ABSTRACT

Title of Dissertation: LABOR MARKET SKILL, FIRMS AND WORKERS

Nicole Nestoriak, Ph.D., 2004

Dissertation Directed By: Professor John C. Haltiwanger, Department of Economics

The role of skilled labor in the modern economy and its importance in explaining trends in wage inequality and productivity has been a focus of a broad strand of research in economics. Labor is an input into production that is very different from capital or materials. One key difference is that workers must decide where they are going to live, and in making that decision, they thereby limit the opportunity set of jobs available to them. In turn, firms must also make a location decision that affects their access to labor and potentially affects their decisions on the technology they will adopt. While many economists have studied issues related to technology adoption and worker skill broadly, the geographic element is rarely developed. This dissertation exploits the variation in the concentration in skilled labor across local labor markets in a sample of U.S. States to study how movements of workers affect the distribution of skill across geography, the investment decisions by firms in reaction to the variation in skill and finally the effect of this variation on worker's wages across local labor markets. Given that skilled labor is an important

force in the economy, variation in the concentration of skilled workers across local labor markets may also play an important role.

The research set out here confirms this hypothesis. Workers locate non-randomly across geography and their movements reinforce the existing distribution of skill across local labor markets. As predicted by a model of endogenous technology, firms react to the skill level of their local labor market. Variation in firm level investment can be partially explained by variation in the availability of skilled labor. The empirical work shows that among a sample of manufacturing firms in 1992, a one standard deviation increase in county skill leads to a 10% increase in firm level investment in computers. Finally, highly skilled workers receive higher wages in metro areas with strong concentrations of skill. Deeper examination of the data shows that this wage gap is largely due to higher returns to skill in highly skilled areas.

LABOR MARKET SKILL, FIRMS AND WORKERS

By

Nicole Nestoriak

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
2004

Advisory Committee:
Professor John C. Haltiwanger, Chair
Professor John M. Abowd
Associate Professor Judith Hellerstein
Associate Professor John Horowitz
Associate Professor John Shea

Preface

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.

Acknowledgements

The completion of this dissertation was made possible through the generous support of advisors and friends. First and foremost, I thank my husband, Kevin Moore, for his encouragement, patience, and willingness to share in my journey through graduate school. I would also like to thank my parents for believing in my abilities and therefore giving me the confidence to undertake this dissertation.

I am deeply indebted to my primary advisor, John Haltiwanger. His enthusiasm for research, thought-provoking conversations, and guidance made graduate school a more enjoyable and beneficial experience. I am also grateful to other members of my dissertation committee: Judy Hellerstein for her careful research practices, John Shea for taking the time to understand my research and helping to formulate ideas more clearly, and John Abowd for his technical expertise.

Finally, my colleagues at the Longitudinal Employer Household Dynamics program at the US Census Bureau provided a great research environment. I am particularly grateful to Hyowook Chiang and Kristin Sandusky for their help in understanding and managing the endless data, but mostly for listening and providing advice.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The role of skilled labor in the modern economy and its importance in explaining trends in wage inequality and productivity has been a focus of a broad strand of research in economics. Labor is an input into production that is very different from capital or materials. One key difference is that workers must decide where they are going to live, and in making that decision, they thereby limit the opportunity set of jobs available to them. In turn, firms must also make a location decision that affects their access to labor and potentially affects their decisions on the technology they will adopt. While many economists have studied issues related to technology adoption and worker skill broadly, the geographic element is rarely developed. This dissertation exploits the variation in the concentration in skilled labor across local labor markets in a sample of U.S. States to further study how mobility decisions of workers affect the distribution of skill across geography, the investment decisions by firms in reaction to the variation in skill and finally the effect of this variation on worker's wages across local labor markets. Given that skilled labor is an important force in the economy, variation in the concentration of skilled workers across local labor markets may also play an important role.

Before studying the effect of variation in the concentration of skilled labor on firms and workers, the next chapter lays out the extent of the variation and how these patterns change over time through movements of workers. Previous research in this area has focused on the characteristics of workers that affect their decision to move across labor markets. A related literature focuses on shocks to local labor markets, the resulting patterns of worker mobility, and the final effect on worker's wages by skill. This chapter takes a different perspective and focuses on the local labor market skill level. Using an extensive database of workers, this chapter focuses on both cross-sectional variation in skilled labor across local labor markets and also on time-series trends in skill level. In addition to testing for statistically significant variation in skill level across counties and metro areas, this essay asks if the extent of this variation is economically important through studying patterns in county level skill both across counties and across time. The research shows that the distribution of county skill appears to be bimodal, suggesting that they are two types of counties as defined by skill. These patterns largely persist over time. The small changes in the distribution of county skill over time are decomposed into the components due to workers who continue to work in the same county over the time period, workers who enter and leave the labor force, and workers who work in different counties in the beginning and end of the time period. This chapter shows that while workers are very mobile, their mobility patterns reinforce the existing patterns in the distribution of county skill.

The next chapter utilizes this variation in the skill level of local labor markets to study endogenous technology. Labor is an input to production different from capital or materials in that workers must decide to locate near their place of employment, resulting in significant variation in the skill level of available workers across areas. Does this variation affect the technology decisions of firms? New technologies boost the productivity of firms, but research has shown that the biggest productivity boost comes from combining new technologies with the appropriately skilled workers. The previous chapter outlines the patterns of skilled labor across geography noting that these skilled workers are not equally available across local labor markets. This chapter continues on this theme, focusing on whether or not firms take variation in the availability of skilled labor into consideration when making investment decisions. A model is developed in which two assumptions about the economy lead to endogenous technology. The first is that skilled labor and technology are complementary inputs. The second assumption is that firms make their investment decision before hiring workers. Under this set of assumptions, the model shows that firms will invest more in highly skilled areas because their probability of finding a skilled worker is higher in a more skilled labor market. The empirical work then tests the model using data on the computer investment decisions of manufacturing firms in 1992 in combination with the data utilized in the previous chapter on local labor market skill. Estimates of the effect of endogenous technology predict that a one standard deviation increase in local labor market skill will lead to roughly a 10% increase in technology investment. The results are shown to be robust to a series of different specifications of the investment equation.

In the final essay, the focus is on the effect of variation in concentrations of skilled labor on worker's wages. Public funding of education has largely been motivated by the belief that the returns to an individual's education reward the greater economy. Potential mechanisms by which these social returns accrue include spillovers to other workers through informal communications and greater investment by firms to take advantage of high worker skill, as outlined above. Both sets of theories predict that workers in high human capital labor markets are more productive. Economists have tested for higher worker productivity in highly skilled labor markets by studying worker's wages. This simple test leads to a variety of measurement issues mostly focused on the unobservable differences of workers in highly skilled labor markets. Workers in these labor markets are, by definition, more highly educated. Given that they are a selected group, it is likely that they also vary on unobservable dimensions that might additionally lead to differences in their wages. Earlier papers have addressed this issue by trying to control for unobservable worker characteristics via a person fixed effect. While this methodology will control for characteristics such as ability, the estimation implicitly restricts the return to these unobservable characteristics to be the same in high and low skill labor markets. This essay addresses this measurement question by directly estimating the return to unobservable worker characteristics separately in high and low skill areas by using variation in the labor market skill of workers who switch jobs. After controlling for education, experience, and unobservable worker characteristics, and allowing for the returns to these characteristics to vary in low and high skill areas, the remaining wage gap between workers in high and low skill areas becomes insignificant. These results do not necessarily suggest that workers in high skill areas are not more productive than those in low skill areas, but rather it is the most skilled workers who receive a bigger productivity boost from locating in a highly skilled area.

Chapter 2: Geographic Dispersion of Human Capital

Analysis of worker migration within the United States has found that high skill workers are highly mobile in absolute terms and relative to low skill workers. Related research into wage and employment dynamics in local labor markets has looked into the response of workers to local demand shocks. Complementing the earlier research on worker migration, studies of local labor market dynamics have found that the high mobility of skilled workers equalizes their wages across local labor markets. While the same cannot be said for low skilled workers, the previous evidence suggests that the labor market for highly skilled workers is potentially national, and at a minimum larger than their local area. This set of facts would lead one to believe that geography is irrelevant in the study of skilled workers. However, the previous research has not shed light on the extent to which highly skilled workers are distributed across geography. Given a national market for highly skilled workers, a firm requiring highly skilled workers would not be limited to any particular labor market, and geography would be irrelevant. However, if skilled labor clusters in particular areas and the mobility patterns of skilled workers reinforces these existing distributions of labor across geography, then the availability of skilled labor and geography more generally remains an important factor for a variety of issues in labor economics. There is little existing research that analyzes the distribution of highly skilled workers across geography. This chapter attempts to fill this gap in the literature by examining the distribution of human capital across geography.

Despite the fact that little is known about the variation in human capital across local labor markets, many economic models depend on this variation or on concentrations of highly skilled workers to explain variation in other worker characteristics or firm behavior. These models include those with increasing returns to human capital accumulation, firm externalities in human capital, and higher growth in skilled areas[1]. In the next two chapters, a model of endogenous technology is developed in which firms in high human capital areas are more likely to invest in technology complementary to high skill workers, and a model with social returns to education motivates a study of variation in worker's wages across areas of different skill types. While the later chapters use micro-data to analyze endogenous technology and geographic variation in wages, this chapter asks a more basic question: is there an economically significant amount of variation in human capital across local labor markets? If all firms face roughly the same skill distribution of workers in their local labor market, then one would not expect to find endogenous technology or any other variation in worker or firm characteristics across local labor markets that is attributable to variation in available local skill. However, evidence which suggests that local labor markets can be distinguished by their worker skill distribution and by different patterns of firm and worker mobility by skill lend support to the class of models which rely on this type of variation.

The first part of this chapter defines a measure of worker human capital that includes all fixed characteristics of the worker that are compensated by wages. Given a definition of the local labor market and of worker skill, how does one determine whether or not there is economically significant variation in worker skill? A first pass

[1] See Acemoglu (1996), Moretti (2002), and Glaeser, Scheinkman, and Schleifer (1995) for examples.

at the data requires testing if there is a statistically significant amount of variation in worker skill across local labor markets. However, this approach does not fully answer whether there is an economically significant level of variation in skill. A local labor market level measure of worker skill is then constructed to compare the skill of local labor markets directly. Using the labor market measure of skill it is possible to look at the distribution of counties by skill and to see how these distributions evolve over time. Finally, beyond the aggregate patterns, it is also possible to decompose changes in the distribution of workers by county skill into components that are due to workers who continue working in the same county over the sample time period, workers who move between counties, entrants and exiters.

The findings in this chapter do not support the hypothesis of a uniformly-distributed, national labor market for high skill workers. The data suggest that there is significant variation in skill across local labor markets and that this variation is persistent. Further, the variation suggests that counties can also be classified as high and low skill. Worker mobility patterns by skill provide further evidence that while workers are mobile, their mobility reinforces the existing distribution of skill across local labor markets.

2.1 Background

The high mobility of highly educated people has been well documented. Greenwood (1975) summarizes the literature nicely. High levels of education are consistently among the characteristics of a person that make him more likely to migrate, and this relationship is stronger the longer the distance of the move. Additionally, highly educated people are more likely to move because of a job. Long

(1988) examines migration patterns across census regions by education. He finds that New England is a net exporter of college-educated individuals while the West Coast is a net importer.

A related literature examines wage and population adjustments in response to local demand shocks. Topel (1986) uses CPS data from the 1970s to study the effect of local demand shocks on worker's wages and employment. He finds that local demand shocks have the greatest effect on older and less-educated workers who are also the least mobile. Bound and Holzer (2000) expand upon Topel's finding using Census data from 1980 and 1990. They also find that local demand shocks are most strongly felt by the less experienced, less educated and black workers. Further, Bound and Holzer argue that the limited mobility of these groups contributed to the well-documented deterioration of their wages over the 1980s.

Moving beyond describing the patterns of migration in the data, Kennan and Walker (2003) estimate a structural model of migration decisions. They find that worker's migration decisions do seem to be influenced by the potential for higher income, or a better locational match. However, they also find that migration does not seem to be influenced by geographic differences in wage distributions. Therefore, worker mobility does not arbitrage the large differences across locations in worker's wages. Instead of the traditional literature, which has focused more on the numbers of workers moving across locations, the authors also measure the cost of mobility off of these regional wage differences. Based on this methodology, they find that the average worker would need to be compensated $250,000 in order to induce him to move. While some of this moving cost may be due to an omitted variables problem, i.e. not including variables which explain a worker's preference for his current location, the results are suggestive of the frictions that exist in worker movement across geography.

2.2 Data

All of the data used in this research are part of the Longitudinal Employer-Household Dynamics program at the Census Bureau. Information on workers comes from the Unemployment Insurance wage records for the selected three states[2]. These files contain person identifiers that allow one to track a worker's earnings over the available period, from 1991 to 1998. The data also contain firm identifiers that allow for an exact link between the UI files and other data sets. The UI wage records contain virtually all business employment for the states included in the analysis, creating a final sample size of 198,644,076 observations representing 37,875,250 people and 3,989,740 firms. The disadvantage of using the UI wage data to characterize workers is the very limited demographic information available. Within the Census bureau, this problem has been partially overcome by combining the UI wage data with other administrative data containing information on date of birth, place of birth, and gender. Additionally, as will be discussed in more detail in the next section, the panel aspect of the data allows one to separate out worker and firm effects.

The local labor market throughout this chapter will be defined as county of work for the employees. There is some limited county of residence information also

[2] Three states were selected on the basis of time-series availability at the time of project inception. This research cannot reveal the identity of the three states used in the analysis due to confidentiality restrictions.

available; however, it only provides information for 1999 and forward. One potential drawback in defining local labor market skill by county of work is that by definition the measure only includes the working portion of the local labor market. This issue of mismeasurement will only cause problems if unemployment rates are large and the distribution of skill among the unemployed varies widely across counties, which seems unlikely. Other geographic information on firms, such as Metropolitan Statistical Area, is also available within the dataset used. The analysis below can directly test whether or not there is significant variation of worker skill by county within MSAs.

As mentioned above, the UI wage records contain identifiers for a worker's firm, but not a worker's establishment. Without the establishment identifier for each of the workers, it becomes difficult to create a measure of local labor market skill, which is an aggregation of individual worker skills. In particular, if a firm has establishments in multiple counties, it is impossible to determine to which county to assign the worker. While multi-unit firms only represent roughly 30% of employment for the states being studied, some algorithm must be used to allocate these workers to the correct county. Fortunately, the ES202[3] files provide additional information that helps to alleviate this problem. In particular, the ES202 lists all establishments, their county location, and the number of employees at each establishment for all of the firms. From this data set, it is possible to assign an employee-weighted county to each firm. While it is impossible to determine which workers are properly assigned, a

[3] The ES202 files are part of the Bureau of Labor Statistics Quarterly Census of Employment and Wages program. These data are provided to the LEHD program directly from the states.

simple tabulation verifies that this procedure correctly identifies the county of work for 91% of the workers[4].

2.3 Characterizing Human Capital

Due to the limited amount of demographic information available, and in particular the lack of information on a worker's education, worker's wages are decomposed into a worker effect, a firm effect, and a time varying effect as follows

$$w_{ijt} = \theta_i + X'_{it}\beta + \psi_{j(i,t)s} + \varepsilon_{it}$$

This decomposition of wages is a variation on the methodology developed by Abowd, Kramarz, and Margolis (1999). Wages are measured on an annual basis (see Abowd, Lengermann, and McKinney (2003) for details on construction of this variable). Human capital is captured in the fixed worker effect, θ, and a quadratic in experience captured within X.[5] Firm characteristics are captured in the time-varying firm effect, ψ.[6] The remaining variables contained in X are a series of gender by year by labor force attachment status dummies. These variables control for the observable time-varying characteristics. Due to the large sample sizes, the wage equation is estimated separately for each state using the conjugate gradient methodology as explained in Abowd, Creecy, and Kramarz (2002). The solution algorithm involves

[4] The county is known for all workers at single unit firms. Nearly 70% of workers in the sample are employed at single unit firms. Workers at multi-unit firms are assigned to the county in which the firm has the greatest number of employees. The percentage of workers correctly assigned can then be determined by calculating the percentage of workers at the multi-unit that are in the county with the greatest employment for that multi-unit.

[5] Experience is equal to the sum of an estimated initial experience and observed experience over the sample time period. Initial experience is equal to age minus imputed years of schooling minus 6. See Abowd, Lengermann, McKinney (2003) for more details.

[6] See section 3.4 for further explanation of the wage decomposition and the decision to use time-varying firm effects.

grouping the data into connected groups. Within each group all but one person and firm effect is identified. In practice, the identification restriction is applied post-processing by setting the group mean of the person and firm effects equal to zero. Results are then pooled together across states weighting by employment in 1992. The worker and firm effects are adjusted by state level mean wages so that they are comparable across states.

The fixed worker effect in this model captures the component of the worker's wages that can be attributed to the worker and reflects any fixed characteristic of the worker that affects his wages. Although no individual level comparison of the worker effect and more traditional measures of skill are done in this chapter, Abowd, Lengermann, and McKinney (2003) have found that there is a positive correlation between the worker effect and years of schooling. In addition to years of schooling, the worker effect should also capture other characteristics of the worker including the quality of the college attended, the major chosen, and the success of the student in school.

In addition to separating out the component of wages attributable to the worker, this wage decomposition has the additional advantage of controlling for area specific fixed effects in the firm effect. One of the difficulties in comparing a worker's wages across areas is that variation in the cost of labor across local labor markets might vary due to varying labor market tightness or cost of living differences. Because firms almost never change location, any fixed characteristics of a physical location should be captured within the fixed firm effect. Further, the worker effect is unlikely to be contaminated with fixed area characteristics due to the high mobility of workers. Within the three state sample, an average of 20% of workers in each county in 1998 had worked in a different county in 1992. Any worker at a firm in which at least one worker had worked in a different county should have a worker effect uncontaminated by the area effect. Given the high degree of worker mobility in the data, most workers should be covered by this condition. A regression of the firm effects on a set of county dummies tests the extent to which regional variation is being absorbed in the firm effects. County dummies explain approximately seven percent of the variation in firm effects. While the county dummies only account for a small proportion of the overall variation, seven percent is approximately twice the amount of variation that counties explain in similar regressions in which the worker effect is the dependent variable.

While some of the following analysis utilizes the micro data, further analysis comparing counties requires a county measure of skill. Two different measures are used extensively throughout the analysis. The first is a simple average of $\widehat{\theta}_{it}$, the estimated worker effect from the wage equation, within the local labor market, labeled $\widehat{\theta}_{lt}^{avg}$. The other measure calculates the percentage of workers within the local labor market that are above the 75th percentile of the overall pooled three state distribution of $\widehat{\theta}_{it}$ for 1992, the reference year. This alternate measure of local labor market skill is labeled $\widehat{\theta}_{lt}^{75}$. As shown in the results, these two measures capture different aspects of the skill distribution.

Table 2.1 summarizes the results of estimating the wage equation with limited time varying firm effects. Looking across the first row, the correlation of log wage with the worker effect is 0.56 and the correlation with the firm effect is 0.50. These results suggest that worker and firm effects are equally important in explaining the variation in log wages. The covariance between the worker and firm effects at the individual level is positive, although small in magnitude at 0.07. The positive covariance between worker and firm effects suggests that high skill workers are more likely to be at employed at high wage firms.

Table 2.1: Results from Wage Regression

	Log wage	Worker effect	Firm effect	XBeta	Residual
Log wage	1	0.5643	0.4958	0.2294	0.4207
Worker effect		1	0.0655	-0.4740	0.0000
Firm effect			1	0.0355	0.0000
XBeta				1	0.0000
Residual					1

2.4 Distribution of Worker Skill across Counties

Using the results from the wage decomposition above, table 2.2 explores the variation in the person effects across counties, metro areas, and industries. The basic equation being estimated across the columns is some variant of

$$\theta_{ilmnt} = \alpha + \gamma_l + \gamma_n + \varepsilon_{it}$$

where γ_l are the set of county dummies, and γ_n are the set of two digit industry dummies. The first set of columns in the table are the p-value and R-squared from a series of regressions of the worker effects on the county dummies alone, separately for each year in the data. Over the sample time period, there is statistically significant variation in the worker effect across counties. While the variation is significant, the county dummies explain only between 2 and 4% of the overall variation in the worker effect. These results are not surprising given that within any given county there is still a tremendous amount of variation in worker skills.

Table 2.2: Regress θ_i on county, and county and 2-digit industry dummies, by year

	1		2	
	p-value	R-Squared	p-value	R-squared
1992 county	0.0001	0.0374	0.0001	0.0656
ind			0.0001	
1993 county	0.0001	0.0357	0.0001	0.0635
ind			0.0001	
1994 county	0.0001	0.0337	0.0001	0.0611
ind			0.0001	
1995 county	0.0001	0.0311	0.0001	0.0580
ind			0.0001	
1996 county	0.0001	0.0284	0.0001	0.0555
ind			0.0001	
1997 county	0.0001	0.0260	0.0001	0.0524
ind			0.0001	
1998 county	0.0001	0.0232	0.0001	0.0478
ind			0.0001	

Somewhat more difficult to explain is the pattern of decreasing R-squared values over time. While in 1992 counties explain 3.7% of the variation in the worker effect, by 1998 counties explain only 2.3% of the variation. This fact alone would suggest that the concentration of like workers into counties is decreasing over time. However, this result may also be driven by the methodology used to quantify worker skill. Given that the worker effect is fixed over for a given worker, the only changes over time in the overall distribution of the worker effects will arise from exiters and entrants. Over the 1990s the labor market grew increasingly strong for workers. Therefore, it is possible that new entrants later in the time period have measured worker effects that still capture, in part, their inherent skill, but that also capture their labor market luck. Extension of the time period into the next decade would allow one to test this hypothesis, but this is beyond the scope of the current research.

The focus throughout the chapter is on the location of workers by county, even though much of the literature on local labor markets uses the metropolitan area as the relevant geographic measure. Because counties nest into metropolitan areas, it is possible to test whether or not there is significant variation in county skill within metro areas by testing for equality of the coefficients on county dummies for counties within metro areas. In the sample, 15 of the 40 metropolitan areas are composed of multiple counties. Tests of equality of the county coefficients are rejected in all of these 15 metro areas. These results suggest that there is significant variation in county skill within metro areas.

Finally, the second set of columns explore a regression of the worker effect on both county and industry dummies. These regressions begin to address the concern that geographic variation in worker skills is primarily due to variation in the industry mix across geography. The direction of causality between firm location and worker location is uncertain. Firms of the same industry may choose a location based on access to transportation, input markets or output markets, and workers may follow firms in order to have access to the jobs they provide. It is also possible that workers make location decisions based on amenities such as proximity to facilities of higher education and that firms choose to locate near these workers in order to have access to the highly skilled labor pool. Determining the comparative importance of these two theories on worker location is beyond the scope of this chapter. However, the extent to which worker location is determined by industry mix is important to many of the theories that rely on geographic variation in worker skills. While the skill mix of workers in an area can never be assumed to be exogenous, if the worker skill mix is completely determined by the industry mix, identification of the effects of skill mix on other worker or firm characteristics will be difficult to achieve.

The second set of columns in table 2.2 begins to address these concerns by testing if county dummies account for a statistically significant amount of skill variation after controlling for industry mix. As is clear from the p-values, county dummies are still significant in these regressions. Industry dummies are also highly significant and the R-squared is much larger with the inclusion of industry dummies. While industry is clearly important in explaining variation in skill, there still remains a role for geographic variation.

2.5 Distribution of County Skill

While the first two tables concentrated on the micro data and the extent to which counties can explain variation in skill, the main focus of this chapter is on how counties vary in skill mix. As mentioned above, two county level measures of skill are computed. The first measure is the percentage of workers in a county who are above the 75^{th} percentile of the overall distribution of worker skill. The second measure is the simple average of worker skill within a county. Summary statistics of both county skill measures are in table 2.3. The first set of columns shows that the average percentage of workers in the top quartile of the distribution by county is 20% in 1992 and increases to 25% in 1998. The increase in county skill over the time period as measured by the fixed worker effect is due to a combination of the exit of less skilled workers and the entry of more skilled workers over the time period. County skill may also be increasing if skilled workers move from larger to smaller counties. The average percentage in the top quartile is generally less than 25%.

Because the reported average is not adjusted by county population, the average percentage in the top quartile below 25% reflects the fact that larger counties have more skilled workers. The pattern in county skill measured as the average of the worker effect within the county is similar. One additional pattern shows up in the second measure, however. The standard deviation of county skill is decreasing over the time period suggesting that counties are becoming more alike over time. In addition to the skill measures, table 3 also includes yearly means and standard deviations of the number of workers in the counties. Average county size increases monotonically, while the standard deviation of county size increases over the time period in a non-monotonic fashion.

Table 2.3: Summary statistics of county skill measures, $\theta_l^{>75}$ *and* θ_l^{mn}

	1		2		3	
	Top quartile skill		Average skill		Number of workers	
year	mean	sd	mean	sd	mean	sd
1992	0.1999	0.0470	-0.1929	0.1119	99712	397476
1993	0.2083	0.0482	-0.1665	0.1127	99075	391473
1994	0.2163	0.0481	-0.1475	0.1120	100637	395591
1995	0.2246	0.0483	-0.1243	0.1062	102105	394539
1996	0.2327	0.0483	-0.1061	0.1036	104429	402364
1997	0.2419	0.0478	-0.0849	0.0997	107168	412440

In order to focus more on the entire distribution of county skill, kernel densities are estimated for each of the skill measures. The kernel density estimate is computed using

$$\hat{f}(x;h) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

where n is the number of observations, h is the bandwidth, X_i is the skill measure and K is the Gaussian kernel. Figure 2.1 contains the kernel density estimate of the distribution of county skill measured by the percentage of workers in the county who are in the top quartile. The solid line represents the distribution in 1991, long dashes represent the distribution in 1995, and the small dashes represents the distribution in 1998. The density emphasizes the tremendous variation in county skill. In 1991, the least skilled county had 10% of its workforce in the top quartile while the most skilled county had 33% of its workforce in the top quartile. The distributions appear to be bi-modal with a second higher skilled mode being much smaller than the first. Formal tests of bi-modality are performed latter in the chapter. Comparing the densities across the three years represented, all counties are becoming more skilled over time. Figure 2.2 repeats the same three densities this time weighting each county by the number of workers in the county in that year. While the overall shape of the distribution still appears to be bi-modal, the location of the modes is shifted to the right. Comparisons of Figure 2.1 and Figure 2.2 make clear the point made earlier that larger counties are more highly skilled.

Figure 2.1: Distribution of θ_l^{75} across counties

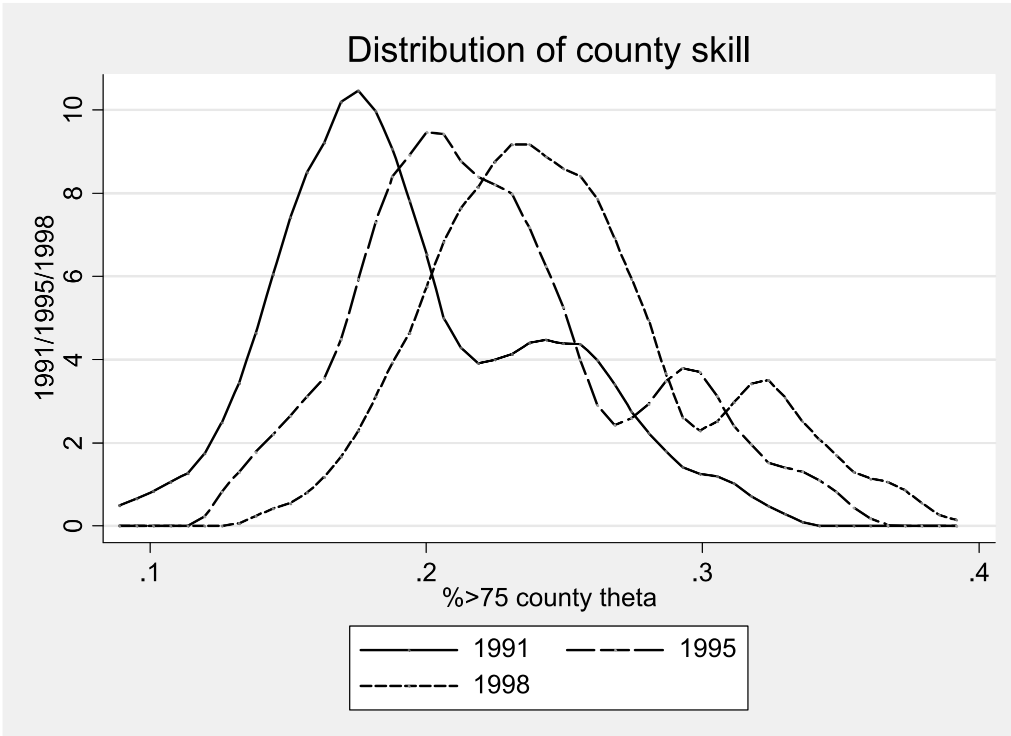


Figure 2.2: Distribution of θ_l^{75} across counties, weighted

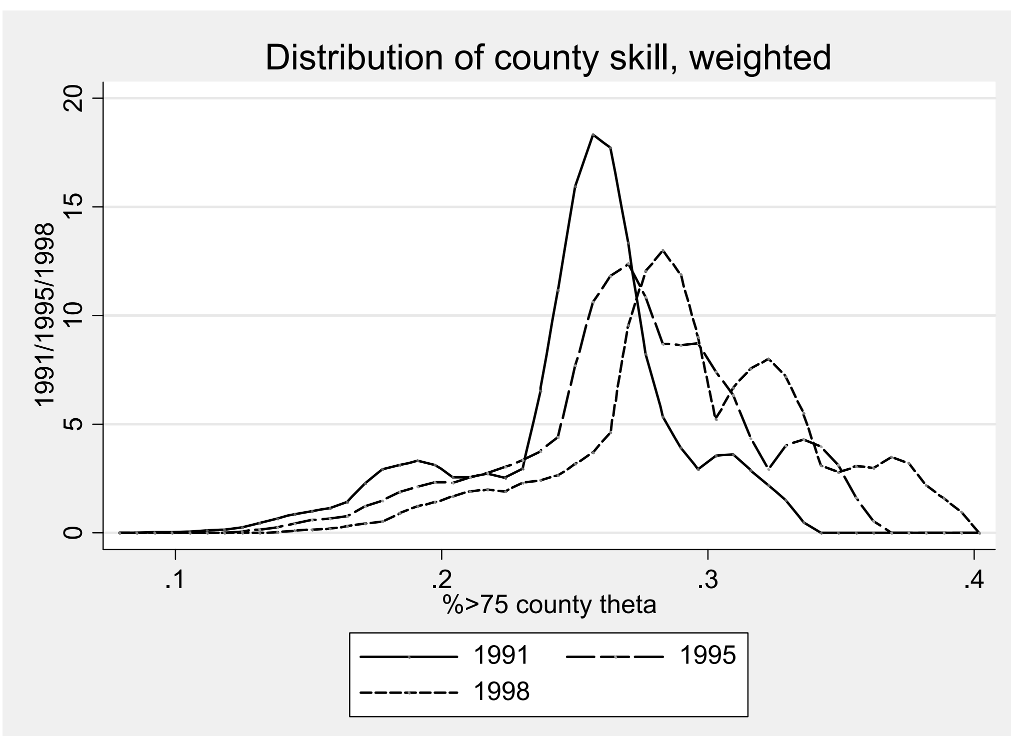


Figure 2.3 repeats the exercise in figure 2.1 for the other measure of county skill, the average of the worker effect in the county. This distribution does not appear to be bi-modal but still maintains the overall shape of a large mode in the left half of the distribution and a long right tail. Additionally, the distribution appears to become more concentrated over time as the mode both becomes larger and narrower. Figure 2.4 is the worker-weighted version of figure 2.3. Again, the correlation between county size and skill is clear as the weighted distribution moves strongly to the right.

Figure 2.3: Distribution of θ_l^{avg} across counties

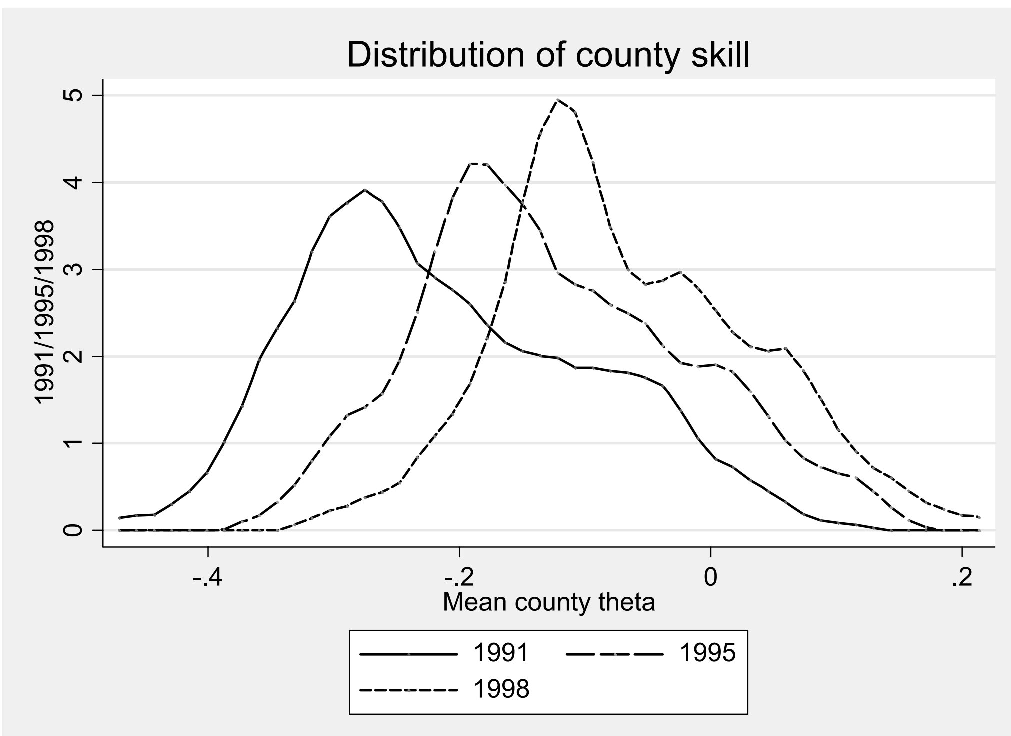
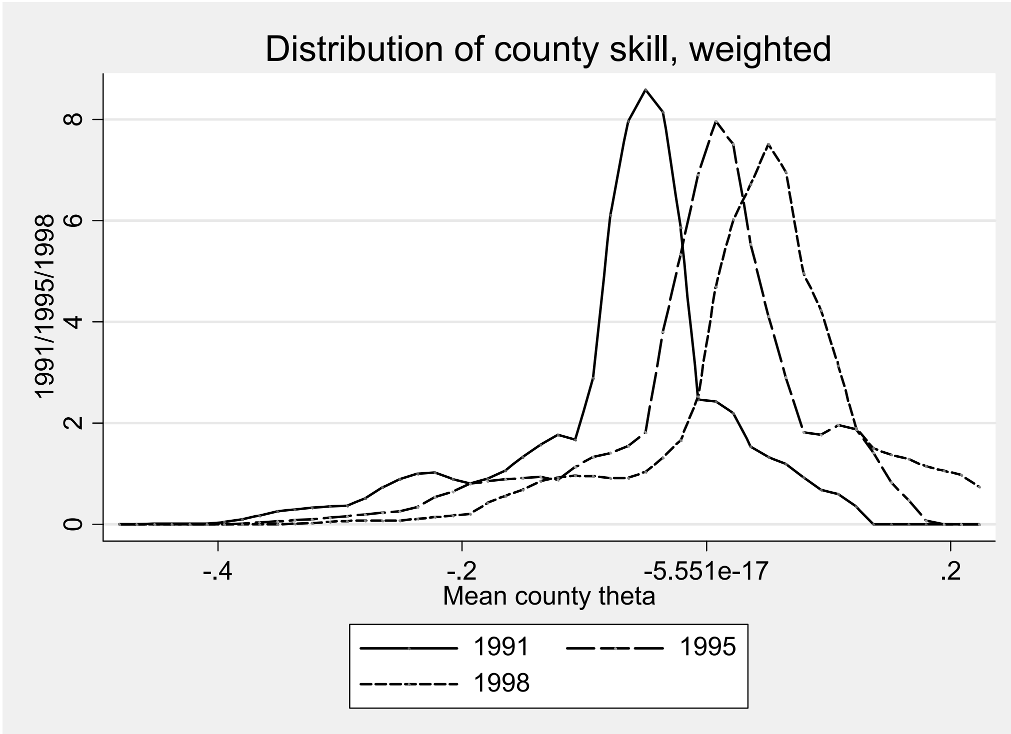


Figure 2.4: Distribution of θ_i^{avg} across counties, weighted



2.5.1 Tests For Bimodality

As mentioned above, the densities of county skill measured by the top quartile appear to be bi-modal. A bi-modal density suggests that the underlying density is composed of two distinct densities with well-separated means.[7] In the context of county skill, this implies that within the overall distribution there are two distinct types of counties, high skill and low skill. Evidence of bimodality would therefore suggest that not only is there a large amount of variation in county skill, but counties tend to fall into one of two distinct types.

[7] See Bianchi (1997) for a similar discussion related to the distribution of per-capita GDP in a cross-section of countries.

Visual inspection of figures 2.1 and 2.2 would lead one to believe that the distribution of county skill is indeed bi-modal. However, the existence of multiple modes in a kernel density estimate depends critically on the bandwidth, h, chosen. While a bandwidth can be optimally chosen by a variety of different rules, there is no test statistic for the bandwidth and therefore no direct test statistic for multiple modes. Silverman (1981) proposes a test for multi-modality based on the relationship between bandwidth and multi-modality that utilizes smoothed bootstrapping. Under the null hypothesis that the density has at most one mode and the alternate hypothesis that the density has more than one mode, the critical bandwidth, h_1^{crit}, is defined as the minimum bandwidth under which the density has at most one mode. If the true underlying density is multimodal, it will require a large critical bandwidth value to smooth a multimodal density into a unimodal density estimate. Therefore, the test calculates the critical bandwidth value for each smoothed bootstrap sample, and the corresponding achieved significance level is given by

$$ASL_{boot} = \Pr_{\hat{g}(\cdot;\hat{h}_1)}\{\hat{h}_1^* > \hat{h}_1\}$$

where $\hat{h}_1$ is the estimated critical bandwidth for 1 mode and $\hat{h}_1^*$ is the critical bandwidth for one mode estimated from the smoothed bootstrap sample drawn from the rescaled density estimate, $\hat{g}(\cdot;\hat{h}_1)$. The test is based off of a rescaled density estimate, $\hat{g}(\cdot;\hat{h}_1)$, as opposed to the estimated null distribution $\hat{f}(\cdot;\hat{h}_1)$, in order to adjust the bootstrapped sample estimate to the original sample variance.[8]

[8] See Efron and Tibshirani (1993) for details.

Applying the above test to the county skill densities, figure 2.5 and 2.6 calculate the relationship between bandwidth and number of modes in 1992 and 1998 using the percentage of workers in the top quartile as the measure of county skill. As is clear in the graphs, as the bandwidth becomes larger, the number of modes becomes smaller. The critical bandwidth is, as mentioned above, the smallest value of the bandwidth that produces a given number of modes. For test purposes here, the focus is on testing the null of at most one mode and at most two modes, versus the alternate hypotheses of more than one mode and more than two modes respectively. Tests for sequentially larger number of modes can be considered in sequence. The associated achieved significance levels, or p-values, are shown in table 2.4. The tests are performed separately for each successive year of the data. Looking at the results in the first column, the null of at most one mode can be rejected at the 10% level in all years and at the 5% level in three out of the eight years. The null of at most two modes, in the second column of the table, is not rejected in any year.

Figure 2.5: Critical Bandwidths, 1992

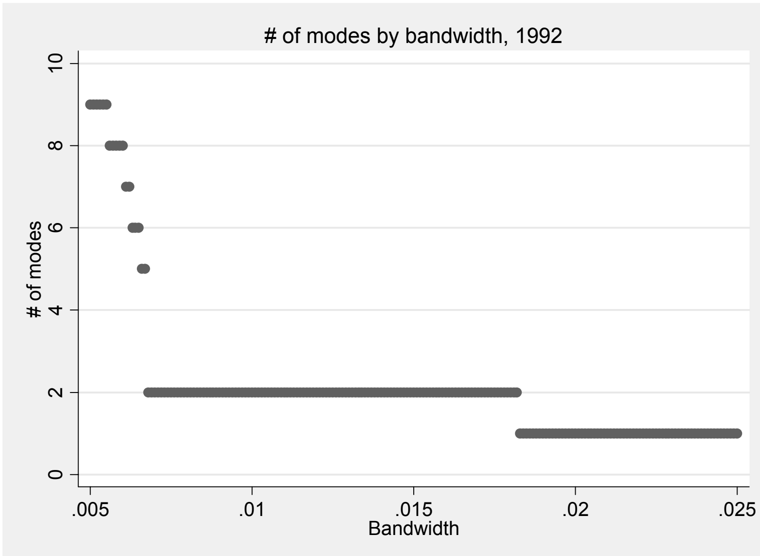


Figure 2.6: Critical Bandwidths, 1998

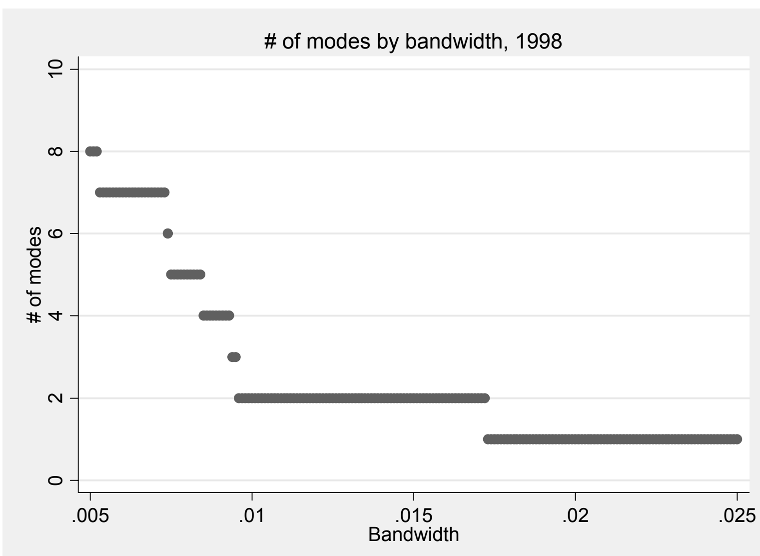


Table 2.4: Test for multimodality
p-value for h0: k modes, h1: > k modes
County skill, top quartile worker effect

modes	1	2
1992	0.059	0.985
1993	0.032	0.406
1994	0.028	0.298
1995	0.064	0.485
1996	0.056	0.283
1997	0.026	0.470
1998	0.083	0.649

The data clearly suggest that the distribution of county skill is bi-modal There is no clear time-series pattern to the significance level of the bi-modal test, so there is no evidence that the two types of counties are converging or diverging over the 8 year time period covered by the sample. Expanding this research to longer time periods, for example across decades using Census data, would be an interesting extension. Still, the evidence composed here is strongly suggestive of two distinct types of counties, thereby suggesting a potentially large role for county skill in explaining other economic facts.

2.6 Comparisons with 1990 Census Data

Because the measures of skill used throughout this chapter are non-standard, it is useful to compare the skill measures computed with LEHD data with more traditional skill measures. In particular, a common proxy for skill is education. The two main data exercises above, exploring the extent to which geography explains variation in skill and examining the distribution of county skill measures, can be replicated using information on worker education using the 100% sample of the long form of the 1990 Decennial Census of Population. In order to be consistent with the rest of the literature, worker skill is equated with having a college degree. The college graduate dummy variable is regressed on county dummies in column 1 of table 2.5. As in table 2.2, there is significant variation in worker education by county. The R-squared of 0.0235 is also consistent with the earlier results. In column 2 the specification includes metro areas, and again both the county and metro area dummies remain significant, suggesting variation in the college worker mix of counties within metro areas. Finally, the specification in column 3 adds industry dummies. As before, both the industry dummies and the county dummies remain significant. However, with college graduation as the measure of worker skill, the R-squared on the regression is much larger at 0.1729 in the new specification and an average of 0.06 in the regressions of table 2. Clearly, college education varies more across industries than does the worker effect.

*Table 2.5: OLS of college grad on county, county*metro and metro, and county and industry dummies using PUMS 1990*

	1		2		3	
	p-value	R-Squared	p-value	R-squared	p-value	R-squared
county	0.0001	0.0235	0.0001	0.0235	0.0001	0.1729
metro			0.0001			
ind					0.0001	

While the worker effect is capturing much of the same thing as the existence of a college degree, there are some important differences. In particular, the worker effect is all of the component of a worker's wages that remains with the worker as he switches firms. In theory, the worker effect should capture characteristics of the worker that are valued in the labor market. In addition to education, it is likely that these characteristics would include ability, quality of education, etc. What the worker effect cannot capture, however, are characteristics whose value varies across firms.

One explanation of the difference between the two dependent variables is that the sorting across industries may be more closely related to education than unobservable worker skills.

In order to compute a county level skill measure analogous to the percentage of workers in the top quartile, the percentage of workers with a college degree is calculated for each county in the LEHD sample using the Census of Population. Panels A and B of table 2.6 provide two comparisons between the Census measure and the LEHD measure of county skill. Because the Decennial Census is only available for years ending in zero and only two out of the three states in the LEHD sample have 1990 data, it is not possible to directly compare the full LEHD sample used here and the Census for the same year. Therefore, two separate comparisons are made. Panel A has correlations between the percentage of college graduates in the 1990 Census and the percentage of workers in the top quartile of worker skill in the two states available in 1990 in the LEHD data. In panel B, the 1990 percentage of college graduates is compared with the county skill measure calculated from LEHD data using the three states in the LEHD data in 1991. Results across the two tables are very similar. The percentage of college graduates is highly correlated with both of the measures of county skill constructed from the LEHD data in either year the comparison is made.

Table 2.6: Comparing Census Data and LEHD Data

Panel A: Comparing 1990 Census data with 1990 LEHD data, two states

	% College grad.	Top quartile skill	Average skill
% College grad.	1	0.7312	0.7210
		(0.0000)	(0.0000)
Top quartile skill		1	0.9297
			(0.0000)
Average skill			1

P-values in parentheses.

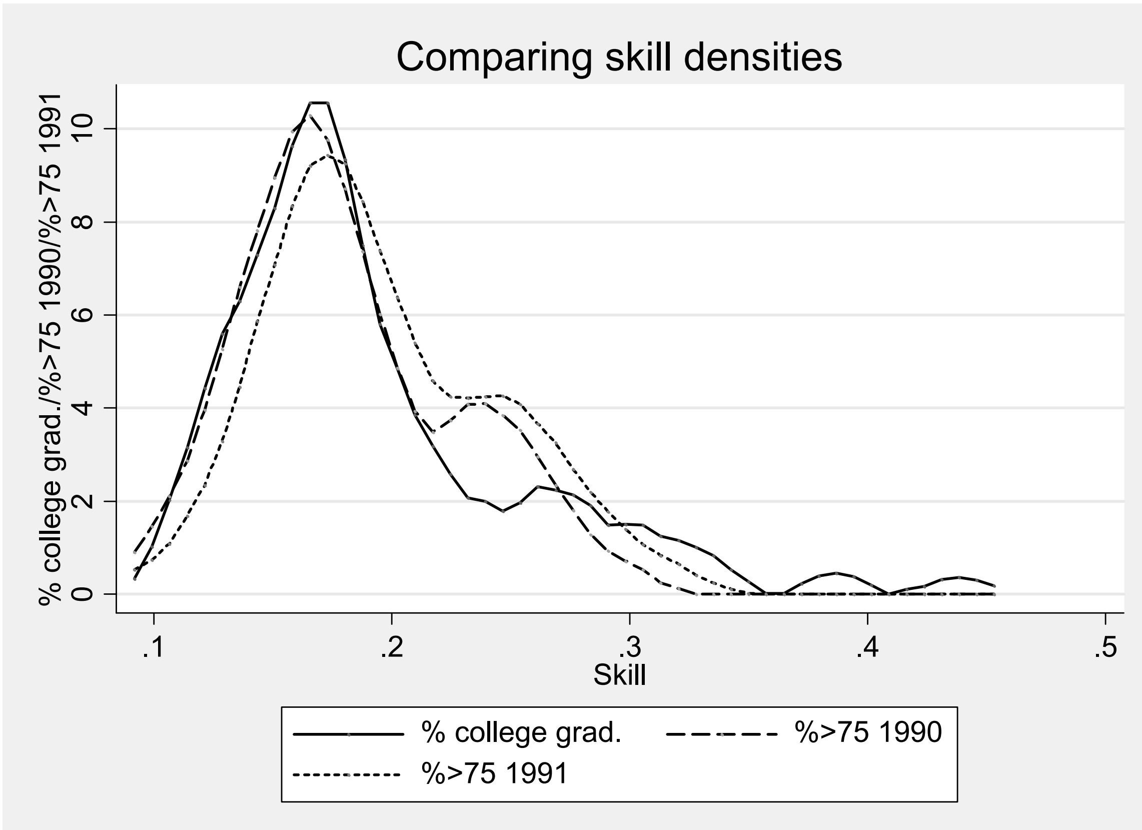
Panel B: Comparing 1990 Census data with 1991 LEHD data, three states

	% College grad.	Top quartile skill	Average skill
% College grad.	1	0.7544	0.7920
		(0.0000)	(0.0000)
Top quartile skill		1	0.9482
			(0.0000)
Average skill			1

P-values in parentheses.

In addition to the simple correlations, figure 2.7 displays kernel density estimates of the 1990 Census skill measure, the top quartile skill measure using 1990 LEHD data, and the top quartile skill measure using 1991 LEHD data. All three densities have similar shapes with a large mode in the left part of the density and a long right tail. The bimodality found in the LEHD data does not seem as prominent in the Census data, although there does appear to be a second smaller mode in the Census data. Considering that these two measures of county skill are constructed from different datasets, using different methodology and different measures of skill, the high level of correlation and similarity between the two datasets is remarkable. This evidence suggests that using the LEHD data will lead to similar conclusions on the geographic distribution of skill as using the Census. Additionally, the main advantage of using the LEHD data in this instance is the coverage that the LEHD data provides both in its near universe of workers and in the availability of data on an annual basis.

Figure 2.7: Comparing 1990 Census numbers with 1990 and 1991 LEHD numbers



2.7 Persistence in County Skill

In addition to studying the distribution of county skill at different points in time, evidence on the persistence of county skill is an important aspect in the argument of why geography matters in studying skilled workers. Other researchers have documented the high levels of mobility of skilled workers. If skilled workers are highly mobile, the skill level of a local labor market may be a transient feature, and therefore further considerations of the effect of clusters of highly skilled worker would need to account for the possibility that their effect is also temporary. The LEHD data are particularly well suited to studying counties over medium-run time frames due to the high frequency of the data.

Before looking at the county level skill measures over time, looking at firm level skill over time allows for a decomposition to give insight into the activity of workers and firms that compose the final measures. Figure 2.8 is composed of a 1% sample of all firms and, for each firm in the sample, measures the percentage of workers in the top quartile of the overall distribution in two adjacent years. The x-axis measures the 1991 value while the y-axis is the 1992 value. In addition, the thin black line is the 45-degree line while the thick gray line is the fitted value from a regression of the 1992 value on the 1991 value. As is clear from the picture, there is a tremendous amount of variation in the skill mix of firms. At one extreme, there are a cluster of firms in which there are no top quartile workers in either period, and a smaller number of firms in which the entire firm is composed of top quartile workers. In between the two extremes, the majority of the observations cluster around the 45-degree line, providing evidence that over the short time horizon there is not much movement in firm skill. In contrast, figure 2.9 plots firm skill in 1991 and 1998. Over this longer time horizon, the clustering of firms around the 45-degree line is less pronounced. In addition, the slope of the fitted line is much less steep than the 45-degree line, suggesting a convergence across firms in their skill mix.

Figure 2.8: Firm level skill, 1991-1992

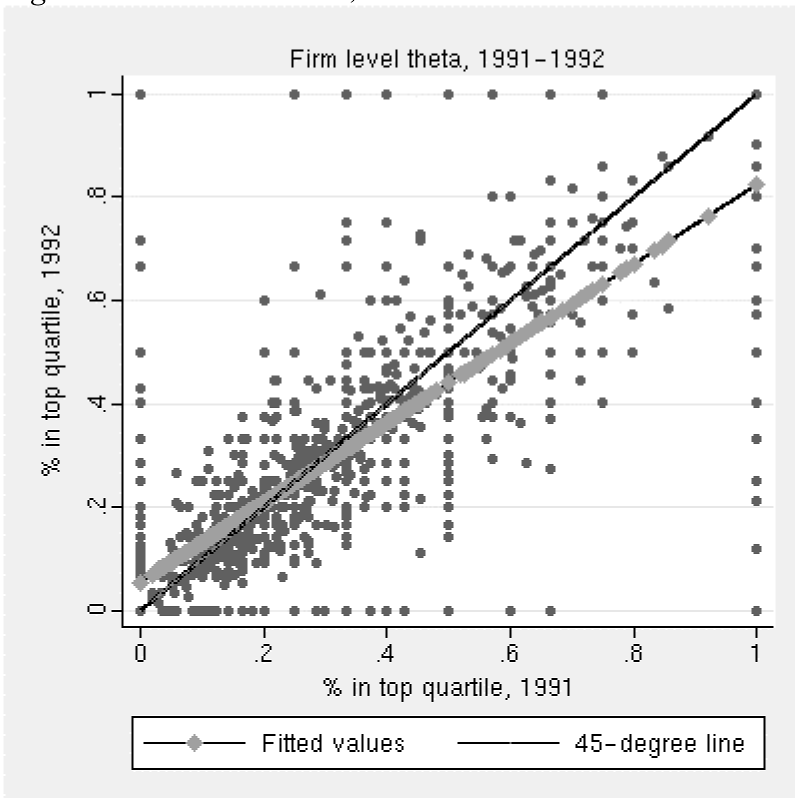
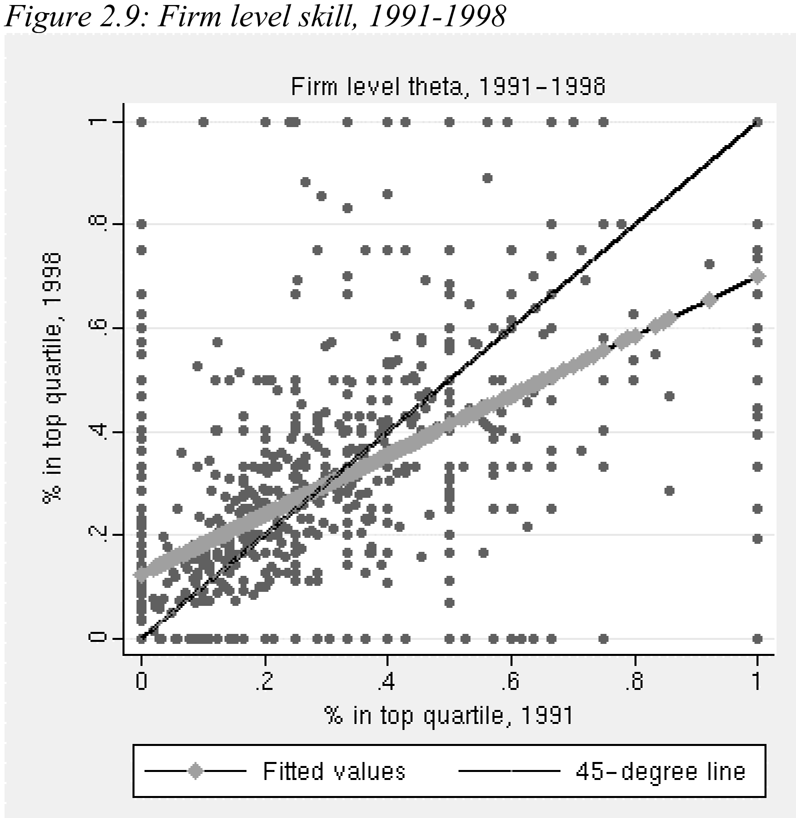


Figure 2.9: Firm level skill, 1991-1998



Figures 2.10 and 2.11 replicate the same two graphs using county level skill measures, in particular the percentage of workers in a county who are in the top quartile of the overall worker effect distribution. The most striking difference between the two sets of graphs is much smaller variation across units and the tighter concentration around the 45-degree line. Given that counties are composed of, on average, thousands of employees while firms are much smaller, this pattern is to be expected. Despite the relatively high amounts of movement in the skill mix of firms, the county level measures remain remarkably persistent. In figure 2.10, the 45-degree line and the fitted line are nearly indistinguishable. Over the longer time period, shown in figure 2.11, the parallel shift of the fitted line above the 45-degree line provides evidence that all of the counties are becoming more skilled at a nearly uniform rate.

Figure 2.10: County level skill, 1991-1992

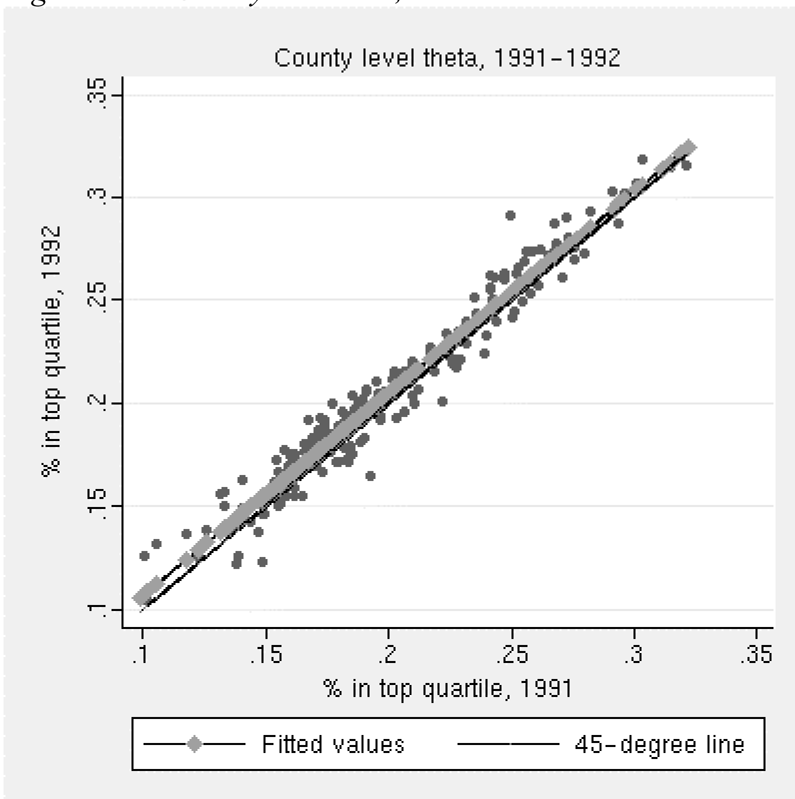
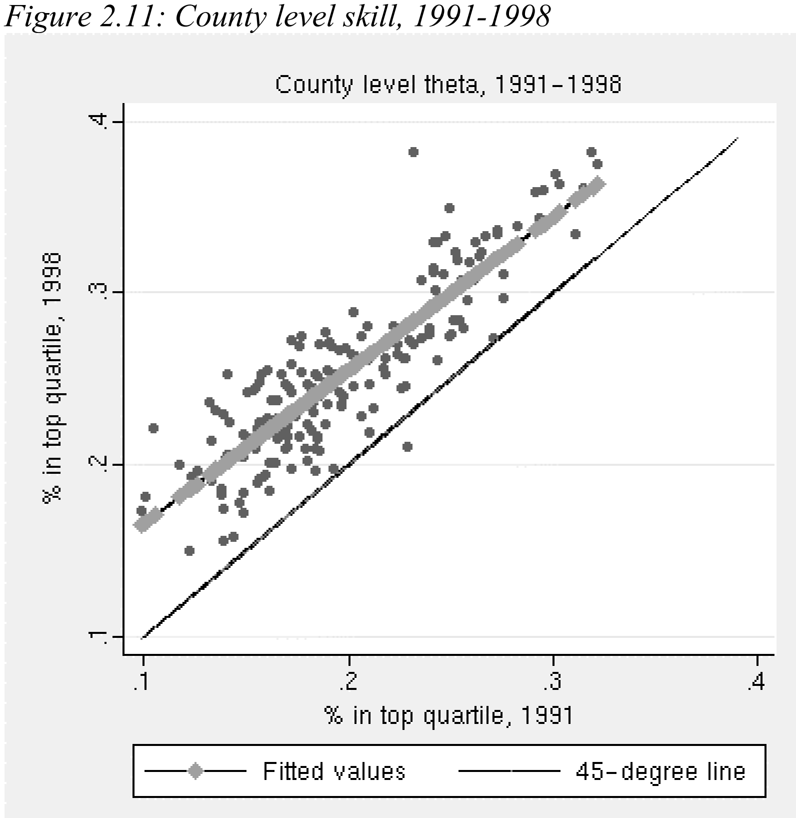


Figure 2.11: County level skill, 1991-1998



Panels A and B of table 2.7 provide more detailed evidence on the movement of counties across different quartiles in the county skill distribution by looking at transition matrices of counties. The rows divide counties into a skill quartile based on their *t* skill level. The columns divide counties into a skill quartile based on their *t+1* skill level. Panel A was created by computing a year-to-year transition matrix (*t* to *t+1*) and taking the average of each of the cells (*t*=1992 to 1997). The 1,1 cell therefore shows that on average from 1992 to 1998, 89% of counties in the bottom quartile in year *t* are in the bottom quartile in year *t+1*. On the opposite end of the spectrum the 4,4 cell shows that 95% of counties in the top quartile of the skill distribution in year *t* are in the top quartile in year *t+1*. Not only is the position of counties among the relative skill distribution persistent over time, but the more highly skilled counties are more persistent than the less skilled counties. Panel B measures the long-range transition matrix looking at county's relative skill ranking in 1992 and 1998. Again, the persistence in the bottom quartile of the skill distribution is much less than in the top quartile of the skill distribution.

Table 2.7: County transition matrices

A: Year to year transitions in quartile position, averaging 1992-1993 through 1997-1998

	1st	2nd	3rd	4th
1st	89.13	10.34	1.07	0.00
2nd	11.05	78.79	9.98	0.36
3rd	0.36	11.05	85.92	3.21
4th	0.00	0.36	3.57	94.83

B: Long-range transitions in quartile position, 1992-1998

	1st	2nd	3rd	4th
1st	68.45	27.81	4.28	0.00
2nd	29.95	47.06	21.39	0.00
3rd	0.00	25.67	64.17	10.70
4th	2.14	0.00	10.70	87.70

Given the well-studied high mobility of skilled workers, these results on the high levels of county skill persistence may initially seem surprising. However, the two facts are not inconsistent. These results suggest that although highly skilled workers may be mobile, they relocate in ways that reinforce the overall distribution of county skill. The next section begins to explore this hypothesis.

2.8 Distribution of Workers within Counties

In order to better understand why county skill is persistent given the high mobility of workers, figures 2.12 and 2.13 look at the distribution of workers by skill within counties that have been classified as high or low skill. Low skill counties are defined as those in the bottom third of the county skill distribution, with less than 17.5% of workers in the top quartile. Similarly, high skill counties are defined as those in the top third of the county skill distribution, with more than 21.2% of workers in the top quartile. Figure 2.12 shows the distributions of worker skill in these two types of counties in 1992. The distribution of workers within low skill counties is on the left, while the distribution of workers within high skill counties is on the right. While the relative position of the two distributions are determined by the cutoffs used to define them, the shape of the distributions provides more precise evidence as to what is different about high skill counties. The mode in the distribution of workers in high skill counties is both narrower and higher, suggesting that the distribution of workers in high skill counties has a much lower variance. Table 2.8 provides further evidence supporting this observation. Correlations between the mean or median and three different measures of within county skill dispersion show that higher skill counties have much lower variation in the top end of the distribution. Figure 2.13 and panel B of table 2.8 repeat these exercises for 1998 with the same conclusions.

Figure 2.12: Within county skill distribution, high and low skill counties, 1992

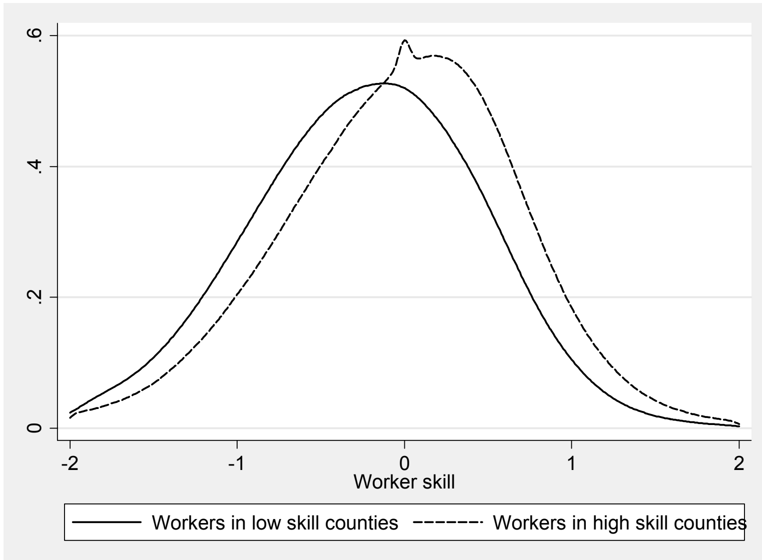


Figure 2.13: Within county skill distribution, high and low skill counties, 1998

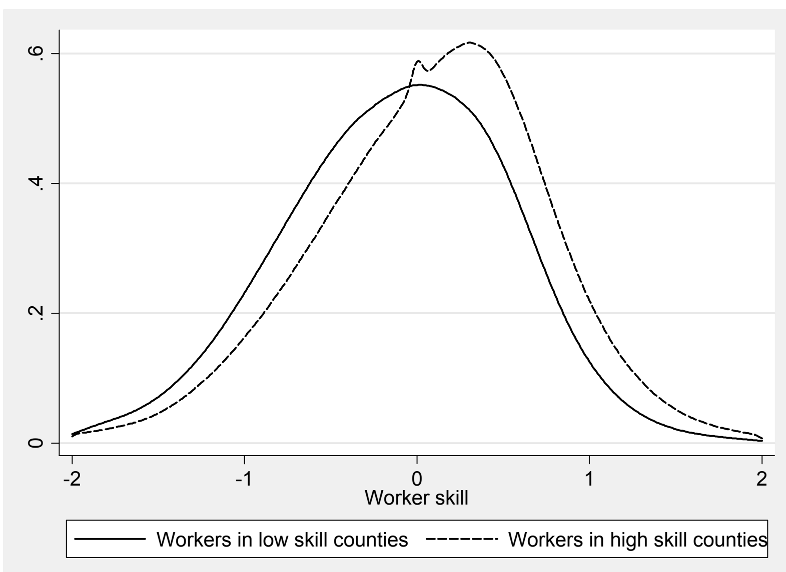


Table 2.8: Within county correlation of mean/median with dispersion measures

Panel A: 1992

	90-10	50-10	90-50
mean	-0.2624	-0.0599	-0.4812
	0.0003	0.4196	0.0001
median	-0.1390	0.1122	-0.5093
	0.0599	0.1293	0.0001

Panel B: 1998

	90-10	50-10	90-50
mean	-0.2132	0.0050	-0.4528
	0.0037	0.9464	0.0001
median	-0.0947	0.1841	-0.4829
	0.2011	0.0124	0.0001

These results suggest that the primary difference between high and low skill counties is in the upper end of the within county skill distribution. All counties have significant numbers of low skill workers, but high skill counties have high concentrations of the most highly skilled workers. As table 2.8 demonstrates, there is no significant correlation between mean or median county skill and dispersion in the bottom half of the within county skill distribution. The overall correlation between mean or median skill and the amount of dispersion within county skill is driven by the upper half of the distribution. As the final column of both panels of table 2.8 show, the correlation between mean or median and the 90-50 within county skill differential is both large and highly significant.

Figures 2.14 through 2.16 further connect the mix of workers with the distribution of county skill by decomposing changes in the distribution of workers by county skill into components due to movements of workers in and out of counties and the changing skill of counties. In the 1992 sample, 45% of workers are in the same county in 1998, 34% of the workers exit the sample by 1998, and 21% are in the sample in both years, but in different counties. The different graphs focus on a different group of workers and their distribution by county skill. Figure 2.14 looks at continuers, figure 2.15 looks at exiters and entrants, and figure 2.16 looks at movers by county skill. Each of these graphs is equivalent to looking at one group of workers in the weighted kernel density estimates for the whole sample as shown in figure 2.2. In figure 2.14, the focus is on continuers, and the figure shows the effect of the changing skill of counties by showing the distribution of the continuers with the 1992 and 1998 county skill level. In figure 2.15, the focus is on entrants and exiters, therefore the two densities are composed of two separate samples of workers. In panel A, differences between the two densities are due to exiters leaving different counties than that which entrants choose to enter. In panel B, the change in the shape of the density is due both to exiters leaving different counties than those entrants enter, and to the changing skill of counties. Similarly, in figure 2.16, the focus is on movers, and, while the sample of workers is identical in the 1992 and 1998 density, the differences between the shape of the two densities in panel A is due to the net movement of workers across counties, while in panel B the differences are due to both the net movement of workers across counties and the counties' skill levels changing.

Figure 2.14 makes clear that the large number of continuers and the similarity in their county skill in 1992 and 1998 accounts for a large part of persistence in county skill. The figure shows how the shape of the density moves due to the upskilling occurring across all counties as shown in earlier figures. As before, the distribution is of a similar shape but shifted to the right. The first mode appears to shrink while the second mode is larger in 1998. Figure 2.15 shows the contribution of entry and exit to the persistence in county skill. Entry and exit is due in part to workers entering and leaving the labor force and in part due to workers moving in and out of the sample of states being studied. Again, in panel A, their distribution in 1992 and 1998, holding county skill constant at its 1992 level, is very similar. It appears that entrants are slightly more likely to enter more highly skilled counties. In panel B, the changing skill of counties has much the same effect that it did among the continuers with a smaller first mode and a larger second mode.

Figure 2.14: Distribution of continuers by county skill
Continuers by 1992 and 1998 county skill

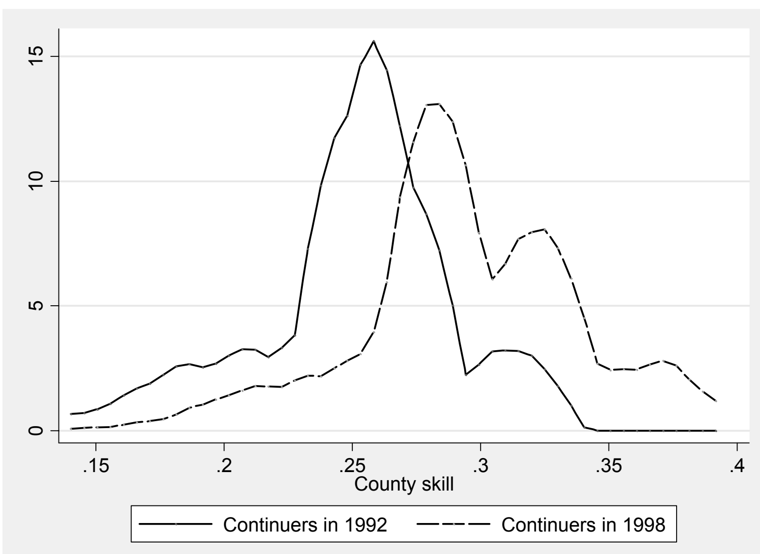
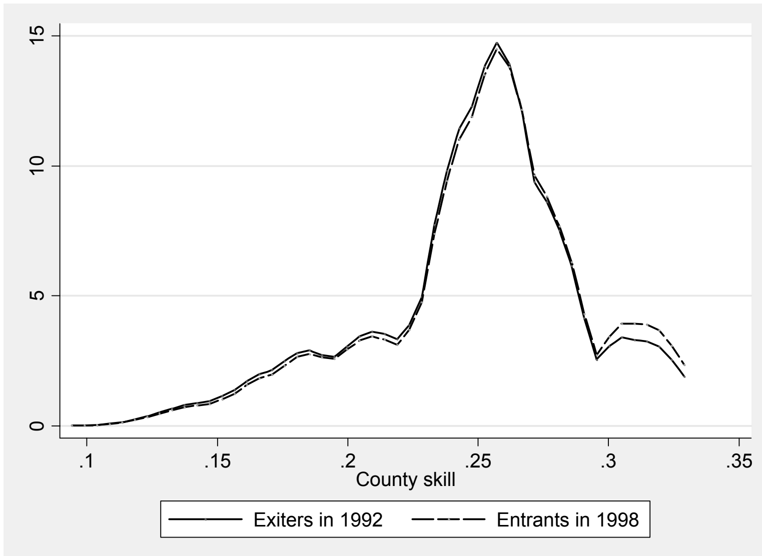


Figure 2.15: Distribution of entrants and exiters by county skill

Panel A: Entrants and exiters by 1992 county skill



Panel B: Entrants by 1992 county skill and exiters by 1998 county skill

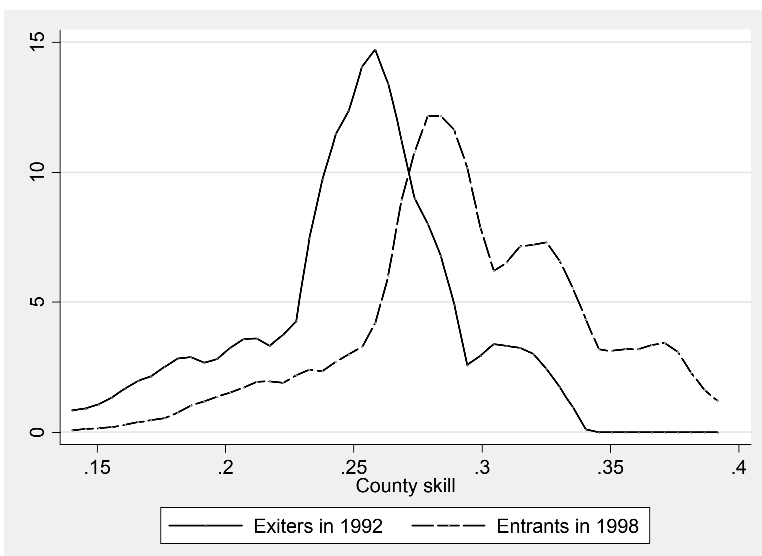


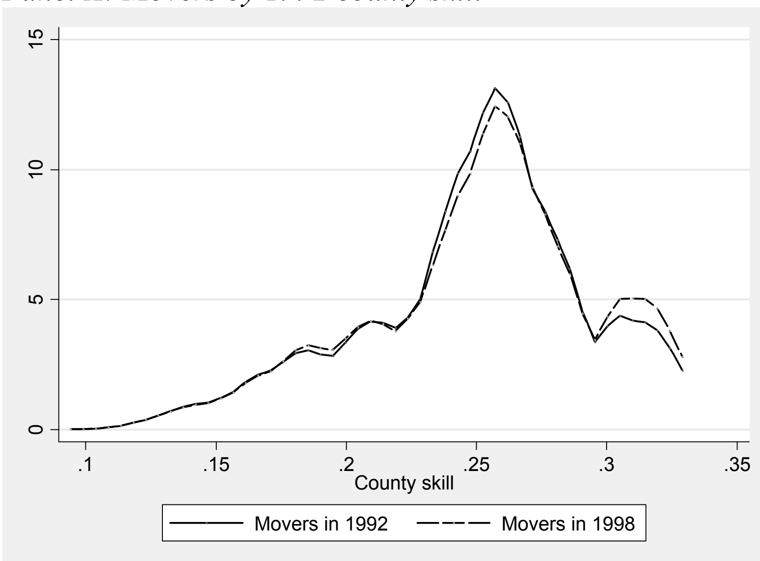
Figure 2.15 shows the contribution of movers to the persistence in county skill. This group is composed of workers in one county in 1992 and in a different county in 1998. A fraction of these moves are likely endogenous switches to a better job, better residence, etc. Another fraction is due to workers who had exogenous

separations at their last job and found employment in a new county. Unfortunately, the data does not provide any direct information that can determine the relative importance of these two different motivations behind moves[9]. Given that a fraction of the moves were likely endogenous, the similarity between the two distributions in panel A is striking. The 1998 peak is slightly lower than the 1992 peak, and there is more mass in the far right of the distribution in 1998. These differences suggest that some fraction of the workers is moving from less to more skilled counties. These patterns are also reinforced in panel B. Similar to the panel B in the two earlier figures, the distribution in 1998 is shifted to the right. In figure 15, the shape of the distribution changes a bit more between the two years. However, these differences are relatively small, and therefore even the mobile workers contribute to the persistence in county skill. These figures suggest that worker mobility had little impact on the changing distribution of county's skills, but that changes in the distribution of county's skills were important.

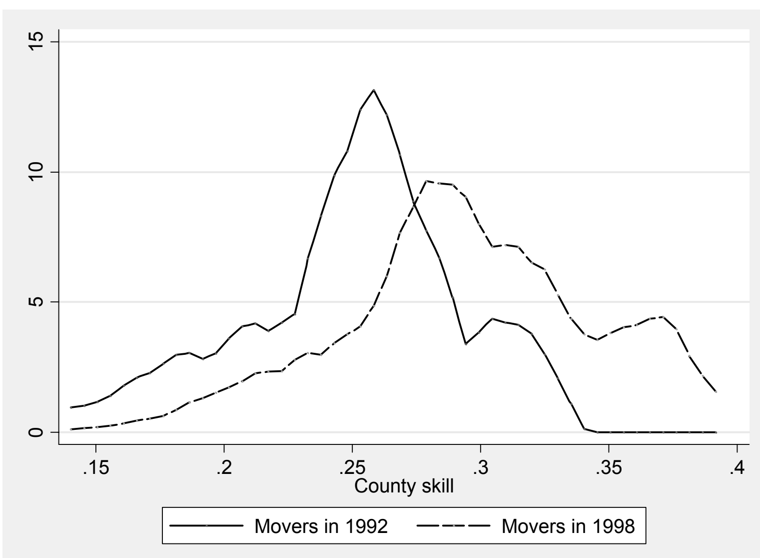
[9] In chapter 4, a sample of movers who were likely to have left their left job exogenously is determined by limiting the sample to workers who left firms that shrunk by 10 to 20%.

Figure 2.16: Distribution of movers by county skill

Panel A: Movers by 1992 county skill



Panel B: Movers by 1992 and 1998 county skill



2.9 Conclusion

Many papers in various strands of economics focus on the effect that skilled labor has on technology, wages, productivity, and other worker and firm outcomes. Despite this focus on skilled labor not much attention has been placed on the distribution of skilled workers across geography. This chapter takes a first step at filling that gap. Not only are there significant differences in the availability of skilled labor across geography, but also these differences persist over time despite the high levels of mobility among workers. While this chapter focuses on changes over the 1990s, further research on this topic could expand this research to focus on a longer time period.

Chapter 3: Endogenous Technology and Local Labor Market Skill

Despite the focus on the effect of new technologies and highly skilled workers in contributing to economic growth and explaining increasing wage inequality, the use of these new technologies and workers is not necessarily widespread. In fact, if one looks at firm level data, it becomes clear that while many firms do appear to be utilizing the latest technologies and most highly skilled workforces, there are also many firms who have not adopted the latest technologies and who have de-skilled their workforce. While the overall trends, particularly in the 1980s, seemed to favor the use of more educated workers, there is heterogeneity behind that trend. What is generating this heterogeneity? This chapter proposes one answer: the highly skilled workforce necessary to implement the latest technologies is an input into production that is very different from capital or materials. The focus on the availability of skilled labor leading to endogenous technology is not new in the economics literature.[10] However, while previous studies of this phenomenon have focused on macro-level changes in the relative supply and demand of skilled labor, this chapter utilizes variation in the availability of skilled labor across local labor markets to determine if otherwise similar firms invest differently in high technology capital depending upon the locally available skill mix.

Because firms must often decide how they would like to conduct their business before they know whom they may be able to hire, they must make their investment decisions based on the type of worker that they expect to be able to hire.

---

[10] See Acemoglu (1998), Kiley (1999), Albrecht and Vroman (2001).

Workers, while mobile, tend to locate near other similar workers, thereby limiting the ability of a firm in a low skill area to attract a high skill workforce. Firms who attempt to recruit more skilled workers potentially face large search costs both in recruiting workers and in the cost of having capital lay idle. The model constructed captures these ideas and provides a framework for the later empirical work. The model is a two-period matching model based on Acemoglu (1996). In the first period, firms know the distribution of workers in their local labor market, but not the worker with whom they will match. Based on this information and their knowledge of their firm type, firms make an investment decision. Worker skill is predetermined in the model. In the second period, workers and firms meet and production takes place. This model sets up the two key equations for estimation: a wage equation and an investment equation. The wage equation takes advantage of a large data set and helps to quantify worker and firm heterogeneity. The investment equation uses results from the wage equation in combination with other firm data to directly address the question laid out above: Do firms consider the skill mix of the local labor market before making a technology investment decision?

The key difference between this research and the existing endogenous technology literature is that this chapter exploits variation in a firm's investment decisions in a single cross section, while previous research has focused on time series trends in the development and implementation of new technologies. Identification in the cross section of the effect of local labor market skill on the decisions of firms is complicated by the fact that the location decisions of both firms and workers are potentially endogenous. This problem is circumvented here by focusing on an exogenous technological shock, the growth of computer technology between the early 1980s and 1990s. The average establishment in the sample used in the empirical work is greater than 10 years old and therefore made its location decision over 10 years ago, based on the technologies and local labor market skill available at that time. In the ten years between when the average firm made its location decision and the cross section in 1992 in which computer investment data is available, computer investment per worker at manufacturing establishments increased by a factor of 20. In contrast, the skill mix of the local labor market remains persistent over time. The introduction of a new technology to a set of older firms allows for identification of the effect on labor market skill on firm's investment decision.

The data needed to address this question empirically require detailed knowledge on firms, workers, and their interaction. For establishments, information on firm type, the amount and type of investment, and location is needed. On the worker side, one must know the skill level of all employees in a local labor market and the skill level at each establishment. A newly developed linked employer-employee data set available at the Census Bureau makes this research possible. State level universe files containing employers and employees allow one to decompose wages into an explained component due to observable time-varying worker characteristics, a fixed worker heterogeneity term, and a firm heterogeneity term. This, in combination with links to the 1992 Census of Manufactures, which provides information on establishment level investment, makes it possible to answer the question set out earlier. Estimates of the effect of endogenous technology predict that a one standard deviation increase in local labor market skill will lead to roughly a

10% increase in technology investment. These results are robust to a series of different specifications including different measures of investment, local labor market skill, and definitions of the local labor market.

3.1 Background

3.1.2 Endogenous Technology

In the late 1970s, the United States saw a large increase in the relative number of college graduates. Traditional factor-demand analysis would suggest that this would be followed by a decrease in the relative wage of skilled workers. On the contrary, throughout the 1980s there was an increase in the price paid for college educated workers. Researchers studying the increased wages paid to more skilled workers have considered many possible explanations including shifts in demand and the effects of trade. Focusing here on the former of the possible explanations, the use of skill biased technology increased the demand for skilled workers faster than the supply of skilled workers was growing, thereby leading to a rise in wages for highly skilled workers while their numbers were also increasing.

Endogenous technology provides a link between these two events. In reaction to the increase in the availability of more skilled workers, businesses began investing more in technologies to utilize this newly available human capital. Models have incorporated this endogenous technology choice in different ways. Acemoglu (1998) and Kiley (1999) set up models in which the economy has two sectors, research and production. The research sector chooses to develop the technologies that will command a high price and that will be demanded by a large number of firms in the production sector. The increase in the availability of skilled workers in the late

1970s, therefore, increased the number of production firms that could potentially use skill-biased technology. The research sector's incentives shifted as the supply of skilled workers increased. As more skill-biased technologies became available, more production firms utilized these technologies, and simultaneously demanded more skilled workers.

A similar set of models (Acemoglu (1999), Albrecht and Vroman (2002), Eudey and Molico (2001)) relies on the fact that firms must commit to a type of investment before meeting workers. Each of the above papers assumes that there are two types of vacancies defined by their technology, and two types of workers defined by skill. Firms decide which type of vacancy to create by determining the probability of meeting a worker appropriate to that technology. The types of vacancies created by firms critically depend upon the skill mix of the workforce. A workforce with a mix of high skill and low skill workers leads to a pooling equilibrium in which only one type of vacancy is created, while a workforce with a greater proportion of skilled workers will lead to a separating equilibrium with some vacancies specifically created for high skill workers. These papers then study the effects these different equilibria have on either wage inequality or productivity within the economy. A similar model with a continuum of skilled vacancies, defined by amount of technology investment, is developed below.

3.1.2 Firm Heterogeneity

Beyond the previous theoretical research on endogenous technology, there is also a fair amount of empirical evidence that firms exercise a choice in the type of technology that they use. Haltiwanger, Lane and Spletzer (2000), using a similar data set to that used here, look at worker characteristics within firms over time. They find considerable heterogeneity across firms in their choice of worker mix even after controlling for detailed industry and other observable firm characteristics. In addition, these firm differences persist over time, suggesting that they are the result of a firm choice and not the result of error on the part of the firm or noise in the data. Abowd, Haltiwanger, Lane, and Sandusky (2001), again using similar data, look at the connections between worker characteristics and firm characteristics. In their analysis of Illinois over the 1990s, they find in a cross sectional analysis that firms with greater levels of technology also have more skilled workforces, and that over time, firms that increase their use of technology also increase their use of skilled workers.

Bresnahan, Brynjolfsson and Hitt (2002) look jointly at worker skill, workplace organization, and technology investment. In their work, they find that firms that invest heavily in information technology not only have more highly skilled workforces, but additionally they have less centralized workplace organizations. The combination of all three of these factors suggests that a firm deciding to invest in high technology must also be willing to invest heavily in changing its workforce to fully take advantage of the new technology. Among their most interesting results, as it pertains to the research set out here, is a regression of log output on labor, capital and a series of four dummies: firms with both high technology and high worker skill, high technology but low worker skill, low technology and high worker skill, and finally low technology and low worker skill. Not surprisingly, the dummy for the high-high mix is large and positive. However, the low-low combination represents the next largest coefficient. These results provide further evidence that firms have a viable alternative to doing business using the latest technology. A combination of workers and capital that complements each other has higher productivity than either using new technology or highly skilled workers singly.

In addition to the cross-sectional heterogeneity in technology usage mentioned above, the time-series variation in relative technology-skill complementarity highlights the point that it is not necessarily optimal for all firms to be using a high skill/high technology mix. Goldin and Katz (1998) examine the evolution of technology-skill complementarity and find that while more recent advances in technology have been skill biased, new technologies adopted in the 19th century were, in fact, biased toward unskilled labor. The shift from artisanal shops to factories using assembly lines may have been an advance in technology, but it certainly involved a reduction in the amount of skilled labor required. They argue that, although skilled labor and capital are complements within the implementation and maintenance of a given technology, they may be complements or substitutes among different technologies.

3.1.3 Local Labor Markets

Finally, a couple of papers study different aspects of the relationship between local labor market characteristics and firm characteristics. Moretti (2002) estimates establishment production functions including the change in the college share outside of the establishment's industry to identify the extent of human capital spillovers. In order to control for various unobservable factors that might influence both establishment productivity and the share of college graduates in the local labor market, Moretti controls for plant, industry by year, and state by year fixed effects. Therefore, identification of the human capital spillovers comes from changes in the college share variable that are correlated with changes in productivity for establishments that survive from 1982 to 1992. Moretti finds that human capital spillovers are responsible for a 0.1% increase in output per year during the 1980s. The key difference between Moretti's model and the one outlined here is that, in Moretti's model, it is assumed that firms and workers are perfectly mobile. Variation in the amount of human capital spillovers across areas continues to exist in equilibrium due to variation in the price of the untraded good, land. In the model below, variation across local labor markets is driven by the limited mobility of workers and establishments. In his empirical work, Moretti also relies on a different data source in which it is impossible to control for the skill of the firm's workers.

Fallick, Fleischman, and Rebitzer (2003) study the relationship between local labor market worker mobility and characteristics of agglomeration economies and of Silicon Valley in particular. They note that one key aspect of Silicon Valley that makes it different than other examples of agglomeration economies is the existence of a California law that makes it impossible for employers to enforce non-compete agreements. This law, in combination with the modularity of technology being developed in Silicon Valley, has led to knowledge spillovers via unusually high mobility of workers between establishments. Their results suggest that Silicon Valley may be a special case of endogenous technology due to the exceptionally common transfer of employees between firms.

3.2 The Model

The endogenous technology model developed here is a two period matching model similar to Acemoglu (1996). While Acemoglu's model focuses on social increasing returns to human capital, the search frictions in the labor market within his model can also be shown to lead to endogenous technology. Within the model developed here, the economy consists of a single autonomous local labor market and exists for two periods. While a full model of this economy would include multiple local labor markets and would allow for endogenous worker and firm mobility, here the larger economy can be thought of consisting of many local labor markets operating independently of one another. Workers vary in their skill level and are exogenously distributed across local labor markets. An individual worker's skill level is determined outside the model. This suggests that workers cannot adjust their skill level within the time frame of a firm choosing its investment level, i.e. a worker with a high school degree cannot obtain a college degree in the time that a firm requires to choose and implement new technology. For simplicity in exposition, there are only two types of worker skill in the model, a high skill level and a low skill level. Extending the model to a continuum of skill types would not affect any of the important results of the model. Firms vary in their predetermined type and in the amount of their capital investment. There is a fixed marginal cost of capital equal to μ. Each firm employs only one worker.

The basic timing of the model is as follows. In period one, firms observe the distribution of workers in their local labor market, but do not know the type of worker with which they will match. The firm decides on a level of capital investment. In the second period, firms and workers are randomly matched to each other. Firms and workers must decide whether to continue with the match and produce or to remain idle for the period. Search costs create quasi-rents for the firm and worker within the match. If production takes place, returns to workers and firms are determined via a Nash bargaining solution. Workers receive a fraction *B* of match surplus and firms receive *1-B*. Match surplus is equivalent to output in this model because the firm's first period investment is a sunk cost. Both workers and firms have zero opportunity cost, and it is assumed that output is nonnegative, implying that workers and firms will accept any division of match surplus and production will occur with all matches. Within this model, the Nash bargaining solution can be shown to be a general solution. The specific case in which *B* is equal to α, labor's share of income under constant returns to scale and Cobb-Douglas production, is equivalent to assuming that factors receive their marginal product.

Production takes place in worker/firm pairs

(1) $$Y_{ij} = A_j h_i^{\alpha} k_j^{1-\alpha}$$

where α is a value between 0 and 1, h_i is worker type *i*'s human capital, k_j is firm *j*'s capital investment, and A_j is firm *j*'s idiosyncratic term meant to capture firm type, i.e. managerial ability, workplace organization, etc. Worker type *i* is equal to either 1 or 2, low skill level and high skill level respectively. A fraction ρ of workers are type 2.

In the second period, the realized returns for a worker *i* and firm *j* are

(2) $$W(h_i, k_j) = Bh_i^{\alpha} k_j^{1-\alpha}$$

(3) $$R(h_i, k_j) = (1-B)h_i^{\alpha} k_j^{1-\alpha}$$

Workers and firms receive share *B* and *1-B* of the match surplus. In the first period, firms and workers know the distribution of worker and firm types. Therefore, under rational expectations, workers' and firms' expectations of their second period earnings are the expected value of the ex-post earnings.

(4) $$W(h_i, \{k_j\}) = Bh_i^{\alpha}(\int A_j k_j^{1-\alpha} dj)$$

(5) $$R(h_1, h_2, k_j) = (1-B)A_j((1-\rho)h_1^{\alpha} + \rho h_2^{\alpha})k_j^{1-\alpha}$$

The random matching process that occurs in the second period translates into uncertainty in first period expected returns for both workers and firms. Because workers don't know either the type or the capital intensity of the firm with which they will match, their expected returns depend on the entire distribution of firm types. Similarly, because firms don't know the skill level of the worker with whom they will match, their expected returns depend upon the distribution of worker types. It is this uncertainty, coupled with the fact that firms must make their investment decision before meeting a worker, that is key to generating endogenous technology in the model. Firms therefore make their investment decision in period one by equating the marginal expected return to capital investment to the marginal cost of capital.

(6) $$(1-B)(1-\alpha)A_j k^{-\alpha}((1-\rho)h_1^{\alpha} + \rho h_2^{\alpha}) = \mu$$

This equation can be solved to find a closed form solution to the firm's investment decision.

(7) $$\Rightarrow k = \left(\frac{(1-B)(1-\alpha)A_j}{\mu}\right)^{1/\alpha}((1-\rho)h_1^{\alpha} + \rho h_2^{\alpha})^{1/\alpha}$$

Comparative statics show

$$(8) \qquad \frac{\partial k}{\partial A_j} > 0 \qquad \frac{\partial k}{\partial \rho} > 0$$

Therefore, the model suggests two main factors that will affect a firm's investment decision. A_j captures a firm's ex-ante heterogeneity in the model. The higher A_j is, the more likely a firm is to heavily invest in capital. Empirically A_j represents factors such as managerial ability or corporate culture that are inherent, semi-fixed characteristics of the firm.. ρ is the proportion of workers in the local labor market that is highly skilled. Firms located in more skill intensive areas should, according to the model, invest more in capital.

There are two key features of the model that lead to endogenous technology: capital-skill complementarity and search frictions. Capital-skill complementarities in the production function imply that a more skilled worker raises the marginal benefit of investing in capital. Therefore, if a firm knew with certainty the skill of the worker to which it would match in the second period, the optimal investment for the firm would increase with the skill of the worker. However, due to search frictions, the firm does not know the type of worker with which it will match when it makes its investment decision in the first period. Therefore, it bases its investment on the expectation of what that worker will be. If there are a large number of highly skilled workers in the local labor market, i.e. a high ρ, the probability that the firm will match with a highly skilled worker increases, and therefore the firm invests more heavily. If there were no search frictions, capital-skill complementarities would lead to a relationship between the skill level of the firm's labor force and the firm's technology choice. However, under this alternate scenario, there would be no relationship between a firm's technology choice and the skill mix of the local labor market after controlling for the firm's skill mix. Therefore, the empirical work can directly test if search frictions are generating endogenous technology.

3.3 Data and Measurement Issues

3.3.1 Workers

All of the data used in this research are part of the Longitudinal Employer-Household Dynamics program at the Census Bureau. The data on workers is identical to that used in the previous chapter. Information on workers comes from the Unemployment Insurance wage records for the selected three states.[11] These files contain person identifiers that allow one to track a worker's earnings within a state over the available period.[12] The data also contain firm identifiers that allows for an exact link between the UI files and other data sets. The UI wage records contain virtually all business employment for the states included in the analysis, creating a final sample size of 198,644,076 observations representing 37,875,250 people and 3,989,740 firms. The disadvantage of using the UI wage data to characterize workers is the limited demographic information available. Within the Census bureau, this problem has been partially overcome by combining the UI wage data with other administrative data containing information on date of birth and gender. Additionally, as will be discussed in more detail in the estimation section, the panel aspect of the data allows one to decompose wages into a worker effect, firm effect, explained component, and a residual.

[11] Three states were chosen on the basis of time-series availability at the time of project inception from the larger set of available states. A full list of the states available and additional information about the LEHD program is available at lehd.dsd.census.gov.
[12] Time periods vary by state, with the latest start date at 1991 and the earliest end date at 1998.

The local labor market throughout this chapter will be defined as county of work for employees. There is some limited county of residence information also available in the data. However, this data is only available for 1999 and forward. Additionally, there are many reasons why the county of work is preferable. The local labor market for this model is defined as the region around a firm from which it can draw potential workers. Given that workers have varying preferences for place of work depending upon disutility of commuting and amenities of particular areas, the places where a worker may potentially be interested in working would best be defined by the current place of work rather than the place of residence.

The county, rather than the metropolitan area is used to define the local labor market, largely because the measure of the local labor market is defined by where individuals work rather than where they live. The metropolitan area definitions are created to capture a center of economic activity and the surrounding areas from which workers commute. However, because this analysis is based on where individuals work, the metropolitan area definitions are less relevant here. Also, as shown in the previous chapter, there is statistically significant variation in county skill within metropolitan areas. Given that local labor market skill is the key independent variable, the greater variation will aid in identification of the effect of local labor market skill. Despite the arguments given for using the county of work to define local labor market skill, as a robustness check, metropolitan area skill is also used as a measure of local labor market skill in the investment equation.

3.3.2 Investment

Information on establishment investment comes from the 1992 Manufacturing Census. In 1992, and Census years prior to 1992, the manufacturing Census included a series of detailed questions on capital expenditures for the Annual Survey of Manufactures (ASM) sample within the Census. The ASM disproportionately samples large establishments and provides sample weights to make the data representative of all establishments in manufacturing.[13] In addition to the sample weights, I also use the total value of shipments to weight the investment equation results in order to make the analysis representative of overall economic activity. The key measure of investment used here focuses on expenditures on new computer equipment. In order to create a measure of computer investment that is standardized across establishments of various sizes, two transformations of the computer investment data are used. In the first measure, the computer investment data is divided by the number of employees at the establishment. A second measure is created by dividing computer investment by total equipment expenditures to create a measure of the technology bias in an establishment's investment decision.

An additional measurement issue arises because the model has only very limited dynamics, and the data on computer investment is only available in a single cross section. The estimation therefore implicitly assumes that computer investment is in a non-durable good. There is some support for this assumption, although deciphering the expected lifecycle of computer equipment is a difficult task. While

[13] The ASM sample consists of a certainty and a sampled component. The certainty component includes all establishments in companies with greater than $500 million in shipments in 1987, accounting for 18,000 establishments, and all establishments with greater than 250 employees, accounting for 10,000 establishments. The remainder of the 27,000 ASM establishments is sampled on the basis of establishment size and industry-level year-to-year volatility in shipments.

computer equipment may not be a non-durable good, existing research shows that it has a short life span that seems to grow shorter as time passes. Within the data, the main problem with modeling computers as a non-durable is the treatment of establishments with zero investment in the data. While these establishments may be low technology firms, they may also be firms who invested heavily in the previous period. The information available in the data makes it impossible to distinguish between these two groups, and approximately half of the firms included in the analysis have zero computer investment.

The group of zero investors presents some potential problems for estimation. Two aspects of computer investment argue in favor of a large number of the zeroes being low-technology firms. The first is, as mentioned above, that the average life span of computer equipment is short. If the life span of a computer is around a year, then there is no problem; the zeroes most likely represent low-tech firms. The second aspect of computer investment that argues in favor of the zeroes being low-tech firms is the sharp increase in computer investment by manufacturing establishments between 1982 and 1992. Dunne, Foster, Haltiwanger, and Troske (2000) have measured the mean level of computer investment per worker in manufacturing to be $40 in 1982, $140 in 1987 and $830 in 1992. These numbers suggest that not many manufacturing firms had invested heavily in computer equipment in prior years. Additionally, the potential uses of computer equipment in manufacturing is increasing over the time period, so that it is likely that a even a firm that invested in previous periods would have to invest in the current period to remain a high technology firm.

In estimating the investment equation, the cross section of data is in essence a mismeasured version of the ideal computer investment measure. As a robustness check to the base estimation, the zero investors are grouped with small investors. Grouping the firms in this fashion assumes that truly high-technology firms invest significantly in every time period, but that it is impossible to distinguish between non-investors and small investors using the available cross-sectional data. A probit is then used to determine to what extent county level skill predicts whether firms are high technology firms, and these results are compared with those from the estimation based more directly on the model above.

Both the Census investment data and the LEHD data are at an establishment level, but the two datasets have different establishment identifiers. The link between the data on investments and workers is available at an Employer Identification Number (EIN) level.[14] The EIN is an administrative unit that for a multi-unit business may be broader than an establishment and as large as a firm. For a single unit firm, the EIN is identical to the establishment. Therefore, the firm level characteristics calculated using the UI data, such as the firm effect and firm level human capital, are at an EIN level so that they an be matched to the investment data. However, because an EIN can have establishments in multiple counties, the EIN is not an acceptable level of aggregation for the firm investment data given that a goal of the analysis is to test the connection between an establishment and the local labor market skill of the county in which it is located. In order to handle this problem, the

[14] While the EIN is the only common identifier between the two datasets, there are other variables in common between the datasets, such as SIC, employment and payroll that could be used as an additional restriction on the match. Using additional restrictions, i.e. requiring that the EIN has the same SIC code in both datasets, has little effect on any of the results.

investment data is aggregated to an EIN-county level, so that the investment numbers reflect data from all the establishments for a given EIN within a given county. This level of aggregation circumvents the problem of trying to match establishments in the UI data to establishments in the Census data when the common identifier is an EIN, a problem that affects about 5% of the observations.[15] In addition, in the analysis here, the key variable is the skill level of the county, thereby making an establishment level match unnecessary. Throughout the rest of this analysis, the EIN-county unit of aggregation will be referred to as the establishment.

3.4 Estimation

The model assumes that wages are determined via rent sharing. This assumption in combination with the form of the production function laid out in the model implies that wages are a function of worker and firm characteristics. Because worker and firm heterogeneity influence the local labor market skill mix and firm investment decisions respectively, it is necessary to characterize this heterogeneity before testing the main implications of endogenous technology laid out in the model. Population based estimates of worker and firm heterogeneity can be determined via a wage decomposition. Taking advantage of what is available in the data, the equation used in estimation is

$$w_{ijt} = \theta_i + X_{it}^{'}\beta + \psi_{j(i,t)s} + \varepsilon_{ijt} \tag{9}$$

This decomposition of wages is a variation on the methodology developed by Abowd, Kramarz, and Margolis (1999), that was also used in the previous chapter.

[15] With computer investment as a percentage of equipment investment as the dependent variable, 95% of the observations are single establishment EIN-county level observations. With computer investment per worker as the dependent variable, 96% of the observations are single establishment EIN-county level observations.

Reiterating the key components of the equation, human capital is captured in the fixed worker effect, θ, and a quadratic in experience captured within *X*. Firm characteristics are captured in the limited time varying firm effect, ψ. The remaining variables contained in *X* are a series of gender by year by labor force attachment status dummies added to control for assumptions necessary to create an annualized wage measure from the quarterly earnings data. These variables also control for the observable time-varying characteristics.

The firm component of a worker's wages depends upon the firm's type and the type of technology being used at the firm. Potentially, firm technology could vary period to period, thereby suggesting fully time-varying firm effects. However, capital investment data is only available in two years, 1992 and 1997.[16] Therefore, a measure of firm type is needed only in the periods prior to investment when the firm is making its investment decision. The fixed firm effect is inappropriate because it will be contaminated with the current investment decision. The limited time-varying firm-effect is a compromise between a fixed firm effect and a fully time varying firm effect. While the fixed time effect is not compatible with the theoretical model laid out above, the fully time varying firm effect would be identified only off of the observations for a firm within a given year. The limited time-varying firm effect contains the best attributes of either strategy, since it compatible with the theory and it is identified off of observations from multiple years. The three sub-periods chosen are as follows: 1991 and earlier (the period before the first investment), 1992 through 1996 (the period before the second investment), and 1997 on (the final period).

[16] Investment data are generally only available during Census years. In this chapter, the 1992 computer investment data and the 1992 and 1997 equipment investment data are used. Computer investment data are not available in 1997.

Identification in the limited time-varying firm effect model requires additional restrictions on the other covariates. In particular, one time effect must be suppressed within each separate sub-period. The time-varying firm effect is generated by creating three separate identifiers for each firm, corresponding to the three different time periods. The three firm effects are identified separately by the observations for that firm within that time period only. Although the firm effects are not generated in a manner which forces them to be correlated over time, analysis of the results from the wage equation show a correlation of approximately 0.7 between firm effects for the same firm across adjacent sub-periods. These results suggest that a component of the firm effect is fixed, and, therefore, that the limited-time varying firm effect is an appropriate way to capture firm heterogeneity in the periods prior to investment.

Due to the large sample sizes, the wage equation is estimated separately for each state using the conjugate gradient methodology as explained in Abowd, Creecy, and Margolis (2002). The results are then pooled across the states included in the analysis, properly adjusting the person and firm effects to control for differences in state level mean wages. Identification of the person and firm effects is then determined by applying a grouping algorithm to the pooled state data. A connected group is determined by taking a firm, then pooling all of the employees of that firm, then taking all of the firms those employees ever worked at, then pooling all employees at the larger set of firms, and so on. The connectedness of the data is generated by the mobility of workers across firms. Within each connected group all but one person or firm effect is identified. For the group of states included in this analysis, 99.9% of the observations are in one connected group.[17] In practice, the identification restriction is applied by setting the mean of the person and firm effects equal to zero for each connected group.

The second equation to be estimated directly tests for endogenous technology at the establishment level. Comparative statics of the model suggest that a firm's technology decision should be increasing in firm type and, most importantly for endogenous technology, in the proportion of workers in the local labor market who are highly skilled. Combining information available from the ASM along with worker and firm heterogeneity terms from the wage decomposition, endogenous technology can be directly tested using

(10) $$k_{jt} = \phi_0 + \phi_1 \psi_{jt-1} + \phi_2 s_{lt-1} + \upsilon_{jt}$$

Taking the components of the equation one by one, k_{jt} is the technology investment variable; ψ_{jt-1}, the limited time-varying firm effect from the wage decomposition (9), is used as a proxy for firm type; and s_{lt} is one of the measures of local labor market skill described below. In the model, k_{jt} is the stock of physical capital at the firm, and because the firm only exists for one time period, it is equivalent to investment in capital. Empirically, k_{jt} is measured as either computer investment per worker or computer investment as a fraction of machinery investment. Focusing on computer investment is ideal for two reasons. The first is that capital skill complementarities are particularly strong with computers. Additionally, investment in computers can more readily be interpreted as a technology choice.

[17] This group represents 99.1% of all workers and 89.3% of all firms in the pooled three state sample.

Computers are a type of investment that is relatively homogenous across industries, and therefore, computer investment is likely to capture similar interactions between firms and workers across the range of industries within manufacturing. While the model laid out is not used to predict what the magnitude of ϕ_2 should be, the model does clearly suggest that ϕ_2 should be positive.

As mentioned above, the estimated firm effect from the wage decomposition is used as a control for firm type in the investment equation. While the model suggests that the firm effect would capture firm type, the firm effect is likely contaminated with other characteristics of the firm. In particular, given the assumption that wages are determined via rent sharing, the firm effect is also capturing aspects of the firm's technology. Because lags of the firm's technology are most likely correlated with their current technology investment decision, it is important to test the extent to which this misspecification might possibly bias the results. This issue is empirically addressed by including a lag of the capital stock per employee, or capital intensity, as an additional control in the investment equation. A more complete discussion and further justification for the empirical strategy can be found in Appendix A. As is shown in the results section, this specification problem does not appear to affect greatly the coefficient on county skill.

3.5 Characterizing Human Capital

Table 3.1 summarizes the results of estimating equation 9 with limited time varying firm effects. These results are the same as table 2.1 in the previous chapter. Reviewing the key results, looking across the first row, the correlation of log wage with the worker effect is 0.56 and the correlation with the firm effect is 0.50. These results suggest that worker and firm effects are equally important in explaining the variation in log wages. The covariance between the worker and firm effects at the individual level is positive, although small in magnitude at 0.07. The positive covariance between worker and firm effects suggests that high skill workers are more likely to be at employed at high wage firms. The results from estimating the wage equation are then used to quantify worker and firm heterogeneity. For the worker, two measures of human capital are used. The first measure uses only the worker effect, θ_i, as was used in chapter 2. The fixed worker effect in this model reflects any fixed characteristic of the worker that affects his wages. Although no individual level comparison of the worker effect and more traditional measures of skill are done in this chapter, Abowd, Lengermann, and McKinney (2003) have found that there is a positive correlation between the worker effect and education. The second measure is constructed as follows

$$s_{it} = \theta_i + \widetilde{X}_{it}\beta$$

where θ_i is the fixed worker effect, and $\widetilde{X}$ is the subset of X that contains the quadratic in experience. This second measure of human capital captures returns to experience in addition to the worker effect.

Table 3.1: Results from Wage Regression

	Log wage	Worker effect	Firm effect	XBeta	Residual
Log wage	1	0.5643	0.4958	0.2294	0.4207
Worker effect		1	0.0655	-0.4740	0.0000
Firm effect			1	0.0355	0.0000
XBeta				1	0.0000
Residual					1

Throughout the rest of the chapter it is necessary to use functions of the individual level skill to define either firm or local labor market level skill measures. While the thetas estimated at the individual level are inconsistent, these functions of theta aggregated to the firm or county level are consistent.[18] Two alternative measures of skill are used interchangeably throughout the chapter. The first is a simple average of either of the human capital measures within the firm or local labor market, θ_l^{mn}, where the subscript denotes the unit of observation and the superscript denotes the method of measurement. The second measure, $\theta_l^{>75}$, calculates the percentage of workers within the firm or local labor market that are above a given threshold of the overall three-state distribution of the human capital measure for a given reference year, chosen as 1992. For the calculations here, the threshold chosen is the 75th percentile, and therefore the measures represent the top quartile of worker human capital.[19] These two measures capture different aspects of the skill distribution that are differentially valued by firms. In a later section of the chapter, the skill measures created here are also compared with the percentage of college graduates by county. Among the advantages of using the LEHD data to develop a skill measure is that it is possible to identify the firm for which a worker works and, therefore, develop a firm-level skill measure.

---

[18] Abowd, Kramarz, and Margolis (1999) show that for firm level averages of the person effect, $\hat{\theta}_j \equiv \frac{1}{N_j} \sum_{(i,t)\in\{J(i,t)=j\}} \hat{\theta}_i$, obey the asymptotic distribution $\hat{\theta}_j \to N(\theta_{j,} \sigma^2_{\theta_j})$ as $N_j \to \infty$ where N_j is the number of observations for firm j and $\sigma^2_{\theta_j} \equiv \frac{1}{N_j^2} \sum \frac{\sigma^2_\varepsilon}{T_i}$ under the assumption that the distribution of firm sizes is constant. Under these same asympotitics, θ_i will not converge to its true value.

[19] The 90th percentile was also used in earlier versions of this research and produced results similar to that of the 75th percentile.

In order to identify the effect of local labor market skill in the investment equation, there must be variation across counties in skill. Figures 2.1-2.4 in chapter 2 depict county level skill over the time period in a series of kernel density estimates for two different measures of county level human capital. Highlighting the key findings in chapter 2, the shape of the distribution in each of the years appears to be bi-modal with a large number of lower skill counties and a smaller concentration of high skill counties. There also appears to be wide variation in the percentage of workers in each county who are in the top quartile of all workers with the least skilled county having 10% of these workers and the most skilled counties having over 30%. Figure 2.3 repeats the kernel density estimates measuring county skill as the mean theta for each county. The distribution in figure 2.3 does not appear to be bi-modal but maintains a similar shape with a larger mass in the lower end of the distribution than in the high end.

In order for firms to be able to predict the skill available in the local labor market, there must be persistence in county skill. Additionally, if county skill is not persistent, then it seems more probable that firms are not limited to the type of workers currently in their local labor market. Again, in chapter 2, Figures 2.10 and 2.11 study short term and long term changes in county skill. The clustering of counties around the 45-degree line is, not surprisingly, very tight at the county level. Over the one year time horizon, the overall increase in skill translates into a fitted regression line that is a small parallel shift of the 45-degree line, emphasizing the persistence in variation discovered in earlier figures. Over the longer time horizon, the overall increase in skill results in a fitted regression line that is above the 45 degree line but also slightly less steep. While workers may be mobile, their mobility patterns reinforce preexisting distributions of skill across counties, thereby suggesting that establishments are limited to the type of workers found in their local labor market.

Finally, chapter 2 also outlined the similarities between these non-standard measures of skill and the percentage of workers with a college degree as calculated from the 1990 Census. The correlation between the two measures is also high, with a correlation coefficient of 0.73. All three measures also have similar shapes to their density, with a large mass in the left tail of the density and a long right tail, although the pattern appears to be most pronounced in the college graduate measure. While the percentage of college graduates in a county is a simple, attractive measure of county skill, it is not an ideal measure. The percentage of college graduates is calculated from responses to the long form of the Census and, as with every other survey, the data is subject to varying response rates by county and respondent error. In addition, the quality of the college attended, the major chosen, and the success of the student in school are not captured in this measure. The worker fixed effect, on the other hand, is a much richer measure of skill. It captures any attribute of a worker that is fixed and that is valued in the labor market, potentially including aspects of worker skill missed by the college graduate measure. While the exact components of what is encompassed in the worker fixed effect are unobservable, there are small differences between the different measures. Therefore, the usage of the worker effect to measure skill should produce results similar to more traditional skill measures, while capturing a richer definition of worker skill. However, as an additional robustness check, the percentage of college graduates by county is used as an alternative measure of local labor market skill.

3.6 Computer Investment

Before moving on to the regression results from the investment equations, a few more comments must be made on the computer investment variables. Figures 3.1 and 3.2 graph the cumulative distribution function for the two constructed computer investment measures, computer investment per worker and computer investment as a fraction of total machinery investment respectively.[20] As is clear in both graphs, nearly half of the sample has zero computer investment. In Figure 3.1, there is a sharp increase in the CDF right after zero, but the graph quickly flattens out and remains flat. In Figure 3.2, there is also a sharp increase in the CDF immediately after zero followed by a much flatter increase. Another sharp increase exists in the CDF exactly at 1, as a small portion of the establishments had all of their machinery investment in computers.

[20] While computer investment/machinery investment is, by definition, bounded between zero and one, computer investment per worker has a very long right tail. Therefore, in order to not disclose the maximum of the distribution, the cumulative density graphed here is truncated at $5,000 of computer investment per worker, which is roughly the 99th percentile of the distribution.

Figure 3.1: Cumulative density of computer investment per employee ($1000)

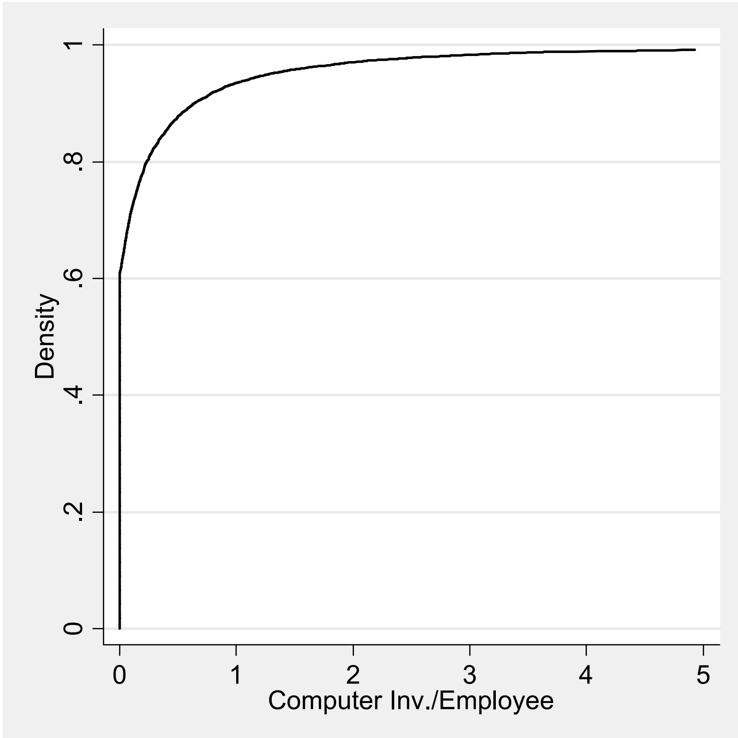
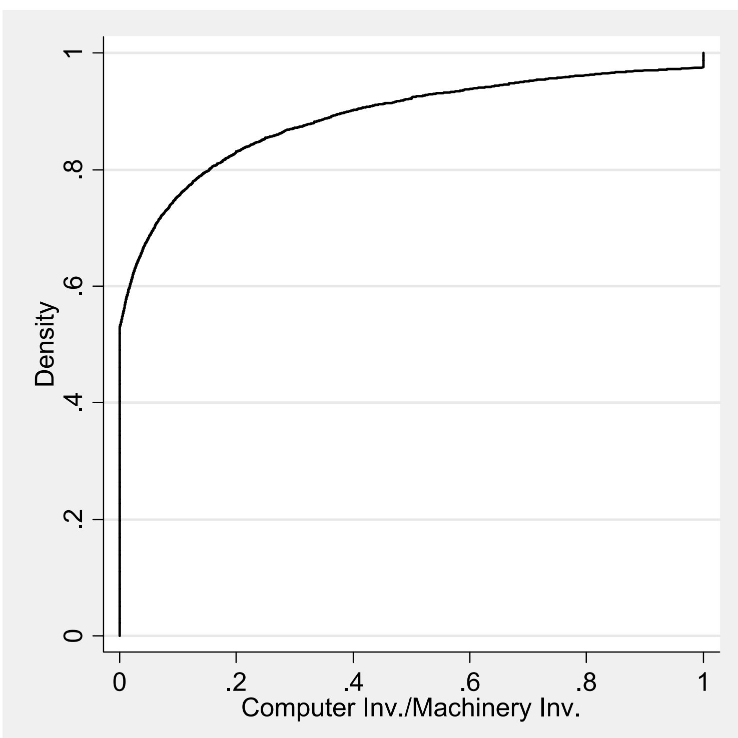


Figure 3.2: Cumulative Density of Computer Investment/Machinery Investment



When considered within a time series perspective, 1992 was a year of strong growth in computer investment. This growth in the use of computers around 1992 also helps to argue in favor of the local skill level being exogenous to the firm. While in equilibrium one would expect that an establishment would choose a location, and therefore local labor market skill, suitable for their technology, the development of a new technology such as computers is an exogenous shock. The median age of a manufacturing establishment used in the analysis is 12 years. Most sample firms therefore made their initial location decisions on the technologies they expected to use over 10 years prior to 1992. In 1982, computer investment per worker in manufacturing was less than 1/20th its 1992 level.[21] In order to account for the large increase in computer investment, the set of technology options must have expanded over the time period, or the price of technology must have fallen to a level where more firms found it profitable to use computers. Regardless, when these firms entered the market, they most likely were unable to predict the change in the technologies that would be available to them, suggesting that their location decision was exogenous to the investment decision being studied here.

3.7 Results

Before turning to the investment equation regressions, table 3.2 lists the summary statistics for the investment variables and results from the wage equation relevant to the investment analysis. The final sample for the investment analysis is the result of a match between the Annual Survey of Manufacturing sample of the 1992 Manufacturing Economic Census and the UI wage data for the selected three

---

[21] See Dunne, Foster, Haltiwanger and Troske (2000).

states. The level of aggregation for this sample is EIN-county. Panel B gives an example of the effects of the sample restrictions and weighting on the key independent variable, mean county skill measured as the percentage of workers in a county who are in the overall top quartile of the theta distribution. If one looks at the unweighted sample of counties, the mean county has approximately 20% of its workforce in the top quartile of the theta distribution with a variance of 0.05. Looking at the sample of firms that match to either of the computer investment variables, the mean county skill is much higher at 25%. The mean county skill is higher in the matched sample because, as shown earlier, larger counties with more firms are more skilled. Weighting county skill by the product of the ASM sample weight and the total value of shipments has little effect on either the mean or variance of county skill.

Table 3.2: Summary Statistics

Panel A:

	Data sample	Number of obs.	Mean	Standard Deviation
Computer Inv. per worker ($1000)	1992 ASM X UI Wage	8339	0.2819	1.4539
Computer Inv. per worker ($1000, weighted)	1992 ASM X UI Wage	8339	0.6448	1.6903
Computer Inv. /Machinery Inv.	1992 ASM X UI Wage	6833	0.1110	0.2272
Computer Inv. /Machinery Inv. (weighted)	1992 ASM X UI Wage	6833	0.1114	0.1844
1991 county skill, $\theta_l^{>75}$	UI Wage	184	0.1957	0.0476
1991 county skill, θ_l^{mn}	UI Wage	184	-0.2096	0.1143
1991 estimated firm effect, ψ	UI Wage	916,896	-0.0138	0.7579
1991 firm skill, $\theta_j^{>75}$	UI Wage	916,896	0.2268	0.2943
1991 firm skill, θ_j^{mn}	UI Wage	916,896	-0.1131	0.6473

Panel B

1991 county skill, $\theta_l^{>75}$	UI Wage	184	0.1957	0.0476
1991 county skill, $\theta_l^{>75}$ (matched CI/MI sample)	UI Wage X 1992 ASM	6833	0.2471	0.0407
1991 county skill, $\theta_l^{>75}$ (weighted, matched CI/MI sample)	UI Wage X 1992 ASM	6833	0.2467	0.0436
1991 county skill, $\theta_l^{>75}$ (matched CI/EMP sample)	UI Wage X 1992 ASM	8339	0.2478	0.0400
1991 county skill, $\theta_l^{>75}$ (weighted, matched CI/EMP sample)	UI Wage X 1992 ASM	8339	0.2471	0.0432

Panel A of table 3.3 lists the results for the first set of regressions with computer investment per worker as the dependent variable. There are six different specifications each including a different set of covariates. The first is the most basic specification using only covariates that are implied by the theoretical model: county level skill and the firm effect as calculated from the wage equation. The county skill measure used throughout the investment equation regressions omits the effect of the establishment's own employees on local labor market skill.[22] The second specification adds in firm level skill as a control, and the third includes both firm skill and the industry dummies. Including firm level skill allows one to distinguish between the model outlined here and a competing model in which there are no search frictions. In this alternate model, firms would be able to meet a worker and then invest in technology leading to a positive coefficient on firm skill and an insignificant coefficient on county skill. Industry controls are necessary because industries locate non-randomly across geography. If highly skilled counties were comprised of industries that are more likely to utilize computers, again the county skill coefficient would be biased upward. Because the key independent variable varies by county but not by observation, robust standard errors based on clustering by county are included in parentheses.

[22] The county skill measure excluding the establishment's contribution is calculated by subtracting the measure of *firm* (ein) skill weighted by the number of workers at that *establishment* (ein-county) from the overall county skill measure.

Table 3.3: Results

Panel A: Computer investment per Worker

	(1)	(2)	(3)	(4)	(5)	(6)
1991 county skill, $\theta_l^{>75}$	8.515** (3.510)	3.867** (1.659)	3.719** (1.837)	1.209* (0.626)	1.266** (0.620)	1.426** (0.698)
1991 estimated firm effect, ψ	0.422*** (0.158)	0.701*** (0.250)	0.734*** (0.240)	0.460*** (0.156)	0.518*** (0.162)	0.427*** (0.140)
1991 firm skill, $\theta_j^{>75}$		4.756*** (1.794)	4.443*** (1.658)	2.926*** (0.847)	3.064*** (0.912)	3.004*** (0.890)
Capital stock per employee						0.001 (0.001)
Constant	-1.635** (0.811)	-1.522** (0.622)	-1.498** (0.648)	-0.558*** (0.197)	-0.602*** (0.202)	-0.677*** (0.245)
Observations	8339	8339	8339	8339	7789	7789
R-squared	0.05	0.16	0.21	0.30	0.31	0.31
Ind. controls	no	no	yes	yes	yes	yes
Cty interaction	no	no	no	yes	yes	yes
Capital sample	no	no	no	no	yes	Yes
% osd	62.86	28.55	27.46	8.93	9.34	10.53

Panel B: Computer Investment/Machinery Investment

	(1)	(2)	(3)	(4)	(5)	(6)
1991 county skill, $\theta_l^{>75}$	0.682*** (0.084)	0.343*** (0.109)	0.099 (0.114)	0.236** (0.091)	0.250*** (0.093)	0.212** (0.087)
1991 estimated firm effect, ψ	-0.059*** (0.015)	-0.038** (0.015)	-0.018 (0.012)	-0.015 (0.013)	-0.010 (0.013)	0.012 (0.014)
1991 firm skill, $\theta_j^{>75}$		0.341*** (0.056)	0.273*** (0.039)	0.255*** (0.042)	0.286*** (0.042)	0.299*** (0.040)
Capital stock per employee						-0.000*** (0.000)
Constant	-0.033* (0.020)	-0.025 (0.023)	-0.012 (0.025)	-0.040 (0.024)	-0.051** (0.024)	-0.033 (0.023)
Observations	6833	6833	6833	6833	6399	6399
R-squared	0.03	0.08	0.15	0.16	0.17	0.18
Ind. controls	no	no	yes	yes	yes	yes
Cty interaction	no	no	no	yes	yes	yes
Capital sample	no	no	no	no	yes	Yes
% osd	29.16	14.68	4.25	10.07	10.69	9.06

Robust standard errors based on clustering by county in parentheses. Weighted by ASM sample weight and total value of shipments. County skill measure excludes establishment's contribution. % osd is the predicted percent change in the dependent variable due to a one standard deviation increase in county skill.
* significant at 10%; ** significant at 5%; *** significant at 1%

The key coefficient of interest, the percentage of workers in a county who are in the top quartile adjusted for the establishment's contribution, is listed first. This coefficient is positive and significant in each of the first three specifications. The magnitude of the coefficient on county skill diminishes as the additional control variables are added to the specification. In order to put an interpretation on the coefficient, the predicted percentage change in the dependent variable due to a one standard deviation increase in county skill is reported in the last row. In the third specification, the regression implies that a one standard deviation increase in county skill leads to a 27% increase in the amount of computer investment per worker. The other two independent variables in the first three specifications are the firm effect from the wage regression and firm level skill. The coefficient on the estimated firm effect is positive and significant once industry controls are included, and the coefficient on firm skill is positive and significant in all of the specifications. Both variables have coefficients smaller in magnitude than county skill.

The specification in column 4 examines the effect of the most highly skilled county on the results. The existence of concentrations of establishments in like industries is well known. While the driving force behind these agglomeration economies may very well be their access to a pool of highly skilled labor and, consistent with the theory laid out above, the success of these agglomeration economies may also be due to a myriad of factors not captured in the model. In order to determine if the results here are driven by a few agglomeration economies existing in highly skilled local labor markets, an interaction term between each of the key coefficients in the model and a dummy variable for the most highly skilled county was included in the fourth regression. While the coefficient on county skill is still positive and significant, the magnitude of the coefficient drops, suggesting that the predicted increase in computer investment per worker from a one standard deviation increase in county skill is around 9%. Throughout most of the rest of the results, the interaction with the most highly skilled county is included in order to present a more conservative estimate of the effect of county skill. Whether or not the forces outlined in the model drive the effect of this most highly skilled county on the results, the inclusion of these interaction terms is necessary because the relationship between county skill and establishment investment in computers is non-linear due to this one county. The sensitivity of the results to such nonlinearities is further explored in the next section.

Columns 5 and 6 address a specification issue related to the estimated firm effect. While the model suggests that technology investment is influenced by firm type, a pure measure of firm type is not available. What is available is the estimated firm effect from the wage decomposition, which captures characteristics of a firm captured in their employee's wages. Under the assumption that wages are determined via rent-sharing, this firm effect should be a good proxy for firm type, but will be contaminated with other characteristics of the firm. In particular, if the estimated firm effect captures characteristics of the firm's previous technology decisions, and if these previous technology decisions are correlated with current technology investment, then the coefficient on the firm effect will be biased. The coefficient on county skill, the key variable for testing endogenous technology, will also be biased to the extent that county skill is correlated with the firm effect. If county skill is orthogonal to the firm effect, then the bias due to this specification problem will only affect the coefficient on the firm effect. In order to determine whether omission of previous technology investment affects the results, the 1991 capital stock is included as an additional covariate.

The capital stock measure is only available for a subset of the firms used in the earlier analysis. Column 5 repeats the specification in column 4 on this subset of firms. Column 6 includes the capital stock measure as an additional control variable. In the fifth column, the coefficient on county skill is slightly higher for the subset of firms for which capital stock information is available. Including the capital stock measure in column 6 increases the coefficient on county skill further. This suggests that any potential bias from using the proxy for firm type and not controlling for the capital stock will depress the coefficient on county skill. Given that the capital stock variable is only available for a subsample of the data, the results in the following tables rely on the full sample and provide a conservative estimate of the effect of county skill on technology investment. While the coefficient on county skill increases with the inclusion of lagged capital stock, the coefficient on the firm effect decreases, as would be expected.

Panel B of table 3.3 repeats the specifications in panel A replacing the dependent variable with the bias of investment toward computers. The pattern in the results is very similar. The coefficient on county skill excluding the establishment's contribution falls as controls for firm level skill and industry are included. However, while the coefficient on county skill is positive in all of the specifications, it is only significant when controlling for firm level skill separately, but not when both firm level skill and industry dummies are included simultaneously in column 3. When the key coefficients in the model are interacted with the most highly skilled county, in column 4, the effect is also different with computer bias as the dependent variable. Here the nonlinearities appear to be depressing the size of the coefficient on county skill. Once the interaction terms are included the size of the coefficient increases dramatically and is once again significant. In this specification, a one standard deviation increase in county skill is predicted to increase the bias in investment toward computers at the establishment level by 10%. The other independent variables in the first three specifications are the firm effect and firm level skill. The coefficient on the firm effect is negative and significant in all of the specifications, which goes against the model predictions, but small in magnitude. The coefficient on firm level skill is positive, significant, and of a greater magnitude than the coefficient on county skill.

Restricting the sample to that for which the capital stock measure is available, in column 5, increases the coefficient on county skill by a small amount, as was the case in panel A. Adding the capital stock as a covariate, however, decreases the magnitude of this coefficient by a small amount, suggesting that omitting previous capital investment inflates previous results. This is contrary to the results in the first panel. However, because the change in the coefficient is small, the specification without a control for the capital stock is still preferred and used throughout the remaining tables. Appropriate caution is necessary when assessing the magnitude of the effect of county skill on computer investment as a fraction of machinery investment.

The two different dependent variables used to test the model, computer investment per worker and the percentage of investment that is in computers, largely lead to the same conclusions as to the effect of county skill on establishment investment. Both dependent variables predict large effects for county skill, which diminish as additional controls are added. The biggest difference between the two specifications is the sensitivity of the results to the most highly skilled county. While including additional interaction terms decreases the explanatory power of county skill in the first set of regressions, it increases the effect of county skill in the second set of regressions. Much of the difference is likely due to the fact that the computer investment bias measure is bounded between zero and one. Regardless of its location, a firm can, at most, concentrate 100% of its investment in computers. Computer investment per worker, on the other hand, is unbounded.

Due to the sensitivity of the results to one county, the preferred base specification is the fourth column of the table, which includes interaction terms with the most highly skilled county. These results suggest that a one standard deviation increase in county skill will lead to an 9% increase in computer investment per worker or a 10% increase in the share of investment in computers. While these effects might seem large, a one standard deviation increase in county skill is equivalent to a five-percentage point increase in the number of workers in a county who are in the top quartile of the overall skill distribution. A five percentage point increase in skill, in which the average county has 25% of its workers in the top quartile, would require a significant reallocation of workers. Still, interpreting the results in this manner is helpful to gauge the importance of county skill in an establishment's investment decision. Regardless of the dependent variable being studied, the effect of a firm's own skill mix, industry, and firm type are also important factors in an establishment's investment decision.

3.8 Robustness Checks

Because a variety of measurement and specification decisions underlie the results in table 3.3, a series of robustness checks are included below to test the sensitivity of the results to these assumptions. As is shown in detail below, the results in table 3.3 are robust to most of these decisions. In instances in which the results are sensitive, potential explanations are provided.

3.8.1 Nonlinearities

Given that the most highly skilled county greatly influences the coefficient on county skill, table 3.4 tests further for nonlinearities in county skill and offers a possible explanation. Column 1 repeats the base specification but includes no interactions with highly skilled counties. Column 2 is identical to the fourth column in the previous table and includes an interaction between the key independent variables in the model and the most highly skilled county. The third column includes interactions with the top 5% of most skilled counties which includes 9 counties. As shown earlier, the inclusion of interactions greatly affects the coefficient on county skill in both panel A and panel B, moving the coefficient in opposing directions in the two panels. Concentrating next on column three, one finds that the coefficient on county skill does not change much when the regression includes an interaction with the top 5% of counties by skill. The results therefore suggest that the nonlinearity with highly skilled counties is concentrated in just one county that happens to be the most skilled one.

Table 3.4: Robustness: Nonlinearities

Panel A: Computer Investment per Worker

	(1)	(2)	(3)	(4)
1991 county skill, $\theta_l^{>75}$	3.719**	1.209*	1.623**	2.508**
	(1.837)	(0.626)	(0.626)	(1.134)
1991 estimated firm effect, ψ	0.734***	0.460***	0.450***	0.693***
	(0.240)	(0.156)	(0.170)	(0.204)
1991 firm skill, $\theta_j^{>75}$	4.443***	2.926***	2.841***	4.376***
	(1.658)	(0.847)	(0.923)	(1.582)
% workers in high skill ind.				4.488***
				(1.336)
Constant	-1.498**	-0.558***	-0.637***	-1.353***
	(0.648)	(0.197)	(0.223)	(0.370)
Observations	8339	8339	8339	8339
R-squared	0.21	0.30	0.27	0.22
County interactions	none	top 1	top 5%	none
% osd	27.46	8.93	11.98	18.52

Panel B: Computer Investment/Machinery Investment

	(1)	(2)	(3)	(4)
1991 county skill, $\theta_l^{>75}$	0.099	0.236**	0.247***	0.168**
	(0.114)	(0.091)	(0.093)	(0.078)
1991 estimated firm effect, ψ	-0.018	-0.015	-0.014	-0.016
	(0.012)	(0.013)	(0.014)	(0.012)
1991 firm skill, $\theta_j^{>75}$	0.273***	0.255***	0.255***	0.276***
	(0.039)	(0.042)	(0.045)	(0.041)
% workers in high skill ind.				-0.248***
				(0.058)
Constant	-0.012	-0.040	-0.042*	-0.020
	(0.025)	(0.024)	(0.025)	(0.019)
Observations	6833	6833	6833	6833
R-squared	0.15	0.16	0.16	0.16
Cty interactions	none	top 1	top 5%	none
% osd	4.25	10.07	10.54	7.16

Robust standard errors based on clustering by county in parentheses. Weighted by ASM sample weight and total value of shipments. Two digit industry dummies included. County skill measure excludes establishment's contribution. % osd is the predicted percent change in the dependent variable due to a one standard deviation increase in county skill.
* significant at 10%; ** significant at 5%; *** significant at 1%

The most highly skilled county is not an outlier in its skill level, which is not far removed from the rest of the distribution. Rather, this one county has both many skilled workers and high levels of computer investment. The final regression in table 3 highlights one potential characteristic of the highly skilled county that may be driving its impact on the results. In the fourth specification, the county's percentage of workers employed in high tech industries, SIC 35 and 36,[23] is included as a covariate. These results are similar to the prior regression that included the interaction terms with the most highly skilled county. As mentioned above, the impact of this one county may be due to aspects of agglomeration economies that are not captured in the endogenous technology model.

3.8.2 Weighting/Firm Size

The dependent variables used throughout the analysis so far all require use of the ASM to obtain information on expenditures on computers. This sample drawn from the Census of Manufactures is disproportionately composed of large firms and is not representative of all manufacturing establishments.[24] In order to make the results representative of the average manufacturing establishment the results must be weighted by the Census ASM sample weight. However, the representative firm in manufacturing is rather small and therefore accounts for only a small fraction of the manufacturing industry's output. Due to this fact, all of the regressions in the earlier tables are weighted by the product of the Census ASM sample weight and the total

[23] SIC 35 and 36 are Industrial and Commercial Machinery and Computer Equipment; and Electronic and Other Electrical Equipment and Components, respectively.
[24] See details in footnote 13.

value of shipments for that establishment in order to make the results representative of a given unit of economic activity.

Table 3.5 repeats the base specification using three different weighting patterns to highlight the effect of weighting on the results. In the first column no weights are used, in the second column the Census ASM weight is used, and in the third column the product of the Census ASM weight and the total value of shipments is used, as in the rest of the analysis. The effect of weighting on the county skill coefficient is similar for either dependent variable although stronger when looking at computer investment per worker. In the unweighted regression the predicted effect of a one standard deviation increase in county skill on the dependent variable leads to a 2% increase in computer investment per worker. Weighting the same regression by the Census ASM weight makes the effect negative, and weighting by the product of the Census ASM weight and total value of shipments increases the effect to 9%. With computer investment per worker, the same effect is 6% with no weighting, 0.5% with Census ASM weighting, and 10% with Census ASM and total value of shipments as the weight.

Table 3.5: Robustness: Weighting/Firm Size

Panel A: Computer Investment per Worker

	(1)	(2)	(3)
1991 county skill, $\theta_l^{>75}$	0.313 (0.340)	-0.590 (0.355)	1.209* (0.626)
1991 estimated firm effect, ψ	0.174*** (0.055)	0.116*** (0.042)	0.460*** (0.156)
1991 firm skill, $\theta_j^{>75}$	0.991*** (0.156)	0.548*** (0.090)	2.926*** (0.847)
Constant	-0.095 (0.080)	0.155* (0.089)	-0.558*** (0.197)
Observations R-squared	8339 0.04	8339 0.03	8339 0.30
Weight	none	Census	Census*TVS
% osd	2.31	-4.35	8.93

Panel B: Computer Investment/Machinery Investment

	(1)	(2)	(3)
1991 county skill, $\theta_l^{>75}$	0.151** (0.066)	0.010 (0.211)	0.236** (0.091)
1991 estimated firm effect, ψ	0.016 (0.013)	0.059** (0.024)	-0.015 (0.013)
1991 firm skill, $\theta_j^{>75}$	0.088*** (0.026)	0.115** (0.046)	0.255*** (0.042)
Constant	-0.005 (0.017)	0.042 (0.056)	-0.040 (0.024)
Observations R-squared	6833 0.06	6833 0.05	6833 0.16
Weight	none	Census	Census*TVS
% osd	6.44	0.45	10.07

Robust standard errors based on clustering by county in parentheses. Two digit industry dummies and interaction terms with high skill county included. County skill measure excludes establishment's contribution. % osd is the predicted percent change in the dependent variable due to a one standard deviation increase in county skill.
* significant at 10%; ** significant at 5%; *** significant at 1%

The differences in the effect of county skill on investment across the regressions is most likely due to the differences in the explanatory power of different size firms. The unweighted sample is disproportionately composed of large firms, and the Census ASM weight corrects for that so that the results reflect the representative firm, which is much smaller. Finally, the product of the Census ASM weight and the total value of shipments shifts the emphasis back to larger firms again. Why does the effect of county skill seem to be larger for larger firms? This effect may be driven by a variety of reasons. First, computer investment at the establishment level is measured with less error in larger firms. The ASM is collected in order to publish aggregate statistics about manufacturing. Because larger firms will drive any aggregate statistic, more effort is focused on collecting data in these large firms. Second, there may be non-linearities in the relationship between county skill and establishment investment in computers. In part, this effect is driven by the fact that larger firms need to hire more workers. Earlier research has shown that in order to get the greatest productivity boost from introducing computers, establishments must integrate computers into much of their operations. Larger establishments with larger operations require more skilled workers in order to integrate computers. In addition, larger establishments invest more per worker than small establishments even when controlling for industry, the firm effect from the wage equation, and firm level skill.

3.8.3 Alternate Skill Measures

Table 3.6 tests the sensitivity of the results to different ways of measuring county skill. Column one repeats the base specification. Column two uses the percentage of workers in a county from the top quartile of the theta distribution by measuring human capital using the sum of the fixed worker effect and the predicted effect of experience from the wage regression. For either dependent variable, the effect of county skill is a bit larger when worker experience is included in the skill measure. The third column uses the mean theta in a county and the fourth column uses the mean of the sum of theta and experience. The results from both of these specifications closely follow the pattern found for the top quartile measures when looking at computer investment per worker. With investment bias towards computers as the dependent variable, the effect of mean county skill is smaller when worker experience is included in the measure of skill.

Table 3.6: Robustness: Alternative Skill Measures

Panel A: Computer Investment per Worker

	(1)	(2)	(3)	(4)	(5)
1991 county skill, $\theta_l^{>75}$	1.209* (0.626)				
1991 county skill, $s_l^{>75}$		2.161** (0.832)			
1991 county skill, θ_l^{mn}			0.534 (0.347)		
1991 county skill, s_l^{mn}				1.219*** (0.406)	
% College Grad					1.841*** (0.527)
Observations R-squared	8339 0.30	8339 0.26	8339 0.27	8339 0.23	8339 0.17
% osd	8.93	13.81	9.46	15.03	17.90

Panel B: Computer Investment/Machinery Investment

	(1)	(2)	(3)	(4)	(5)
1991 county skill, $\theta_l^{>75}$	0.236** (0.091)				
1991 county skill, $s_l^{>75}$		0.352*** (0.118)			
1991 county skill, θ_l^{mn}			0.129*** (0.036)		
1991 county skill, s_l^{mn}				0.093* (0.050)	
% College Grad					0.243*** (0.062)
Observations R-squared	6833 0.16	6833 0.14	6833 0.15	6833 0.14	6833 0.13
% osd	10.07	13.05	13.21	6.65	13.67

Robust standard errors based on clustering by county in parentheses. Two digit industry dummies, a constant term, and interactions with high skill county are included all the specifications. Firm level skill is included in the first four specifications. Weighted by ASM sample weight and total value of shipments. % college graduates in county calculated from 1990 Census data. County skill measure excludes establishment's contribution in columns 1-4. % osd is the predicted percent change in the dependent variable due to a one standard deviation increase in county skill.
* significant at 10%; ** significant at 5%; *** significant at 1%

Finally the fifth column uses the percentage of college graduates in a county calculated using the 1990 Census data. For either dependent variable, these results are remarkably close to that found in the base specification. While the impact of county skill is a bit larger, this specification does not include a control for firm level skill because it is not possible to compute using Census data. Additionally, because there is no information on firm skill, the firm's contribution to the county skill measure cannot be excluded as it is in the other columns. The key difference between the college measure and the others is that the college measure is calculated using individual's responses about their education from the Census, while the other skill measures are derived from a wage regression using administrative data. Additionally, this measure of county skill uses 1990 data while the other county skill measures all use 1991 data. While the differences between the results using the employer-employee matched data from LEHD and those using the Decennial Census data may be small, the matched data is necessary for two primary reasons. The first is that the magnitude of the results are deceptively large when using the Decennial Census measure because one cannot control for firm level skill. The second reason is that the firm effect can be computed only from the matched data, and is a control in the investment equation mandated by the model.

The small differences in the predicted change in the dependent variable due to a one standard deviation increase in the county skill measure across the five specifications suggests that the effect being found is not the result of a particular way of measuring skill. The results in the final column are the strongest support for this claim, given that they are calculated in a different way, from a different data source.

3.8.4 Alternate Local Labor Market Measure

County of work has been used as the measure of the local labor market throughout the analysis. While a county is a desirable measure of the local labor market for the reasons listed above, the metropolitan area is also commonly used to define a local labor market. Table 3.7 includes a comparison between the base specification, in column 1, and one in which the local labor market is defined by the metropolitan area[25] in column 2. Because the metropolitan areas are not exhaustive, the counties outside of any metropolitan area are included in one pooled non-metro area group in column 2. Columns 3 and 4 exclude establishments in these non-metro areas. Using either dependent variable and whether or not the non-metro areas are included, the results are largely the same across all of the specifications. With computer investment per worker, the results are a bit stronger when using the metropolitan areas, especially when the non-metro areas are dropped from the analysis. In panel B, the results with computer investment over machinery investment are a bit smaller when the metropolitan areas are used, whether or not the non-metro areas are excluded from the analysis. The effect of a one standard deviation increase in local labor market skill is a bit misleading because there is less variance in metropolitan area skill than there is in county skill. Regardless, using the metropolitan area as the measure of the local labor market produces results very similar to those produced using county as the local labor market measure.

[25] The metropolitan area used in this analysis is either the Metropolitan Statistical Area or the Primary Metropolitan Statistical Area of a Consolidated Metropolitan Statistical Area. For the states included here, there are 40 metropolitan areas.

Table 3.7: Robustness: Alternative Local Labor Market Measure

Panel A: Computer Investment per Worker

	(1)	(2)	(3)	(4)
1991 county skill, $\theta_l^{>75}$	1.209*		1.770**	
	(0.626)		(0.819)	
1991 msa skill, $\theta_m^{>75}$		1.677***		2.482***
		(0.569)		(0.759)
1991 estimated firm effect, ψ	0.460***	0.457***	0.497***	0.498***
	(0.156)	(0.158)	(0.156)	(0.162)
1991 firm skill, $\theta_j^{>75}$	2.926***	2.908***	2.926***	2.914***
	(0.847)	(0.837)	(0.872)	(0.865)
Constant	-0.558***	-0.663***	-0.712***	-0.880***
	(0.197)	(0.219)	(0.248)	(0.278)
Observations	8339	8339	7757	7757
R-squared	0.30	0.30	0.31	0.31
Non-metro areas included	yes	yes	no	no
% osd	8.93	9.85	13.06	14.59

Panel B: Computer Investment/Machinery Investment

	(1)	(2)	(3)	(4)
1991 county skill, $\theta_l^{>75}$	0.236**		0.240**	
	(0.091)		(0.114)	
1991 msa skill, $\theta_m^{>75}$		0.224**		0.212*
		(0.091)		(0.114)
1991 estimated firm effect, ψ	-0.015	-0.015	-0.016	-0.016
	(0.013)	(0.013)	(0.014)	(0.014)
1991 firm skill, $\theta_j^{>75}$	0.255***	0.258***	0.249***	0.252***
	(0.042)	(0.041)	(0.044)	(0.043)
Constant	-0.040	-0.038	-0.039	-0.033
	(0.024)	(0.025)	(0.032)	(0.033)
Observations	6833	6833	6324	6324
R-squared	0.16	0.16	0.16	0.16
Non-metro areas included	yes	yes	no	no
% osd	10.07	7.62	10.28	7.21

Robust standard errors based on clustering by county in parentheses. Weighted by ASM sample weight and total value of shipments. Two digit industry dummies and interaction terms with high skill county included. County and MSA skill measure excludes establishment's contribution. % osd is the predicted percent change in the dependent variable due to a one standard deviation increase in county skill.
* significant at 10%; ** significant at 5%; *** significant at 1%

3.8.5 Probit

As mentioned above, the computer investment data used for this analysis comes from a single cross section. Implicitly, the estimation assumes that computers are a non-durable. As a robustness check to the base specification, establishments are placed into a low investor and a high investor group, where the high investors are in the top quartile of either computer investment per worker or computer investment as a share of machinery investment. Table 3.8 repeats the base specification, excluding firm skill in column one and with firm skill in column two, and shows the results of the probit without firm skill in column three and with firm skill in column four. The probit predicts that a one standard deviation increase in county skill will increase the likelihood that a firm is high tech, as measured using computer investment per worker, by 2 to 9% depending on whether or not firm skill is included. While the results in columns one and two appear larger, the specifications are measuring different things and are impossible to directly compare. However, they do both suggest that county skill plays a role in computer investment. In Panel B, the base specification predicts between a 10 and 17% increase in computer investment over machinery investment from a one standard deviation increase in county skill, while the probit predicts between a 19 and 24% increase in the likelihood that a firm is high-tech.

Table 3.8: Robustness: Probit

Panel A: Computer Investment per Worker

	(1) ols	(2) ols	(3) probit	(4) probit
1991 county skill, $\theta_l^{>75}$	2.907*** (0.823)	1.209* (0.626)	0.783* (0.447)	0.192 (0.442)
1991 estimated firm effect, ψ	0.367*** (0.140)	0.460*** (0.156)	0.326*** (0.058)	0.384*** (0.055)
1991 firm skill, $\theta_j^{>75}$		2.926*** (0.847)		0.943*** (0.173)
Constant	-0.464** (0.191)	-0.558*** (0.197)		
Observations R-squared	8339 0.16	8339 0.30	8338	8338
% osd	21.46	8.93	9.43	2.31

Panel B: Computer Investment/Machinery Investment

	(1) ols	(2) ols	(3) probit	(4) probit
1991 county skill, $\theta_l^{>75}$	0.403*** (0.096)	0.236** (0.091)	1.744*** (0.464)	1.325*** (0.439)
1991 estimated firm effect, ψ	-0.028** (0.014)	-0.015 (0.013)	-0.003 (0.065)	0.023 (0.061)
1991 firm skill, $\theta_j^{>75}$		0.255*** (0.042)		0.650*** (0.122)
Constant	-0.032 (0.025)	-0.040 (0.024)		
Observations R-squared	6833 0.13	6833 0.16	6832	6832
% osd	17.25	10.07	24.37	18.58

Robust standard errors based on clustering by county in parentheses. Weighted by ASM sample weight and total value of shipments. Two digit industry dummies and interaction terms with high skill county included. County skill measure excludes establishment's contribution. % osd is the predicted percent change in the dependent variable due to a one standard deviation increase in county skill.
* significant at 10%; ** significant at 5%; *** significant at 1%

3.8.6 Alternate Dependent Variables

As a final robustness check, machinery investment per worker, a separate question on the ASM from computer investment, is used as the dependent variable in table 3.9. Machinery investment per worker is asked of all establishments in Census years, therefore the results using machinery investment per worker are computed for both years for which human capital data is also available, 1992 and 1997. Because machinery investment data is collected for all establishments, the Census ASM sample weight does not apply here, and results are weighted only by the total value of shipments for that establishment. The results in Panel A, machinery investment per worker in 1992, and in Panel B, machinery investment per worker in 1997, both follow the same broad pattern. The first specification uses theta as the measure of county skill. The second also uses theta and additionally includes interactions between the key coefficients and the most skilled county. The third uses the sum of theta and worker experience to measure county skill and the fourth adds in interactions with the most skilled county. Contrary to the earlier results using computer investment as the dependent variable, the coefficient on county skill is negative in the first two specifications of either panel, suggesting that firms in counties with large numbers of workers with high worker effects are less likely to invest in machinery. However, the coefficient is positive, yet still not significant, in the last two specifications, in which county skill is measured using the sum of the worker effect and worker experience. The difference in the pattern in these results compared to the earlier results using computer investment is likely due to the fact that machinery investment is very heterogeneous. While machinery investment includes computer investment, it is also comprised of much older technologies. These older technologies may disproportionately require worker experience. While the endogenous technology model does not fit as well using machinery investment per worker as the dependent variable, table 8 provides some evidence that the results laid out in the earlier tables using computer investment are not an artifact of the computer investment data.

Table 3.9: Robustness: Machinery Investment per Worker

Panel A: Machinery Investment per Worker, 1992

	(1)	(2)	(3)	(4)
1991 estimated firm effect, ψ	11.099*** (2.881)	11.422*** (3.151)	9.934*** (2.811)	10.546*** (3.004)
1991 county skill, $\theta_l^{>75}$	-4.284 (23.345)	-6.152 (31.787)		
1991 firm skill, $\theta_j^{>75}$	16.583*** (5.097)	15.520** (5.997)		
1991 county skill, $s_l^{>75}$			24.047 (24.973)	29.898 (45.600)
1991 firm skill, $s_j^{>75}$			4.645 (5.230)	4.646 (6.072)
Constant	3.531 (5.830)	4.017 (8.156)	-0.514 (5.998)	-2.028 (10.846)
Observations R-squared	56563 0.37	56563 0.38	56563 0.37	56563 0.37
County interactions	none	top 1	none	top 1
% osd	-1.88	-2.69	9.12	11.34

Panel B: Machinery Investment per Worker, 1997

	(1)	(2)	(3)	(4)
1996 estimated firm effect, ψ	21.307*** (3.454)	21.126*** (3.572)	19.825*** (3.853)	19.557*** (4.009)
1996 county skill, $\theta_l^{>75}$	-12.988 (12.822)	-16.042 (12.271)		
1996 firm skill, $\theta_j^{>75}$	19.713*** (4.888)	24.481*** (5.441)		
1996 county skill, $s_l^{>75}$			5.639 (14.509)	6.484 (18.056)
1996 firm skill, $s_j^{>75}$			14.814*** (3.118)	16.597*** (3.518)
Constant	10.450*** (3.810)	10.179** (4.165)	5.779 (4.259)	5.052 (5.375)
Observations R-squared	55604 0.25	55604 0.25	55604 0.25	55604 0.25
County interactions	none	top 1	none	top 1
% osd	-4.13	-5.11	1.55	1.79

Robust standard errors based on clustering by county in parentheses. Weighted by total value of shipments. Two digit industry dummies and interaction terms with high skill county included. County skill measure excludes establishment's contribution. % osd is the predicted percent change in the dependent variable due to a one standard deviation increase in county skill.
* significant at 10%; ** significant at 5%; *** significant at 1%

3.9 Conclusion

There is tremendous heterogeneity in the technology employed by firms, even in well-defined industries. One potential cause of this heterogeneity is endogenous technology driven by the variation and persistence of human capital across different local labor markets. The research here builds a matching model capturing the effects of local labor market worker skill on establishment investment decisions. By taking advantage of a unique employer-employee matched dataset, the results begin to quantify the effects of local labor market skill on establishment technology.

The best estimates of the effect of county skill on an establishment's investment predicts that a one standard deviation increase in county skill will lead to an 9% increase in computer investment per worker and an 10% increase in the investment share of computers for a representative unit of economic activity. Weighting the results and thereby shifting the emphasis between smaller and larger firms does affect the results. The effect of an increase in county skill is not nearly as large for a representative establishment. This outcome suggests that county skill has a greater impact on the investment decisions of larger firms. However, the results are robust to different ways of measuring county skill, different measures of the local labor market, and different functional forms of the specification. The results are sensitive to the type of investment undertaken by the establishment. When county level skill is measured by including the effect of worker experience, one finds positive yet insignificant results with machinery investment per worker as the dependent variable, likely due to the fact that capital-skill complementarity is not as strong with machinery investment. The pattern found in the results is consistent with other research on technology adoption. Productivity enhancements from the usage of computers require widespread changes in an establishment. These changes require a large investment in skilled workers. The research here suggests that firms are more willing to make the investment in computers if the necessary workforce is available.

One area for further research is to explore how endogenous technology affects the dynamics of worker location. While the empirical work here uses investment in computers in a relatively early time period to ensure that workers are exogenously distributed in reference to firms' likelihood of investing in computers, data in later periods can be used to examine workers' reactions to firms' investments. The results here suggest that it is in the best interest of the high-tech firm and the highly skilled worker to locate in high skill areas. As the usage of technology increases in an area, does one also see an increase in the concentration of skilled workers?

Chapter 4: Estimating External Returns to Education

Public funding of education has largely been motivated by the belief that there are positive externalities associated with education. Different policy makers have cited a variety of positive externalities associated with education, including a more informed voting citizenry and more productive workforce. Economists have focused mostly on the effect that raising the education level of all workers has on individual worker wages. The focus on wages is due in part to the fact that information on wages is readily available, but also because wages are one of the best measures available for capturing the productivity of individual workers. While there is a strong empirical relationship between the level of education and level of wages in a local labor market, identifying the causal relationship of the effect of education on wages is much more difficult. Workers may be paid higher wages in highly educated areas because the overall level of education raises their productivity, or workers in highly educated areas may also be paid more because they are a selected group of workers. Workers are very mobile and are likely to select into the location that rewards their skills the best. Workers in high education areas are by definition more educated, but they might also vary in non-random ways on other unobservable dimensions. Additionally, both the observable and unobservable characteristics of workers may receive different returns in differently skilled areas. The econometric strategy employed here directly addresses the potential nonrandom selection by allowing for variation in the returns to both observable and unobservable characteristics in high and low skill local labor markets.

While other research into the external returns to education have allowed the returns to observable characteristics to vary across differently skilled local labor markets and used a fixed worker effect to control for unobservable individual skill, specifications using this estimation strategy do not allow the returns to unobservable skills to be different in high and low skill local labor markets. Given that others have found that the returns to education vary in high and low skill areas, it is likely that the returns to unobservables also vary. This possibility is likely to bias results toward finding an effect of higher metro area education on individual wages for two reasons. The first is that more highly educated workers are likely to also be higher ability workers, and therefore, areas with high levels of education are also likely to have strong concentrations of high ability workers. Additionally, given that highly educated areas compensate education more generously, it is likely that they also compensate ability more generously. Therefore, the finding that workers in high skill metropolitan areas have higher wages may be due to the fact that these workers are more skilled among unobservable dimensions and are also rewarded more generously for these skills.

This chapter decomposes the wage gap between workers in low and high skill metropolitan areas in a series of different specifications. The first specification estimates the raw wage gap between workers in high and low skill areas. The second adds observable characteristics of workers and their firms. In the third specification, the returns to observable skill are allowed to vary between high and low skill metropolitan areas. Next, fixed worker effects are included to control for unobserved worker characteristics. The final specification, and the new contribution of this chapter, allows for different returns to these unobservable characteristics in high and low skill areas.

The econometric strategy here focuses on directly estimating separate returns to observable and unobservable skill in metropolitan areas with different levels of skilled workers. The permanent component of the error term is interacted with an indicator for high skill metropolitan area to allow for different returns to unobservable skill. A quasi-differenced wage equation is then estimated via non-linear instrumental variables as described in Gibbons, Katz, Lemieux, and Parent (2002). This estimation strategy decomposes the wage gap between high and low skill metropolitan areas into different returns to observable skills, different returns to unobservable skills, and a part that is not attributable to the characteristics of the worker. This final piece of the wage decomposition most closely aligns with the earlier estimates of the effect of metro area's skill on worker wages.

Results from this last specification suggest that both the observable and unobservable characteristics of workers may receive higher returns in more highly skilled metropolitan areas. After controlling for these characteristics, the coefficient on the indicator for a highly skilled metropolitan area is no longer significant. These results are not necessarily a refutation of models in which workers in more highly skilled areas are more productive. Rather, the results here suggest that the impact of a concentration of skilled workers may be felt more strongly for the most skilled workers. The next section discusses the existing literature on measuring the external returns to education and the theoretical motivation behind the estimation. The following sections then outline the estimation strategy, the data used, and finally the results.

4.1 Measuring the Social Return to Higher Education

Rauch (1993) is one of the earlier papers that attempts to directly estimate the social return to high levels of human capital by using variation in skill across metropolitan areas in the United States. He motivates his research using a variant of a modeled originally developed by Roback (1982). In the model, the consumer iso-utility curve and the firm iso-cost curve determine the equilibrium level of wages and rents within each city. Consumers prefer higher wages and lower rents, while firms prefer lower wages and lower rents. Workers and firms therefore sort themselves across cities so that they receive the same utility/face the same total costs in each city. Within this model, the level of education of a metropolitan area is considered a site characteristic that potentially increases the productivity of workers. The spatial equilibrium is maintained, however, through variation in rents across local areas. Rauch then uses a reduced form version of the model to empirically test if the level of education has an effect on either the wages or rents in an area. As predicted from the model, metropolitan areas with higher levels of education have higher wages and higher rents. Rauch then provides further evidence that his results are not due to other omitted characteristics of metropolitan areas or higher unobserved ability of workers. He addresses the self-selection issue by testing the additional implications of higher returns to unobserved ability and therefore selection of high ability workers into high skill areas. In particular, Rauch argues that if self-selection of high ability workers into high skill areas is important, then according to the Roy model, there should also be higher wage dispersion, for which he finds no empirical evidence.

Acemoglu and Angrist (2000) extend this earlier literature by focusing on the identification of the causal effect of aggregate levels of education on an individual's earnings. They identify two classes of models that can potentially generate social returns to education. The first they refer to as non-pecuniary externalities. These theories motivate externalities by focusing on the interchanges among workers within a city. The second class of theories is classified as pecuniary externalities. Within this class of models, firms take advantage of the skilled labor pool by investing more heavily in complementary technologies.[26] Both classes of models generate a positive relationship between workers' wages and the level of skill in the area. Therefore, their empirical work cannot support one set of theories over the other.

In their empirical work, they argue that while the correlation between the level of education and the level of earnings across countries or states is undeniably positive, the direction of the causality is not as clear. While high skill areas might lead to high levels of earnings, it is also possible that individuals in areas with high levels of earnings have a stronger taste for education. They rely on cross-state and time-series variation in compulsory schooling laws to identify exogenous variation in the levels of schooling across states and across time. Using instrumental variables, they find a much smaller effect of education on earnings than Rauch. However, the compulsory schooling instruments are likely to pick up variation in schooling at the high school level, since it is the decision to leave high school that is affected by the

[26] Chapter 3 provides an empirical test of this theory.

laws, but will not detect variation in higher levels of schooling. It is still possible that there is an external effect of education at higher levels of education.

Moretti (2004) focuses on the effect of the share of college graduates within a metropolitan area on individual wages using a variety of estimation strategies. His theoretical framework lays out two competing effects of high levels of human capital. The first is the spillover effect, similar to the non-pecuniary theories mentioned in Acemoglu and Angrist. The second effect is imperfect substitution between more and less skilled workers. If there is a standard downward sloping demand curve for workers, less skilled workers will have higher wages if the concentration of highly skilled workers increases. Conversely, more skilled workers will have lower wages. Given his empirical strategies, it is impossible to separately identify the effect of the spillover and the effect of imperfect substitution between types of workers. In order to address this issue, he generates separate estimates of the combined spillover/imperfect substitution effect for differently skilled workers.

Using NLSY data, he estimates a model that simultaneously controls for unobserved ability and a worker-city match by including a fixed individual-city match effect. Identification of the social returns to education is off of an annualized measure of the increase in the college share in a metropolitan area between 1990 and 2000 for workers who remain in the same metropolitan area. Additionally, Moretti uses Census data to estimate the social returns to education, addressing two additional concerns. First, he directly controls for city level productivity shocks, using a Katz & Murphy demand shift measure. Second, he instruments for college share, using the age structure of the metropolitan area and an indicator for the existence of a land grant college. Moretti's results are similar to those estimated by Rauch. While his results control for the selection issue by controlling for an individual-city match, identification of social returns to education is coming from the change to the metropolitan area's college share. The only data available to compute college share are from the 1990 and 2000 Censuses. Moretti then interpolates between these two data points to construct an annual time series of metropolitan area college share. Identification of the social returns to education is off of variation in the annual changes of this measure.

4.2 Theoretical Model and Estimation Strategy

The existing theoretical literature on social returns to education suggests that a worker's productivity varies across metropolitan areas based on the skill level of the metro area in which the worker is located. Under the assumption that firms make zero profits, this variation in productivity leads to variation in workers' wages. The following model, based heavily on earlier models by Lemieux (1998) and the comparative advantage model within Gibbons, Katz, Lemieux and Parent (2002),[27] expands on this premise by decomposing the difference in a worker's wages across metropolitan area skill levels into components due to observable worker characteristics, unobservables, and a level difference. In the context of the education externalities literature, there is not a consensus on how worker characteristics are rewarded in differently skilled metro areas. In Moretti's theoretical work, increases

[27] Lemieux (1998) focuses on wage differences between workers in the union and non-union sectors while Gibbons, Katz, Lemieux, and Parent (2002) focus on industry and occupation wage differentials. Gibbons, Katz, Lemieux and Parent (2002) extend the comparative advantage model to include learning of worker's productivity over time. This extension of the model provides for endogenous movement of workers across sectors. Learning is not included in the present model, in part, because of the estimation requirement of an additional period of wage information for each of the workers in the panel.

in the number of highly educated workers have two competing effects: a spillover effect and imperfect substitution. Therefore, on net, the effect of concentrations of highly skilled workers will be felt most strongly for less skilled workers. One could imagine, however, in the pecuniary models laid out by Acemoglu and Angrist (2000) that more skilled workers might be rewarded more, as they possess the skills most complementary to the increased investment in physical capital. Regardless of the direction in which returns are rewarded differently, if these differences exist, workers are likely to select into the metro area type in which they have a comparative advantage. Not accounting for the differential returns and selection of workers into areas based on their observable and unobservable skills will lead to biased estimates of the wage gap between high skill and low skill metro areas.

To be more precise on how worker characteristics might affect their productivity and therefore their wages, a worker's log wages can be expressed as

(1) $$y_{ijt} = X_{it}\beta_j + \psi_{ij}$$

where *y* is worker *i*'s log wages in metro area *j* in period *t*, *X* is composed of observable worker characteristics, and ψ is an error term. ψ can be further decomposed as

(2) $$\psi_{ij} = c_j + b_j\theta_i$$

where *c* is a metro area specific intercept and θ are unobservable worker characteristics that are differentially valued across metro areas as evidenced by the *j* subscript on *b*.

While the above model could be estimated allowing the coefficients to vary across metropolitan areas, in the empirical work metropolitan areas are classified for simplicity into two skill types, high or low. Combining the wage equations for high and low skill metropolitan areas leads to

$$(3) \qquad \ln w_{it} = H_{ijt} y_{it}^{H} + (1 - H_{ijt}) y_{it}^{1-H} + u_{it}$$

where H are the high skill metro areas, (1-H) are the low skill metro areas and u is an idiosyncratic error term. Following Lemieux (1998), a crucial identification assumption for panel data estimation is that u is mean zero conditional on θ, and all leads and lags of x and H:

$$(4) \qquad E(u_{it} \mid x_i, H_i, \theta_i) = 0$$

where $x_i=\{x_{i1},\ldots,x_{iT}\}$ and $H_i=\{H_{i1},\ldots H_{iT}\}$ where T is the length of the panel. This strict exogeneity assumption rules out models in which workers move between metropolitan areas on the basis of productivity shocks.[28] This assumption and the resulting implications for worker mobility between metropolitan areas are the main impetus for focusing some of the regressions below on workers who are likely to have left their previous jobs involuntarily.

Focusing on the two sectors and setting c_{1-H}=0, β_{1-H} =1 and b_j= $\beta_j\, k$, the wage equation becomes

$$(5) \qquad \ln w_{it} = H_{it}\partial_H + H_{it} X_{it} \beta_H + (1 - H_{it}) X_{it} + H_{it}\theta_i \beta_H k + (1 - H_{it})\theta_i k + u_{it}$$

where ∂_H is the wage gap between H and *(1-H)* metropolitan areas, k is the proportionality factor, i.e. the relative importance of unobservable skills, and u is the exogenous error term. This specification allows for different returns to unobservables between high and low skill metropolitan areas, but restricts this relationship to be

28 See Lemieux (1998) for further discussion of which wage models are and are not consistent with the strict exogeneity assumption.

proportional to the difference to the returns to observables in high and low skill metropolitan areas.

It is important to note that the θ used here is not directly comparable to the θ used in the earlier chapters. The θ from the two preceding chapters is estimated from a wage equation first specified in Abowd, Kramarz, Margolis (1999) and captures all fixed characteristics of the worker valued by the firm. Their wage equation also contains a term capturing fixed firm characteristics that affect wages, thereby separately capturing worker and firm heterogeneity. The θ in equation 1 above is similar in that it captures fixed worker characteristics, but observable skills are already controlled for, so these fixed worker characteristics are limited to unobservables. In addition, the above equation does not separately control for firm effects, but in the empirical work below, does include a set of firm characteristics.[29] Separating out the observable and unobservable worker characteristics allows for different returns to the two different types of worker skills. While by definition, the only difference between the two types of characteristics is what is and is not observable to the econometrician, in practice, characteristics unobservable to the econometrician are likely to be difficult for employers to quantify. Allowing for different returns between observable and unobservable characteristics gives an empirical test of whether or not these two different types of skills are valued differently in the labor market.

---

[29] While adding a firm heterogeneity term to the above equation is a potentially interesting extension of the research set out here, it is left for future work. Among the additional issues that extending the research in this direction brings up is that measuring firm effects requires a longer panel of data and the metro area education measures are only available in Census years.

Equation 5 as written above is not estimable because of the θ term, which is by definition unobservable. In models in which the return to unobservables are held constant across differently skilled areas, this term can be dealt with by using standard panel data techniques using either a fixed worker effect or first-differencing the wage equation. These techniques, however, do not allow for the returns to unobservables to vary. Both Lemieux (1998), focusing on the union/non-union wage gap, and Gibbons, Katz, Lemieux, and Parent (2002), focusing on industry and occupation wage differentials, have provided variants on these standard estimation strategies which identify all of the parameters of this type of model. In order to derive the final estimating equation, one must first solve equation 5 for the unobservable skill term.

$$(6) \qquad \theta_i = \frac{\ln w_{it} - (H_{it}\partial_H + H_{it}X_{it}\beta_H + (1-H_{it})X_{it} + u_{it})}{k(H_{it}\beta_H + (1-H_{it}))}$$

Equation 6 is then lagged one period and the resulting equation is plugged into equation 5 resulting in the final equation

$$(7) \qquad \ln w_{it} = H_{it}\partial_H + H_{it}X_{it}\beta_H + (1-H_{it})X_{it} + k(\frac{H_{it}\beta_H + (1-H_{it})}{H_{it-1}\beta_H + (1-H_{it-1})})$$
$$(\ln w_{it-1} - (H_{it-1}\partial_H + H_{it-1}X_{it-1}\beta_H + (1-H_{it})X_{it-1})) + e_{it}$$

$$(8) \qquad e_{it} = u_{it} - (\frac{H_{it}\beta_H + (1-H_{it})}{H_{it-1}\beta_H + (1-H_{it-1})})u_{it-1}$$

Estimating equation 7 via non-linear least squares will yield inconsistent estimates due to the correlation of lagged wages with the error term. Both Lemieux (1998) and Gibbons et al (2002) suggest similar instrumenting strategies to overcome this endogeneity problem. In particular, the panel data is further utilized to find instruments correlated with lagged wages that are uncorrelated with innovations to current wages. Both papers suggest using the interaction between sector affiliation in time t and $t-1$, or in the context of the current model, metro area skill type in t and $t-1$. Under the strictly exogeneity assumption, that u_{it} is exogenous to the leads and lags of the independent variables, the sector histories will also be uncorrelated with u_{it}. Changes in sector history, however, are likely to be correlated with lagged wages, as workers are likely to move between sectors in response to their wage levels. Sector history remains a valid instrument as long as workers are not moving between sectors in response to productivity shocks. In this case, the workers are responding to changes in their expected wages in sectors, and their lagged wages and current wages are likely to be correlated with the change in sector affiliation, violating the exogeneity restriction necessary for a valid instrument.

Within this chapter, the concern that productivity shocks are inducing the movement of workers across metropolitan areas are addressed by estimating equation 7 separately for a sample of workers likely to be involuntary switchers. In particular, the data used here contains information not only on a worker's wages but also on the worker's firm. If the worker's firm suffered a large job destruction rate in the period that the worker left the firm, it is likely that the worker is moving involuntarily as the result of a mass layoff. Other papers have used similar strategies focusing on the level of job destruction to define periods likely to be mass layoffs.[30]

In addition to directly measuring the differences in the returns to unobservables across differently skilled metropolitan areas, this chapter differs from the existing literature on education externalities in other ways. Most of the existing literature uses either the average education level or the percentage of college graduates within a metropolitan area to define its skill. Here, metropolitan areas are

[30] See Jacobson, LaLonde, and Sullivan (1993) and Bowlus and Vilhuber (2002).

divided into two groups, high and low skill. These two groups were chosen in order to limit the number of metro area types being considered. This approach has advantages and disadvantages. While some of the information about metropolitan areas is lost, chapter 2 highlighted the fact that local labor markets can largely be qualified as falling into one of two skill groups. One clear disadvantage is that it is more difficult to quantify the effect of a given increase in the share of college graduates on the wage premium of high skill areas. Future research can directly test whether one approach fits the data better than the other.

The other key difference is that others have largely focused on identifying the effect of city skill off of longitudinal changes in the skill of an area and focused on workers who remain in the same city. In the results that follow, the variation in city skill is all in the cross-section, and its effect is measured off of workers who switch the metropolitan area in which they work. There is some concern that workers who change jobs are a non-random sample of workers. This concern is partially addressed by the focus on the sample of workers who were likely to have faced mass layoffs. As for the source of variation in city skill, there are advantages and disadvantages to focusing on cross-sectional differences. The main disadvantage is that city skill is largely a permanent characteristic of a city and is highly correlated with other characteristics of the city, which are likely to effect wages, such as industrial mix. Some of the characteristics are directly controlled for in the wage equations below, but the potential of additional omitted city characteristics remains. The main advantage to using the cross-sectional variation is that this is where the bulk of the variation exists. In chapter 2, it was found that over the 1990s all local labor markets appear to becoming more skilled at similar rates. Additionally, variation in city skill in the cross section is likely due, in part, to historical factors that are exogenous to workers current decisions on where to work. Changes to the college share, on the other hand, are likely more strongly correlated with workers current decisions.

4.3 Data

All of the data used in this research are part of the Longitudinal Employer-Household Dynamics program at the Census Bureau. Information on workers comes from the Unemployment Insurance wage records for the selected three states.[31] These files contain person identifiers that allow one to track a worker's earnings. The UI wage records contain virtually all business employment for the states included in the analysis. The disadvantage of using the UI wage data to characterize workers is the very limited demographic information available. Within the Census bureau, this problem has been partially overcome by combining the UI wage data with other administrative data containing information on date of birth, place of birth, and gender. Additionally, this research utilizes a matched sample with the Decennial Census that allows for a richer set of controls for observable skill.

The quasi-first-differencing estimation strategy requires a panel of two jobs for each worker for estimation. The sample of workers was chosen as the set of all workers in the three states who matched to the 2000 Decennial Census Sample, held two jobs between 1999 and 2001, and were between the ages of 25 and 65. The date range was chosen to restrict the sample to be close to 2000 so that the education

[31] Three states were selected on the basis of time-series availability at the time of project inception. This research cannot reveal the identity of the three states used in the analysis due to confidentiality restrictions.

information for individual workers and for metropolitan areas remained accurate. Choosing additional years further from 2000 would require additional assumption on the evolution of both individuals' and metropolitan areas' skill.

The UI wage records contain information on a workers quarterly wages but do not have information regarding the hours or weeks worked. Only earnings at the dominant job is used where the dominant job is defined as the job at which the worker had the greatest earnings within a quarter. An annualized wage measure is constructed for each worker at each job as follows: if a worker worked one full quarter at a job his annual wage is four times is full quarter earnings, if a worker worked two full quarters at a job his annual wage is 2 times the sum of the two full quarters of earnings, if a worker worked 3 full quarters at a job then the annual wage is four thirds the sum of the three full quarters of earnings, if a worker worked 4 full quarters at a job then the annual wage is the sum of the four full quarters of earnings.[32] The number of full quarters used to construct the annual wage is then included as an additional control in all of the wage equations.

Following the approach used by Gibbons et. al (2002), a skill index is created to control for the observable skills of workers. Including the skill index allows for a rich characterization of workers' skills without increasing the number of parameters to be estimated in the final wage equation. The index is created from a regression of log wages on human capital characteristics (high school, some college and college graduate indicators, potential experience and its square) and controls (metro area, gender, sic division, indicator for a small firm, time dummies, and indicators for the number of full quarters used to calculate the wage measure). The fitted value from

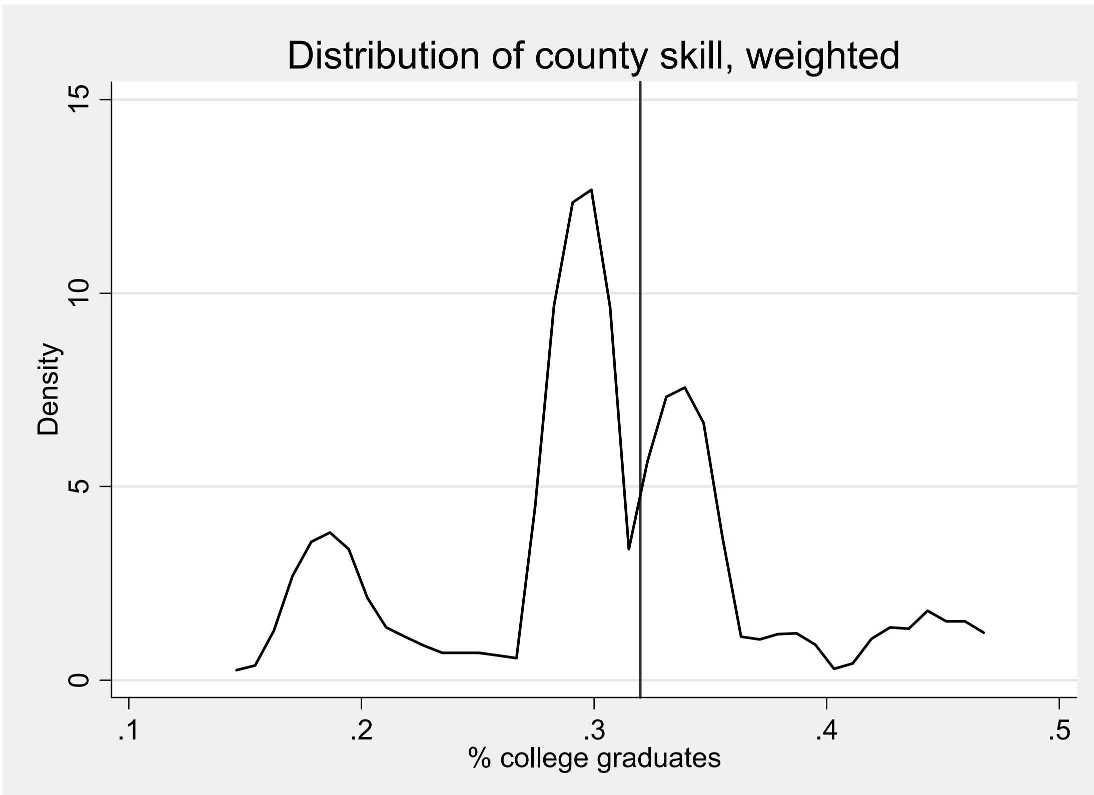
[32] Quarter t is defined as a full quarter if a worker is employed by the same firm in t-1, t, and t+1.

the above regression is then calculated using the workers' values for human capital characteristics and average values over the sample for the controls. The measure is then demeaned. Coefficients from the wage regression are listed in Appendix B.

A basic model of social returns to education suggests that an individual worker's wages are a function of his own characteristics and the education level of other workers. The existing theoretical literature provides little direction in defining the set of other workers whose education level is relevant. Within the empirical literature, different geographic regions used for this purpose include the country, the state, and the metropolitan area. In order to be comparable with the existing literature, this chapter also focuses on the education level of the metropolitan area. Choosing the metropolitan area suggests that an individual is most influenced by the workers with whom he is likely to interact. This definition is consistent with both the non-pecuniary spillovers literature in which workers interactions are the key and with the pecuniary spillovers literature in which businesses make investment decisions based on the skill level of the available labor pool.

Counties which do not fall into metro areas are aggregated to state-level non-metro areas. In the current specification, metropolitan areas are classified as high skill or low skill areas. In order to classify metro areas by skill, the share of workers who are college graduates is calculated for each metro area. Figure 4.1 plots an employee-weighted kernel density estimate of this metro area skill measure. As can be seen in the graph, the distribution appears to be bi-modal. The minimum between the two modes occurs at 32% of the workers in a metro area with a college degree. This number provides a natural break between different types of metro areas and is therefore used as the cutoff skill level to define high and low skill metro areas. The results are not directly comparable to the existing literature because the metropolitan area skill is not a continuous variable in the following regressions.

Figure 4.1: Distribution of employee weighted college share across metropolitan areas.



<u>*4.4 Results*</u>

Table 4.1 provides the mean and standard deviation of the variables used in the regressions. As can be seen from the table, roughly half of the workers are in a metropolitan area that can be classified as high skill and half are in a low skill metropolitan area. Movements across the metropolitan areas in the two jobs for each worker are surprisingly symmetrical. Around 41% of workers are in a high skill metropolitan area in each of their two jobs and 42% are in a low skill metropolitan area in each job. Similarly, 8% of workers move from a high to a low skill metropolitan area and 9% move from a low to a high skill metropolitan area. The skill index is mean zero by construction. Time dummies are provided for each job to control for the fact that the workers wages are coming from different time periods and aggregate wage levels may be changing over the three-year time period. Finally, wage controls capture the number of full quarter earnings used to construct each of the annual earnings measures for each job. Over half of the jobs have earnings from four full quarters.

Table 4.1: Summary Statistics

Variable	Mean	Std Dev
Log annual wage	10.4183	0.7769
High skill msa	0.4978	
Low skill msa	0.5022	
High high msa	0.4142	
High low msa	0.0803	
Low high msa	0.0869	
Low low msa	0.4185	
Worker characteristics:		
Skill	0.0000	0.3098
Experience	18.9595	8.9612
Male	0.5257	
Firm characteristics:		
Agr, mining, constr	0.0641	
Manufacturing	0.1592	
TCU	0.0666	
Wholesale	0.0702	
Retail	0.1272	
FIRE	0.0884	
Services	0.4041	
Public Admin	0.0202	
Time dummies:		
99:1	0.0708	
99:2	0.0723	
99:3	0.0840	
99:4	0.1178	
00:1	0.1295	
00:2	0.1374	
00:3	0.1246	
00:4	0.0997	
01:1	0.0559	
01:2	0.0649	
01:3	0.0431	
Wage controls:		
1 fq job	0.1672	
2 fq job	0.1471	
3 fq job	0.1087	
>=4 fq job	0.5770	

Table 4.2 lists some basic regression results that measure the effect of controlling for different worker and firm characteristics on the wage differential between high and low skill metropolitan areas. All regressions include time dummies and controls for the number of quarters used to construct the wage measure. The first column only includes a high skill metro area indicator along with the wage controls

and time dummies. The raw wage differential measured from this equation is 25% and is strongly significant. The second column includes both worker and firm characteristics. In this regression the wage differential is reduces to 16%. In the third column, the coefficient on the individual workers skill level is allowed to vary in high and low skill metropolitan areas.   While the coefficient on county skill suggests that the returns to skill are 6% higher in high skill metropolitan areas than low skill metropolitan areas that have an implicit coefficient of one, the wage differential remains unchanged from the earlier specification. Column four shows the results from a fixed person effects regression. Controlling for unobservable worker skills reduces the size of the wage differential substantially to approximately 4%. The differential between the returns to observable skills in low and high skill metropolitan areas decreases to 0.5% in this specification. Although this regression controls for unobservable skills, the returns to unobservable skills are restricted to being the same in high and low skill metropolitan areas.

Table 4.2: OLS, Fixed Effect, Difference in Differences

	1	2	3	4	5	6
	raw	comp	ols	fe	ols	
					job1	job2
Intercept	10.6584	10.3679	10.3680		10.1309	10.3106
	0.0049	0.0063	0.0063		0.0087	0.0078
High skill msa	0.2498	0.1605	0.1605	0.0350		
	0.0017	0.0014	0.0014	0.0019		
Low skill msa	0.0000	0.0000	0.0000	0.0000		
Skill		1.0000				
		0.0023				
High skill msa*skill			1.0558	1.0049	1.0677	1.0654
			0.0048	0.0014	0.0070	0.0068
Low skill msa*skill			1.0000	1.0000	1.0000	1.0000
High-high msa					0.1871	0.1865
					0.0022	0.0022
High-low msa					0.1388	0.1163
					0.0038	0.0038
Low-high msa					0.1111	0.1564
					0.0037	0.0037
Male		0.3829	0.3825		0.3828	0.3797
		0.0015	0.0015		0.0020	0.0020
Female		0.0000	0.0000		0.0000	0.0000
Firm characteristics:						
Agr, mining, constr		0.2421	0.2415	-0.0277	0.2381	0.2430
		0.0057	0.0057	0.0056	0.0087	0.0076
Manufacturing		0.2336	0.2324	0.0172	0.2369	0.2259
		0.0052	0.0052	0.0050	0.0081	0.0069
TCU		0.2484	0.2481	-0.0052	0.2511	0.2361
		0.0056	0.0056	0.0054	0.0086	0.0074
Wholesale		0.2433	0.2426	-0.0097	0.2412	0.2399
		0.0056	0.0056	0.0052	0.0085	0.0074
Retail		-0.1694	-0.1700	-0.1643	-0.1858	-0.1506
		0.0053	0.0053	0.0050	0.0082	0.0070
FIRE		0.2842	0.2836	-0.0182	0.2758	0.2820
		0.0055	0.0055	0.0053	0.0084	0.0072
Services		0.0588	0.0579	-0.0822	0.0400	0.0751
		0.0051	0.0051	0.0046	0.0078	0.0066
Public Admin		0.0000	0.0000	0.0000	0.0000	0.0000
Small Firm Indicator		-0.1384	-0.1387	-0.0627	-0.1491	-0.1228
		0.0016	0.0016	0.0014	0.0022	0.0023
Person effect						
R-Squared	0.07	0.34	0.34	0.89	0.28	
Observations	819802	819802	819802	819802	819802	

Standard errors are listed directly below the parameter estimates. The coefficient on skill-metro area interaction is normalized to one in low skill metro areas.

Table 4.2: continued

	Difference in differences		
	job1	job2	diff
high-high	0.1871	0.1865	-0.0006
high-low	0.1388	0.1163	-0.0225
			0.0219
			0.0054
	job1	job2	diff
low-high	0.1111	0.1564	0.0453
low-low	0	0	0
			0.0453
			0.0052

Columns 5 and 6 show the results from a specification in which each of the right hand side coefficients is interacted with an indicator for job 1 or job 2. In lieu of the high skill metropolitan area indicator, there are four indicators for the four different groups of worker movement across metropolitan area types: high to high, high to low, low to high, low to low (omitted group). These coefficients follow a predictable pattern and will be discussed more fully in the difference in differences estimates below. The returns to observable skill are similar to the earlier OLS results. The remaining coefficients are for the firm characteristics and are consistent in sign and magnitude across all specifications except for the person fixed effects regression.

From these estimates difference in differences estimates are calculated and listed at the end of the table. The first difference is across job 1 and 2 (columns 5 and 6) separately for each type of worker movement, i.e. high to high skill metro area, high to low skill metro area, and low to high skill metro area. Workers who move between jobs that are both in low skill areas are the omitted group. This difference separates out the component of the change in a workers wages when changing jobs that is due to the type of metropolitan area in which the two jobs are located. The second difference is across types of movement done separately by the originating metro area type. This difference separates out the effect of changing jobs for workers in low skill areas who move to high minus those who stay in low or workers in high skill areas who stay in high minus those who move to low. Constructing the difference in difference in this fashion allows for variation in worker types and job opportunities across the originating metro area skill types.

If movement across metropolitan area types were exogenous, measuring the effect of a high skill metropolitan area on a workers wages would be the same whether the worker moved into or out of a high skill area. Focusing on the difference in difference results in bold at the bottom of table 4.2, workers who move from a low to a high skill metropolitan area receive a 4.5% wage premium over those who stayed in a low. However, workers who stay in a high skill metropolitan area receive a 2.2% wage premium over those who move into a low skill metropolitan area. Given that the returns to high skill metropolitan areas are higher for workers who move toward high skill metropolitan areas, it is likely that worker mobility across metropolitan areas is not completely exogenous. Put differently, workers who leave a high skill metropolitan area receive less of a penalty for leaving than the bonus received by workers who leave low skill metropolitan areas. This pattern of wage changes suggests that workers currently in high skill metropolitan areas only move to low skill metropolitan areas when they find jobs with wages that are attractive relative to their current earnings in a high skill metropolitan area.

There are multiple ways to address endogenous mobility. Gibbons et al (2002) directly instrument for the choice of sector in time t by using the interaction between sector choice in time t-1 and t-2 as instruments. With the data set used here, it is unclear how to construct a long time series of metro area affiliations. As an alternative approach, the sample of workers is subset to the sample of moves that are likely to be exogenous[33]. The administrative data provides no information from the worker as to why the worker left a job. However, the employee-employer match aspect of the data can be utilized to define firms that appear to be laying off large portions of the workforce. Previous research has similarly used periods of large job destruction to identify displacement activity.[34] These sharp reductions in the number of employees are likely to lead to worker separations that are exogenous to the decisions of the worker. Two additional samples of workers are constructed to try to capture these exogenous separations: workers who leave a firm at time t in which the firm's employment falls by 20% between t and t+1 and workers who leave a firm that shrinks by 10% over the same time period. The samples are 4% and 8%, respectively, of the size of the original sample. While the 20% cutoff is more in line with the literature on mass layoffs, the 10% cutoff is likely to be more representative of the full sample of workers.

Table 4.3 repeats the specifications in columns 5 and 6 of table 4.2 for the new samples described above. The corresponding difference in differences is shown at the bottom of the table. For the sample with 20% job loss at the firm, workers who move from a low to a high skill metro area receive a 3.5% wage premium over those who remain at low skill metropolitan areas, while workers who remain in a high skill

[33] Lemieux (1998) uses a similar approach to identify a union premium with a dataset that contains information on the reason why a worker left his last job as reported by the worker.
[34] Jacobson, LaLonde, and Sullivan (1993) identify a "mass layoff" sample by focusing on workers who leave a firm in which "the firms' employment in the year following their departure [in the early and mid 1980s] was 30-percent or more below their maximum levels during the late 1970's." Lengermann and Vilhuber (2002) identify a mass layoff by focusing on workers who leave a firm in which there is a 30% reduction in jobs from one quarter to the next. The measure used here more closely follows Lengermann and Vilhuber.

metropolitan area receive a 1.6% wage premium. Similarly, for the sample with 10% job loss at the firm, workers who move from a low to a high skill metropolitan area receive a 4.0% wage premium while workers who remain in a high skill metropolitan area receive a 3.0% wage premium over those who leave. The difference between the two estimates is smaller for either of these two samples suggesting that a greater portion of the moves in the layoff samples is the result of involuntary separations. While the difference between the two estimates is similar for the two layoff samples, the levels of the premiums are smaller in the sample defined by firms shrinking 20%. Other research that has focused on the earnings outcomes of workers finds similar results with wage decreases resulting from a layoff being on the order of 25%.[35] In the second sample, the premiums are more in line with those found in table 4.2, and the difference between the workers who leave and enter metropolitan areas is similar. Due to the tradeoffs in using each of the samples, both are used to test the sensitivity of the final results.

[35] See Jacobson, LaLonde, and Sullivan (1993). In addition to using a different definition of mass layoffs, Jacobson et al also motivate their research with a theoretical model that is different from the one being considered here.

Table 4.3: Difference in Differences, Layoff Sample

	1	2	3	4
	jdr>0.2		jdr>0.1	
	job1	job2	job1	job2
Intercept	9.9107	10.1457	9.9836	10.2051
	0.0531	0.0317	0.0330	0.0222
High skill msa*skill	1.0724	1.0791	1.0893	1.0897
	0.0271	0.0283	0.0187	0.0188
Low skill msa*skill	1.0000	1.0000	1.0000	1.0000
High-high msa	0.2041	0.2087	0.1986	0.2053
	0.0082	0.0081	0.0057	0.0057
High-low msa	0.1833	0.1715	0.1699	0.1467
	0.0145	0.0145	0.0098	0.0098
Low-high msa	0.1402	0.1749	0.1277	0.1680
	0.0145	0.0145	0.0099	0.0099
Male	0.4101	0.4024	0.3951	0.3868
	0.0078	0.0078	0.0053	0.0054
Female	0.0000	0.0000	0.0000	0.0000
Firm characteristics:				
Agr, mining, constr	0.3963	0.3354	0.3621	0.3160
	0.0522	0.0300	0.0325	0.0212
Manufacturing	0.3776	0.2742	0.3141	0.2350
	0.0520	0.0292	0.0321	0.0204
TCU	0.3598	0.2811	0.3362	0.2564
	0.0537	0.0316	0.0334	0.0220
Wholesale	0.4214	0.2675	0.3555	0.2413
	0.0530	0.0311	0.0329	0.0216
Retail	-0.0686	-0.1680	-0.1016	-0.1469
	0.0525	0.0300	0.0323	0.0208
FIRE	0.5820	0.3528	0.4824	0.3304
	0.0532	0.0308	0.0328	0.0213
Services	0.2254	0.1485	0.1808	0.1233
	0.0515	0.0282	0.0317	0.0198
Public Admin	0.0000	0.0000	0.0000	0.0000
Small firm indicator	-0.1268	-0.1103	-0.1336	-0.1129
	0.0076	0.0079	0.0053	0.0056
R-squared	0.27		0.27	
Observations	62114		125854	

Standard errors are listed directly below the parameter estimates. The coefficient on skill-metro area interaction is normalized to one in low skill metro areas.

Table 4.3: Difference in Differences, Layoff Sample, continued

	Difference in differences						
	jdr>0.2				jdr>0.1		
	job1	job2	diff		job1	job2	diff
high-high	0.2041	0.2087	0.0046	high-high	0.1986	0.2053	0.0067
high-low	0.1833	0.1715	-0.0118	high-low	0.1699	0.1467	-0.0233
			0.0164				**0.0300**
			0.0205				0.0139
	job1	job2	diff		job1	job2	diff
low-high	0.1402	0.1749	0.0347	low-high	0.1277	0.1680	0.0403
low-low	0	0	0	low-low	0	0	0
			0.0347				**0.0403**
			0.0205				0.0139

Table 4.4 provides estimates of equation 3 in which comparative advantage in the unobservables across metropolitan area types is allowed by incorporating an interaction between the high skill metro area dummy and the worker effect. The first column contains results using the full sample of workers. In the second column, firm controls are dropped in order to determine if the wage premium of high skill metropolitan areas is affected by the composition of firms. In the third column, non-metro areas are excluded from the analysis. The purpose of this sensitivity check is two fold. First, non-metro area skill is measured over all counties within a state. This type of area is not directly comparable to a metro area because the counties are not necessarily contiguous and because they are likely to be more heterogeneous than the counties in a metro area. In addition, much of the research on education spillovers focuses on metropolitan areas because by definition metro areas have higher concentrations of workers. Because workers are more concentrated, their interactions amongst themselves are likely to be more frequent than among workers in a non-metro area. Finally, column 4 focuses on a sample in which both the first and second job of the worker has a wage measure that was constructed from four full quarters of earnings.

Table 4.4: Non-linear Two Stage Least Squares

	1 all	2 nofirm	3 msa	4 longjob
High skill msa	-0.0355	0.0360	-0.0044	-0.1588
	0.0497	0.0483	0.0532	0.0764
High skill msa * skill	1.0066	0.9997	1.0034	1.0183
	0.0047	0.0046	0.0050	0.0072
Low skill msa * skill	1.0000	1.0000	1.0000	1.0000
High skill msa * worker effect	1.0144	1.0160	1.0146	1.0162
	0.0003	0.0003	0.0003	0.0006
Firm characteristics:				
seinsmall	-0.0623		-0.0594	-0.0683
	0.0014		0.0015	0.0023
Agr., mining, constr.	-0.0259		-0.0322	-0.0294
	0.0056		0.0060	0.0083
Manufacturing	0.0192		0.0057	0.0197
	0.0050		0.0053	0.0072
TCU	-0.0035		-0.0145	0.0050
	0.0054		0.0057	0.0079
Wholesale	-0.0080		-0.0170	-0.0095
	0.0052		0.0056	0.0076
Retail	-0.1607		-0.1692	-0.1378
	0.0050		0.0054	0.0074
FIRE	-0.0167		-0.0221	-0.0225
	0.0053		0.0056	0.0077
Services	-0.0795		-0.0865	-0.0680
	0.0046		0.0050	0.0066
R-squared	0.59	0.59	0.59	0.67
Observations	409901	409901	377878	127378

Standard errors are listed directly below the parameter estimates. The coefficient on skill-metro area interaction is normalized to one in low skill metro areas.

Focusing on the full sample in column 1, the coefficient on the high skill metro area indicator captures the wage differential between high and low skill metropolitan areas after allowing for different returns to both observable and unobservable skills. The coefficient is negative but not significantly different from zero. The coefficient on the interaction between high skill metropolitan area and the index of the workers observables skills is greater than one, but not significantly. These results suggest that workers in high skill and low skill metro areas receive slightly larger returns to their observable and unobservable skills. The coefficient on the interaction between the high skill metro area indicator and the worker’s fixed effect is significantly different from one, suggesting the gap between the returns in high and low skill metro areas is greater for unobservable skills than it is for observable skills. While the relative returns to unobservable skill are restricted to be proportional to the relative returns to observable skill, this coefficient provides the proportionality constant between the two types of returns. The remaining coefficients on firm characteristics are of similar sign and magnitude in comparisons across the columns and with the earlier person fixed effects results in table 4.2.

The results estimated on the remaining samples in table 4 show mixed results. In column 2, the firm characteristics are dropped from the specification. While the dummy for high skill msa is still not significantly different from zero, it is positive. The coefficient on the interaction of high skill metro area and skill index is now less than one. Despite the fact that the key coefficients are not significant, the firm characteristics appear to play an important role suggesting that differences across metro areas may be due, in part, to firm characteristics. Excluding non metro areas in column 3 does not appear to have much affect on either of the two key variables. Focusing on jobs in which workers have greater attachment in column 4 does lead to different results. The coefficient on the high skill metro area dummy is large and negative. The interaction of high skill metro area and the skill index is also significantly greater than one suggesting that the returns to skill are higher in high skill metro areas. Although not reported in here, similarly restricting the earlier fixed effect specification to this sample does not affect the results in the same way. It appears that the nonlinear estimation is more sensitive to the wage measure. Although the results are sensitive to the long job restriction, the basic pattern of the results is the same and, therefore, the remaining estimation relies on the full sample. The proportionality factor measuring the gap between high and low skill metro areas on the returns to unobserved ability relative to observed ability is consistently greater than one and very similar across all the specifications.

While the coefficient measuring the wage gap between high and low skill metro areas seems inconsistent with the earlier specifications and most of the literature, this measure of the wage gap is different than the other measures and difficult to interpret. In particular, this coefficient measures the wage gap after allowing for different returns to observable and unobservable skills. Given that the wage gap is a residual measure of the wage differences between workers in high and low skill metro areas after accounting for worker and firm characteristics, one would expect that the results in table 4 would be different from the earlier results. These results suggest that more skilled workers, on both observable and unobservable dimensions, receive greater benefits from working in a high skill area than less skilled workers, although the results are imprecisely measured. Given that some of the wage gap between the two types of areas may be accounted for by different returns to unobservables, one would expect the wage gap to be smaller as measured in table 4.4. The negative coefficients and the inconsistency across columns are surprising, however. Clearly, after accounting for the worker and firm characteristics and allowing the returns to worker skills to vary, the remaining wage gap is not as well defined. Still, higher returns to both observable and unobservable skills suggest specific patterns of spillovers or additional investment that disproportionately benefits more skilled workers.

As mentioned above, the wage equation as estimated contains a lagged wage term that is instrumented for with sectoral history dummies that consist of interactions between a workers metro area skill in the two periods. Traditional tests of the power of the instruments used in linear two stage least squares involve a test of the significance of the excluded instruments in the first stage regressions. The appropriate first stage equation in non linear two stage least squares, however, involves regressing the endogenous variable on the gradient of the second stage error term with respect to the structural parameters.[36] In lieu of estimating the appropriate, highly non-linear first stage equation, the linear first stage equation is estimated as an approximation. The F tests of the excluded instruments are presented in table 4.5. Across all of the different samples, the sectoral history dummies strongly predict lagged wages.

Table 4.5: F-Test of Excluded Instruments

	all	men	nofirm	msa	longjob
F-stat	70.05	19.77	103.03	90.25	20.26

Table 4.6 contains results from the two additional samples designed to focus on workers who are more likely to have exogenously left their previous employer. The results in column 1 are different than the earlier coefficients. The wage gap between high and low skill metro areas is now large and positive, although still not significantly different from zero. The coefficient on the interaction between high skill metro area and observable skills is not significantly different from one suggesting that the returns to these skills are the same in high and low skill metro

[36] See the appendix of Gibbons, Katz, Lemieux, and Parent (2002) for a more complete explanation.

areas. The coefficient on the interaction between high skill metro area and the worker fixed effect is significantly different from one, but the interpretation is difficult given that it is proportional to the difference in the returns to observable skills which are not different in high and low skill metro areas. In contrast with column 1, the results in column 2 are more similar to those in table 4.4. Although the coefficient on the interaction between high skill metro area and observable skills is not significantly different from 1, all of the coefficients are of the same sign and magnitude as those in table 3. These results provide somewhat contradictory evidence on the importance of focusing on exogenous moves. The odd results in column 1 may also be due to the fact that workers leaving firms with mass layoffs receive a wage penalty. Previous research on mass layoffs by Jacobson et al (1993) attribute this wage penalty to loss of firm-specific knowledge or a reshuffling of workers into lower quality job matches. Neither of these theories is currently incorporated into the wage model laid out above. Extension of the model in these directions in order to incorporate the empirical finding that wage losses occur with separations after mass layoff is left to future research. Given that the sign and magnitude of the coefficients in column 2 follows the same patterns as that in table 4, the insignificant sign on the interaction of high skill metro area and observable skill may be due to the smaller sample size. Table 4.7 provides the F-statistics of the excluded instruments from the linearized first stage regression. Even in the smaller samples used in table 6, the sectoral history dummies strongly predict lagged wages.

Table 4.6: Non-linear Instrumental Variables, Layoff Sample

	1	2
	jdr>0.2	jdr>0.1
High skill msa	0.1652	-0.1775
	0.2129	0.1399
High skill msa*skill	0.9866	1.0205
	0.0203	0.0133
High skill msa *	1.0157	1.0156
worker effect	0.0010	0.0007
Firm characteristics:		
Small firm indicator	-0.0728	-0.0677
	0.0053	0.0036
Agr., mining, constr.	0.0595	-0.0011
	0.0238	0.0160
Manufacturing	0.0987	0.0330
	0.0230	0.0153
TCU	0.0719	0.0205
	0.0248	0.0164
Wholesale	0.0643	0.0058
	0.0239	0.0158
Retail	-0.0790	-0.1379
	0.0234	0.0154
FIRE	0.0758	0.0018
	0.0244	0.0161
Services	0.0084	-0.0525
	0.0219	0.0145
R-square	0.52	0.57
Observations	31033	62903

Standard errors are listed directly below the parameter estimates.

Table 4.7: Test of Excluded Instruments, Layoff Sample

	jdr>0.2	jdr>0.1
F-stat	11.83	17

The results in tables 4.4 and 4.6 are also different than those found by Moretti in which the returns to changes in the college share were felt most by low skill workers. Moretti's work suggests that although the spillover is important for all worker groups, as the number of high skill workers increase their returns decrease, and therefore the net effect of an increase in college share is smaller for more skilled workers. The results above may be different for a variety of reasons. First identification of the importance of metro area skill is determined by utilizing the variation in a cross section as opposed to over time, as in Moretti's work. Therefore, in Moretti's specification, imperfect substitution between workers of different skill types is likely playing a larger role. Within his specification, influxes of more skilled workers need to be absorbed by the metro area. It is possible that in the short run the imperfect substitution plays a more dominant role, while the cross section is identifying a long run phenomenon. In the results above, the greater the fraction of high skill workers, the greater the return to skill, suggesting that the productivity of high skill workers increases as their numbers increase. In addition, identification in Moretti's specification is determined by workers who remain in a metro area while in the above results identification is determined by workers who move across metro areas of different skill types. These two different samples may be different among a variety of other dimensions.

4.5 Conclusion

Many researchers have found that workers receive higher wages in highly skilled metro areas. Often, this correlation is interpreted as an external return to education. However, workers in more skilled areas may differ on more than just their observable skills. While others have attempted to control for unobservable characteristics of workers that affect their wages, this is the first research that allows the returns to both observable and unobservable worker characteristics to vary across high and low skill metro areas. Using a panel dataset of workers, non-linear instrumental variables is used to estimate a wage equation with controls for high skill metro area, an interaction between high skill metro area and skill, and an interaction between high skill metro area and the fixed person effect. Identification of a separate return to unobservable skill, captured by the person effect, is determined by workers who change jobs and metro areas.

Allowing for variation in these returns does affect the remaining wage gap between high and low skill metro areas. Workers in high skill areas appear to receive higher returns for both their observable and unobservable characteristic, although the effects are not precisely measured. In addition, the employer-employee match aspect of the data is utilized to construct a sample of workers who are likely to have been laid off of their last job. This group of workers has returns across high and low skill metro areas that are more consistent with exogenous movement of workers. The pattern of returns to observable and unobservable skill found in the wage regressions suggests a pattern of spillovers from education that disproportionately favors highly skilled workers. It is left for future research to determine which models of social returns to education are consistent with this pattern in the data.

Chapter 5: Conclusion

The preceding chapters focus on different aspects of geographic concentrations of skilled workers. The second chapter focused on describing patterns of skilled workers across geography and how these patterns evolve over time. Despite the fact that workers are highly mobile, they choose locations in non-random patterns. Therefore, the overall distribution of skill across local labor markets remains constant although there is a continuous reshuffling of workers. Given the variation in worker skill across geography, the potential exists for location to play an important role in a variety of different models. While there are many models in which considering the U.S. labor market as one entity remains appropriate, there are others in which the variation in the availability of skilled labor will play an important role.

The two following chapters then outline models that fit the latter category. The third chapter outlines a model in which worker skill and technology investment are complementary. Firms must make investment decisions before hiring workers, and therefore choose a level of investment based on the type of worker they expect to be able to hire. Within this framework, firms invest more in technology in high human capital areas. The empirical work testing this basic prediction of the model affirms this result. A one standard deviation increase in county skill leads to a 10% increase in the amount of computer investment of manufacturing firms in 1992. While the empirical work is focused on one industry in one time period, it is likely that the results are more broadly applicable to other instances in which firms face an exogenous technological shock. In the longer term, firms are likely to choose locations with workers suitable to the best technology. In addition to explaining variation in firm's technology decisions across the United States, the theory is likely also applicable to explaining variation in technology investment across countries.

Finally, the fourth chapter looks at how variation in the concentration of skilled labor in metropolitan areas affects worker wages. The correlation between the overall level of worker wages and the skill level of a metropolitan area is a well-established empirical fact. Two models that would explain this relationship are models in which there are information spillovers and models such as the one outlined in the preceding chapter in which firms' investment depends upon the skill level of the local labor market. However, testing this implication of either of the models above is complicated by the fact that while there is correlation between worker wages and the skill level of the local labor market, it is more difficult to identify a causal effect. Workers are mobile and are likely to choose to work in locations in which their skills are rewarded the most highly, i.e. where they have comparative advantage. Previous research has accounted for this possibility by allowing the returns to workers observable characteristics to vary across metropolitan areas. This chapter extends upon that research by allowing for the returns to unobservables to also vary across metropolitan areas. The results show that workers are more highly compensated for both their observable and unobservable skills in highly skilled metro areas. The remaining wage gap between workers in high and low skill metro areas becomes insignificant. It is left for future research to determine what this pattern of worker returns implies in the theoretical models.

This dissertation has outlined two models in which the variation in skilled labor across local labor markets plays an important role in the economy. Clearly, there are possibly many other models in which this variation would also be important. Further research explaining aggregate trends in productivity and wages might benefit from accounting for variation in the allocation of workers across local labor markets.

Appendix A: Investment Equation Specification

The model suggests that the firm type needs to be included as a control in the capital investment equation. While a pure measure of firm type is not available, we do measure ψ, the firm effect from the wage equation, and use it in estimating equation 10. In order to get a handle on the extent of the potential bias due to this problem, it is first necessary to distinguish between capital investment and capital stock. Due to its limited dynamics, it is impossible to distinguish between investment and capital stock within the model. Empirically, however, it is likely that wages are influenced by the sum of previous period's investments of the worker's firm, or the capital stock. Keeping this distinction in mind, it is clearer if equation 13.10 is re-written as

$$I_{jt} = \phi_0 + \phi_1(\ln A_j^{1/\alpha} + \ln k_{jt-1}) + \phi_2 \hat{s}_{lt-1} + \upsilon_{jt}$$

in which the term in parentheses is the firm effect from the wage equation, I_{jt} is computer investment and k_{jt-1} is the capital stock in the previous period. This equation, which is the one that can be estimated with the available data, is misspecified due to the inclusion of lagged capital stock. The true equation can be re-written as

$$I_{jt} = \phi_0 + \phi_1(\ln A_j^{1/\alpha} + \ln k_{jt-1}) + \phi_2 \hat{s}_{lt-1} - \phi_1 \ln k_{jt-1} + \upsilon_{jt}$$

Written in this form, it becomes clear that equation 3.10 contains an omitted variable, k_{jt-1}. Given that k_{jt-1} is correlated with ψ and with s_{lt-1}, all of the coefficients in the model are potentially biased. While there is no available information on the stock of computer investment in the previous period, a measure of the stock of machinery equipment is available in 1991.[37] The stock of machinery equipment includes the stock of computer equipment, but is a much broader measure. While not the ideal measure, adding the capital stock variable will help to determine the extent to which the specification problem mentioned above is effecting the results, and in particular, the coefficient on county skill. Results are shown in the last column of table 3 and discussed in the text.

Additional biases to the results will occur if one assumes that computer investment is correlated over time in a way that is not captured by either the firm effect, the firm's skill level, nor the local labor market skill level, i.e. serially correlated technology shocks, such that

$$\upsilon_t = \lambda \upsilon_{t-1} + \eta_t$$

[37] For details on the creation of this variable see Chiang (2003). The basic methodology involves using the reported capital stock measure in Census years and applying the perpetual inventory method for non-Census years.

in which λ captures the serial correlation and η is the current period innovation. The coefficient on the firm effect in the original specification will be biased upward because, as shown above, the firm effect contains information on the previous period's investment, which will be positively correlated with the serially correlated component of the error term.

To some extent, the inclusion of the capital stock as an additional control will help to alleviate this problem. If serial correlation of the error term is caused by serially correlated technology shocks, the coefficient on the capital stock measure will be biased upward but there should be no direct effect on the other coefficient's included in the estimation. Further, serial correlation of the error term and therefore the upward bias on the capital stock coefficient will likely bias the coefficients on the other variables downward.

Appendix B: Constructing the Skill Index

	All	Men
intercept	9.5366	9.6838
	0.0076	0.0106
experience	0.0327	0.0469
	0.0004	0.0005
experience^2	-0.0007	-0.0010
	0.0000	0.0000
high school	0.2169	0.2297
	0.0026	0.0035
some college	0.4360	0.4429
	0.0025	0.0033
college	0.9092	0.9511
	0.0026	0.0034
high skill msa	0.1605	0.1827
	0.0014	0.0020
male	0.3829	
	0.0015	
female	0.0000	
Firm characteristics:		
Agr, mining, constr	0.2421	0.2960
	0.0057	0.0078
Manufacturing	0.2336	0.2492
	0.0052	0.0074
TCU	0.2484	0.2497
	0.0056	0.0079
Wholesale	0.2433	0.2476
	0.0056	0.0079
Retail	-0.1694	-0.0920
	0.0053	0.0076
FIRE	0.2842	0.2838
	0.0055	0.0081
Services	0.0588	0.1065
	0.0051	0.0073
Public Admin	0.0000	0.0000
seinsmall	-0.1384	-0.1135
	0.0016	0.0022
Time dummies:		
99:1	-0.3971	-0.3757
	0.0046	0.0064
99:2	-0.3451	-0.3302
	0.0046	0.0064
99:3	-0.2649	-0.2397
	0.0045	0.0062
99:4	-0.2308	-0.2223
	0.0044	0.0060

00:1	-0.2700	-0.2462
	0.0043	0.0060
00:2	-0.2099	-0.1924
	0.0043	0.0059
00:3	-0.1707	-0.1562
	0.0043	0.0060
00:4	-0.1930	-0.1785
	0.0045	0.0061
01:1	-0.1453	-0.1284
	0.0051	0.0071
01:2	-0.0183	0.0056
	0.0050	0.0069
01:3	0.0000	0.0000
Wage controls:		
1 fq job	-0.2918	-0.2849
	0.0022	0.0031
2 fq job	-0.1879	-0.1926
	0.0025	0.0034
3 fq job	-0.1265	-0.1259
	0.0027	0.0037
>=4 fq job	0.0000	0.0000
RSquare	0.34	0.30
NOB	819802	430982

Bibliography

Abowd, John M., Robert H. Creecy and Francis Kramarz. 2002. Computing person and firm effects using linked longitudinal employer-employee data. LEHD Technical Paper 2002-6. (April).

Abowd, John M., John Haltiwanger, Julia Lane and Kristen Sandusky. 2001. Within and between changes in human capital, technology, and productivity. Mimeo. (October).

Abowd, John M., Francis Kramarz, and David Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67(2), 251-334.

Abowd, John M., Paul Lengermann, and Kevin McKinney. 2003. The measurement of human capital in the U.S. economy. LEHD Technical Paper 2002-9. (August).

Acemoglu, Daron. 1996. A microfoundation for social increasing returns in human capital accumulation. *Quarterly Journal of Economics* August, 111:779-804.

—. 1998. Why do new technologies complement skills? Directed technical change and wage inequality. *Quarterly Journal of Economics*, 113:1055-1089.

—. 1999. Changes in unemployment and wage inequality: An alternative theory and some evidence. *American Economic Review*, 89:1259-1278.

Acemoglu, Daron, Joshua Angrist. 2002. How large are human capital externalities? Evidence from compulsory schooling laws. *NBER Macro annual 2000*. pp9-59.

Albrecht, James and Susan Vroman. 2002. A matching model with endogenous skill requirements. *International Economic Review,* 43(1):283-305.

Bianchi, Marco. 1997. Testing for convergence: evidence from non-parametric multimodality tests. *Journal of Applied Econometrics* 12:393-409.

Bound, John, and Harry J. Holzer. 2000. Demand shifts, population adjustments, and labor market outcomes during the 1980s. *Journal of Labor Economics* 18(1):20-54.

Bowlus, Audra, and Lars Vilhuber. 2002. Displaced workers, early leavers, and re-employment wages. LEHD Technical Paper 2002-18. (November).

Bresnahan, Timothy F., Erik Brynjolfsson and Lorin M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics,* 117(1):339-376.

Chiang, Hyowook. 2003 Learning by doing, worker turnover, and productivity dynamics. Mimeo.

Dunne, Timothy, John Haltiwanger, Lucia Foster, and Kenneth Troske. 2000. Wage and productivity dispersion in U.S. manufacturing: The role of computer investment." NBER Working Paper No. 7465. (January).

Efron, Bradley and Robert J. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman & Hall

Eudey, Gwen and Miguel Molico. 2001. Production synergies, technology adoption, unemployment, and wages. Finance and Economics Discussion Series, Federal Reserve Board. (July).

Fallick, Bruce C., Charles A. Fleischman, and James B. Rebitzer. 2003. Job-hopping in Silicon Valley: The micro-foundations of a high technology cluster. Paper presented at 2003 NBER Personnel Economics Summer Institute. (July).

Gibbons, Robert, Lawrence Katz, Thomas Lemieux, and Daniel Parent. 2002. Comparative advantage, learning, and sectoral wage determination. NBER Working Paper No. 8889. (April).

Glaeser, Edward L., Jose A. Scheinkman, and Andrei Schleifer. 1995. Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economics* 36:117-143.

Goldin, Claudia and Lawrence F. Katz. 1998. The origins of technology-skill complementarity. *Quarterly Journal of Economics,* 113(3):693-732.

Greenwood, Michael J. 1975. Research on internal migration in the United States: a survey. *Journal of Economic Literature* 13(2):397-433.

Haltiwanger, John, Julia Lane, and James Spletzer. 2000. Wages, productivity, and the dynamic interaction of businesses and workers. NBER Working Paper No. 7994. (November).

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. Earnings losses of displaced workers. *American Economic Review* 83(4):685-709.

Kennan, John and James R. Walker. 2003. The effect of expected income on individual migration decisions. NBER Working Paper No. 9585. (March).

Kiley, Michael. 1999. The supply of skilled labor and skill-biased technological progress. *Economic Journal*. 109:708-724.

Lemieux, Thomas. 1998. Estimating the effects of unions on wage inequality in a panel data model with comparative advantage and nonrandom selection. *Journal of Labor Economics* 16(2):261-291.

Lengermann, Paul and Lars Vilhuber. 2002. Abandoning the sinking ship: the composition of worker flows prior to displacement. LEHD Technical Paper 2002-11. (August).

Long, Larry E. 1988. *Migration and residential mobility in the United States*. Russell Sage Foundation.

Moretti, Enrico. 2002. Human capital spillovers in manufacturing: Evidence from plant-level production functions. NBER Working Paper No. 9316. (October).

—. 2004. Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics* 121:175-212.

Rauch, James E. 1993. Productivity gains from geographic concentration of human capital: evidence from the cities. *Journal of Urban Economics* 34:380-400.

Roback, Jennifer 1982. Wages, rents and the quality of life. *Journal of Political Economy* 90:1257-1278.

Silverman, B. W. 1981. Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society* 43(1):97-99.

Topel, Robert H. 1986. Local Labor Markets. *The Journal of Political Economy* 94(3):S111-S145.