06
TransCommunication	991,212	14%	2%	-3%	-4%	7%	0.09	0.06
Wholesale Trade	1,066,376	11%	0%	-7%	0%	7%	0.13	0.06
Retail Trade	2,138,239	20%	0%	-5%	1%	4%	0.12	0.04
FIRE	1,111,889	21%	3%	-5%	-1%	6%	0.14	0.05
Services	4,998,570	41%	0%	-4%	-4%	7%	0.08	0.06
Source: LEHD data for CA, IL, MD and NC.								

Table 5: Firm Entry and Exit Rates								
		Entrants		Exiters		Continuers		
Sector	Number of Unique Firms in 1992 & 2003	Proportion	Mean firm fixed effect*	Proportion	Mean firm fixed effect*	Proportion	Mean firm fixed effect*
AgricFishForestry	50,825	32%	-0.28	39%	-0.23	29%	-0.23
Mining	2,135	45%	0.38	35%	0.28	20%	0.35
Construction	155,195	33%	0.02	45%	0.05	22%	0.1
Manufacturing	107,200	36%	0.2	37%	0.15	27%	0.18
TransCommunication	56,355	35%	0.18	45%	0.16	20%	0.21
Wholesale Trade	143,414	36%	0.11	43%	0.12	21%	0.11
Retail Trade	263,093	40%	-0.23	41%	-0.24	20%	-0.22
FIRE	120,763	33%	0.13	46%	0.2	22%	0.13
Services	686,606	31%	0.05	49%	0.04	19%	0.01

Source: LEHD data for CA, IL, MD and NC.
* Means are employment weighted.

Table 6: Decompositions of Changes in Log Real Earnings Distributions Between 1992 and 2003: All Sectors								
Measure	**2003 (1)**	**Sector Distribution (2)**	**(2) +Worker entry and exit (3)**	**(3) + Change in observable worker characteristics (4)**	**(4) + Firm entry and exit (5)**	**(5) +Sorting of firms and workers (6)**	**1992 (7)**	**Change from 1992 to 2003 (8)**
10^{th} percentile	9.379	9.393	9.403	9.245	8.261	9.082	9.276	0.103
50^{th} percentile	10.341	10.347	10.357	10.209	10.059	10.188	10.271	0.070
90^{th} percentile	11.322	11.322	11.337	11.254	11.299	11.242	11.155	0.167
90-10 difference	1.943	1.929	1.934	2.009	3.038	2.160	1.879	0.064
90-50 difference	0.981	0.975	0.980	1.045	1.240	1.054	0.884	0.097
50-10 difference	0.962	0.954	0.954	0.964	1.798	1.106	0.995	-0.033
Source: LEHD data for CA, IL, MD and NC.								

Table 7: Decomposition of Changes in Log Earnings Inequality Measures

Measure	2003 (1)	Change in Statistic when also Accounting for: Worker entry and exit (2)	Change in observable worker characteristics (3)	Firm entry and exit (4)	Sorting of firms and workers (5)	1992 (6)	Change from 1992 to 2003 (7)
Panel a: Sectors with Declining Inequality							
Agriculture/Forestry/Fisheries							
10th percentile	9.119	-0.006	0.207	0.264	-0.197	8.837	0.282
50th percentile	9.864	-0.006	0.139	0.038	-0.02	9.703	0.161
90th percentile	10.695	-0.011	0.124	-0.109	0.097	10.556	0.139
90-10 difference	1.576	-0.004	-0.083	-0.373	0.294	1.719	-0.143
90-50 difference	0.831	-0.005	-0.015	-0.147	0.117	0.853	-0.022
50-10 difference	0.746	0	-0.068	-0.226	0.177	0.866	-0.12
Mining							
10th percentile	10.018	0.012	0.224	1.337	-1.374	9.993	0.025
50th percentile	10.734	0.009	0.109	0.293	-0.36	10.741	-0.007
90th percentile	11.323	0.006	-0.026	0.126	-0.029	11.392	-0.069
90-10 difference	1.305	-0.006	-0.25	-1.21	1.345	1.399	-0.094
90-50 difference	0.589	-0.003	-0.135	-0.167	0.332	0.651	-0.062
50-10 difference	0.716	-0.003	-0.115	-1.043	1.014	0.748	-0.032
Construction							
10th percentile	9.605	-0.014	0.158	0.56	-0.362	9.409	0.195
50th percentile	10.439	-0.02	0.12	0.155	-0.079	10.353	0.086
90th percentile	11.201	-0.023	0.065	0.031	0.007	11.125	0.076
90-10 difference	1.596	-0.01	-0.093	-0.53	0.369	1.716	-0.12
90-50 difference	0.761	-0.003	-0.055	-0.124	0.086	0.772	-0.011
50-10 difference	0.835	-0.006	-0.038	-0.405	0.283	0.944	-0.109
Retail							
10th percentile	9.049	-0.006	0.111	0.479	-0.363	8.932	0.117
50th percentile	9.894	-0.01	0.142	0.131	-0.062	9.834	0.06
90th percentile	10.847	-0.011	0.119	-0.017	0.068	10.778	0.069
90-10 difference	1.798	-0.005	0.008	-0.496	0.431	1.846	-0.048
90-50 difference	0.953	-0.001	-0.023	-0.148	0.13	0.944	0.009
50-10 difference	0.845	-0.004	0.031	-0.348	0.301	0.902	-0.057

Table 7: Decomposition of Changes in Log Earnings Inequality Measures (continued)							
		Change in Statistic when also Accounting for					
Measure	2003 (1)	Worker entry and exit (2)	Change in observable worker characteristics (3)	Firm entry and exit (4)	Sorting of firms and workers (5)	1992 (6)	Change from 1992 to 2003 (7)
Panel b: Rising Inequality							
Manufacturing							
10th percentile	9.628	0.005	0.199	1.349	-1.294	9.547	0.081
50th percentile	10.439	0.007	0.212	0.13	-0.247	10.381	0.058
90th percentile	11.415	0.002	0.128	-0.076	-0.011	11.199	0.216
90-10 difference	1.787	-0.003	-0.071	-1.425	1.283	1.652	0.135
90-50 difference	0.975	-0.004	-0.084	-0.205	0.236	0.818	0.157
50-10 difference	0.811	0.002	0.013	-1.219	1.047	0.834	-0.023
TransCommunication							
10th percentile	9.678	-0.013	0.172	0.702	-0.69	9.702	-0.024
50th percentile	10.587	-0.012	0.09	0.134	-0.173	10.605	-0.018
90th percentile	11.326	-0.005	0.065	0.011	-0.045	11.22	0.106
90-10 difference	1.649	0.008	-0.107	-0.691	0.645	1.518	0.131
90-50 difference	0.74	0.007	-0.025	-0.122	0.128	0.615	0.125
50-10 difference	0.909	0.001	-0.081	-0.568	0.517	0.903	0.006
Wholesale Trade							
10th percentile	9.636	-0.003	0.172	0.462	-0.379	9.55	0.086
50th percentile	10.459	-0.004	0.148	0.022	-0.09	10.376	0.083
90th percentile	11.473	-0.004	0.077	-0.377	0.292	11.275	0.198
90-10 difference	1.837	-0.001	-0.096	-0.839	0.672	1.725	0.112
90-50 difference	1.014	0	-0.071	-0.399	0.382	0.899	0.115
50-10 difference	0.823	-0.001	-0.025	-0.44	0.289	0.826	-0.003
FIRE							
10th percentile	9.695	-0.006	0.18	0.79	-0.66	9.526	0.169
50th percentile	10.521	-0.007	0.162	0.139	-0.076	10.323	0.198
90th percentile	11.648	-0.006	0.091	-0.098	0.201	11.296	0.352
90-10 difference	1.952	0	-0.088	-0.888	0.861	1.77	0.182
90-50 difference	1.127	0.001	-0.071	-0.237	0.277	0.973	0.154
50-10 difference	0.825	-0.001	-0.018	-0.651	0.584	0.797	0.028
Services							
10th percentile	9.352	0.021	0.134	1.265	-0.993	9.241	0.111
50th percentile	10.353	0.023	0.108	0.186	-0.076	10.269	0.084
90th percentile	11.328	0.012	0.04	0.002	0.106	11.163	0.165
90-10 difference	1.975	-0.009	-0.094	-1.264	1.099	1.922	0.053
90-50 difference	0.975	-0.011	-0.068	-0.184	0.182	0.894	0.081
50-10 difference	1.001	0.001	-0.026	-1.079	0.917	1.028	-0.027

Source: LEHD data for CA, IL, MD and NC.
Notes: Entries in columns (2) - (5) report the change in the measure when the factor is either assumed not to have occurred as in worker and firm entry and exit or replaced by its value in 1992 as in observed worker characteristics and the conditional distribution of worker matches (θ) given a firm level of pay (ψ).

Table 8: Kullback-Leibler Distance Measure Decompositions					
		Change in Statistic when also Accounting for			
	Change 1992 to 2003	Worker entry and exit	Change in observable worker characteristics	Firm entry and exit	Sorting of firms and workers
Sector	(1)	(2)	(3)	(4)	(5)
AgricFishForestry	0.094	0.099	0.013	0.046	0.011
Mining	0.016	0.016	0.071	0.666	0.066
Construction	0.032	0.040	0.006	0.095	0.021
Manufacturing	0.036	0.035	0.070	0.318	0.055
TransCommunication	0.025	0.025	0.041	0.191	0.034
Wholesale Trade	0.028	0.029	0.028	0.204	0.058
Retail Trade	0.019	0.022	0.015	0.134	0.043
FIRE	0.093	0.098	0.025	0.166	0.031
Services	0.024	0.031	0.014	0.190	0.057

Source: LEHD data for CA, IL, MD and NC.

Table 9: Sector and Sample Mobility between 1992 and 2003

Mobility Type	Number	Percent	Mean θ	Mean xb 1992	Mean xb 2003	Mean ψ 1992	Mean ψ 2003
Switchers: Switch Sectors between 1992 & 2003	2,523,443	9.58%	0.08	1.08	1.37	0.04	0.08
Entrants: Out of 1992 Sample, In 2003 Sample	10,982,559	41.70%	-0.05	.	1.26	.	0.01
Exiters: In 1992 Sample, Out of 2003 Sample	7,859,437	29.84%	0.02	1.15	.	0.04	.
Stayers: Remain in Sector between 1992 & 2003	4,974,338	18.89%	0.24	1.16	1.38	0.09	0.09

Source: LEHD data for CA, IL, MD and NC.

Table 9: Sector and Sample Mobility between 1992 and 2003

Mobility Type	Number	Percent	Mean Unobserved skill	Mean Observed skill 1992	Mean Observed skill 2003	Mean Firm effect 1992	Mean Firm effect 2003
Switchers: Switch sectors between 1992 & 2003	2,523,443	9.58%	0.08	1.08	1.37	0.04	0.08
Exiters: Out of 1992 sample and in 2003 Sample	10,982,559	41.70%	-0.05	.	1.26	.	0.01
Entrants: In 1992 sample and out of 2003 Sample	7,859,437	29.84%	0.02	1.15	.	0.04	.
Stayers: Remain in sector between 1992 & 2003	4,974,338	18.89%	0.24	1.16	1.38	0.09	0.09

Source: LEHD data for CA, IL, MD and NC.

Table 10: Decomposition of Changes in Log Earnings Inequality Measures Results with California Excluded							
		Change in Statistic when also Accounting for					
Measure	2003 (1)	Worker entry and exit (2)	Change in observable worker characteristics (3)	Firm entry and exit (4)	Sorting of firms and workers (5)	1992 (6)	Change from 1992 to 2003 (7)
Agriculture/Forestry/Fisheries							
10th percentile	9.195	-0.015	0.174	0.263	-0.180	8.976	0.219
50th percentile	9.980	-0.018	0.140	0.063	-0.033	9.869	0.111
90th percentile	10.825	-0.045	0.087	-0.080	0.102	10.715	0.110
90-10 difference	1.630	-0.030	-0.087	-0.343	0.282	1.739	-0.109
90-50 difference	0.845	-0.027	-0.053	-0.143	0.135	0.846	-0.001
50-10 difference	0.785	-0.003	-0.034	-0.200	0.147	0.893	-0.108
Mining							
10th percentile	9.930	0.002	0.265	0.649	-0.474	9.907	0.023
50th percentile	10.653	0.002	0.187	0.160	-0.296	10.703	-0.050
90th percentile	11.139	0.002	0.139	-0.002	-0.106	11.159	-0.020
90-10 difference	1.209	0.000	-0.126	-0.651	0.368	1.252	-0.043
90-50 difference	0.486	0.000	-0.048	-0.162	0.190	0.456	0.030
50-10 difference	0.723	0.000	-0.078	-0.489	0.178	0.796	-0.073
Construction							
10th percentile	9.600	-0.013	0.168	0.552	-0.375	9.434	0.166
50th percentile	10.422	-0.018	0.155	0.140	-0.087	10.289	0.133
90th percentile	11.184	-0.021	0.085	0.007	0.023	11.042	0.142
90-10 difference	1.584	-0.008	-0.083	-0.545	0.398	1.608	-0.024
90-50 difference	0.762	-0.003	-0.070	-0.133	0.110	0.753	0.009
50-10 difference	0.822	-0.005	-0.013	-0.412	0.288	0.855	-0.033
Manufacturing							
10th percentile	9.648	0.003	0.249	1.548	-1.547	9.543	0.105
50th percentile	10.383	0.003	0.243	0.096	-0.219	10.270	0.113
90th percentile	11.264	0.003	0.151	-0.247	0.078	11.084	0.180
90-10 difference	1.616	0.000	-0.098	-1.795	1.625	1.541	0.075
90-50 difference	0.881	0.000	-0.092	-0.343	0.297	0.814	0.067
50-10 difference	0.735	0.000	-0.006	-1.452	1.328	0.727	0.008
TransCommunication							
10th percentile	9.686	-0.003	0.195	0.268	-0.254	9.555	0.131
50th percentile	10.455	-0.004	0.158	-0.092	-0.044	10.344	0.111
90th percentile	11.453	-0.004	0.085	-1.040	0.863	11.250	0.203
90-10 difference	1.767	-0.001	-0.110	-1.308	1.117	1.695	0.072

Table 10: Decomposition of Changes in Log Earnings Inequality Measures
Results with California Excluded

Measure	2003 (1)	Change in Statistic when also Accounting for: Worker entry and exit (2)	Change in observable worker characteristics (3)	Firm entry and exit (4)	Sorting of firms and workers (5)	1992 (6)	Change from 1992 to 2003 (7)
90-50 difference	0.998	0.000	-0.073	-0.948	0.907	0.906	0.092
50-10 difference	0.769	-0.001	-0.037	-0.360	0.210	0.789	-0.020
Wholesale Trade							
10th percentile	8.960	-0.009	0.094	0.578	-0.502	8.869	0.091
50th percentile	9.864	-0.011	0.156	0.117	-0.090	9.774	0.090
90th percentile	10.829	-0.016	0.126	-0.144	0.122	10.736	0.093
90-10 difference	1.869	-0.007	0.032	-0.722	0.624	1.867	0.002
90-50 difference	0.965	-0.005	-0.030	-0.261	0.212	0.962	0.003
50-10 difference	0.904	-0.002	0.062	-0.461	0.412	0.905	-0.001
Retail							
10th percentile	9.715	-0.009	0.183	0.908	-0.780	9.538	0.177
50th percentile	10.497	-0.010	0.183	0.127	-0.098	10.272	0.225
90th percentile	11.621	-0.007	0.105	-0.174	0.232	11.260	0.361
90-10 difference	1.906	0.002	-0.078	-1.082	1.012	1.722	0.184
90-50 difference	1.124	0.003	-0.078	-0.301	0.330	0.988	0.136
50-10 difference	0.782	-0.001	0.000	-0.781	0.682	0.734	0.048
FIRE							
10th percentile	9.340	-0.026	0.143	0.252	-0.194	9.222	0.118
50th percentile	10.303	-0.027	0.127	0.059	-0.035	10.197	0.106
90th percentile	11.240	-0.022	0.062	-0.163	0.183	11.087	0.153
90-10 difference	1.900	0.004	-0.081	-0.415	0.377	1.865	0.035
90-50 difference	0.937	0.005	-0.065	-0.222	0.218	0.890	0.047
50-10 difference	0.963	-0.001	-0.016	-0.193	0.159	0.975	-0.012
Services							
10th percentile	9.832	-0.070	0.094	0.129	-0.183	9.816	0.016
50th percentile	10.464	-0.083	0.069	0.013	-0.055	10.452	0.012
90th percentile	11.066	-0.063	0.064	0.004	-0.002	10.969	0.097
90-10 difference	1.234	0.007	-0.030	-0.125	0.181	1.153	0.081
90-50 difference	0.602	0.020	-0.005	-0.009	0.053	0.517	0.085
50-10 difference	0.632	-0.013	-0.025	-0.116	0.128	0.636	-0.004

Source: LEHD data for IL, MD and NC.
Notes: Entries in columns (2) - (5) report the change in the measure when the factor is either assumed not to have occurred as in worker and firm entry and exit or replaced by its value in 1992 as in observed worker characteristics and the conditional distribution of worker matches (q) given a firm level of pay (y).

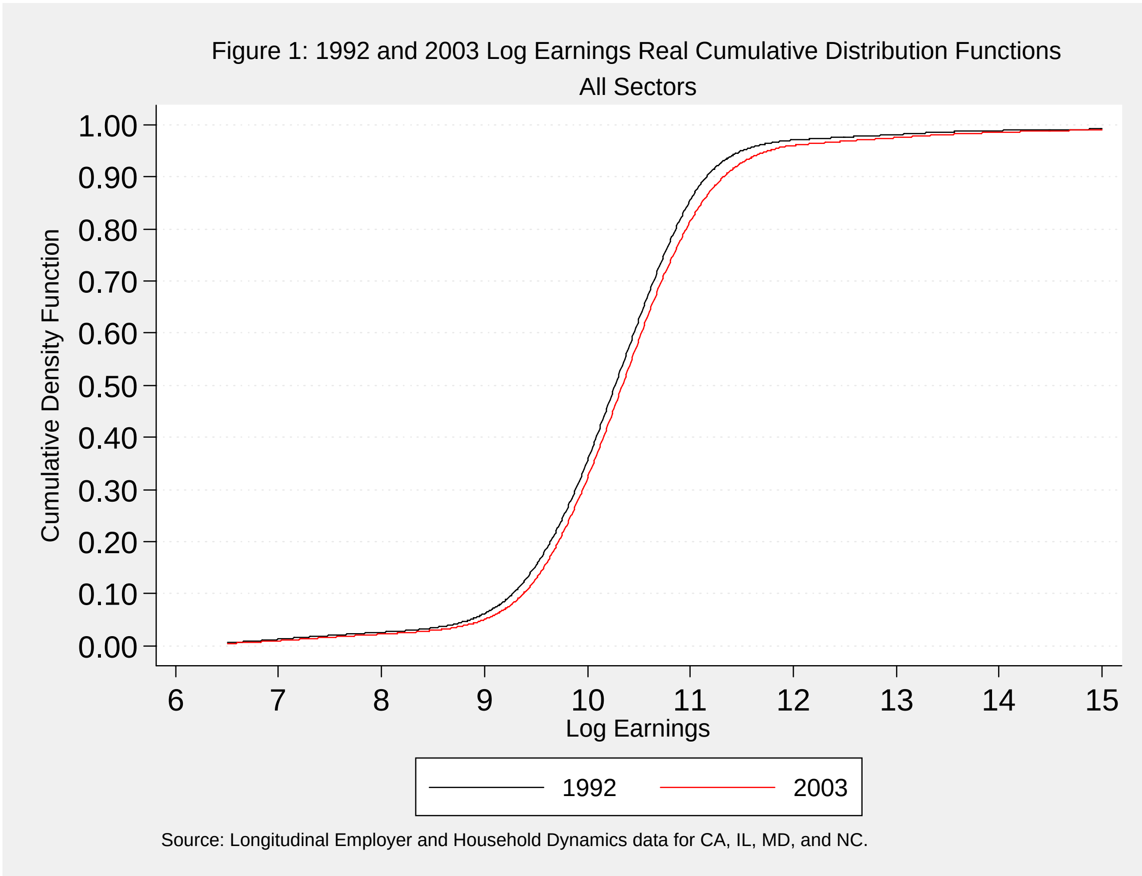
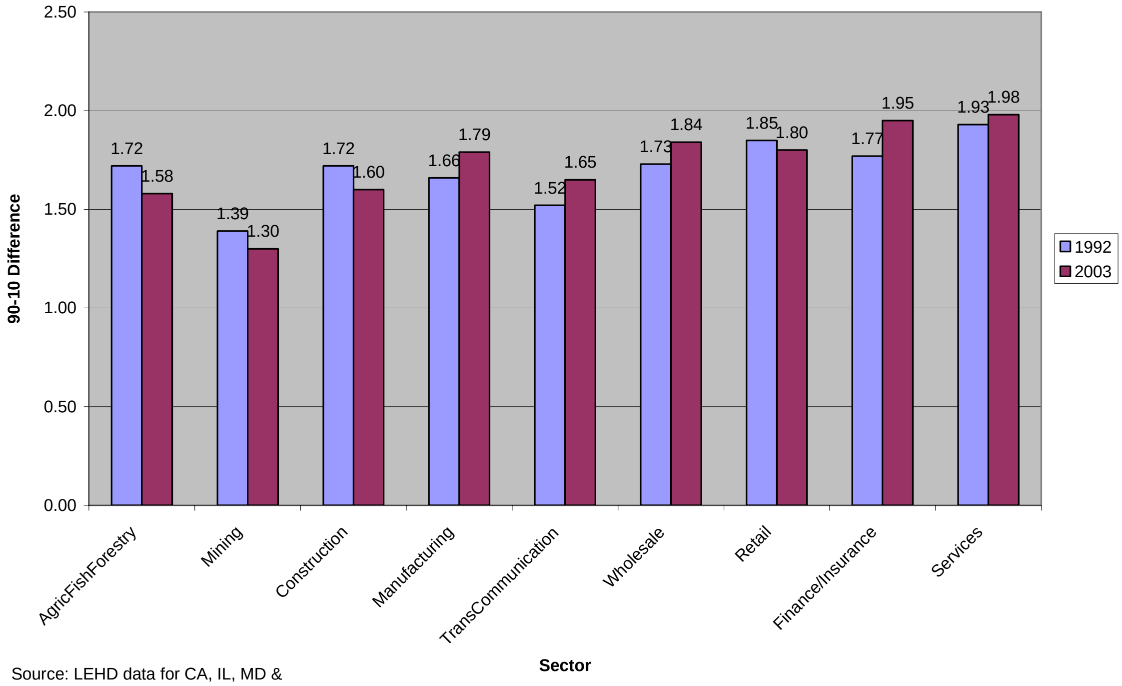


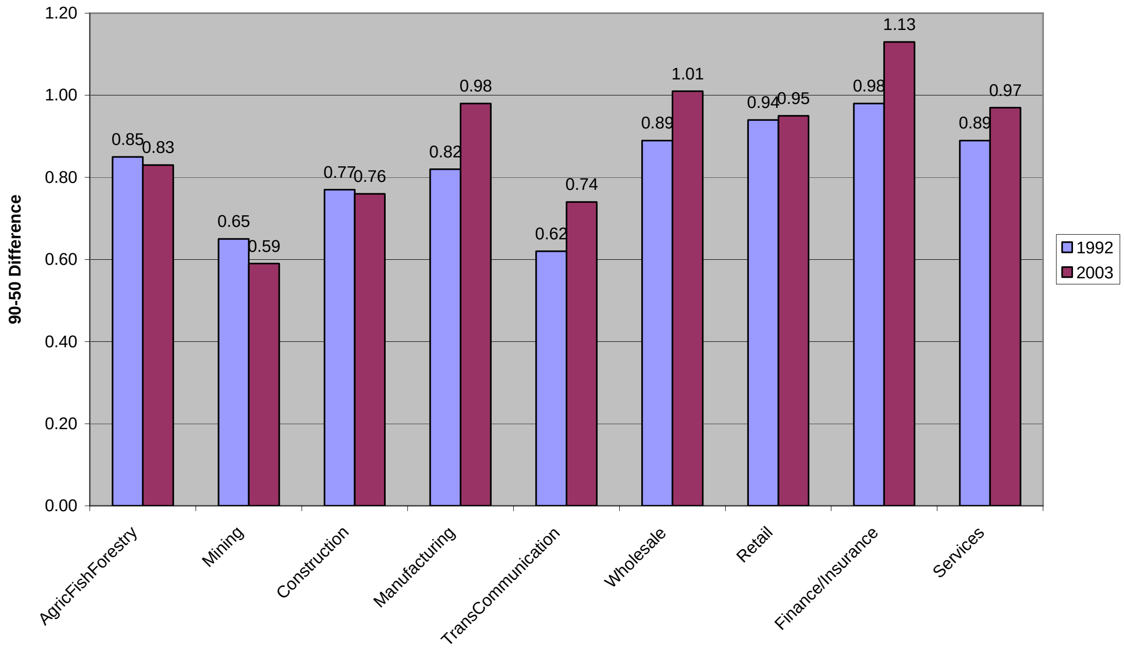
Figure 1: 1992 and 2003 Log Earnings Real Cumulative Distribution Functions
All Sectors
Cumulative Density Function
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
6
7
8
9
10
11
12
13
14
15
Log Earnings
1992
2003
Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.

Figure 2: 90-10 Log Earnings Differences by Sector, 1992-2003



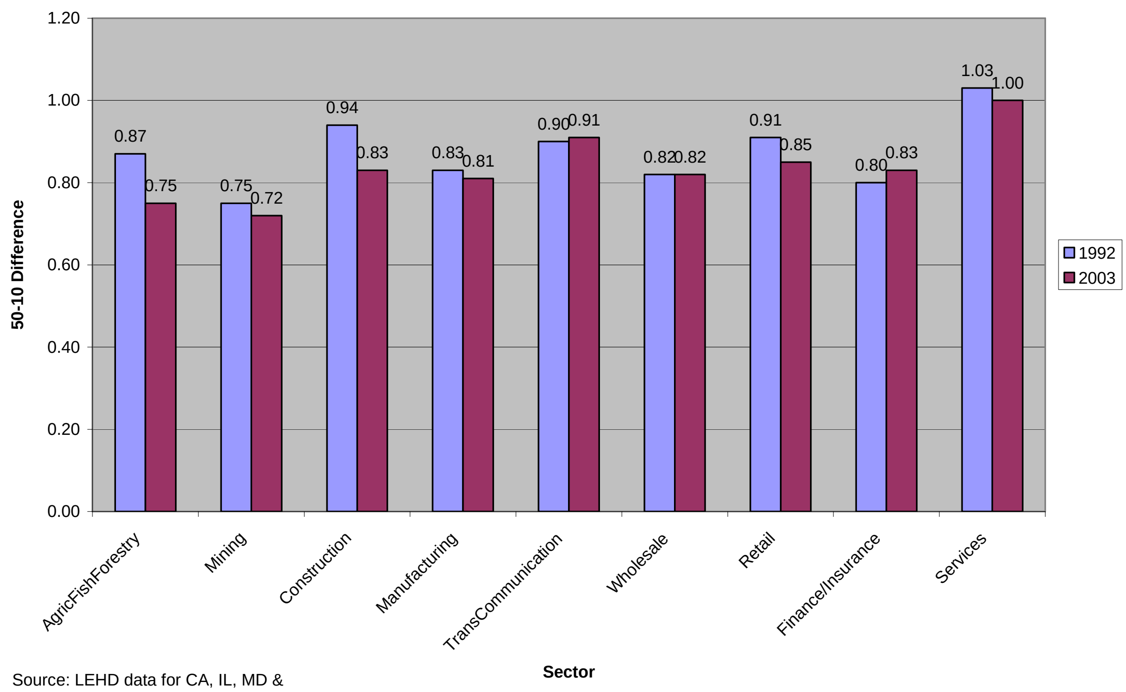
Source: LEHD data for CA, IL, MD & NC

Figure 3: 90-50 Log Earnings Differences by Sector, 1992-2003



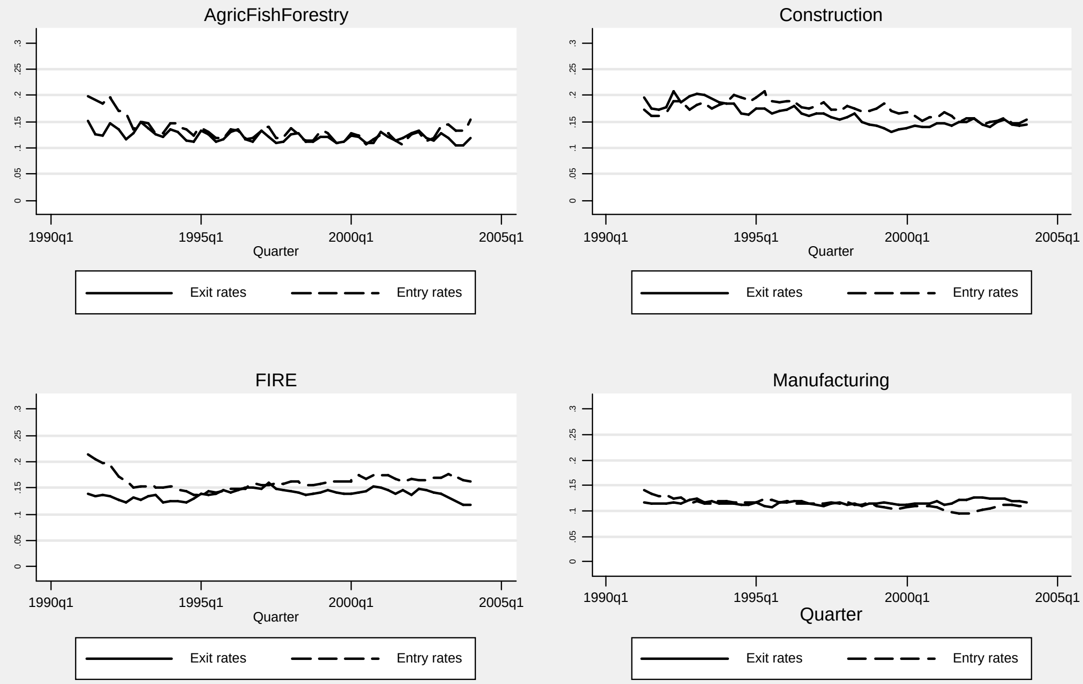
Source: LEHD data for CA, IL, MD &

Figure 4: 50-10 Log Earnings Differences by Sector, 1992-2003



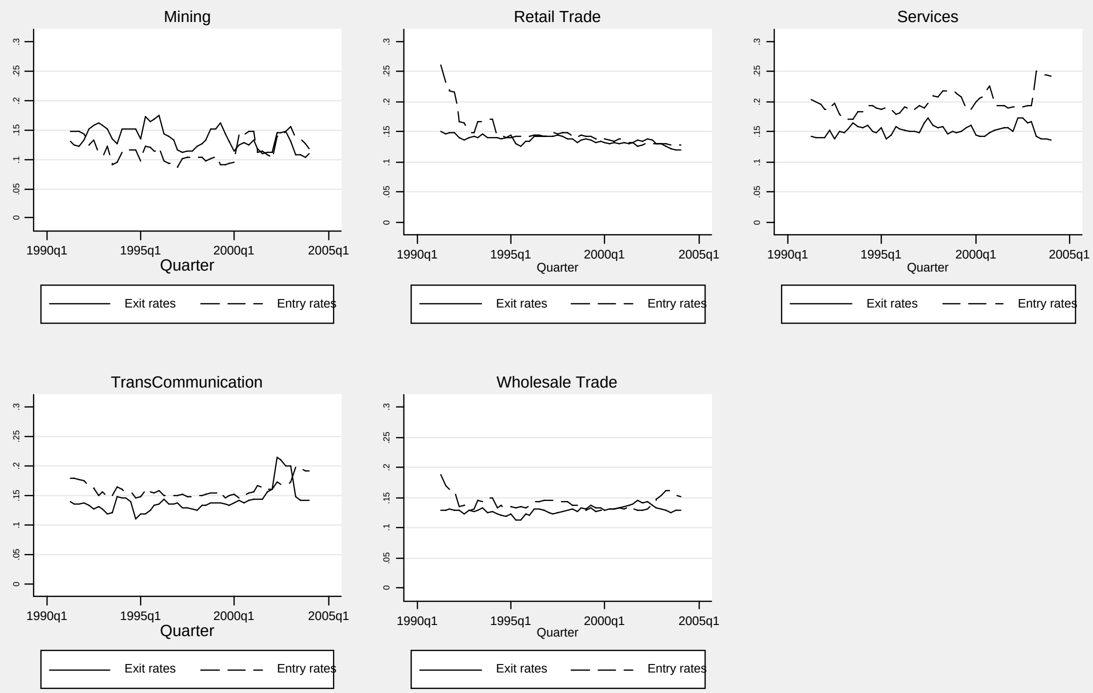
Source: LEHD data for CA, IL, MD &

Figure 5: Quarterly Firm Entry and Exit Rates by Sector
Panel A



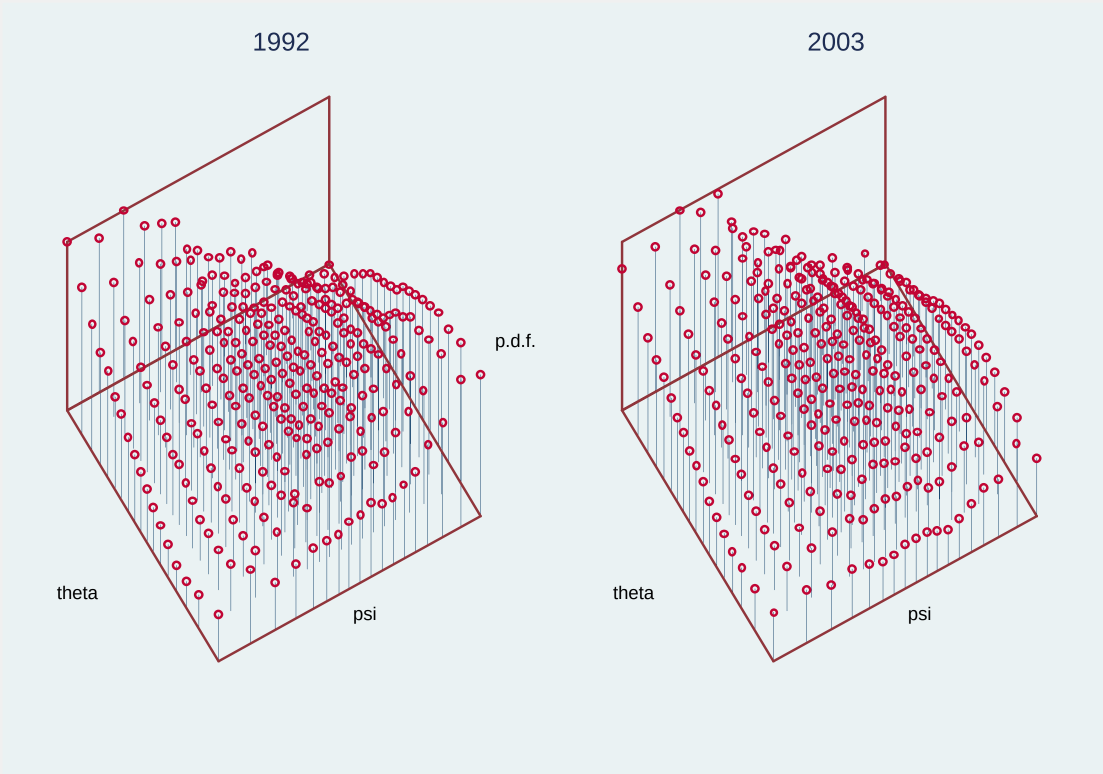
Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.

Figure 5: Quarterly Firm Entry and Exit Rates by Sector
Panel B



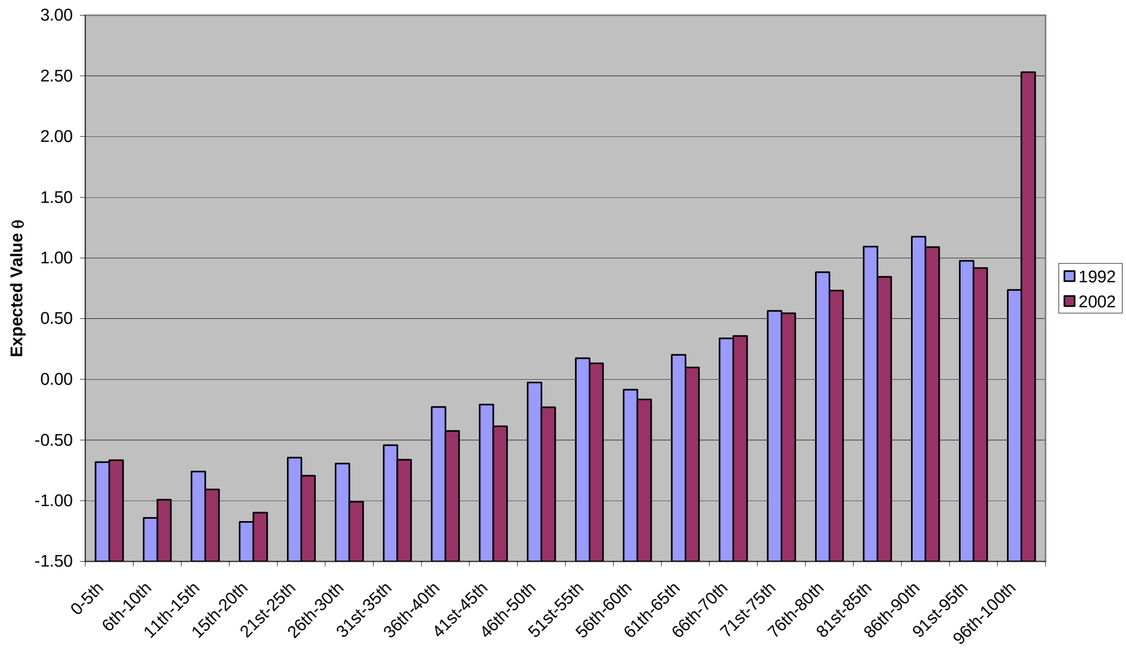
Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.

Figure 6: Joint Distribution of Worker Human Capital (theta) and Firm Pay Policy (psi) Match



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD & NC.

Figure 7: Expected Value of θ by Percentile of ψ



means.

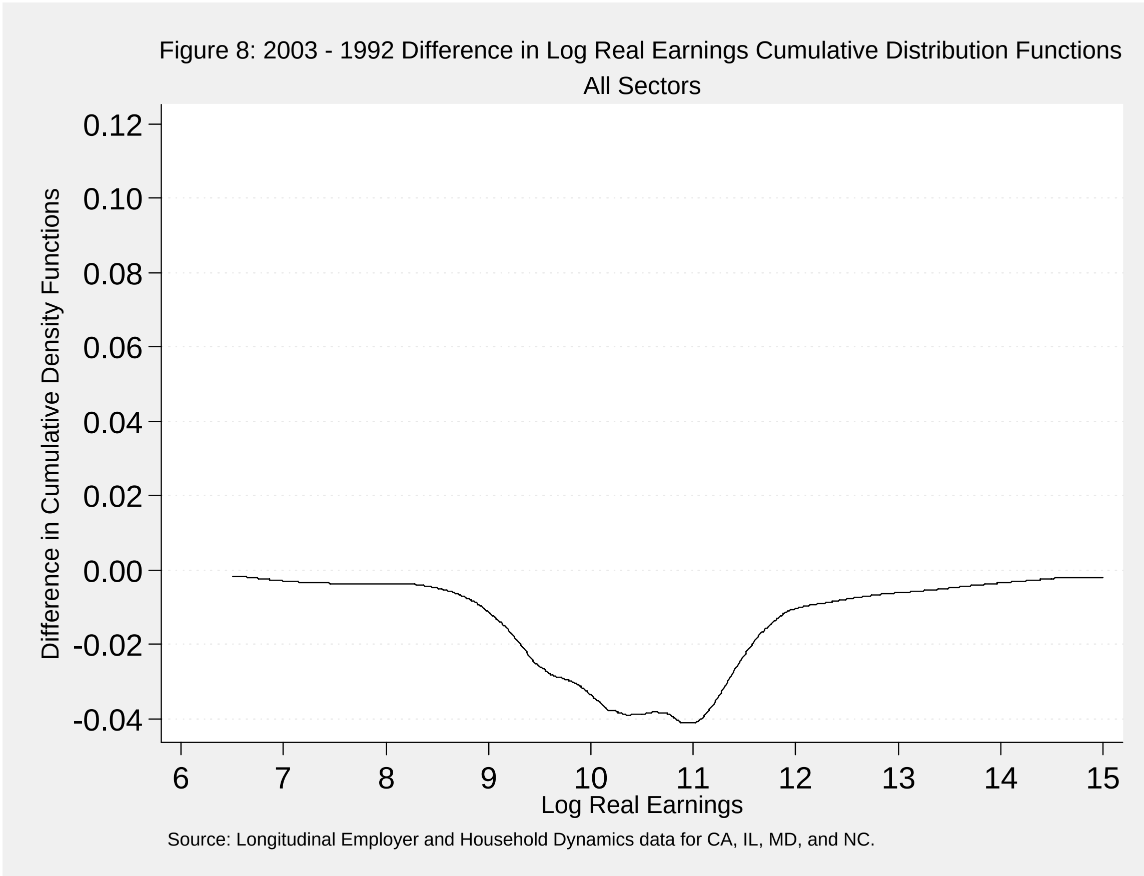
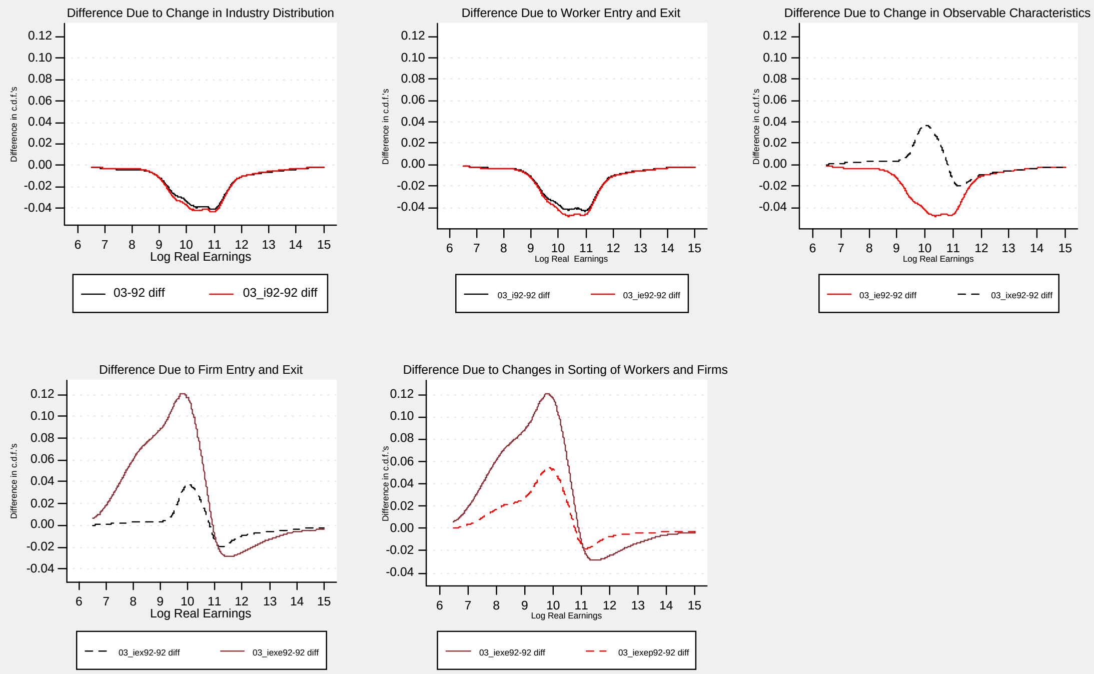


Figure 8: 2003 - 1992 Difference in Log Real Earnings Cumulative Distribution Functions
All Sectors
Difference in Cumulative Density Functions
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
-0.04
6
7
8
9
10
11
12
13
14
15
Log Real Earnings
Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.

Figure 9: Decomposition of Differences in Log Real Earnings Cumulative Distribution Functions
All Sectors



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_i92, 03_ie92, 03_iex92, 03_iexe92, and 03_iexep92-92 refers to the 2003 distribution with the industry employment distribution,industry employment distribution - set of workers, industry employment distribution - set of workers - observable characteristics, industry employment distribution - set of workers - observable characteristics - set of firms, and industry employment distribution- set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

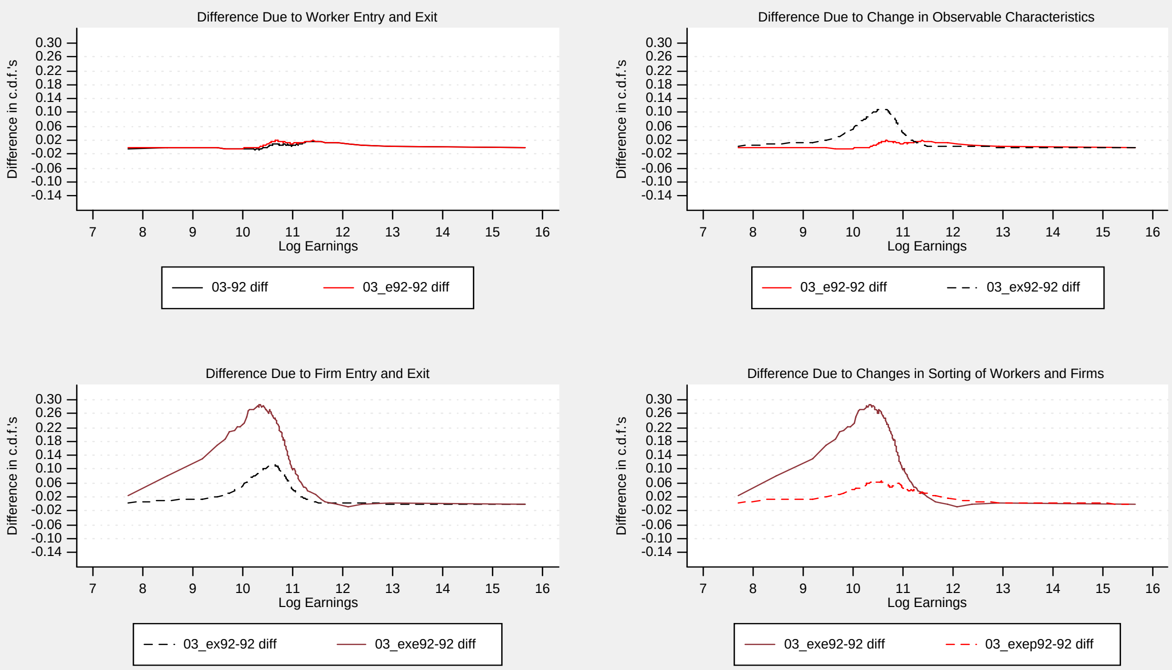
Appendix B: Decompositions of Log Real Earnings by Sector

Figure B1

Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

Figure B2

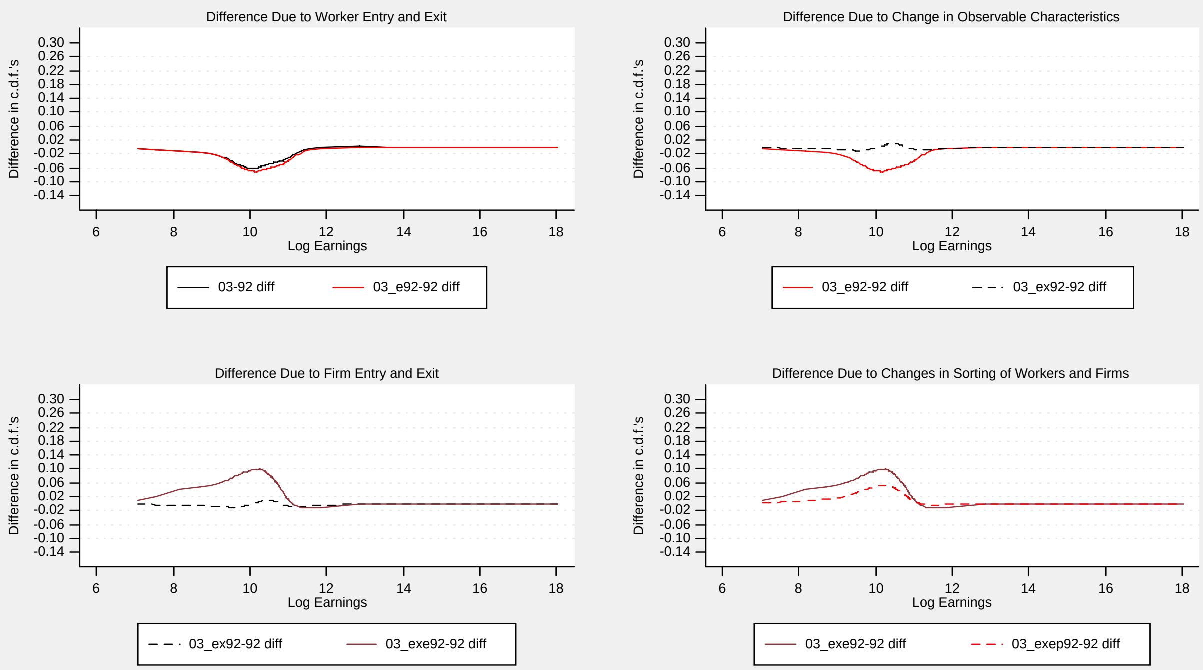
Decomposition of Differences in Log Real Earnings Cumulative Distribution Functions
Mining Sector



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

Figure B3

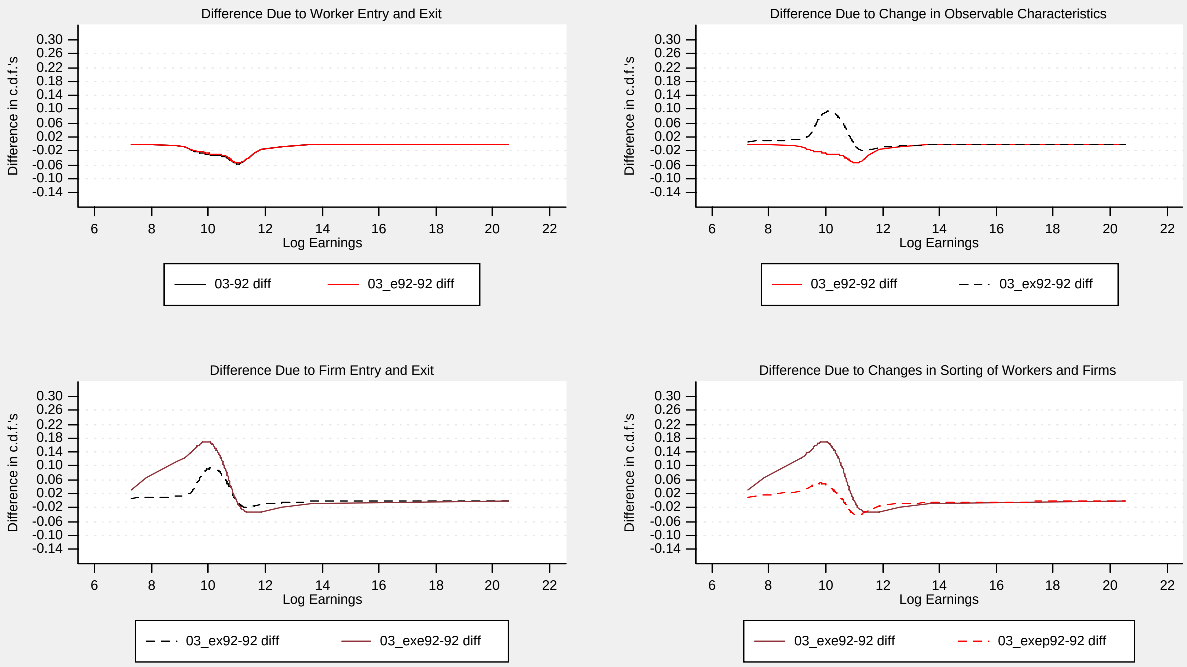
Decomposition of Differences in Log Real Earnings Cumulative Distribution Functions
Construction Sector



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

Figure B4

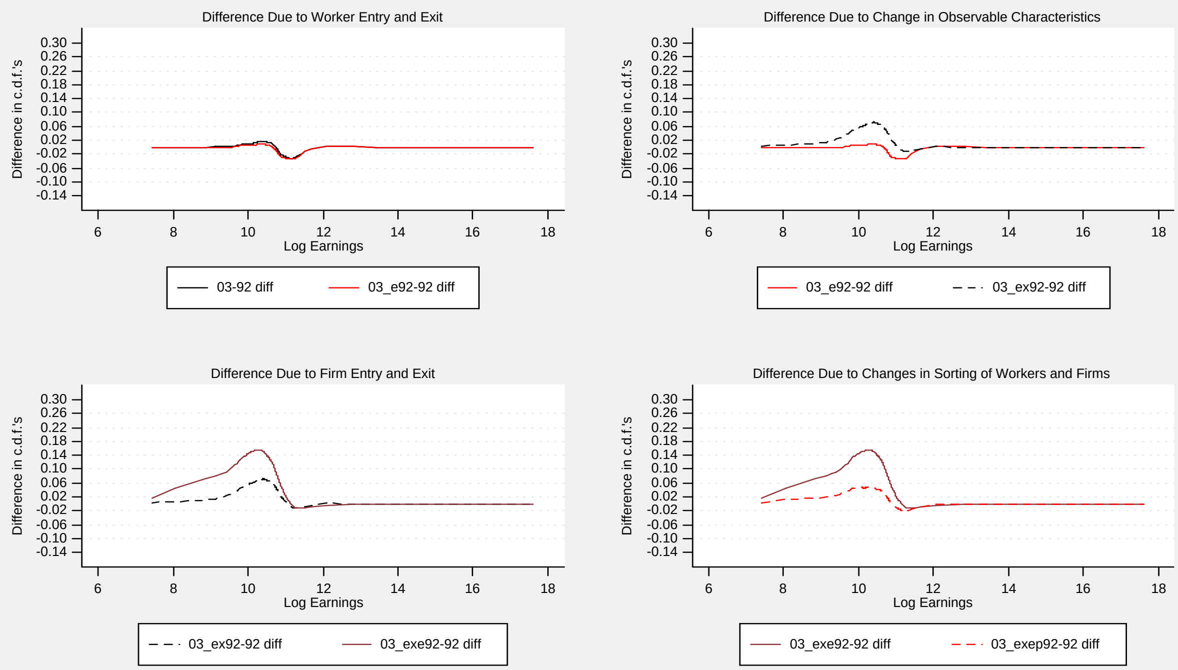
Decomposition of Differences in Log Real Earnings Cumulative Distribution Functions
Manufacturing Sector



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

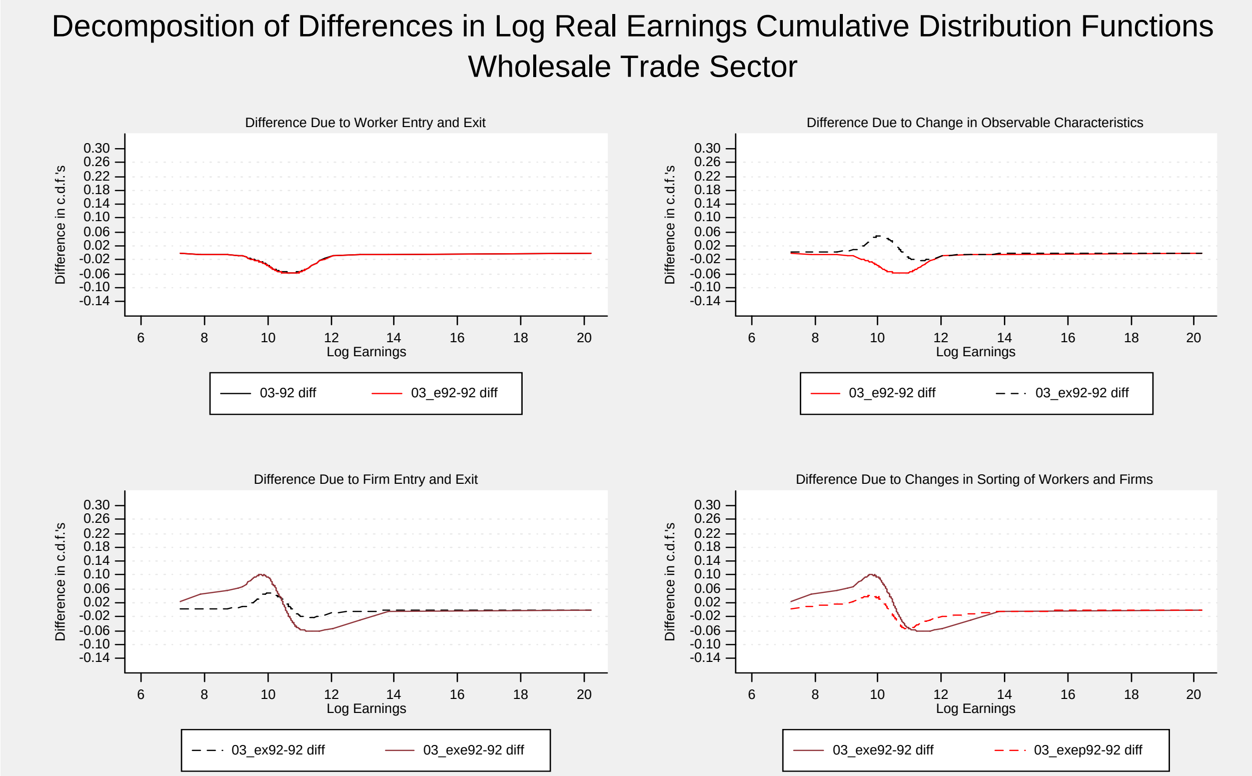
Figure B5

Decomposition of Differences in Log Real Earnings Cumulative Distribution Functions
TransCommunication Sector



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

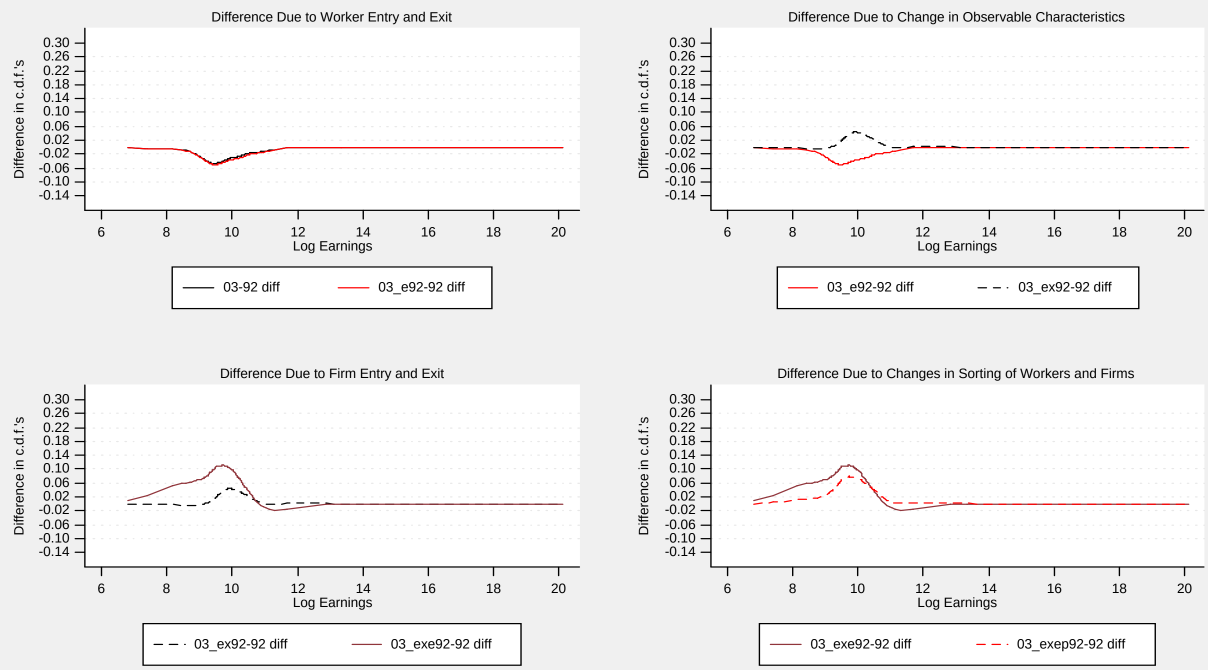
Figure B6



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

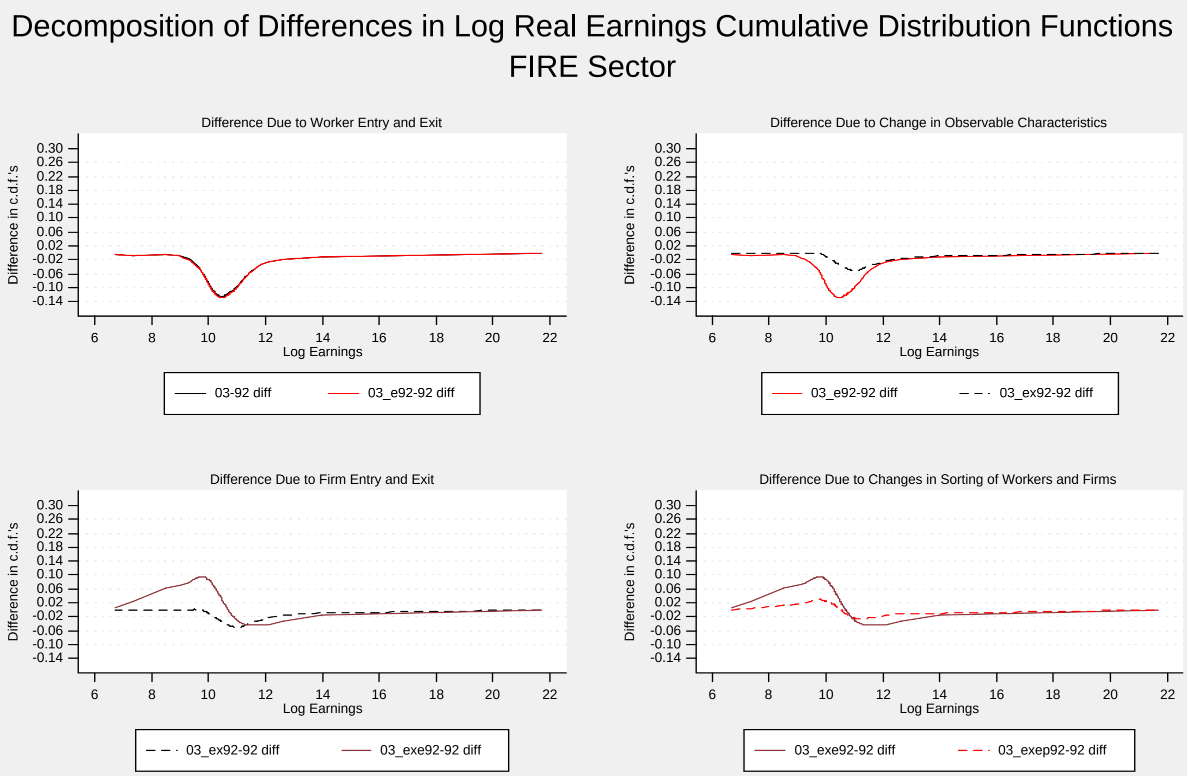
Figure B7

Decomposition of Differences in Log Real Earnings Cumulative Distribution Functions
Retail Trade Sector



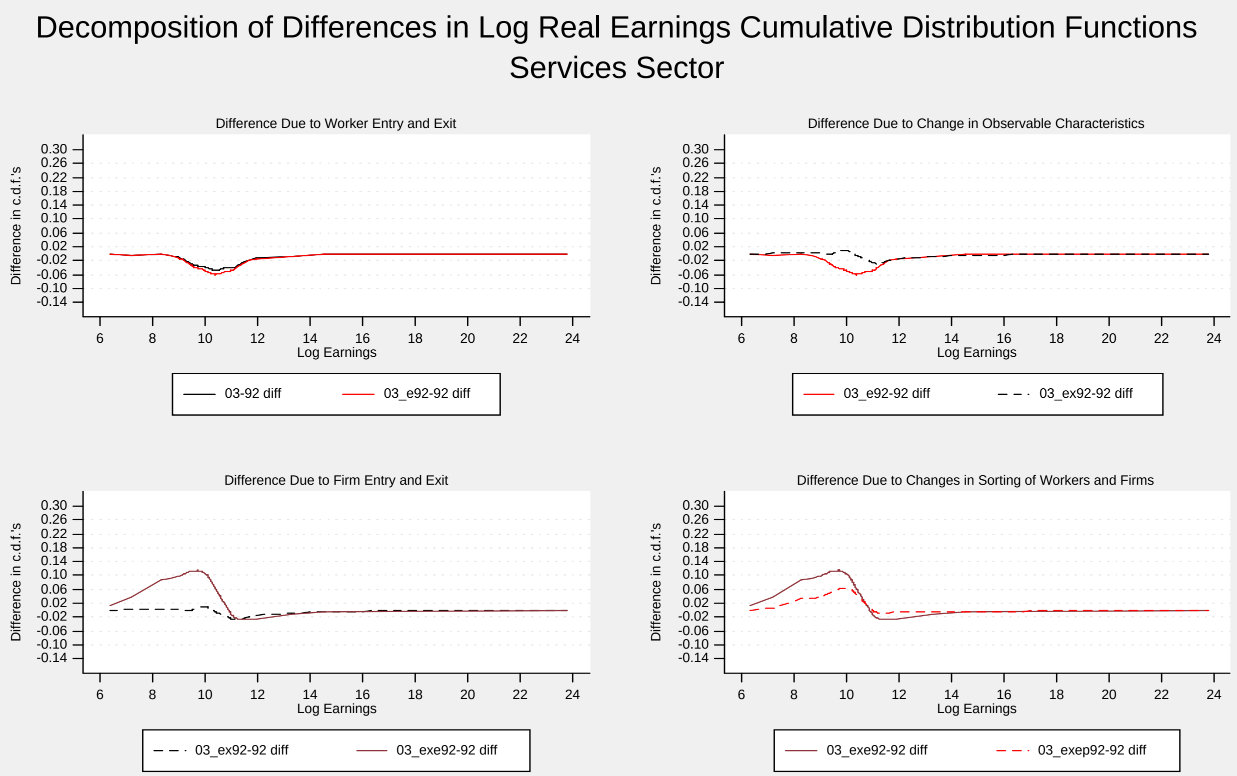
Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

Figure B8



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.

Figure B9



Source: Longitudinal Employer and Household Dynamics data for CA, IL, MD, and NC.
Notes: The notation 03_e92, 03_ex92, 03_exe92, and 03_ixep92-92 refers to the, set of workers, set of workers - observable characteristics, set of workers - observable characteristics - set of firms, and set of workers - observable characteristics - set of firms - conditional distribution of theta given psi, respectively, the same as in 1992.