Preliminary Draft

Decomposing the Sources of Earnings Inequality: A Five Industry Study

December 2004

Fredrik Andersson, Elizabeth Davis, Matthew L. Freedman, Julia I. Lane, Brian P. McCall, L. Kristin Sandusky

1. Introduction

Disentangling the sources of earnings inequality has long been a challenge. The literature has provided both demand and supply side explanations – such as 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, much of what is known is derived from analyses of surveys of workers. In this paper we use matched administrative data on both firms and workers from five industries spanning the spectrum of economic activity to partially remedy this gap in knowledge.

This paper advances knowledge on the way in which firms affect changes in earnings inequality in several ways. Indirect evidence suggests that production processes have become more biased towards using skilled workers. Most changes in production modalities within industries are due to firm entry and exit rather than due to changes in the production processes of existing firms, and here we directly quantify the contribution of such job reallocation to earnings inequality. Other evidence suggests that changes in workforce composition, and in particular workforce skill and experience, may also be important, and hence we quantify the impact of worker reallocation on the earnings distribution. Finally, there is also ample evidence that different firms pay different wage premia – whether it is due to unionization, compensating differentials, or rent-sharing – and thus the reallocation of workers across firms can potentially contribute to changes in the earnings distribution. We directly examine the degree to which changes in the matching of workers and firms affect earnings inequality.

This paper, which examines changes in the earnings distribution over a relatively short period of time, does not address the issue of changes in the price of skill due to skill-biased technological change. Rather, because our focus is on the contribution of firms to changes in earnings inequality, we focus specifically on explaining changes in the earnings distribution within, rather than across, five industries - financial services, retail food, semiconductors, software, and trucking. These industries span the gamut of the economic spectrum: low-wage and high-wage, union and non-union, manufacturing and services, regulated and competitive. We use our new employer-employee data first to document the levels of worker and firm reallocation across the industries, and then to tease out the separate contributions of firm turnover, worker turnover, and firm pay strategies to different parts of the earnings distribution. After a brief review of the literature, description of the industries analyzed in the study, and

Deleted: , particularly workforce skill and experience

Deleted: ;

Deleted: ;

Deleted: ;

discussion of the data, we present some basic empirical facts that were about changes in earnings distributions in these five industries. Since we are interested in ascertaining the relative importance of various factors in explaining changes over time in the earnings distribution, we then develop an econometric method for decomposing changes in earnings distributions when employer-employee matched data are available. The remaining sections of the paper describe the results of performing these decompositions and summarized what we have learned.

Deleted: following

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 contributing factors interact in very different ways to generate similar outcomes. For example, while changes in the workforce are the main contributing factor to earnings inequality changes in the semiconductor industry, the changing structure of the firm is the critical component in financial services. Even in industries in which the earnings distribution did not change, such as retail food and software, the lack of observable change actually resulted from powerful changes in offsetting factors. The one constant across all of our industries is that, 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 is only on the consequences of within-industry reallocation, these results suggest caution when searching for simple answers to questions raised by complex economic phenomena.

Deleted: – like

Deleted: –

Deleted: the offsetting effects of different

Deleted: ;

2. Background

Despite a vast literature that attempts to disentangle the sources of increased inequality in the 1980s and 1990s, there is still not complete consensus. Although 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), Card and DiNardo argue that changes in the minimum wages and declines in unionization were important factors. Fortin and Lemieux (1997) also examine the impact of de-regulation in the 1980s predominantly in the transportation, communication and banking industries. Lemieux (2004) also points to changes in the composition of the workforce as an important contributor to earnings inequality changes.

Although almost all of the literature is driven by an analysis of worker-based surveys – most notably the Current Population Survey, there is some evidence that suggests that changes in the distribution of wages across industries 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 (1995, 1991) find that firm size is an important determinant of wages, and that wage inequality has shifted both among and within manufacturing plants. Abowd, Kramarz and Margolis (1999) examine the relationship between "high wage firms": firms that seem to pay a wage premium or markup; and "high wage workers," who earn a premium, mostly likely as a return to unobserved skill.

Each of these analyses points to the importance of firm effects on the earnings distribution. And, because firms are both heterogeneous in the way in which they organize themselves, and very persistent in their production techniques, firm entry and exit likely play important roles in shaping this distribution. Firms deliberately choose different workforces, different pay structures, and different ways of producing goods and services, and this choice is remarkably persistent over time (Haltiwanger, Lane, Spletzer, 2001; Abowd, Kramarz, and Margolis (1999); Foster, Haltiwanger, and Krizan (1998); Dunne, Foster, Haltiwanger, and Troske (2000)).

Certainly there is ample room for changing firm and industry structure to alter the economic landscape within an industry. Davis, Haltiwanger, and Schuh (1996) document the large magnitude of job creation and destruction and the dominance of idiosyncratic factors in accounting for the observed large pace of job reallocation. 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. This entry and exit pattern may well be due to firm learning and selection of new technologies (Jovanovic (1982), Ericson and Pakes (1995)), which Aghion (2001) argues is an important determinant of earnings inequality. Firms that fail to adopt appropriate technologies are less likely to survive; Foster,

Deleted: are
Deleted: to
Deleted: –
Deleted: ,
Deleted:
Deleted: ,
Deleted: ,
Deleted: and
Deleted: , amongst others
Deleted: Certainly f
Deleted: :

Haltiwanger, and Krizan (1998) note that a common finding in the emerging literature is that low productivity is an excellent predictor of exit.

3. Insights from Five Industries

Our five industries provide us in a very real sense with a microscope that enables us to decompose these very different factors – particularly firm entry and exit. The industries have experienced different economic forces, and combination of forces, during the 1990s. The trucking industry has seen enormous deregulation, resulting in the replacement of large, unionized firms, by smaller, non-unionized firms. The retail food industry, while becoming steadily more concentrated in the 1990s also faced new types of competition from other types of retailers, such as Wal-mart and Costco. The story of semiconductor manufacturing is one of steady dis-integration as design, fabrication and assembly split up, and new fab-less design companies emerged. And in financial services, the number of banks dropped by 30% in the 1990s, while enormous waves of mergers and acquisitions have buffeted the industry.

This turbulence is reflected in the amount of worker turnover. Only 47% of workers who were in financial services in 1992 were still there by 1998, and only 30% of 1992 workers in the software industry remained there until 1998. The changing structure of the industry affected job turnover as well - 30% of semiconductor workers in 1992 worked for firms that no longer existed by 1998. That compares with 40% of workers in the trucking industry. The pay structure of firms in some industries changed substantially, although in others it changed very little. The average firm in software paid 34% above the economy-wide average in 1992; by 1998, this had risen to 43% more. Yet there was little change in the semiconductor industry, and only slight changes in retail food.

Deleted: ;

Deleted: -

Deleted: ,

Deleted: as did

Deleted: as well – in others not at all.

Deleted: software

The five industries used in this study span the economic spectrum – representing industries in manufacturing and services, high technology and low technology. In addition, the industries have been affected in widely differing ways in the 1990s by the factors that have been identified in the literature – technological change, changes in the demand for skill, changes in industry structure, changes in the levels of competition from imports, the erosion in the minimum wage and decline in union density.

In the *financial services* industry, a number of broad trends have been important. Most notably, the industry has seen steady deregulation culminating in repeal of the Glass-Steagall

Deleted: ,

Deleted: notably:

Act, massive technological change (including the introduction of the ATM, increase in use of call centers, and introduction of on-line services), and heightened competition within and between sub-industries. Trade and unionization effects are much less important. National micro-data suggest the relative proportion of individuals employed in higher-paying occupations in the industry has increased while the share of those in lower-paying positions has decreased gradually (Demsetz 1997). Union membership density in the banking industry, for example, is less than one percent.[1]

Substantial technological change has occurred in the *retail food* industry, but it is generally not "skill-biased" toward higher skilled workers. The use of certain technologies such as scanning may have increased the productivity of workers like cashiers, who are generally low skilled. King and Park (2002), however, found little evidence of store-level productivity gains due to adoption of information and internet-based technologies. In addition, the industry has undergone major changes in terms of market structure and competition both from restaurants and from "power retailers" outside of food retailing such as Wal-Mart (Wal-Mart is now the largest food retailer in the United States). Market shares held by leading retailers have risen markedly in the past decade, and consolidation has likely affected pricing behavior among firms in the industry. However, whether increased concentration has had any implications for wages or wage inequality is less clear. Finally, this industry, which employs a large number of low-wage workers, has seen quite substantial changes in the real value of the minimum wage in the 1990s, which has come amid a widespread decline in unionization, increases in store hours and size, and increased use of scanning technology (Budd and McCall, 2001).

Two main trends have dominated the *semiconductor industry*: globalization and technological change. Although rapid growth of the semiconductor industry over the 1990s has put upward pressure on earnings, especially for certain types of engineers, the spread of semiconductor activities to less developed countries and liberal immigration policies for high-tech workers have put downward pressure on earnings in the United States. Technological change has been quite lumpy in semiconductor fabrication plants, which display discrete jumps in technology as each generation of plant implements a host of technological advancements. Skill requirements, training, work organization, occupational distribution are highly sensitive to technology inputs within the plant, although compensation is less sensitive to technology inputs.

Deleted: act
Deleted: ;
Deleted: ;
Deleted: -
Deleted: .
Deleted: .
Deleted: combined with
Deleted: .

[1] For a discussion of industry change, and a matched pair case study, see Hunter (1999) and Hunter et al. (2001).

Deleted: -

Case study evidence on within-plant wage and employment trends suggests a key role for plant or firm specific compensation decisions, and these are likely to be driven by both plant characteristics (i.e. technology, product, product mix) and workforce composition.

Deleted: –

Deleted: this is

Deleted: level

The issues in the *trucking* industry are very different than those in the other industries. First, workforce composition differences that have been so marked in other industries are much less important in the trucking sector, simply because there are very few women or minorities in trucking. In addition, firm specific factors are likely to be much more important: many of the workers in the industry earn most of their income in a piecework fashion based on mileage rates, which are similar within firms, but differ across firms. The most likely driver of wage change is differences across firms – some firms take a "high road" approach – with higher quality freight, carried by higher skilled and paid drivers, others take a "low road" approach. There have also been substantially changes in compensation practices over time: a case study of the J.B. Hunt Company revealed that there was a conscious decision to raise mileage rates dramatically in an effort to reduce turnover and improve productivity. In effect, Hunt was able to transform itself from a 'low road' to a 'high road' firm. The results of this change have been exactly what the firm hoped for; namely, lower turnover, higher productivity, and increased profitability.

Deleted: ,

Deleted: s

So, in sum, changing industry structure and changing firm dynamics are common themes in each of the industry. Other factors have buffeted the industries in very different ways. One industry - semiconductors – has been subject to substantial import competition; technological change in two sectors - financial services and retail food - should affect earnings, but may not affect them in one other sector - semiconductors. Firm-specific compensation practices are important in trucking and semiconductors, while unionization and minimum wage changes are important only in retail trade. Changing industry structure is a common theme across all industries.

Deleted: two industries – financial services and retail food – have been affected by technological change that should affect earnings differentials, while one technological change in semiconductors likely will not.

Deleted:

Deleted: ;

4. The LEHD Data

In this paper, we take advantage of a new database created by the Longitudinal Employer and Household Dynamics (LEHD) Program at the U.S. Census Bureau. These data enable us to match workers with past and present employers, together with employer and worker characteristics (Abowd, Lane and Prevost, 2000). This database consists of quarterly records of the employment and earnings of almost all individuals from the unemployment insurance

Deleted: the

Deleted: ,

Deleted: that

Deleted: s

systems of a number of US states in the 1990s[2]. These type of data have been extensively described elsewhere (Haltiwanger, Lane and Spletzer, 2000), 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 the full universe of employers and workers[3], movements of workers across earnings categories and across employers can be accurately tracked. 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, there is substantial internal evidence from the LEHD program that workers do not accurately identify their industry at the major industry level (preliminary results suggest that the accuracy is only about 65%), much less at the detailed industry level.

Deleted: ,

Deleted: ,

Deleted: -

Because all jobs held by all workers 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. Because both of these measures describe different facets of the labor market, we will use both in our analysis. In particular, 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. We also compare earnings measures in each industry looking at all jobs with the measures based on dominant jobs to illustrate both within and across industry differences.

Deleted: -

Deleted: analyse

[2] 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.

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

Earnings refer to quarterly earnings, and we have no information on either wage rates or hours and weeks worked. There are a large number of ways the LEHD data may be used to characterize the distribution of annual earnings and of human capital for each sector. Both the measure used and the observations included in these distributions may be varied. Our particular measures and sample selection not only help to illustrate similarities and differences in the distribution of earnings and human capital across sectors and over time, but also shed light on whether earnings and human capital differences can be found across different groupings of workers. In addition, these measures help us to understand the labor force composition of each industry in terms of the share of full time workers as well as job changers and multiple jobholders. More detail on the LEHD data is provided in the appendix.

Deleted: Although the primary objective is

Deleted: both

Deleted: we also hope to learn

Deleted: The statistics provided in these datasets should provide information on the distributions of earnings and human capital for various groups of workers.

Deleted: should also

5. Basic Empirical Facts for the Five Industries

a) Describing the earnings distribution

An important strength of this analysis is the ability to compare earnings levels, differences, and changes across five very different detailed industries. However, as the literature has emphasized, the identification of trends in inequality often depends heavily on which earnings measures are used (Lerman et al.). We used both the LEHD data as well as data from the Current Population Survey's (CPS) monthly outgoing rotation group samples to establish some basic facts about the industries, and report them in Table 1. We examine three basic earnings measures – annual earnings from workers whose main jobs are in the industries; annual earnings from all jobs in the industry (which will substantially affect those industries, such as retail food, whose jobs are often second jobs for their workforce); and average weekly earnings from the CPS.

Deleted: ,

Deleted: and

The differences across industries in terms of earnings levels are quite stark. For example, median earnings in software are over three times as high as in retail food if the measure is earnings from dominant jobs, and the same holds at both the 90th and 10th percentiles; but ten times as high if the measure is earnings when the job in the industry is the one from which the worker derived the most income. The pattern is exacerbated when we simply examine the earnings distribution based on all jobs (including secondary jobs) – the earnings differences for the lowest earnings workers in retail food are only one tenth of those in software.

It is common in the study of earnings inequality (and changes) across industries to compare log earnings, and that is what is presented in the second set of panels in Table 1. We immediately see that trends in earnings inequality are quite disparate across our five industries. Earnings inequality actually declined or stayed roughly the same in two industries: retail food and trucking. By contrast, earnings inequality increased (for dominant jobs) in financial services, semiconductors and software. However, some of this is clearly due to changes in workforce composition; – when we narrow the focus to examine the change in inequality for 30-50 year old males (who are also less likely to hold multiple jobs and work part-time), earnings inequality increased in every industry except trucking, and the differences between dominants jobs and all jobs (due to multiple job holding) are substantially reduced. The most remarkable result from Table 1, however, is the lack of volatility in the earnings distribution given the enormous amount of economic change that firms in each of these industries has faced.

Deleted: The best way to

Deleted: is

Deleted: S

Deleted: however

Deleted: simply

Figures 1 through 5 plot empirical estimates of the cumulative distribution functions of annual earnings for each of the five industries for 1992 and 1998, and confirm the results presented in the table. While there has been modest change in the distributions of earnings in all high-skill industries (financial services, semiconductors, and software), there is on a slight change in the earnings distributions of workers in the lower-skilled industries of retail food and trucking.

Deleted: Financial

Deleted: Services

Deleted: Semiconductors

Deleted: Software

Deleted: Retail

Deleted: Food

Deleted: Trucking

b) The role of worker and firm reallocation

The amount of economic turbulence in each industry is substantial, and documented in Table 2. Briefly, about 10% of jobs are reallocated from firm to firm within each industry in a year – the sole exception being semiconductors, with only a 6% reallocation rate – and these rates have increased since 1994. Workers are reallocated across firms at even higher rates of around 20% - although again, the reallocation is lower in semiconductors. Matching this is an annual accession rate of around 40% in each industry. The potential for this kind of worker reallocation to change earnings inequality is evident from the fact that for most industries, approximately 80% of accessions come from outside the industry. The only industry that bucks this trend is financial services, where only 60% come from outside the industry. Similarly, about 5% of workers in each period find themselves working for new businesses – the firm entry and exit rates range from about 2% to 7% of employment.

Deleted: – except for financial services, What impact does this have on the workforce? The bottom line on changes in the workforce composition between 1992 and 1998 is described in Table 3. Fewer than half of the workers who were in the financial services industry in 1992 were still there in 1998.; That compares with fewer than one third of the software industry workers. The differences in retention patterns across the earnings distribution are quite remarkable as well. The industry with the lowest retention rate for workers in the bottom percentile – retail food – had the highest retention rate for workers in the top percentile. It is clear that the relatively high average turnover rates in the retail food industry are driven primarily by very high turnover at the bottom end of the distribution. In the software industry, by contrast, long-term retention rates at both ends of the earnings distribution are quite similar.

Deleted: less

We would like to investigate what underlying forces are responsible for these observed changes (or lack of changes) in the earnings distributions across the Sloan industries – particularly focusing on the the amount due to worker entry and exit from the industry, changes in the worker's observable characteristics, firm entry and exit from the industry, and intra-industry reallocation of different types of workers among different types of firms. The next section develops the econometric methods we use to perform this decomposition.

6. 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 is 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).

Deleted: See

Let θ be a variable representing an individual's (unobserved) productivity that is assumed to be constant over time. Further, let ψ represent a firms (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.

$$dP_t(\varepsilon,\theta,\psi,x) = f_t(\varepsilon,\theta,\psi,x)dt \,. \qquad (1)$$

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 firm and worker there are stayers (*s*), leavers (*l*) and new entrants (*n*) Now, we can rewrite the joint distribution in (1) at time period 1 as

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

Deleted:

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

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

Deleted: 2

Deleted:

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 be no exit and entry of workers. In that situation, (3) becomes

Deleted: While

Deleted: t

Deleted: i

Deleted: 2

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

Deleted: 3

Deleted: ¶

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 firms that ending up staying in the industry.

Deleted:

Deleted:

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

Deleted: ¶

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

Deleted: ¶

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

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

Deleted: 4

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

Deleted: 3

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

Deleted: 5

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:

Deleted: of had the set

Deleted: been

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

Deleted: 6

Finally, after we have restricted the set of firms and workers to be the same as in period 1, we can still 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,

Deleted:

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

so we can define

Deleted: ¶

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

where the a superscript refers to holding the allocation of workers to firms constant. (

Deleted: 6)

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,

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

Deleted: 7

We can then define the counterfactual probabilities by

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

Deleted: 8

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

Deleted: 9

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

and

Deleted: ¶

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

Deleted: 10

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

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

Deleted: 11

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

Deleted: :

$$\begin{aligned} &E_2(r(y)) - E_1(r(y)) = \\ &\int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx - \int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_1(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx = \\ &\left(\int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx - \int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2^{w}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx \right) + \\ &\left(\int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2^{w}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx - \int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2^{w,x}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx \right) + \\ &\left(\int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2^{w,x}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx - \int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2^{w,x,e}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx \right) + \\ &\left(\int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,s))\, f_2^{w,x,e}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx - \int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_2^{w,x,e,a}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx \right) + \\ &\left(\int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,s))\, f_2^{w,x,e,a}(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx - \int_{\{\varepsilon,\theta,\psi,x\}\in D} r(g(\varepsilon,\theta,\psi,x))\, f_1(\varepsilon,\theta,\psi,x)\,d\varepsilon\, d\theta\, d\psi\, dx \right) \end{aligned} \tag{16}$$

where D denotes the domain of $(\varepsilon, \theta, \psi, x)$. Note that in (11) we have $r(y) = I(y \in A)$ where I is the indicator function.

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. Recall that a linear panel data model with fixed firm and individual effects such as described in Abowd, Kramarz, and Margolis (1999) was used to estimate the determinants of log earnings using the entire LEHD data. Thus, the function g has the following form

$$y_{it} = g(\varepsilon, \theta, \psi, \mathbf{x}) = \mathbf{x}_{it}\boldsymbol{\beta} + \theta_i + \psi_{j(i,t)} + \varepsilon_{it} \qquad (17)$$

Deleted: 12

Thus, we assume that no structural change has occurred over this time period.

Since our focus is on changes in earnings distributions over time within five particular industries, from this estimation we need to calculate $dP_t(\varepsilon, \theta, \psi, \mathbf{x})$ for each industry. In the LEHD data all exogenous variables (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 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 $\mathbf{x}$ within each of the five industries in our sample. Since the LEHD data set is extremely large, we accomplished this task by discretizing the 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:

$$y_{it}^d = \mathbf{x}_{it}\boldsymbol{\beta} + \theta_i^d + \psi_{j(i,t)}^d + \varepsilon_{it}^d . \qquad (18)$$

Deleted: E

Deleted: has been discretized

Deleted: which

Deleted:

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

7. Decomposition Results

In this section we discuss the results of the earnings decompositions for each of the five industries investigated in this paper. Table 4 presents the results of decomposing various earnings statistics changes while Table 5 presents the results of decomposing the Kullback Leibler distance measure (see DiNardo, Fortin and Lemieux, 1996) which measures the distance between two distributions.[4] For each of the five industries, Figures 6-11 present sequential decompositions of the difference in 1998 and 1992 cumulative distribution functions for annual earnings 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.

Deleted: See

The first striking result upon examining Table 4 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 retail food industry, for example, the 90/10 ratio dropped only slightly from 1.70 in 1992 to 1.62 in 1998. An examination of gross effects would suggest that all the change occurred in the 90th percentile – because log earnings in the 10th percentile was 9.10 in 1992, and 9.10 in 1998, while log earnings in the 90th percentile in 1992 was 10.80 but had dropped to 10.72 by 1998. 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.

Deleted: Retail

Deleted: Food

Deleted: And an

Reading across the 10th percentile rows in each of the industry panels in Table 4, a comparison of the first and second columns reveals that, despite the high levels of worker churning documented in Table 2, the churning was among workers of the same average skill level (Θ), resulting in no change in earnings for workers in that decile. A similar result is found for workers in the 90th percentile. That this is true in every industry suggests that, by and large,

---

[4] The Kullback-Leibler measure of the distance between two distributions f_1 and f_2 is

$$J = \int \left[f_1(w) - f_2(w) \right] \ln \frac{f_1(w)}{f_2(w)} dw.$$

workforce quality within each industry is quite persistent, which is consistent with work by Haltiwanger, Lane and Spletzer (2001).

An analysis of the second and third columns reveals that, holding Θ constant, the aging of the workforce (and the associated returns to experience) acted to reduce earnings (NOTE: Is this correct or reversed?) at both ends of the distribution by essentially similar orders of magnitude, so had very little impact on the 90/10 ratio. 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 earnings in the financial services industry by over 25 log points, compared with under 15 log points in retail food.

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. Column four (compared to column three) shows that if no firm entry or exit had occurred between 1992 and 1998, 90/10 earnings differentials would have swung by as much as 53 points in software and 49 points in financial services – the two industries which have been characterized by enormous amounts of economic turbulence. Notably, firm entry and exit typically acted to increase earnings at the bottom end of the distribution and decrease 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 negative impact on earnings for workers in the 10th percentile workers in all industries, but only had a negative for workers in the 90th percentile in retail and trucking.

Turning to more aggregate summary statistics, the Kullback-Leibler measure of the distance between the 1992 and 1998 earnings distributions, which are presented in Table5, there is some evidence that worker entry and exit tended to shrink the distance between the 1992 and 1998 earnings distributions in the software and semiconductor industries. Changes in observable characteristics tended to widen the distance between the two distributions for the financial services, semiconductor, and software industries while narrowing the difference in the retail food and trucking industries. Finally, firm entry and exit narrowed and sorting of workers among firms widened the distance between the 1992 and 1998 earnings distributions across all industries.

Deleted: --

Deleted: --

Deleted: was

To summarize, earnings distributions changed differently among the five industries we examined over the 1992 to 1998 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 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 of the five industries.

8. Conclusions

In this paper, we used unique employer-employee matched panel data from the Longitudinal Employer Household Dynamics Program to explore changes in the earnings distributions across five industries. We hoped to advance knowledge on the way in which firms affect changes in earnings inequality by focusing on within industry changes in earnings inequality. Because the evidence suggests that a major way in which industries change production modalities is through firm entry and exit rather than changes in the production processes of existing firms, we directly quantified the contribution of such job reallocation to earnings inequality. We also directly examined the way in which changes in workforce composition -- particularly workforce skill and experience -- affecting the industry specific earnings distribution. Finally, we directly examined the degree to which changes in the matching of workers and firms affect earnings inequality. All the industries that we examined had substantial amounts of worker entry and exit (in common with other U.S. industries), albeit very different experiences with respect to firm entry and exit, as well as very different workforce skill levels.

Deleted: ,

Deleted: data

Deleted: ,

Deleted: which

While there were differences across industries in the magnitudes and directions of change in various aspects of the earnings distribution over the 1992 to 1998 time period, our earnings decompositions found that some 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 1992, there were enormous, albeit offsetting, changes in the factors contributing to earnings change.

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 ample opportunities for firms to change their workforce composition, industry workforces remain, by and large, very similar.

On the other hand, firm entry and exit tended to reduce the dispersion of earnings for all industries. While 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, sorting of workers based on the "human capital" measures over time tended to increase the dispersion of industry earnings distributions between 1992 and 1998.

Deleted:

Our findings suggest that even when earnings distributions seem superficially not to change, the large 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. Clearly, future research could examine other industries to see whether these general trends are more pervasive or are limited to the particular industries examined in this paper. Nevertheless, 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.

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

Aghion, Philippe (2001) "Schumpeterian Growth Theory And The Dynamics Of Income Inequality", *Econometrica*, forthcoming

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

Budd, John W. and Brian P. McCall. 2001. "The Grocery Stores" Wage Distribution: A Semi-parametric Analysis of the Role of Retailing and Labor Market Institutions," *Industrial and Labor Relations Review*, 54(2): 483-501.

Brown, Clair, ed. 1997. *The Competitive Semiconductor Manufacturing Human Resources Project: Final Report (Phase I)*. CSM–38. Institute of Industrial Relations, University of California, Berkeley.

Brown, Clair, and Benjamin Campbell. 1999a. "The Impact of Technological Change on the Internal Labor Market Structure of a Semiconductor Firm." Sloan CSM-HR Working Paper, University of California, Berkeley.

Brown, Clair, and Benjamin Campbell. 1999b. "The Impact of Technological Change on Work and Wages." Center for Work, Technology, and Society Working Paper, University of California, Berkeley. Forthcoming in *Industrial Relations.*

Brown, Clair and Benjamin Campbell. 2001 "Technical Change, Wages, and Employment in Semiconductor Manufacturing" *Industrial and Labor Relations Review 54*(2A).

Brown, Clair, Greg Pinsonneault, and Dan Rascher. 1999. "The Use of New Technology and Human Resource Systems in Improving Semiconductor Manufacturing Performance." Center for Work, Technology, and Society, University of California, Berkeley.

Campbell, Benjamin, and Vince Valvano. 1997. "Wage Structures in the Electronics Industry." In Clair Brown, ed., *The Competitive Semiconductor Manufacturing Human Resources Project: Final Report (Phase I).* Berkeley, Calif.: Institute of Industrial Relations and Engineering Systems Research Center.

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

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

Demsetz, Rebecca S. 1997. "Human resource needs in the evolving financial sector." *Current Issues in Economics and Finance* Vol. 3, No. 13. New York: Federal Reserve Bank of New York.

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

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.

Haltiwanger, Lane and Spletzer. 2001.

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

Hunter, Larry W. 1999. "Transforming retail banking: Inclusion and segmentation in service work." Pp. 153-192 in Peter Cappelli (ed.), *Employment Strategies: Why Similar Employers Manage Differently*. Oxford University Press.

Hunter, Larry W., Annette Bernhardt, Katherine L. Hughes, and Eva Skuratowicz. 2001. "It's Not Just the ATMs: Technology, Firm Strategies, Jobs, and Earnings in Retail Banking." *Industrial and Labor Relations Review 54* 2A.

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, L. and D. 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.

King, Robert and Tim Park. 2002. "Modeling Scale Economies in Supermarket Operations: Incorporating the Impacts of Store Characteristics and Information Technologies," working paper.

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.

Levi, F. and R.J. Murnane. 1992. "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," *Journal of Economic Literature*, 30: 1333-1381.

Table 1: Earnings Levels, Differences and Changes: 1992 – 1998							
				90/10 ratio			
Sector	90th	Median	10th	(Levels)	(Logs)	Change in log inequality	
						All Workers	30-50 year old males
Financial Services							
Dominant Jobs	$98,708	$35,250	$16,296	6.06	1.80	0.16	0.25
All Job	$72,165	$21,313	$2,438	29.60	3.39	0.23	0.39
CPS Weekly	$1442	$520	$260	5.54	0.047		
Retail Food							
Dominant Jobs	$44,348	$19,912	$8,216	5.40	1.69	-0.01	0.06
All Job	$31,268	$4,923	$632	49.50	3.90	0.05	0.03
CPS Weekly	$700	$300	$120	5.833	.03		
Semiconductors							
Dominant Jobs	$118,993	$52,040	$22,578	5.27	1.66	0.15	0.21
All Job	$106,611	$40,716	$8,212	12.98	2.56	-0.02	0.07
CPS Weekly	1750	$769	$336	5.21	.54		
Software							
Dominant Jobs	$128,292	$62,328	$28,582	4.49	1.50	0.03	0.10
All Job	$101,657	$37,785	$5,076	20.03	3.00	0.04	0.10
CPS Weekly	$1634	$813	$346	4.73	.66		
Trucking							
Dominant Jobs	$55,342	$33,678	$15,094	3.67	1.30	-0.13	-0.06
All Job	$44,307	$10,872	$844	52.48	3.96	0.00	-0.14
CPS Weekly	$1000	$532	$250	4	.24		

Table 2: Worker and firm reallocation: Levels and Changes: 1998							
	Annual Job reallocation rate	Annual Worker Reallocation Rates		Proportion of accessions coming from		Firm reallocation rates (employment weighted)	
		Churning rate	Accession rate	Within industry	Manufacturing	Entry	Exit
Financial Services	10.48%	17.87%	37.39%	37.50%	1.96%	7.79%	7.68%
--Change since 1994	24.22%	14.60%	17.02%	23.52%	17.06%		
Retail Food	11.54%	22.21%	40.90%	17.08%	2.77%	5.45%	5.23%
--Change since 1994	36.44%	11.26%	1.77%	89.49%	-8.74%		
Semiconductors	6.23%	14.25%	29.69%	21.79%	15.78%	2.34%	2.53%
--Change since 1994	17.44%	36.37%	43.98%	75.16%	9.79%		
Software	11.37%	19.37%	41.74%	14.28%	7.13%	5.48%	4.74%
--Change since 1994	9.80%	12.34%	16.25%	23.33%	-30.53%		
Trucking	10.01%	18.30%	39.73%	21.84%	4.93%	3.33%	3.90%
--Change since 1994	9.52%	2.01%	3.06%	57.16%	9.79%		

Table 3: Proportion of 1992 workforce still employed in industry in 1998

	In 10th percentile in 1992	In 90th percentile in 1992	Overall
Financial Services	34.00%	56.33%	47.28%
Retail Food	13.44%	60.35%	35.10%
Semiconductors	36.69%	40.66%	38.82%
Software	20.91%	36.68%	30.47%
Trucking	15.06%	38.60%	34.57%

Table 4: Decomposing Changes in the Log Real Earnings Distribution

Log Real Earnings	1998	Counterfactuals: 1998 with 1992 Worker Entry and Exit	Change in Observable Worker Characteristics	Firm Entry and Exit	Sorting of workers and firms	1992
Financial Services						
Mean	10.61	10.62	10.46	10.34	10.44	10.49
Standard Deviation	1.28	1.28	1.41	1.66	1.46	1.27
10th Percentile	9.70	9.70	9.40	9.10	9.40	9.60
25th Percentile	10.10	10.10	9.90	9.70	9.84	10.00
50th Percentile	10.50	10.50	10.30	10.20	10.29	10.40
75th Percentile	10.98	11.00	10.78	10.70	10.70	10.80
90th Percentile	11.59	11.60	11.34	11.51	11.30	11.30
"90/10" Ratio	1.89	1.90	1.94	2.42	1.90	1.70
"90/50" Ratio	1.09	1.10	0.84	1.21	1.10	1.01
Retail Food						
Mean	10.02	10.03	9.88	9.77	9.88	10.04
Standard Deviation	1.16	1.15	1.22	1.33	1.22	1.19
10th Percentile	9.10	9.10	8.98	8.67	8.97	9.10
25th Percentile	9.50	9.50	9.30	9.20	9.31	9.50
50th Percentile	9.91	9.98	9.80	9.66	9.80	10.00
75th Percentile	10.40	10.43	10.20	10.20	10.22	10.50
90th Percentile	10.72	10.74	10.60	10.60	10.60	10.80
"90/10" Ratio	1.62	1.64	1.62	1.93	1.63	1.70
"90/50" Ratio	0.81	0.76	0.62	0.80	0.94	1.00
Semiconductor						
Mean	10.86	10.87	10.71	10.68	10.76	10.75
Standard Deviation	0.85	0.85	1.07	1.22	1.01	0.92
10th Percentile	10.00	10.00	9.79	9.63	9.87	9.90
25th Percentile	10.40	10.40	10.13	10.10	10.25	10.30
50th Percentile	10.90	10.90	10.60	10.60	10.70	10.70
75th Percentile	11.30	11.30	11.10	11.10	11.10	11.10
90th Percentile	11.70	11.72	11.53	11.57	11.52	11.50
"90/10" Ratio	1.70	1.72	1.74	1.94	1.65	1.60
"90/50" Ratio	0.80	0.82	0.63	0.97	0.92	0.80
Software						
Mean	11.05	11.06	10.92	10.88	10.87	10.87
Standard Deviation	0.95	0.95	1.08	1.25	1.13	1.04
10th Percentile	10.20	10.22	10.00	9.78	9.90	10.00
25th Percentile	10.60	10.60	10.43	10.40	10.40	10.40
50th Percentile	11.04	11.10	10.90	10.85	10.80	10.80
75th Percentile	11.40	11.44	11.30	11.32	11.22	11.20
90th Percentile	11.82	11.86	11.70	12.01	11.62	11.60

"90/10" Ratio	1.62	1.64	1.70	2.23	1.72	1.60
"90/50" Ratio	0.78	0.76	0.60	1.11	0.77	0.80
		Trucking				
Mean	10.41	10.42	10.34	10.28	10.30	10.39
Standard Deviation	0.90	0.90	1.03	1.21	1.07	0.91
10th Percentile	9.60	9.60	9.40	9.12	9.40	9.50
25th Percentile	10.08	10.10	9.90	9.80	9.90	10.00
50th Percentile	10.40	10.40	10.30	10.20	10.30	10.40
75th Percentile	10.70	10.70	10.60	10.60	10.60	10.70
90th Percentile	10.98	11.00	10.90	11.00	10.90	11.00
"90/10" Ratio	1.38	1.40	1.50	1.88	1.50	1.50
"90/50" Ratio	0.58	0.60	0.50	0.70	0.70	0.70

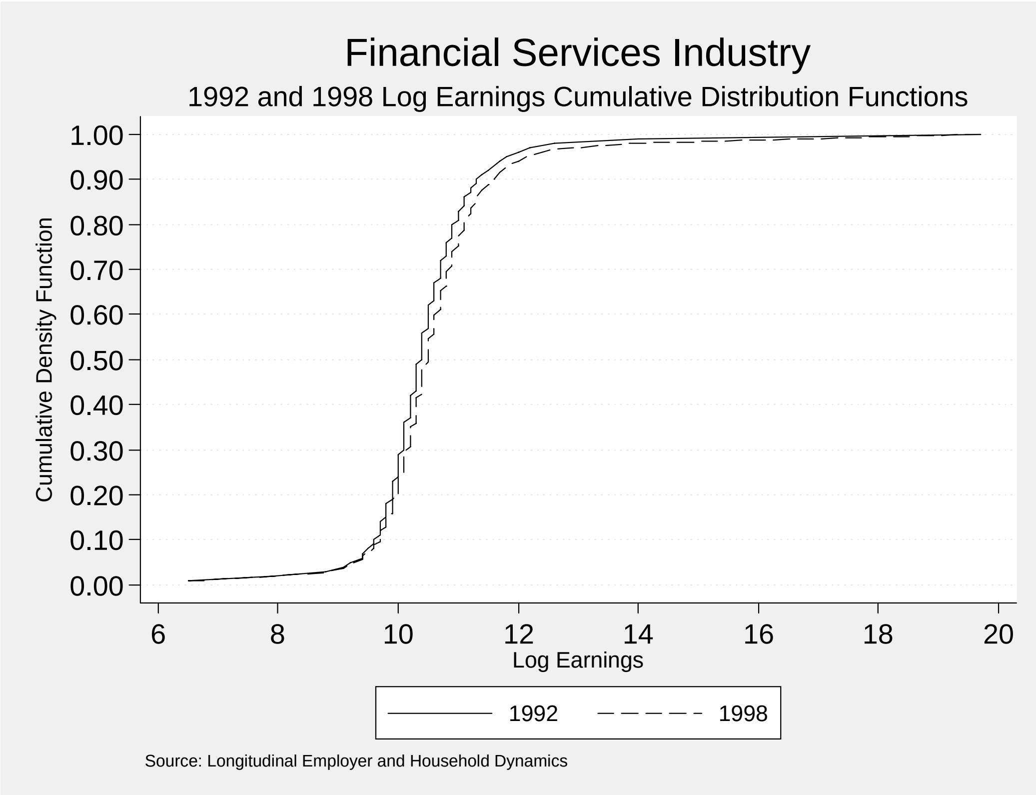
Source: Longitudinal Employer and Household Dynamics.

Table 5: Kullback- Leibler Distance Measure Decompositions

	Counterfactuals: 1998 with 1992				
Industry	**1998**	**Worker Entry and Exit**	**Change in Observable Worker Characteristics**	**Firm entry and exit**	**Sorting of workers and firms**
Financial Services	0.0355	0.0380	0.0342	0.1589	0.0392
Retail Food	0.0047	0.0042	0.0816	0.1961	0.0793
Semiconductors	0.0787	0.0843	0.0655	0.1252	0.0241
Software	0.1117	0.1287	0.0127	0.1371	0.0151
Trucking	0.0201	0.0203	0.0381	0.1451	0.0825

Source: Longitudinal Employer and Household Dynamics data.
Notes: All Kullback-Leibler distance measures are relative to the 1992 distribution.

Figure 1



Financial Services Industry
1992 and 1998 Log Earnings Cumulative Distribution Functions
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
8
10
12
14
16
18
20
Log Earnings
1992
1998
Source: Longitudinal Employer and Household Dynamics

FIGURE 2

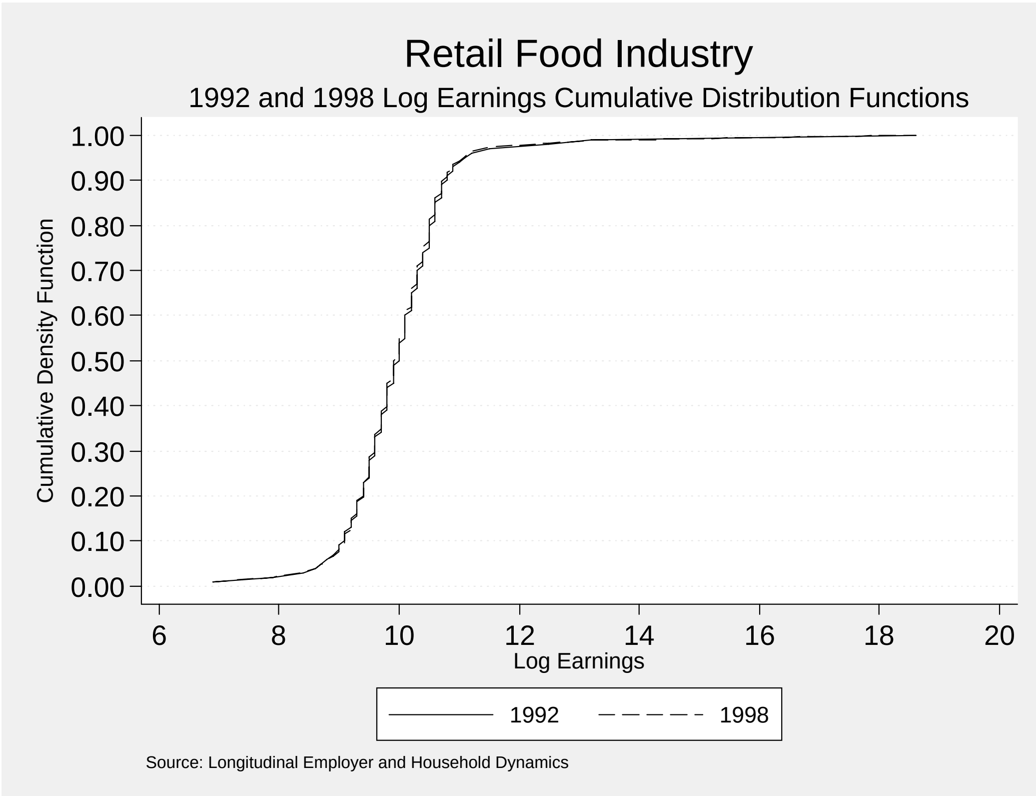

Retail Food Industry
1992 and 1998 Log Earnings Cumulative Distribution Functions
Cumulative Density Function
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
6
8
10
12
14
16
18
20
Log Earnings
1992
1998
Source: Longitudinal Employer and Household Dynamics

FIGURE 3

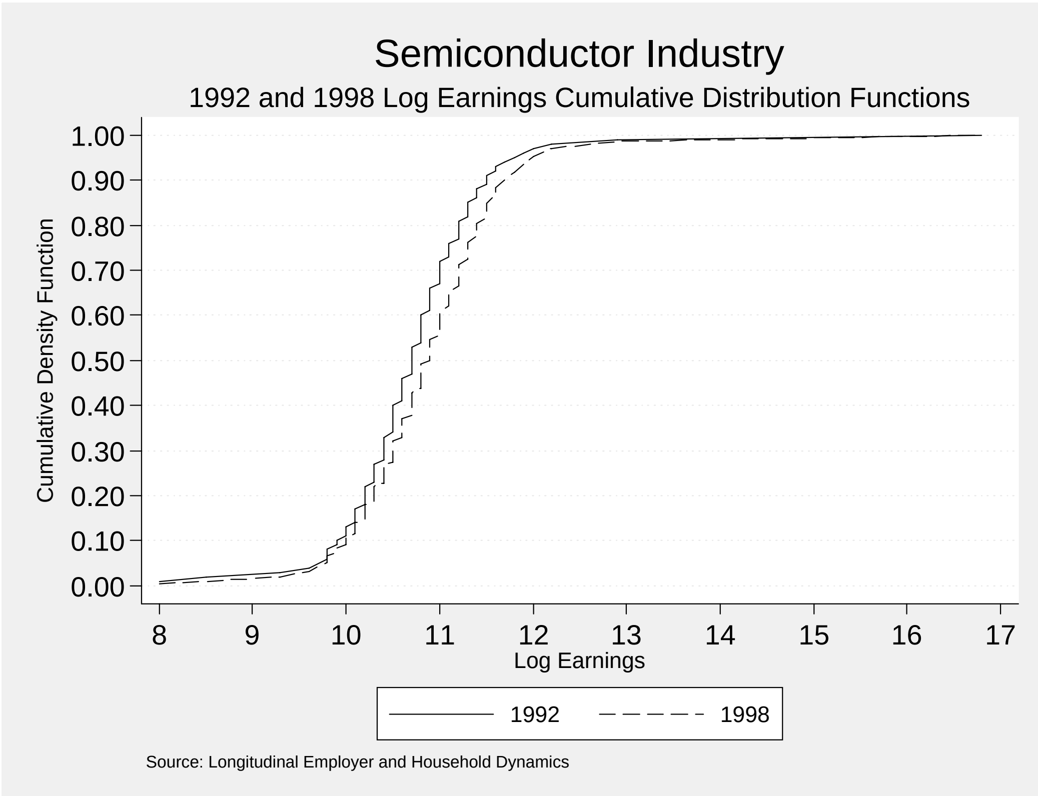

Semiconductor Industry
1992 and 1998 Log Earnings Cumulative Distribution Functions
Cumulative Density Function
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
8
9
10
11
12
13
14
15
16
17
Log Earnings
1992
1998
Source: Longitudinal Employer and Household Dynamics

FIGURE 4

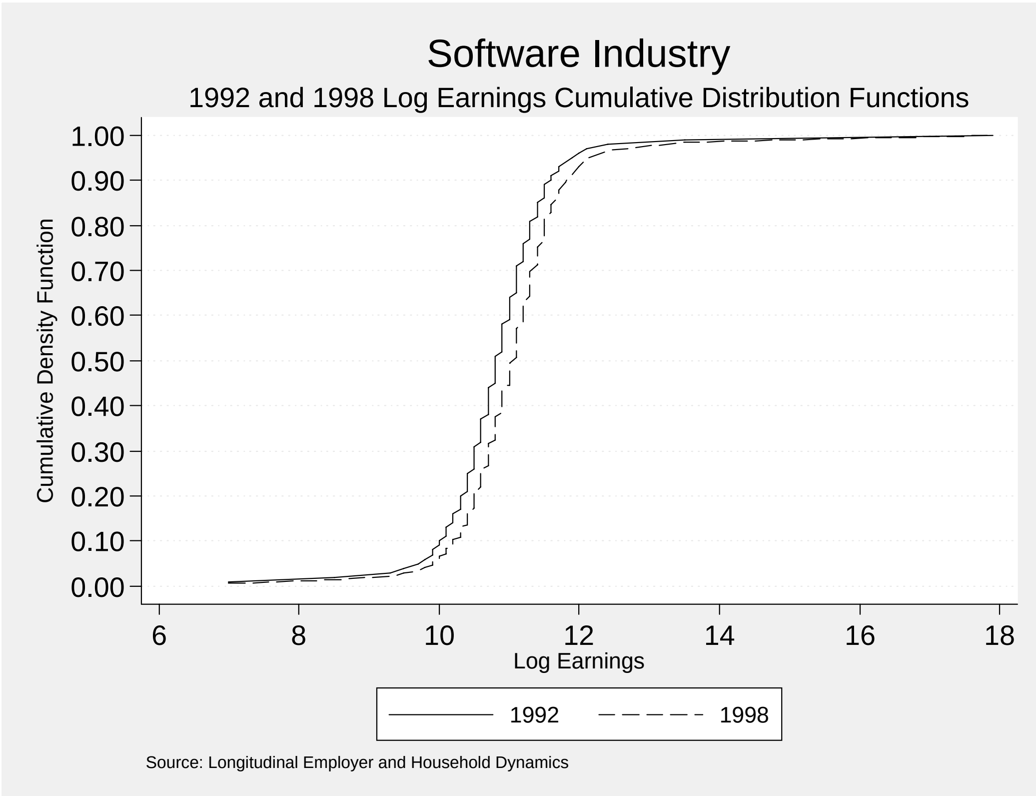
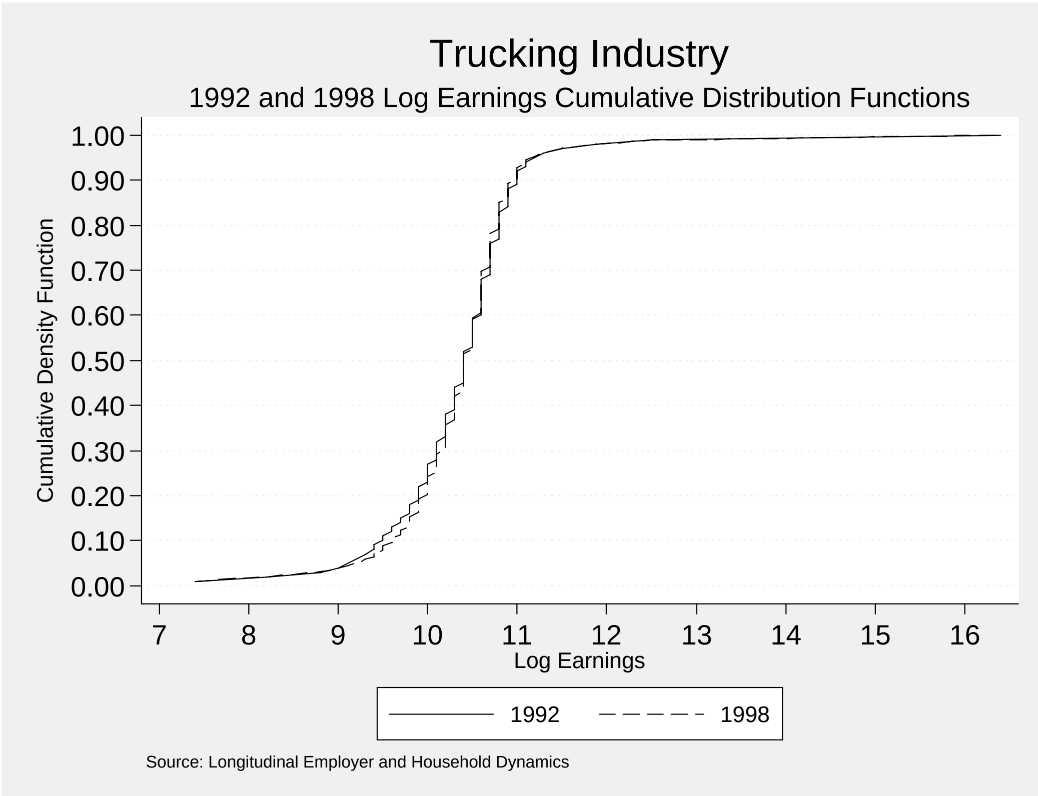

Software Industry
1992 and 1998 Log Earnings Cumulative Distribution Functions
Cumulative Density Function
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
6
8
10
12
14
16
18
Log Earnings
1992
1998
Source: Longitudinal Employer and Household Dynamics

FIGURE 5


Trucking Industry
1992 and 1998 Log Earnings Cumulative Distribution Functions
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
7
8
9
10
11
12
13
14
15
16
Log Earnings
1992
1998
Source: Longitudinal Employer and Household Dynamics

Appendix

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
Given turnover, etc, 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 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. This indicator is the same indicator used in the QWI and AWI statistical processing.

Measuring Earnings:
The statistics provide information on the distribution of two different earnings measures. The first measure is log real “annualized” earnings, and the second is the sum of quarterly earnings at each job.

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.

Log Real Annual Earnings
We measure annual earnings as simply the sum of earnings in each quarter. Depending on the method used this might be the sum of earnings on the dominant job only, or the sum of earnings for each job (multiple job holders).

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.

Which Workers Have a Human Capital Measure?
When computing the worker and firm fixed effects, only dominant job spells held by workers who:
1. are between 18 and 70 years old

2. imputed to work full time at that job
are used to compute these 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 Groups of Workers:
The datasets include year and sector specific earnings and human capital distributions summary statistics for 3 groups of workers or jobs.
Sample one includes all workers
a. with a "dominant employer" in that sector, where the dominant employer is defined as the employer contributing the most to the workers annual earnings, .
b. who are imputed to work full time in that year
c. who we have identified as likely working at the end of the first quarter of the year.
d. Who have earnings (real) of at least $250. in at least one quarter of the year.
e. The earnings measure used for sample one is log real annualized earnings.

Sample 2 includes all workers
a. with a "dominant employer" in that sector, where the dominant employer is defined as the employer contributing the most to the workers annual earnings, .
b. Who have earnings (real) of at least $250. in at least one quarter of the year.
c. The earnings measure used for sample two is log real annualized earnings.

Sample 3 include all workers
a. with a "dominant employer" in that sector, where the dominant employer is defined as the employer contributing the most to the workers annual earnings, .
b. Who have earnings (real) of at least $250. in at least one quarter of the year.
c. The earnings measure used for sample three is log real annual earnings (sum of quarterly earnings on dominant jobs).

Sample 4 (observation is now a JOB) includes all jobs held by workers
a. who have at least one employer in one of the five Sloan sectors and
b. who have earnings (real) of at least $250 in at least one quarter of the year.
c. The earnings measure used for sample 4 is log real annual earnings (sum of quarterly earnings on the job in question)

Comparison of Samples:

Dominant jobs Held By Full-time Point-in-time workers vs. All Dominant Jobs
Table 1 shows, for each sector and for each sample, counts of the number of workers with an earnings measure or human capital measures in 1997. All included workers have earnings of at least $250. in at least one quarter of the year. Recall that Sample One covers dominant jobs for workers imputed to work full time and to be working at the end of quarter one. The share of jobs that meet these restrictions (as a share of all dominant jobs, or Samples 2 and 3) varies notably by sector. In semiconductors, for example, eighty-two percent of all dominant jobs held at any point in the year are held by "full-time" workers who are working at then end of quarter one.

This fraction is substantially higher than in retail food. Here, only fifty-five percent of dominant jobs are help by full time workers employed at the end of quarter one. The shares for Trucking, Financial Services and Software all lay between these two extremes. As noted above, only workers who are imputed to work full time at least once in the period of time covered by the LEHD data have values for the human capital measures. Thus, these worker and job counts will almost always be smaller than worker and job counts for the earnings measures. Recalling that all human capital counts are conditional on a worker having worked full time at least once, it is not surprising that the current full time and point in time restrictions have a smaller impact on the count of dominant jobs for the human capital measures than we observe for the earnings measures. This is true in all sectors. For retail food, the share of dominant jobs held by full-time workers working at the end of quarter one is now seventy-eight percent, given that the worker is observed to work full time at least once. The share in Semiconductors, however, essentially remains unchanged by this condition.

All Dominant Jobs Vs. All Jobs

Regardless of whether we consider the count of jobs with an earnings measure or the count of jobs held by workers with human capital measures, the fraction by which the job count increases when we include all jobs as opposed to dominant jobs only is identical in each sector. A comparison of these counts across sectors provides information on the relative amount of job changing and multiple jobholding in each sector relative to other sectors. Surprisingly, the amount of variation across sectors in this fraction is not as striking. The count of all jobs is between 127 percent of dominant jobs (in semiconductors) and 140 percent of dominant jobs (in trucking). Lastly, recall that the count of dominant jobs for the human capital measures is smaller than the count of dominant jobs for earnings measures because only workers who have worked full time at least once have human capital measures. However, regardless of the difference in magnitude between the two counts, the fraction by which the job count increases when we include all jobs is identical. This suggests that, within each sector, workers who have worked full time at least once are perhaps neither more nor less likely to change jobs or to hold multiple jobs.

Table 1

Sector	Sample Restrictions	Earnings Counts	Share of all Dominant Jobs	Human Capital Counts	Share of All Dominant Jobs
Financial Services	1	1,067,455	0.68	1,057,018	0.77
	2 and 3	1,559,526	1	1,370,891	1
	4	2,149,686	1.38	1,897,232	1.38
Retail Food	1	497,692	0.55	473,932	0.78
	2 and 3	907,557	1	611,043	1
	4	1,248,475	1.38	847,663	1.39
Semiconductors	1	84,199	0.82	83,877	0.83
	2 and 3	102,789	1	100,552	1
	4	130,765	1.27	127,946	1.27
Software	1	171,838	0.71	171,203	0.75
	2 and 3	243,198	1	228,970	1
	4	327,317	1.35	309,013	1.35
Trucking	1	202,505	0.63	200,286	0.70
	2 and 3	320,848	1	285,189	1
	4	448,744	1.40	402,205	1.41
Numbers are for LEHD "short" panel - CA, IL, MD, NC for 1977					