ABSTRACT

Title of dissertation: ESSAYS ON EMPLOYER-EMPLOYEE RELATIONSHIPS AND FIRM PERFORMANCE

Hyowook Chiang, Doctor of Philosophy, 2005

Dissertation directed by: Professor John C. Haltiwanger
Department of Economics

Recently, the heterogeneity of workers has been documented and emphasized as a very important feature of labor in various economics fields. Labor is one of the key inputs in the production process, and it is quite different from other inputs in that no single worker can be treated the same as any other worker given their unique characteristics. Research on business performance has only recently begun to pay attention to worker heterogeneity. The most important reason for this is a lack of data that incorporate both business level information about production processes, and rich information about the individuals that work in each business. Many studies of productivity have treated all workers as homogeneous and use the total number of workers (or hours worked) as a single variable representing labor input. Studies using micro level databases could only differentiate labor input into production workers and non-production workers or skilled labor and unskilled labor. This dissertation exploits the heterogeneity of labor and variation in human resource management systems, and tries to understand their impact on firm performance and various outcomes.

I use a newly developed employer-employee matched database to examine the impacts of human resource practices on firm outcomes. First, I show that firms with lower rates of worker turnover have higher productivity and "learn" faster than those with higher worker turnover. Moreover, I develop new instruments to show that learning by doing and turnover have causal effects on productivity. Second, I show that firm performance is tightly linked with workforce quality and worker turnover. Strikingly, workforce quality and worker turnover independently contribute to firm survival even after taking productivity into account. Lastly, I assess the fit between firm-level internal labor markets and firm diversification in the U.S. financial services sector. Drawing on the "resource-based view" of firm strategy, I hypothesize that firms with stronger ILMs are more likely to diversify. I find that firms with lower churn, lower wage dispersion, and greater opportunities for workers inside the firm tend to be those that diversify more subsequently.

ESSAYS ON EMPLOYER-EMPLOYEE RELATIONSHIPS AND FIRM PERFORMANCE

by

Hyowook Chiang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2005

Advisory Committee:

Professor John Haltiwanger, Chair/Advisor
Dr. Julia Lane
Professor Gordon Phillips
Professor Michael Pries
Professor John Shea

PREFACE

This document reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It 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, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author 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.

This dissertation is dedicated to my parents.

ACKNOWLEDGMENTS

I owe my gratitude to all the people who have made this dissertation possible and because of whom my graduate study has been one that I will cherish forever.

First of all, I would like to thank my advisor, Professor John Haltiwanger for his continuous support and help over the past years. He also gave me a chance to work with beautiful micro data at Longitudinal Employer Household Dynamics Program of the U.S. Census Bureau. He always encouraged me from my early stage when I struggled to get my ideas until now when I am getting close to finishing my dissertation. It has been my pleasure to work with and learn from such a wonderful person.

I would also thank Professor John Shea. Without his valuable comments and suggestions, this dissertation would not have been finished. It was he who gave me my first graduate lecture. Since then I have got tremendous insights and energy from him. Professor Michael Pries also gave me a lot of valuable comments and guidance. Thanks are due to Dr. Julia Lane and Professor Gordon Phillips for agreeing to serve on my dissertation committee and for sparing their invaluable time reviewing the manuscript. I also benefited from helpful comments and advice from Fredrik Andersson, Matt Freedman, Namsuk Kim, Nicole Nestoriak, Kristin Sandusky, and members of the LEHD program. Without their help and LEHD's collegeal working environment, I could not be at this point.

I owe my deepest thanks to my family. My parents have always stood by me and guided me throughout my life. They have been always the source of my energy. I also had very influential discussions with my borther who gave me very insightful ideas from his business experience. My sister always encouraged and supported me during my graduate study. I thank all the members of my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Recently, the heterogeneity of workers has been documented and emphasized as a very important feature of labor in various economics fields. Labor is one of the key inputs in the production process, and it is quite different from other inputs such as capital and materials in that no single worker can be treated the same as any other worker given their unique characteristics including skill levels, age, labor market experience, gender, and so on. Labor economists have been trying to understand how these individual characteristics affect each worker's performance, especially with respect to their wages. Research on business performance, such as firm/establishment productivity studies, have only recently begun to pay attention to worker heterogeneity. The most important reason for this is a lack of data that incorporate both business level information about production processes and rich information about the individuals that work in each business. Many studies of productivity have treated all workers as homogeneous and use the total number of workers or total hours worked as a single variable representing labor input in the production function. Studies using micro level databases could only differentiate labor input into production workers and non production workers or skilled labor and unskilled labor. Past work that has tried to exploit labor heterogeneity has been mostly case studies.[1] This dissertation exploits the heterogeneity of labor and variation in human resource management systems by different businesses, and tries to understand their impact on firm performance and various outcomes.

I use the unique employer-employee matched data recently developed by the Longitudinal Employer Employee Dynamics (LEHD) program of the U.S. Census Bureau to measure human resource management systems in businesses. This new dataset allows one to derive information about each business' workforce. We can measure not only net worker flows but also gross worker

[1] There are some exceptional studies that use more representative samples, notably Ichniowski, Shaw, and Prennushi (1997), Black and Lynch (2001), and Bresnahan, Brynjolfsson, and Hitt (2002).

flows. Workforce quality can be measured in more rigorous ways, and we can exploit not only the first moment of firm wages but also the entire distribution of wages. These measures of within-firm human resource practices can be linked to different business level information such as the inputs/outputs of businesses' production processes, the entry/exit of establishments/firms, and the scope of businesses' activities. I take full advantage of the data, especially its longitudinality and representativeness, to derive more general results on the impact of human resource policies on business performance.

Chapter 2 investigates the effects of "learning by doing" at the establishment level in U.S. manufacturing. The existing literature on learning suggests that firms learn from their past production activity. I argue that firm learning is not only affected by past output, but also by worker turnover, as learning is embodied in workers. High worker turnover will make learning by doing more difficult because firms lose workers who have "learned" about the production process from their past activity. I integrate the LEHD data with detailed business information contained in the Longitudinal Research Database (LRD). This enables us to look deeply inside businesses and to characterize and measure worker flows, workforce composition, and thus the nature of learning, as well as to explore the relationship between establishment-level productivity and learning by doing. The empirical analysis reveals that firms with lower rates of worker turnover have higher productivity and "learn" faster than those with higher worker turnover given the same amount of past production experience. Moreover, I develop new instruments, based on local downstream demand, firm wage policy, regional hiring conditions of related industries, firm vintage, and firm location, that significantly affect past production experience and worker turnover but are not necessarily affected by firm productivity shocks. With these instruments I show that learning by doing and turnover have causal effects on productivity.

Chapter 3[2] investigates the relationship between firm performance with workforce quality and worker turnover in five selected industries: semiconductors, software, retail food, trucking, and financial services. Our sample integrates LEHD data with the Economic Censuses, which permits

[2]This chapter draws heavily upon joint work with Ben Campbell, John Haltiwanger, Larry W. Hunter, Ron Jarmin, Nicole Nestoriak, Tim Park, and Kristin Sandusky (2004).

us to examine both the product market side and the labor market side of business operations. We find that measures of productivity, workforce quality, and worker turnover are highly correlated across businesses. High productivity businesses have a large share of high human capital workers and also have low worker churn. Survival is a function of all of these factors - businesses with high productivity, low churn, and high human capital are more likely to survive. Surprisingly, workforce quality and worker turnover are found to have independent effects on firm survival even after controlling for productivity. While there are some common patterns, we found it helpful to take into account the idiosyncratic factors present in the industries under investigation. All of these industry-specific idiosyncratic factors make one cautious about drawing general inferences from the analysis. While this caution is warranted in some respects, the common patterns are more striking than the idiosyncratic factors. Workforce quality, worker churn, and firm performance are related across all the industries studied and virtually all classifications of businesses that we have considered.

Chapter 4[3] assesses the fit between firm-level internal labor markets (ILMs) and firm diversification in the U.S. financial services sector. The sector comprises a number of related sub-industries and recent deregulation has allowed firms to construct increasingly diversified portfolios of activities across these sub-industries. Banking deregulation in particular has loosened geographic restrictions on firm activities. Drawing on the "resource-based view" of firm strategy, I hypothesize that firms with stronger ILMs are more likely to diversify. I find support for this view in analysis of data from the LEHD program matched to the Longitudinal Business Database (LBD). Firms with lower net turnover, lower wage dispersion, and greater opportunities for workers inside the firm tend to be those that diversify more subsequently.

Finally, Chapter 5 provides the conclusion to the thesis.

[3]This chapter draws heavily upon joint work with Cheryl Grim, John Haltiwanger, Larry W. Hunter, Ron Jarmin, Nicole Nestoriak, Kristin Sandusky, and Jeongil Seo (2005).

Chapter 2

Learning by Doing, Worker Turnover, and Productivity Dynamics

How do firms[1] improve performance? There are a number of possible explanations for firm-level productivity growth. "Learning by doing," the notion that unit costs are a decreasing function of cumulative production, is a leading candidate. Traditional models that incorporate learning by doing assume that the learning process is a function of cumulative gross activities such as output, investment, or employment. However, if there is no difference in two firms' initial productivity levels and each faces the same demand conditions, then this assumption will generate the same learning processes for the two firms since their output (or employment) decisions will be identical. In this case, productivity levels will differ only because firms are different in their ages or their inherent abilities. It is not possible for less productive firms in earlier periods to catch up with more productive firms in later periods. However, this is not realistic, since we often see firms' relative performances within a cohort change dramatically.[2]

One possible source of change in idiosyncratic productivity over time may be differential rates of accumulation in the learning process. Some firms may be slower to learn than others. This could occur if there is variation in managerial skill. Alternatively, when the frequency of production interruptions varies across firms, we may observe differential rates of speed in firm learning. These factors might cause firms to differ in their ability to convert "doing" into "learning". As Lucas (1993) points out, a firm's learning can be done by the management, the workforce, or the organization as a whole. If it is the workforce that is doing the learning, then learning is embodied in workers. High worker turnover will make learning by doing more difficult, since firms lose workers who have "learned" about the production process from their past activity. In

[1] In this chapter, the unit of observation is the establishment (or plant). However, I use "firm" and "establishment/plant" interchangeably, although, in principle, firms can have more than one establishment (multi-unit firms).

[2] See Baily, Hulten, and Campbell (1992).

other words, these firms fail to convert "doing" into "learning." Given that there is tremendous heterogeneity in firms' worker flows and turnover patterns, assuming that learning is simply a function of cumulative activities can miss very important heterogeneity that can explain variation of productivity among firms within an industry.

The above arguments are closely related to the concepts of "general" and "firm-specific" investment in human capital[3]. Purely general training received by a worker within a given firm is defined as investment that raises the potential productivity of the worker in other firms by as much as it is raised within the firm providing the training. Purely specific training raises the worker's productivity within the firm providing the training, but leaves his productivity unaffected in other firms. General capital is completely embodied in the worker, but the productivity of specific capital is jointly dependent on the productive characteristics embodied in the worker and the characteristics of other firm-specific inputs[4]. Becker (1993) provides the following discussion on the relationship between specific human capital and worker turnover:

> Turnover becomes important when costs are imposed on workers or firms, which are precisely the effects of specific training. Suppose a firm paid all the specific training costs of a worker who quit after completing the training. ... he would have been receiving the market wage and a new employee could be hired at the same wage. If the new employee were not given training, his marginal product would be less than that of the one who quit since presumably training raised the latter's productivity. Training could raise the new employee's productivity but would require additional expenditures by the firm. In other words, a firm is hurt by the departure of a trained employee because an equally profitable new employee could not be obtained.

I develop new ways of measuring the effects of "learning" that incorporate both cumulative "doing" and worker turnover. Turnover functions as if it were a depreciation factor of "doing"

[3]Losing experienced workers means losing both general and specific human capital of those workers. However, I am more interested in losing "firm-specific" human capital because the same level of "firm-specific" human capital is not easy to replenish just by paying the same amount of compensation to newly hired workers.

[4]See Willis (1986).

and essentially decreases the magnitude or lowers the speed of "learning". These measures can be constructed only if one has longitudinal and universal data on employment history and firm activity. In contrast to the existing literature on learning, this paper shows that learning is not only affected by past output, but also by worker turnover within firms. The basic idea is that high worker turnover will make learning by doing more difficult. Using this approach, I estimate that firms with historically lower rates of turnover "learn" faster than those with higher turnover given the same amount of past output.

However, both cumulative activity and worker turnover are likely to be correlated with unobserved firm productivity shocks. "Learning," as measured by cumulative past output, can be correlated with current productivity shocks when there is serial correlation in productivity shocks. The literature on learning by doing has mostly ignored this problem.[5] Also, workers may leave their jobs when they expect a negative productivity shock at the firm where they currently work. This line of causality runs counter to my previous argument that worker turnover will affect firms' productivity negatively. I develop new instruments, based on local downstream demand, firm wage policy, and regional hiring conditions of related industries, that significantly affect past production experience and worker turnover but are not necessarily affected by firm productivity shocks. With these instruments I show that learning by doing has causal effects on productivity, which is the most important contribution of this paper.

I integrate a new employer-employee matched data developed by the Longitudinal Employer Household Dynamics (LEHD) Program at the U.S. Census Bureau with detailed business information from the Longitudinal Research Database (LRD) to create a data set that enables us to look more deeply inside businesses and to characterize and measure worker flows, workforce composition, and thus the nature of learning, as well as to explore the relationship between establishment-level productivity and learning by doing.

In this study, I focus on the U.S. manufacturing industry, since we see substantial establishment level productivity growth in this sector. Foster, Haltiwanger, and Krizan (2001) find that for

[5] One exception is Benkard (2000). He uses world GDP, the price of oil, and a time trend as demand shift instruments.

the manufacturing industry, within-plant productivity growth explains about 50% of industry-wide productivity growth.

The paper is organized as follows. In Section 2.1, I review related literature on learning by doing. In Section 2.2, I discuss the estimation methods. Section 2.3 describes the data sets used in my analysis and explains how I construct my sample. In Section 2.4, I present initial results, and in Section 2.5 I try to identify whether there is a causal relationship between key variables. Section 2.6 concludes.

2.1 Learning By Doing

"Learning by doing," which occurs when unit costs decrease with cumulative production, was first observed in the aircraft industry by Wright (1936), who found that unit labor inputs in airframe production declined with the total number of airframes of the same type previously produced. Arrow (1962) argues that learning is the product of experience and that it can only take place through the attempt to solve a problem during activity. He uses cumulative gross investment as an index of experience instead of cumulated gross output. Given that new machines produced and put into use are capable of changing the environment in which production takes place, he argues that learning happens with continually new stimuli.

Rapping (1965) uses Liberty shipbuilding data during World War II and takes a production function approach to show that cumulated output has a significant effect on productivity advances during wartime. Sheshinski (1967), working under the assumption of disembodied technical progress and using a constant elasticity of substitution (CES) production function, shows with cross-sectional U.S. and international manufacturing data that efficiency growth is correlated with the level of cumulated investment (and output).

Early research on the topic of "learning by doing" is in many aspects limited. One obvious problem lies in the data. Studies on the aircraft industry use military production data. Sheshinski (1967) uses aggregate (two digit) and state level U.S. manufacturing data. Therefore, sample sizes are quite small. Since data are not longitudinal, it is not possible to identify firm births or to calculate cumulative gross investment (output). Sheshinski uses the gross book value of capital as an index for learning. The cross-country data are two digit aggregate manufacturing data over ten years. Imposing homogeneous production technology across countries is also very restrictive.

Bahk and Gort (1993) use the U.S. Census Bureau's LRD, which is plant level, longitudinal data for U.S. manufacturing. One can identify a plant's birth, so the calculation of cumulated gross activity is straightforward. Furthermore, the sample size using the LRD increases significantly to 2,150 plants over a 14-year period, and production functions can be estimated by four-digit SIC. Bahk and Gort decompose learning by doing into organizational learning, capital learning, and manual task learning. They estimate effects of firm-specific learning-by-doing while controlling for variation of general human capital with the average wage rate. Since identifying birth is possible with the LRD, they focus only on new plants and their histories following birth. However, a sample of only new plants may not be representative, and estimating production functions with this sample may result in sample selection bias. Below I explain an alternative way to utilize the entire LRD, including not only new plants but also continuing plants. Bahk and Gort find that plant-specific learning effects are important, but they do not take into account potential endogeneity of learning with respect to productivity shocks.[6]

As Argote *et al.* (1990) and Benkard (2000) point out, traditional learning models define experience simply as cumulative past output (or investment):

$$E_t = E_{t-1} + Q_{t-1}, \tag{2.1}$$

This assumes that recent production and more-distant past production are equally important in determining a firm's current efficiency. For example, the conventional literature assumes that production during the Henry Ford era in the early 20th century is as important as production last year for current productivity. Argote *et al.* allow for the possibility of depreciation of knowledge

[6] It is well known that, in typical production function estimation, input variables are likely correlated with unobserved productivity shocks, resulting in inconsistent estimates. Moreover, if there is a serial correlation in productivity shocks, then cumulative output is also likely correlated with current productivity. This implies that the OLS estimate of the impact of learning will be inconsistent.

and define experience as follows:

$$E_t = \delta E_{t-1} + Q_{t-1} \tag{2.2}$$

This specification allows for the possibility that learning does not persist. Argote *et al.* (1990) find with wartime Liberty ship production data that learning depreciates quickly. Benkard (2000), using production data for the Lockheed L-1011 TriStar, finds evidence supporting organizational forgetting, which occurs when a firm's stock of production experience depreciates over time. However, "depreciation" or "forgetting" are very abstract concepts.[7] Moreover, past authors' implicit assumption that there is a constant rate of depreciation or organizational forgetting over time is quite strong.

In this paper, I adopt the idea of depreciating learning, but instead of estimating a depreciation rate or forgetting rate, I explicitly measure variables that I believe are a main source of depreciation of learning. I test whether "learning" defined in this fashion can explain a firm's productivity variation better than the traditional measure. The depreciating factor I examine is the worker turnover rate. One index for "learning" I use is defined as follows:

$$E_t = (1 - tr_{t-1})(E_{t-1} + Q_{t-1}) \tag{2.3}$$

where tr_{t-1} is the worker turnover rate.[8] The reason why I use turnover as a source of depreciation of learning is explained below.

The precise meaning of "learning" has not been clear in the previous literature on learning by doing. The literature implicitly assumes that the firm itself is doing the "learning." Hence, if two firms are identical in their past gross activities, then they should be identical in their levels of "learning." As mentioned in the previous section, a firm's learning can be undertaken by the management, the workforce, or the organization as a whole. Which part dominates a firm's learning

[7]Although Benkard (2000) does not explicitly explore the sources of forgetting, he suggests that turnover and layoffs may lead to losses of experience. Argote *et al.* (1990) investigate turnover effects on productivity in their study and find that the turnover effect is not significant. However, because there is tremendous heterogeneity among various industries, and because they exclusively focus on war time ship production while I study more broad manufacturing industry, I do not try to draw a general conclusion from their study.

[8]A formal definition of worker turnover will be provided in Section 2.3.

is an empirical question and may depend on industry characteristics, etc. If it is the management or the organization that learns, then the usual focus on gross activity may not be misleading. However, if it is the workforce that does the learning, then heterogeneity in worker flows among firms is a potentially important determinant of learning. "Learning" from the firm's perspective is just the sum of each individual worker's "learning." High turnover (i.e. the loss of "learned" employees and/or accession of new and inexperienced workers) will make it harder for a firm to convert its "doing" into "learning." Only firms that retain their "learned" workers - those who have accumulated important knowledge from their past production activities - can fully convert their "doing" into a stock of "learning." In any case, as emphasized in Reichheld (1996), it is not companies but individuals that actually learn, and their learning takes time.

A firm's "learning", in this sense, might be thought of as the sum of each worker's "human capital" within that firm. As is well known, human capital can be decomposed into "general" and "firm-specific" human capital. What is lost from failing to retain experienced workers should be "firm-specific" human capital, since the same level of "general" human capital is easy to replenish by paying the same amount of compensation to new workers. New workers accumulate "firm-specific" human capital during the production process. This point is well understood by Bahk and Gort (1993), who use the average wage rate to control for general human capital when they try to identify firm-specific learning by doing. I also try to separate out effects of "general" and "specific" learning. However, instead of using the wage rate to proxy for general human capital, I use new estimates of human capital developed by Abowd, Lengermann, and McKinney (2003)[9].

However, retaining all workers may not be an optimal strategy for a firm. Separating with workers who are not well matched to the firm could indeed enhance productivity. Also, hiring workers with different working experience or new knowledge is another way of accessing new technology, just as investing in a new capital good is a way to adopt new technology. So, if this matching effect dominates the learning effect, then it is not clear whether worker turnover is necessarily bad for firms' productivity.

[9]The estimation method of Abowd, Lengermann, and McKinney is discussed in Section 2.3.

2.2 Empirical Method

There are at least two ways to estimate learning by doing using the production function approach. To illustrate the differences between these approaches, consider the following simple Cobb-Douglas production function with an index of "experience" or "learning" incorporated as in equation (2.4):

$$Q_{j,t} = AK_{j,t}^{\alpha_k} L_{j,t}^{\alpha_l} M_{j,t}^{\alpha_m} E_{j,t}^{\alpha_e} \tag{2.4}$$

where $Q_{j,t}$ is the gross output of establishment j in year t, $K_{j,t}$ its real capital stock at the end of year t, $L_{j,t}$ its total hours worked, $M_{j,t}$ its real materials input, and $E_{j,t}$ the index of "learning".

The first approach adopted in previous literature, such as Argote, *et al.* (1990), Bahk and Gort (1993), and Benkard (2000), is to estimate equation (2.4) directly to determine the coefficient of "learning" together with the elasticities of the usual input variables. The second approach, which is the one I take here, involves a two-step procedure.[10] In the first step, a productivity measure is generated ignoring factors other than standard input variables. In the second step, the effects of learning on firm productivity are estimated.

As will be emphasized in the Section 2.3, our final sample for estimating the impact of learning by doing is quite small.[11] Estimation of the production function using this sample raises several potential problems. First, estimates in a small sample are likely to be imprecise. This is more serious when we allow elasticities to vary by detailed industry within the entire manufacturing sector. Second, this sample may not be representative of all manufacturing firms or of all young firms.[12] In this case, if all firms (within the same detailed industry) have access to the same production technology, then estimates of elasticities may be inconsistent.[13] Moreover, due to endogeneity of inputs and learning with respect to productivity, OLS estimation of the production function will generate spurious estimates of elasticities and an inconsistent estimate of the learning

[10] For example, see Levinsohn and Petropoulos (2001) and Pavcnik (2002).

[11] In most of my specifications below, the sample size is less than 10,000. This is a quite small number compared to size of the LRD.

[12] See Figure 2.2 in the next Section.

[13] If the production technology of young firms are different from that of old firms, then one need not worry about this problem. However, the first problem still exists. Ideally, if we had complete information on the production history of all existing firms, then it would be best to estimate equation (2.4) directly.

effect.

On the other hand, the two-step approach uses information on all firms in the manufacturing sector in the first step. Hence, we can get reasonable productivity estimates from the first step. Since the final sample is representative of all young establishments, as is shown in Section 2.3, the second step will generate consistent estimates of the impact of learning among young businesses. Given the focus of our sample on young businesses, I follow the second approach.

Existing studies use two different methods to measure multifactor productivity. The index number approach involves no estimation; input elasticities are calculated from cost shares, invoking Shephard's Lemma. The other approach is econometric and is based on production function estimation. In this study, I will use the index number method.[14]

Armed with productivity measures from the first step, we are ready to estimate learning by doing effects in the second step. The basic equation estimated to determine the effects of learning on productivity is

$$tfp_{j,t}^{v,s} = \gamma_0 + \gamma_s + \gamma_v + \gamma_t + \gamma_h h_{j,t} + \gamma_e e_{j,t} + u_{j,t} \tag{2.5}$$

where $tfp_{j,t}^{v,s}$ is the log of TFP in year t of firm j born in year v and belonging to industry s; γ_s is a four digit industry effect; γ_v is the vintage effect; γ_t is the year effect; $h_{j,t}$ is a human capital measure; and $e_{j,t}$ is the log of the stock of learning measured either by equation (2.1), equation (2.2), or a simple turnover rate. I include γ_s to control for between-industry heterogeneity, given that our sample includes all manufacturing industries. The vintage effect will capture any effect embodied in the capital of each vintage and the time effect will capture economy-wide technological progress. With both vintage and time effects in the model, I implicitly control for firm age. If firms learn their production process merely by aging, then the "learning" variable will not have a

[14]The econometric approach implicitly assumes that standard input variables and those variables not included in the first step but used as regressors in the second step are orthogonal to each other. Therefore, there may be an "omitted" variable bias if this assumption is not valid. The index number approach relies on a constant returns to scale technology assumption. Hence, this method will over- or under-estimate productivity when there is increasing or decreasing returns to scale. However, under constant returns to scale and profit maximization (either static or dynamic), the estimates of learning effects will be consistent if we use the productivity measure based on the index number approach. For detailed methodology, see Appendix

significant effect on productivity since firm age will capture most of the learning effect. However, if learning is done through the problem solving process during production, and not by just aging, then the "learning" variable will have a significant impact on productivity.[15]

To see whether firms with lower worker turnover have higher productivity levels than those with higher turnover (controlling for past output), I estimate

$$tfp_{j,t}^{v,s} = \gamma_0 + \gamma_s + \gamma_v + \gamma_t + \gamma_h h_{j,t} + \gamma_e e_{j,t} + \gamma_s SR_{j,t} + u_{j,t} \tag{2.6}$$

where $e_{j,t}$ is the log of past output measured by equation (2.1) and $SR_{j,t}$ is the separation rate. The latter will be defined in Section 2.3. Given that worker turnover may affect not only the level of productivity but also the "speed" of learning, I estimate

$$tfp_{j,t}^{v,s} = \gamma_0 + \gamma_s + \gamma_v + \gamma_t + \gamma_h h_{j,t} + (\gamma_{eh} dh_{j,t} + \gamma_{el} dl_{j,t})\, e_{j,t} + u_{j,t} \tag{2.7}$$

where $dh_{j,t}$ ($dl_{j,t}$) is a dummy variable that takes a value of one when the separation rate is high (low) and $e_{j,t}$ is again the log of past output. I interpret γ_{eh} and γ_{el} as measuring the speed of learning.

2.3 Data

2.3.1 Sources

In this study, I use two datasets constructed by the U.S. Census Bureau. One is the Longitudinal Research Database (LRD), which contains annual data on U.S. manufacturing establishments collected in the Census of Manufactures (CM) and the Annual Survey of Manufactures (ASM). The U.S. Census Bureau conducts the CM in years ending in "2" and "7" and the ASM in each of the 4 years between the CMs. The ASM is based on a sample drawn from the census universe of approximately 300,000-400,000 establishments. The ASM sample is updated in years ending in "4"

[15] Bahk and Gort (1993) did not simultaneously estimate the impact of their traditional learning measure and firm age (which is just the difference between the calendar year and the firm's vintage), since those variables were highly correlated with each other. In this sense, the evidence of learning I present below is stronger than that presented by Bahk and Gort (1993).

and "9." The design of the ASM assures that large establishments are included with certainty and that small establishments are rotated out of the panel at the end of the five-year period. Both the CM and ASM contain data on output, inputs, establishment ID, firm ID, industry classification, and geographic codes. Information is available on factors of production such as employment, payroll, supplementary labor costs, worker hours, cost of fuels and electricity, cost of materials, capital expenditures, and inventories; and on measures of output, such as the value of shipments and value added. The basic unit of observation in the LRD is the "establishment", which is defined as a "single physical location" engaged in one of the categories of industrial activity in the Standard Industrial Classification (SIC) system. Information from the LRD is rich enough to estimate production functions. However, while the LRD can generate gross job flows data by establishment, one cannot generate the worker flows data that is crucial to constructing my preferred measure of "learning." For this purpose, we need a dataset that stores a complete work history for each individual worker.

			Earnings			
Person	Plant	year	Q1	Q2	Q3	Q4
1	X	1994	4500	4500	4800	4800
1	X	1995	5000	3500		
1	Y	1995		2000	6500	6900
2	Y	1994	1800	1800	1800	1800
2	Y	1995	2000			
2	Z	1995			2500	3000
3	Z	1994	5500	5500	5500	5500
3	Z	1995	6000	6000	6000	6000
4	X	1994	3700	3700	3800	3800
4	X	1995	4000		4200	4300

Table 2.1: Structure of the LEHD Dataset

The other main dataset used here was developed by the Longitudinal Employer-Household Dynamics (LEHD) Program at the Census Bureau. This data set integrates information from state unemployment insurance (UI) data, ES202 data, and Census Bureau economic and demographic data in a manner that permits the construction of longitudinal information on workforce composition at the firm level.[16] Every state in the U.S., through its Employment Security Agency, collects quarterly employment and earnings information to manage its unemployment compensation program. This database enables LEHD to construct quarterly longitudinal data on employees as shown in Table 2.1. The sample size is large and information is more accurate than survey based data. Since the data are universal, movements of individuals to different employers and their consequences for earnings can be tracked. It is also possible to construct longitudinal data using the employer as the unit of analysis. Furthermore, individual UI wage records can be linked to Census data to obtain information such as date of birth, place of birth, and gender for almost all workers in the sample. LEHD staff have exploited the longitudinal and universal nature of the dataset to estimate fixed worker and firm effects jointly using the "human capital" model discussed below.

It is important to note at this point that the units of observation in the two data sets are not identical. While the LRD is at the establishment level (Permanent Plant Number, or PPN), the business level identifiers on UI files are State Employer Identification Numbers (SEINs), which do not necessarily match the establishment level identifiers. I explain how I matched the two data sources in Appendix C.

As mentioned in the previous section, I want to control for the effect of general human capital on a firm's productivity. I use the measures defined in Abowd, Kramarz and Margolis (1999) and in Abowd, Lengermann, and McKinney (2003). These measures are based upon a statistical decomposition of the wage for a worker into a person effect, a firm effect, and an effect due to time varying person characteristics including general labor market experience. The person effect is the portable component of a worker's wage and, as such, is a good summary measure of the general skills of a worker (indeed, studies have shown that it is highly correlated with direct measures of skills such as education).[17] Specifically, the core estimation model used is

$$w_{ijt} = \theta_i + x_{it}\beta + \psi_j + \varepsilon_{ijt}. \tag{2.8}$$

[16]For the desciption of LEHD data, see Abowd, Lengermann, and McKinney (2003).

[17]See Abowd, Lengermann, and McKinney (2003).

The dependent variable is the log real wage rate of individual i working for employer j at time t. The first component is a time invariant person effect, the second is the contribution of time varying observable individual characteristics, the third is the firm effect, and the fourth component is the statistical residual, orthogonal to all other effects in the model. In the empirical analyses, I control for workforce quality using a "human capital" estimate, denoted by "h", that is defined as the sum of the fixed worker effect and the experience component (i.e. $h_{it} = \theta_i + x_{it}\beta$). I use the mean of "$h$" across all workers at a firm in the estimation.

Variables from the LRD are constructed as follows. Real output is the total value of shipments less inventory investment (finished goods and work-in-progress), deflated by the four-digit SIC output deflator. Real material costs are the sum of the costs of materials, parts, resales, fuels, electricity, and contract work, deflated by the four-digit materials deflator. Total (quality adjusted) hours worked is measured as the product of total production hours and the ratio of total wages to wages paid to production workers. Industry level deflators are available from the NBER/CES Productivity Database constructed by Bartelsman, Becker, and Gray[18].

Initial establishment equipment and structure capital stocks are the reported book values of machinery and building assets, deflated by the ratio of book to real values for the corresponding two digit industry published by the BEA. Given the initial capital stock values, I use the perpetual inventory method forward and backward to construct capital stock series.[19] There were several changes in capital stock information available in the CM and the ASM during 1980s and 1990s. Until 1987, book values of equipment and structures are reported separately in all CMs and ASMs. However, after 1987 capital is not reported at all in the ASM. 1992 is the last CM year when book values of both equipment and structures are reported separately. In the 1997 CM, only the sum of equipment and structures is reported. These changes restrict my ability to build capital stock and

[18]See Bartelsman and Gray (1996) for detailed description.

[19]The perpetual inventory method is as follows:

$$K_t = (1 - \delta_t)\, K_{t-1} + I_t$$

where δ_t is the three-digit industry depreciation rate, K_{t-1} is the real capital stock (equipment or structures) of the previous year, and I_t is real investment.

thus productivity (more specifically, total factor productivity (TFP)) series. For establishments that are observed in 1992, I use the 1992 real capital stock (equipment and structures) as initial values. For those observed in 1997 but not in 1992, I use the real capital stock in 1997, broken into equipment and structures based on the relative amount of cumulative investment in equipment and structures. For these two types of plants, I can use the perpetual inventory method forward and backward.[20] For establishments observed only in the ASM years, I am not able to build a capital stock series at all.

The turnover measure I use in this paper is the separation rate. The separation rate ($sr_{j,t}$) of firm j at year t is defined as

$$sr_{j,t} = \frac{separation_{j,t}}{employment_{j,t-1} + employment_{j,t}} \tag{2.9}$$

For formal definitions, see the Appendix. The separation rate measures the fraction of firm's workers who had previous production experience who left their jobs. The conventional way of defining worker and job flows uses average employment as a denominator (see Davis, Haltiwanger, and Schuh (1996)). I use the sum of employment instead of average of employment over two years in order to bound flow rates between zero and one. Below, I use the worker turnover measure as a depreciation factor, and this factor needs to be bounded by one from above. Using past output and turnover rates, I define the following measures that capture "learning" or "experience" at the firm level:

$$E_{1,j,t} = E_{1,j,t-1} + Q_{j,t-1} \tag{2.10}$$

$$E_{2,j,t} = \log\left(E_{1,j,t}\right) \tag{2.11}$$

$$E_{3,j,t} = \left(1 - sr_{j,t}\right)\left(E_{3,j,t-1} + Q_{j,t-1}\right) \tag{2.12}$$

$$E_{4,j,t} = \log\left(E_{3,j,t}\right) \tag{2.13}$$

Equation (2.10) is a traditional measure of learning, which is a simple sum of all past outputs (as

[20] Suppose establishments are observed in 1991, 1992, 1993, 1996, 1997, and 1998. For these establishments, I use 1992 capital stocks as initial values to calculate the 1991-1993 capital stock series, and use the 1997 capital stocks to calculate the 1996-1998 capital stock series.

in equation (2.1)). A more complicated and structured measure is equation (2.12). It incorporates both cumulative output and the history of worker turnover.

2.3.2 Sample Selection

The datasets I use in the estimation are much smaller than either the LRD or the UI samples. First, while the LRD covers all states, the LEHD has not yet incorporated the UI datasets of all states, and therefore estimates of human capital are only available for 7 states. Second, the UI data are available only for the 1990s, while the LRD covers the 1980s and 1970s as well. Third, the LRD only covers the manufacturing sector, while the UI data cover all sectors. Hence, the sample size of the matched dataset is smaller than either of the two main datasets. In addition, the learning-by-doing hypothesis requires that one calculate cumulated gross output (or investment) as well as the history of turnover. Given that the UI data are only available for the 1990s,[21] I can calculate cumulative activity only for businesses born in the 1990s (the "no-left-censoring" condition) with two or more years of reports (the "no-hole" condition).

Figure 2.1 shows the age distribution of the final matched sample. Observations are highly concentrated on very young businesses and in the Census year 1997. 45% of my sample businesses are one year old and 80% are three years old or younger. Given the short average age of businesses in the sample, the estimation that I conduct relies more heavily on cross-sectional variation than on time series variation.

Figure 2.2 shows three kernel density estimates of the size distribution, where size is measured by the log of real output in 1997. The first estimate ("All Plants") is based on all establishments in 1997 CM. The second ("All Young Plants") is based on all establishments born in 1990 or later from the 1997 CM. The last ("Matched Plants") is based on the final matched sample used in my empirical analysis.[22] The average size of young plants is smaller than that of all plants. However, the size distribution of the matched sample is very similar to that of all young plants.

[21] To be more specific, available years of UI data are different among the 7 states. Some states' data are available from 1990 or 1991 to the present, while others' data are available only from 1994 or 1995 to the present.

[22] All three samples are constructed from the seven states that are included in the human capital estimation.

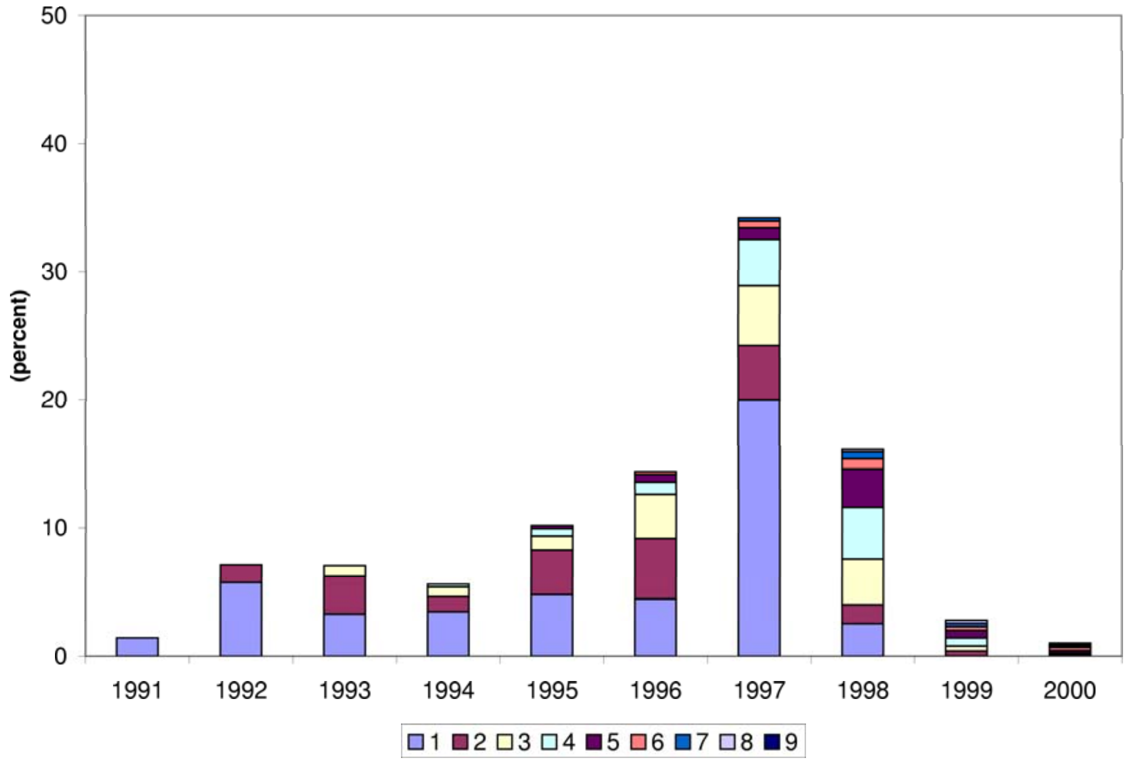


Figure 2.1: Age Distribution by Year

Therefore the matched sample seems to be quite representative of all young plants, although the matched sample (9,041 observations) is significantly smaller than the sample of all young establishments (81,909 observations).[23] Among 81,909 observations of all young plants, I can construct TFP measures for only 42,217 observations mainly due to the difficulty of constructing capital stock measures. Among these 42,217 observations, only 23,236 observations are matched to the LEHD data and have human capital estimates and turnover measures. Finally only 9,041 observations have complete histories since their birth (the "no-hole" condition).

Figure 2.3 shows how productivity evolves over time for young businesses in the matched sample. TFP is persistent over time, but not extremely persistent. This is consistent with the idea that there is a significant time-varying component as well as a permanent component in plant level productivity. Figure 2.4 shows the evolution of worker turnover at young establishments. Again,

[23] I also constructed a larger sample using imputed outputs to construct learning for firms with incomplete histories. This sample is similar to the matched sample in terms of the size distribution, but has more older firms. Empirical results using this sample are similar to those using the matched sample.

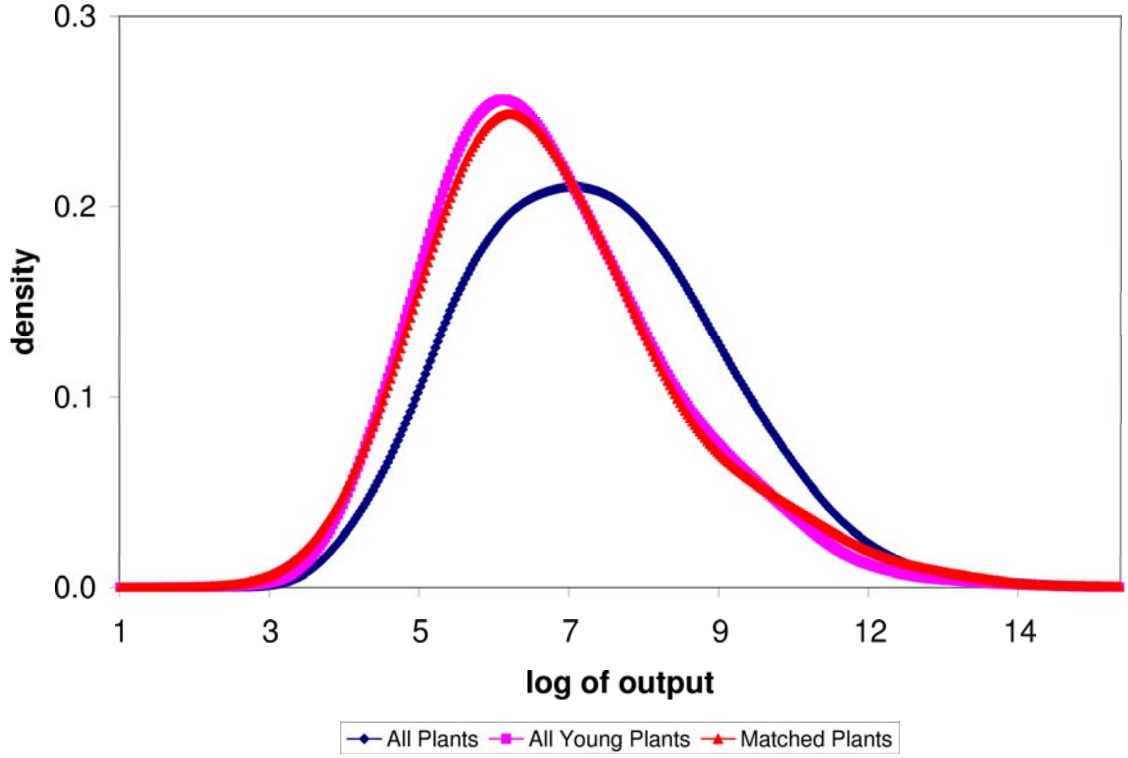


Figure 2.2: Kernel Density of Size in 1997

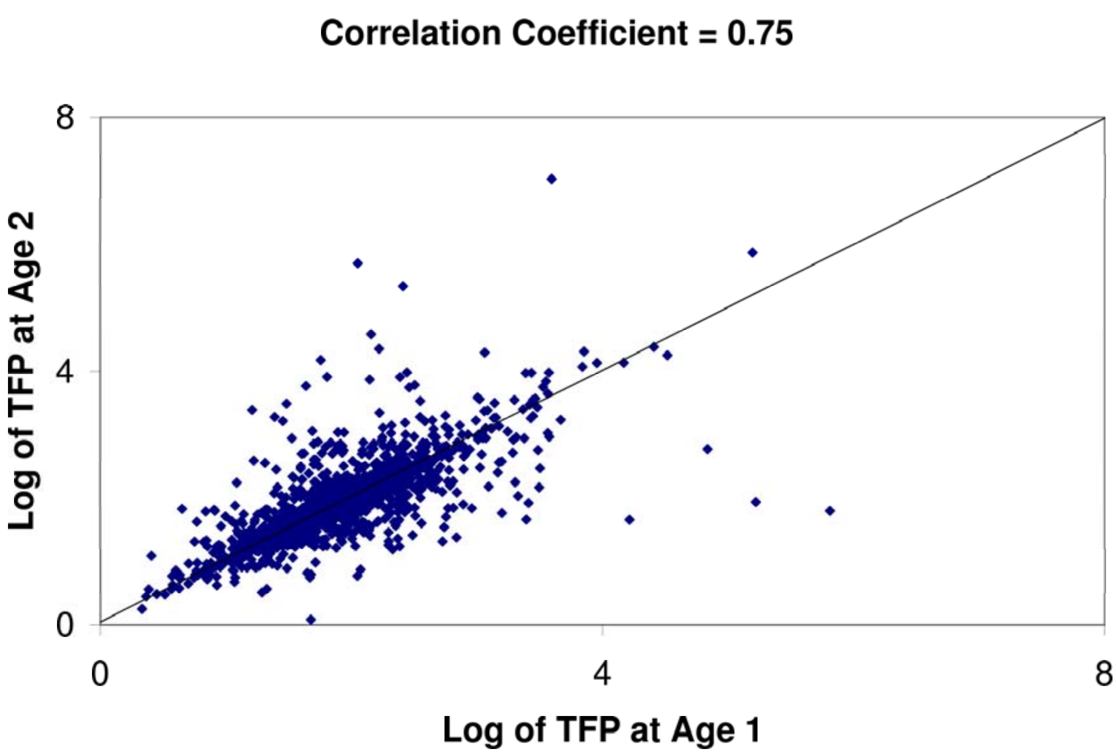


Figure 2.3: Persistence and Heterogeneity of Productivity

there is persistence in turnover rates and significant heterogeneity among businesses. One can also see many observations off the 45^0 line. Persistence is weaker in worker turnover rates than in plant level productivity.

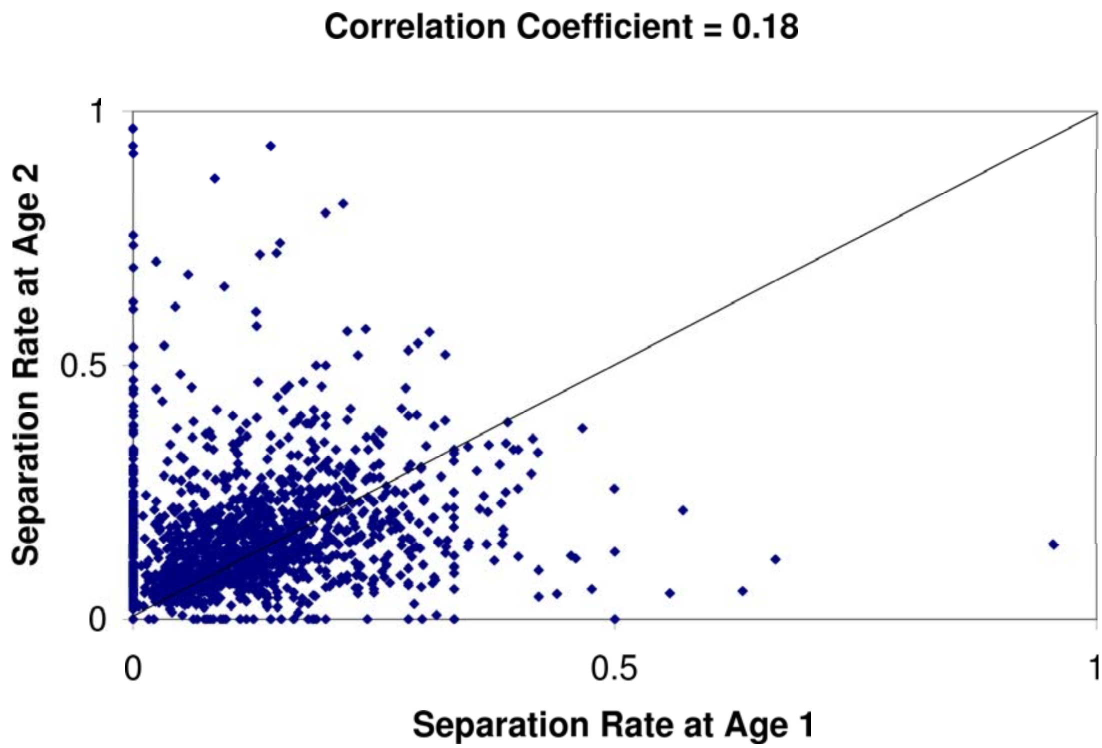


Figure 2.4: Persistence and Heterogeneity of Separation Rate

Figures 2.5 shows the evolution of a traditional measure of learning. This measure exhibits very high persistence. There are no observations below the 45^0 line in Figure 2.5 by construction, since traditionally measured learning cannot decline. Negative productivity growth at the firm level cannot be explained with this learning measure, since the stock of learning is always at least as high in the current period as it was in the previous period. However, there are some observations below the 45^0 line in Figure 2.6, which shows the evolution of my preferred measure of learning incorporating worker turnover. This measure allows for learning to depreciate as well as accumulate over time. If a firm has low production activity and high worker turnover in a given period, then it is possible that the stock of learning is lower in the current period than in the previous period. In principle, this measure of learning can explain negative productivity growth at the plant level.

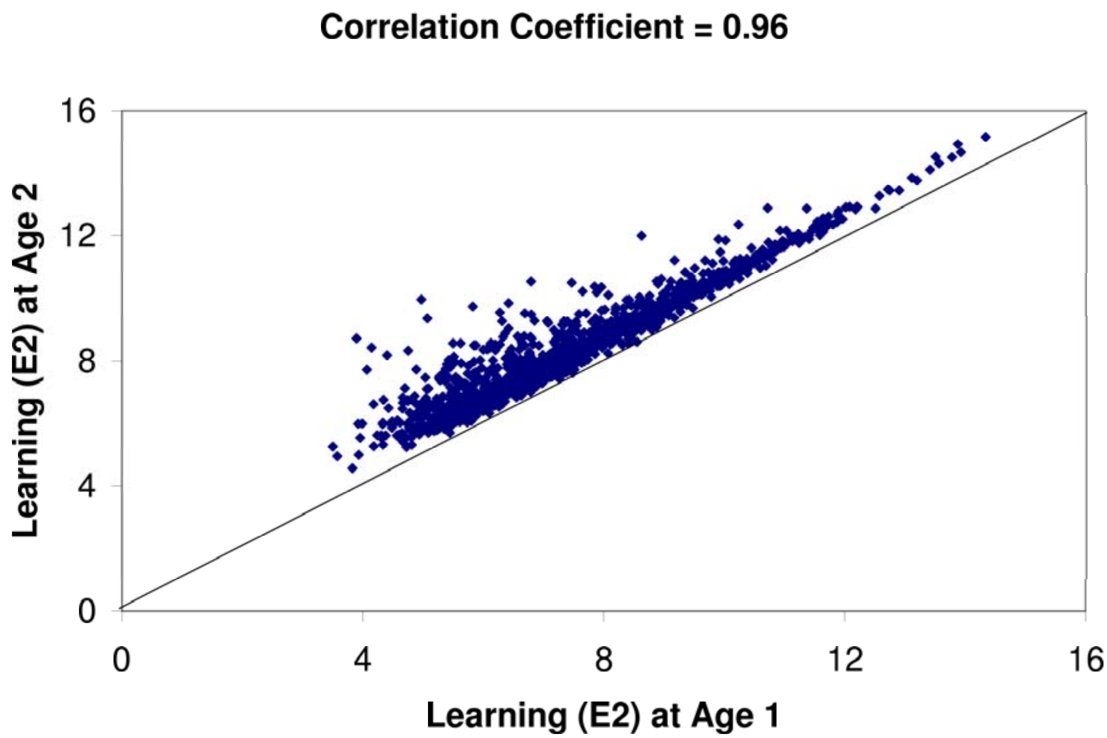


Figure 2.5: Persistence and Heterogeneity of Learning (E2)

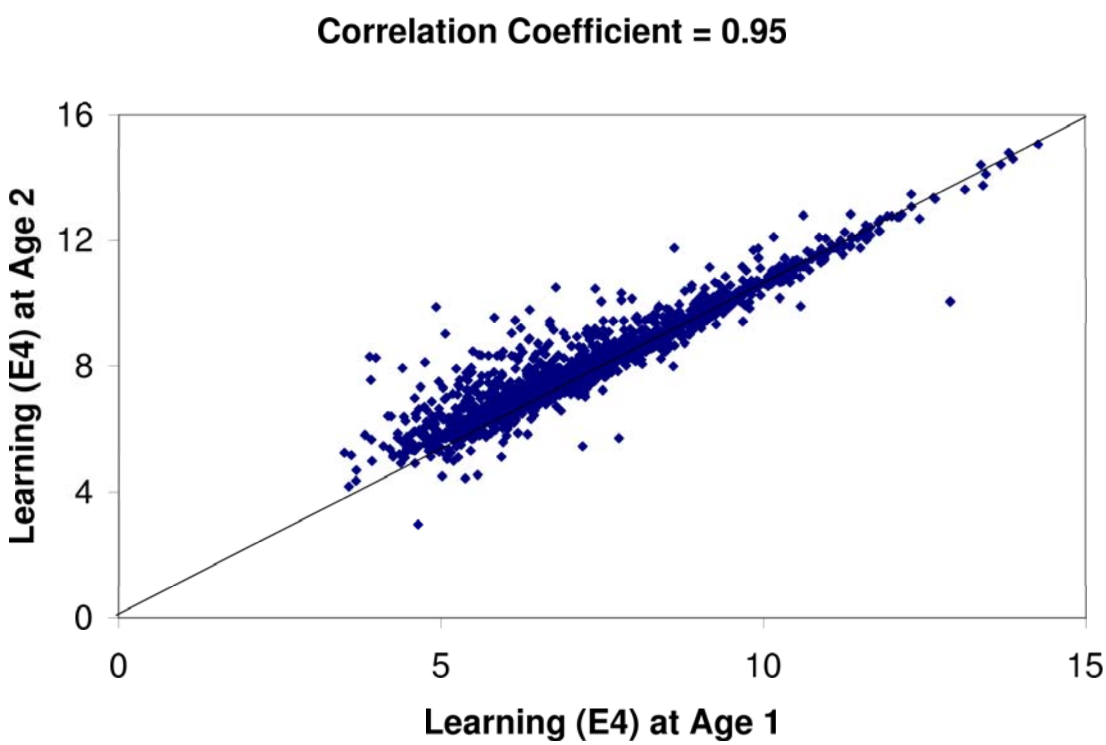


Figure 2.6: Persistence and Heterogeneity of Learning (E4)

Even though the learning measure in Figure 2.6 is very persistent, it is less persistent than the traditional measure, since it has additional variation resulting from turnover.

2.4 Empirical Findings

As explained in the previous section, I estimate the impact of learning on TFP, controlling for four digit SIC, calendar year, vintage, and human capital in all estimations[24]. Regressions (1)-(5) are based on estimation of equation (2.5),

$$tfp_{j,t}^{v,s} = \gamma_0 + \gamma_s + \gamma_v + \gamma_t + \gamma_h h_{j,t} + \gamma_e E_{j,t} + u_{j,t}$$

where $tfp_{j,t}^{v,s}$ is the log of TFP in year t of firm j born in year v and belonging to industry s; γ_s is a four digit industry effect; γ_v is the vintage effect; γ_t is the year effect; $h_{j,t}$ is a human capital measure; and $E_{j,t}$ is the log of stock of learning measured either by traditional learning index (equation (2.11)), my preferred learning index (equation (2.13)), or a simple turnover rate (equation (2.9)).

	1	2	3	4	5	6	7	8	9	10
Human Capital (h)	0.061	0.060	0.049	0.048	0.061	0.057	0.054	0.058	0.054	0.054
	(0.015)**	(0.015)**	(0.008)**	(0.008)**	(0.015)**	(0.015)**	(0.016)**	(0.016)**	(0.016)**	(0.016)**
Learning (E2)	0.010				-0.088	0.009	0.010	-0.054		
	(0.004)*				(0.018)**	(0.004)*	(0.004)*	(0.023)*		
Learning (E4)		0.014			0.097			0.062		
		(0.004)**			(0.018)**			(0.023)**		
Separation Rate (SR)			-0.098			-0.207		-0.122		
			(0.022)**			(0.039)**		(0.052)**		
Separation Rate: High				-0.020			-0.046			-0.015
				(0.006)**			(0.010)**			(0.038)
Learning (E2): High SR									0.006	0.007
									(0.004)	(0.005)
Learning (E2): Low SR									0.012	0.011
									(0.004)**	(0.004)**
Observations	9,041	9,036	23,236	23,236	9,036	9,041	9,041	9,036	9,041	9,041
R-squared	0.51	0.51	0.54	0.54	0.51	0.51	0.51	0.51	0.51	0.51

Standard errors in parentheses
Control variables include year, vintage, and four digit SIC
* significant at 5%; ** significant at 1%

Table 2.2: Basic Estimation Results

[24]Coefficients on industry, calendar year, and vintage are not reported to save space, but are available from the author.

In Table 2.2, I report basic estimation results. In all cases, the impact of human capital on TFP is highly statistically significant with the expected sign. In specification (1), the traditional measure of learning (E_2) has a positive and significant impact on TFP even after controlling for vintage and time effects (i.e. age effects)[25]. This finding is consistent with the idea of learning by "doing", not just learning by "aging". In the second specification, a new learning measure (E_4), which depreciates over time with worker turnover, also has a positive and significant effect on TFP. Moreover, the effect of this new measure is larger than that of the traditional measure. High (current) separation rates are associated with low productivity both in parametric and non-parametric specifications ((3) and (4), respectively)[26].

In regression (5), the traditional measure of learning (E_2) is entered together with the new measure of learning (E_4). The traditional measure, which assumes permanent learning, enters with a negative sign whereas the measure that allows learning to depreciate with turnover has a positive and significant effect on TFP. The two measures are highly correlated, and each has an individually positive and significant effect on productivity. The difference between the two coefficients in (5) suggests that worker turnover reduces a firm's experience, and that this loss has a negative effect on productivity. Given the same amount of past output, firms with lower past turnover rates are more likely to have high productivity.

Instead of interacting past outputs and turnover in a complicated and restrictive way, we can examine effects of cumulative output and turnover separately. Regressions (6)-(8) are based on equation (2.6),

$$tfp_{j,t}^{v,s} = \gamma_0 + \gamma_s + \gamma_v + \gamma_t + \gamma_h h_{j,t} + \gamma_e E_{2,j,t} + \gamma_s SR_{j,t} + u_{j,t}$$

[25]The coefficients associated with learning variables are elasticities. In regression (1), if we double past output then productivity increases by one percent. If one firm produced the same amount of output in its first and second year of business, TFP will increase by one percent between the second and third year.

[26]In equation specification (4), establishments with separation rates higher than the overall median level are classified as high separation establishments. Note that the sample used in (3) and (4) is larger than the other samples. As explained in Section 2.3.2, many observations are dropped due to incomplete histories (the "no-hole" condition).

where $E_{2,j,t}$ is the log of past output and $SR_{j,t}$ is the separation rate. In this case, both measures are statistically significant with expected signs, as shown in specification (6) and (7). Past output and turnover have independently significant contributions to TFP. In (7), I examine whether there is any intercept difference of TFP between high and low turnover firms, and find that high turnover firms have significantly lower TFP than low turnover firms. In column (8), I enter current turnover together with the two measures of learning. Controlling for past output, firms with historically low turnover rates (high E_4) and low contemporaneous turnover are more productive. In specifications (9) and (10), I examine differences in the speed of learning (as measured by the slope coefficient on past output) between firms with high and low current turnover based on equation (2.7):

$$tfp_{j,t}^{v,s} = \gamma_0 + \gamma_s + \gamma_v + \gamma_t + \gamma_h h_{j,t} + (\gamma_{eh} dh_{j,t} + \gamma_{el} dl_{j,t}) E_{2,j,t} + u_{j,t}$$

where $dh_{j,t}$ ($dl_{j,t}$) is a dummy variable taking one when the separation rate is high (low) and $E_{2,j,t}$ is the log of past output. In both regressions, the relative magnitude of slope coefficients and intercepts are as expected. In specification (9), the slope difference is statistically significant (with a t ratio of 4.58). In specification (10), where we allow both the intercept and slope to vary by firm turnover type, the slope difference is not statistically significant.

2.5 Robustness and IV Estimation

2.5.1 Endogeneity

Equation (2.6) implicitly assumes that there is a causal relationship running from cumulative output and worker turnover to firm productivity. However, both cumulative output and worker turnover may be correlated with firm productivity shocks. If this is the case, then the coefficient estimates reported above are inconsistent. There is a strong possibility of endogeneity between past experience and productivity heterogeneity across firms. It may be that firms that produced a lot in past years are more productive because of learning by doing, but it also could be that firms that produced a lot in past years did so because they were more productive to begin with.

Meanwhile, turnover may also be endogenous. Workers may leave their jobs in anticipation of a negative productivity shock in the firm where they currently work. Firms and workers may agree to separate if the match quality and thus productivity are low. If managerial ability is important in terms of both its direct contribution to productivity and its indirect contribution through match quality, then managerial ability will affect firm productivity as well as worker turnover. If this is the case, the estimation will only pick up correlations between two proxies for managerial ability. This endogeneity will result in inconsistent estimates. Thus, we need exogeneous instruments that affect past output and/or worker turnover but are not correlated with firm productivity shocks.

2.5.2 Instruments for Past Output

The instruments I adopt here are local demand shift instruments, based on Shea (1993a, 1993b) and Syverson (2004). Suppose that there are two plants of equal productivity in 1994. Assume one plant is exposed to high demand, and another to low demand, between 1995 and 1998. Let demand be equal in 1999. If there is learning by doing, then the plant with high demand (and thus high output) between 1995 and 1998 should have higher productivity in 1999 than the plant with low demand between 1995 and 1998. If there is no learning by doing, then there may still be a correlation across plants between 1995-1998 output and 1999 productivity, because productivity is persistent and because plants with more favorable productivity shocks after 1995 will produce more output after 1995. However, this correlation should disappear once we instrument 1995-1998 output with our demand-shift variable. For demand-shift instruments, I follow the input/output approach suggested in Shea (1993a, 1993b). Moreover, given the short panel aspect to available plant-level output and input data, I want to explore geographic variation within industries as in Syverson (2004).

The first step is to identify industries with a substantial local demand component. I follow Syverson's approach by defining local demand using BEA's Component Economic Areas (CEA). CEAs are collections of counties centered on Metropolitan Statistical Areas (MSAs). The BEA uses various indicators[27] of economic linkage to insure that CEAs are cohesive economically. There

[27] County-to-county commuting flows and regional newspaper circulation are examples of information used to construct CEAs.

SIC	Industry	< 60miles	< 100miles
3273	Ready-mixed Concrete	0.95	0.97
2951	Asphalt Paving Mixtures and Blocks	0.87	0.94
2711	Newspapers: Publishing, Or Publishing And Printing	0.89	0.93
2048	Prepared Feeds And Feed Ingredients For Animals And Fowls	0.66	0.78
2411	Logging	0.53	0.75
3281	Cut Stone And Stone Products	0.64	0.68
3272	Concrete Products, Except Block and Brick	0.49	0.64
3829	Measuring and Controlling Devices, Not Elsewhere Classified	0.62	0.63
2653	Corrugated And Solid Fiber Boxes	0.44	0.59
2448	Wood Pallets And Skids	0.43	0.58
3842	Orthopedic, Prosthetic, And Surgical Appliances And Supplies	0.55	0.57
2741	Miscellaneous Publishing	0.52	0.56
2439	Structural Wood Members, Not Elsewhere Classified	0.46	0.55
2869	Industrial Organic Chemicals, Not Elswhere Classified	0.47	0.54

Table 2.3: Local Industries

are 348 CEAs that are mutually exclusive and exhaustive of the land mass of the U.S.[28] The mean area of CEAs is 10,165 square miles, the mean radius is 57 miles. From Census Bureau's 1997 Commodity Flow Survey, I can calculate statistics on mileage of shipments by industry. In Table 2.3, I list industries that deliver over 50% of shipments (by weight) within 100 miles.

Among these industries, some do not have suitable downstream demand shift instruments. For industry A to be a good downstream instrument to industry B, A should have a high demand share of B's output, but the share of A's cost accounted for by industry B should be low. Shea (1993a, 1993b) presents both direct and ultimate cost and demand shares, where ultimate shares incorporate indirect input-output linkages. I use the maximum between the direct and ultimate shares to define the demand/cost share ratio between upstream and downstream industries.[29] Based on the 1992 input-output table, I find downstream industries, listed in Table 2.4, that are suitable instruments for the upstream industries listed in Table 2.3.[30]

[28] For reasons why a CEA might be better unit of local market than other geographic aggregations, see Syverson (2004).

[29] Shea suggests a demand share of over 15% and a ratio of demand share to cost share above 3 as a criteria on instruments. Here I choose 10% demand share and demand to cost share ratio of 2. For more details, see Shea (1993a, 1993b).

[30] Matt Freedman kindly helped me to find these downstream industries using the 1992 I/O table.

SIC	Instruments	DDS	UDS	DCS	UCS	DS	CS	DS/CS
3273	15 Building construction	0.40	0.25	0.05	0.02	0.40	0.05	8.06
	17 Construction special trade contractors	0.46	0.28	0.05	0.02	0.46	0.05	9.15
2951	17 Construction special trade contractors	0.55	0.33	0.02	0.01	0.55	0.02	30.75
	16 Heavy construction other than building	0.22	0.12	0.03	0.01	0.22	0.03	8.25
	15 Building construction	0.20	0.13	0.01	0.00	0.20	0.01	23.73
2048	02 Agricultural production livestock and animal specialties	0.75	0.12	0.18	0.03	0.75	0.18	4.06
2411	15 Building construction	0.00	0.15	0.07	0.03	0.15	0.07	2.27
	17 Construction special trade contractors	0.00	0.14	0.06	0.02	0.14	0.06	2.40
3281	15 Building construction	0.35	0.21	0.05	0.02	0.35	0.05	7.03
	17 Construction special trade contractors	0.33	0.20	0.05	0.02	0.33	0.05	6.58
3272	17 Construction special trade contractors	0.41	0.25	0.05	0.02	0.41	0.05	8.13
	15 Building construction	0.29	0.18	0.05	0.02	0.29	0.05	5.89
	16 Heavy construction other than building	0.20	0.11	0.04	0.02	0.20	0.04	5.07
3829	49 Electric, gas, and sanitary services	0.22	0.12	0.03	0.02	0.22	0.03	7.69
	37 Transportation equipment	0.18	0.14	0.06	0.04	0.18	0.06	3.14
2653	20 Food and kindred products	0.25	0.25	0.06	0.03	0.25	0.06	4.17
3842	80 Health Services	0.41	0.45	0.06	0.03	0.45	0.06	7.04
2439	15 Building construction	0.47	0.29	0.07	0.03	0.47	0.07	7.05
	17 Construction special trade contractors	0.45	0.27	0.06	0.02	0.45	0.06	7.46
2869	80 Health Services	0.08	0.15	0.04	0.02	0.15	0.04	3.69

Note: DDS is the direct demand share, UDS is the ultimate demand share, DCS is the direct cost share, and UCS is the ultimate cost share. DS=Max(DDS,UDS) and CS=Max(DCS,UCS)

Table 2.4: Local Industries with Relevant Downstream Industries

To be more specific, suppose there is a 3 year old firm which belongs to a locally operating industry A. The traditional learning variable is constructed as cumulative output over the past three years. I use the aggregate employment (or payroll) of all firms that belong to downstream industry B and that are operating in the same CEA as this firm during the past three years as an instrument for this learning variable.[31] These instruments are calculated using employment and payroll information from Business Register datasets. If there is more than one downstream industry, then I calculate the weighted mean, where the weights are demand shares.

2.5.3 Instruments for Turnover

When workers transition from one job to another, they do not randomly switch to any available job. Instead, this transition is driven by workers' skill and career choices, as well as by firms' desire to find workers with the right skills. As input-output tables produced by the BEA are used to study inter-industry linkages, we can similarly view employment-to-employment (E-to-E)

[31] I only use lagged values of local downstream activity. Contemporaneous activity of a downstream industry may not be orthogonal to the measured productivity shock if factor utilization is procyclical, even if downstream demand is orthogonal to the firm's true productivity shock.

transitions as inter-industry linkages in the labor market. If we know the origins and destinations of worker movements, then we can identify which industries supply which industries with labor. If industry A hires a significant number of workers who have left industry B, and as long as A's hiring is not related to B's productivity shocks, we can use hiring data for A to instrument for worker turnover in B.[32] If hiring demand is high for industry A, workers in industry B will get more job offers, resulting in more turnover in industry B.

Given the longitudinal and universal nature of the LEHD data, I can measure E-to-E flows for all workers and thus find origin and destination of worker flows. In the IV estimation, I use weighted hiring rates of destination industries within the state in which the firm is located.[33] In calculating E-to-E flows, I use male workers aged 25 to 55, since flows of this group are likely to be less affected by noisy flows out of the pool of unemployed.[34]

Workers may also leave their jobs if they get higher wage offers. In this case, high-wage firms have low worker turnover, since workers in these firms have a low chance of being offered higher wages from other firms[35]. Therefore, to the extent that firms' wage policies are permanent characteristics and not the result of current productivity levels, we can use firm wages as an instrument for turnover. For this purpose, the firm fixed effect, ψ, derived from the human capital estimation, is preferable to the wage itself. Average wages reflect not only firm wage policy, but also the average level of workforce quality. The firm fixed effect, on the other hand, is estimated controlling for workforce quality.[36]

[32]For this reason, I exclude all four-digit industries that belong to the same two-digit industry of the origin industry as possible instruments. Also, I exclude SIC 5812 (Eating Places) and SIC 7363 (Help Supply Services) since there are always high E-to-E flows to these two industries from any industry.

[33]Weights are average flows in 1996-2000 from origin industries.

[34]For more details, see Appendix E.

[35]Burdett and Mortensen (1998) derive a non-degenerate wage distribution resulting from this job switching motive. They derive wage distributions when firms are identical as well as when they are different in their productivity levels. In the latter case, wage offers turn out to be a monotonic transformation of underlying firm productivity.

[36]One nice feature of the firm effect is that it varies by firm (although it does not vary over time), and this will make the first stage estimation more accurate than when we use more aggregated instruments. However, if the firm effect estimate reflects variations other than firm wage policy, such as rent-sharing, then it may be correlated with the firm's productivity shocks. If this component is more important than firm wage policy, then my argument is

Finally, I use firm location-by-vintage dummies as instruments. The idea of using them as instruments is that they may reflect entrepreneurial type, which affects worker turnover patterns as well as other personnel policies.[37] When one decides to start a business, one must choose where to locate one's plant and when to start operating. To the extent that these decisions reflect firm or managerial type, we can use them as instruments. I use county as a location unit.[38] County-by-vintage dummies may not be perfectly exogenous if there exist region-specific productivity shocks that are persistent. However, county-by-vintage effects are time invariant. As with firm effect estimates, they may be correlated with the persistent component of productivity shocks, but are less likely to be correlated with the short run or transitory component.

2.5.4 IV Estimation Results

I have two different IV estimations. The first specification instruments only for learning variables. The second instruments only for the turnover measure. The reason for the separate estimations is that using local demand instruments for past output sharply restricts sample size. Since I use UI data from only seven states, we are already quite restrictive in the geographic dimension. In addition, we have a very small set of "local" industries that have good downstream demand-shift instruments. Therefore, the sample size when I instrument both past output and the turnover measure simultaneously is very small (fewer than 600 observations). Results in this case are very imprecise, and neither past output nor turnover is significant. However, if we instrument only for the turnover measure, the IV sample is almost the same size as the original OLS sample.[39] Meanwhile, if we ignore the turnover effect on productivity and use a proxy to control for workforce quality, we do not need LEHD information and therefore need not restrict the sample to seven

weaker. It is hard to measure the relative importance of firm wage policy *ex ante* and this is an empirical question.

[37] Haltiwanger, Lane, and Spletzer (2000) propose the idea that unobserved underlying types of firms generate persistent differences in firm behavior.

[38] Ideally, I would like to find a single or a small number of variables that can proxy for firm or managerial type. County-by-vintage dummies create more than a thousand dummy variables, so I lose quite a large number of degrees of freedom in the first stage regression.

[39] Instruments are unavailable for a few firms with no firm effect estimate, as well as for firms for which there are no flows to industries other than industries belonging to the same two digit SIC.

states. In the following, I use the simplest specification, as in equation (2.6), where past output and turnover affect productivity separately.

	1	2	3	4	5	6
	OLS	IV	IV	OLS	IV	IV
Past Output (E2)	0.052	0.269	0.246	0.053	0.273	0.249
	(0.006)**	(0.078)**	(0.072)**	(0.007)**	(0.079)**	(0.073)**
Log(Payroll/worker)	0.064	0.236	0.218	0.066	0.242	0.223
	(0.007)**	(0.062)**	(0.058)**	(0.007)**	(0.063)**	(0.059)**
Factor Utilization				0.019	0.066	0.061
				(0.011)	(0.021)**	(0.020)**
Observations	7,142	7,142	7,142	7,142	7,142	7,142
R-squared	0.33	0.22	0.24	0.33	0.22	0.24

	2	3		5	6
			First Stage		
Log(Cumulative weighted employment)	0.062			0.061	
	(0.008)**			(0.008)**	
Log(Cumulative weighted payroll)		0.059			0.059
		(0.007)**			(0.007)**
Observations	7,142	7,142		7,142	7,142
R-squared	0.801	0.801		0.804	0.805
Partial R-squared	0.008	0.009		0.008	0.009

Instrumented: LearnB
Standard errors in parentheses
Control variables include year, vintage, and four digit SIC
* significant at 5% level; ** significant at 1% level

Table 2.5: IV Estimation: The Impact of Past Output on Productivity

Table 2.5 reports IV results when we instrument for the traditional "learning" variable and ignore the impact of turnover on productivity. Here, I use national level data from the LRD and use payroll per worker to control for workforce quality.[40] In regressions (2)-(3), learning still has a significant effect in the IV specification. In the first stage, the local downstream demand-shift variables have a strongly positive impact on cumulative activity.[41] Even though the point estimate is much larger in the IV case than in the OLS case, the t ratio is smaller due to much higher standard errors. Also, note that the sample is reasonably large. Although we restrict the sample to industries that are locally traded and that have good downstream instruments, we

[40]Bahk and Gort (1993) use payroll per worker to control for workforce quality.

[41]Note that the first stage R^2 is very high. However, one should be careful in interpreting the result. As mentioned in Shea (1997), a high first stage R^2 does not necessarily imply that the local downstream demand shift instrument is highly relevant for past output. Since other controls (in this case, the year effect, vintage effect, and industry effect) may contribute to this high first stage R^2. I report the partial R^2 as suggested by Shea to check instrument relevance. It is much lower than the first stage R^2.

have data for all fifty states. Overall, these results support the learning by doing hypothesis. In regressions (4)-(6), I add a measure of cyclical utilization that is defined by the ratio of energy usage to capital stock normalized by its establishment level time average. One can argue that even if I instrument for past output, measured TFP can be high with persistent demand and procyclical factor utilization. Results show that past output significantly enhance current productivity even after controlling for factor utilization.

	1	2	3	4	5	6	7
	OLS	IV	IV	IV	IV	IV	IV
Human Capital (h)	0.057	0.043	0.057	0.048	0.050	0.056	0.056
	(0.015)**	(0.017)*	(0.018)**	(0.016)**	(0.016)**	(0.016)**	(0.016)**
Separation Rate (SR)	-0.207	-0.920	-0.217	-0.696	-0.589	-0.267	-0.245
	(0.039)**	(0.306)**	(0.441)	(0.247)**	(0.250)*	(0.096)**	(0.095)*
Past Output (E2)	0.009	0.007	0.009	0.008	0.008	0.009	0.009
	(0.004)*	(0.004)	(0.004)*	(0.004)	(0.004)*	(0.004)*	(0.004)*
Observations	9,041	9,039	9,041	9,039	9,039	9,041	9,041
R-squared	0.51	0.49	0.51	0.50	0.51	0.51	0.51

	2	3	4	5	6	7
			First Stage			
Firm Effect (psi)	-0.065		-0.066	-0.066		
	(0.005)**		(0.005)**	(0.005)**		
Weighted Job-to-Job Hiring Rate		0.436	0.454			0.376
		(0.052)**	(0.052)**			(0.376)**
Weighted Hiring Rate				-0.041		
				(0.005)**		
County*Vintage					1.40#1	1.36#3
					(0.000)#2	(0.000)#4
Observations	9,039	9,041	9,039	9,039	9,041	9,041
R-squared	0.14	0.13	0.15	0.15	0.27	0.27
Adjusted R-squared	0.10	0.09	0.11	0.11	0.13	0.13

Instrumented: Separation Rate
Standard errors in parentheses
Control variables include year, vintage, and four digit SIC
* significant at 5% level; ** significant at 1% level
#1 $F_{(1096,7543)}$ statistic, #2 p-value
#3 $F_{(1096,7542)}$ statistic, #4 p-value

Table 2.6: IV Estimation: The Impact of Turnover on Productivity

Table 2.6 reports seven IV specifications where we instrument only for the separation rate. I use the firm effect on wages, weighted job-to-job hiring rates of destination industries, weighted hiring rates of destination industries, and county-by-vintage dummies as instruments. The job-to-job hiring rate is defined as the ratio of new workers hired from other firms to total employment. Thus, the difference between hires and job-to-job hires represents new hires from the pool of unemployed. In the first IV estimation, I use only the firm wage effect as an instrument. The separation rate still has a negative and significant impact on productivity, although the standard error is now much larger than in the OLS estimation. Past output is now significant at only the 10% level. In the first stage, firms with a high wage premium enjoy low worker turnover. The second IV specification uses the weighted job-to-job hiring rate of destination industries as the instrument for turnover. In the first stage, firms with higher job-to-job hiring demand for their workforce suffer from high turnover. However, the impact of separation on productivity in the second stage is not significant. The third IV specification uses the firm wage effect and the weighted job-to-job hiring rate of destination industries as instruments. The first stage R^2 is now slightly higher than those of the first and the second IV specifications. The firm wage effect and the weighted job-to-job hiring rate of destination industries have significant impacts on the separation rate with expected signs. In the second stage, the separation rate again has a negative and significant impact on productivity with a lower standard error. The fourth IV specification finds that the weighted overall hiring rate of destination industries (including hires from the pool of unemployed) has a negative effect on turnover.[42] Nonetheless, the second stage still implies that past output and turnover significantly affect productivity. The last two IV specifications use county-by-vintage dummies as instruments. Note that the standard error for the coefficient on the separation rate is quite low, reflecting a high first stage R^2. In the first stage, we can see that county-by-vintage dummies are highly relevant for worker turnover, even when I adjust R^2 for degrees of freedom[43].

2.6 Summary

The existing literature on learning by doing suggests that firms learn from their past production activity. In this paper, I argue that firm learning is affected not only by past output, but also by worker turnover, as learning is embodied in workers. High worker turnover makes learning by doing more difficult, because firms lose workers who have "learned" about the production process from their past activity.

[42] I do not have any good explanation on this finding. One possibility is that a high value of hiring from unemployment may indicate a depressed labor market with few quits.

[43] However, when I do overidentifying restrictions tests, they suggest that I have "too many instruments".

I integrate unique employer-employee matched data from the LEHD Program with detailed business information from the LRD. This matched dataset enables us to look deeply inside businesses and to characterize and measure worker flows, workforce composition, and thus the nature of learning. The data also allow us to explore the relationship between establishment-level productivity and learning by doing. I show that worker turnover has a significant negative effect on productivity. Firms with relatively low turnover not only have a relatively high level of productivity, but also learn more quickly. Turnover patterns of different worker groups are quite different. Firms with the most stable workforce are the most productive as well as the fastest learners.

Learning and turnover appear to have causal effects on productivity, although sample size issues preclude me from showing this with my preferred specification. I develop new instruments based on local downstream demand, firm wage policy, regional hiring conditions of related industries, firm location, and firm vintage. These instruments significantly affect past production experience and worker turnover, but are not necessarily affected by firm productivity shocks. With these instruments, I present evidence that learning by doing has causal effects on productivity.

In the IV estimations, I could not instrument both past output and turnover because of a small sample problem. The main reason for this is that the instrument for past output is restricted to "local" downstream industries. Note that I only use the UI information of seven states. In the future, I will be able to include more states, and therefore will be able to solve this small sample problem. Although I instrument only past output, the IV estimation strongly supports learning by doing effects on firm productivity. However, the IV estimation for turnover gives us mixed results. E-to-E worker flows, which is my preferred instrument, support the idea that worker turnover has a negative and causal impact on productivity only when it is used with other instruments that are less likely to be exogenous to firm productivity shocks. I leave these problems for future research.

Chapter 3

Firm Performance, Workforce Quality, and Workforce Churning

What factors contribute to the success and survival of a business? Underlying this important question are even more basic questions: what constitutes a business, how has this changed over time, and have the factors contributing to success changed? In the old economy (or at least the common view of the old economy) a business is typically characterized by economists as a production location that combines inputs like capital, labor and materials to produce goods. The benchmark old economy model permits little role for differences across businesses in how they organize themselves, especially within industries.[1]

For the new economy (or perhaps even a new view of the old economy) there is a perspective that a business is more difficult to define in terms of just the standard measures of outputs and inputs. From this new view, the mix and scale of these inputs may differ substantially across businesses even if they are producing similar products and services. Moreover, the key inputs for a new economy business that are often emphasized are human capital and high tech physical capital, and recent evidence suggests that even businesses in narrowly defined sectors combine these two key inputs in a myriad of ways. In a related fashion, the increasing emphasis on producing services

[1] To be more precise, research in industrial organization has a long tradition of studying the organization of firms in terms of the mix of establishments that constitute a firm. While the issues of horizontal and vertical integration of firms are closely related to the issues emphasized in this study, our focus is more on the heterogeneity across establishments and firms inside narrowly defined industries, which has only recently become the focus of the literature on industry evolution. The neglect of within industry heterogeneity is a common limitation of the statistics on businesses produced by the statistical agencies. Indeed, the definition of a specific industry is often based in part on the presumption that the production processes for the businesses classified in the same industries are similar. Interestingly, the recently developed NAICS emphasizes this basis for defining industries relative to the earlier SIC basis. In some ways, the findings and the perspectives from this and related studies that emphasize the heterogeneity in firm performance within narrowly defined industries raises fundamental questions about this conceptual basis for defining industries.

as opposed to goods has raised questions about what constitutes the boundary of a firm and the role of location as a defining characteristic of a firm. For a bricks and mortar business that produces or sells physical goods, it is clear that the physical location of a business is a fundamental characteristic of the business. For a producer of services or even high tech goods like software, the information revolution has changed the importance of physical proximity in the organization of the production of the service or the good.

The development of longitudinal business databases for the U.S. and other countries has provided a rich new perspective on how to think about what constitutes a business. These databases permit the measurement of firms and establishments and include measures of productivity along with the underlying measures of outputs and inputs. The dominant empirical finding is the overwhelming importance of idiosyncratic factors in terms of firm performance measured in a variety of ways. Within narrowly defined industries, some businesses are growing while others are shrinking, some businesses are entering while others are exiting, some businesses are increasing the skill mix of their workforce while others are decreasing the skill mix, some are adopting advanced technologies and some are not, some are increasing their capital intensities while others are decreasing their capital intensities, and some are exhibiting increases in productivity while others are exhibiting decreases in productivity. Moreover, the heterogeneity in key firm performance outcomes like growth, survival and productivity are linked to enormous heterogeneity in how firms are organizing themselves.

The recent literature has shown that all of this heterogeneity is important not only for understanding micro business dynamics but also for understanding aggregate dynamics. In particular, the recent literature has shown that aggregate productivity growth depends critically on the efficiency of the ongoing churning of jobs and firms. That is, a large fraction of aggregate productivity growth is accounted for by reallocation of outputs and inputs from less productive to more productive businesses. In this context, a key part of the productivity enhancing churn is generated by the entry and exit of businesses as entering businesses tend to be much more productive than the exiting businesses they displace. These entry and exit dynamics are intimately connected since for each wave of entrants there are selection and learning dynamics that shape the evolution of young businesses. That is, in the first several years for each cohort, we observe many young businesses fail, with those that fail being the least productive and those that survive exhibiting rapid growth in both activity and productivity, suggesting a form of learning-by-doing.

A key theme of this study is to explore these issues for five selected industries: financial services, retail food, semiconductors, software and trucking. These industries span services and goods producing industries as well as traditional and high-tech industries. Our objective is to combine the case study contextual information about the detailed workings of the industry with the insights from the new longitudinal matched employer-employee datasets that are at the core of this analysis.[2] Relative to the findings of the recent literature discussed above, an additional contribution is that much of the above referenced literature exploits longitudinal business databases that did not have much information about the characteristics of workers and in turn the human resource practices of businesses. In the analysis of this study, we focus our attention on the connection between firm performance and measures of workforce churning and workforce quality.

At first glance the role of workforce quality in firm performance appears to be straightforward. That is, a higher quality workforce should result in higher productivity. Part of our analysis is to explore whether this straightforward relationship holds. However, another aspect of the relationship between workforce quality and firm performance is not so straightforward. That is, it is not clear that a higher quality workforce should lead to greater survival. The reason is that a higher quality workforce may lead to higher productivity but not necessarily higher profitability. Simply put, a higher quality workforce will yield higher value-added output but also, in principle, a higher wage bill. However, there may be a number of factors that imply that a higher quality workforce is positively correlated with profitability and survivability. For one, there may be some wage compression so that wages do not fully reflect worker quality. As such, businesses that attract and retain the best workers will in turn have higher profits as well. In a related fashion, a higher quality workforce may be correlated with key unobserved factors that are related to the success

[2]This chapter is part of a collaborative project between the Sloan industry centers for these selected industries and the LEHD project at the U.S. Bureau of the Census.

of a business. Suppose for example that high quality owners/managers are even more productive if they surround themselves with a high quality workforce. This type of complementarity between managerial ability and worker ability would lead to a positive correlation between workforce quality and survival.[3]

In terms of the relationship between worker churning and firm performance, it is important to recognize that some worker churning is undoubtedly part of a healthy firm. Both life cycle factors and turnover associated with efficient matching (referred to as life cycle turnover and match turnover in what follows) may be fully efficient. Life cycle factors inducing worker turnover include workers entering and exiting the labor market for a variety of reasons (new entrants, exits for schooling/children/retirement, re-entry after schooling/children). Efficient matching turnover results when workers and firms make a match about whose quality they are initially uncertain (i.e., is the worker well-suited for the job relative to outside options? is the worker happy with the job relative to outside options?)

Beyond life cycle and efficient match turnover, there may be inefficient turnover related, for example, to workforce practices and/or managerial practices and abilities. Evidence from existing studies suggests that firm performance varies dramatically within industries and it often speculated that such variance reflects managerial ability or deliberate choices regarding managerial or workforce practices (see, e.g., Haltiwanger, Lane and Spletzer (2003)). Along these lines, one potentially important characteristic of a good manager often emphasized in the literature is the ability to attract and retain good workers (Ichinowski and Shaw (1997), Black and Lynch (2001)). Workforce practices that contribute to an attractive work environment and loyalty to the firm likely include wage and benefit practices that reward loyalty (e.g., upward sloping wage-firm specific tenure profiles), the provision of training that benefits the worker and the firm, and the use of effective teams for production so that workers have a voice in and are rewarded for the productivity of their teams. All of these practices should reduce worker turnover but as noted above should not eliminate worker turnover. As such, in what follows we explore the relationship between our

[3] Some evidence in favor of this idea using LEHD data is provided in Abowd et. al. (2003). In the latter paper, firms with high quality workforce are found to have higher market value.

measure of churning and firm performance.

This chapter is organized as follows: The next section discusses how we define and measure key variables. Of particular importance is our ability to make significant improvements to past measures through use of a linked employer-employee database. At the same time, we also identify those economic factors that still pose measurement challenges to researchers. Section 3.2 uses these measures to generate a set of basic empirical facts about each of the five Sloan industries. Section 3.3 shows basic patterns across entering, exiting, and continuing establishments. Section 3.4 takes a closer look at the complex relationship between traits of a business' workforce (such as worker skill and turnover) and productivity. In Section 3.5, we explore whether patterns that are true for the sector as a whole are also true for various sub-groupings of businesses within the industry. In this context, we explore some of the key idiosyncratic features of the characteristics of the businesses in each of our five industries. The final section summarizes and highlights the significance of our empirical findings.

3.1 Measurement Challenges

Two main types of data we use in this study are the 1992 and 1997 Economic Censuses and the Longitudinal Employer Household Dynamics (LEHD) databases. Variables available from Economic Censuses are revenue, employment, payroll, establishment identifier, and firm identifier. Given that we have two Economic Censuses we can identify establishments' entry and exit behavior. For 1992 establishments, we can identify whether they survive or exit until 1997. For 1997 establishments, we can identify whether they are new entrants or existed in 1992 (incumbents). It is possible to identify not only entry/exit of establishments but also entry/exit of parent firms so that mergers/acquisitions and firm entry/exit can be quantified and analyzed. From LEHD, we can get measures of worker turnover and workforce quality at the business level.

The longitudinal employer-employee data that we use in this study permits an unprecedented look inside businesses. In what follows, we provide some exploratory analysis of the relationship between measures of firm performance (measured here by proxies for productivity and by survival) and measures of workforce churning and workforce quality. For the latter we use indirect measures in some cases as well so while we offer a rich new perspective, we also must recognize the measurement challenges and limitations of this analysis.

What do we measure well? We measure the entry and exit of establishments and the organization of establishments into firms well. We measure revenue, employment, job flows, worker flows, earnings, and workforce composition well. For firm performance, the measurement of entry and exit dynamics is important, as a key indicator of performance is survival.

What do we measure less well? To start, our measures of productivity (like much of the micro and aggregate literature) are crude at best. Our measure of productivity in what follows is gross output per worker, where gross output is measured as gross revenue deflated with a detailed industry deflator. This is a crude measure of labor productivity. This crude measure is closely related to the measures of gross output per unit of labor that are published by the BLS (who typically use gross revenue data from Census as the primary source data for gross output) and that are used extensively in the literature. For some industries, gross output per worker is not a bad proxy for productivity. For example for the manufacturing sector, a variety of studies have shown that at the establishment-level, labor productivity measured in this manner is highly correlated with value-added per worker and even carefully measured multi-factor productivity (with careful treatment of the measurement of output and inputs including physical capital, labor and materials). However, for the non-goods producing industries, gross output per worker measures of productivity are often problematic. Recent studies have shown that in many service industries, measures of labor productivity based upon gross output per worker at the aggregate level have yielded implausible negative productivity growth in the 1990s (see, e.g., Corrado and Slifman (1999) and Gullickson and Harper (2002)). The problems with gross output per worker are especially severe in those industries where the product or service is difficult to measure. A related problem is that in some sectors it is especially difficult to allocate the output of a firm to individual establishments. In our case, these problems are particularly severe in the financial services sector and in what follows we explore the limitations of our measures for this industry.

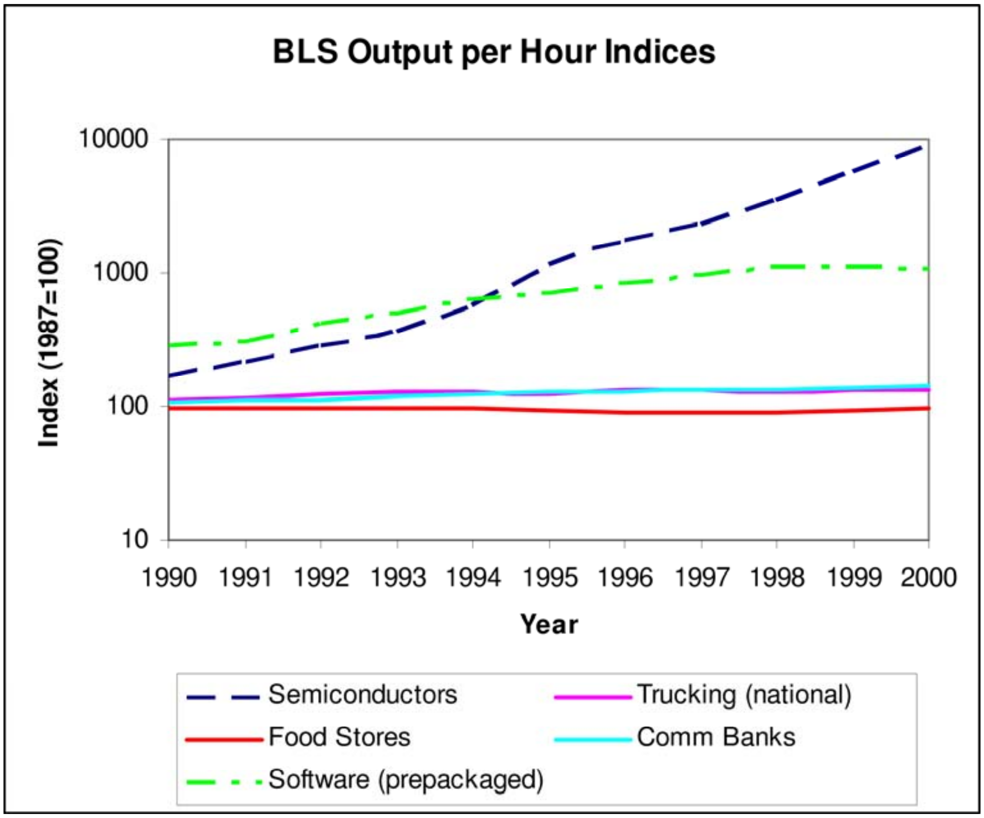


Figure 3.1: BLS Output Per Hour Indices

To gain some perspective on the challenges of measuring productivity for our industries, Figure 3.1 depicts the BLS output per hour index for key 4-digit industries that are part of the five somewhat broader sectors that are the focus of this study. A log scale for the vertical axis is used because of the dramatic increases in the productivity index for the semiconductor industry. The latter is largely driven by the tremendous decreases in semiconductor price index measures that take into account the enormous efficiency/quality improvements in semiconductors (via hedonic price indices). At the other end of the scale, the official BLS indices suggest little or even declining productivity for food stores, commercial banks and trucking. As noted above, it is not uncommon to find modest or even declining productivity for many non-goods producing industries in the 1990s. An open question is the extent to which this poor productivity performance is real or reflects measurement difficulties especially in difficult to measure sectors like financial services.[4]

[4]As will become apparent below, our biggest problem with productivity measurement is also with financial

Another related problem is that our revenue measure is gross revenue. While for some industries, we can measure value-added at the firm-level for a sample of firms (especially for manufacturing industries) we focus our attention on gross revenues since this measure is readily available for all businesses. Given our focus on the impact of the entry and exit of firms and establishments this is important, as value-added measures are often not available for small and young businesses. Value-added per worker would be the preferred concept but a number of studies have shown that value-added per worker is highly correlated with gross output per worker across firms within the same industry (see, e.g., Foster, Haltiwanger and Krizan (2001)). An obvious limitation is that gross output per worker measures in levels (not growth rates) are not comparable across industries. In what follows, this limitation will be self-evident for the retail food industry. In retail food, we measure gross revenue per worker not taking into account the cost of the goods sold (as we do not measure gross margins at the micro level). Much of the gross revenue in retail food (and in retail more generally) is accounted for by the cost of goods sold. As such, we find that gross revenue per worker is very high relative to gross revenue per worker in the software and semiconductor industries, which is quite misleading. For the most part, we focus on the growth of revenue per worker or we only consider variation within industries, so that this problem with measurement levels across industries is not relevant.

In what follows, our primary focus is on the relationship between firm performance (measured as revenue per worker and survival) and measures of workforce quality and workforce churning. For workforce quality, we take advantage of the measures of human capital developed by Abowd, Lengermann and McKinney (2003). These measures are based upon a statistical decomposition of the wage for a worker into a person effect, a firm effect and time varying person characteristics including general labor market experience. The person effect is the portable component of a worker's wage and as such is a good summary measure of the general skills of a worker (and indeed studies have shown that it is highly correlated with direct measures of skills such as

services. We should note in this regard that BLS uses the gross revenue measures that we use for all of our sectors except for financial services (for the latter they attempt to measure the service flow from financial service providers). Even with their alternative approach, there are anomalous results for the financial services sector.

education). In what follows, we use two measures of workforce quality: overall human capital, measured as the person effect plus the labor market experience component, and the person effect by itself. It may be that the contribution of these different components of human capital interacts with firm performance differently across businesses and industries.

Our measures of human capital are available at the individual worker level but the focus here is on measures of workforce quality at the business level. We construct summary measures of these measures at the business level by estimating kernel densities for every business in the LEHD dataset. Using such kernel densities a wide range of summary measures of the distribution of workforce quality are available. Here we use two simple measures based upon the fraction of workers at a business that are above the economy-wide median human capital measure (using 1997 as the reference year for the economy-wide median). We compute this fraction for both the overall human capital measure and the person effect.

We are also interested in exploring the role of workforce churning. For this purpose, we focus on a measure of excess worker reallocation that is likely to be related to internal labor markets and workforce practices. This measure at the business level is given by:

$$\frac{Accession_t + Separation_t - |Employment_t - Employment_{t-1}|}{(Employment_t + Employment_{t-1})/2}$$

This measure captures the component of worker turnover or reallocation that is in excess of that needed to accommodate any net changes in the number of workers in the business. Whether it represents any excess in an efficiency sense is an open theoretical question and part of our investigation. This is a topic we take up in detail below.

Before proceeding, it is important to discuss the unit of observation for this analysis. In this chapter, the unit of observation is typically the establishment. That is, for performance we measure the productivity and survival of establishments. However, our data permit linking the establishments to the parent firms and many of our exercises exploit this information. For example, we distinguish between entering establishments that are new firms and entering establishments for existing firms. For measuring revenue, employment, payroll, firm linkages, and survival the primary sources of information are the Economic Censuses. Our workforce quality and workforce churning measures are developed from the matched employer-employee datasets from the LEHD project. We integrate these measures at the establishment-level with our Census based measures by matching LEHD data to Census data at the EIN by County by 2-digit SIC level of aggregation. For most businesses, this match is at the establishment-level. When the match is at a higher level of aggregation (e.g., for a firm that has multiple establishments in the same county and same industry), we assume that the workforce quality and workforce churning are the same across establishments in the EIN by County by 2-digit industry cell.

3.2 Basic Facts

Sector	Year	Revenue /worker ($)	Churning Rate (%)	Human Capital (%)	Person Effect (%)	Employment (Number)	Payroll /worker ($)
Financial	1992	143,814	16.8	48.3	57.1	18.9	24,433
Services	1997	117,857	15.9	61.6	63.1	18.2	23,397
Retail	1992	138,176	28.7	31.2	46.7	16.1	9,343
Food	1997	140,355	24.3	40.6	50.3	16.6	9,068
Semiconductors	1992	141,306	13.2	56.6	48.4	82.4	26,873
	1997	555,483	13.2	65.7	53.9	84.5	28,188
Software	1992	116,952	20.2	72.3	74.1	19.0	35,220
	1997	139,924	17.1	79.0	77.0	23.0	38,671
Trucking	1992	97,891	26.9	54.5	39.3	13.9	17,547
	1997	99,313	21.3	67.4	46.0	14.1	17,307

Table 3.1: Means of Core Variables

We begin by reporting basic facts about firm performance and measures of workforce quality and churning across time and across industries. Table 3.1 reports mean values of core variables for the five industries under consideration by year[5]. In terms of labor productivity (measured again as real gross revenue per worker), one can see a huge increase in measured productivity for semiconductors, a significant increase in software, modest increases in trucking and retail food but a significant decrease for financial services. However, the results for financial services may be due

[5]In all the succeeding tables, we refer to our measure of gross output per worker as revenue/worker. As noted in the text, the measure is gross revenue deflated with an industry deflator per worker. For earnings per worker, we measure this in real terms by deflating payroll with the CPI and dividing by the number of workers at the business.

to poor measurement of productivity. High revenue per worker in retail food reflects use of gross revenue rather than gross margins. The huge increase in revenue per worker for semiconductors is driven substantially by a dramatically decreasing price index for semiconductors, reflecting the use of hedonic price indexes for this industry. Hedonics are used for only a handful of industries in the price deflators produced by the statistical agencies, including semiconductors. As has been noted elsewhere in the literature, the enormous changes in the characteristics of semiconductors (and other key IT products) over this period of time have led to rapid rises in measured productivity in key IT industries.[6]

Churning rates are high in retail food and trucking and lower in semiconductors and financial services. The average variation across industries is substantial with churning in retail food and trucking almost twice the rate in semiconductors. Churning rates are somewhat lower in 1997 than in 1992, which might reflect both trend and cyclical effects.

Businesses in software and semiconductors have, on average, a high fraction of high human capital workers, whether the overall human capital or person effect measure is used. In contrast, retail food has low human capital. Trucking has surprisingly high overall human capital but low person effect human capital. The implication is that the surprisingly high human capital of trucking is being driven by high experience as opposed to high general skills. In contrast, financial services has very high person effect human capital but relatively low overall human capital, suggesting that labor market experience is low in financial services. All of the industries exhibit substantial increases in human capital over the 1992-97 period for both overall and person effect human capital.

The average size of establishments is much larger in semiconductors than in the other industries and especially small for trucking and retail food establishments. Average size has been relatively stable over time with modest increases in establishment size between 1992 and 1997 in software.

Earnings are highest in software and semiconductors followed by financial services, trucking and retail food. Workers in software have average earnings that are about four times larger than

[6]There has been some recent work at BEA exploring hedonics for software but this research has not been incorporated into the official statistics.

in retail food. The low earnings per worker in retail food is not surprising but also helps remind us that the level of revenue per worker is a highly misleading indicator of value-added per worker. If workers are roughly paid their marginal products, we would have expected that the rank ordering of revenue per worker to be roughly the same as the rank ordering of payroll per worker, but this is far from the case.[7]

Sector	1992		1997	
	Type	Share	Type	Share
Financial Services	Exiter	39.5	Continuer	65.4
	Continuer	60.5	Entrants	34.6
Retail Food	Exiter	38.6	Continuer	66.7
	Continuer	61.4	Entrants	33.3
Semiconductors	Exiter	25.2	Continuer	70.0
	Continuer	74.8	Entrants	30.0
Software	Exiter	38.7	Continuer	47.4
	Continuer	61.3	Entrants	52.6
Trucking	Exiter	40.1	Continuer	63.3
	Continuer	59.9	Entrants	36.7

Table 3.2: Entry/Exit (Establishments)

Table 3.2 reports shares of continuers, exiters and entrants amongst establishments in 1992 and 1997 respectively. By exit of an establishment in this context, we mean that the establishment truly exited - i.e., is no longer in operation. Similarly, by entry of an establishment we mean that the establishment did not previously exist. The 1992 continuer rates can be interpreted as survival rates since they reflect the fraction of establishments in 1992 that survived until 1997. Survival rates are around 60% for all sectors except semiconductors where 75% of establishments survived until 1997. The 1997 continuer rates can be quite different from the 1992 rates depending on the pace of entry and exit. For example, in 1997 only 52.6 percent of software establishments are continuers even though 61.3 percent of establishments survived from 1992 to 1997. This difference

[7]Of course, it is an open empirical question whether productivity per worker and payroll per worker are similarly ranked as many factors may impact their relationship.

reflects the enormous entry rate in software over this period - the entry rate is so large that there are more "new" than "old" establishments in software in 1997. The entry rate in software is much higher than those of other sectors, which are in the 30-37% range. The high entry rate in software is obviously associated with the IT boom in late 1990s.

Sector	1992			1997		
	Type		Share	Type		Share
	Estab	Firm		Estab	Firm	
Financial Services	Exiter	Exiter	20.6	Entrants	Entrants	18.0
		Continuer	18.9		Continuer	16.6
	Continuer	Different	7.8	Continuer	Different	8.1
		Same	52.7		Same	57.4
Retail Food	Exiter	Exiter	34.0	Entrants	Entrants	27.4
		Continuer	4.6		Continuer	6.0
	Continuer	Different	4.8	Continuer	Different	4.4
		Same	56.6		Same	62.3
Semiconductors	Exiter	Exiter	22.4	Entrants	Entrants	28.6
		Continuer	2.7		Continuer	1.4
	Continuer	Different	6.1	Continuer	Different	5.5
		Same	68.8		Same	64.5
Software	Exiter	Exiter	34.9	Entrants	Entrants	48.4
		Continuer	3.8		Continuer	4.2
	Continuer	Different	4.1	Continuer	Different	3.2
		Same	57.3		Same	44.2
Trucking	Exiter	Exiter	37.9	Entrants	Entrants	34.6
		Continuer	2.2		Continuer	2.1
	Continuer	Different	2.5	Continuer	Different	2.0
		Same	57.4		Same	61.2

Table 3.3: Entry/Exit (Establishments and Firms)

Table 3.3 shows more detailed results on entry/exit incorporating parent firms' entry/exit and merger/acquisition. Entry and exit of firms is conceptually different than the entry and exit of establishments and reflects administrative, organizational and ownership changes. Formally, entry and exit of firms here reflects changes in the firm identification number. A firm obtains a firm identification number if it is truly a new firm (starts up with all new establishments) or has undergone an administrative or organizational change requiring a new firm identification number. Similarly, a firm identification exit occurs when all establishments of the firm cease operations, the firm is sold in its entirety, and/or it undergoes a organizational change (e.g., changes legal form of operation or changes from a single-unit firm to a multi-unit firm).

Caution must be used in combining the establishment entry and exit concepts with the firm entry and exit concepts used here. An establishment may exit because its parent firm closes or because its parent firm closes some of its establishments (downsizing, restructuring, etc.). In both cases, the establishment exits - the open question is whether the parent firm is also ceasing operations or changing structure. In addition, surviving establishments may change ownership. New establishment entrants might be associated with an existing firm or represent a totally new firm. Again, in both cases, new establishments are true establishment entry but they may be part of a new firm or a new establishment of an existing firm. In Table 3.3, we denote exiter/exiter as cases where the establishment exited and the firm identification exited as well, and we denote entrants/entrants as cases where the establishment entered and the firm id entered as well.

Based on Table 3.3, it is clear that most entry and exit of establishments is associated with the entry and exit of firms. This reflects the fact that most firms have only one establishment and the distinctions between establishment and firm entry and exit are not important. However, in the financial services sector there is very high entry and exit of establishments from continuing firms. In terms of ownership changes of establishments, the pace is relatively modest but with the highest rates in financial services and in semiconductors. The high entry and exit rates of establishments for continuing firms and the high ownership changes for establishments in financial services reflects the rapid restructuring process and M&A in this sector during 1990s.

	Revenue /Worker	Churning Rate	Human Capital	Person Effect	Size	Payroll /Worker
Revenue /Worker	1.000	**-0.048**	**0.045**	**0.049**	0.005	**0.571**
Churning Rate		1.000	**-0.096**	**0.097**	**-0.033**	**-0.116**
Human Capital			1.000	**0.499**	**-0.047**	**0.179**
Person Effect				1.000	0.009	**0.126**
Size					1.000	**0.206**
Payroll /Worker						1.000

Table 3.4: Correlations (Financial Services)

	Revenue /Worker	Churning Rate	Human Capital	Person Effect	Size	Payroll /Worker
Revenue /Worker	1.000	**-0.053**	**0.118**	**0.039**	**-0.081**	**0.452**
Churning Rate		1.000	**-0.235**	**0.070**	-0.004	**-0.115**
Human Capital			1.000	**0.416**	**0.223**	**0.337**
Person Effect				1.000	**0.260**	**0.150**
Size					1.000	**0.288**
Payroll /Worker						1.000

Table 3.5: Correlations (Retail Food)

	Revenue /Worker	Churning Rate	Human Capital	Person Effect	Size	Payroll /Worker
Revenue /Worker	1.000	**-0.092**	**0.380**	**0.302**	**0.133**	**0.547**
Churning Rate		1.000	**-0.135**	0.041	**-0.116**	**-0.156**
Human Capital			1.000	**0.642**	**0.097**	**0.516**
Person Effect				1.000	0.027	**0.371**
Size					1.000	**0.283**
Payroll /Worker						1.000

Table 3.6: Correlations (Semiconductors)

	Revenue /Worker	Churning Rate	Human Capital	Person Effect	Size	Payroll /Worker
Revenue /Worker	1.000	**-0.081**	**0.313**	**0.222**	**0.062**	**0.671**
Churning Rate		1.000	**-0.121**	0.015	-0.008	**-0.122**
Human Capital			1.000	**0.530**	**0.099**	**0.497**
Person Effect				1.000	**0.114**	**0.340**
Size					1.000	**0.174**
Payroll /Worker						1.000

Table 3.7: Correlations (Software)

	Revenue /Worker	Churning Rate	Human Capital	Person Effect	Size	Payroll /Worker
Revenue /Worker	1.000	**-0.072**	**0.282**	**0.086**	**-0.077**	**0.594**
Churning Rate		1.000	**-0.223**	0.012	**-0.063**	**-0.131**
Human Capital			1.000	**0.344**	**0.137**	**0.431**
Person Effect				1.000	**0.072**	**0.127**
Size					1.000	**0.248**
Payroll /Worker						1.000

Table 3.8: Correlations (Trucking)

Tables 3.4 - 3.8[8] report correlation coefficients among core variables for each of the industries. All values in this table are calculated using deviations from mean values at the four-digit SIC level. A bold item in the table indicates that the correlation is statistically different from zero. The patterns in the table make sense for the most part but also help highlight some of the measurement challenges. For all sectors, we observe that labor productivity is positively correlated with human capital (both overall and person effect) as well as with payroll per worker and is negatively correlated with churning. In a related fashion, in all sectors we observe that earnings per worker is positively correlated with human capital (both overall and person effect) as well as with employer size but is negatively correlated with churning. We also find that employer size is positively correlated with human capital (both overall and person effect). In terms of the relationship between human capital and churning rates, there is some tendency for overall human capital to be inversely correlated with churning but some weaker evidence that the person effect is positively correlated with churning. Putting these results suggests that the negative relationship for the overall measure presumably reflects the fact that businesses that are dominated by experienced workers are less likely to exhibit churning.

There are some exceptions to these patterns. For example, in retail food and trucking measured labor productivity is inversely correlated with size. Moreover, the magnitudes of the correlations vary substantially. For example, the correlation between labor productivity and human

[8]In all tables, coefficients in bold indicate statistical significance at the five percent level.

capital is very high in semiconductors and software and close to zero in financial services. The generally weaker correlations of measured productivity with other variables in financial services are presumably related to the difficult measurement challenges in measuring productivity in this sector. The finding that at least the sign of the cross sectional correlations between productivity and the other variables is consistent with other industries suggests that there is some information content in the revenue data collected for the financial services sector. In what follows, we will see that there are some sub-industries within the financial services sector where the results on productivity are more sensible.

The main focus of this chapter is on the relationship between productivity, human capital and churning where we find clear and striking patterns. Businesses with higher human capital are also more productive as expected. The finding that businesses with higher churning rates are less productive suggests that the high churning businesses within an industry may be experiencing inefficiently high churning. That is, even if there is, as suggested in the introduction, a more complex nonlinear relationship between churning and efficiency, these results suggest that the overall relationship is negative.[9] Note that even if the relationship between efficiency and churning is negative, there may not be a negative relationship between churning and profitability. For example, it may be to reduce churning, firms must pay higher wages so there is a tradeoff between efficiency and costs. We cannot address this issue directly but since we examine the relationship between survival and churning this provides an indirect means of investigating the impact on profitability.

3.3 Basic Patterns Across Entering, Exiting and Continuing Establishments

In this section, we explore the role of entry and exit more fully with a focus on the connection between survival, workforce quality and workforce churning. We begin this analysis with some basic facts about the differences in the core variables we have been exploring across entering, exiting and incumbent establishments. In all cases, exit of an establishment here means that the

[9] In the future, we will investigate whether we can find a nonlinear relationship.

establishment has ceased operations and entry of an establishment is a new establishment that had not previously operated at that location.

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Exiter (1992)	**0.171**	**-0.133**	**-0.918**	**-0.251**	**-0.158**
	(0.007)	(0.008)	(0.063)	(0.020)	(0.015)
Continuer (1992)	**0.118**	**0.060**	**-0.773**	**-0.091**	**0.026**
	(0.006)	(0.007)	(0.045)	(0.018)	(0.013)
Entrants (1997)	**-0.124**	**-0.020**	**0.220**	-0.004	**-0.058**
	(0.008)	(0.009)	(0.059)	(0.017)	(0.015)
Continuers (1997)	0.000	0.000	0.000	0.000	0.000
	(-)	(-)	(-)	(-)	(-)
R-squared	0.335	0.347	0.616	0.042	0.028
N	80,308	60,125	1,277	11,674	19,767

* Dependent variable is log of revenue/worker and standard errors are in parenthesis.

** Controls: Four digit SIC

Table 3.9: Productivity Difference (Continuing/Entry/Exit Establishments)

Table 3.9 shows the differences in means of productivity across exiting, entering and continuing establishments in 1992 and 1997 for the five industries. These differences in means are generated from a simple linear regression of measured labor productivity on entry, exit and continuer dummies with the omitted group the continuers in 1997. Calculating the differences in means in this fashion provides a transparent way of identifying whether differences in means are statistically significant.

In semiconductors and software, we observe that the large increases in overall productivity are associated with both dramatic productivity increases for continuing businesses and entrants with much higher productivity than the exiting establishments they are replacing. Even in these sectors with rapid within business growth (as evidenced by the growth rate in productivity for continuers), it is striking that the productivity gap between entering and exiting businesses for both of these sectors is greater than the productivity gap between incumbents in 1992 and 1997. For example, the productivity gap between entering and exiting businesses is about 140 log points in semiconductors while the productivity gap between continuers in 1992 and 1997 is 77 log points. This finding indicates that the contribution of net entry is disproportionate in semiconductors.

For retail food and trucking, entrants are less productive than incumbents but still substantially more productive than exiting businesses. In both of these sectors the net entry gap substantially exceeds the growth rate in productivity for continuers so net entry's contribution is again disproportionate in these sectors. Indeed, in retail food the continuing establishments exhibited negative measured productivity growth. While the latter might reflect measurement difficulties with the gross revenue per worker measure of productivity, it is still striking that the net entry effect is so large and positive.

For financial services, the results imply that continuing businesses exhibited substantial negative productivity growth and that entering businesses are less productive than the exiting businesses they are displacing. Further, the exiting businesses seem to be more productive than incumbents. While these results are consistent with the overall drop in revenue per worker documented in Table 3.2, as an account of industry dynamics, the findings are implausible. It is more likely that measured revenue per worker in the sector is not representative of establishment-level productivity. Even in the cross section (i.e., same year), the results are puzzling, which suggests that the difficulties are not simply a matter of having the wrong price deflator or some other factor that is mismeasured over time. Rather, comparisons of revenue per worker across businesses within the same year do not provide accurate representations of productivity differences across the financial services sector.

Table 3.10 depicts the differences in mean productivity across businesses taking into account entry and exit of establishments, entry and exit of firms and ownership changes. For most sectors, exiting establishments of exiting firms are typically the least productive and entering establishments of continuing firms are more productive than entering establishments of new firms. In terms of ownership change, continuing establishments that change ownership are more productive both before and after the ownership change than establishments that did not change ownership.

The productivity differences between those that changed ownership and those that did not are especially large in the trucking industry. These findings taken together with earlier results suggest ownership changes are important in productivity dynamics and also that the productivity dynamics for entering and exiting establishments are closely linked to firm structure. In general, we

Estab/Firm (Year)	Financial Services	Retail Food	Semiconductors	Software	Trucking
Exiter/Exiter	-0.002	**-0.142**	**-0.923**	**-0.289**	**-0.174**
(1992)	(0.009)	(0.008)	(0.067)	(0.021)	(0.015)
Exiter/Continuer	**0.410**	0.017	**-0.703**	**0.177**	**0.301**
(1992)	(0.010)	(0.019)	(0.169)	(0.050)	(0.050)
Continuer/Different	**0.063**	**0.220**	**-0.646**	0.086	**0.383**
(1992)	(0.014)	(0.018)	(0.116)	(0.048)	(0.047)
Continuer/Same	**0.135**	**0.058**	**-0.765**	**-0.095**	0.021
(1992)	(0.007)	(0.007)	(0.047)	(0.018)	(0.013)
Entrants/Entrants	**-0.102**	**-0.029**	**0.235**	-0.019	**-0.065**
(1997)	(0.010)	(0.009)	(0.060)	(0.018)	(0.016)
Entrants/Continuer	**-0.105**	**0.077**		**0.275**	**0.219**
(1997)	(0.011)	(0.017)		(0.044)	(0.052)
Continuer/Different	**0.062**	**0.149**	0.220	**0.129**	**0.301**
(1997)	(0.014)	(0.020)	(0.118)	(0.050)	(0.052)
Continuer/Same	0.000	0.000	0.000	0.000	0.000
(1997)	(-)	(-)	(-)	(-)	(-)
R-squared	0.344	0.349	0.617	0.054	0.038
N	80,308	60,125	1,277	11,674	19,767

* Dependent variable is log of revenue/worker and standard errors are in parenthesis
** Controls: Four digit SIC
*** Small number of entrants/continuer (1997) in semiconductors preclude reporting statistics

Table 3.10: Productivity Difference and Ownership Change

find that the positive net entry effect associated with establishments is driven disproportionately by the exit of very low productivity establishments that are single-unit firms and the high productivity of entering establishments of existing firms.

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Exiter (1992)	**0.043**	**0.083**	**0.049**	**0.108**	**0.148**
	(0.003)	(0.004)	(0.022)	(0.011)	(0.008)
Continuer (1992)	**0.024**	**0.053**	0.003	**0.049**	**0.058**
	(0.003)	(0.003)	(0.015)	(0.010)	(0.007)
Entrants (1997)	**0.062**	**0.061**	**0.048**	**0.078**	**0.098**
	(0.003)	(0.004)	(0.020)	(0.009)	(0.008)
Continuers (1997)	0.000	0.000	0.000	0.000	0.000
	(-)	(-)	(-)	(-)	(-)
R-squared	0.021	0.013	0.008	0.011	0.020
N	77,464	57,056	1,242	11,031	18,769

* Dependent variable is churning rate and standard errors are in parenthesis
** Controls: Four digit SIC

Table 3.11: Churning Differences (Continuing/Entry/Exit Establishments)

Table 3.11 reports differences in the mean churning rates across entering, exiting and continuing establishments. The overall decrease in the churning rate over time is driven by a fall in the churning rate for continuing establishments. Entering establishments have higher turnover than incumbents and exiting businesses have higher turnover than continuers. The differences in churning rates between entering and exiting establishments on the one hand and continuers on the other hand are especially large in the software and trucking industries.

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Exiter (1992)	**-0.137**	**-0.132**	-0.038	**-0.068**	**-0.185**
	(0.002)	(0.003)	(0.024)	(0.007)	(0.006)
Continuer (1992)	**-0.131**	**-0.092**	**-0.089**	**-0.072**	**-0.133**
	(0.002)	(0.002)	(0.015)	(0.006)	(0.005)
Entrants (1997)	0.001	**-0.039**	**0.057**	-0.006	**-0.071**
	(0.002)	(0.003)	(0.022)	(0.006)	(0.007)
Continuers (1997)	0.000	0.000	0.000	0.000	0.000
	(-)	(-)	(-)	(-)	(-)
R-squared	0.383	0.092	0.073	0.055	0.143
N	43,661	32,163	886	4,784	8,664

* Dependent variable is human capital and standard errors are in parenthesis
** Controls: Four digit SIC

Table 3.12: Human Capital Differences (Continuing/Entry/Exit Establishments)

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Exiter (1992)	**-0.061**	**-0.052**	-0.013	**-0.014**	**-0.083**
	(0.002)	(0.003)	(0.021)	(0.007)	(0.006)
Continuers (1992)	**-0.043**	**-0.027**	**-0.030**	-0.010	**-0.051**
	(0.002)	(0.002)	(0.013)	(0.006)	(0.005)
Entrants (1997)	**0.030**	0.002	**0.125**	**0.046**	**0.029**
	(0.002)	(0.003)	(0.020)	(0.006)	(0.007)
Continuers (1997)	0.000	0.000	0.000	0.000	0.000
	(-)	(-)	(-)	(-)	(-)
R-squared	0.142	0.031	0.097	0.034	0.048
N	43,741	32,316	886	4,787	8,658

* Dependent variable is person effect and standard errors are in parenthesis
** Controls: Four digit SIC

Table 3.13: Person Effect Differences (Continuing/Entry/Exit Establishments)

In Table 3.12, differences in means of workforce quality measured using the overall human capital by entering, exiting and continuing establishments are reported. Exiting businesses have lower human capital than survivors except for semiconductors and entrants have lower quality than incumbents except for semiconductors. The semiconductors sector shows a very different pattern on workforce quality in that the entering establishments have much higher workforce quality than incumbents. In all sectors, continuing establishments exhibited substantial increases in human capital. The analogous differences for the person effects are reported in Table 3.13. For the most part, the results mimic the results in Table 3.12. One notable exception is that entering establishments have higher person effects than incumbents suggesting that the human capital advantage of incumbents is primarily through experience.[10] In a related fashion, observe that the gap between the human capital of the entering and continuing establishments in semiconductors is even larger when human capital is measured in terms of the person effect, suggesting that entering establishments may have lower worker experience (or younger workers) on average.

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Exiter (1992)	**-0.057**	**-0.196**	-0.085	**-0.331**	**-0.291**
	(0.007)	(0.006)	(0.043)	(0.021)	(0.013)
Survivor (1992)	0.005	**0.063**	-0.001	-0.029	**0.043**
	(0.006)	(0.005)	(0.030)	(0.018)	(0.012)
Entrants (1997)	**-0.169**	**-0.157**	0.001	**-0.089**	**-0.224**
	(0.007)	(0.007)	(0.040)	(0.018)	(0.014)
Continuers (1997)	0.000	0.000	0.000	0.000	0.000
	(-)	(-)	(-)	(-)	(-)
R-squared	0.141	0.071	0.010	0.030	0.061
N	80,299	60,013	1,276	11,673	19,762

* Dependent variable is log of real payroll per worker and standard errors are in parenthesis
** Controls: Four digit SIC

Table 3.14: Payroll Per Worker Effect Differences (Continuing/Entry/Exit Establishment)

One data limitation present in Tables 3.12 and 3.13 relative to earlier tabulations is that the number of establishments for which we can measure the human capital variables is smaller than the overall sample. For example, in semiconductors there are 1242 establishments in 1992 and 1997 used in Table 3.11 to depict churning differences across establishments, but only 886 establishments in 1992 and 1997 in Tables 3.12 and 3.13 for which we can measure human capital. Both as a cross check and for independent interest, Table 3.14 shows the differences across continuing, entering and exiting establishments for payroll per worker, which we can measure for the full sample of establishments. One pattern that is similar is that exiting establishments have lower payroll per worker than incumbents. We also find that entrants have lower payroll per worker than incumbents. This latter pattern is consistent with the findings of overall human capital differences between incumbents and entrants.

[10] It is not a previously known result that older firms have older workers and need further investigation.

3.4 Market Selection: The Role of Productivity, Workforce Quality and Worker Churning

Economic models of industry evolution (e.g., Jovanovic (1982), Ericson and Pakes (1996), and Hopenhayn (1992), Melitz (2003)) suggest that market selection will be an ongoing process in industries given the myriad of changing economic conditions and the role of idiosyncratic factors determining business success. Idiosyncratic shocks to demand, costs and efficiency generate idiosyncratic profitability outcomes across establishments. Differences in managerial ability yield idiosyncratic profitability outcomes in addition to the impact of idiosyncratic shocks through a variety of channels - good managers make good choices in their responses to shocks, in their choices of goods and services to produce, in their choices of business location, in their choices of the mix of inputs (including the mix of workers), and in their managerial practices (including human resource practices).

The recent literature using longitudinal business databases has provided substantial empirical support for these models of industry evolution as a number of authors have found that businesses with low measured efficiency (via measures of productivity) and that are young and/or small are more likely to exit (see, e.g., Bartelsman and Doms (2001) and Caves (1998) for recent surveys of these findings). Our value-added to this literature is that we measure workforce quality and workforce churning as well at the business level and, as discussed in the introduction of this chapter, these are factors that are also potentially related to business survival. In this section, we present a model to help motivate our approach and frame the results as well as discuss our value-added to previous literature.

3.4.1 Illustrative Model: Workforce Quality and Survival

Consider the following two period model. Firms hire one worker and face the following revenue function:

$$R_{i,t} = A_{i,t} h_i \tag{3.1}$$

where $R_{i,t}$ is the revenue of the firm i in period t, $A_{i,t}$ its idiosyncratic revenue shock, and h_i the level of human capital of its sole worker. In this model, $A_{i,t}$ can take the form of a productivity shock and/or a price shock for a firm. At the beginning of period one, firms hire one worker who can be either a high human capital worker (h^H) or a low human capital worker (h^L). Firms pay high human capital workers and low human capital workers fixed wages of w^H and w^L, respectively. The initial revenue shock can take three values, high (A^H), normal (A^M), and low (A^L). Each revenue shock occurs with the same probability. Moreover, for simplicity, assume that A^M is the mean point of A^H and A^L. In period one, firms produce and sell their products as well as pay their workers. At the end of period one, firms have to make a stay/exit decision for period two before the revenue shocks in period two are realized. Revenue shocks in period two follow a Markov process with the following transition matrix:

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix} \tag{3.2}$$

where $P_{1,2} = 1/2$ is the probability of realizing A^M conditional on receiving the initial shock of A^H. To make the model more relevant to our exercise, we will assume the following equalities:

$$A^M h^H = A^H h^L \tag{3.3}$$

$$A^L h^H = A^M h^L \tag{3.4}$$

Firms with low human capital workers can have the same amount of revenue (or revenue per worker since firms have only one worker) as firms with high human capital workers only if the former receive more favorable shocks than the latter.

Firms are risk neutral and hence they make a stay/exit decision based on the expected profit of period two. If the expected profit is zero, then firms decide to stay with probability p and to exit with probability $1 - p$. Low human capital workers are paid $A^M h^L$, the unconditional mean of their value of marginal product, and high human capital workers are paid $A^M h^H - \alpha$ where $\alpha \in \left[0, \left(A^M - A^L\right) h^H\right)$.[11] One way of interpreting α is that high human capital workers have

[11] We can think α as a wage compression parameter, although it measures wage compression between firms instead

some firm specific skills and that firms and workers share the returns to the specific skills. The importance of having α is that the relative wages of high human capital workers may be lower than the relative values of their marginal products. In other words, the wage differential is smaller than the productivity differential between high human capital workers and low human capital workers. Given the above assumptions, expected profits conditional on different combinations of shock values and human capital are in equations 3.5 - 3.12.

$$
\begin{aligned}
E(\pi_{i,2}|A_{i,1}=A^H,h_i=h^L) &= \frac{1}{2}[A^H+A^M]h^L - A^M h^L = \frac{1}{2}(A^H-A^M)h^L > 0 && (3.5)\\
E(\pi_{i,2}|A_{i,1}=A^M,h_i=h^L) &= \frac{1}{4}[A^H+2A^M+A^L]h^L - A^M h^L = 0 && (3.6)\\
E(\pi_{i,2}|A_{i,1}=A^L,h_i=h^L) &= \frac{1}{2}[A^M+A^L]h^L - A^M h^L = -\frac{1}{2}(A^M-A^L)h^L < 0 && (3.7)\\
E(\pi_{i,2}|A_{i,1}=A^H,h_i=h^H) &= \frac{1}{2}[A^H+A^M]h^H - (A^M h^H - \alpha) \\
&= \frac{1}{2}(A^H-A^M)h^H + \alpha > 0 && (3.8)\\
E(\pi_{i,2}|A_{i,1}=A^M,h_i=h^H) &= \frac{1}{4}[A^H+2A^M+A^L]h^H - (A^M h^H - \alpha) = \alpha \\
&= 0 \text{ if } \alpha = 0 && (3.9)\\
&> 0 \text{ if } \alpha > 0 && (3.10)\\
E(\pi_{i,2}|A_{i,1}=A^L,h_i=h^H) &= \frac{1}{2}[A^M+A^L]h^H - (A^M h^H - \alpha) = \alpha - \frac{1}{2}(A^M-A^L)h^H \\
&< 0 \text{ if } \alpha < \frac{1}{2}(A^M-A^L)h^H && (3.11)\\
&> 0 \text{ if } \alpha > \frac{1}{2}(A^M-A^L)h^H && (3.12)
\end{aligned}
$$

Low human capital firms will stay if they get the high initial revenue shock A^H, stay with a probability p and exit with a probability $1-p$ if they get the medium revenue shock A^M, and exit if they get the low revenue shock A^L. High human capital firms will also stay if they get the most favorable initial revenue shock. If high human capital workers are paid the unconditional mean of the value of their marginal product (that is $\alpha = 0$), then they will stay with probability p and exit with probability $1-p$ if they get normal revenue shock, and exit with the low revenue shock.

However, if high human capital workers are paid a lower wage ($\alpha > 0$), then firms will make a different stay/exit decision. When $0 < \alpha < \frac{1}{2}(A^M - A^L)h^H$, high human capital firms

of within firms (as wage compression measures typically do).

will stay with the normal initial revenue shock, A^M, but exit with the low initial revenue shock, A^L. If high human capital workers are paid much lower ($\alpha > \frac{1}{2}(A^M - A^L)h^H$), then firms will stay notwithstanding which shock is realized. By assumption, we may observe the same level of revenue per worker in high and low human capital firms (see equations (3.3) and (3.4)). It is only when the wages of high and low human capital workers are the closest ($\alpha > \frac{1}{2}(A^M - A^L)h^H$) that firms with low human capital are more likely to exit given the same level of revenue per worker (equations (3.6) vs. (3.12)). In other cases, we should see more high human capital firms exit than low human capital firms.

The above model shows that there exists the possibility that, under some circumstances, variables other than productivity (in this case, human capital) can affect firm survival even after taking productivity into account. In principle, the role of human capital on survival conditional on productivity can go either way. The intuition behind this result is as follows. The fact that high and low human capital firms have the same level of labor productivity means that high human capital firms are hit by worse revenue shocks than low human capital firms. If workers are paid the expected value of their marginal product, then we should see more high human capital firms exit given the same labor productivity. However, if there is any mechanism that reduces the wage differential between high and low human capital workers (such as wage compression or rent sharing), then the results may change depending on the size of wage differential. This differential could be affected by forces such as the bargaining power of firms and workers. In any case, more productive firms are more likely to stay given the quality of the workforce.

3.4.2 Illustrative Model: Churning and Survival

Consider the following two period model. Firms hire one worker and face the following revenue function:

$$R_{i,t} = A_{i,t}H_{i,t} \tag{3.13}$$

where $R_{i,t}$ is the revenue of firm i in period t, $A_{i,t}$ its idiosyncratic revenue shock, and $H_{i,t}$ the level of human capital of its sole worker. In this model, $A_{i,t}$ can take the form of a productivity shock and/or a price shock for a firm. At the beginning of period one, all firms hire one worker with the same human capital level G.[12] The initial revenue shock can take three values, high (A^H), normal (A^M), and low (A^L). Each revenue shock occurs with the same probability. Moreover, for simplicity, assume that A^M is the mean point of A^H and A^L. In period one, firms produce and sell their products as well as pay their workers. At the end of period one, firms have to make a stay/exit decision for period two before the revenue shocks in period two are realized. Revenue shocks in period two follow a Markov process with the following transition matrix:

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix} \tag{3.14}$$

where $P_{1,2} = 1/2$ is the probability of realizing A^M conditional on receiving the initial shock of A^H.

Now suppose that there are two types of firm/worker matches. One is the internal labor market (ILM) type (or low chur type), where workers are trained in period one and retained by the end of period two. The cost of training is I and trained workers' human capital level increases by S in period two.[13] Training cost I is an increasing function of the number of firms taking ILM type. However, individual firms are too small and hence take I as given. The other is spot market (SM) type (or high churn type), where firms hire new workers each period and no workers are trained. One could say that the ILM type firms "make" workers while the SM type firms "buy" workers.

Firms are risk neutral and hence they make a stay/exit decision based on the expected profit of period two. If the expected profit is zero, then firms decide to stay with probability p and to exit with probability $1-p$. In SM type firms, workers are paid $A^M G$, the unconditional mean of their value of marginal product, in both periods. However, in ILM type firms, workers are offered upward sloping wage profiles. Suppose that firms and workers equally share the burden of the training cost. Workers are paid $A^M G - 0.5 \times I$ in period one and $A^M G + 0.5 \times A^M S$ in period

[12] This is general human capital.

[13] This is specific human capital only productive to the firm that trained this worker.

two. Given the above assumptions, expected profits conditional on different revenue shock values are in equations (3.15) - (3.20).

$$\begin{aligned}
E(\pi_{i,2}|A_{i,1} = A^H, SM) &= \frac{1}{2}[A^H + A^M]G - A^M G \\
&= \frac{1}{2}(A^H - A^M)G > 0 \qquad (3.15) \\
E(\pi_{i,2}|A_{i,1} = A^M, SM) &= \frac{1}{4}[A^H + 2A^M + A^L]G - A^M G \\
&= 0 \qquad (3.16) \\
E(\pi_{i,2}|A_{i,1} = A^L, SM) &= \frac{1}{2}[A^M + A^L]h^L - A^M G \\
&= -\frac{1}{2}(A^M - A^L)G < 0 \qquad (3.17) \\
E(\pi_{i,2}|A_{i,1} = A^H, ILM) &= \frac{1}{2}[A^H + A^M](G+S) - A^M G - \frac{1}{2}A^M S \\
&= \frac{1}{2}(A^H - A^M)G + \frac{1}{2}A^H S > 0 \qquad (3.18) \\
E(\pi_{i,2}|A_{i,1} = A^M, ILM) &= \frac{1}{4}[A^H + 2A^M + A^L](G+S) - A^M G - \frac{1}{2}A^M S \\
&= \frac{1}{2}A^M S > 0 \qquad (3.19) \\
E(\pi_{i,2}|A_{i,1} = A^L, ILM) &= \frac{1}{2}[A^M + A^L](G+S) - A^M G - \frac{1}{2}A^M S \\
&= -\frac{1}{2}(A^M - A^L)G + \frac{1}{2}A^L S >< 0 \qquad (3.20)
\end{aligned}$$

SM type firms will stay with certainty if they get the high initial revenue shock A^H, stay with a probability p and exit with a probability $1-p$ if they get the medium revenue shock A^M, and exit if they get the low revenue shock A^L. ILM type firms will stay with certainty if they get the most favorable initial revenue shock or normal revenue shock. With the low revenue shock, ILM type firms may stay if return to training is large (high S) or differential in revenue shocks is small.

The above model shows that low churning firms are more likely to stay than high churning firms conditional on productivity and human capital. The intuition behind this result is as follows. ILM type (or low churning) firms pay less than the (expected) marginal product of their workers in period one to cover part of the training cost. However, in period two, the same workers can earn more in this firm than in any other firms because these workers have accumulated specific human capital with training. In other words, workers who receive training at ILM type firms in period one have an incentive to remain with their employers in period two. At the same time, firms have higher expected profits because they share extra profits that arise from specific human capital accumulated with worker training.

Does the fact that ILM type firms are more likely to survive mean that all potential entrants would choose ILM type human resource practice? To answer this question, we need to compare values of adopting ILM and SM types for potential entrants. Given the initial distribution and transition process of revenue shock, and choice of stay/exit, one can calculate expected value of potential ILM type firms and SM type firms as in equations (3.21) and (3.22).

$$\begin{aligned} V(SM) &= \frac{1}{3}[A^H G - A^M G + \frac{1}{2}(A^H - A^M)G] \\ &+ \frac{1}{3}[A^M G - A^M G + 0] \\ &+ \frac{1}{3}[A^L G - A^M G] \\ &= \frac{1}{6}(A^H - A^M)G \qquad (3.21) \\ V(SM) &= \frac{1}{3}[A^H G - A^M G - \frac{1}{2}I + \frac{1}{2}(A^H - A^M)G + \frac{1}{2}A^H S] \\ &+ \frac{1}{3}[A^M G - A^M G - \frac{1}{2}I + \frac{1}{2}A^M S] \\ &+ \frac{1}{3}[A^L G - A^M G - \frac{1}{2}I] \\ &= \frac{1}{6}(A^H - A^M)G - \frac{1}{2}I + \frac{1}{6}(A^H + A^M)S \qquad (3.22) \end{aligned}$$

Potential entrants are indifferent to two types of HRM practice when values expressed in equations (3.21) and (3.22) are equal to each other. It happens when

$$I = \frac{1}{3}(A^H + A^M)S. \qquad (3.23)$$

Because training cost is an increasing function of the number of firms adopting ILM type HRM practice, free entry will make training cost adjust such that potential firms are indifferent to either type. Therefore, higher probability of survival does not mean *ex ante* higher value to potential entrants.

Note that under this condition (3.23), the expected wage of workers in ILM type firms is always higher than that of workers in SM type firms. However, if workers are risk averse and not allowed to borrow, then it is possible that workers are also indifferent to working in either type of firms because SM firms pay flat wage while ILM firms offer upward sloping wage profiles. Based on these illustrative theoretical models, the next section describes the empirical model for firm survival.

3.4.3 Empirical Model

Firms choose to exit if the present discounted value (PDV) of their future profits is negative. The PDV of future profits is thus the sufficient statistic to predict individual firms' stay/exit decisions. However we do not observe this statistic directly. Hence variables that affect future profits have been used to predict firm survival/failure. As mentioned above, productivity, size, and/or age are found to be strong indicators of firm survival. In the spirit of the model described in Section 3.4.1 and 3.4.2, we try to see whether there are other variables than these that independently contribute to firm survival. The empirical model used in this exercise is

$$EXIT_{i,t} = f\left(LP_{i,t}, L_{i,t}, Single_{i,t}, H_{i,t}, Chr_{i,t}, ...\right) \tag{3.24}$$

where $EXIT_{i,t}$ is a binary variable that is equal to 1 if firm i exits at the end of period t and 0 if it stays, $LP_{i,t}$ is labor productivity measured by revenue per worker (to capture efficiency of the production), $L_{i,t}$ is the size of firm i measured by log employment (to incorporate size and/or age effects), $Single_{i,t}$ is a single-unit dummy variable for firm structure, $H_{i,t}$ is the summary measure of firm level human capital, and $Chr_{i,t}$ is the churning rate. The human capital and churning rate variables capture firms' human resource management practices and are the key variables of interest in this regression. Finally, we also include four-digit industry dummies to control for industry heterogeneity.

3.4.4 Estimation Results

In Tables 3.15 - 3.17, we present results from estimating the determinants of exit of establishments (using a probit estimation) based upon the exit and survival of establishments from 1992 to 1997

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Single unit dummy	**-0.131**	**0.039**	-0.093	**-0.098**	**-0.112**
	(0.011)	(0.008)	(0.059)	(0.025)	(0.020)
Size	**-0.078**	**-0.072**	-0.019	-0.017	**-0.049**
	(0.003)	(0.005)	(0.016)	(0.010)	(0.008)
Revenue/worker	**0.066**	**-0.110**	-0.015	**-0.038**	**-0.067**
	(0.005)	(0.006)	(0.036)	(0.019)	(0.011)
Churning	**0.100**	**0.069**	**0.494**	**0.219**	**0.206**
	(0.032)	(0.019)	(0.175)	(0.064)	(0.033)
Human Capital	**-0.135**	**-0.151**	**0.261**	0.064	**-0.170**
	(0.027)	(0.023)	(0.105)	(0.066)	(0.036)
N	23160	15682	428	2044	4318

* Estimation is based on Probit with exit/stay as a dependent variable.

** Standard errors are in parenthesis

*** Controls: Four digit SIC

Table 3.15: Probability of Exit of Establishment (Using Overall Human Capital)

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Single unit dummy	**-0.140**	**0.039**	-0.086	**-0.098**	**-0.103**
	(0.011)	(0.008)	0.060	(0.025)	(0.020)
Size	**-0.077**	**-0.073**	-0.021	-0.017	**-0.050**
	(0.003)	(0.005)	(0.016)	(0.010)	(0.008)
Revenue/worker	**0.065**	**-0.113**	0.013	-0.033	**-0.076**
	(0.005)	(0.006)	(0.036)	(0.019)	(0.011)
Churning	**0.144**	**0.098**	**0.404**	**0.214**	**0.234**
	(0.031)	(0.019)	(0.164)	(0.063)	(0.033)
Person Effect	**-0.219**	**-0.091**	0.051	-0.013	**-0.173**
	(0.025)	(0.022)	(0.115)	(0.063)	(0.038)
N	23160	15682	428	2044	4318

* Estimation is based on Probit with exit/stay as a dependent variable.

** Standard errors are in parenthesis

*** Controls: Four digit SIC

Table 3.16: Probability of Exit of Establishment (Using Person Effect)

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Single unit dummy	**-0.136**	**0.039**	-0.1	**-0.12**	**-0.153**
	(0.009)	(0.007)	(0.062)	(0.022)	(0.019)
Size	**-0.076**	**-0.089**	**-0.056**	**-0.033**	**-0.082**
	(0.002)	(0.003)	(0.013)	(0.006)	(0.005)
Revenue/worker	**0.017**	**-0.078**	0.053	0.003	**-0.032**
	(0.004)	(0.005)	(0.043)	(0.014)	(0.009)
Churning	**0.055**	**0.039**	0.085	**0.048**	**0.082**
	(0.008)	(0.008)	(0.069)	(0.017)	(0.012)
Payroll/worker	**-0.038**	**-0.095**	-0.04	**-0.127**	**-0.116**
	(0.005)	(0.006)	(0.059)	(0.013)	(0.011)
N	39474	29170	605	4982	9442

* Estimation is based on Probit with exit/stay as a dependent variable.

** Standard errors are in parenthesis

*** Controls: Four digit SIC

Table 3.17: Probability of Exit of Establishment (Using Payroll Per Worker)

as a function of initial conditions in 1992. All subsequent probits in this chapter are of this general functional form. Table 3.15 uses the overall human capital measure, Table 3.16 uses the person effect measure and Table 3.17 uses the payroll per worker measure to capture workforce quality. The sample size in Table 3.17 is, as noted above, larger since we can measure payroll per worker for more establishments.

We begin our discussion with Table 3.15. Consistent with much of the literature, we find that larger businesses are less likely to fail and, except for financial services, high productivity businesses are less likely to fail. We also control for firm structure with a single-unit dummy. In most sectors, single unit establishments are less likely to fail after controlling for all of the other factors. While this finding might seem surprising, recall that it is after controlling for size, productivity, churning and workforce quality. Moreover, it is consistent with the Holmes and Schmitz (1995) hypothesis that single-unit firms may be, holding other factors constant, less willing to close since closing down the establishment implies closing down the firm while this is not the case for establishments belonging to a multi-unit firm.

Even after controlling for these important factors, we find that workforce quality and worker churning contribute independent explanatory power to accounting for survival. In particular, as predicted in the previous illustrative model, higher churning businesses are more likely to fail. This result is the most robust finding in Table 3.15 in that this result holds for all industries and the impact of churning is large in magnitude. For example, ten percentage points increase in churning increases the likelihood of failure by almost five percentage points in the semiconductor industry. These results are striking because they reflect the independent contribution of churning even after taking into account the impact of productivity (albeit crudely measured) on survival. These results suggest that high churning businesses are low profitability businesses. Appropriate caution on the causal interpretation here (and for all the results of these multivariate probits) is needed of course. It may be that the businesses are of low profitability, which leads to churning prior to exit.

Establishments with low churn tend to have relatively more long-tenured workers who have accumulated large amounts of firm-specific skills. If firms and workers share the returns to firm-specific human capital, then workers may be willing to accept a smaller share of the rents when their employers are in diffult situations. Because the component of workers' wages coming from their firm specific skills is not compensated by other firms when workers quit their current jobs, workers may prefer receiving a lower wage from their current employer over getting an even lower market wage from other firms. Low churning firms are also likely to have more attached workers, to be family business, to have a very strong team-style work environment, or to be unionized. In these cases, workers may try harder to overcome their firms' current difficulties, or perhaps firms face a tougher decision when considering shutting down their operations.

Human capital yields mixed results on failure with high human capital businesses in financial services, retail food and trucking less likely to fail, but high human capital businesses in semiconductors are more likely to fail. In software, the coefficient is also positive but insignificant. However, the measure of human capital here is overall human capital including the contribution of experience, and it may be that worker experience is less important in semiconductors and software. Alternatively, it is worth emphasizing that these results should be interpreted in the context of already controlling for productivity (and other factors like churning). Of course, one explanation for the independent contribution of human capital is that our measures of productivity are poor, so that human capital is capturing part of the influence of true (unmeasured) productivity.

However, we should not be surprised with the estimation results when we consider the predictions of our model. Recall that, in principle, the impact of human capital on survival conditional on productivity is ambiguous. If wage compression is significant, high human capital firms are more likely to survive conditional on productivity. Otherwise, we expect that they are more likely to fail, because observing the same productivity from both high and low human capital firms implies that high human capital firms suffered more serious negative revenue shocks. We can conjecture that semiconductors and software are sectors where we see low wage compression relative to other sectors. The key idea from the theoretical model is that given the same productivity level, high and low human capital firms can have different expectations about their future profit streams. Firms make their stay/exit decisions based on their expectations, not on their current profits or productivity levels.

High human capital can contribute not only to improving the efficiency of production processes, but also to solving many other problems. Some portion of human capital is problem solving skills, including overcoming difficult financial situations. This set of skills may not be fully reflected by high productivity, but can play a very critical role when firms are in trouble.

Turning to Table 3.16, which uses person effects as an alternative measure of human capital, the results are largely the same as in Table 3.15, but now the effect of human capital is negative for four of the five sectors and positive but insignificant for semiconductors. Thus, part of the story for software and semiconductors is that the experience component of human capital works differently in those industries. Finally, Table 3.17 shows the results using a cruder but more widely available measure of human capital - namely payroll per worker. Using this more widely available measure we find that in all sectors, high payroll per worker firms are less likely to fail and this effect is statistically significant in four of the five sectors.

Table 3.18 reports results of probit analyses where we allow a possibility of nonlinear relationship between churning and survival. According to matching literature, not only high churning but also too low churning is bad for firm performance. We classify firms into low, medium, and

	Financial Services	Retail Food	Semiconductors	Software	Trucking
Single unit dummy	**-0.143**	**0.039**	-0.094	**-0.097**	**-0.118**
	(0.011)	(0.008)	(0.059)	(0.025)	(0.020)
Size	**-0.074**	**-0.072**	-0.018	-0.015	**-0.047**
	(0.003)	(0.005)	(0.016)	(0.010)	(0.008)
Revenue/worker	**0.060**	**-0.111**	-0.004	-0.037	**-0.066**
	(0.005)	(0.006)	(0.036)	(0.019)	(0.011)
Churning (Low)	**0.117**	-0.013	-0.052	-0.055	**-0.053**
	(0.014)	(0.011)	(0.044)	(0.034)	(0.023)
Churning (Medium)	**0.023**	**-0.027**	**-0.137**	**-0.068**	**-0.115**
	(0.008)	(0.009)	(0.043)	(0.023)	(0.016)
Human Capital	**-0.172**	**-0.156**	**0.244**	0.059	**-0.164**
	(0.027)	(0.023)	(0.103)	(0.066)	(0.037)
N	23,160	15,682	428	2,044	4,318

* Estimation is based on Probit with stay/exit as a dependent variable.

** Standard errors are in parenthesis

*** Controls: Four digit SIC

Table 3.18: Probability of Exit of Establishment (Investigation of Nonlinear Effect of Churning)

high (omitted group) firms where we use the 33rd percentile and the 67th percentile as cutoff values. Except for financial services, we can observe that firms with medium level of worker churning is more likely to survive than both high and low churning firms. This finding shows that there is a nonlinear relationship between firm survival and worker churning.

3.5 A Deeper Look By Sector

The findings in the prior two sections show that firm performance as measured by entry, exit, and growth of continuing establishments is closely connected to workforce quality and workforce churning. The general finding is that businesses that are high productivity businesses are also high worker quality and low churning businesses. Moreover, all three of these factors independently contribute to survival - businesses with high productivity, high worker quality (especially via the person effect) and low worker churning are more likely to survive. While these patterns are reasonably common across the sectors, we know that each of these sectors has its own dynamic idiosyncratic factors. In this section, we explore some of these factors on a sector by sector basis.

3.5.1 Securities Brokers

In financial services, the structure of the industry has changed dramatically following deregulation and, as we have discussed at some length, measurement of firm performance in financial services is especially problematic. However, the measurement problems in financial services on these grounds differ across specific industries. In the banking industry, revenue numbers are especially problematic (and indeed are not used for official BLS statistics - BLS uses an indirect way to measure value-added in the banking industry). However, revenue numbers are potentially more reliable and sensible indicators of activity in the securities brokers industry. In the latter, brokers are largely providing a transaction service and revenues will reflect the amount of transactions services provided.

Establishment	Revenue/ Worker	Estab/Firm (Year)	Revenue/ Worker		Probability of Exit
Exiter (1992)	**-0.405**	**Exiter/Exiter**	**-0.442**	**Single unit dummy**	0.017
	(0.040)	**(1992)**	(0.052)		(0.039)
		Exiter/Survivor	**-0.294**	**Size**	**-0.092**
		(1992)	(0.052)		(0.011)
Survivor (1992)	**-0.266**	**Survivor/Different**	-0.008	**Revenue/worker**	-0.012
	(0.033)	**(1992)**	(0.071)		(0.023)
		Survivor/Same	**-0.260**	**Churning**	**0.170**
		(1992)	(0.035)		(0.042)
Entrants (1997)	**-0.254**	**Entrants/Entrants**	**-0.139**	**Person Effect**	**0.442**
	(0.034)	**(1997)**	(0.040)		(0.099)
		Entrants/Continuer	**-0.339**		
		(1997)	(0.046)		
Continuers (1997)	0.000	**Continuer/Different**	**0.303**		
	(-)	**(1997)**	(0.071)		
		Continuer/Same	0.000		
		(1997)	(-)		
R-squared	0.030	**R-squared**	0.042		
N	4,097	**N**	4,097	**N**	1,734

* Standard errors are in parenthesis

** Controls: Four digit SIC

Table 3.19: Productivity Difference and Probability of Exit of Establishment (Securities Brokers)

To get some sense of the sensitivity of results to more narrowly defined industries in financial services, in Table 3.19, we repeat the exercises from Tables 3.9, 3.10 and 3.16 respectively for the securities brokers industry (SIC 6211). The first column of Table 3.19 repeats the Table 3.9 exercise examining productivity differences across entering, exiting and continuing establishments.

The second column of Table 3.19 repeats the Table 3.10 exercise showing related productivity differences taking into account firm entry and exit and ownership changes as well. Recall that it is important here to distinguish conceptually between how entry and exit of establishments are defined vs. the entry and exit of firms. The third column of Table 3.19 repeats the Table 3.16 exercise, where we explore determinants of the probability of exit of establishments.

Overall, we find results that are very different than for the overall financial services industry, being more consistent with our prior hypotheses (and with results in other industries). In particular, we find that for securities brokers there is positive overall productivity growth between 1992 and 1997. This finding is driven both by substantial productivity growth of incumbents and by entering establishments being substantially more productive than exiting establishments. Ownership change for securities brokers is associated with higher productivity for continuing establishments both before and after the ownership change. In addition, the exiting establishments of exiting firms have especially low productivity, suggesting that the market selection of these establishment/firms is particularly productivity enhancing for the securities broker industry. Larger businesses are less likely to fail in this industry as predicted, and businesses with higher churning are less likely to fail. The remaining effects are mostly statistically insignificant, although there is a seemingly anomalous result that high human capital security broker businesses are more likely to fail. This latter result needs more exploration but might reflect the prevalence in this industry of very small startups with security broker "stars", who subsequently quit and destroy the establishment/firm.

3.5.2 Integrated vs. Fabless Semiconductor Establishments

With the growth of foundries in Asia and an ensuing rise of fabless semiconductor companies, the entrants into the domestic semiconductor industry in the 1990s are primarily fabless establishments (semiconductor design establishments without production capability). Fabless establishments employ primarily engineers, as opposed to integrated, or fabbed, establishments that employ engineers, technicians and operators. As a result, employment at entering establishments is skewed towards high-human capital workers. While the fabless startups are high human capital employers, they are also smaller and riskier than their integrated counterparts so the characteristics and dynamics of the young fabless establishments will look quite different from large, established traditional fabbed establishments.

The changes in the composition of the semiconductor industry have resulted in the workforce being more educated as the industry employs a higher proportion of workers with engineering degrees. As a result of this change in industry composition, we would expect revenue per worker, payroll per worker, and human capital all to increase between 1992 and 1997. As a first pass, we can examine the results in earlier sections in light of these composition changes. Observe that (real) revenue per worker in the semiconductor industry increased fourfold from 1992 to 1997 (Table 3.1); while payroll per worker increased 5% during this period. Low turnover rates in the industry (churning rates of 13.2%) were unchanged during this period, while mean establishment size increased slightly. The overall level of human capital increased significantly over the period. These findings are consistent with a shift in industry composition away from integrated, fabbed firms towards engineer-dominated fabless firms. Also the revenue based productivity measure may be less reliable for fabless producers since the timing between receipts and inputs may be quite different in this part of the industry.

While it is difficult to identify the fabless and fabbed establishments precisely in our data, we know key characteristics of fabless and fabbed establishments. For one, most entrants in the domestic industry after 1987 are fabless establishments. Second, fabless establishments are much smaller than fabbed establishments. Case study analysis by the Sloan semiconductor industry center suggests that all fabbed establishments have at least 300 employees. As such, we have used this information to classify establishments as "fabbed" and "fabless" based upon vintage and size. That is, any establishment that entered after 1987 and is less than 300 employees is classified as a "fabless" establishment while all others are classified as "fabbed" establishments.

Results using this classification are reported in Tables 3.20 - 3.23. We find that our so-called fabless establishments are indeed much smaller, more human capital intensive and have higher revenue per worker (in spite of the potential timing problems between receipts and inputs).

Year	Group	Revenue /worker ($)	Churning Rate (%)	Human Capital (%)	Person Effect (%)	Employment (Number)	Payroll /worker ($)
1992	Fabless	156,657	13.7	59.7	54.2	19.7	26,704
	Fabbed	136,145	13.0	55.8	46.9	103.5	26,931
1997	Fabless	681,979	15.5	67.3	59.6	28.7	28,330
	Fabbed	446,645	11.3	64.7	50.4	132.5	28,065

Table 3.20: Means of Core Variables in Semiconductor Industry by Type

Year	Group	Establishments	Employment	Sales
1992	Fabless	25.2	6.0	4.6
	Fabbed	74.8	94.0	95.4
1997	Fabless	46.2	15.7	12.6
	Fabbed	53.8	84.3	87.4

Table 3.21: Shares of Activity in Semiconductor Industry by Type

	Year	Group	Revenue/Worker	Churning	Person Effect
Exiter	**1992**	**Fabless**	**-0.925**	0.036	0.014
			(0.099)	(0.035)	(0.036)
		Fabbed	**-0.854**	**0.066**	-0.020
			(0.077)	(0.027)	(0.025)
Survivor	**1992**	**Fabless**	**-0.723**	0.022	0.037
			(0.077)	(0.026)	(0.024)
		Fabbed	**-0.739**	0.005	**-0.041**
			(0.051)	(0.017)	(0.015)
Entrants	**1997**	**Fabless**	**0.246**	**0.056**	**0.132**
			(0.062)	(0.021)	(0.021)
Continuers	**1997**	**Fabless**	**0.157**	0.025	0.014
			(0.075)	(0.026)	(0.022)
		Fabbed	0.000	0.000	0.000
			(-)	(-)	(-)
R-squared			0.618	0.010	0.111
N			1,276	1,241	885

* Estimation is based on Probit with stay/exit as a dependent variable.

** Standard errors are in parenthesis

*** Controls: Four digit SIC

**** Small number of fabbed entrants preclude reporting coefficients

Table 3.22: Differences of Core Variables in Semiconductor Industry by Type

	Group	
Single unit dummy		**-0.102**
		(0.061)
Size	**Fabless**	0.015
interacted with:		(0.035)
	Fabbed	-0.024
		(0.018)
Revenue/worker	**Fabless**	0.028
interacted with:		(0.045)
	Fabbed	0.007
		(0.038)
Churning	**Fabless**	0.344
interacted with:		(0.260)
	Fabbed	**0.414**
		(0.214)
Person Effect	**Fabless**	-0.158
interacted with:		(0.224)
	Fabbed	0.090
		(0.135)
N		428

* Estimation is based on Probit with exit/stay as a dependent variable.
** Standard errors are in parenthesis (bold at 10 percent level)
*** Controls: Four digit SIC

Table 3.23: Probability of Exit of Establishment in Semiconductor Industry by Type

Interestingly, despite these skill and productivity advantages, average payroll per worker is no higher at fabless establishments. Fabless establishments increased dramatically relative to fabbed establishments in terms of the number of establishments. However, they still account for relatively small shares of employment and sales by 1997 accounting for about 16 percent of employment and 13 percent of sales. Fabless establishments that enter between 1992 and 1997 are especially high productivity and high human capital establishments but the continuing fabless establishments did not exhibit much skill upgrading (they were already high skill in the first place). Fabless establishments that exit between 1992 and 1997 are low productivity but not especially low skill (again reflecting the high human capital of fabless establishments).

In examining the multivariate probit analysis of exit, not much is significant although fabbed establishments that have low churning are less likely to exit with results statistically significant at the 10 percent level. All of the rest of the effects are not statistically significant. The latter may reflect the relatively small sample size for the semiconductor industry so that multivariate analysis that also classifies plants into separate categories is pushing the data very hard. The weak multivariate results may also reflect the changing industry structure in semiconductors that implies amongst other things that factors such as size and human capital play different roles for fabbed and fabless establishments. Put more broadly, the changing industry structure may imply that it has become more difficult to characterize the determinants of survival in this industry.

3.5.3 National vs. Regional vs. Local vs. Single-Unit Establishments

In many industries, dynamics depend on whether the establishment is part of a large, national firm (with many establishments across many states) or a regional, local or single-unit firm (the firm has only one establishment). For at least two of the sectors that we are examining, these effects are likely to be especially important: retail food and trucking. In retail food, there has been a shift of activity towards large, national chains. In trucking, there is a clear bifurcation between national trucking companies that provide both national and local transportation services vs. smaller locally oriented businesses.

Sector	Year	Group	Revenue /worker ($)	Churning Rate (%)	Human Capital (%)	Person Effect (%)	Employment (Number)	Payroll /worker ($)
Retail Food	1992	Single Unit	132,445	27.7	29.6	44.6	7.9	8,393
		MU local	142,056	28.5	30.8	47.6	18.5	10,555
		MU regional	157,224	30.3	38.4	50.4	48.3	12,704
		MU national	154,871	33.9	31.8	49.5	40.2	11,404
	1997	Single Unit	130,326	24.1	39.2	47.1	8.0	8,241
		MU local	153,723	24.6	41.1	50.9	21.2	10,444
		MU regional	150,421	23.9	45.8	55.5	40.0	10,989
		MU national	166,788	25.1	41.2	54.0	38.8	10,728
Trucking	1992	Single Unit	93,794	28.2	51.5	38.7	9.7	16,491
		MU local	103,611	19.8	56.5	39.0	29.7	21,533
		MU regional	113,856	20.2	61.3	41.8	38.0	24,049
		MU national	149,193	15.3	71.6	42.9	57.3	28,189
	1997	Single Unit	95,595	22.0	65.5	44.7	10.1	16,410
		MU local	112,057	16.0	70.8	48.0	34.6	23,345
		MU regional	130,787	18.5	73.7	51.1	37.6	23,693
		MU national	138,707	14.2	77.1	53.3	59.3	25,818

Table 3.24: Means of Core Variables by Type

To investigate these differences for these two sectors, we classified establishments into one of four groups: single units, local establishments (establishment part of a multi-unit firm that operates in only one state), regional establishments (establishment part of a multi-unit firm that operates in 2-5 states) and national establishments (establishment part of a multi-unit firm that

Sector	Year	Group	Establishments	Employment	Sales
Retail Food	1992	Single Unit	68.5	33.5	27.0
		MU local	11.9	13.7	13.3
		MU regional	7.8	23.3	25.0
		MU national	11.9	29.6	34.7
	1997	Single Unit	66.1	32.0	26.0
		MU local	11.0	14.1	13.1
		MU regional	5.5	13.3	12.9
		MU national	17.4	40.7	48.0
Trucking	1992	Single Unit	88.1	61.3	54.5
		MU local	3.5	7.4	7.9
		MU regional	2.7	7.3	7.4
		MU national	5.8	24.0	30.2
	1997	Single Unit	89.1	63.5	56.8
		MU local	2.9	7.0	7.9
		MU regional	2.7	7.1	8.5
		MU national	5.3	22.4	26.8

Table 3.25: Shares of Activity by Type

operates in 6 or more states).[14] Tables 3.24 - 3.27 report empirical results for these two sectors using this more detailed classification. For both sectors, most establishments are single units but large, national establishments account for a disproportionate share of activity. For retail food, the share of activity accounted for by national establishments has grown dramatically. In 1992, national establishments accounted for 35 percent of sales but by 1997 they accounted for 48 percent of sales.

In retail food, national and regional establishments are more productive, are larger, pay higher wages and are more human capital intensive. The productivity ranking between these two types reversed between 1992 and 1997 with regional establishments relatively more productive in 1992 and national establishments more productive in 1997. Observe that although national establishments gained a productivity advantage over regional establishments by 1997, national establishments still pay less than regional establishments in 1997. Both regional and national establishments exhibit greater churning than local and single unit establishments.

Two thirds of retail food stores are single units accounting for one third of employment and a quarter of sales. National firms increased their share of stores by over 50 percent and increased

[14]This classification has been used by Foster et al. (2004) to study the selection and learning dynamics in the retail trade sector in the 1990s. They find that this distinction is very important across the retail trade sector and especially in the department and general merchandise store industries.

	Year	Group	Revenue/Worker		Churning		Person Effect	
			RF	TR	RF	TR	RF	TR
Exiter	**1992**	**Single Unit**	**-0.440**	**-0.591**	**0.065**	**0.214**	**-0.104**	**-0.175**
			(0.013)	(0.041)	(0.007)	(0.022)	(0.004)	(0.012)
		MU local	**-0.384**	**-0.581**	**0.059**	**0.116**	**-0.094**	**-0.155**
			(0.023)	(0.080)	(0.012)	(0.044)	(0.007)	(0.024)
		MU regional	**-0.220**	**-0.350**	**0.033**	0.063	**-0.057**	**-0.141**
			(0.031)	(0.082)	(0.015)	(0.044)	(0.009)	(0.023)
		MU national	**-0.291**	0.032	**0.164**	0.059	**-0.035**	**-0.147**
			(0.027)	(0.064)	(0.014)	(0.035)	(0.008)	(0.019)
Survivor	**1992**	**Single Unit**	**-0.302**	**-0.406**	**0.022**	**0.120**	**-0.087**	**-0.147**
			(0.013)	(0.041)	(0.006)	(0.022)	(0.004)	(0.011)
		MU local	**-0.254**	**-0.280**	**0.042**	0.056	**-0.051**	**-0.149**
			(0.017)	(0.061)	(0.009)	(0.033)	(0.005)	(0.017)
		MU regional	**-0.040**	**-0.204**	**0.084**	**0.086**	**-0.034**	**-0.099**
			(0.019)	(0.068)	(0.010)	(0.037)	(0.005)	(0.019)
		MU national	0.022	0.019	**0.083**	0.015	**-0.043**	**-0.076**
			(0.017)	(0.054)	(0.008)	(0.029)	(0.005)	(0.014)
Entrants	**1997**	**Single Unit**	**-0.323**	**-0.487**	**0.046**	**0.159**	**-0.062**	**-0.071**
			(0.013)	(0.042)	(0.007)	(0.022)	(0.005)	(0.013)
		MU local	**-0.286**	-0.174	**0.047**	0.037	**-0.031**	**-0.073**
			(0.027)	(0.114)	(0.014)	(0.062)	(0.008)	(0.032)
		MU regional	**-0.220**	-0.040	**0.059**	**0.104**	0.023	-0.033
			(0.037)	(0.086)	(0.019)	(0.046)	(0.010)	(0.026)
		MU national	**-0.190**	**-0.176**	**0.052**	0.053	0.005	-0.001
			(0.023)	(0.067)	(0.012)	(0.036)	(0.006)	(0.019)
Continuers	**1997**	**Single Unit**	**-0.387**	**-0.433**	**-0.021**	**0.056**	**-0.071**	**-0.103**
			(0.013)	(0.041)	(0.006)	(0.022)	(0.004)	(0.011)
		MU local	**-0.262**	**-0.262**	-0.003	0.035	**-0.032**	**-0.059**
			(0.017)	(0.061)	(0.009)	(0.033)	(0.005)	(0.016)
		MU regional	**-0.165**	**-0.213**	-0.011	0.040	0.009	-0.015
			(0.022)	(0.068)	(0.011)	(0.036)	(0.006)	(0.018)
		MU national	0.000	0.000	0.000	0.000	0.000	0.000
			(-)	(-)	(-)	(-)	(-)	(-)
R-squared			0.363	0.050	0.016	0.025	0.056	0.067
N			60,013	19,762	56,960	18,765	32,315	8,657

* Estimation is based on Probit with stay/exit as a dependent variable.

** Standard errors are in parenthesis

*** Controls: Four digit SIC

Table 3.26: Differences of Core Variables by Type

	Group	Retail_Food	Trucking
Single unit dummy		**-0.375**	**-0.274**
		(0.030)	(0.086)
Size	**Single Unit**	**-0.056**	**-0.049**
interacted with:		(0.008)	(0.010)
	MU local	**-0.111**	**-0.122**
		(0.009)	(0.027)
	MU regional	**-0.043**	-0.013
		(0.012)	(0.027)
	MU national	**-0.090**	**-0.043**
		(0.009)	(0.017)
Revenue/worker	**Single Unit**	**-0.081**	**-0.074**
interacted with:		(0.007)	(0.011)
	MU local	**-0.122**	**-0.093**
		(0.009)	(0.026)
	MU regional	**-0.150**	**-0.110**
		(0.011)	(0.026)
	MU national	**-0.183**	**-0.079**
		(0.009)	(0.020)
Churning	**Single Unit**	**0.197**	**0.252**
interacted with:		(0.023)	(0.035)
	MU local	0.000	**0.472**
		(0.047)	(0.197)
	MU regional	-0.094	0.009
		(0.079)	(0.172)
	MU national	-0.085	0.082
		(0.053)	(0.145)
Person Effect	**Single Unit**	**-0.084**	**-0.144**
interacted with:		(0.025)	(0.042)
	MU local	**-0.203**	-0.087
		(0.054)	(0.192)
	MU regional	**-0.219**	-0.210
		(0.099)	(0.157)
	MU national	**0.223**	**-0.399**
		(0.074)	(0.111)
N		15,700	4,319

* Estimation is based on Probit with exit/stay as a dependent variable.
** Standard errors are in parenthesis
*** Controls: Four digit SIC

Table 3.27: Probability of Exit of Establishment by Type

employment and payroll shares by around 40%, while regional firms' shares of employment and sales fell by about 50%. It may be that part of this reflects some regional firms becoming national over this period of time with an associated change in establishment status along these lines.

In terms of dynamics, skill upgrading occurred in all types of establishments but especially amongst single-units. It may be that this is the only way such establishments could keep up with the large, national chains. In terms of productivity, national establishments increased their productivity advantage both by selection (there is an especially large gap between exiting and entering national establishments) and by smaller productivity losses for continuing establishments. National establishments that exit exhibit especially high churning. In terms of multivariate results, the selection effects on productivity are especially large for national establishments consistent with the above finding of a large gap between exiting establishments and incumbents. Somewhat surprisingly large national establishments that are more human capital intensive are more likely to exit. Perhaps this reflects a shift in the composition of large, national establishments over this period of time.

In trucking, national establishments stand out as being larger, more skill intensive, more productive and less prone to churning. Relative shares of establishments, employment, and sales among these firms remained more or less unchanged (although national establishments lost some market share from 1992 to 1997). In terms of dynamics, national establishments lost some of their productivity advantage over time as selection dynamics (the gap between entering and exiting establishments.) worked in the wrong direction: All types of establishments exhibited skill increases and decreases in churning. In terms of the multivariate exit probits, single-unit establishments and local establishments are especially adversely impacted by churning, and national establishments are especially adversely impacted by low human capital. All types of trucking establishments are adversely impacted by low productivity.

3.5.4 Small vs. Large Software Establishments

In the software industry, there are both small, custom-designed software producers and also very large pre-packaged software producers. To explore this facet of the industry in a simple manner, we classified establishments into small and large based upon whether or not employment is above or below the mean (about 20 workers). Tables 3.28 - 3.31 report the results for the software industry classified on this dimension.

Year	Size	Revenue /worker ($)	Churning Rate (%)	Human Capital (%)	Person Effect (%)	Employment (Number)	Payroll /worker ($)
1992	Small	114,757	20.3	71.4	72.8	4.9	34,184
	Large	128,584	19.5	73.8	76.3	94.1	40,713
1997	Small	136,875	17.0	77.5	75.4	4.9	37,382
	Large	152,333	17.4	81.1	79.3	96.8	43,913

Table 3.28: Means of Core Variables in Software Industry by Type

Year	Size	Establishments	Employment	Sales
1992	Small	84.1	21.5	18.2
	Large	15.9	78.5	81.8
1997	Small	80.3	17.0	13.8
	Large	19.7	83.0	86.2

Table 3.29: Shares of Activity in Software Industry by Type

In both 1992 and 1997, large software producers account for only about 20 percent of the establishments but more than 80 percent of the sales. Large software producers have higher revenue per worker, pay higher wages, are more skill intensive and have slightly lower churning. In terms of dynamics, the gap between large and small stayed relatively constant over this time although the productivity gap between large and small widened largely through greater productivity gains for continuing, large establishments. A larger gap between entering and exiting small establishments for small firms worked offset some of these trends. In terms of multivariate exit probits, productivity has a greater impact for large establishments while churning has an especially large adverse impact for small establishments.

	Year	Size	Revenue/Worker	Churning	Person Effect
Exiter	**1992**	**Small**	**-0.412**	**0.079**	**-0.058**
			(0.029)	(0.016)	(0.010)
		Large	**-0.203**	0.018	-0.016
			(0.048)	(0.026)	(0.012)
Survivor	**1992**	**Small**	**-0.244**	0.006	**-0.050**
			(0.027)	(0.015)	(0.008)
		Large	**-0.126**	0.031	**-0.020**
			(0.037)	(0.020)	(0.009)
Entrants	**1997**	**Small**	**-0.157**	**0.043**	0.009
			(0.027)	(0.015)	(0.008)
		Large	0.003	0.017	**0.035**
			(0.041)	(0.022)	(0.010)
Continuers	**1997**	**Small**	**-0.181**	**-0.053**	**-0.055**
			(0.028)	(0.015)	(0.008)
		Large	0.000	0.000	0.000
			(-)	(-)	(-)
R-squared			0.050	0.013	0.048
N			11,673	11,030	4,786

* Estimation is based on Probit with stay/exit as a dependent variable.
** Standard errors are in parenthesis
*** Controls: Four digit SIC

Table 3.30: Differences of Core Variables in Software Industry by Type

	Size	
Single unit dummy		**-0.094**
		(0.025)
Size	**Small**	-0.027
interacted with:		(0.025)
	Large	-0.019
		(0.019)
Revenue/worker	**Small**	-0.029
interacted with:		(0.020)
	Large	**-0.052**
		(0.026)
Churning	**Small**	**0.304**
interacted with:		(0.078)
	Large	0.020
		(0.112)
Person Effect	**Small**	-0.065
interacted with:		(0.073)
	Large	0.125
		(0.116)
N		2,045

* Estimation is based on Probit with exit/stay as a dependent variable.
** Standard errors are in parenthesis
*** Controls: Four digit SIC

Table 3.31: Probability of Exit of Establishment in Software Industry by Type

3.6 Summary

Firm performance is tightly linked with workforce quality and worker turnover. Measures of productivity, workforce quality and worker turnover are highly correlated across businesses. High productivity businesses have a large share of high human capital workers, whether measured in terms of general skills or experience, and also have low churning of workers. Survival is a function of all of these factors - businesses with high productivity, low churning and high human capital (especially in terms of general skills) are more likely to survive. The patterns of these results vary substantially across the five industries. For example, worker churning is especially important in semiconductors while human capital is especially important in the trucking industry. Some of the patterns observed are due to inherent measurement problems associated with measuring firm performance. In particular, as others have noted, it is difficult to measure output and in turn productivity in the financial services sector. We found that the standard measure of revenue per worker as a measure of labor productivity works reasonably in all sectors but financial services. However, even here, this measure has reasonable properties in selected subindustries like securities brokers.

While there are some common patterns, we found it helpful to take into account the idiosyncratic factors present in the industries under investigation. Some of the anomalous patterns for the semiconductor industry, for example, are apparently driven by the changing composition to fabless semiconductor establishments. Such establishments are by construction recent entrants but are also small and highly human capital intensive thus changing the dynamics in that industry. In retail food and trucking there are large disparities between the characteristics and dynamics of establishments that are part of national chains vs. small, local establishments. In a related but different way, there are large differences between small and large software producers.

All of these industry-specific idiosyncratic factors make one cautious about drawing general inferences from the analysis. A natural reaction is to say how can we possibly compare the results for high tech semiconductors with the changing composition to fabless establishments with retail food establishments that are facing the wave of entry of new establishments from large, national firms. While this caution is warranted in some respects, at the end of the day we are struck by the common patterns rather than by the idiosyncratic factors. Workforce quality, worker churning and firm performance are related in the most basic and sensible of ways across all the industries studied and virtually all classifications of businesses that we have considered.

Chapter 4

Internal Labor Markets and Diversification Strategies in Financial Services

The theoretical perspective that has come to be known as the resource-based view of the firm suggests that sustainable competitive advantage often originates inside the firm, and that strategy at the firm level is therefore driven by firm-specific resources and capabilities. Human resources hold a prominent position in these resource-based theories of the firm. To date, however, few empirical studies have assessed the role human resources plays in driving firm strategies, largely because large-sample data on firm-level human resources are difficult to come by.

In this chapter, I take advantage of the development of new Census Bureau data sets developed out of the Longitudinal Employer-Household Dynamics (LEHD) Program and the Center for Economic Studies (CES) to explore the linkages between firm-level human resources and one aspect of firm strategy: diversification. The resource-based view of the firm suggests that diversification arises as firms attempt to leverage non-tradable firm-specific resources, among them human resources. We explore this possibility by examining recent diversification activity in the relatively newly deregulated American financial services industry to investigate whether characteristics of firms' internal labor markets influence their subsequent diversification activities.

4.1 Background

4.1.1 A Resource-Based View of Diversification

What sparks firm diversification, and why do diversified firms exist? Studies of diversification have long been a mainstay of economics as well as strategic management research (Montgomery, 1994; Ramanujam and Varadarajan, 1989; Hoskisson and Hitt, 1990). Economic theory generally assumes that firms are organized with a single product focus and face a homogeneous factor market (Scherer, 1980). Based on those assumptions, a market power view (Edwards, 1955) of diversification emphasizes the benefits a firm may reap at the expense of its competitors and customers.

More skeptical views offered by agency theorists emphasize the benefits that diversification offers to firm managers themselves, often at the expense of its shareholders (Jensen, 1986; Shleifer and Vishny, 1989).

The resource-based view of the firm (Barney, 1991; Penrose, 1959; Wernerfelt, 1984) suggests a different perspective, emphasizing firm resources and capabilities as the principle basis for strategy, including diversification activity. The resource-based view begins with the idea that firms are heterogeneous with respect to resources and capabilities that are not perfectly mobile across firms (Teece, Pisano and Shuen, 1997; Barney, 1991, 1986; Montgomery, 1994; Amit and Shoemaker, 1993; Hoskisson and Hitt, 1990; Teece, 1982; Penrose, 1959). Resources are stocks of available factors that are owned or controlled by the firm, including physical, intangible, and financial resources (Chatterjee and Wernerfelt, 1991). Firm capabilities refer to a firm's capacity to deploy resources, usually in combination, using organizational processes, to affect a desired end (Amit et al., 1993).

In order for firm-specific resources and capabilities to generate competitive advantage, they must be valuable, relatively rare, and relatively inimitable or immobile (Barney, 1986; 1991), enabling the firm to earn rents. The effectiveness of firm strategies depends on the utilization and exploitation of existing resources (Dierickx and Cool, 1989; Wernerfelt, 1984; Teece, 1980; Penrose, 1959). To the extent that firms have pools of underused resources, these create unique, firm-specific opportunities for exploitation (Chandler, 1990; Teece, 1980; Penrose 1959; Mahoney and Pandian, 1992; Montgomery, 1994).

Diversification is one such strategy for exploiting existing firm-specific resources: firm diversification can be understood as a process through which managers first identify resources that are unique to their firm, and then decide in which markets those resources can earn the highest rents (Teece, Pisano, and Shuen 1997). Some firm resources are 'indivisible' (Penrose, 1959) and therefore 'sticky' (Teece, Pisano, and Shuen, 1997), and, particularly if they are intangible, difficult or impossible to trade in the market. Firms with these kinds of resources may seek to deploy them in product markets through diversification.

4.1.2 Diversification and Human Resources

One general extension of the resource-based view of the firm is that intangible resources, such as knowledge, are more likely to produce a sustainable competitive advantage than tangible resources (Hitt, Bierman, Shimizu, and Kochhar, 2001; Teece, Pisano, and Shuen, 1997), because other firms will find it more difficult to imitate firm-specific processes associated with value creation (Lippman and Rumelt, 1982; Nelson and Winter 1982; Dierickx et al., 1989). Foremost among intangible resources are human resources: the accumulated skills of the firm's employees in the context of the firm's practices for organizing work.

The resource-based view suggests that human resources have implications for diversification strategy (Penrose 1959; Teece 1982; Montgomery and Hariharan 1991; Lei, Hitt and Bettis 1996). Most firm knowledge and other intangible resources reside in firm employees (Hitt et al. 2001). Firms can be expected to exploit these resources. To the extent that it serves to leverage firm resources in other market segments (Wernerfelt, 1984), diversification has the potential to move a firm toward more extensive utilization of its human resources. This is especially true where human resources create knowledge and information, which from a perspective internal to the firm are quasi-public goods that can be exploited at close to zero marginal cost.

Despite their potential salience, previous studies have not directly assessed the extent to which human resources affect diversification strategies. As Farjoun (1994: p.187) noted, "empirical studies have primarily focused on R&D and advertising or other 'tangible' assets, essentially avoiding the simple observation that business organizations ultimately consist of people." For example, Chatterjee and Wernerfelt (1991) considered R&D and advertising intensity in predicting the types of market firms choose to enter, while Schoenecken and Cooper (1998) showed that R&D and marketing activities influence entry timing. Because they can be leveraged at low marginal cost, R&D and marketing capabilities are found to generate diversified expansion (Montgomery and Hariharan 1991).

The paucity of empirical research on human resources can be ascribed primarily to difficulties in measurement, particularly in measuring human capital (Steffy and Maurer 1988; Ichniowski, Shaw and Prennushi 1997). There are some hints that human resources matter. Studies of law firms indicated linkages between firm strategy and leverage of human resources (Sherer 1995) and between leverage and firm performance (Hitt et al. 2001). And Farjoun (1994) showed empirically that diversification across industries was more likely where the industries had related "human resource profiles," or clusters of occupations.

The majority of research on diversification in the financial services industry has focused on the implications of diversification, such as the impact diversification has on firm performance. Few have analyzed the types of firms that have diversified. We are trying to explore the relationship between human resource practices and the extent of diversification based on the resource-based view.[1]

In this study we are able to take advantage of newly available data from the Census Bureau to extend our understanding of the relationship between human resources and diversification. LEHD data enable us to construct firm-level measures of human resources for a large sample of firms. These data allow us to investigate empirically the connections between firm resources - specifically, human resources - and subsequent diversification activity of firms.

4.1.3 Deregulation and Diversification: Financial Services

Our study is set in the U.S. financial services sector. This sector is an especially good venue for examination of the effects of human resources on diversification. Until recently, regulation constrained firms from a full range of diversification activity; many of these regulatory constraints disappeared over the course of the 1990s. This meant that firms developed human resources over a period in which diversification activity was limited; the relaxation of those limits therefore provides us with an opportunity to assess the extent to which human resources are associated with subsequent diversification activity.

U.S. firms were long prevented from engaging in activities across sub-sectors of financial services, primarily by regulation associated with the Glass-Steagall Act of 1933. The Gramm-

[1]The implications of firm-specific human resource practices on diversification have never been investigated to our knowledge.

Leach-Bliley Act of 1999 formally removed a large set of regulatory restrictions on banks by explicitly permitting financial "holding companies" (and their subsidiary firms) to participate in brokerage activities, underwriting, and the provision of financial advice.[2] This deregulation occurred to some extent after the fact; U.S. banks had been expanding their business beyond lending and deposits and toward provision of a broader set of financial services for years prior to the passage of Gramm-Leach-Bliley. In fact, deregulation began in earnest in the late 1980s, when the Federal Reserve Board began allowing commercial banks to enter the investment banking industry - first allowing commercial banks to underwrite corporate bonds in 1989 (for example, Gande, Puri, and Saunders (1999)), so that by the early 1990s commercial banks began to gain a meaningful share of the investment banking market.

Cross-sectoral activity represents one mode of diversification in financial services. A second mode is geographic. While in some sectors (such as brokerage and insurance firms) firms have operated on a national scale for some time, this was not true in other sectors, particularly in the banking industry, where various kinds of regulation restricted geographic diversification. Prior to 1970, for example, branch banking even within state boundaries was somewhat limited, and all states prohibited interstate branching.[3]

Over the following two decades, restrictions on intrastate and interstate branching gradually eased. Intrastate deregulation first allowed holding companies to own multiple banks, then allowed these holding companies to integrate these banks as members of a single branch system. In 1975, Maine provided the first opportunities for interstate banking, by allowing holding companies from other states to acquire banks in Maine. Over the 1980s, many states established arrangements in which their banks could be bought by banks from either selected states or all other states. In 1994, Congress passed the Riegle-Neal Interstate Banking and Branching Efficiency Act, allowing full interstate banking, and by 1997 all states but Texas and Montana (each of which passed legislation opting out of Riegle-Neal) permitted complete interstate banking.

[2]See Fay (2000) for a discussion of Gramm-Leach-Bliley.

[3]See Kroszner and Strahan (1999) for a more complete discussion of the timing and effects of the geographic deregulation of banking.

Deregulation was accompanied by mergers and acquisitions. Over the late 1980s and 1990s, large firms acquired smaller ones, expanding their reach across both geography and scope of activities. The overall sweep of deregulation in the 1990s allowed firms to operate nationally and across many financial sectors for the first time in several decades. To take one example: in 1998 Citicorp anticipated Gramm-Leach-Bliley by merging with Travelers Group, itself the result of acquisitions and mergers of such businesses as the investment banks Salomon Inc., Smith Barney, and Drexel Burnham Lambert, Travelers Life and Annuity in insurance, the property and casualty divisions of Aetna, and the retail brokerage and asset management operations of Shearson Lehman. By 2004, Citigroup had credit card customers in every U.S. state and its expansive branch banking network served retail customers in 22 states.

4.1.4 Diversification, and Human Resources in Internal Labor Markets

The deregulation wave of the 1980s and 1990s opened up previously non-existent opportunities for diversification. The mergers, acquisitions, and greenfield growth that produced these increasingly diversified firms reflect firms' searches for new customers and enhanced market power. But this begs the question: which firms were likely to diversify?

Our framework suggests that firm resources could play an important role in determining diversification strategy. Specifically, we have suggested that firms with greater intangible assets in the form of human resources are more likely than others to seek to leverage these assets through diversification. These intangible assets include firm-specific skills. As Montgomery and Wernerfelt (1988) noted, sources of value that are not firm-specific are insufficient to allow firms to enter industries where more specialized factors are required. While general skills can create value, these values do not sustain competitive advantage or create valuable resources that yield economic rents, because they are freely tradeable. Tradeability thus has clear implications for diversification strategy: the value of nontradeable assets, or resources, cannot be realized in factor markets. In order to tap their rent earning potential, owners of such assets must deploy them in product markets (Dierickx and Cool, 1989). Similarly, Montgomery and Wernerfelt (1988) note, following

Williamson (1985), that standard theory suggests that value arising from firm-specific skills will be deployed internally, and that such circumstances should be associated with diversification.

We suggest that firms whose human resources reflect greater levels of firm-specific skills and capabilities are more likely to diversify. For two reasons, these firms are likely to be those with robust internal labor markets (ILMs). First, internal labor markets encourage the development of firm-specific skills. Firm-specific skills are especially important because they are more likely than general skills to be associated with the slack resources that diversification seeks to exploit. Second, firms with strong internal labor markets are more likely to have valuable human-resource derived intangible resources and capabilities beyond the skills of the workers themselves: team-level, unit-level, and organizational knowledge, and accumulated social capital (Cappelli, 2004). This reasoning leads us to hypothesize that firms with stronger ILMs are more likely to diversify subsequently. Such diversification may take two forms: operating in a more extended geographic range and offering services in more sub-sectors of financial services.

4.2 Data Sources

4.2.1 The Longitudinal Business Database (LBD)

We draw our diversification measures in the industry from the LBD, for which a detailed description is available in Jarmin and Miranda (2002). A few points about its construction are useful here. The LBD is created by linking data from annual business register files. The Census Bureau's business register, the Standard Statistical Establishment List (SSEL), is a continuously updated database of basic information about all employer business establishments in the U.S., and the Center for Economic Studies maintains annual snapshot SSEL files from 1975 onward. Currently, the LBD contains very good longitudinal linkages for all employer business establishments in the U.S. from 1975 to 2000. These linkages provide an exact measure of establishment age for all establishments born after 1975. The LBD contains basic information on establishment employment, payroll, location, industrial classification and firm affiliation. The LBD contains numeric establishment identifiers that allow it to easily be matched to other Census Bureau establishment level datasets that contain more detailed survey based information. The LBD also contains numeric firm identifiers that allow researchers to aggregate the establishment level data up to the company level. We make use of this approach in this chapter.

4.2.2 The Longitudinal Employer-Household Dynamics (LEHD) Program

We also exploit new Census Bureau data from the LEHD Program. The LEHD Program integrates information from state unemployment insurance data and Census Bureau economic and demographic data in a manner that permits the construction of longitudinal information on workforce composition at the firm level. This Program represents a substantial investment made by the Census Bureau in order to permit direct linking of its demographic surveys (household-based instruments) with its economic censuses and surveys (business and business unit-based surveys).

The unemployment insurance (UI) wage records are discussed elsewhere (see Burgess, Lane and Stevens, 2000). Every state in the U.S., through its Employment Security Agency, collects quarterly employment and earnings information to manage its unemployment compensation program. These data enable us to construct quarterly longitudinal information on employees. The advantages of UI wage record data are numerous. The data are frequent, longitudinal, and potentially universal. The sample size is generous and reporting for many data items is more accurate than survey based data. The advantage of having a universe as opposed to a sample is that movements of individuals to different employers and their consequences for earnings can be tracked. It is also possible to construct longitudinal data using the employer as the unit of analysis.

Perhaps the main drawback of the UI wage record data is the lack of even the most basic demographic information on workers (Burgess, Lane and Stevens 2000). Links to Census Bureau data overcome this for two reasons. First, individual wage records can be integrated with administrative data at the Census Bureau containing information such as date of birth, place of birth, and gender for almost all the workers in the data. Second, LEHD staff have exploited the longitudinal and universal nature of the dataset to develop measures of workforce quality using the methodology described in detail in Abowd, Creecy and Kramarz (2002) and in Abowd, Lengermann and

McKinney (2003).

The LEHD Program now houses data from more than thirty states. In this chapter, however, our attention to the role of internal labor markets in accounting for the evolution of diversification in the financial services in the 1990s requires extensive LEHD data from the early 1990s. We have data on all establishments and all firms in the financial services sector (defined precisely below) in the LBD from 1992 through 2000. We also have LEHD data on all establishments and all firms in the financial services sector for three large selected states. The crosswalk between these files is based on a common business-level identifier and the match rate between these files is extremely high.

4.2.3 Diversification in Financial Services

We investigate geographic and industry diversification in the U.S. financial services industry. We focus on financial services because of the unique opportunities that deregulation of the industry presented in the 1990s. The deregulation of the industry acts as a form of a natural experiment during our sample period - that is, financial services firms saw the opportunity set change dramatically in response to regulatory changes that can for our purposes be viewed as exogenous. Our analysis explores which firms changed their diversification in response to this deregulation as a function of the ILM structure of the firm.

Table 4.1 lists the 4-digit 1987 SIC codes we used to identify establishments in financial services, and in both the LBD and LEHD data we can measure activity at the establishment-level. Our analysis of diversification, however, is conducted at the firm-level, and thus we aggregate our establishment activity to construct firm measures. The Census Bureau maintains firm corporate structure of all establishments in the U.S. using a definition of operational control, giving a common identifier to any establishment under the operational control of a parent firm. In what follows, we exploit this rich characterization of the corporate structure to determine financial services firms - all the financial services establishments under a common firm identifier. Put differently, we focus on only the financial services components of firms.[4] We define a financial services firm as a firm-

[4]Note that large firms that are not thought of primarily as financial services firms (for example, auto companies)

1987 SIC Code	Description
6021	National Commercial Banks
6022	State Commercial Banks
6029	Commercial Banks NEC
6035	Savings Institutions (Fed)
6036	Savings Inst (Not Fed)
6061	Credit Unions (Fed)
6062	Credit Unions (Not Fed)
6081	Branches of Foreign Banks
6099	Functions Related to Deposit Banking
6111	Federal Credit Agencies
6141	Personal Credit Inst
6153	Short Term Business Credit Inst
6159	Miscellaneous Business Credit
6162	Mortgage Bankers & Loan Correspondents
6163	Loan Brokers
6211	Security Brokers and Dealers
6221	Commodity Contracts Brokers and Dealers
6231	Security and Commodity Exchanges
6282	Investment Advice
6289	Securities Exchange Services
6311	Life Insurance
6321	Accident and Health Insurance
6324	Hospital & Medical Service Plans
6331	Fire Marine and Casualty Insurance
6351	Surety Insurance
6361	Title Insurance
6371	Pension, Health and Welfare Funds
6399	Insurance Carriers
6411	Insurance Agents, Brokers, and Service
6712	Offices of Bank Holding Companies

Table 4.1: SIC codes in financial services

year observation comprising establishments in any of the 4-digit SIC industries listed in Table 4.1, in the firm in that year. We create diversification measures for these financial services firms based on LBD data for the period 1992-2000 (though our analysis of diversification will focus on the period 1997-2000).

Table 4.2 shows counts of single- and multi-unit financial services firms (hereafter firms) for

may have substantial financial services components. We have examined the diversification measures using all components of the firms (including the non-financial services components) and we obtain similar basic patterns. However, given that the deregulation in the financial services industry primarily impacted the financial services components of firms, we focus on the financial services components only.

Year	Single Units		Multi-Units		Total
	Number of Firms	Percent of Total	Number of Firms	Percent of Total	
1992	157,959	90.3	17,059	9.7	175,018
1993	163,575	90.9	16,471	9.1	180,046
1994	166,126	91.4	15,698	8.6	181,824
1995	164,606	91.7	14,984	8.3	179,590
1996	169,070	92.6	13,577	7.4	182,647
1997	181,277	91.6	16,684	8.4	197,961
1998	188,965	92.1	16,258	7.9	205,223
1999	193,377	92.6	15,530	7.4	208,907
2000	195,645	92.9	14,881	7.1	210,526
Total	**1,580,600**	**91.8**	**141,142**	**8.2**	**1,721,742**

Source: Authors' calculations on the LBD.

Table 4.2: Single- and multi-unit firm counts

Year	Single Units		Multi-Units		Total
	Number of Establishments	Percent of Total	Number of Establishments	Percent of Total	
1992	157,959	46.6	180,713	53.4	338,672
1993	163,575	48.3	175,345	51.7	338,920
1994	166,126	47.7	182,072	52.3	348,198
1995	164,606	47.2	184,452	52.8	349,058
1996	169,070	47.8	184,299	52.2	353,369
1997	181,277	47.6	199,253	52.4	380,530
1998	188,965	48.8	198,534	51.2	387,499
1999	193,377	48.4	206,200	51.6	399,577
2000	195,645	48.8	204,887	51.2	400,532
Total	**1,580,600**	**47.9**	**1,715,755**	**52.1**	**3,296,355**

Source: Authors' calculations on the LBD.

Table 4.3: Single- and multi-unit establishment counts

Year	Single Units		Multi-Units	
	Percent of Total Payroll	Percent of Total Employment	Percent of Total Payroll	Percent of Total Employment
1992	13.90	15.09	86.10	84.91
1993	16.17	16.92	83.83	83.08
1994	13.64	16.97	86.36	83.03
1995	13.41	17.02	86.59	82.98
1996	14.97	18.19	85.03	81.81
1997	15.09	17.03	84.91	82.97
1998	16.80	18.73	83.20	81.27
1999	16.03	18.84	83.97	81.16
2000	15.56	19.05	84.44	80.95
All Years	**15.16**	**17.58**	**84.84**	**82.42**

Source: Authors' calculations on the LBD.

Table 4.4: Percent of payroll and employment represented by single- and multi-unit firms

the U.S. over the period 1992-2000, drawn from the LBD data for all 50 states plus the District of Columbia. The number of single-unit firms grew by over 20% during this period, even as the number of multi-unit firms in financial services dropped by more than 10% in the same time frame. The drop in multi-unit firms is attributable to the substantial pace of consolidation activity between medium- and large-sized firms in the industry. Table 4.3 shows establishment counts over the same period, indicating that the number of establishments in both single- and multi-unit firms grew over the period, with slightly higher growth among the establishments that did not belong to multi-unit firms. Table 4.4 shows that the relatively small share of multi-unit firms in the sector account for the vast share of sector activity: over 80% of both employment and payroll are represented by multi-unit firms.

We create five different diversification measures: three simple measure of overall diversification, and two measures of relatedness in diversification. The simpler measures are equal to one minus a basic Herfindahl index: industry diversification (ind_div); county diversification ($county_div$); and state diversification ($state_div$). We also use weighting to account for the role of larger establishments in diversification activity. LBD data for payroll are quite reliable and thus we prefer payroll-weighting to employment-weighting. We construct all measures using payroll weights (in the aggregate, employment and payroll, as Table 4.4 suggests, represent roughly similar shares of activity).

We illustrate the construction of the measure for industry diversification. We calculate total payroll (pay_{it}) for firm i in year t, the total payroll (pay_{jit}) for establishments (e) operating in industry j in firm i in year t, and the payroll share (s_{jit})for establishments operating in industry j in firm i in year t.

$$
\begin{aligned}
pay_{it} &= \sum_{e \in i} pay_{eit} && (4.1)\\
pay_{jit} &= \sum_{e \in i \cap e \in j} pay_{eit} && (4.2)\\
s_{jit} &= \frac{pay_{jit}}{pay_{it}} && (4.3)
\end{aligned}
$$

We use these measures to create basic Herfindahl indices (in this case, payroll-industry) for firm i in year t.

$$H_{it}^{industry} = \sum_{j \in i} (s_{jit})^2 \tag{4.4}$$

From this measure we create our index, *ind_div*, for firm i in year t.[5]

$$ind\_div_{it} = 1 - H_{it}^{industry} \tag{4.5}$$

County and state diversification are calculated similarly.

We also calculate measures of relatedness in diversification, using distance-based diversification indices weighted by payroll. We call these measures *geog_dist_div* for geographic diversification and *ind_dist_div* for industry diversification. For these measures, we also begin by calculating payroll and payroll shares. With respect to geography, we then proceed to identify the "core" county c of firm i in year t. The "core" county is defined as the county with the highest payroll share in

[5]Our measurement approach to diversification follows in the spirit of Gollop and Monahan (1991) both for the basic measures and the distance-relatedness measures of diversification.

firm i in year t. From here, we create the following diversification index:

$$H_{it}^{county} = \sum_{j \in i} (1/d_{ce}) \, (s_{jit})^2 \tag{4.6}$$

where d_{ce} is $1 +$ the distance from the center of the county where establishment e is located and the "core" county c. This enables the construction of the variable *geog_dist_div* for firm i in year t.

$$geog\_dist\_div_{it} = 1 - H_{it}^{county} \tag{4.7}$$

For our "distance"-based diversification index weighted by payroll for industry, we also calculate total payroll (pay_{it}) and payroll share. We then proceed to identify the "core" industry (j) of firm i in year t. The "core" industry is defined as the industry with the highest payroll share in firm i in year t. From here, we create the diversification index

$$H_{it}^{industry} = \sum_{j \in i} (1/d_{je}) \, (s_{jit})^2 \tag{4.8}$$

where $d_{je} = 1$ if firm i operates only in one 4-digit industry, 2 if establishment e operates in the same 3-digit industry as the firm "core" industry j, 3 if establishment e operates in the same 2-digit industry as the firm "core" industry j, and 4 otherwise. This enables the creation of the *ind_dist_div* diversification index for firm i in year t:

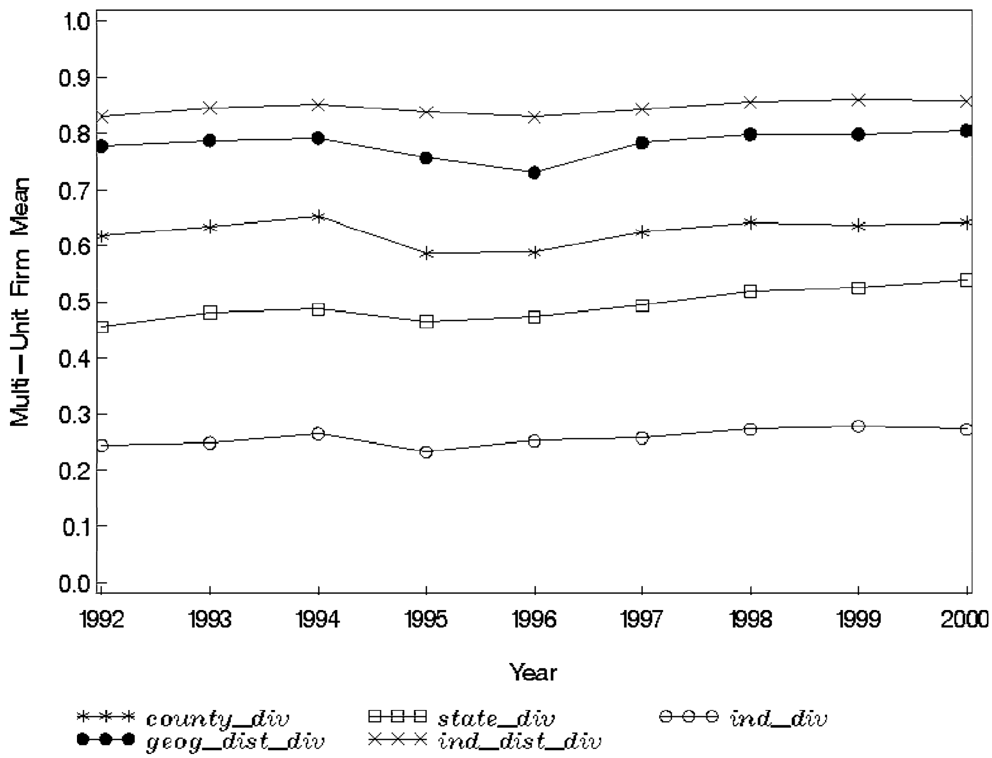
$$ind\_dist\_div_{it} = 1 - H_{it}^{industry} \tag{4.9}$$

All diversification measures are bounded in the interval [0,1] and are equal to 0 for single-unit and other completely non-diversified firms.

Year	County Diversification		State Diversification		Industry Diversification		Geographic Distance		Industry Distance	
	All firms	Long-term continuers	All firms	Long-term continuers	All firms	Long-term continuers	All firms	Long-term continuers	All firms	Long-term continuers
1992	0.618	0.668	0.456	0.520	0.244	0.273	0.778	0.814	0.831	0.860
1997	0.626	0.648	0.494	0.516	0.258	0.279	0.784	0.797	0.844	0.858
2000	0.642	0.678	0.541	0.568	0.273	0.308	0.806	0.839	0.859	0.887

Source: Authors' calculations on the LBD.

Table 4.5: Mean of firm diversification measures weighted by total firm payroll, selected years

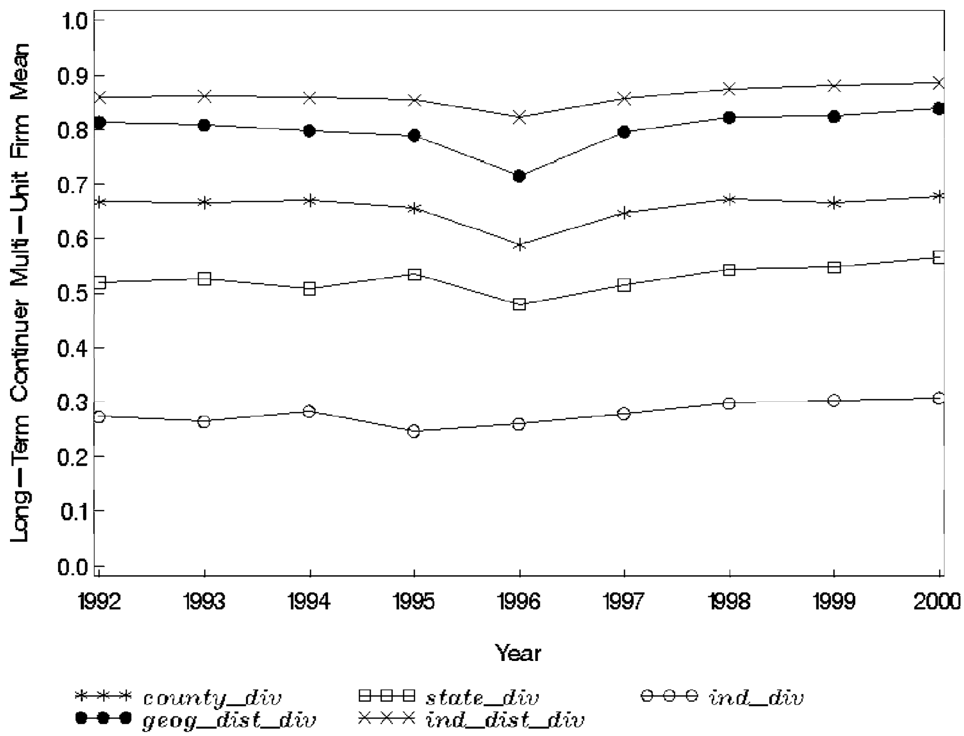


Source: Authors' calculations on the LBD.

Figure 4.1: Mean of multi-unit firm diversification measures weighted by total firm payroll

Figure 4.1 shows the annual mean of all five firm-level diversification measures weighted by the total payroll of the firm for multi-unit firms (Table 4.5 shows the data for each measure for key years in this period in more detail). Average diversification levels of the firms in each of our five indices show modest growth for the period 1992-2000, consistent with the stylized facts for the sector. Firm-level geographic diversification, whether measured at the county or the state level, and industry diversification, are highest at the end of the period. Firms also appear to be decreasing the extent to which their activities are related; the distance indices by geography and by industry are also higher at the end of the period.

Our analytical strategy will be to focus on changes in firm diversification for the period 1997-2000, enabling us to use data on the characteristics of firms' internal labor markets from the period preceding this window. Moreover, our focus on 1997-2000 implies that we are examining changes from perspective pre- and post-passage of the Gramm-Leach-Bliley Act. We earlier noted



Source: Authors' calculations on the LBD.

Figure 4.2: Mean of long-term continuer multi-unit firm diversification measures weighted by total firm payroll

that some aspects of deregulation, and considerable diversification activity, clearly pre-date the Act. Figures 4.1 and 4.2, however, show that over the 1997-2000 period the trends in the industry continue to point toward increasingly diversified organizations, operating over increasingly distant geographies and across somewhat less related industries.

The composition of our sample of firms changes over time as a result of entry and exit activity. From our sample, we define long-term continuing firms as firms that appear in the sample for every year from 1992-2000. There are 79,840 single-unit and 10,192 multi-unit long-term continuer financial services firms. Figure 4.2 displays all five firm-level diversification measures weighted by the total payroll of the firm for long-term continuer multi-unit firms. Over the period we are studying, the long-term continuers in the sample show greater increases in diversification levels by the various measures, a fact that is also consistent with our account of existing firms' pursuit of increasingly diversified activities in the 1990s (again, see Table 4.5 for more detail).

While the mean level of diversification in financial services increased by any measure over our period, firms did not follow identical strategies. Perhaps not surprisingly, most firms experience little change in their diversification levels over the period. Our sample is, however, characterized by considerable heterogeneity in strategies even over the relatively short time window we have chosen, and includes not only firms with varying levels of diversifications, but firms that decreased as well as increased their range of activities both geographically and sectorally. Examination of plots of diversification measures for both 1997 and 2000 helps to make this clearer, as shown in Figure 4.3.

We can further consider firm diversification strategies by decomposing changes in the various measures of diversification, separating the roles of continuers from changes generated by firm entry and exit. Let the aggregate weighted average of firm-level index be given by:

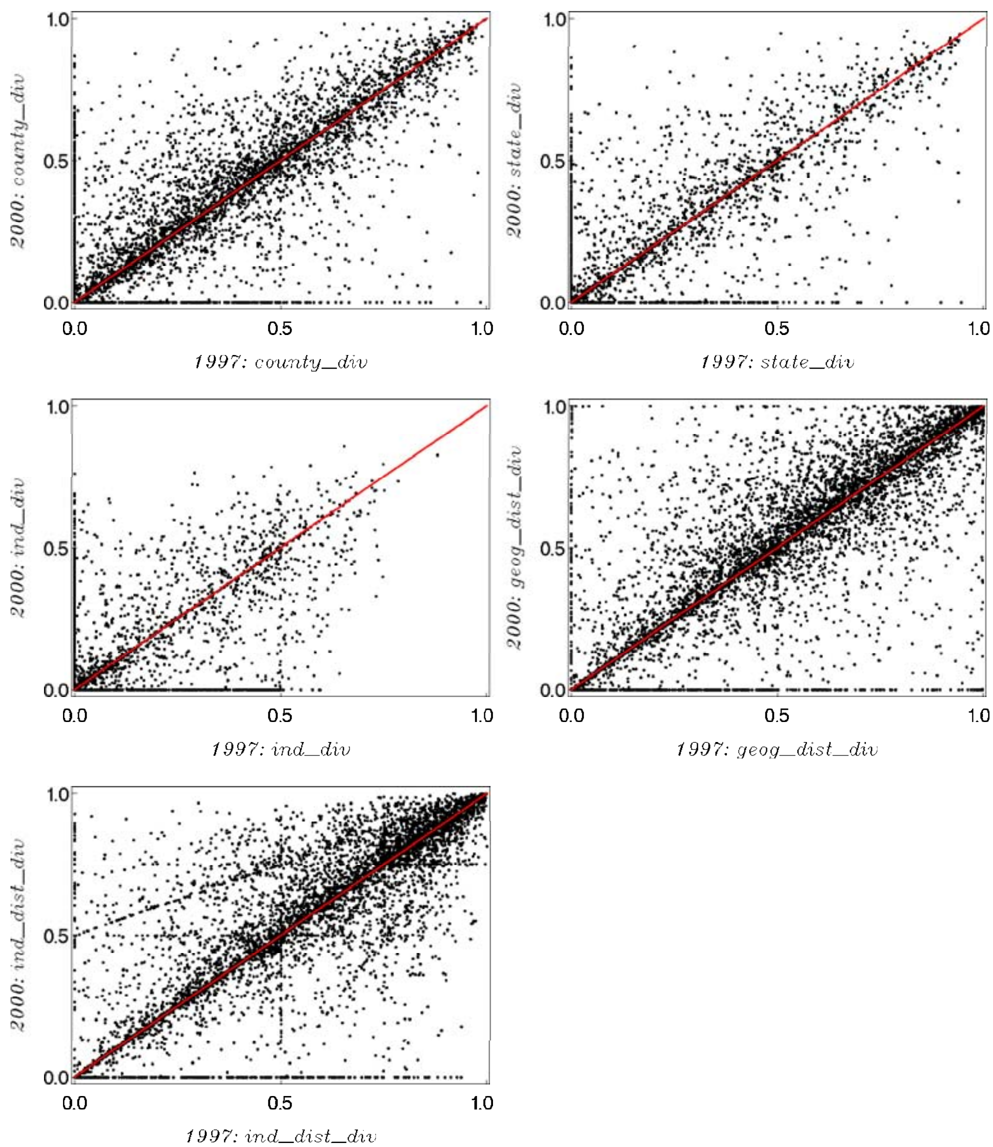
$$D_t = \sum_i \theta_{it} D_{it} \tag{4.10}$$

where D_t is the share-weighted average diversification index, D_{it} is the diversification index for firm i, and θ_{it} is the share of firm i. Consider the following decomposition:

$$\begin{aligned} \Delta D_t &= \sum_{i \in C} \theta_{it-1} \Delta D_{it} + \sum_{i \in C} (D_{it} - D_{t-1}) \Delta \theta_{it} + \sum_{i \in C} \Delta D_{it} \Delta \theta_{it} \\ &+ \sum_{i \in N} \theta_{it} (D_{it} - D_{t-1}) - \sum_{i \in X} \theta_{it-1} (D_{it-1} - D_{t-1}) \end{aligned} \tag{4.11}$$

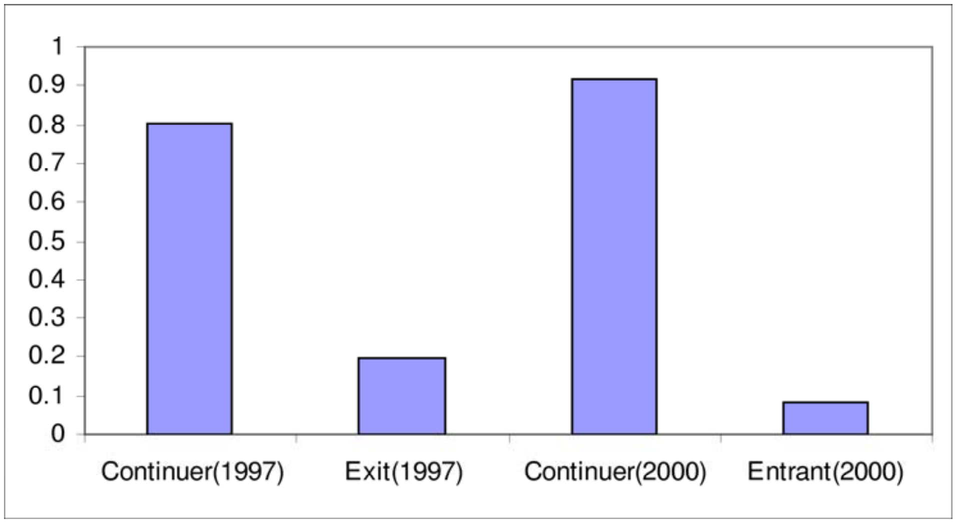
The sets C, N, and X respectively represent the set of continuing firms, entering firms, and exiting firms. This decomposition involves four terms: a within-firm effect, a between-firm effect, a cross effect, and a net entry effect. We define firm entry and exit in terms of changes in the firm identification code, and as such, a firm that is acquired will result in an "exit" of a firm. In what follows, since our firm-level diversification measures are based upon using payroll as a measure of activity, we use firm shares of total industry payroll as weights in the aggregation and decomposition.

Decomposition of changes in our measures for the period 1997-2000 clearly reveals that the increase in diversification levels over this period is generated by continuing firms. Figure 4.4 shows that most of the activity in our sample takes place in continuing firms, but that entry and exit are



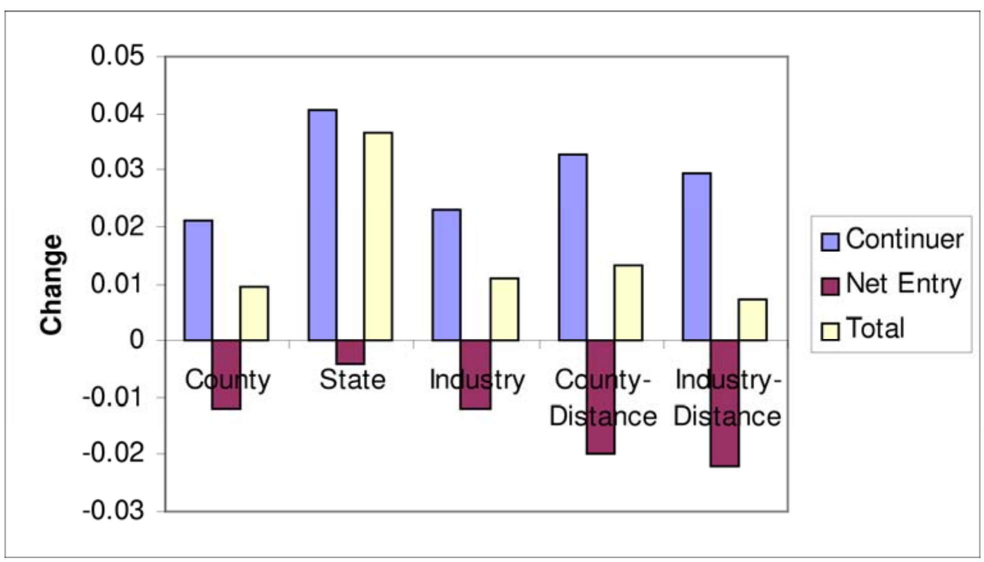
Source: Authors' calculations on the LBD.

Figure 4.3: Scatter plots of multi-unit firm diversification measures



Source: Authors' calculations on the LBD and LEHD data.

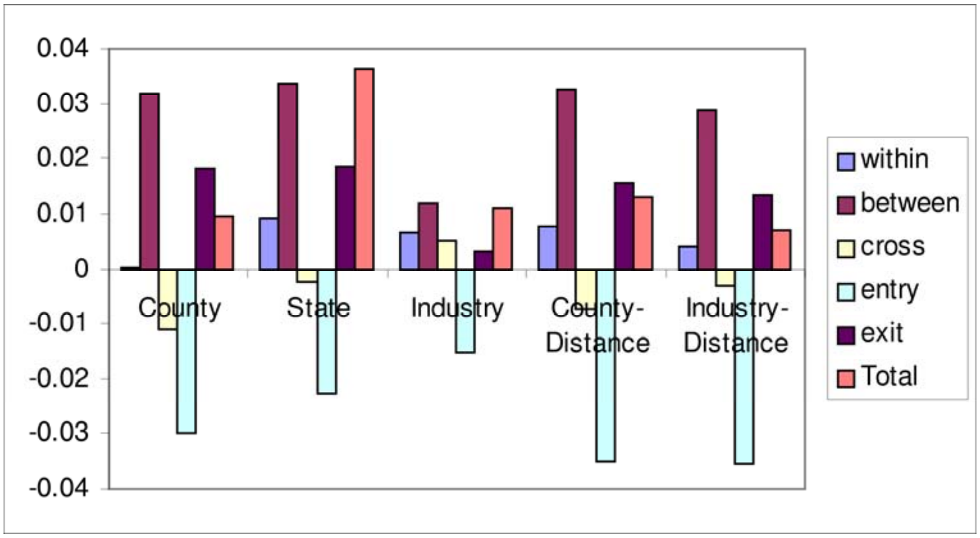
Figure 4.4: Shares of payroll for continuing, entering, and exiting firms



Source: Authors' calculations on the LBD and LEHD data.

Figure 4.5: Diversification decomposition (1997-2000 change): continuers, net entry, and total

substantial. Of the firms in the 1997 sample, about 20% (by payroll weight) exit over the period, while entrants in the 1997-2000 period account for only about 10% of the payroll weight in 2000. As Figure 4.5 shows, continuing firms increased their diversification levels over this period, while net entry actually generated a decline, and activity among continuers was especially prominent for diversification across state boundaries.



Source: Authors' calculations on the LBD and LEHD data.

Figure 4.6: Diversification decomposition (1997-2000 change): within, between, cross, entry, exit, and total

The decline attributable to net entry reflects mainly the fact that entrants to the industry tend not to be diversified (consistent with the resource-based view, these firms have few specific resources to leverage). While firms that exit are less diversified than the firms that continue, entrants are even less diversified than exiters, so the impact of net entry is negative. We find it striking that continuing businesses exhibit a pronounced increase in diversification over the 1997-2000 period. Figure 4.6 provides a fuller characterization of the dynamics of continuing firms by showing all of the components of the above decomposition. We find that the within terms (changes at the firm level weighted by initial shares) and especially the between terms (changes in the shares weighted by initial diversification) are both positive for all measures. Thus, the contribution of continuers arises both because the average continuing firm exhibited an increase in diversification and also because the firms that were already highly diversified in 1997 increased their share of activity.

4.3 Human Resources in Internal Labor Markets

4.3.1 Basic concepts and measurement

In this section we turn our attention from basic facts on diversification to our measures of human resources in internal labor markets. The LEHD data do not permit us to observe firm practices directly; rather, we have indicators of internal labor market outcomes that result from internal firm processes and practices. The LEHD data allow us to construct a number of indicators of the strength of firm-level internal labor markets and thus of firm resources. Each of these measures can be thought of as an outcome of ILMs, and likely to be associated with the resources that diversification seeks to exploit. We focus on three indicators of internal labor markets derived from the LEHD data: firm-level "churning" (worker turnover rates in excess of net changes); the extent to which wage-tenure profiles at firms slope upward; and the dispersion of wages within firms. We construct these measures initially at the establishment-level and then aggregate the measures to the firm level using appropriate employment weights.

First, consider the role of worker turnover in this context. The sorts of firm-specific skills that can be leveraged through diversification are likely to be acquired on the job, in firms that have relatively low worker turnover (Fairris, 2004). Moreover, it has long been argued that low quit rates are one feature of firms with strong internal labor markets (Doeringer and Piore, 1971). In this chapter we look specifically at worker "churning." We measure worker churning at the establishment-level as:

$$churning_t = \frac{Accessions_t + Separations_t - |Employment_t - Employment_{t-1}|}{0.5\,(Employment_t + Employment_{t-1})} \tag{4.12}$$

This measure captures the component of worker turnover or reallocation that is in excess of that needed to accommodate any net changes in the number of workers in the business. We have this measure on a quarterly frequency for every establishment, and aggregate it to an annual firm level by taking appropriate employment-weighted averages. We expect that firms with high mean rates of worker churning (chr) are less likely to accumulate firm-specific skills over time. Firms with relatively low churn rates, in contrast, are more likely to develop the sorts of skills that can be leveraged through diversification.

A second important aspect of internal labor markets is that they comprise opportunities to advance inside the firm. Such advancement is also conducive to the development of resources that can be leveraged through diversification. We do not have specific measures of advancement through job ladders, often seen as a key feature of internal labor markets (Pfeffer and Cohen, 1984), but the LEHD data allow us to construct a proxy for these kinds of opportunities through the identification of wage-tenure profiles inside the firm. Workers in firms with relatively strong internal labor markets are likely to have wage-tenure profiles that slope more sharply upward within the firm, as they are rewarded for seniority with higher-ranking and better-paying jobs.

Our second indicator of the strength of the firm's internal labor market is the mean growth of workers' wages over their period of employment. The LEHD data allow us to construct a profile of the "within-job-wage-growth" ($wjwg$) for each establishment in the firm. We focus on the five year period preceding 1997, taking the mean wage gains of all newly hired workers who begin spells of employment during the period 1992-1996, inclusive, and who have tenure for five or more years. In other words, our measure wjwg gives the average firm wage-tenure profile, built from the first five years of individual workers' tenure at the firm conditional on tenure lasting for at least five years.

A third feature of many internal labor markets is wage compression. Firms with strong internal labor markets are likely to feature less dispersion of wages across workers with similar jobs and skill levels. In internal labor markets, wages are set in part by bureaucratic rules and may not perfectly reflect forces in the external market. Such rules may reflect norms of equity, or arise for reasons of administrative convenience. Freeman (1982), for example, shows that unionized firms are more likely to feature wage compression for observationally equal workers, and unionization is also associated with the existence of strong internal labor markets (Kalleberg et al. 1996). Pfeffer and Langton (1993) show that wage compression is positively related to cooperation among workers. Such cooperation provides opportunities to build individual-specific skills, and has further effects because firm knowledge is embedded not simply in individuals' skills, but in "routines" (Nelson and Winter, 1992) and in relationships between individuals (Kogut and Zander, 1992). Routines and social capital are not transferred easily to other organizations; firms with valuable capabilities in these areas may seek to leverage them through diversification. Our measure of within-firm wage dispersion ($diff$) is the ratio of earnings of the worker at the 90th percentile in the firm to that of the worker at the 10th percentiles (expressed in logarithmic form).

The effects of these different aspects of internal labor markets on subsequent diversification may also be complementary. For example, if a relatively small number of workers stay with the firm long enough to accumulate skills, the effects of steep wage profiles may be less than if most workers tend to remain at the firm. Thus high rates of turnover (or our measure, churning) may tend to dampen the effects of high levels of wage growth. Similar reasoning applies to wage dispersion. We suggested that low levels of wage dispersion, particularly controlling for human capital, are more likely to be associated with skill development and accumulation of social capital that can be leveraged through diversification. This relationship should be stronger in firms with relatively low rates of turnover. Finally, we also expect the relationship between wage dispersion and steep wage profiles to be complementary. We expect that the negative effects of wage dispersion should be dampened by steeper wage profiles. That is, dispersion in firms where individuals have the opportunity to make wage gains should not have the same kinds of negative effects on skill accumulation, cooperation, and social capital as would dispersion in firms in which individuals do not have these kinds of opportunities.

Our approach will be to use measures of these indicators, constructed at the level of the firm, to predict subsequent firm-level diversification activity. The underlying premise here is that internal labor markets in financial service firms developed over time, perhaps in part as human resources strategies consciously chosen, in part as responses to institutional pressures, and in part due to idiosyncratic factors (which may have in turn induced firms to adopt alternative human resource practices). While we clearly recognize that firm-level differences in our indicators of ILMs are driven by many possible factors and inherently endogenous, our empirical strategy is to take advantage of the changes in the regulatory environment to identify the impact of ILMs on diversification. That is, in the 1990s deregulation and technological changes provided new opportunities to firms. The open question is not whether or why financial services diversified on average but rather, which firms increased diversification. Our working hypothesis is that the firms with well-established ILMs were in a better position to take advantage of these new opportunities, and thus we use the 1992-1996 outcomes of ILM processes at the firm level to predict changes in firm-level diversification activity in the 1997-2000 period. It is, of course, the case that deregulation and diversification began prior to 1997, but our approach is designed to relate changes from t to $t + k$ based upon initial conditions in period t.

We would like to note the merits and drawbacks of our ILM indicators. Because these indicators are outcome variables but not human resource management policy variables, we have to be cautious not to conclude and accept blindly that ILM policies can provide sustainable competitive advantage. Two firms with very different HRM practices may have the same level of measured ILM indicators that we use. However, according to the resource-based view, our ILM indicators are likely good proxies for the firm-specific, underused, rare, inimitable, and untradable "resources" that firms can exploit for diversification. As noted in Barney, Wright, and Ketchen (2001), individual HRM practices can be imitable but HRM systems and routines may not be because they are intangible assets accumulated over time and because they cannot be replicated. Hence, regardless of firms' intended HRM practices, these outcome variables may be good proxies for the "resources" of a particular firm which may be unique to that firm and contribute to the creation of specific human capital skills. Moreover, we measure our ILM indicators over a relatively long period of time so that they capture both firm HRM policies and actual human resource assets assuming that long term trends in these indicators better reflect firms' HRM policies.

We use another set of variables to control for other firm characteristics, both those related to general features of the firm, and those related to human resources. Firm size ($lnsize$) is measured by the average employment (in logs) of the firm (restricted, as noted above, to the financial services establishments of that firm) from 1992-96.[6] Firm $growth$ is measured by average quarterly employment growth over the period 1992-96 and firm age ($firmage1997$) is measured by the age in years of the oldest establishment in the firm as of 1997. Because longitudinal firm linkages are currently under development, exact measures of firm age are not yet available on the LBD. However, other work at CES (Becker et. al. 2004 and Davis et. al. 2004) has shown that using the age of the oldest establishment owned by a firm is a very good approximation. We control for the "home" state of each firm: our LEHD data are taken from three large states and we include dummy variables indicating which of the three states employs the largest share of employees. We also control for the chief sub-industry in which in each firm operates, including dummy variables for the 4-digit industries listed in Table 1 which take on a value of "1" for the sub-industry employing the largest share of the firm's workers, and a value of "0" otherwise. We control for these features of our firms since each might plausibly be related both to internal labor market characteristics and to diversification.

The LEHD data also allow us to control for other demographic features of sample firms' workforces. We control for the share of female workers, *shr_fem*, taken from the LEHD data. We also control for the firm's employment of high-skilled workers by including a measure for the share of high-skill workers, *shr_high*, derived from the LEHD data, and based upon the measures of workforce quality developed by Abowd *et al.* (2003). The workforce quality measures are based upon a statistical decomposition of the wage for a worker into a person effect, a firm effect and time varying person characteristics including general labor market experience. The person effect is the portable component of a worker's wage and as such is a good summary measure of the general skills of a worker (and indeed studies have show that it is highly correlated with direct measures of skills such as education). Using this person effect, we construct summary measures of the skill

[6]In this draft, figures in Tables 4.7 and 4.8 refer to employment only in the three LEHD states. Estimates in the regression models reported in Tables 9, 10, and 11 refer to national employment levels. Subsequent drafts will use the national figures in all tables.

distribution of the firm based upon the fraction of workers the firm has in the quintiles of the person effect distribution (where the quintile thresholds are based upon all workers in the financial services sector).

All of these measures of workforce quality and workforce composition are controls that help in the interpretation of the ILM measures. Our measure of within firm wage dispersion, for example, might be thought to reflect differences in the mix of workers at the firm. Controlling for workforce composition implies we are capturing the variation across firms in our ILM measures holding these composition measures constant. In effect, we are able to examine internal labor market effects controlling for general skills; this is consistent with our theoretical argument that rests on the effects of firm-specific skills and cooperation.

We construct these measures for all establishments, and aggregate them to the firm level for all establishments in the three large states for which we have these measures for the 1992-96 period. While we have constructed these measures for all firms, in what follows much of the analysis focuses on financial services firms that have at least five employees (cumulatively) in our three states. While we have found that our empirical results are robust to the inclusion of all firms, many of our measures (e.g., churning, dispersion) are inherently noisy for very small firms (e.g., a firm with one worker).

In matching our ILM firm-level measures from the LEHD to the LBD diversification measures, we focus on firms that have at least one establishment in our LEHD states. However, it should be emphasized that the diversification measures we use for these firms are the national diversification measures. We are thus using the ILM measures from these three states, derived from the observed dynamics of workers and firms for the period 1992-1996, as proxies for the ILM behavior for the entire national firm.[7]

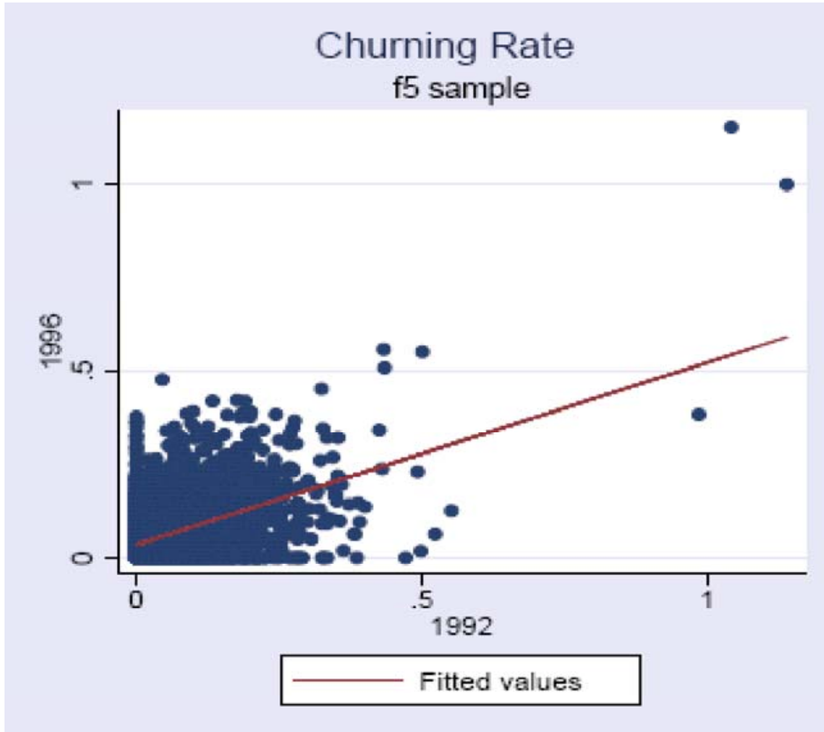


Figure 4.7: Scatter plot of firm-level churning

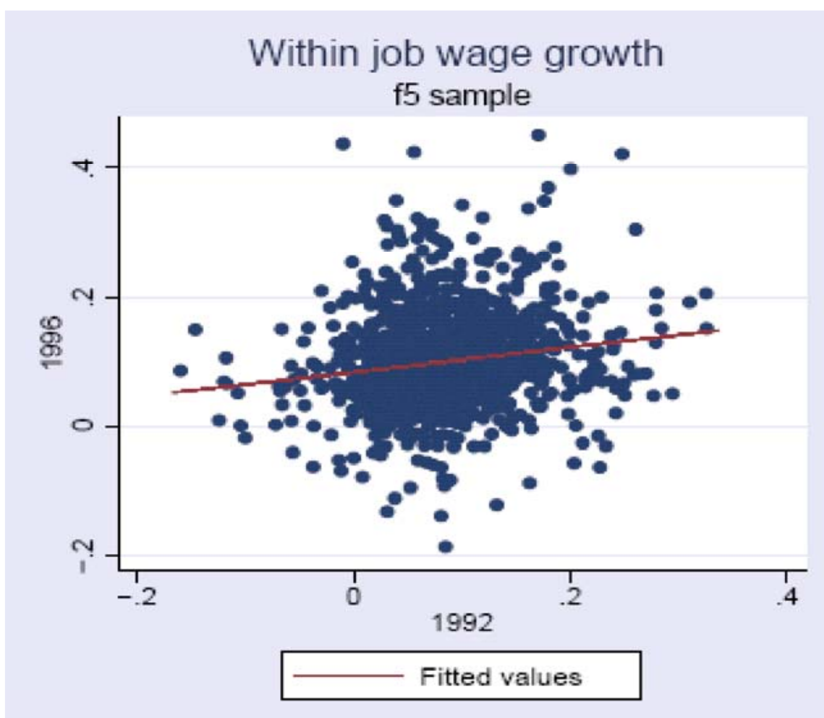


Figure 4.8: Scatter plot of firm-level wage-tenure profiles

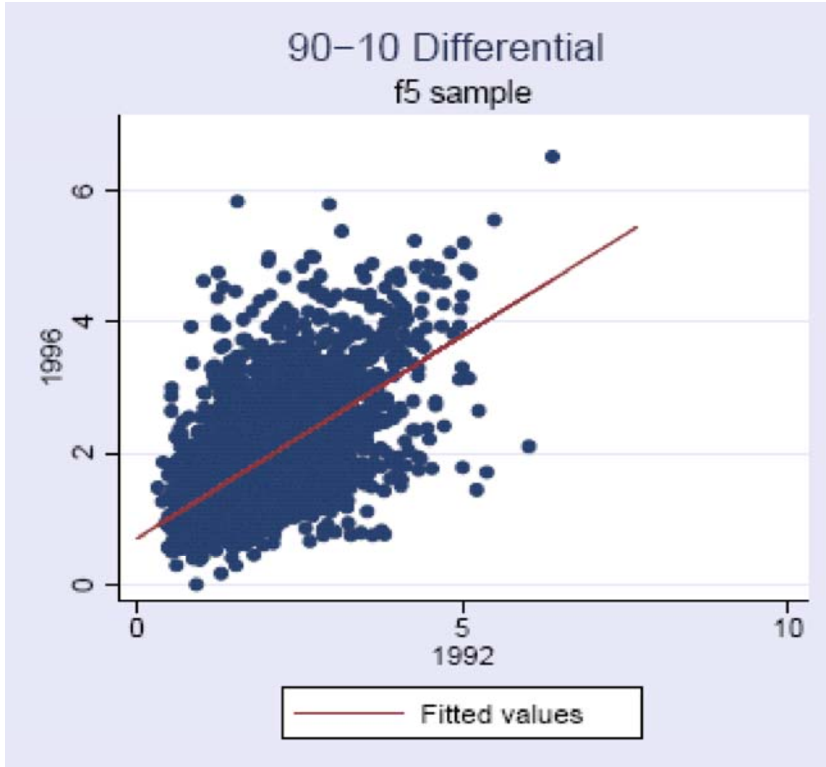


Figure 4.9: Scatter plot of within firm wage dispersion (log difference between 90th and 10th percentile within firm

4.3.2 Basic Facts about ILMs

There is substantial heterogeneity across firms in the measures we have chosen to represent outcomes of ILMs. Figures 4.7, 4.8, and 4.9 reports scatter plots of the churning measure, the within firm wage tenure profiles, and within firm wage dispersion. There is evidence for heterogeneity of each measure. For churning, there is substantial mass from very low rates up to a rate of 0.5 (a fifty percent turnover rate abstracting from net growth is very large). There is substantial mass of wage-tenure profiles from slightly negative to more than 20 percent, and there is substantial mass in the 90-10 log differential from just above zero to more than 400 log points.

We first ask whether our measures are likely to characterize consistent aspects of firm strat-

[7] In future drafts, we plan on examining the behavior of those firms that have most of their activity in our three states as a robustness check. However, even in this case, the national diversification of the firm is clearly the issue of interest.

Independent Variable	Definition
firmage1997	Firm age in 1997
growth	Average net employment growth
size	Average number of full quarter workers
lnsize	Average log number of full quarter workers
shr_fem	Average share of female workers
shr_low	Average share of low human capital workers
shr_high	Average share of high human capital workers
shr_hw	Average share of high wage workers
wjwg	Average within job wage growth (five years) for new hires
chr	Average churning
diff	Average within firm 90-10 log wage differential
Dependent Variable	**Definition**
county_div	Change in geographic diversification at county level
state_div	Change in geographic diversification at state level
ind_div	Change in diversification at industry level
geog_dist_div	Change in geographic diversification at county level, weighted by distance
ind_dist_div	Change in diversification at industry level, weighted by relatedness

Notes:
The independent variables in this table are five-year (1992-1996) averages.
The dependent variables indicate the change in the indices (construction described in the text) from 1997 to 2000.

Table 4.6: Summary of variable definitions

egy and human resource policies. If these indicators vary considerably from year to year, they are likely to be poor representations of internal labor markets or the development of specific skills and capabilities. If, on the other hand, the measures are reasonably stable over the period, then it is more likely that they are capturing some firm-level approach to human resources. The scatter plots demonstrate that there is substantial persistence in each of the measures, suggesting that the variation we have detected reflects part of the firm's long run approach to human resources. Because the variables indicate relatively consistent aspects of firm-level internal labor markets, we reduce the complexity of our analysis by constructing new variables for each indicator by taking the annual mean level of each of our measures for the period 1992-1996. Variable names and their definitions are summarized in Table 4.6.

Statistics	*chr*	*diff*	*shr_hw*	*wjwg*	*shr_low*	*shr_high*	*growth*	*shr_fem*	*size*	*lnsize*
Mean	0.07	1.89	0.35	0.09	0.25	0.26	0.07	0.64	63.99	2.74
Std. Dev.	0.07	0.72	0.22	0.09	0.16	0.16	0.24	0.20	441.93	1.12
N	8,775	8,775	8,775	5,370	7,963	7,857	8,775	8,775	8,775	8,775
Correlation										
chr	**1.00**									
diff	**-0.06**	**1.00**								
share_hw	**-0.11**	**0.04**	**1.00**							
wjwg	**0.05**	**0.07**	**0.09**	**1.00**						
shr_low	**-0.16**	**0.12**	**-0.36**	**-0.09**	**1.00**					
shr_high	**-0.03**	**0.24**	**0.55**	**0.17**	**-0.40**	**1.00**				
growth	**-0.06**	0.01	**0.09**	**0.03**	**-0.04**	**0.13**	**1.00**			
shr_fem	**-0.08**	**-0.14**	**-0.46**	**-0.08**	**0.17**	**-0.42**	**-0.11**	**1.00**		
size	**0.04**	**-0.06**	**0.04**	0.01	**-0.07**	**-0.03**	**-0.03**	**0.00**	**1.00**	
lnsize	**0.17**	**-0.26**	**0.04**	0.00	**-0.26**	**-0.22**	**-0.10**	**0.06**	**0.47**	**1.00**

Note: All variables as defined in Table 6. Correlations in bold are statistically significantly different from zero at p<0.05.

Table 4.7: Unweighted Summary Statistics

Statistics	*chr*	*diff*	*shr_hw*	*wjwg*	*shr_low*	*shr_high*	*growth*	*shr_fem*	*size*	*lnsize*
Mean	0.09	1.60	0.41	0.09	0.18	0.22	0.02	0.64	3115.84	6.37
Std. Dev.	0.05	0.48	0.19	0.04	0.10	0.11	0.16	0.13	4666.29	2.29
N	8,775	8,775	8,775	5,370	7,963	7,857	8,775	8,775	8,775	8,775
Correlation										
chr	**1.00**									
diff	**0.20**	**1.00**								
share_hw	**-0.25**	**-0.09**	**1.00**							
wjwg	**0.07**	**0.20**	**0.05**	**1.00**						
shr_low	**-0.05**	**0.10**	**-0.46**	**-0.15**	**1.00**					
shr_high	**0.07**	**0.42**	**0.52**	**0.27**	**-0.55**	**1.00**				
growth	**-0.16**	0.00	0.02	0.02	**-0.06**	**0.09**	**1.00**			
shr_fem	**-0.08**	**-0.23**	**-0.58**	**-0.11**	**0.29**	**-0.55**	**-0.01**	**1.00**		
size	**-0.02**	**-0.04**	**0.26**	**0.05**	**-0.21**	**0.07**	**-0.07**	**-0.12**	**1.00**	
lnsize	**0.08**	**-0.17**	**0.27**	**0.06**	**-0.29**	**-0.02**	**-0.11**	**-0.09**	**0.74**	**1.00**

Note: All variables as defined in Table 6. Weighted statistics are weighted by average employment from 1992-96. Correlations in bold are statistically significantly different from zero at p<0.05.

Table 4.8: Weighted Summary Statistics

Tables 4.7 and 4.8 report summary statistics for our sample on both weighted and unweighted bases, as well as correlations between the variables. In what follows, we focus on analysis on a weighted basis. For our three key ILM measures, the weighted statistics show an average churning rate of around 9 percent, an average within firm wage tenure profile over the first five years of tenure of 9 log points, and an average within firm wage differential of 160 log points. Consistent with the scatter plots in Figures 4.7 - 4.9, the reported standard deviations show substantial variation. It is also worth noting that the average firm has about 63 workers in the unweighted sample but the employment-weighted average is over 3000. Thus, the average worker in the financial services sector works at a very large firm even though the average firm is relatively small.

Tables 4.7 and 4.8 also show that many of our control variables are strongly correlated. In the analysis to follow, we control only for gender composition and workforce quality, but here we also show the correlations with more basic measures such as the share of high wage workers. The latter is highly correlated with the share of high skill workers and inversely correlated with the share of low skill workers and the share of female workers. These correlations in controls suggest caution in interpreting the effects of any single control variable in subsequent multivariate models.

We also find that several of our control variables are associated with our internal labor market indicators: wage profiles are steeper in firms with lower shares of female workers, and steeper in firms with relatively more workers with high levels of human capital. Wage dispersion is also positively correlated with the share of high-human capital workers. Churning is higher in firms that employ smaller shares of high-human capital workers. Churning is also negatively associated with net growth; that is, growing firms tend to have lower churn rates.

The relationships between the internal labor market variables suggest that the three indicators do not represent a single construct. Churning and wage dispersion are positively correlated; that is, firms with higher turnover rates tend to be those with greater wage dispersion. But firms with relatively steep wage profiles also tend to have slightly higher rates of churn, and considerably more wage dispersion. This is not necessarily surprising: Fairris (2004), for example, shows that quit rates are actually increased by internal opportunities in circumstances where workers compete for such opportunities rather than having them awarded on the basis of seniority. Seniority-based opportunities, on the other hand, may be more characteristic of the ideal-type labor markets described by Doeringer and Piore (1971) featuring lower levels of wage dispersion.

Thus our indicators represent different aspects of internal labor markets, and are not necessarily associated with one another in practice. The data do not support an interpretation in which we would combine these indicators into a single scale representing the overall effects of a strong internal labor market. Instead, we investigate whether each indicator may have its own effects on the development of firm-level resources that lend themselves to leveraging through diversification. We then turn to a consideration of complementarity which focuses on the effects of internal labor markets rather than on the adoption of practices.

4.4 Analysis

Our analytical strategy is to estimate OLS equations with the changes between 1997 and 2000 in the various measures of diversification as dependent variables. As independent variables, we use the constructed five-year means of our internal labor market and control variables. Recall that these means are estimated for the years 1992-1996, fully preceding any subsequent changes in diversification levels. We weight observations by payroll in all estimated models.

	(1) *county_div*	(2) *state_div*	(3) *ind_div*	(4) *geog_dist_div*	(5) *ind_dist_div*
(Constant)	0.132**	0.237**	0.008	0.090**	0.052**
	(0.024)	(0.025)	(0.025)	(0.023)	(0.014)
firmage1997[b]					
1-6 years	-0.026	-0.020	-0.039	-0.010	-0.007
	(0.025)	(0.025)	(0.026)	(0.023)	(0.014)
7-10 years	-0.021	-0.031	-0.038	-0.006	0.011
	(0.020)	(0.020)	(0.020)	(0.018)	(0.011)
11-20 years	0.018*	0.005	0.085**	-0.007	-0.007
	(0.009)	(0.009)	(0.009)	(0.008)	(0.005)
log_size	-0.019**	-0.025**	-0.007**	-0.011**	-0.007**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Growth	0.054*	-0.057*	0.044	0.078**	0.103**
	(0.025)	(0.025)	(0.026)	(0.023)	(0.014)
shr_fem	0.109**	0.143**	0.025	0.099**	0.048**
	(0.024)	(0.024)	(0.024)	(0.022)	(0.014)
shr_high	-0.018	0.019	-0.012	-0.136**	-0.052**
	(0.022)	(0.023)	(0.023)	(0.021)	(0.013)
N	4818	4818	4818	4818	4818
$\mathbf{R^2}$	0.21	0.27	0.14	0.20	0.21

[a] Industry dummies and state dummies are not reported. Standard errors are in parentheses;

[b] over 20 years old is omitted; * $p < 0.05$, ** $p < 0.01$

Further note: For this and subsequent regression models we measure firm size with the overall (national) level of employment in the firm.

Table 4.9: OLS results for diversification measures: control model

Table 4.9 reports results from a control model, before estimation of the variables of interest. (For simplicity we do not report the state and industry dummies.) Examining the overall fit of these models shows clearly that most of the change in diversification levels over the period is attributable to factors we have not measured. The goodness of fit is somewhat similar for each of the measures; cross-industry diversification has a considerably worse fit, and state diversification a somewhat better fit; the other three are roughly similar. The overall explanatory power of the models is reasonable for firm level cross sectional regressions, especially given that the dependent variables are changes in the measure of interest.

As we have noted, there are high correlations amongst the controls so we interpret the results in Table 4.9 with some caution. We chose to operationalize age by dividing the sample into four cohorts; this provided a better fit than a linear specification. "Middle-aged" firms (11-20 years old) are more likely to have positive changes in county-level and industry-level diversification over the period; other effects of age are insignificant. Firm size (as measured by number of workers) is negatively associated with changes in diversification levels by any of our measures. This result may indicate some inertia associated with large firms; it is also true that larger firms are more diversified and thus start from a higher baseline level in 1997 so we may be capturing a catch-up effect.[8]

Results for average growth rate are mixed. Firms that grew extensively in the period 1992-1996 tend to increase their diversification levels as indicated by distance measures, but growth in this earlier period is actually negatively related to the unweighted changes across state boundaries. It is possible that firms that experienced growth during the earlier period may already have begun to diversify before 1997 and that subsequent changes indicate further expansion into the areas that were entered in the earlier period.

Holding other factors fixed, the share of "high human capital" workers is negatively associated with changes in diversification levels for three out of the five measures. This is not inconsistent

[8]We have also estimated models using other specifications for firm size (including groupings and more complex non-linear functions). Some alternatives result in slightly improved fit. The basic findings both with respect to our study variables and with respect to the negative effect of size are robust.

with our theoretical perspective, which suggests that human resources are likely to be leveraged through diversification where they carry firm-specific value that cannot be realized in other ways. This human capital indicator is a measure of general skills, and thus firms may not have opportunities to exploit such skills through diversification. Interestingly, we find that, again holding other factors fixed, a higher share of female workers is associated with a greater increase in diversification. We offer no ready interpretation of this latter finding; we included this measure of workforce composition as a control variable. The impact here, however, is large enough that this finding is worthy of more investigation.[9]

In Table 4.10 we present results for models which include the control variables and our main variables of interest. Results for the internal labor market variables are consistent with our hypothesized relationships, and F-tests suggest that the three indicators contribute to the fit of each of the five models. Results for the control variables are relatively stable in comparison to Table 4.9 (there are some changes in the estimates for age effects, suggesting relationships between internal labor markets and firm age).

Table 4.10 shows that each of the three measures of internal labor market strength is significantly associated with changes in each of the diversification indices, and the relationships are in the expected direction. Churning, our measure of net turnover, is negatively associated with changes in diversification by all five measures; four of these are statistically significant. More extensive wage differentials are also negatively associated with subsequent diversification, and the effects are statistically significant with respect to all five of our diversification measures. And steepness of wage profiles is positively associated with changes in diversification in each of the five models.

In terms of the magnitudes of the effects, the effects we have detected are important but account for a relatively small fraction of the variation in the distribution of changes in diversification. For each of our diversification change measures, a one standard deviation change is about 0.15. The coefficient estimates from Table 4.10 and the summary statistics from Tables 4.7 and 4.8 imply that a one standard deviation change in churning (of about .05) produces a change in

[9]For example, the share of female workers may be a proxy for diversification in activities on other dimensions (e.g., more likely to have part time workers, more likely to be in an urban area).

	(1) *county_div*	(2) *state_div*	(3) *ind_div*	(4) *geog_dist_div*	(5) *ind_dist_div*
(Constant)	0.206**	0.294**	0.074**	0.155**	0.097**
	(0.024)	(0.025)	(0.025)	(0.023)	(0.014)
firmage1997[b]					
1-6 years	-0.054*	-0.032	-0.052*	-0.031	-0.023
	(0.024)	(0.025)	(0.025)	(0.023)	(0.014)
7-10 years	-0.056**	-0.052**	-0.060**	-0.032	-0.009
	(0.019)	(0.020)	(0.020)	(0.018)	(0.011)
11-20 years	0.000	-0.011	0.065**	-0.024**	-0.019**
	(0.008)	(0.009)	(0.009)	(0.008)	(0.005)
log_size	-0.021**	-0.027**	-0.009**	-0.014**	-0.008**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Growth	0.011	-0.091**	0.013	0.045	0.080**
	(0.024)	(0.025)	(0.025)	(0.023)	(0.014)
shr_fem	0.071**	0.105**	-0.028	0.057**	0.020
	(0.023)	(0.024)	(0.024)	(0.022)	(0.014)
shr_high	-0.069**	0.005	0.013	-0.148**	-0.065**
	(0.023)	(0.024)	(0.024)	(0.022)	(0.013)
chr	-0.473**	-0.412**	-0.075	-0.178**	-0.132**
	(0.057)	(0.059)	(0.060)	(0.054)	(0.033)
diff	-0.044**	-0.055**	-0.076**	-0.052**	-0.033**
	(0.006)	(0.007)	(0.007)	(0.006)	(0.004)
wjwg	0.721**	0.345**	0.387**	0.568**	0.425**
	(0.049)	(0.050)	(0.051)	(0.046)	(0.028)
N	4818	4818	4818	4818	4818
$\mathbf{R^2}$	0.26	0.30	0.17	0.24	0.26
Δ $\mathbf{R^2}$	0.05**	0.03**	0.03**	0.04**	0.05**

[a] Industry dummies and state dummies are not reported. Standard errors are in parentheses;
[b] over 20 years old is omitted; ΔR^2 **:** vs. control only model; * $p < 0.05$, ** $p < 0.01$

Table 4.10: OLS results for diversification measures: baseline model

the various dispersion change measures that ranges from about one-seventh of a standard deviation (for county-level geographic diversification) to about four percent of a standard deviation (for distance-weighted industry-level diversification). Analogously, increasing the wage dispersion in a firm by one standard deviation (about .48) results in an increase in diversification growth that is between one-tenth (distance-weighted industry diversification) and one-fourth (unweighted industry diversification) of a standard deviation. Finally, a one standard deviation change in within-job wage growth (about .04) is linked to a change in diversification that ranges between one-fifth (county geographic diversification) and one-tenth (state geographic diversification) of a standard deviation. Again, while much of the variation remains unexplained, accounting for as much as a quarter of the standard deviation of the variation we are seeking to explain suggests that the effects we have captured are important.

One interesting question is the role of distance or relatedness in this context. The pattern of coefficients in Table 4.10 for county-based diversification changes shows that the diversification measure that weights by distance yields slightly smaller effects for wage-tenure profiles and churning than does the measure that does not weight by distance. The effects of wage compression are slightly larger. For the industry based measure, weighting the diversification measure for distance/relatedness yields a larger effect for churning; one that is statistically significant. The effects of wage-tenure profiles are only slightly larger; the effects of wage compression, smaller. Overall, the patterns are similar enough that it is difficult to argue that the results hinge on weighting by distance or relatedness.[10]

The results in Table 4.10 are consistent with our resource-based view of diversification. Firms with low turnover, relative wage compression, and steep within-firm wage profiles are likely to have the sorts of firm-specific resources that can be leveraged through diversification. The results are especially interesting in light of the results in Tables 4.7 and 4.8, which showed that steep wage profiles are actually associated with higher levels of churning and more wage dispersion inside the firm. The results suggest that each of the three measures may indicate the development of firm-specific resources in internal labor markets, though firms that develop resources through worker retention and wage compression may not tend to be the same firms that have steep wage profiles. One possibility is that there may be two kinds of paths to the development of firm-level human resource capabilities: one that focuses on rewarding worker loyalty and cooperation; and a second which focuses on tournament-like structures that encourage worker effort (Lazear and Rosen, 1981).

Next we turn our attention to possible complementarities between internal labor market

[10]In future drafts, we plan to pursue empirical exercises that will help us disentangle the role of diversification and distance/relatedness. The current results should be viewed primarily as a robustness check on alternative measures of diversification.

indicators. Here our analytical strategy follows that of MacDuffie (1995), who argued that multiplicative interactions are one way to estimate complementary effects of different aspects of human resource and production practices. We form interaction terms between each pair of internal labor market indicators. Before doing so, we subtract the mean value of each of the three indicators from each score, "centering" the variables. This procedure reduces multicollinearity without altering the structural relationships among the variables, and allows straightforward interpretations of main effects in the same models that include interactions (Jaccard, Turisi, and Wan 1990).

Results for interactions are displayed in Table 4.11. We estimated both models which added individual interaction terms to each of the models in Table 4.10, and models in which we entered all three terms simultaneously. In Table 4.11 we report only the estimates for the models including all three interaction terms. Results for models which added only one term at a time (available on request) were substantially similar. The effects of churning and wage profiles appear to depend on one another, and the interaction, as expected, is negative and significant for four of the five measures of change diversification (the exception is unweighted industry diversification). Churning and wage compression have complementary effects in the expected direction in three of the five models. The results least consistent with our expectations are those for the interaction between wage profiles and wage compression; while three of the five models suggest complementarity in the expected direction, one does not show statistically significant relationships, and one (for state diversification) has the opposite sign from the one we hypothesized. Overall, the results in Table 4.11 provide some support for our conjecture that internal labor market indicators would have complementary effects on changes in diversification.

4.5 Discussion, Future Directions, and Conclusions

We find that continuing firms in the financial services industries have substantially increased diversification in the latter half of the 1990s. This increase in diversification is on both industrial and geographic dimensions. The increased diversification is not surprising given the changes in the regulations faced by financial services firms on both of these dimensions.

	(1) county_div	(2) state_div	(3) ind_div	(4) geog_dist_div	(5) ind_dist_div
(Constant)	0.203**	0.296**	0.061*	0.141**	0.086**
	(0.024)	(0.025)	(0.025)	(0.023)	(0.014)
firmage1997[b]					
1-6 years	-0.054*	-0.032	-0.048	-0.024	-0.019
	(0.024)	(0.025)	(0.025)	(0.023)	(0.014)
7-10 years	-0.057**	-0.053**	-0.057**	-0.030	-0.007
	(0.019)	(0.020)	(0.020)	(0.018)	(0.011)
11-20 years	-0.000	-0.011	0.067**	-0.020*	-0.017**
	(0.009)	(0.009)	(0.009)	(0.008)	(0.005)
log_size	-0.021**	-0.027**	-0.010**	-0.014**	-0.009**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Growth	0.009	-0.089**	-0.000	0.033	0.070**
	(0.024)	(0.025)	(0.025)	(0.023)	(0.014)
shr_fem	0.076**	0.103**	-0.010	0.073**	0.034*
	(0.023)	(0.024)	(0.024)	(0.022)	(0.014)
shr_high	-0.074**	0.001	0.028	-0.132**	-0.055**
	(0.023)	(0.024)	(0.024)	(0.022)	(0.014)
chr	-0.468**	-0.396**	-0.147*	-0.256**	-0.189**
	(0.059)	(0.061)	(0.062)	(0.056)	(0.034)
diff	-0.044**	-0.056**	-0.074**	-0.053**	-0.033**
	(0.006)	(0.007)	(0.007)	(0.006)	(0.004)
wjwg	0.767**	0.292**	0.598**	0.649**	0.545**
	(0.063)	(0.065)	(0.066)	(0.060)	(0.037)
chr×wjwg	-3.372**	-1.682*	0.472	-1.761*	-1.471**
	(0.788)	(0.816)	(0.821)	(0.745)	(0.455)
chr×diff	0.118	-0.002	0.385**	0.575**	0.408**
	(0.107)	(0.110)	(0.111)	(0.101)	(0.062)
wjwg×diff	0.026	0.133*	-0.346**	-0.078	-0.148**
	(0.061)	(0.063)	(0.063)	(0.057)	(0.035)
N	4818	4818	4818	4818	4818
$\mathbf{R^2}$	0.26	0.30	0.18	0.25	0.27

[a] Industry dummies and state dummies are not reported. Standard errors are in parentheses;
[b] over 20 years old is omitted; * $p < 0.05$, ** $p < 0.01$

Table 4.11: OLS results for diversification measures: interaction model

Our analysis sought to identify characteristics of firms that increased diversification the most. We find substantial heterogeneity in the distribution of changes in diversification, suggesting that features of firms would help to account for this variation. We have hypothesized that internal labor markets help firms to develop resources that can be exploited through diversification, and thus serve as a potential factor that would explain the variation. We test this hypothesis by using outcomes of internal labor market processes for the first half of the 1990s to help predict which firms increased diversification the most in the second half of the 1990s.

We find strong evidence in favor of this resource based view of diversification. Firms with strong internal labor markets as evidenced by steep wage-tenure profiles, low churning of workers and low within firm dispersion of wages increased diversification substantially more than their counterparts without these features. While we explain a relatively small fraction of the overall variation in the distribution of changes in diversification, our results are economically important, robust and statistically significant.

We find mixed evidence with respect to the complementary effects of our three indicators of ILMs. The various aspects of strong internal labor markets do not necessarily "bundle" together to reflect a coherent package or system of practices. Wage profiles, turnover rates, and wage compression exist somewhat independently of one another. Yet our evidence does suggest that the impact of each of these different aspects of internal labor markets on subsequent diversification strategies depends in part on the other aspects; nine of the fifteen interactions we examined empirically were significant and ran in the direction our theory suggested.

The findings suggest investigation of a second set of questions regarding the mode of diversification. Firms may increase their levels of diversification (whether geographic or sectoral) via two processes. They may diversify through acquisition of firms already operating in desired geographies or sectors. Alternatively, they may diversify through greenfield expansion: opening offices and branches in new geographies, or creating service offerings across sectors in which they previously did not. It seems possible that strong ILMs will be especially associated with greenfield expansion. While some existing resources must be deployed in order to integrate and manage acquisitions, more slack resources are required for pure entry into new markets.

Future analyses should also investigate factors that have the potential to moderate the above relationships. We looked at complementarity between aspects of ILMs. But there may be other relationships that also matter. For example, relationships between strong ILMs and diversification may be stronger where human resources are relatively more valuable. Thus ILMs for higher-skilled workers may be more closely associated with diversification strategies and directions than will ILMs for lower-skilled workers. Second, these relationships may vary with firm size. Firms are likely to be heterogeneous with respect not only to human resources but to other kinds of valuable resources. These slack resources may vary directly with firm size: larger firms are more likely to have slack resources that they can exploit through diversification. It is possible that the existence of other slack resources can strengthen the relationships between ILMs and diversification strategy: where firms have valuable, specific human resources and other slack resources, firms are especially likely to choose diversification as a means for appropriating the value of those resources.

Chapter 5

Conclusion

This dissertation investigates the role of various measures of human resource policies of businesses on their performance or behavior using the LEHD database. The lack of proper data has precluded a more thorough investigation of this relationship in the past. I utilize a recently developed employer-employee matched database in various ways to examine the human resource aspects of business. I link these data to detailed business level information such as the LRD and the LBD. Using these integrated data, we can study the relationship among business workforce practices, performance, and behavior.

Chapter 2 focuses on the effects of "learning by doing" on establishment level total factor productivity in U.S. manufacturing. Worker turnover has not been extensively investigated in the existing literature on learning by doing. I argue that firm learning is not a function of past output, but also of worker turnover, as learning is embodied in workers. Using rich information from the LEHD and the LRD, I show that firms with lower rates of worker turnover have higher productivity and "learn" faster than those with higher worker turnover given the same amount of past production experience. Moreover, I develop new instruments, based on local downstream demand, firm wage policy, regional hiring conditions of related industries, firm vintage, and firm location, that significantly affect past production experience and worker turnover but are not necessarily affected by firm productivity shocks. With these instruments, I show that learning by doing and turnover have causal effects on productivity.

In Chapter 3, I examine the relationship between firm performance with workforce quality and worker turnover in five selected industries utilizing the information from the LEHD and the Economic Censuses. We find that measures of productivity, workforce quality, and worker turnover are highly correlated across businesses. Moreover, we find that firms with high workforce quality and low worker turnover are more likely to survive even after controlling for productivity. While there are some common patterns, there are also different idiosyncratic aspects in different industries. We find that the different industries under investigation have different relationships between these measures. However, the common patterns are more striking than the idiosyncratic factors. Workforce quality, worker churning, and firm performance are related across all the industries studied and virtually all classifications of businesses that we have considered.

Chapter 4 investigates the fit between firm-level internal labor markets (ILMs) and firm diversification in the U.S. financial services sector. We find that continuing firms in the financial services industries substantially increased diversification in the latter half of the 1990s. The increase in diversification was expected given the changes in the regulations faced by financial services firms. Our analysis seeks to identify the characteristics of firms that increase diversification the most. We hypothesize that internal labor markets help firms to develop resources that can be exploited through diversification, and thus serve as a potential factor that would explain the variation. We find strong evidence in favor of this resource based view of diversification. Firms with strong internal labor markets, as evidenced by steep wage-tenure profiles, low worker churn, and low within firm wage dispersion, increased diversification substantially more than their counterparts without these features.

This dissertation has focused on the relationship between the human resource policies of businesses and their performance. In three chapters, we have shown that human resource practices are critical in understanding firm behavior and performance. These are only a few ways of looking inside businesses, and there are many other ways to investigate this relationship. More importantly, fundamental questions such as why firms have different human resource management systems have not been answered in this dissertation, and future research should try to address these questions.

Appendix A

Productivity Measure in Chapter 2

The index number method using cost-shares does not involve any regressions. Based on the assumption of constant returns to scale and Shephard's lemma, one can calculate elasticities of inputs using cost shares. Cost shares are measured at the four-digit industry level and equal the average of the current and previous year's shares. To do this, I first calculate four-digit industry level total costs as follows:

$$TC_{k,t} = pe_{k,t} \times KE_{k,t} + ps_{k,t} \times KS_{k,t} + PAY_{k,t} + MATCOST_{k,t} + ENERGY_{k,t}$$

where $TC_{k,t}$ is the total cost for industry k in year t, $pe_{k,t}$ is the rental price of equipment for industry k in year t, $KE_{k,t}$ its equipment stock, $ps_{k,t}$ the rental price of structures, $KS_{k,t}$ its structures stock, $PAY_{k,t}$ its total payroll, $MATCOST_{k,t}$ its total materials costs except for energy, and $MATCOST_{k,t}$ its total energy costs.[1]. Then the cost share of, say, the equipment stock is calculated as

$$\alpha ke_{k,t} = \left(\frac{pe_{k,t} \times KE_{k,t}}{TC_{k,t}} + \frac{pe_{k,t-1} \times KE_{k,t-1}}{TC_{k,t-1}} \right) \times 0.5$$

where $\alpha ke_{k,t}$ is industry k's equipment share in year t. Cost shares for structures($\alpha ks_{k,t}$), total hours worked($\alpha l_{k,t}$), materials($\alpha m_{k,t}$), and energy($\alpha en_{k,t}$) are calculated in the same manner. With cost shares calculated, total factor productivity (in logs) is obtained by

$$tfp_{j,k,t} = q_{j,t} - \alpha ke_{k,t} \times ke_{j,t} - \alpha ks_{k,t} \times ks_{j,t} - \alpha l_{k,t} \times l_{j,t} - \alpha m_{k,t} \times m_{j,t} - \alpha en_{k,t} \times en_{j,t}$$

[1]BLS has two digit industry level data on the following variables: Capital income ($EQKY$ and $STKY$), the real productive capital stock ($EQPK$ and $STPK$), and the ratio of capital input to the productive capital stock ($EQKC$ and $STKC$, 1987=100). I calculate rental prices pe and ps as

$$\begin{aligned} pe &= \frac{EQKY}{EQPK \times EQKC} \times 100 \\ ps &= \frac{STKY}{STPK \times STKC} \times 100. \end{aligned}$$

Industry level input variables are in the NBER-CES Manufacturing Industry Database constructed by Bartelsman, Becker, and Gray.

where $ke_{j,t}$ is the log of equipment capital stock for plant j in year t, $ks_{j,t}$ its log of structures, $l_{j,t}$ its log of total hours worked, $m_{j,t}$ its materials input, and $en_{j,t}$ its energy input.

Appendix B

Estimation of Learning in Chapter 2

In estimating learning by doing effects on firm productivity, I adopt a two step approach where I derive the TFP measure in the first step and estimate learning by doing effects on TFP in the second step. Moreover, I follow the index number approach to get the TFP measure instead of using econometric methods. The index number approach depends on two assumptions: constant returns to scale production technology with respect to standard input variables (such as capital, labor, and materials) and profit maximization by the firm.

Suppose a firm is naive or a static profit optimizer who does not take into account the benefit of extra output of current period in the following periods. Given the production technology as in equation (2.4)[1], the firm's optimization problem is as follows:

$$\max_{K_t, L_t} \; pAK_t^{\alpha_k} L_t^{\alpha_l} E_t^{\alpha_e} - rK_t - wL_t \tag{B.1}$$

where p is output price, K_t real capital stock at the end of year t, L_t its total hours worked, E_t the index of "learning", r capital rental price, and w wage rate. The first order conditions are

$$\alpha_k pAK_t^{\alpha_k - 1} L_t^{\alpha_l} E_t^{\alpha_e} = r \tag{B.2}$$

$$\alpha_l pAK_t^{\alpha_k} L_t^{\alpha_l - 1} E_t^{\alpha_e} = w \tag{B.3}$$

Combining equations (B.2) and (B.3), one can get

$$(\alpha_k + \alpha_l)\, pAK_t^{\alpha_k} L_t^{\alpha_l} E_t^{\alpha_e} = rK_t + wL_t \tag{B.4}$$

From equations (B.2), (B.3), and (B.4), we have

$$\frac{\alpha_k}{\alpha_k + \alpha_l} = \frac{rK_t}{rK_t + wL_t} \tag{B.5}$$

$$\frac{\alpha_l}{\alpha_k + \alpha_l} = \frac{wL_t}{rK_t + wL_t} \tag{B.6}$$

[1]For simplicity, I assume that production funtion has two inputs, capital and labor.

Under constant returns to scale ($\alpha_k + \alpha_l = 1$), input elasticities and cost shares are equal to each other.

What if a firm is a dynamic optimizer who takes into consideration the future benefit of current output through learning by doing? The dynamic programming which a firm faces is

$$V(E_t) = \max_{K_t, L_t} pAK_t^{\alpha_k} L_t^{\alpha_l} E_t^{\alpha_e} - rK_t - wL_t + \beta V(E_{t+1}) \tag{B.7}$$

$$s.t. \;\; E_{t+1} = E_t + AK_t^{\alpha_k} L_t^{\alpha_l} E_t^{\alpha_e} \tag{B.8}$$

The first order conditions are (after simple algebra)

$$(1 + \beta \frac{\partial V}{\partial E_{t+1}}) \alpha_k pAK_t^{\alpha_k} L_t^{\alpha_l} E_t^{\alpha_e} = rK_t \tag{B.9}$$

$$(1 + \beta \frac{\partial V}{\partial E_{t+1}}) \alpha_l pAK_t^{\alpha_k} L_t^{\alpha_l} E_t^{\alpha_e} = wL_t \tag{B.10}$$

Adding equations (B.9) and (B.10), one gets

$$(1 + \beta \frac{\partial V}{\partial E_{t+1}})(\alpha_k + \alpha_l) pAK_t^{\alpha_k} L_t^{\alpha_l} E_t^{\alpha_e} = rK_t + wL_t \tag{B.11}$$

From equations (B.9), (B.10), and (B.11), we have

$$\frac{\alpha_k}{\alpha_k + \alpha_l} = \frac{rK_t}{rK_t + wL_t} \tag{B.12}$$

$$\frac{\alpha_l}{\alpha_k + \alpha_l} = \frac{wL_t}{rK_t + wL_t} \tag{B.13}$$

They are exactly the same as equations (B.5) and (B.6) in "static" case. Therefore, under constant returns to scale technology and profit maximization (regardless of "static" or "dynamic"), index number approach based on cost shares will give us reasonable estimates for input elasticities and thus TFP measure.

Now assume that we have the following data generating process:

$$y = X_1\beta_1 + X_2\beta_2 + u \tag{B.14}$$

where X_1 includes capital stock, labor, and materials and X_2 represents learning, workforce quality, and other control variables. Assuming that both X_1 and X_2 are orthogonal to disturbance term

u,[2] OLS will generate consistent estimates. Then the usual partialing out arguments imply

$$\hat{\beta}_1 = (X_1' M_{X_2} X_1)^{-1} X_1' M_{X_2} y \quad \text{(B.15)}$$

$$\hat{\beta}_2 = (X_2' M_{X_1} X_2)^{-1} X_2' M_{X_1} y \quad \text{(B.16)}$$

where $M_x = I - X(X'X)^{-1} X'$. Under cost minimization and constant returns to scale with respect to capital, labor, and materials, elasticities corresponding to them are equal to their cost shares where total cost is the sum of capital rental cost, payroll, and the materials cost. Therefore, we can use cost shares as consistent estimates of elasticities. In terms of equation $(B.14)$, we have consistent estimates for β_1, expressed by

$$\overline{\beta}_1 = \beta_1 + \nu \quad \text{(B.17)}$$

where the probability limit of ν is zero. From equation $(B.14)$ and equation $(B.17)$, the second step estimation equation is

$$\overline{y} = X_2 \beta_2 + \epsilon \quad \text{(B.18)}$$

where $\overline{y} = y - X_1 \overline{\beta}_1$ and $\epsilon = u - X_1 \nu$. The estimate of β_2 from equation $(B.18)$ is

$$\overline{\beta}_2 = \beta_2 + (X_2' X_2)^{-1} X_2' u - (X_2' X_2)^{-1} X_2' X_1 \nu. \quad \text{(B.19)}$$

It is easy to show that the probability limit of $(\overline{\beta}_2 - \beta_2)$ is zero. Therefore, the two step approach, which uses the index number method to get TFP in the first step, can consistently estimate the impact of learning by doing on productivity.

[2]This is not a realistic assumption. But it is fine for the discussion in this section.

Appendix C

Sample Construction in Chapter 2

The units of observation in the two data sets (the LRD and the LEHD database) are not identical. While the LRD is at the establishment level (Permanent Plant Number, or PPN), the business level identifiers on UI files are State Employer Identification Numbers (SEINs), which do not necessarily match the establishment level identifiers. Although one can impute establishment level data, the establishment level identifiers (SEINUNITs) are still different from those in the LRD. To solve this problem, we need a common identifier for both data sources. Fortunately, there are two supplementary sources to enable successful matches. One is the Census Bureau's Business Register, previously known as the Standard Statistical Establishment List (SSEL). The other is ES202 data, available from each state. The variables I use are the (Federal) Employer Identification Number (EIN)[1] as well as the county and state codes that identify geographic information. Since the LRD and the Business Register have common establishment level identifiers (Census File Number, or CFN), one can match LRD to the Business Register by CFN to get information on the EIN and geographic codes. The UI data and the ES202 data also share the state level employer identifiers (SEINs and SEINUNITs), so one can match the UI data with the ES202 data using SEIN to get information on the EIN, county, and state.

Some caution is warranted when the EIN is used as a business identifier. The EIN is a unique identifier for single unit establishments. However, when more than one establishment are under common ownership (multi-units), then those establishments may have the same EIN. In any case, the EIN is an identifier for potentially more aggregated business units than establishments. This is also the case for LEHD data sources. Hence, we have to deal with at least some level of aggregation before matching. I use the EIN/state/county combination for the level of aggregation when using data from the LEHD. However, establiments under the same EIN/state/county

[1] An EIN entity is an administrative unit that the IRS has assigned a unique identifier for use in tax reporting.

combination may have different birth years. Since my analysis relies on measuring the history of establishments, working at this aggregate level is unsatisfactory, when studying productivity in the LRD. Instead of aggregating information of the LRD up to this EIN/state/county level, I choose to work at the establishment level of the LRD. To work at this level, however, I need somewhat restrictive assumption on workforce flows of establishments within the same EIN/state/county. I assume that all establishments sharing a common EIN/state/county have the same workforce characteristics and worker flows. Under this uniformity assumption, every establishment under the same EIN/state/county has the same level of workforce quality, the same worker turnover rates, etc. This way, we can track an establishment's history of activity exactly and roughly follow its pattern on workforce characteristics and worker flows.

Appendix D

Turnover Measures in Chapter 2

This section describes concepts used to construct the turnover measure. Worker turnover rates are derived from the UI data of the LEHD. While the LRD variables are annual, the UI earnings history data is quarterly. Hence, I have to develop measures comparable to the annual LRD variables. A worker i is regarded as "employed" at firm j in year t if he has a valid UI wage record for at least two consecutive quarters in year t. Restrictions on consecutive quarters are made so that worker flows are not heavily affected by those with very short spell jobs. A separation of worker i from firm j occurs in year t when he is not "employed" at firm j in year t while he was "employed" at firm j in year $t-1$. At the firm level, employment at firm j at year t is the total number of workers who are "employed" at firm j in year t. Likewise, separation at firm j at year t is the total numbers of workers who separate from firm j in year t. Formal definitions are described below.

D.1 Individual concepts

D.1.1 Quarterly variables

In the following, t refers to the sequential quarter.

Flow employment (m): individual i employed (matched to a job) at some time during period t at employer j

$$m_{i,j,t} = \begin{cases} 1, & \text{if } i \text{ has positive earnings at employer } j \text{ during period } t \\ 0, & \text{otherwise} \end{cases} \tag{D.1}$$

Beginning of quarter employment (b): individual i employed at the end of $t-1$, beginning of t

$$b_{i,j,t} = \begin{cases} 1, & \text{if } m_{i,j,t-1} = m_{i,j,t} = 1 \\ 0, & \text{otherwise} \end{cases} \tag{D.2}$$

D.1.2 Annual variables

In the following, t refers to the year and q refers to the quarter. Let $b_{i,j,t,q}$ refer to beginning-of-quarter employment status. Then there is a one-to-one mapping between this definition and the one defined in equation $(D.2)$ using sequential quarters.

Flow employment (emp): individual i employed at some time during year t at employer j

$$emp_{i,j,t} = \begin{cases} 1, & \text{if } b_{i,j,t,2} = 1 \text{ or } b_{i,j,t,3} = 1 \text{ or } b_{i,j,t,4} = 1 \\ 0, & \text{otherwise} \end{cases} \tag{D.3}$$

This definition of annual employment requires some attachment of employer-employee relationship. More specifically, it requires at least two consecutive positive earnings record during a year. Given our emphasis on the learning by doing process, this attachment requirement should not be a very restrictive condition.

Separations (s): individual i separated from employer j during year t

$$s_{i,j,t} = \begin{cases} 1, & \text{if } emp_{i,j,t-1} = 1 \text{ and } emp_{i,j,t} = 0 \\ 0, & \text{otherwise} \end{cases} \tag{D.4}$$

D.2 Employer concepts

Annual employment (number of jobs) for employer j during year t

$$EMP_{j,t} = \sum_i emp_{i,j,t} \tag{D.5}$$

Annual separations for employer j during year t

$$S_{j,t} = \sum_i s_{i,j,t} \tag{D.6}$$

Annual separation rate for employer j during year t

$$SR_{j,t} = \frac{S_{j,t}}{EMP_{j,t-1} + EMP_{j,t}} \tag{D.7}$$

Appendix E

Instruments for Turnover Measures in Chapter 2

In the IV estimation, I use the weighted job-to-job hiring rate and the weighted overall hiring rate as instruments. For example, suppose workers who leave industry A either find no jobs (unemployed) or jobs in industry B or industry C. Over 1996-2000, suppose the average numbers of workers who move from industry A to industries B and C are 1,000 and 2,000, respectively. In year 2000, suppose industry B has employment of 20,000 workers, 4,000 new hires, and 2,000 job switchers among new hires. Industry C has employment of 20,000 workers, 5,000 new hires, and 3,000 job switchers among new hires. This hiring rate is defined as the ratio of new hires to average employment. Therefore, hiring rates are 20% for industry B and 25% for industry C. The job-to-job hiring rate is defined as the ratio of job switchers among new hires to average employment. Job-to-job hiring rates are 10% for industry B and 15% for industry C. Industry A's weighted hiring rate is $23\%(0.33 \times 0.2 + 0.67 \times 0.25)$ and its weighted job-to-job hiring rate is $13\%(0.33 \times 0.1 + 0.67 \times 0.15)$.

BIBLIOGRAPHY

[1] Aaronson, Daniel and Daniel Sullivan, "Growth in Worker Quality," *Economic Perspectives*, Fourth Quarter 2001, 25, 53-74.

[2] Abowd, John M., Robert H. Creecy and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Cornell University working paper, 2002.

[3] Abowd, John M., John Haltiwanger, Julia Lane, Kevin McKinney and Kristin Sandusky, "Within and Between Firm Changes in Human Capital, Technology, and Productivity," Working paper, 2004.

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

[5] Abowd, John M., Paul Lengermann, and Kevin McKinney, "The Measurement of Human Capital in the U.S. Economy," U.S. Census Bureau preliminary working paper, 2002.

[6] Abowd, John M., John Haltiwanger, Ron Jarmin, Julia Lane, Paul Lengermann, Kristin McCue, Kevin McKinney, and Kristin Sandusky, "The Relationship Between Human Capital, Productivity and Market Value: Building Up form Micro Evidence," U.S. Census Bureau preliminary working paper, 2002.

[7] Amit, R. and J. H. P. Schoemaker, "Strategic Assets and Organizational Rent," *Strategic Management Journal*, January 1993, 14, 33-46.

[8] Argote, Linda, Sara L. Beckman, and Dennis Epple, "The Persistence and Transfer of Learning in Industrial Settings," *Management Science*, February 1990, 36, 140-54.

[9] Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *Review of Economic Studies*, June 1962, 29, 155-173.

[10] Aw, Bee Yan, Xiaomin Chen, and Mark Roberts, "Firm-Level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing," *Journal of Development Economics*, October 2001, 66, 51-86.

[11] Bahk, Byong-Hyong and Michael Gort, "Decomposing Learning by Doing in New Plants," *Journal of Political Economy*, August 1993, 101, 561-83.

[12] Baily, Martin Neil, Eric J. Bartelsman and John Haltiwanger, " Downsizing and Productivity Growth: Myth or Reality?" *Small Business Economics*, August 1996, 8, 259-278.

[13] Baily, Martin Neil, Eric J. Bartelsman and John Haltiwanger, "Labor Productivity: Structural Change and Cyclical Dynamics," NBER Working Paper No. 5503, 1997.

[14] Baily, Martin N., Charles Hulten, and David Campbell, "Productivity Dynamics in Manufacturing Plants," *Brooking Papers on Economic Activity, Microeconomics*, 1992, 187-249.

[15] Barney, Jay B., "Strategic Factor Markets: Expectations, Luck, and Business Strategy," *Management Science*, October 1986, 32, 1231-41.

[16] Barney, Jay B., "Firm Resources and Sustained Competitive Advantage," *Journal of Management*, March 1991, 17, 99-120.

[17] Barney, Jay B., Mike Wright, and David J. Ketchen, Jr., "The Resource-Based View of the Firm: Ten Years After 1991," *Journal of Management*, 2001, 27 (6), 625-641.

[18] Bartelsman, Eric J., and Phoebus J. Dhrymes, "Productivity Dynamics: U.S. Manufacturing Plants, 1972-86," *Journal of Productivity Analysis*, January 1998, 9, 5-34.

[19] Bartelsman, Eric J. and Mark Doms, "Understanding Productivity: Lessons from Longitudinal Microdata," *Journal of Economic Literature*, September 2000, 38, 569-594.

[20] Bartelsman, Eric J. and Wayne Gray, "The NBER Manufacturing Productivity Database," NBER Technical Working Paper No. 205, 1996.

[21] Batt, Rosemary, Larry Hunter, and Stefannie Wilk, "How and When Does Management Matter? Job Quality and Career Opportunities for Call Center Workers," in *The Future of Low Wage Work,* editors: Eileen Appelbaum, Annette Bernhardt, and Richard J. Murnane, New York: Russell Sage Foundation, 270-315.

[22] Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education,* 1993, Third edition, Chicago and London, University of Chicago Press.

[23] Becker, Randy, John Haltiwanger, Ron Jarmin, Shawn Klimek, and Dan Wilson, "Micro and Macro Data Integration: The Case of Capital," Working Paper, 2004.

[24] Benkard, C. Lanier, "Learning and Forgetting: The Dynamics of Aircraft Production," American Economic Review, September 2000, 90, 1034-54.

[25] Berger, Allen N., "The Economic Effects of Technological Progress: Evidence from the Banking Industry." Federal Reserve Board Finance and Economic Discussion Series, 2002-50, 2002.

[26] Berger, Allen N., Rebecca S. Demsetz, and Philip E. Strahan, "The consolidation of the financial services industry: Causes, consequences, and implications for the future." Journal of Banking and Finance, February 1999, 23, 135-194.

[27] Berger, Allen N. and Loretta J. Mester, "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions?1997" *Journal of Banking and Finance*, July 1997, 21, 895-947.

[28] Berger, Allen N. and Loretta J. Mester, "What Explains the Dramatic Changes in Cost and Profit Performance of the U.S. Banking industry?" Wharton Financial Institutions Center Working Paper, 99-10, 1999.

[29] Betancourt, Roger, and David Gautschi, "The Outputs of Retail Activities: Concepts, Measurement and Evidence from U.S. Census Data," *Review of Economics and Statistics*, May 1993, 75, 294-301.

[30] Black, Sandra and Lisa M. Lynch, "Human Capital Investments and Productivity," *American Economic Review*, May 1996, 86, 263-267.

[31] Black, Sandra and Lisa M. Lynch, "The New Workplace: What Does it Mean for Employers?" Industrial Relations Research Association Papers and Proceedings, 1998, 60-67.

[32] Black, Sandra E. and Lisa M. Lynch, "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity." *Review of Economics and Statistics*, August 2001, 83, 434-445.

[33] Black, Sandra E. and Lisa M. Lynch, "What's Driving the New Economy? The Benefits of Workplace Innovation." NBER Working Paper No. w7479, 2002.

[34] Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt, "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence." NBER Working Paper No. 7136, 1999.

[35] Burdett, Kenneth and Dale T. Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, May 1998, 39, 257-73.

[36] Burgess, Simon, Julia Lane, and David Stevens, "Job Flows, Worker Flows and Churning," *Journal of Labor Economics*, July 2000, 18, 473-502.

[37] Caballero, Ricardo, and Mohamad Hammour, "The Cleansing Effect of Recessions," *American Economic Review*, December 1994, 84, 1350-1368.

[38] Caballero, Ricardo J and Mohamad L. Hammour, "On the Timing and Efficiency of Creative Destruction," *The Quarterly Journal of Economics*, August 1996, 111, 805-852.

[39] Campbell, Ben, Hyowook Chiang, John Haltiwanger, Larry W. Hunter, Ron Jarmin, Nicole Nestoriak, Tim Park, and Kristin Sandusky, "Firm Performance, Workforce Quality and Workforce Churning," Working Paper, 2004.

[40] Cappelli, Peter, "Why Do Employers Retain At-Risk Workers? The Role of Social Capital," *Industrial Relations*, April 2004, 43, 421-447.

[41] Caves, Richard E., "Industrial Organization and New Findings on the Turnover and Mobility of Firms," *Journal of Economic Literature*, December 1998, 36, 1947-1982.

[42] Chandler, Alfred. D., *Scale and Scope : The Dynamics of Industrial Capitalism*, Cambridge, MA, Harvard University Press, 1990.

[43] Chatterjee, Sayan and Birger Wernerfelt, "The Link Between Resources and Type of Diversification: Theory and Evidence," *Strategic Management Journal*, January 1991, 12, 33-48.

[44] Chiang, Hyowook, Cheryl Grim, John C. Haltiwanger, Larry W. Hunter, Ron Jarmin, Nicole Nestoriak, Kristin Sandusky, and Jeongil Seo, "Internal Labor Markets and Diversification Strategies in Financial Services, Working Paper, 2004.

[45] Corrado, Carol and Lawrence Slifman, "Decomposition of Productivity and Costs," *American Economic Review*, May 1999, 89, 328-32.

[46] Davis, Steven J., and John Haltiwanger, "Gross Job Flows," in *Handbook of Labor Economics*, Volume 3B, editors: Orley Ashenfelter and David Card, 1999, New York, Elsevier Science.

[47] Davis, Steven J., John C. Haltiwanger, and Scott Schuh, *Job Creation and Destruction*, 1996, Cambridge, MIT Press.

[48] Davis, Steven J., John Haltiwanger, Ron Jarmin, Javier Miranda, Alfred Nucci, and Kristin Sandusky, "The Dynamics of Young Businesses," Working Paper, 2004.

[49] Dierickx, Ingermar and Karel Cool, "Asset Stock Accumulation and Sustainability of Competitive Advantage," *Management Science*, December 1989, 35, 1504-1511.

[50] Disney, Richard, Jonathan Haskel, and Yvla Heden, "Restructuring and Productivity Growth in U.K. Manufacturing," mimeo. 2001.

[51] Doeringer, Peter, and Michael J. Piore, *Internal Labor Markets and Manpower Analysis*, Lexington, Mass., D.C. Heath, 1971.

[52] Doms, Mark, Ron Jarmin, and Shawn Klimek, "IT Investment and Firm Performance in Retail Trade," mimeo, 2001.

[53] Dwyer, Douglas, "Productivity Races I: Are Some Productivity Measures Better Than Others?" Center for Economic Studies Working Paper, CES 97-2, 1997.

[54] Dwyer, Douglas, "Technology Locks, Creative Destruction, and Non-Convergence in Productivity Levels" *Review of Economic Dynamics*, April 1998, 1, 430-73.

[55] Dumas, Mark W, "Productivity Trends in Two Retail Trade Industries, 1987-95." *Monthly Labor Review*, July 1997, 35-39.

[56] Edwards, C. D., "Conglomerate Bigness as a Source of Power," in *Business Concentration and Price Policy*, Princeton, Princeton University Press, 1955.

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

[58] Fairris, David, "Internal Labor Markets and Worker Quits," *Industrial Relations*, July 2004, 43, 573-594.

[59] Farjoun, Moshe, "Beyond Industry Boundaries: Human Expertise, Diversification and Resource-related Industry Groups," *Organization Science*, May 1994, 5, 185-199.

[60] Fay, S, "Recent Activity Resulting from the Gramm-Leach-Bliley Act," *New England Banking Trends*, Summer 2000, 24, 1-8.

[61] Foster, Lucia, John Haltiwanger, and C.J. Krizan, "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," in *New Developments in Productivity Analysis*, editors: Edward Dean, Michael Harper and Charles Hulten, Chicago: University of Chicago Press, 2001

[62] Foster, Lucia, John Haltiwanger, and C.J. Krizan, "Market Selection, Restructuring and Reallocation in the Retail Trade Sector in the 1990s," mimeo, 2004.

[63] Freeman, Rechard B., "Union Wage Practices and Wage Dispersion within Establishments," *Industrial and Labor Relations Review*, October 1982, 36, 3-21.

[64] Frei, Frances X., Patrick T. Harker, and Larry W. Hunter, "Inside the Black Box: What Makes a Bank Efficient?" *Performance of Financial Institutions*, editors: Patrick T. Harker and Stavros A. Zenios, 1999, Cambridge, Cambridge University Press.

[65] Gande, Amar, Manju Puri, and Anthony Saunders, "Bank Entry, Competition, and the Market for Corporate Securities Underwriting," *Journal of Financial Economics*, October 1999, 54, 165-195.

[66] Gant, John, Casey Ichniowski, and Kathryn Shaw, "Social Capital and Organizational Change in High-Involvement and Traditional Work Organizations," *Journal of Economics and Management Strategy*, Summer 2002, 11, 289-328.

[67] Gollop, Frank and James Monahan, "A Generalized Index of Diversification: Trends in U.S. Manufacturing," *Review of Economics and Statistics*, May 1991, 73, 318-330.

[68] Griliches, Zvi, and Haim Regev, "Firm Productivity in Israeli Industry: 1979-1988," *Journal of Econometrics*, January 1995, 65, 175-203.

[69] Gullickson, William and Michael J. Harper, "Bias in aggregate productivity trends revisited," *Monthly Labor Review*, March 2002, 125, 32-40.

[70] Haltiwanger, John C., "Measuring and Analyzing Aggregate Fluctuations: The Importance of Building from Microeconomic Evidence," *Federal Reserve Bank of St. Louis Review*, May/June 1997, 55-77.

[71] Haltiwanger, John C., and Ron Jarmin, "Measuring the Digital Economy." *Understanding the Digital Economy*, editors: E. Byrnjolfsson and B. Kahin, 2000, Cambridge, MIT Press.

[72] Haltiwanger, John C., Julia I. Lane, and James R. Spletzer, "Wages, Productivity, and the Dynamic Interaction of Business and Workers," NBER Working Paper No. w7994, 2000.

[73] Haltiwanger, John C., Julia I. Lane, and James R. Spletzer, "Productivity Differences Across Employers: The Role of Employer Size, Age, and Human Capital" Working Paper, 2001.

[74] Hitt, Michael A., Leonard Bierman, Katshuhiko Shimizu, and Rahul Kochhar, "Direct and Moderating Effects of Human Capital on Strategy and Performance in Professional Service Firms: A Resources-Based Perspective," *Academy of Management Journal*, February 2001, 44, 13-28.

[75] Holmes, Thomas J. and James A. Schmitz, Jr., "A Theory of Entrepreneurship and Its Application to the Study of Business Transfers," *The Journal of Political Economy*, April 1990, 98, 265-294.

[76] Holmes, Thomas J. and James A. Schmitz, Jr., "On the Turnover of Business Firms and Business Managers," *The Journal of Political Economy*, October 1995, 103, 1005-1038.

[77] Hopenhayn, Hugo A. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, September 1992, 60, 1127-1150.

[78] Hoskisson, Robert E. and Michael A. Hitt, "Antecedents and Performance Outcomes of Diversification: A Review and Critique of Theoretical Perspectives," *Journal of Management*, June 1990, 16, 461-509.

[79] Ichniowski, Casey and Kathryn Shaw, "The Effects of Human Resource Systems on Productivity: An International Comparison of U.S. and Japanese Plants," *Management Science*, 1999, 45, 704-722.

[80] Ichniowski, Casey and Kathryn Shaw, "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices," *Journal of Economic Perspectives*, Winter 2003, 17, 155-178.

[81] Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi "The Effects of Human Resource Management Practices on Productivity," *American Economic Review*, June 1997, 86, 291-

313. Reprinted in *Personnel Economics*, editors: Edward P. Lazear and Robert McNabb, Cheltenham, UK: Edward Elgar Publishing, Ltd., (forthcoming).

[82] Jaccard, James, Robert Turrisi, and Choi K. Wan, *Interaction Effects in Multiple Regression*, Newbury Park, CA, Sage, 1990.

[83] Jarmin, Ronald, Shawn Klimek, and Javier Miranda, "Firm Entry and Exit in the U.S. Retail Sector, 1977-1997," mimeo, 2001.

[84] Jarmin, Ron S. and Javier Miranda, "The Longitudinal Business Database," Center for Economic Studies, U.S. Bureau of the Census, Working Paper, Working Paper No.CES-WP-02-17, 2002.

[85] Jensen, Michael C., "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *American Economic Review*, May 1986, 76, 323-29.

[86] Jorgenson, Dale W., Frank M. Gollop, and Barbara M. Fraumeni, *Productivity and US Economic Growth*, 1987, Harvard University Press, Cambridge, MA.

[87] Jorgenson, Dale W. and Kevin J. Stiroh, "Raising the speed limit: US economic growth in the information age," OECD Economics Department Working Paper 261, OECD Economics Department, 2000.

[88] Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, May 1982, 50, 649-70.

[89] Kogut, Bruce and Udo Zander, "Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology," *Organization Science*, August 1992, 3, 383-397.

[90] Kroszner, Randall S. and Philip E. Strahan, "What drives deregulation? Economics and politics of the relaxation of bank branching restrictions," *Quarterly Journal of Economics*, November 1999, 114, 1437-1467.

[91] Lazear, Edward and Sherwin Rosen, "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, October 1981, 89, 841-864.

[92] Lei, David, Michael A. Hitt and Richard Bettis, "Dynamic Core Competences Through Meta-Learning and Strategic Context," *Journal of Management*, 22(4), 549-569.

[93] Levinsohn, James and Wendy Petropoulos, "Creative Destruction or Just Plain Destruction: The U.S. Textile and Apparel Industries since 1972," June 2001, Working paper.

[94] Lippman, S. and R. P. Rumelt, "Uncertain Imitability: An Analysis of Interfirm Difference in Efficiency Under Competition," *Bell Journal of Economics*, Autumn 1982, 418-438.

[95] Liu, Lili, and James R. Tybout, "Productivity Growth in Chile and Columbia: The Role of Entry, Exit and Learning," in *Industrial Evolution in Developing Countries: MicroPatterns of Turnover, Productivity and Market Structure*, Mark J. Roberts and James R. Tybout, 1996, New York, Oxford University Press for the World Bank, 73-103.

[96] Longitudinal Employer-Household Dynamics Program Technical Paper No. TP-2002-05

[97] Lucas, Robert E. Jr., "Making a Miracle," March 1993, Econometrica, 61, 251-272.

[98] MacDuffie, John Paul, "Human Resource Bundles and Manufacturing Performance: Organizational Logic and Flexible Production Systems in the World Auto Industry," *Industrial and Labor Relations Review*, January 1995, 48, 197-221.

[99] Mahoney, Joseph T. and Rajendran Pandian, "The Resource-Based View Within the Conversation of Strategic Management," *Strategic Management Journal*, June 1992, 13.

[100] McKinsey Global Institute, *U.S. Productivity Growth 1995-2000*, October 2001, Washington, D.C., McKinsey Global Institute.

[101] Melitz, Marc, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, 71, 1695-1725.

[102] Melitz, Marc, "When and How Should Infant Industries Be Protected?" Forthcoming in the *Journal of International Economics*

[103] Montgomery, Cynthia A., "Corporate Diversification," *Journal of Economic Perspectives*, Summer 1994, 8, 163-178.

[104] Montgomery, Cynthia A. and S. Hariharan, "Diversified Expansion by Large Established Firms," *Journal of Economic Behavior and Organization*, January 1991, 15, 71-89.

[105] Montgomery, Cynthia A. and Birger Wernerfelt, "Diversification, Ricardian Rents, and Tobin's q," *RAND Journal of Economics*, Winter 1988, 19, 623-632.

[106] Nakamura, Leonard, "The Measurement of Retail Output and the Retail Revolution," Federal Reserve Bank of Philadelphia, Working Paper 98-5, 1998.

[107] Nelson, Richard and Sydney Winter, *The Evolutionary Theory of Economic Change*, Cambridge, MA, Belknap Press of Harvard University Press, 1992.

[108] Nguyen, Sang and Michael Ollinger, "Mergers and Acquisitions and Productivity in the U.S. Meat Products Industries: Evidence from the Micro Data," CES-WP-02-07, March 2002

[109] Olley, G. Steven, and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, November 1996, 64, 1263-1297.

[110] Pavcnik, Nina, "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *Review of Economic Studies*, January 2002, 69, 245-76.

[111] Penrose, Edith T., *The Theory of the Growth of the Firm*, Oxford, Oxford University Press, 1959.

[112] Pfeffer, Jeffrey and Nancy Langton, "The Effect of Wage Dispersion on Satisfaction, Productivity, and Working Collaboratively: Evidence from College and University Faculty," *Administrative Science Quarterly*, September 1993, 38, 382-407.

[113] Pfeffer, Jeffrey and Yinon Cohen, "Determinants of Internal Labor Markets in Organizations," *Administrative Science Quarterly*, December 1984, 29, 550-72.

[114] Ramanujam, Vasudevan and P. Varadarajan, "Research on Corporate Diversification: A Synthesis," *Strategic Management Journal*, November/December 1989, 10, 523-551.

[115] Rapping, Leonard, "Learning and World War II Production Functions," *The Review of Economics and Statistics*, February 1965, 47, 81-86.

[116] Reichheld, Frederick F. and Thomas Teal (Contributor), *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value,* July 2001, Harvard Business School Press.

[117] Scherer, Frederic M., *Industrial Market Structure and Economic Performance*, Chicago, IL, Rand McNally College Publishing Co., 1980.

[118] Schoenecker, Timothy S. and Arnold C. Cooper, "The Role of Firm Resources and Organizational Attributes in Determining Entry Timing: A Cross-Industry Study," *Strategic Management Journal*, December 1998, 19, 1127-1143.

[119] Shea, John, "The Input-Output Approach to Instrument Selection," *Journal of Business and Economic Statistics*, April 1993a, 11, 145-55.

[120] Shea, John, "Do Supply Curves Slope Up?" *The Quarterly Journal of Economics*, February 1993b, 108, 1-32.

[121] Shea, John, "Instrument Relevance in Multivariate Linear Models: A Simple Measure," *The Review of Economics and Statistics*, May 1997, 79, 348-52.

[122] Sherer, Peter D., "Leveraging Human Assets in Law Firms: Human Capital Structures and Organizational Capabilities," *Industrial and Labor Relations Review*, July 1995, 48, 671-691.

[123] Sheshinski, Eytan, "Tests of the 'Learning by Doing' Hypothesis," *Review of Economics and Statistics*, November 1967, 49, 568-78.

[124] Shleifer, Andrei and Robert W. Vishny, "Takeovers in the '60s and the '80s: Evidence and Implications," *Strategic Management Journal*, Winter Special 1991, 12, 51-59.

[125] Sieling, Mark, Brian Friedman, and Mark Dumas, "Labor Productivity in the Retail Trade Industry, 1987-99," *Monthly Labor Review*, December 2001, 3-14.

[126] Steffy, Brian D. and Steven D. Maurer, "Conceptualizing and Measuring the Economic Effectiveness of Human Resource Activities," *Academy of Management Review*, April 1988, 13, 271-286.

[127] Steindel, Charles and Kevin Stiroh, "Productivity: what is it and why do we care about it?," Staff Reports 122, Federal Reserve Bank of New York, 2001.

[128] Syverson, Chad, "Market Structure and Productivity: A Concrete Example," *Journal of Political Economy*, November 2004, 112, 1181-1222.

[129] Teece, David J., "Economics of Scope and the Scope of the Enterprise," *Journal of Economic Behavior and Organization*, September 1980, 1, 223-247.

[130] Teece, David J., "Towards an Economic Theory of the Multiproduct Firm," *Journal of Economic Behavior and Organization*, March 1982, 3, 39-61.

[131] Teece, David J., Gary Pisano and Amy Shuen, "Dynamic Capabilities and Strategic Management," *Strategic Management Journal*, August 1997, 18, 509-533.

[132] Triplett, Jack E. and Barry P. Bosworth, "The Solow Productivity Paradox: What Do Computers Do to Productivity?" *Canadian Journal of Economics*, April 1999a, 32, 309-34.

[133] Triplett, Jack E. and Barry P. Bosworth, "Economic Statistics, the New Economy, and the Productivity Slowdown." *Business Economics*, April 1999b, 34, 13-17.

[134] Triplett, Jack E. and Barry P. Bosworth, "Numbers Matter: The U.S. Statistical System and a Rapidly Changing Economy." *Brookings Policy Brief*, 63, July 2000.

[135] Triplett, Jack E. and Barry P. Bosworth, "What's New About the New Economy?: IT, Economic Growth and Productivity," *International Productivity Monitor*, Spring 2001a, A

longer version of this paper appeared in *The Impact of the Internet Revolution on International Economic Relations and Society*, Tokyo Club Papers No. 14, editors: Heidemarie C. Sherman, 2001.

[136] Triplett, Jack E. and Barry Bosworth, "Productivity in the Services Sector," in *Services in the International Economy*, editor: Robert M. Stern, 2001b, University of Michigan Press.

[137] Triplett, Jack E. and Barry P. Bosworth, "Productivity Measurement Issues in Services Industries: "Baumol's Disease" Has Been Cured," *Federal Reserve Bank of New York Economic Policy Review*, September 2003, 9, 23-33.

[138] Tybout, James R., "Heterogeneity and Productivity Growth: Assessing the Evidence." in *Industrial Evolution in Developing Countries: MicroPatterns of Turnover, Productivity and Market Structure*, Mark J. Roberts and James R. Tybout, 1996, New York, Oxford University Press for the World Bank, 43-72.

[139] Wernerfelt, Birger, "A Resource-based View of the Firm," *Strategic Management Journal*, April-June 1984, 5, 171-80.

[140] Williamson, Oliver E., *The Economic Institutions of Capitalism*, New York, Free Press, 1985.

[141] Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in *Handbook of Labor Economics,* Volume 1, editors: O. Ashenfelter and R. Layard, New York, Elsevier Science.

[142] Wright, T. P., "Factors Affecting the Cost of Airplanes," *Journal of the Aeronautical Sciences*, February 1936, 3, 122-28.