ABSTRACT

Title of dissertation: FIRM OWNERS AND WORKERS:
AN ANALYSIS OF IMMIGRANTS AND
ETHNIC CONCENTRATION

Mónica García-Pérez, Doctor of Philosophy, 2009

Dissertation directed by: Professor John Haltiwanger
Department of Economics

This dissertation consists of three chapters examining the important role of firm and coworker characteristics, as well as the use of social networks, in labor markets. The first paper investigates the effect of firm owners and coworkers on hiring patterns and wages. Immigrant-owned firms are more likely to hire immigrant workers. This prevalence is especially strong for Hispanic and Asian workers. We also find that the probability that a new hire is a Hispanic is higher for immigrant firms. On wage differentials, the results illustrate that much of the difference between the log annual wages of immigrants and natives can be explained by immigrants' propensity to work in non-native owned firms, which pay the lowest average wages. Interestingly, though, native workers holding a job in immigrant firms are paid less than immigrant workers. The last section examines the potential mechanisms for these findings. It explores the importance of job referral and use of networks for migrants in labor markets. We consider the theoretical implications of social ties between owners and workers in this context. Firms decide whether to fill their vacancies by posting their offers or by using their current workers' connections.

Next, we explore the patterns of immigrant concentration relative to native workers at the establishment level in a sample of metropolitan areas. Immigrants are much more likely to have immigrant coworkers than are natives, and are particularly likely to work with others from the same country of origin, even within local markets. The concentration of immigrants is higher for recent immigrants and interestingly for older immigrants. We find large differences associated with establishment size that cannot be explained solely by statistical aggregation. Exploring the mechanisms that underlie these patterns, we find that proxies for the role of social networks, as well as the importance of language skills in the production process, are important correlates of immigrant concentration in the workplace.

FIRM OWNERS AND WORKERS: AN ANALYSIS OF IMMIGRANTS AND ETHNIC CONCENTRATION

by

Mónica Isabel García-Pérez

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
2009

Advisory Committee:
Professor John Haltiwanger, Chair/Advisor
Professor John Shea
Professor Seth Sanders
Professor Judith Hellerstein
Professor Howard Leathers

Dedication

To Chavela and Nene ...

This thesis is dedicated to my wonderful parents, Isabel y Luis, who have raised me to be the person I am today and sacrificed a lot to offer me the means to reach my dreams. You have been with me every step of the way, through good times and bad. Thank you for all the unconditional love, guidance, and support that you have always given me, helping me to succeed and instilling in me the confidence that I am capable of doing anything I put my mind to. ¡Gracias, los amo!

Acknowledgments

It is a pleasure to thank to many people who made this thesis possible. I owe my gratitude to all of them and because of whom my graduate experience has been one that I will cherish forever.

I cannot overemphasize my gratitude and love to my husband Darin, who did more than his share around the house as I sat at the computer. Without his support, and gentle prodding, I would still be trying to write the introduction of my first paper. With his enthusiasm, his patience, his revisions and his inspiration, he helped me to overcome the hardest moments of the creative process. I would have been lost without him. Thank you for your love amor.

I would like to express my deep and sincere gratitude to my advisor, Prof. John Haltiwanger for his valuable advice and great encouragement as well as for his excellent guidance and assistance for this research. His wide knowledge and his logical way of thinking have been of great value for me. I'd like to thank him for giving me an invaluable opportunity to work on challenging and extremely interesting projects over the past three years. It has been a pleasure to work with and learn from such an extraordinary individual.

I am deeply grateful to Prof. John Shea for offering advises in difficult moments and always have the door open when I needed him. I am also indebted for the amount of time and effort he has spent reading and correcting my work. I also wish to express my warm thanks to Professor Seth Sanders who inspired me with his questions and comments. His discussions around my work and initial explorations have been very helpful for this study.

My most sincere gratitude to Kristin Sandusky for being next to me and kindly grants me her time for answering my unintelligent questions during the time of this research. I am much indebted to Kristin for her valuable insights on the data details and her emails with great tips.

I also thank to Dr. Jose Tessada, Dr. Jeanne LaFortune, Prof. Judy Hellerstein, and the participants in brownbag seminars in the University of Maryland.

I am also grateful to my many student and work colleagues for providing a stimulating and fun environment in which to learn and grow. To Helena Schweiger for sharing great times at the beginning of my career and offering me a lot of support.

I would also like to acknowledge help and support from some of the staff members, especially to Vickie Fletcher, Elizabeth Martinez, and Terry Davis, who provided me with all the help and advice in all the administrative process during the career. I am thankful to Dorothea Brosious for providing the LaTeX thesis template.

Finally, I would like to extent my gratitude to Jeremy Wu, Fredrik Andersson and all the LEHD staff at the Census Bureau for their kind support.

I owe my deepest thanks to my family. Words cannot express my gratitude.

Disclaimer

This work is unofficial and thus has not undergone the review accorded to official Census Bureau publications. The views expressed in the paper are those of the authors and not necessarily those of the U.S. Census Bureau or the U.S. Department of the Treasury.

All papers are screened to ensure that they do not disclose confidential information. Persons who wish to obtain a copy of the paper, submit comments about the paper, or obtain general information about the series should contact Sang V. Nguyen, Editor, Discussion Papers, Center for Economic Studies, Bureau of the Census, 4600 Silver Hill Road, 2K132F, Washington, DC 20233, (301-763-1882) or internet address sang.v.nguyen@census.gov.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CBO	Characteristics of Business Owners
SSEL	Standard Statistical Establishment List
BR	Business Register
LEHD	Longitudinal Employer-Household Dynamics
MSA	Metropolitan Statistical Area

Chapter 1

Introduction

Over the last several decades, labor markets in many cities in the US have absorbed large inflows of new immigrants. During the same period, numerous empirical studies have analyzed the effect of immigration in the host economy. In the early 90s the consensus was that there is only a small effect of immigration on native economic outcomes (Grossman [1982]). However, since the late 90s, the consensus moved toward a significant effect of foreign-born migration on natives (Borjas [1994]). Recent surveys on the economics of immigration [Borjas, 2003, 2005], Friedberg and Hunt,1995; Card, 2001; Card and Lewis,2005; Card, 2006) conclude that the impact of immigration on the wages and employment is still unclear.

As of 2007, immigrant workers represented 15% of the U.S. population. The impact of large inflows of immigrants and their assimilation into the host economy has been a primary objective of analysis in the labor literature. How such large flows of workers are incorporated into the labor market and interact with various businesses and workers is of special interest. An alternative literature has focused on how firms respond to an inflow of immigrants. The key question is no longer one of job supply but also one of job distribution. Lewis and Card [2005] and Beaudry et al. [2006] look at an exogenous local unskilled labor supply change and find that areas with higher concentration of immigrants have employed higher number of unskilled workers and increased productivity at the same time. Their findings also suggest a small impact of immigration on natives' relative wages.

In this analysis, the role of business owners in the patterns of hires and earnings in the labor market is relevant. In particular, some studies have found that the type of manager recruiting new workers is a determinant in the workforce composition of the business. In a extensive analysis of race and ethnic segregation across workplaces in the U.S., Hellerstein and Neumark [2007] find that a large degree of segregation remains even after accounting for metropolitan area, education and occupation. In a follow-up paper, they explore the role of residential networks in these patterns, and found preliminary evidence of its relevance for low-educated and low-English-ability workers.

On the other hand, many other authors have analyzed the direct effect of the type of manager on the type of worker in the firm. For instance, Carrington and Troske [1995] and Giuliano and Ransom [2008] have found that females and blacks are disproportionately employed by female and black supervisors respectively. Meanwhile, Stoll et al. [2004] found that black businesses receive more applications from black workers and employ more black workers than other businesses. Giuliano and Ransom [2008] found a causal relation between the race of managers and workers using panel data of a retail store. They control for the unobserved characteristics that can also affect the race of the coworkers and hires in a firm. Although the primary determinants of the racial composition of new hires are workplace and location characteristics, manager race also stands as a significant component. Nevertheless, this second group of analyses have been mainly focused on black versus white issues and particular industries.

In the sociology literature, there have been a limited number of studies that provide some insights on the tendency of immigrants to work for immigrant firms. For instance, in Los Angeles in 1989, 30 percent of employed Koreans held jobs in firms owned by fellow Koreans even though Koreans composed only one percent of the Los Angeles County population.[1] According to Cardenas and Hansen [1988], during the 1980s, Mexican immigrant employers were most likely to hire Mexican, whether legal or undocumented, and were more likely to evaluate their quality favorably. Porter and Wilson [1980] find two relevant patterns in the Cuban immigration to Miami during the 1960s. First, Cubans worked with other Cubans. Second, almost one-third of the Cubans worked for Cuban employers. The phenomenon of immigrants hiring immigrants is not limited to coethnic relationships between employees and employers. Other researchers have found that employers from one immigrant group often hire workers from other ethnic/racial groups.[2] Immigrant entrepreneurs can take advantage of their language, cultural background and affinities to have access to different ethnic groups. Their immigrant status can give them privileged access to sources of labor less available to native entrepreneurs. Immigrant entrepreneurs routinely employ coethnics (including relatives) at rates vastly above chance levels.[3]

Making use of unique longitudinal and cross-sectional micro level databases, this thesis examines the role of owners, coworkers and networks, focusing on the

[1] Min [1989].
[2] Light [2006].
[3] Massey [1999], Massey et al. [1987].

importance of immigration and race/ethnicity on hiring patters, the scope for segregation and wage differentials. The main contributions of the research presented in this document are providing new stylized facts on the immigration issue and evidence on the role of social networks in labor markets[4].

The outline of the thesis is as follows. Chapter 2 analyzes the effect of the birthplace of firm owners and coworkers on hiring patterns and wages. Using a unique matched sample from an employer-employee administrative database and a survey of characteristics of firm owners, this chapter studies the impact of the type of employers and individual coworkers (native versus immigrant workers, and ethnic/racial groups) on firm hiring patterns and workers' average log wages. We connect owner and firm characteristics (place of birth, size and industry) with workers' characteristics (wage, age, education, and place of birth) to test different assumptions about firm hiring patterns and the wage differentials of workers of different types. Given the unique features of the matched database, the data allows asking whether the odds that a worker of a particular group is hired are related to the types of owners and coworkers, and whether there exist wage premia associated with being an immigrant and working for or with other immigrants.

Our results suggest that immigrant owners are three percentage points more likely to hire other immigrants than native owners, even after controlling for industry, firm size, geographic concentration of immigrants in the population, population density, and the legal form of organization of the firm. Looking at ethnic/race groups, immigrant owners are 3 to 4 percentage points more likely than native own-

[4]For an extensive analysis on job information networks see Ioannides and Loury [2004].

ers to hire Asians and Hispanics versus blacks. Both types of owners, immigrants and natives, hire white non-Hispanic workers, but native owners have a higher probability of having white workers as new hires. These results are based on linear probability models as well as multinomial logit specification that accounts for the simultaneity of choosing from among different types of workers. Among our strongest findings are the existence of a persistent pattern of hiring similar types and the effect of the share of dissimilar coworkers on the likelihood of hiring a particular individual. For instance, the increase of the share of similar coworkers at the time of recruitment by 100 workers increases the probability of hiring a worker of a type by around 60%. The probability is smaller if we look at the effect of the fraction of coworkers of other different types. Additionally, this probability depends on whether the employer is immigrant versus native. Immigrant businesses show higher chances of hiring a new immigrant, Hispanic or Asian worker compared to native businesses, even after controlling for whether they have similar workforce distribution at the time of a new recruitment.

Later, after controlling for owner's race, our results are similar. Hispanic and Asian owners are 2.5 percentage points more likely to hire their own type (Hispanic and Asian workers respectively) than white and black owners. Given the lack of representation of native Hispanic and Asian owners in the data, we were not able to control for the cross categories race-birthplace. To the best of our knowledge, no previous study has analyzed the link between employer and coworkers' birthplaces and employees' employment opportunities and wages. This research provides initial steps on that branch of analysis.

Chapter 3 presents descriptive evidence on immigrant segregation at the workplace and analyzes the mechanisms that drive immigrant concentration. We have unique matched employer-employee data for a large number of states in the US that permits quantifying the extent of and covariates of the workplace concentration of immigrants. A lack of suitable data has limited economists' ability to address these questions. The paper has two broad objectives. The first is primarily descriptive. The descriptive findings show that immigrants are much more likely to have immigrant coworkers than are natives. This pattern is driven partly by the geographic concentration of immigrants, but the patterns hold true even within local labor markets. At the same time, most immigrants do have native coworkers; only a small share work in immigrant-only workplaces. The concentration of immigrants is higher for recent immigrants and, conditional on recent arrival, for older immigrants. Part of the assimilation process is a movement towards more interaction with natives in the workplace over time, and younger immigrants are more likely to work with natives. We find large differences associated with firm size: concentration is much higher in smaller firms, but is far from zero even in the largest firms. We also find substantial variation in the extent of immigrant concentration across industries even after controlling for a detailed set of location, employer and employee characteristics.

Second, our finding that the allocation of immigrants across workplaces is far from random raises the question of what drives this workplace concentration. Both the existing literature and our descriptive findings suggest that it is important to consider how businesses hire their employees and the choices that businesses make about the skill mix of their workforce. One relevant issue here is the role that language skills play in governing interactions among employees and between employees and customers. A second issue is the role of social networks in the process that matches workers and firms. A third issue is human capital - the sorting and concentration of immigrants in the workplace may reflect sorting by skills. In the second part of the paper, we explore the role of these factors. We find evidence that immigrants with primarily immigrant coworkers are likely to have coworkers who live in the same residential tract. This pattern is robust to inclusion of controls for other closely related factors such as residential segregation. We also find evidence that immigrant workers with poor English speaking ability and low education are more likely to work with immigrant coworkers.

Our findings suggest that social connections and social capital may be important for understanding workplace concentration, employment opportunities and wage differentials. Continuing this line of thought, Chapter 5 offers the conclusions and discusses on the main factors that can explain the previous findings. It is intended to focus on the role on networks in the labor markets, and the connection of our findings with previous empirical and theoretical literature. It also describes the key issues to be considered to develop in appropriate theory.

Chapter 2

Does It Matter Who I Work For And Who I Work With? The Impact Of Owners And Coworkers On Wages And Hiring

2.1 Introduction

This paper analyzes the effect of the birthplace of firm owners and coworkers on hiring patterns and wages. As of 2007, immigrant workers represented 15% of the U.S. population. The impact of large inflows of immigrants and their assimilation into the host economy has been a primary area of study in the labor literature. How such large flows of workers are incorporated into the labor market and interact with various businesses and workers is of special interest. The role of business owners in the patterns of hires and earnings in the labor market has played an important role in this literature. In particular, some studies have found that the type of manager recruiting new workers is a determinant of the firm's workforce composition. For instance, Carrington and Troske [1995] and Giuliano and Ransom [2008] have found that females and blacks are disproportionately employed by female and black supervisors respectively. Meanwhile, Stoll et al. [2004] found that black businesses receive more applications from black workers and employ more black workers than other businesses.

Using a unique matched sample from an employer-employee administrative database and a survey of characteristics of small-firm owners, this study analyzes the impact of the type of employers and individual coworkers (natives versus immigrants, or ethnic groups) on firm hiring patterns and workers' average log wages. Firm types are defined by the type of owner (immigrant-owned versus native-owned), while ‘’coworker” refers to the fraction of same-kind fellow workers holding a job in the same firm. The share of immigrant coworkers in the firm is called the *coworker index*.[1] We connect owner and firm characteristics (place of birth, size and industry) with workers' characteristics (wage, age, education, and place of birth) to test different assumptions about firm hiring patterns and the wage differentials of workers of different types. Given the unique features of the matched database, the data allows asking whether there exist wage premia associated with being an immigrant and with working for or with other immigrants.

The type of a new hire can be affected by the type of employer in different ways. First, social networks, segregated by race or similar background, could be used by job seekers and by employers when looking for new candidates. Ethnic communities provide a network for immigrant entrepreneurs to find workers, to sell ethnic goods, and to obtain credit. Second, matching productivity generated by employer-employee similarity could motivate owners to employ same-kind individuals. In certain industries the use of a common language may be important for productive efficiency. Third, employer tastes might bias them to employ workers

[1] In this Chapter, the expressions *firm type* and *owner type* are used to explain that firm's owners correspond to one of the following groups: native-only, immigrant-only, and mix owned firms.

of a similar kind. Employer discrimination could generate scope for segregation.[2] However, coworker effects could compensate for the presence of employer discrimination. In fact, for all types of owners the share of similar coworkers increases the probability of being hired in the firm. We also control for specific characteristics in the firm, such as the fraction of English speakers, to identify the possible scope for matching productivity. This paper focuses on the importance of social ties in the process of recruitment when firms use current employees' social connections to help find and identify new candidates. However, employers may use this mechanism differently for different worker types, depending on their ability to take advantage of their workers' connections. For instance, given their cultural, linguistic, and social backgrounds, immigrant employers have an advantage, compared to natives, in exploiting their immigrant workers' social connections.

Our results suggest that immigrant owners are three percentage points more likely than native owners to hire other immigrants, even after controlling for industry, firm size, geographic concentration of immigrants in the population, population density, and the legal form of organization of the firm. Looking at ethnic/race groups, immigrant owners (Hispanic/Asian owned firms) are 3 to 4 percentage points more likely than native owners (white and black owned firms) to hire Asians and Hispanics versus blacks and whites. Both, native and immigrant owners, hire white non-Hispanic workers, but native owners have a higher probability of having white workers as new hires. These results are based on both linear probability models and a multinomial logit specification that accounts for the simultaneity of choosing from

[2] Lang [1986]

different types of workers.

Among our strongest findings are the existence of a persistent pattern of hiring similar types and the smaller effect of the share of dissimilar coworkers on the likelihood of hiring a particular individual. For instance, the share of similar coworkers at the time of recruitment increases the probability of hiring a worker of a type by around 60%. The probability is higher when the owner is similar to the new hired. Additionally, this probability is different whether the employer is immigrant versus native. Immigrant businesses show higher chances of hiring a new immigrant, Hispanic or Asian compared to native businesses, even after looking whether they have similar workforce distribution at the time of a new recruitment.

To study the wages of employees, one must understand the role of employers in wage-setting, which necessitates gathering wage data by employer and having detailed information about the employer. Immigrant workers tend to have lower average wages than native workers. Many authors have used a human capital approach to explain that wage gap and have found that skill accounts for almost two thirds of the wage difference between Hispanics and white Non-Hispanics.[3] Meanwhile, the residual unexplained wage gap has traditionally been used to claim the existence of racial/ethnic discrimination in the labor market. Other authors have found that industry wage-differentials are to a very large extent explained by the characteristics of workers and the contribution of industry to wage setting is much smaller after looking at both person and that industry effects.[4] However, these studies don't rule

[3] Borjas [1994], Trejo [1997], Chiswick [1978], Borjas [2003] among others.

[4] Abowd et al. [1999]

out a significant impact of firm-level effects on wage formation.[5] The results in this paper suggest that much of the difference between the log annual wages of immigrants and natives comes from immigrants' propensity to work in non-native owned firms, which pay the lowest average log annual wages. Interestingly, though, native workers holding a job in immigrant firms are paid less than immigrant workers. After controlling for typical human capital variables, full-time immigrant workers earn about 8% less than native workers ($3,293 less each year). When working for native employers this difference increases to 11%. Meanwhile, immigrant workers earn 10% more than native workers in immigrant owned firms ($4,398 more each year).

Recent work has used the idea of networks in the labor market to explain labor market inequalities as a function of differential social capital (social resources, network structures, network resources). Minority individuals are generally connected to other minority-group workers who cannot provide them with the opportunity to change their employment outcomes. Hispanics and blacks are disadvantaged because they are likely to match with same-kind job contacts, and end up working in lower wage workplaces where other Hispanics and blacks work (Elliot [2001]).

To the best of our knowledge, no previous study has analyzed the link between employer and coworkers' birthplaces and employees' employment opportunities and wages in a large set of industries and geographic locations. This research provides initial steps on that branch of analysis. These findings suggest that social connections and social capital may be important for understanding employment opportunities

[5] These authors obtained that the average of the difference in wages paid to an identical worker employed at two different firms in France was 20%-30%.

and wage differentials.

The remainder of the paper is organized as follows. Section 2.2 and section 2.3 review previous work on the relation between workers and types of firms, ethnic economies and 'ethnic matching' between supervisors and employees, the usage of networks, and network effects on hiring procedures and workers' wages. It also discusses the importance of analyzing small businesses when looking at the impact of immigration. Section 2.4 examines the data and presents basic descriptive statistics on owners' and workers' characteristics. Next, section 2.5 presents preliminary information on workers' average earnings by worker type and by different levels of coworker shares. Section 2.6 is divided in two sections. The first part analyzes whether the type of employer and coworker characteristics affect the composition of new hires in firms. The second part evaluates the impact of firm owner type on employees' log annual earnings controlling for worker human capital. Section 2.7 concludes.

2.2 Literature Review

Because no single theory exists to explain the effect of firm owners and coworkers on hiring patterns and wages, we draw on the literature of several related fields to motivate our hypotheses on the subject. Those literatures include ethnic economy theories dealing with ethnic/immigrant concentration, theories of firm wage differentials and hiring procedures, and network theories.

Immigrants tend to work in low-wage/low-productivity firms, low-pay occupations, and in firms with a high percentage of immigrant workers.[6] Some researchers have found occupational and ethnic coworker concentration in the United States (Andersson et al. [2007], Patel and Vella [2007], and Light [2006]) and in other countries (Barr and Oduro [2000] and Andersson and Wadensjó [2001]). The literature has attempted to explain workers' concentration by skill, race, and sex.[7] Hellerstein and Neumark [2007] analyzed ethnic segregation in the United States and found a substantial degree of segregation in the workplace. They claim that even though workplace segregation partially results from residential segregation (*spatial mismatch* explanation) and from ethnically correlated skills, there seem to be other mechanisms that suggest the presence of *immigrant social connection* effects (local residential networking). In an extensive analysis of racial and ethnic segregation across U.S. workplaces, they found that a large degree of segregation remains even after controlling for metropolitan area characteristics, and that very little of this segregation can be explained by observed differences in education and occupations. Language, however, seems to be a significant factor for immigrant segregation. Lang [1986]'s theory provides an explanation for worker segregation by language groups. When there are transaction costs associated with employees of different language groups working together, there is scope for segregation. Employers of each language group have incentives to fully segregate to avoid the cost of needing employees who can be the bridge between different language groups.

Despite findings on immigrant concentration at different levels, we cannot be

[6] Borjas [1994], Borjas [2003], Andersson et al. [2007], and Andersson et al. [2008].
[7] Kremer and Maskin [1996], Hellerstein and Neumark [2003].

sure that immigrants are more likely to work for immigrant bosses and that such a pattern would affect individuals' labor market outcomes. There is no evidence that immigrant-owned businesses are distributed (or concentrated on) differently across specific industries, firm sizes, or skills, than native businesses, and that this distribution is correlated with the distribution of immigrant workers across industries, sizes, and skills.

A recent group of studies analyzes the matching process between managers and workers by racial group. Giuliano et al. [2006] found a significant effect of race and ethnicity on hiring procedures. For example, in locations with large Hispanic populations, Hispanic managers tend to hire more Hispanics and fewer whites than white non-Hispanic managers. In a more recent analysis, Giuliano and Ransom [2008] looks at the effect of manager ethnicity on hires, separations and promotions across different occupations in a U.S. retail firm. Whites were more likely to leave stores where managers were Hispanics than when they were white. Their work is very relevant, although they only focus on a very particular retail firm. Their studies do not consider the coworker effect. That is, they don't study the effect of the fraction of similar coworkers holding a job in the firm on the probability a particular type of worker is hired.

There has not yet been a connection established between owner's birthplace and the type of workers employed at a firm or these workers' earnings. Nevertheless, the literature discusses motivations for supervisor-employee matching. First, firm owners could have preferences for employing individuals of their own type or with the same background. Second, the types of goods offered by immigrant firms may differ from those offered by native firms. If immigrants specialize in producing ethnic goods, immigrant workers have a comparative advantage over native workers in these firms. The differences between products can result in different worker composition.[8] However, none of these reasons have obvious predictions of workers' earnings. That an employer has a preference for a certain group does not necessarily imply higher wages for that group. The distribution of workers and employers in the market also affects the labor market equilibrium.

In the sociology literature, there have been a limited number of studies that provide some insights on the tendency of immigrants to work for immigrant firms. For instance, in Los Angeles in 1989 30 percent of employed Koreans held jobs in firms owned by fellow Koreans even though Koreans composed only one percent of the Los Angeles County population.[9] According to Cardenas and Hansen [1988], during the 1980s, Mexican immigrant employers were most likely to hire Mexicans, whether legal or undocumented, and to evaluate their quality favorably. Porter and Wilson [1980] find two relevant patterns in the Cuban immigration to Miami during the 1960s. First, Cubans worked with other Cubans. Second, almost one-third of the Cubans worked for Cuban employers. The phenomenon of immigrants hiring immigrants is not limited to coethnic relationships between employees and employers. Other researchers have found that employers from one immigrant group often hire workers from other ethno/racial groups.[10] In Los Angeles, during the nineties, 51% of the garment factories were owned by Asians with most of their employees being

[8] Andersson and Wadensjó [2001]

[9] Min [1989].

[10] Massey [1999], Massey et al. [1987].

Hispanics. Ethnic networks alone cannot expand the supply of coethnic-accessible jobs. Generally, the number of jobs offererd by ethnic-specific owned firms is not equal to the number of possible candidates from the same ethnic group in the local community. Business leaders from ethnic groups whose rates of entrepreneurship are higher than other groups find it difficult to limit hiring to members of their own groups. Ethnic crossover can expand the economic opportunities provided by immigrant-owned businesses. Immigrant workers often join networks that cross ethnic boundaries. Using the Garment Industry in Los Angeles as an example, Light [2006] analyzes immigrant ownership economies consisting of immigrant employers plus their immigrant but not coethnic employees. He finds that this type of economy explains part of the garment industry's growth during early 1990s in Los Angeles.

The cited studies have been limited to small samples from particular geographic areas and specific groups of firms and immigrants. Most of them also focus on a particular period of time, with a cross-sectional view of the distribution of workers and firms. These analyses tended not to look beyond the segregation aspect to analyze the possible causes and consequences of those patterns. Unlike previous studies, this paper uses a representative group of areas, firms, industries and workers, and it analyzes the flow of hiring and the effect of employer-employee type matches on wages. The underlying hypothesis in the analysis is that workers and employers make different use of their social connections in the market, given their specific characteristics, such as race/ethnicity and immigration status, which leads to a particular hiring pattern by each firm. Immigrant firms, for instance, would have an advantage over native firms when using their immigrant current workers as a channel to find new workers.

On wage effects, previous research has suggested that much of the unexplained variation in wages among employees is linked to characteristics of their firms, such as size and industry.[11] Not only do individual characteristics explain wage differentials between immigrants and natives, but potentially so do other characteristics, such as the birthplace or ethnicity of employers and coworkers. Unfortunately, most wage databases come from household surveys of individuals (Decennial Census and CPS), rather than from establishment surveys of wage-paying employers; they provide little employer-specific information, except for industry and, in some cases, firm size.

2.3 On the use of social networks

Recent work has suggested that supervisor-employee ethnic matching could result from the use of networks.[12] On the one hand, according to several sociological studies on the ethnic economy, ethnic solidarity serves to provide entrepreneurs with privileged access to immigrant labor and to legitimize paternalistic work arrangements (Sanders and Nee [1987] and Model [1997]). Different firms have different recruitment processes, generating an initial sorting of worker types. On the other hand, networks can also have an impact on wages, providing better matches and more opportunities to the individual. Ethnic networks can generate informal sources

[11] [Groshen, 1990, 1991a,b], Abowd et al. [1999], Abowd et al. [2004] among others.

[12] Networks is not a new concept in the literature. For an extensive analysis on job information networks see Ioannides and Loury [2004]. Sociologists have investigated the origins and creation of social networks for more than 40 years. Rees[1966] draws attention to differences among workers and their use of available information (formal and informal sources). Job referral is also extensively used in the labor market, as well as family networks (Granovetter [1995]).

of capital formation and captive markets, making these firms more self-sufficient and flexible (Volery [2005]). Social capital becomes another form of capital resource.[13]

Individual's social networks are likely to have an impact on labor market outcomes (Simon and Warner [1992]). The differential use of social networks does not provide the same access to information and opportunities to all individuals, offering a better relative position to those agents with better social connections or better use of their social networks. Recent literature has moved away from spatial mismatch model in explaining inequality across ethnic/race groups towards theories that include how social networks affect urban inequality [Hellerstein and Neumark, 2007, Hellerstein et al., 2008a]. Life-chances depend not only on individual resources but also on network characteristics reflecting the resources of network members. In this context, personal networks are then considered an additional determinant of inequalities (Light [2006]).

How do these mechanisms affect our groups of analysis? What is different about particular types of workers and firms such as immigrant/racial groups? Although the comparison between whites and blacks has been long discussed, immigrant status can be crucial for understanding group differences in informal job matching and labor outcomes. Two important characteristics of the immigrant community are relevant for these implications. First, Borjas [1994] pointed out that immigrants tend to be less educated, to have poor English language skills, and to lack domestic experience. Second, immigrants rely heavily on social networks for

[13] Social capital in its simplest form is a social network of strong and weak social ties (Light and Gold [2000]).

finding jobs and geographically reallocate (Massey et al. [1987]).

Previous literature has also discussed racial and ethnic differences in informal job matching (Elliot [2001], Holzer [1987]). These differences arise because informal channels permit race and other characteristics in the network to play a more prominent role in the hiring process than it does when formal mechanisms are used. As noted by Elliot [2001], one of the puzzles during 1980s and 1990s was the worsening position of less educated blacks in the labor market while the economy was absorbing thousands of new immigrant workers. Surprisingly, these new workers had, on average, similar characteristics to blacks: low formal education and high geographic segregation. So the question of job distribution became a first order issue, especially in the topics of immigration and immigrant assimilation. Research on this puzzle has focused on the use of social networks by different groups for finding employment [Waldinger, 1997], while the role of prospective employers in the use of these mechanisms has been ignored.

Our empirical analyses sheds light on the impact of networks on immigrants. Considering the tendency of workers to refer their own, the immediate effect of network is the reproduction of the workforce composition across time as shown in this chapter. Our results in the following chapter support the hypothesis that social networks play an important role in workplace concentration. The tendency of social networks to be racially/ethnically homogeneous - exacerbated by individual's immigration status- increases the probability that workers would refer same-type candidates and that same-type employers would tend to hire from shame-type groups. Immigrant employers can take better advantage of their immigrant employees in hiring than native employers.

The differential use of job referrals by employers is also evident when we examine who is hired and how the wages are distributed in the firm. Immigrants will tend to be hired more by immigrant firms with a high share of immigrant workers than by native firms with high share of immigrant workers.

Immigrant entrepreneurs can take advantage of their language, cultural background and affinities to have access to different ethnic groups. Their immigrant status can give them privileged access to sources of labor less available to native entrepreneurs. Immigrant entrepreneurs routinely employ coethnics (including relatives) at rates vastly above chance levels. The most important network relationships are based on kinship, friendship, and paisanaje (the feeling of belonging to a common community of origin).[14] Immigrant economies rely upon networks to locate jobs. On the one hand, referrals by friends or coworkers remove some of the uncertainty associated with finding a job with unfamiliar employers and increase the chance of finding a better job match. On the other hand, immigrant entrepreneurs tend to rely on their current employees to help fill their vacancies. Workers tend to refer individuals that are 'similar' to them, from the same group, or with the same characteristics. Referral coworkers could also provide informal training, show the new worker how to perform the job, and have a good interaction with the new hire. Moreover, referral coworkers indirectly accept responsibility for new hires. Employers realize that this practice is beneficial for them as well. Little cost or effort need be expended when new workers are located through employee contacts.

[14] Massey[1980].

Previous empirical findings show that Hispanic men report more frequent use of friends and relatives for job search than non-Hispanic whites, and are also significantly more likely to have obtained their most recent job through personal contacts. Hispanics use informal contacts 32.8 percent more often than white non-Hispanics and blacks.[15] Recent Latino immigrants are more likely than blacks or Latino natives to use personal contacts to find jobs.[16] Weak English skills explain much of this difference. However, this difference comes not only from the use of job networks by workers, but also from a greater reliance on referrals in small workplaces in combination with a concentration of recent immigrants in small firms. Employers also have a role in this process given that firms' hiring procedures will affect individuals' likelihood of receiving offers from jobs heard about through friends and relatives.

2.3.1 Small firms

Our focus on small/medium firms[17] is motivated by two observations. First, in larger firms, the separation between ownership and management could detach the firm's hiring process from owner characteristics. As Haltiwanger [2006] points out, however, in small firms the decision process is likely dependent on owner ability and characteristics. When dealing with each worker, small firm owners could project their tastes and managerial abilities onto the hiring and production processes of the firm. Since it is usually the business owner who makes such choices, the identification of the person responsible for hiring decisions is easier and more relevant for small

[15] Holzer[1987b], Smith [2000].

[16] (Elliot [2001]).

[17] We consider small/medium firms those with less than 500 employees.

firms.

Second, immigrant workers are more likely than natives to work in small firms. In Chapter 3 we find that there is a significant market segmentation that appears in any detailed distribution of workers in firms. Immigrants are more likely to be employed in firms with less than 10 employees 70% of immigrants work for small firms. Meanwhile, more than 60% of native workers are employed at firms with more than 100 employees. The labor force changes generated by immigration inflows are thus borne primarily by smaller, younger firms. These firms are more sensitive to immigration shocks. If we only look at aggregate numbers (including small and big firms), immigration effects will be obscured.

2.4 Data and Measures

2.4.1 Sources

In this paper, we use three different databases to match owners' characteristics to workers' characteristics. First, we use the Characteristics of Business Owners Survey (CBO) from 1992, and then match this survey with administrative data from the IRS (Business Register) for the years 1992 to 1996. To obtain workers characteristics, we use information from the Longitudinal Household-Employer Dynamics (LEHD) database for the years 1992 to 1996. In this section, we give a brief description of each database and their limitations, and discuss how we construct relevant variables used in the regressions.

The *Characteristics of Business Owners (CBO)*, later renamed the Small Business Owner (SBO)database, is produced by the Bureau of the Census. The 1992 release of CBO was the final version of this survey, which formerly was conducted every five years. The survey for the 1992 CBO's release was conducted in 1996, along with the economic census. Therefore, the questions in the survey refer to the business' and owners' information for years 1992 and 1994. The CBO is a supplement to the Survey of Minority-Owned Business Enterprises (SMOBE) and Survey of Women-Owned Businesses (WOB). The survey universe considered was ''any business which files an IRS form 1040, Schedule C (individual proprietors or self-employed persons); form 1065 (partnership); or form 1120S(Subchapter S corporation) in 1992."[18] It considers as business owners those who filed business tax forms as owners of the firm, excluding non-S corporations, with at least 500 dollars in yearly business receipts, and with the largest employment size category equal to five hundred. Note that non-S corporations generally have investors, not decision-making owners, and thus this group is not in the CBO survey's universe. However, excluding non-S corporations often excludes the largest employers, making comparisons of small and large business owners difficult. The CBO provides details about both business owners and their businesses. The unique firm identifier is the CFN (Census File Number). At the cross-sectional level this number is unique for each firm.

According to a CBO publication cited in of the Census [1997], almost 62% of the 78,147 firms' surveys [19] and 59% of the 116,589 owners' surveys were returned.

[18] Characteristics of Business Owners 1992:CBO092-1. U.S. Bureau of the Census (September 1997) and Headd [1999].

[19] This is translated into 63% of the 41,297 employer firm surveys.

One possible reason for this low rate of reply is the difficulty of finding owners of exiting firms after 3-4 years. Almost 70% of all businesses present in 1992 were still in operation in 1996. This rate is lower for minority-owned firms (around 66%).

We use employer firms in our sample. When sampling weights are used, the survey indicates that in 1992, 20% of owners were in firms with employees. According to the minority-firm surveys, women, Asian, Pacific Islander, American Indian, black, and Hispanic owners were typically underrepresented in the larger employment size classes. Hispanic-owned firms were 3.68% of all employer firms, but just 2.04% of firms with 100 or more employees. Additionally, 90.6% of business owners were born in the United States, while 9.4% percent were foreign born. [20] The percentage of native-owned firms was higher in the case of larger firms (94.5%). In this paper we focus only on employer firms.

On average, there exists more than one owner per firm. In the CBO(1992), more than 52% of firms are employer firms, and almost 41% of this group have only one owner. Employer firms tend to have more owners than non-employer firms. Based on previous research using the CBO, [21] we consider the CBO as a sample of firms even though it is essentially a sample of firm owners. The resulting complication is that we need to make assumptions to identify the owner characteristics for multiple-owner firms. As a first attempt, we consider three types of firms: only-native-owned, only-immigrant-owned, and mix-owned. Using this classification, more than 85% of employer firms have 1 or 2 owners for all types.

[20] A foreign born is an individual that was born outside the USA. CBO has a particular question on whether the owner was born in the US or abroad.

[21] Carrington and Troske [1995].

In order to identify the characteristics of the owners of a particular firm (particularly immigration status and race), we follow the work of previous research based on the CBO (Carrington and Troske [1996]). For single-owner firms, the identification is straightforward. Meanwhile, for multi-owner firms the mode is used. The number of hours per week spent at the business was used to break ties.

This database has some limitations. First, in the 1992 survey the CBO's sample universe omits chapter C corporations. This group of corporations corresponds to bigger businesses; therefore, comparison between small and large businesses in the CBO must be done with care. Second, even though we have each firm's average payroll, we know nothing about the interfirm distribution of payroll between different types of workers. Third, this survey has zero information on human capital or occupational characteristics of workers. We try to overcome some of these limitations by merging CBO with data from Bureau of Labor statistics (UI and ES202) as described below.

The second database used in this paper is the Census Bureau's *Standard Statistical Establishment List* (SSEL) or *Business Register* (BR).[22] This data has more complete information on firms given that the source of the SSEL is at the administrative level. This database works as a register of active employer business

[22]Walker [1997] has an extensive discussion on the Census Bureau's Business Register. The initial source of information on businesses is the IRS(Parker and Spletzer [2000]). The SSEL receives three main files from IRS; the Business Master File (BMF), with information on name, addresses and legal form of organization; the Payroll Tax Return File (Form 941) containing quarterly payroll and first quarter employment (including March 12th employment); and the Annual Business Income Tax Return Files with information on receipts/revenues, industry classification. For all three sources, EIN is the primary business' id.

establishments[23] in the United States and its territories. The unit of information is an enterprise, which can be associated with one or more establishments and with one or more EIN entities (Employer Identification Number).[24] In this paper we concentrate on those businesses organizations associated with only one EIN and one establishment, known as *single-establishment enterprises* or *single-unit firms*.[25] All of the small firms in this chapter correspond to single-unit establishments. The assumption that firm owners are the ones making the main contracting decisions in a firm is more plausible in firms with only one establishment than otherwise. In the case of younger and smaller firms, this restriction does not exclude many firms.[26] Additionally, businesses have a CFN (Census File Number) as an identifier, which is unique for single-unit businesses. To follow the firm across time, the longitudinal identifier for each firm is called alpha, and corresponds to the first 6 digits of firms' EIN. In the sample, we only follow firms that survived the entire period 1992 to 1996. Because most non-surviving firms did not respond to the CBO survey and the weights are constructed such that this pattern is considered, the weighted results are not impacted by this exclusion.[27]

We take data on industry, legal form of organization and employment from the SSEL files. See Appendix B for specific description of these variables. Because of the time difference between the year of information and the year in which the CBO

[23] Active employer business establishments are those with payroll at anytime during the past three years, or with an indication that the business expects to hire employees in the future.

[24] An EIN entity is an administrative unit assigned by IRS for tax purpose. Under the Federal Insurance Contributions Act (FICA) every organization with paid employees has to obtain an EIN.

[25] All the matches between CBO(1992) and SSEL(1992) are in this category.

[26] Haltiwanger et al. [2005].

[27] Headd [1999].

survey was conducted, information on employment and sales are from the SSEL dataset.[28] We use the common unique firm identifier (CFN) to match CBO with SSEL.[29] We then follow the firm across time until 1996.[30]

The second set of information is associated with the characteristics of workers. This information comes from the *Longitudinal Employer-Household Dynamics* database. Information on workers comes from the Unemployment Insurance wage records for a group of states[31] and the ES202 data.

Based on availability, we use data from eight states for the years 1992 to 1996. The sample includes states with high immigrant concentration and low immigrant concentration areas. These files contain person identifiers that allow researchers to track a worker's quarterly earnings within a State across years. We sum over quarters to obtain each worker's annual earnings. This database also contains firm identifiers that allow for an exact link between the UI files and other data sets. The business level identifiers in UI files are State Employer Identification Numbers (SEINs). Therefore, one can match the UI data with the ES202 data, using SEIN to get information on the EIN, and compare it with the data previously matched using CBO(1992) and Business Register. For single-unit firms, the units of observation at the firm level used for CBO, SSEL and LEHD are generally similar.

[28]The CBO is a retrospective survey. The response rate is affected by the survival rate of the firm and the extent to which owners can accurately recall past information.

[29]We use businesses' CFN, which are the Census Bureau's preferred intra-year, cross-dataset link. The CFN contains the EIN firm identifier and is unique for single-unit firms.

[30]To illustrate the groups of firms included in both databases, we include a short discussion on firms matching rate in the Appendix A.

[31]More detailed analysis on these records is presented in Abowd et al. [2006], and additional information on date of birth, place of birth, and gender are obtained for almost all workers in the sample after linking UI wage records to Census data. 98% of all private, non-agricultural employment is covered by the employer reports.

The UI wage records contain virtually all business employment for the sample states (for private non-farm firms). Earnings reports from these records are more accurate than survey-based earnings data, and one can obtain information for each worker in a specific firm (or establishment).

Using this database, we follow firms across time from 1992 to 1996 using the unique identifier within the state. We end up using only those firms that survived the entire period and did not change ownership. This group represents about 67% of the initial set of firms in 1992.[32] Finally, the data set used in this study is unique in the sense that it contains data from each firm on output and inputs used in the production process, as well as data on earnings and some demographic characteristics of each worker in the firm. We use the years 1992 to 1996 for the analysis mainly because information about owners' place of birth (i.e. being born in or outside the US) is only available in the Characteristics of Business Owners Survey in 1992. Our data tracks the total payroll and workforce composition of each firm from 1992 to 1996.

The drawback of using UI data is its lack of certain demographic information on workers, such as education and occupation. However, the staff at the LEHD has overcome this limitation by imputing education using administrative data from the Census Bureau containing information such as date of birth, place of birth, geographic area, industry, and sex. In this chapter, we use this imputed information on education,[33] which has been used in previous work on the LEHD. This variable is

[32] Few firms were dropped because, initially, the survey's rate of response was highly correlated with the firms survival rate, so that most of the firms with information in the survey are surviving businesses.

[33] See Lengermann et al. [2004] for details on the imputation.

a proxy for individuals' human capital. We are aware that the lack of occupational information could be a relevant drawback of the data given that prior research has documented an important role for occupational segregation in creating different workers' wage gaps. We might think that immigrants tend to concentrate in low-skilled occupations relative to natives. However, as Troske [1999] and Carrington and Troske [1995] point out, occupations and job titles are less likely to be sharply defined in small firms, and as a result there could be less occupational segregation in small firms compared to large firms. Despite this limitation, we have to keep in mind that we can account for other workers' characteristics, such as age, sex and imputed education. Given that workers have varying preferences for place of work depending on the disutility of commuting and amenities of particular areas, the areas where they would be willing to work are better represented by their actual place of work than their place of residence. Therefore, we need data on individuals' place of work. Location of the firm is obtained using the LEHD database.

2.4.2 Construction of ex post weights

A relevant technical issue that arises in the process of using different databases, especially when a survey is included, is the change of sample frame used by the survey database. Additionally, for smaller geographic areas, differences in industry and geographic information along with differences in the scope of industries covered lead to dissimilarities between the universe considered by the LEHD data and surveys based on the Economic Census.[34]

In the design of the CBO survey, four panels were created in addition to divisions by employer status (employer versus non-employer), 2-digit industry and state. These panels consider racial categories using the owners' social security information and the categories: Asian, Asian-American / Pacific Islander, Hispanic, Black, and White. These groups were created by the Survey on Minority Businesses. Therefore, small firms and minority-owned firms are over-represented in the survey.

The difference between the universe and sampling frames used in the CBO survey implies that our matched analysis sample will not be representative. Specifically, the sample frame used in the CBO will over-represent small, minority-owned businesses when linked with the UI database. To deal with this issue, we follow Abowd et al. [2007] and build *ex post* weights that control for the firms' size, 2-digit industry code, legal form of organization, and employer status. We follow previous research in that we first construct the fractions of firms each the category in the universe of ES-202. The universe of ES-202 is single-unit firms with more than one employee (coworker shares can be computed only for these firms), not in Agriculture, Mining, nor Public Administration, and less than one thousand employees, and are in Economic Census in-scope industries in 1992. This represents the numerator in the ex post weight. Then, we compute the same fractions for the final matched data and use each fraction as the denominator of the ex post weight. This weight has the property that the distribution of employment by each category reflects the size

[34]LEHD database covers partially agriculture and public administration industries. Surveys based on the Economic Census tend to over-represent businesses in areas with high density population.

distribution of the ES-202 considered universe.

The second section of the adjustment procedure involves the construction of an inverse Mills ratio. We use a probit estimation that considers the probability of being matched as a function of log employment, legal form of organization, owner's place of birth (in or out the US), and log of sales per employee to generate the propensity scores. This section intends to account for the CBO survey's sampling frame and the possible selection bias generated by the effect of unobservables on firms exiting from the universe considered to design the sample of the CBO survey. The *ex post* weights are included in all regressions. For more details and unweighted summary statistics see appendix D.

Before using our approach the matched sample under-represent small, minority-owned businesses (see appendix D.1). After the match, and without considering the re-weighting process, we would be under-representing minority groups in small size firms. The sample of firms offering unemployment benefits are relatively of bigger size. After applying our new weight, we try to recover some of the original distribution in the CBO sample. There is a lower representation of sole proprietorship after matching the original sample with the UI database without using the new weights.

2.4.3 Firms

To compare the full CBO sample to the final matched sample used in the analysis, we look at descriptive statistics for a set of variables. The final match uses LEHD information from 8 states,[35] which include high and low immigration

[35]Those states with available data in 1992 are included.

states. For these states we obtain workers' and firms' information. Firms from the agriculture, mining and public administration sectors are not included. Additionally, only single-unit businesses are considered. The original matched sample in the analysis has 7,200 firms, representing 339,040 workers from 1992 to 1996. All results are weighted by the adjusted-weight discussed in section 2.4.2.

Table 2.1 shows two blocks of summary statistics. One block (CBO-SSEL) contains the employer firms matched from the CBO survey and the BR, while the second block (Sample(CBO-LEHD)) contains the final matched sample, consisting of the subset of CBO-SSEL data matched to the LEHD. For each block, this table presents the distribution of firm type across firm size categories and sectors, together with the average number of owners, average share of immigrant workers, de-meaned average log sales per employee, average percentage of immigrants in the county in which the firm is located and in the counties surrounding this location, and the percentage of each type of owner. Total population and the share of immigrant workers are constructed from the public 1990 Census, and are based on all Census counties surrounding the location of the firm. Immigrant firms have a higher proportion of immigrants in the local population than native and mixed firms. Because immigrants also tend to be geographically segregated, we will use this variable to control for differences in firms' local workforce.

In the final matched sample, the average immigrant-owned firm employs 38% immigrant workers. The distribution of firms across sectors and sizes for each type of firm by owner birthplace is very similar, except for the tendency of immigrant-owned firms to be in retail or services, and this distribution is only slightly changed after matching the original database with the LEHD database.

From the table we observe that immigrant-owned firms' log sales per employee is slightly higher than native-owned firms. Actually, on average, native owned firms have the lowest log labor productivity. In general, firms are concentrated in size categories with fewer than 50 employees. Meanwhile, regardless their owner type, firms are highly concentrated in the sectors Services, Retail, Manufacturing and Construction. Sole proprietorships represent more than 50% of immigrant and native firms. Mixed-owned firms tend to be larger in size with respect to the other groups. These firms are mainly Partnerships and Corporations.

Table 2.1: Descriptive Statistics - CBO(1992) and Sample/Matched Firms

	CBO[1]					**Matched Sample(CBO-LEHD)**				
Distribution/Type of firm	**Mix**	**Imm**	**Nat**	**Unk**	**ALL**	**Mix**	**Imm**	**Nat**	**Unk**	**ALL**
Size (%)										
2-4	26.52	49.30	42.51	46.43	44.66	16.67	37.27	33.60	33.29	33.79
5-9	18.06	20.90	21.35	20.91	21.05	18.75	21.53	20.97	20.33	20.79
10-19	23.48	14.68	16.69	15.00	15.90	18.75	18.44	18.02	17.40	17.88
20-49	18.18	10.70	12.10	10.95	11.59	28.13	14.25	16.72	16.89	16.58
50-99	6.94	2.97	4.56	4.26	4.26	9.90	5.67	6.65	6.78	6.59
100+	6.82	1.46	2.79	2.45	2.54	7.81	2.84	4.03	5.31	4.37
Sector (%)										
Construction	6.49	5.07	12.74	10.41	10.58	5.73	4.71	13.49	9.38	10.06
Manufacturing	20.26	10.35	13.89	13.59	13.38	25.00	15.93	17.65	18.36	17.76
Transp. & Utility	5.45	2.83	7.43	6.73	6.43	5.21	2.77	7.58	6.91	6.34
FIRE	19.22	19.73	17.14	19.24	18.36	18.75	22.18	19.03	22.24	20.84
Retail	17.14	29.68	19.73	23.17	22.47	14.58	29.85	16.42	21.34	20.83
Wholesale	6.10	3.19	6.18	5.21	5.36	7.81	3.55	5.30	4.44	4.70
Services	25.32	29.16	22.89	21.65	23.43	22.92	21.02	20.54	17.33	19.49
Legal Form (%)										
Sole Proprietorship	-	52.40	51.43	29.15	50.48	-	49.12	56.81	37.18	50.64
Partnership	28.51	12.58	12.26	18.69	12.97	25.00	12.26	10.46	13.77	12.99
Corporation [2]	71.94	35.10	25.01	52.16	37.36	75.94	38.61	32.72	49.35	36.37
l(sales/employment) [3]	11.64	11.60	11.48	11.54	11.54	11.73	11.59	11.57	11.63	11.64
	(1.17)	*(1.17)*	*(1.05)*	*(1.15)*	*(1.11)*	*(1.08)*	*(1.19)*	*(1.05)*	*(1.18)*	*(1.13)*
Imm. in the neighborhood[4]	13.47	22.15	12.43	15.06	12.90	14.03	21.17	11.34	15.07	12.80
In MSA	92.01	96.50	90.07	80.60	92.41	91.30	97.37	93.01	81.71	92.85
Average Number of Owners	3.58	1.57	1.88	1.80	1.84	3.78	1.55	1.95	1.83	1.87

Continued on next page.

Table 2.1: Descriptive Statistics - CBO(1992) and Sample/Matched Firms (continued)

	CBO[1]					Matched Sample(CBO-LEHD)				
Distribution/Type of firm	**Mix**	**Imm**	**Nat**	**Unk**	**ALL**	**Mix**	**Imm**	**Nat**	**Unk**	**ALL**
Average Share of imm. Workers	.	.	.	.	.	33.00	38.05	11.55	28.00	26.00
Unweighted dist. of firms	2.03	14.94	45.87	37.16	100.00	2.84	18.10	42.54	36.53	100.00
Weighted dist. of firms	2.01	13.50	78.60	5.89	100.00	2.84	12.89	77.92	6.40	100.00
RACE/ETHNICITY										
Hispanic	10.39	18.56	2.99	4.35	4.69	10.09	19.20	3.41	4.15	5.66
Asian	10.60	35.68	1.34	5.20	5.07	15.50	51.70	2.55	3.02	9.41
Black	0.78	1.56	2.22	2.60	2.16	0.79	0.69	1.48	1.20	1.36
White	78.23	44.20	93.45	87.85	88.08	73.62	28.41	92.56	82.30	83.57
# of Firms unweighted			41,297					7,985		
# of Observations unweighted			1,655,750					339,040		

Note: Statistics based on weighted outcomes unless the contrary is indicated.
(1) Single-unit firms that matched with SSEL.
(2)Only S- Corporation .
(3)Source SSEL: Sales (total receipts/sales), and employment (Employment March12th). Numbers in parenthesis are standard deviations.
(4)Using Census 1990, computed percentage of immigrant population in the counties including the firm and surrounding.

The average number of owners (owner type) is similar in the original and matched samples. The average number of owners by owner birthplace is similar, except, as expected, for mix-owned firms which by definition have two or more owners. Table 2.1 illustrates that these patterns are similar in the original CBO sample and the final matched CBO-LEHD sample. In the matched sample, Asian-owned firms are over-represented, while white immigrant owners are underrepresented. However, as in the original CBO sample, immigrant-owned firms are mainly owned by Hispanics and Asians, while most of the native-owned firms have white owners.

In the original matched data there is a percentage of firms with unknown owners' place of birth. We decide to exclude this group from further analysis. Given that, on average, the characteristics of this unknown group are similar to the rest of the sample (see Appendix(C) for t-tests and a chi-square analysis), we don't expect this exclusion to affect our findings.

We drop firms with less than two employees. Given that female labor participation is characterize for additional elements different to the ones analyze here we only consider male workers.[36] Workers should have at least one coworker, and the analysis of earnings is net of other labor supply factors that could affect female workers differently. After these restrictions, the final sample is reduced to *4,478 firms* and *214,398 workers* from 1992 to 1996.

[36]The effect of networks for female immigrants is also a very important analysis. According to Massey et al. [1987], Mexican female immigrants tended to arrive and go directly to specific industries such as babysitter and meat packaging. The variation at such detailed level is not enough in our data, so we cannot disentangle industry effect versus owner effect. Given the particularity in the way female labor enter the market, there could be additional unobserved elements affecting the likelihood of hire an immigrant woman that we cannot consider in this aggregate analysis.

2.4.4 Workers

Among the relevant workers' characteristics available in our data are age, immigration status (place of birth), date of entry in the US (date of SSN application), education, quarterly earnings, and race. We sum over quarters to obtain each worker's annual earnings, and then compute real earnings based on 1992 dollars. The data set used for the analysis includes all male workers with positive earnings.

On the distribution of workers, Table (2.2) and Figure (2.1) show the proportions of workers by age, race, sex, education, owner type,size, and sector, as well as, mean age, education and earnings, for all workers and for immigrants and natives. Foreign workers represent almost 24% of the sample.

Similar to previous studies, on average, foreign born workers tend to be less educated, younger and tend to have lower income than native workers (Borjas [1994]), although these differences are not large in our sample. The fraction of workers across age categories, however, is similar for both types of workers in age categories 40 years and more.

Table 2.2: Descriptive Statistics - Characteristics of Workers

	Individual		
	IM	**US**	**ALL**
MEAN (std)			
Age	34.01	34.14	34.11
	(13.33)	*(12.02)*	*(13.13)*
Education	13.04	13.16	13.13
	(2.76)	*(2.94)*	*(2.79)*
Log(annual earnings)	8.30	8.32	8.33
	(1.87)	*(1.68)*	*(1.84)*
DISTRIBUTION (%)			
AGE			

Continued on next page.

Table 2.2: Descriptive Statistics - Characteristics of Workers (continued)

	Individual		
	IM	**US**	**ALL**
Under 25	18.09	24.43	22.91
25-39	51.64	43.03	45.10
40+	30.26	32.54	31.99
EDUCATION			
High School Dropout	8.89	7.41	7.77
High School Graduate	59.27	59.26	59.26
Some College Education	30.81	32.00	31.71
College Graduate	1.04	1.33	1.26
SECTOR			
Construction	7.92	18.37	15.85
Manufacturing	37.08	26.13	28.76
Transportation and Utilities	3.81	7.11	6.31
Wholesale	14.17	14.16	14.16
Retail	19.33	16.63	17.28
FIRE	1.36	1.58	1.52
Services	16.34	16.03	16.10
SIZE			
2-4	2.18	1.65	1.78
5-9	4.70	4.23	4.34
10-19	9.27	9.32	9.31
20-49	19.32	21.23	20.77
50-99	18.96	18.91	18.92
100+	45.57	44.66	44.88
RACE			
White	17.36	75.07	61.18
Hispanic	47.21	4.85	15.04
Asian	22.76	0.93	6.18
Black	1.71	11.21	8.92
Other	10.88	4.67	6.16
TYPE OF OWNER			
Immigrant	42.70	12.10	19.46
Mixed	8.34	5.67	6.31
Native	48.96	82.23	74.22
RACE OF OWNER			
Asian	20.64	4.74	7.15
Black	0.85	0.96	0.95
Hispanic	15.58	4.94	6.24
White	64.92	89.36	85.68
Part-time	43.67	37.39	38.90

Continued on next page.

Table 2.2: Descriptive Statistics - Characteristics of Workers (continued)

	Individual		
	IM	**US**	**ALL**
In MSA	97.23	83.38	86.71
All	24.06	75.94	100

Note: Number of observations equal to 214,398 workers. Statistics based on weighted outcomes. Standard Deviations in parenthesis. Male workers with positive earnings in a year. Log annual wage in 1992 dollars.

We can compare our sample of workers with the distribution and characteristics of workers from IPUMS 1990 (see appendix E.1 we find interesting differences. To build the comparable sample, we only look at male workers, older than 16, and not working in Agriculture, Mining nor Public Administration sectors. One main difference is the average year of school between our sample and IPUMS. In IPUMS, both immigrant and natives have more year of schooling. Our sample has a very low proportion of college graduate workers (natives and immigrants). This low representation of this group could be driven by the over representation of small firms in our sample versus IPUMS database, and the types of workers that these firms hire.[37] Natives have higher wages, but the wage differential between natives and immigrants is higher in IPUMS (around 12%) than in our sample (around 1%). Interestingly, natives in our sample are younger than the national average. By race, the distribution of workers is very similar. The proportion of immigrants in our sample is around 24% versus 13% for the national average. In sum, our sample contains younger male workers with low educational attaintments.

[37]Ipums database does not include firm's size. Therefore, we cannot control for the size of the firms.

The share of workers with a high school diploma or less is over 60% for both immigrants and natives. Immigrants are more concentrated in the high school dropout and high school graduate categories. Looking at sectoral distribution, both foreign and native workers are concentrated in Construction, Manufacturing, Retail and Services, with natives more likely to be in Construction and immigrants in Manufacturing.[38] Foreigners are more likely to be working for immigrant owners than native workers. 43% of immigrant workers are employed in immigrant firms and 49% are employed in native firms. Asian and Hispanic-owned firms employ more immigrant workers than the average firm. More than forty percent of immigrant employees are hired by immigrant owners (around 43%).

Most of the immigrants are Hispanics or Asians, while natives are mainly either white or black. Although there is a fraction of native-Hispanic and native-Asian workers, these proportions are less than 5%. The racial and ethnic categories follow the SSA codes, which form a set of mutually exclusive and collectively exhaustive categories. I also include information on whether the worker is full or part time. A worker is full time if he or she has worked during the full year (worker has positive earnings all four quarters). Most of the survey corresponds to information from firms located in MSAs. However, we include a variable that identifies those firms and workers located outside a MSA. Almost 90% of the workers holds jobs in a firm located inside a MSA.

Looking at place of birth in detail, Mexican, Salvadorian, Indian, and Chinese

[38]One explanation for this pattern is that informal and undocumented immigrants workers are not largely covered by the database.

workers are the most represented immigrant groups in the data. At the national level, these are also the largest immigrant groups in the US according to Census 1990. In the data, native owners employ almost 75% of the total workforce.

2.4.5 Measuring coworker share

As described further in 3 below, we calculate the immigrant coworker share by considering all workers at the firm aside from the sample worker using the following formula:

$$COW_{ij} = \frac{1}{emp_j - 1} \sum_{emp_j}^{k \neq i} I_k \tag{2.1}$$

Where I_k is one when the worker is an immigrant. Therefore, this measure equals the fraction of immigrant coworkers of an employee in a firm. This measure is generally used in concentration analysis.[39] Here I use it as an indication of workforce composition in the firm.

2.5 Analysis of New Hires, Earnings of Workers and Skill Distribution

2.5.1 New Hires

For the analysis of hiring procedures, we look at the type, race and ethnic composition of new hires by type of owner. During the period of analysis (1992-

[39] Hellerstein and Neumark [2007]; Aslund and Skans [2005a], Aslund and Skans [2005b].

1996), there were 147,373 new hires. We identify a new hire in the data by following a firm and looking at those workers that accessed the sample during the period of analysis. We track information on each new worker. Table (2.3) shows the distribution of new hires by type of owner. While new hires include a large share of natives for every type of owner, the proportions of newly hired immigrants for immigrant and mixed-owned firms (more than 30%) are almost three times the proportion of immigrants hired in native-owned firms (almost 12%).

The second section of Table (2.3) displays the composition of new hires by race and ethnicity. Hispanics and Asians correspond to more than 35% of immigrant-owned firms' new hires. Again, this represents almost three times the proportion hired by native firms. Both immigrant and native firms hired more new workers later in the sample period as the economy recovered from the 1991-1992 recession (see Figures 2.1 to 2.4). The main diagonal shows that immigrant-owned firms hire more immigrants (33.33%) than the average firm (14.80%)), while native-owned firms hire more natives (88.44%) than the average firm (85.20%).

Table 2.3: Average Race and Ethnic Composition of New Hires by Owner's Type

	Owner Type			
Worker type/ race/ethnicity	Immigrant	Mixed	Native	All
Immigrant	**33.30**	37.10	11.56	14.80
Native	66.70	**62.90**	88.44	85.20
Hispanic	**20.14**	22.32	10.14	11.50
Asian	**16.09**	12.78	2.65	4.26
White	48.20	49.80	**73.61**	70.40
Black	6.49	8.41	**8.68**	8.46

Note: Number of Observations equal to 147,373. Male workers with positive earnings in a year. The other race/ethnic groups represent 0.5% of the sample. Results are not shown.

If we further look at the distribution of new hires across owner's race, the findings are stronger. Table 2.4 shows a strong correlation between the race of the owner and the racial/ethnic composition of the new hires.[40] Asian, Black and Hispanic-owned firms hire their own type more than twice as often as the average firm. Asian-owned firms also hire Hispanic in a large number.

Table 2.4: Average Race and Ethnic Composition of New Hires by Owner's Race

Worker / Owner	Asian	Black	Hispanic	White	All
Asian	**23.75**	2.50	4.27	2.70	4.25
Black	6.29	**38.54**	10.33	8.04	8.46
Hispanic	20.10	10.52	**35.46**	9.10	11.50
White	39.03	44.43	42.24	**75.39**	70.42

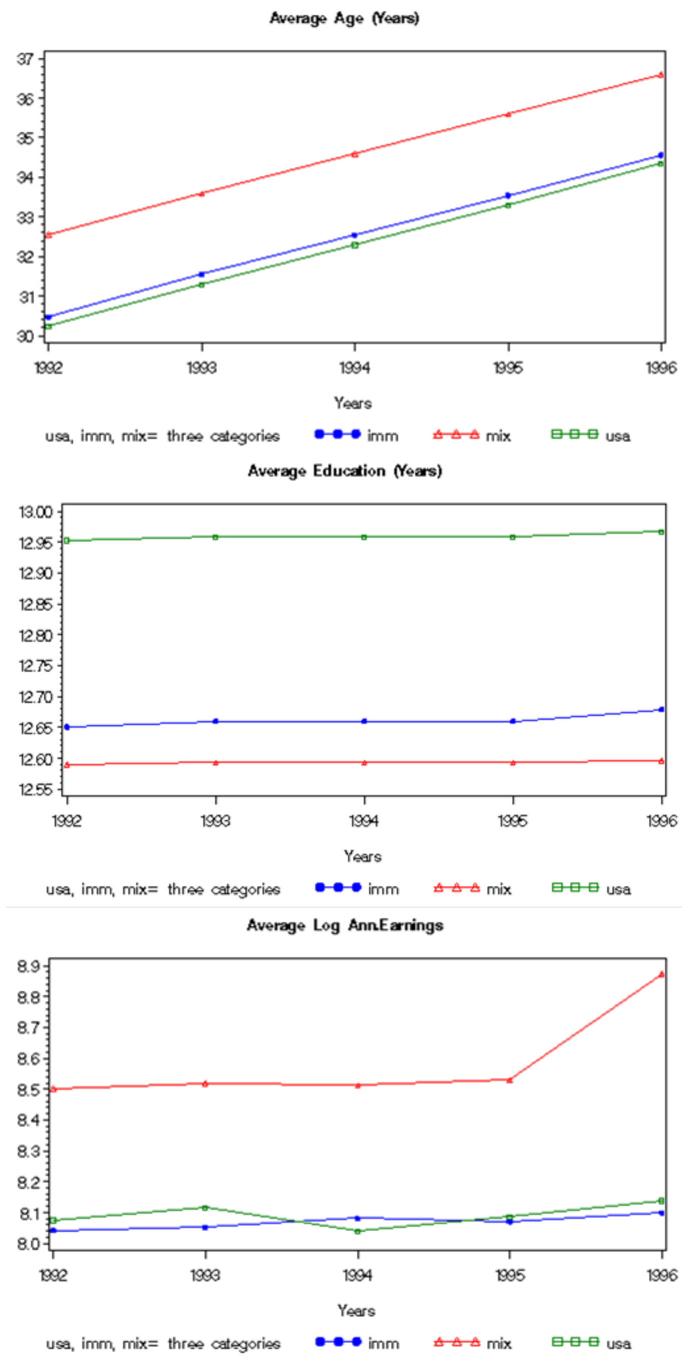
Note: Number of Observations equal to 147,373. Male workers with positive earnings in a year. The other race/ethnic groups represent 0.5% of the sample. Results are not shown.

2.5.2 Earnings of Workers

In this section, we look at workers' earnings. On average, immigrant workers have lower wages than natives. Most of the explanations given by the literature are based on human capital formation. Immigrants have lower host country abilities and generally less education than natives. However, even after controlling for some of these characteristics, immigrants tend to receive lower wages than observationally similar natives (Borjas [1994]). But do workers receive different wages than their counterfactual group regardless of who they work for? To answer this question we undertake two different exercises. First, we look at the average real log annual

[40]Giuliano et al. [2006] find a similar correlation when they look at the race of the hiring manager and the racial composition of the new hires in different establishments of a retail store.

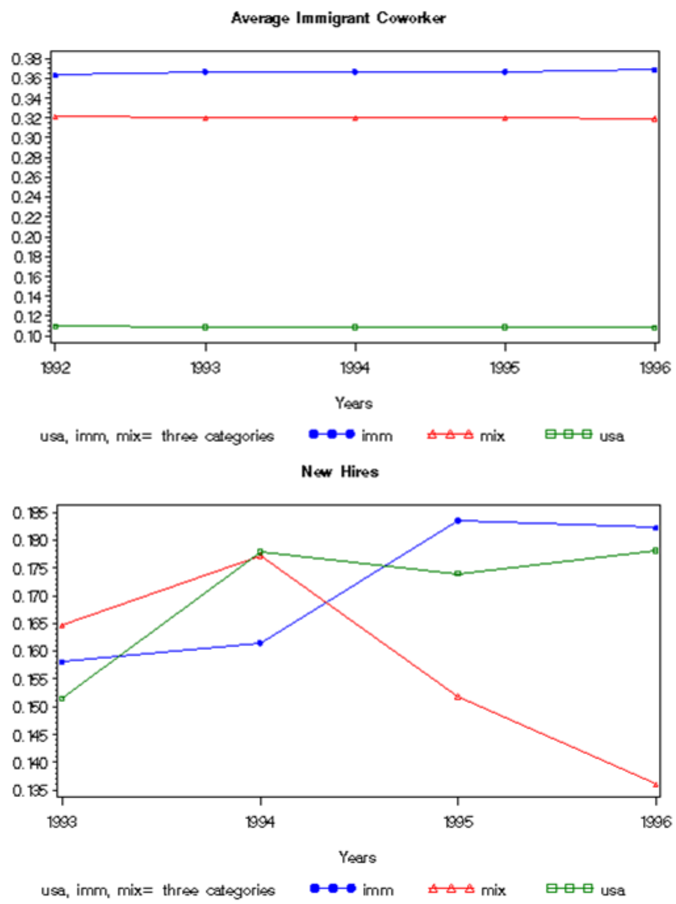
Figure 2.1: Workforce Characteristics of Immigrant, Mix and Native Firms



Note: Weighted share and percentage. Base on years 1992-1996.

earnings of each worker type across owner types. We also look at these statistics for different groups of firms defined by the fraction of similar coworkers in the firm. This

Figure 2.2: Workforce Characteristics of Immigrant, Mix and Native Firms Continuation



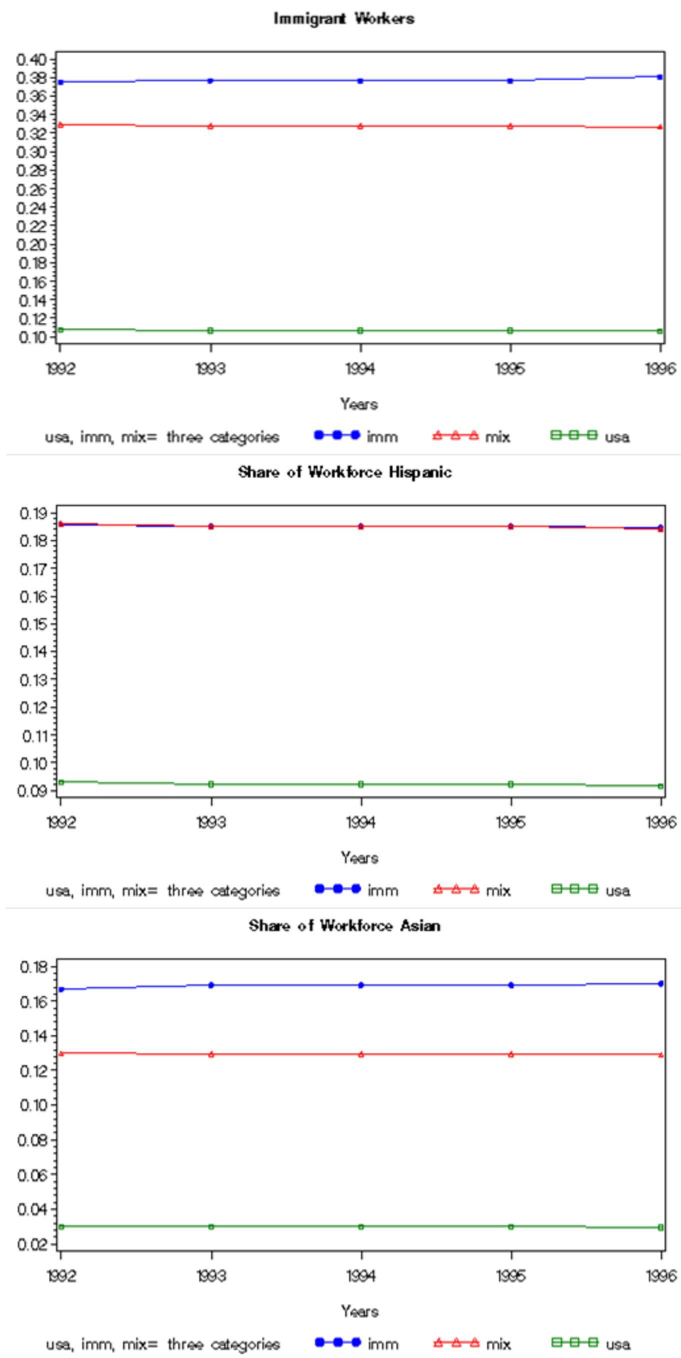
Note: Weighted share and percentage. Base on years 1992-1996.

analysis is a first look at the impact of firm owner types on earnings. Second, we estimate owner type wage effects after controlling for a number of firm and worker characteristics, and evaluate the sources of wage differentials.

The natural log of real annualized earnings of each worker comes from LEHD-UI records.[41] Table (2.5) shows how average wages change according to the type

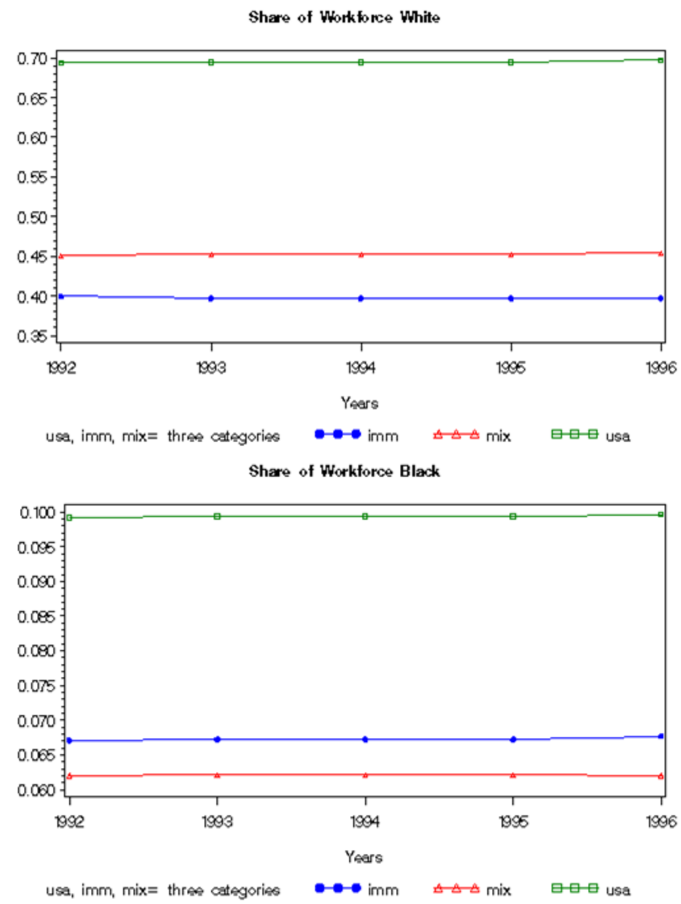
[41] When we take the average log annual earnings for each type of firm, we find that it is slightly below the log of annual payroll per employee in the SSEL database. According to internal documentation on the ES202/SSEL joint project, annual payroll in SSEL files includes non-wage payments, such as benefit payments, retirement pension funds, annuity funds, supplemental benefit funds, etc, which are not included in the UI files.

Figure 2.3: Workforce Characteristics of Immigrant, Mix and Native Firms Continuation



Note: Weighted share and percentage. Base on years 1992-1996.

Figure 2.4: Workforce Characteristics of Immigrant, Mix and Native Firms Continuation



Note: Weighted share and percentage. Base on years 1992-1996.

of owner. The last column of the table shows the t-test computed for worker type wages for each owner type. A t-test can reject the null hypothesis that the mean of immigrant worker wages and the mean of native worker wages are the same at the 90% level.

Table 2.5: Mean Earnings by Owner and Worker Type

Variable=log(annual earnings)	**(%)**	**Mean**	**STD**	**T-test**
owner = Immigrant				
Immigrant	50.30	8.35	*1.47*	
Native	49.70	8.12	*1.67*	
All	**100.00**	**8.23**	**1.64**	**24.20**
owner = Mix				
Immigrant	35.94	8.52	*1.86*	
Native	64.06	9.04	*1.71*	
All	**100.00**	**8.71**	**1.82**	**-16.07**
owner = Native				
Immigrant	15.87	8.32	*1.73*	
Native	84.13	8.38	*1.88*	
All	**100.00**	**8.37**	**1.73**	**-5.83**

Note:STD indicates standard deviation. Log annual wage in 1992 dollars. Using workers during the period 1992-1996.(*)T-tests are computed on the difference between average wages of immigrant and native workers for each specified owner type.

Looking at Table (2.5) we notice three relevant outcomes for wage differential analysis. First, immigrants are paid slightly less by native than by immigrant owners. On average, they are paid the lowest when working for native owners. Second, native workers are paid significantly less in immigrant owned businesses. Third, on average native owned firms pay more than immigrant owned firms. Fourth, mix-owned firms significantly pay less to immigrant workers. However, these firms employ a lower proportion of immigrant workers than immigrant-owned firms.

In sum, immigrant workers end up receiving lower log annual earnings than native workers. If we combine the first three outcomes, we can see that much

Table 2.6: By Similar Coworker Share: Mean Earnings by Owner and Worker Type

	Coworker Share					
	Below the median			Above the median		
Variable=log(annual earnings)	**(%)**	**Mean**	**STD**	**(%)**	**Mean**	**STD**
owner = imm						
Native	33.64	7.37	*1.71*	66.36	7.67	*1.53*
Immigrant	66.69	7.98	*1.68*	33.31	8.19	*1.34*
all	**48.09**	**7.74**	**1.51**	**51.91**	**7.82**	**1.70**
owner = mix						
Native	26.42	7.90	*1.79*	73.58	8.39	*2.98*
Immigrant	66.36	8.67	*1.63*	33.64	6.91	*1.21*
all	**39.07**	**8.32**	**1.75**	**60.93**	**8.13**	**1.98**
owner = usa						
Native	6.96	7.74	*1.81*	93.04	8.38	*1.89*
Immigrant	91.48	7.80	*1.77*	8.52	7.68	*1.96*
all	**20.33**	**7.78**	**1.79**	**79.67**	**8.31**	**1.92**

Note: STD indicates standard deviation. Log annual wage in 1992 dollars. Statistics based on estimation sample: all male individuals working between 1992 and 1996.

of the difference between the log annual wages of immigrants and natives comes from immigrants' propensity to work in immigrant owned firms. These firms pay the lowest wages, and the difference in immigrant earnings between immigrant and native firms is small. Additionally, native owned firms pay immigrant workers less than native workers (see Table(2.5)).

It is important to highlight the relevance of having actual earnings of each employee at the firm level, so we can exploit these variations to identify the effect of owner types on individuals' wages. Therefore, individual level wages are used in the regressions analyzed in the next sections. Table (2.5) would not be possible if we didn't have data on both employers and employees' characteristics. Our unique database allows us to compare average earnings between workers of different types holding a job in the same type of firm, and workers of the same type (native or immigrant) working for different types of owners.

We now perform a similar exercise, but separating firms by the share of coworkers similar to the worker called "similar coworker share" (see Table 2.6). This measure is different from the measure of immigrant coworker share defined previously, in that here we define the similar coworker share as the share of workers that are of a similar type to a particular worker in a specific firm. For instance, the coworker share of a native worker is the share of native born workers in the firm excluding the worker. The second column (%) shows the percentage of workers of each type in the firm accordingly below or above the similar coworker share median. We can see in the table that the previous findings in Table 2.5 remain valid. Foreign-born employers pay the lowest wages, on average. However, for businesses with coworker share below the median, immigrant employees working for immigrant employers are paid slightly more than immigrant employees working for native employers. Additionally, workers are paid more when working with similar coworkers. When workers' similar coworker share is below the median, employers pay lower annual wages. More than 65% of sample businesses have a mixed workforce, that is, the share of immigrant coworkers is neither one nor zero ($0 < share < 1$).

These tables do not control for individuals' characteristics, so we don't know the profiles of native and foreign employees holding jobs in these businesses. Nevertheless, these findings are striking. Immigrant owners pay the lowest on average. Furthermore, they to pay natives less than the rest of the market. This motivates the question of what type of native workers work for immigrant employers.

2.5.3 Sorting by Skill

Sorting by skill is a possible cause of sorting by owner type. The incentive to combine workers of identical skills within the same firm has been documented previously (Kremer and Maskin [1996]). Job descriptions and skill requirements are also a concern as characteristics of employers and employees are correlated. Additionally, if firms of different types have different skill mix productivity, that is, they use a combination of workers' skills and capital differently, then the differences in the probability of hiring a specific type of worker could be motivated by the capital/labor firm's decisions. For instance, immigrant owners could use labor more intensively than native businesses, or could hire more low-skilled workers than native firms. Immigrants, Hispanics, and other minority groups have lower skill on average so they may tend to work in low-skill sectors and low-skill jobs regardless of the owner type. Immigrant owners, on the other hand, may tend to concentrate in low-skill sectors because they also have low skill levels. For both group, the mayority of the firms are in the 'Low-Skill Industries'. Almost 30% of immigrant-owned firms belong to the 'High-Skill Industries' group, while more than 45% of native-owned firms belong to this group.

Table 2.7: Worker types distribution by owner's skill requirement

	Low-Skill Industries				High-Skill Industries			
Worker / Owner	**Immigrant**	**Native**	**Mixed**	**All**	**Immigrant**	**Native**	**Mixed**	**All**
Immigrant	38.60	11.60	27.50	15.40	33.70	9.20	41.60	11.50
Native	61.40	88.40	72.50	84.60	66.30	90.80	58.40	88.50
Race/ethnicity								
Hispanic*	19.23	10.12	16.78	11.40	17.73	6.44	20.05	7.43
Asian	18.22	2.63	8.23	4.64	17.09	2.97	24.30	4.34
Black	5.38	6.83	7.62	6.68	6.45	10.84	4.45	10.42
White (non-hispanic)	48.71	76.18	61.38	72.47	52.22	75.49	42.07	73.28
All	**70.83**	**54.16**	**62.24**	**100.00**	**29.17**	**45.84**	**37.76**	**100.00**

Note: Using Census 1990 information on workers' education attainment by industry, industries are separated into High Skill and Low Skill. High skill refers to those industries in which more than 50% of workers have at least a high school diploma. Otherwise we define the industry as low skill. (*) Hispanic refers to all races with ethnic group Hispanic. The group Other includes Native American and otherwise unclassified racial groups. Native-American workers represented only 0.5% of the total sample.

Table (2.7) shows workers' distribution by owner's skill requirement. The skill requirement for a firm is computed using Census 1990 data after compiling the share of workers by industry at the 2-digit level that have low educational attainment(less than high school) and high educational attainment(more than high school). High skill industries are those in which more than 50% of workers have at least a high school diploma. The remaining industries are low skill. The idea is to illustrate whether specific owner and worker types are concentrated in a particular skill group.

Not surprisingly, the table shows that firms in low-education industries have higher fractions of immigrant workers than firms in high-education industries. Immigrant firms continue to have a bigger proportion of immigrant workers, except for mix-owned businesses. Results are similar breaking down by workers' race. However, it is worth mentioning that immigrant-owned firms are more than 60% of the group of low-skill firms.

To account for part of this pattern, in the regressions below we include the share of workers in the firm in four education categories: high school dropouts, high school graduate, some college, and college graduate.

2.6 Regression Analysis

The ideal data to analyze the effect of owners, coworkers, and social connections on individual labor market outcomes requires information on individuals' labor market histories, earnings, and, specifically, the employer's source of ex-ante information about the job seekers that apply to its open vacancies. With this information we would be able to measure the actual hiring policies that firms use to find new workers.

Unfortunately, we don't have detailed data on hiring procedures used by firms. However, we do have a good deal of valuable information on the firms and workers. Workers can be divided into different categories by birth place or by race/ethnicity[42] to infer workers' and candidates' likely social connections. This, together with information on the type of owner, will help us infer the use of social ties in the firm's hiring process and its effect on workers' earnings. More specifically, network structure refers to the number of ties an individual has (Smith, 2000).

In this paper, we try to identify the impact of networks by using the proportion of coworkers who are potentially tied to a newly hired worker. Besides identifying the type of owner for whom the employee works, I use the proportion of similar employees in the firm at the time the new worker is hired as a measure of the network link between coworkers, employers, and the new worker.

Following each firm from 1992 to 1996, we obtain the number of employees who work for the firm and their earnings. We also have the total number of workers possessing any given set of demographic characteristics at each period of time. Following the definition of networks used in previous literature, we compute the share of similar coworkers for each new hire at each firm in each period, assuming that a similar birthplace or ethnicity implies at least a weak network connection between individuals.[43]

[42]White, Black, Hispanic, and Asian.

[43]At this point, it is worth to mention that even though immigrants are very diverse and it is a group that reflects a multiple gamma of ethnic/cultural backgrounds, not necessarily captures by the denomination of being foreign-born, it is also true that immigrants tend to have similar

A key challenge in linking owners and employees is that the characteristics of both owners and employees may be correlated with other characteristics of a workplace and its location. Section 2.5.3 above gives preliminary evidence on sorting by skill. The correlation between owner and employee types could also be a result of residential segregation of workers and owners (spatial mismatch). Job descriptions and skill requirements are also a concern, as characteristics of employers and employees are correlated. Immigrants, and in particular Hispanics, tend to be low skilled and therefore are likely to work in low-skilled sectors and low-skilled jobs regardless of the owner type. However, at the same time, immigrant owners could tend to concentrate in low-skill sectors, perhaps because they also have low skill levels.

Because the proportions of immigrants are unequally distributed across sectors and regions, we control for the 2-digit industry and geographic location of each firm. There exist sectors such as Retail, Services and Construction where immigrants represent a significant proportion of the workforce [44]. We also see this pattern in the geographic distribution of the immigrant population. For instance, according to Census 2000, Los Angeles and New York represent more than 30% of the total immigrant population in the country. To account for these concerns we need to control for fixed attributes of the workplace and the local labor market, and also for local trends in labor pool demographics. Therefore, we estimate the model controlling for

strategies to enter into the labor market regardless of their cultural background. Using migrant networks is one common factor among foreign-born workers, especially for new immigrants (Porter and Wilson [1980], Light [2006]).

[44] This can be also related to the fact that these sectors are also highly represented by relatively smaller firms than in Manufacturing, for instance.

characteristics of the firm (F_j) and local community (Z_j). These controls include the immigrant workforce population and population density in the local community, 2-digit industry code dummies, firm size (log of reported employment), and legal form of organization. We also include the share of the firm's workers in the four education categories discussed previously.

Previous research has noted the impact of English language ability in the use of networks and the level of wages for immigrant workers.[45] We capture this feature by interacting the 2-digit industry dummy with an English speaker dummy [46] This interaction is a proxy that intends to capture whether language is used differently in different industries.

In the wage regressions, we also control for individual characteristics (X_j), including worker's age, education and a dummy for working full time.[47] The composition of the labor pool might also be affected by changes over time in labor supply and demand. For example, white natives may be more likely to work in low-wage retail jobs when labor markets are weak. Therefore, we also include a dummy variable for each of the years in the sample (M_t) to control for national fluctuations in the labor market.

The identification strategy exploits variation across owner types for otherwise similar firms. By controlling for a rich set of firm characteristics we can narrow the possible alternative explanations for any residual correlation between owner type

[45]Hellerstein and Neumark [2007] and Hellerstein et al. [2008a]. For additional analysis, see 3 below.

[46]We identify a group of countries where English is the main language, and use this information to identify the worker as English speaker or otherwise.

[47]A worker with positive time during the complete year is considered full quarter worker or full year worker.

and worker outcomes.

2.6.1 Analysis of firms hiring patterns

This section starts by looking at the hiring patterns of the firm, estimating a model that predicts the probability that a newly hired employee is an immigrant. Firm hiring decisions indirectly reflect the way owners use current employees to help fill their job vacancies. We use a linear probability model to estimate the likelihood that a newly hired worker is of a particular type (immigrant or from a specific race/ethnic group).[48]

$$\Pr(\text{new hire:group}_i)_{kjt} = \\ c + B_1 * O_j + \delta * W_{jt-1} + B_2 * O_j * W_{jt-1} + \Phi * F_j + Z * Z_{kj} + T * M_t + \epsilon_{kjt} \quad (2.2)$$

Where k, j and t designate the worker, firm type, and time respectively. O_j is a vector of dummy variables for owner type (defined by immigration status or race). If i refers to the group of immigrant workers, we use as the reference group firms owned by immigrants. B_1 represents the vector of coefficients associated with the impact of owner type on hiring. The elements of this vector are expected to be negative when the omitted group is the same type as the new hire. For instance, the coefficient on native owners would be negative if immigrant-owned firms are more likely to hire new immigrant workers. W_{jt-1} corresponds to the vector of the proportion of workers of

[48] We use a linear probability model over a Probit (Logit) model because we don't need to restrict the sample to firms that hire at least one new worker of each type. This restriction could introduce sample selection bias because firms with zero hiring could have a completely different policy than those with a least one new hire.

type each type i at the firm in the previous period. An interaction between owner type and W_{jt-1} is included to asses differences in use of current employees' networks across owner types. I also control for firm characteristics F_j (a vector of variables measured at the firm level), year dummies M_t, and local community information and state dummies Z_{kj}.

In a regression with both owner type and coworker share included, the estimated coefficient on owner type will capture only the direct impact of owner type on hiring, not the total effect, which will include both the direct effect and the indirect effect coming through owner type's effect on coworker share. The use of employee referrals can be correlated with the type of owner and can affect hiring patterns if owners have the tendency to hire same-group individuals. When employees tend to refer same-group workers, the owner type's effect may be amplified. If we believe that the share of similar coworkers is a good proxy for social connections, these exercises illustrate the combined result of owner effects and hiring patterns.

We assume that the error (ϵ_{kjt}) in equation 2.2 is independent and identically distributed across firms, but not within firms. To correct for non spherical disturbances, we estimate Huber-White robust standard errors clustered by firm. This procedure is used in all subsequent estimations. We cluster the errors by firm since firms in the sample may have hired more than one worker and thus may have repeated observations.

For purposes of analysis, we estimate different versions of equation (2.2) and look at the impact of the addition of controls on the estimates of B_1 and B_2. The first regression includes only year dummies; subsequent specifications add controls one by one. Most of the literature on hiring networks argues that current workers' referrals are more important to firm hiring patterns than owners' personal networks. Owners are likely to hire individuals from their residential area. However, current workers have a larger and more diverse set of connections that can be exploited by the firm. We are not able to disentangle these effects directly. Nevertheless, by allowing owners of different groups to make use of their workers' social ties differently, the estimated interaction effects can measure the ability of owners to use social ties.

Table (2.8) shows the probability of a new hire being an immigrant given the characteristics of the firm, its community and the share of immigrant coworkers in the firm. Controlling only for year dummies, native owners are 25 percentage points less likely to hire a new immigrant worker than immigrant firms (column 1). This difference is significantly reduced, to 3.5 percentage points, when we include the share of immigrant coworkers (column 2). Controlling for year and industry dummies, the share of immigrant coworker positively affects the likelihood of an immigrant being hired. The inclusion of the share of English speaker and its interaction with industry dummies decreases the impart of the share of immigrant coworkers on the probability of being hired. This covariates controls for whether language is used differently in different industries (column 3). For instance, a Mexican restaurant would probably hire Mexicans or Spanish speaker because of the type of service they offer and type of frequent consumers. The use of language can be different in a industry where workers don't need to communicate with each other, so language differences are not obstacle in the production process. Given the results, it seems that firms use language in different ways, affecting the likelihood of an immigrant

Table 2.8: Linear Estimates of the Effect of Owner Type on the Probability that a New Hire is an Immigrant

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(FE)**
Owner Mix	-0.0519***	-0.041***	-.0034**	-0.0037**	-0.00313**	
	0.0057	*0.0065*	*0.001*	*0.001*	*0.001*	
Owner Native	-0.2358***	-0.0351***	-0.0342***	-0.033***	-0.0254***	
	0.0032	*0.0031*	*0.0009*	*0.0004*	*0.0014*	
% Imm. Coworkers		0.9961***	0.782***	0.7724***	0.7132***	0.6715***
		0.0056	*0.002*	*0.0101*	*0.0234*	*0.0435*
% Imn. Coworkers				-0.0125**	-0.0094**	
* Owner Mix				*0.005*	*0.005*	
% Imm. Coworkers				-0.0711***	-0.0378***	
* Owner Native				*0.003*	*0.0041*	
Corporation					-0.00085*	
					0.0033	
Sole Prop.					0.0026	
					0.003	
log(employment)					0.003	
					0.002	
Share of workers					0.0021**	
with HSD (firm)					*0.0004*	
Share of workers					-0.0012	
with HSG (firm)					*0.001*	
Share of workers					0.005	
with SCG (firm)					*0.006*	
Pop. % immigrant					0.0162**	
in neighborhood(+)					*0.0068*	
Population in					0.0004***	
neighborhood(+)					*0.00*	
In MSA					-0.005***	
					0.0009	
Constant	0.4081***	0.0211***	0.0989***	0.0969***	0.0285**	0.1575*
	0.099	*0.0039*	*0.002*	*0.0016*	*0.0069*	*0.781*
Dummies						
year	yes	yes	yes	yes	yes	yes
Industry	yes	yes	yes	yes	yes	-
Indus*English Spkr	-	-	yes	yes	yes	-
R-Square	*0.29*	*0.32*	*0.34*	*0.35*	*0.38*	*0.41*

Note: Reference group is immigrant firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. (+) Neighborhood is defined counties adjacent to the county where the firm is located. Population in 100,000's. FE represents the firm fixed-effect model. ***significant at 1%, ** significant at 5%, * significant at 10%.

being hired and reducing the impact of immigrant coworkers in the firm.

There is a positive and significant impact on the probability of the new hire being an immigrant when the proportion of workers in the firm with low education (high school dropout) increases. The owner effect diminishes and the difference in the probability of hiring a Hispanic between immigrant and native owners is 2.5 percentage points (column 4). The coworker effect is smaller too, although it is still significant. The interaction effects between owner type and coworker share decrease slightly when others controls are included, although the results are similar. The effect of immigrant coworker share is smaller in mix and native owned firms than in immigrant owned firms. Immigrant employers can take advantage more efficiently of their current immigrant workers than other types of employers. The increment of immigrant coworker share by 1 percentage point increases this likelihood by 0.71-0.67. The inclusion of other characteristics of the firm and the local community has a smaller impact on the relative likelihood of native versus immigrant owners hiring a new immigrant worker.

We should be cautious when analyzing these results. We include a vast series of covariates to control for all possible observables that can be correlated with employer and employee effects. However, the presence of unobservables correlated with firm and worker interactions could bias the results. As another exercise, we compute the firm fixed-effect version of the model by including firm dummies. The last column of Table (2.8) shows the results. The impact of share of immigrant coworkers in the firm at the time of the new hire remains positive, high, and significant.

2.6.2 Hiring Process by Race/Ethnicity

We next consider the determinants of the probability that a new hire comes from a particular race/ethnic group: white, black, Hispanic and Asian. That is, we estimate equation (2.2), setting i equal to a particular racial category. Tables 2.9 and F.1 show the effects of owner types and shares of type i coworker, and other types of coworker, at the time of hiring on the probability that a new hire is Hispanic, Asian, white, or black respectively.

2.6.2.1 **Worker Race**

The likelihood of a new worker being Hispanic or Asian significantly decreases when the employer is native. This result holds even after including a exhaustive list of controls(Tables 2.9 and 2.10). The direct impact of owner type is reduced, however, once we control for the share of Hispanic coworkers. For instance, having a one percentage point increment of Hispanics as current employees in the firm increases the probability that a new hire is Hispanic (by up to 0.88 in immigrant owned firms). The impact of Hispanic coworkers is smaller for native owned firms.

Table 2.9: Linear Probability Estimates of the Effect of Owner Type on the Probability that a New Hire is Hispanic

			Hispanic			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**FE**
Owner Mix	0.0214***	0.012	0.0077	0.014	-0.0694	
	0.0041	*0.01*	*0.005*	*0.0357*	*0.054*	
Owner Native	-0.0903***	-0.0872***	-0.0412***	-0.0245**	-0.0172**	
	0.003	*0.003*	*0.003*	*0.001*	*0.001*	
Hispanic Cowkrs			0.9441***			
			0.0054			
Asian Cowkrs				-0.628***	-0.526***	-0.504***
				0.013	*0.023*	*0.029*

Continued on next page.

Table 2.9: Linear Probability Estimates of the Effect of Owner Type on the Probability that a New Hire is Hispanic (continued)

	Hispanic					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**FE**
White Cowkrs				-0.703***	-0.681***	-0.596***
				0.0095	*0.0075*	*0.0197*
Black Cowkrs				-0.879***	-0.725***	-0.616***
				0.0123	*0.0134*	*0.0212*
Hispanic Cow*			-0.0621**			
Owner Mix			*0.031*			
Asian Cow*				0.1060**	0.093**	
Owner Mix				*0.0483*	*0.034*	
White Cow*				0.0185	0.0175	
Owner Mix				*0.0384*	*0.0434*	
Black Cow*				0.137	0.105	
Owner Mix				*0.0845*	*0.0945*	
Hispanic Cow*			-0.0869***			
Owner Native			*0.043*			
Asian Cow*				-0.113***	-0.102***	
Owner Native				*0.0295*	*0.0243*	
White Cow*				-0.148**	-0.124**	
Owner Native				*0.062*	*0.056*	
Black Cow*				-0.094*	-0.097*	
Owner Native				*0.04*	*0.04*	
log(employment)					0.0013**	
					0.00	
Share of workers					0.0012***	
with HSD (firm)					*0.0005*	
Share of workers					0.0025***	
with HSG (firm)					*0.0003*	
Share of workers					0.0030***	
with SCG (firm)					*0.0008*	
Work. Pop.					0.0561**	
Total.$_1$					*0.0245*	
Work. Pop.					*0.094****	
%Hisp $_1$					*0.002*	
Constant	0.1920***	0.1243***	0.9702***	0.8454***	0.9511***	0.8411***
	0.0032	*0.009*	*0.08*	*0.1616*	*0.171*	*0.201*
year dummies	yes	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes	-
State dummies	-	-	yes	yes	yes	-
Other controls(+)	-	-	-	-	yes	-
p-value	*0.0001*	*0.002*	*0.0001*	*0.003*	*0.003*	*0.01*
R-Square	*0.22*	*0.29*	*0.31*	*0.34*	*0.42*	*0.35*

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

The effect of the share of Hispanic coworkers is positive regardless the type of the owner. However, the effect is smaller than the baseline effect on the Hispanic-owned firms (column 3). Columns 4 includes the effect of all races coworker share on the likelihood of being hired. Other races coworker shares affect negatively the probability of a new hire is Hispanic. Interestinly though, Asian coworker share is less negative when the firm is mix-owned. Column 5 includes other firm and local community characteristics. Their inclusion decreases the average effects, but do not change the directions of the results.

In section (2.5.3) we discussed the distribution of workers by average industry-level skill requirement. As a proxy to control for this effect, we include the firm's share of workers in four education categories and the fraction of workers of similar type in the local community. The results show that a higher share of low-educated workers in the firm increases the probability that the new worker is Hispanic. We also include the share of workers of each racial group in the local labor force. The inclusion of these shares decreases the impact of the coworker shares.

Looking at Asian new hires (Table 2.10), we again find that native employers are less likely to hire Asian workers. The inclusion of additional controls reduces the difference in probability of hiring an Asian between immigrant and native owned firms. Another interesting result is that Asians are less likely to be hired in firms with bigger proportion of workers with education attainment below the high school level.

Table 2.10: Linear Probability Estimates of the Effect of Owner Type on the Probability that a New Hire is Asian

	Asian					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**FE**
Owner Mix	-0.029***	-0.0280***	0.002	0.001	0.007	0.1948
	0.0032	*0.0032*	*0.003*	*0.002*	*0.002*	*0.254*
Owner Native	-0.1245***	-0.1114***	-0.054**	-0.052**	-0.06**	
	0.002	*0.002*	*0.002*	*0.002*	*0.014*	
Hispanic Cowkrs				-0.743***	-0.712***	-0.622***
				0.0079	*0.008*	*0.008*
Asian Cowkrs			0.8194***			
			0.031			
White Cowkrs				-0.7876***	-0.741***	-0.6715***
				0.0071	*0.0072*	*0.0074*
Black Cowkrs				-0.8795***	-0.8214***	-0.7631***
				0.0083	*0.0081*	*0.0083*
Hispanic Cow*				-0.0197	-0.0556	
Owner Mix				*0.0303*	*0.041*	
Asian Cow*			0.007			
Owner Mix			*0.002*			
White Cow*				-0.0746***	-0.076***	
Owner Mix				*0.0022*	*0.0022*	
Black Cow*				0.032	0.0404	
Owner Mix				*0.0446*	*0.0536*	
Hispanic Cow*				-0.0064	-0.0064	
Owner Native				*0.0185*	*0.0185*	
Asian Cow*			-0.152***			
Owner Native			*0.013*			
White Cow*				-0.0158	-0.031	
Owner Native				*0.0154*	*0.0221*	
Black Cow*				-0.0076	-0.0095	
Owner Native				*0.02*	*0.02*	
log(employment)					0.0014**	
					0.00	
Share of workers					-0.0012**	
with HSD (firm)					*0.0001*	
Share of workers					-0.000	
with HSG (firm)					*0.00*	
Share of workers					-0.0013**	
with SCG (firm)					*0.00*	
Work. Pop.					0.024*	
% Imm.1					*0.001*	
Work. Pop.						
%Hisp 1						
Work. Pop.					0.043**	
% Asian 1					*0.01*	
Constant	0.1495***	0.112**	0.065**	0.095***	0.094***	0.097
	0.0018	*0.0562*	*0.02*	*0.01*	*0.01*	*0.081*
year dummies	yes	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes	-
State dummies	-	-	yes	yes	yes	-

Continued on next page.

Table 2.10: Linear Probability Estimates of the Effect of Owner Type on the Probability that a New Hire is Asian (continued)

	Asian					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**FE**
Other controls(+)	-	-	-	-	yes	-
p-value	*0.0001*	*0.002*	*0.0001*	*0.003*	*0.003*	*0.01*
R-Square	*0.26*	*0.31*	*0.34*	*0.35*	*0.37*	*0.38*

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

Whites and blacks are more likely to be hired by native firms (See Tables (F.1) and (F.2)). However, the probability that a new hire is black or white depends on the share of blacks or whites in the firm at the time of the recruitment process. The significance of the immigrant owner effect on black hiring vanishes when I include the black coworker share in the regression. The column FE shows the results of the regression after including firm fixed effects. The impact of similar coworkers decreases slightly but is still high and significant. The largest change in coefficients caused by the inclusion of fixed effects is the drop in the impact of white coworkers on the probability of that a new hire is black.

We also experiment with estimating a multinomial logit model to account for the posibility that employers may simultaneously choose among different types of workers. The estimation sample is then restricted to firms that hire at least one worker of each race group during the period 1992-1996. This restriction eliminates more homogeneous firms. The new sample contains 2,662 firms out of the original sample of 4,478 firms. We investigate how the owner type and shares of different types of workers at the time of hiring affect the type/race of the new hire. We estimate a model[49] that aims to reveal whether the birthplace of the employer affects the likelihood that a new worker is of the same type as opposed to other types, conditional on having accessed to the firm during the period of analysis and controlling for the characteristics of the worker and the firm.

$$\Pr(\text{new hire is worker type: } i)_{kjt} = \frac{exp(c^i + B_1^i * O_j + \delta^i * W_{jt-1} + \Phi^i * F_j + Z^i * Z_{kj} + T^i * M_t + \epsilon^i_{kjt})}{\sum_{s=1}^{5} exp(c^s + B_1^s * O_j + \delta^s * W_{jt-1} + \Phi^s * F_j + Z^s * Z_{kj} + T^s * M_t + \epsilon^s_{kjt})} \tag{2.3}$$

with $i = 1, ..., 4$ for the four race groups: white, black, Asian, and Hispanic. This procedure makes very strong assumptions with respect to the relevance of other alternatives. The odds ratio of any two options is assumed independent of the other alternatives. This feature is important to consider when more than two alternatives are included. To test the Independence of Irrelevant Alternatives assumption, we conduct a Hausman test by excluding each outcome category in turn. The test indicates that I cannot reject the null hypothesis that the odds of one outcome happening are independent of other alternatives. Additionally, we perform Wald tests for combination of categories. The tests reject the null hypotheses that all coefficients associated with a given pair of outcomes are zero (except intercepts). We cluster the errors by firm since observations within firms are not independent. The results for this regression are shown in Tables (2.11) and (2.12).

[49] I specifically estimate a mixed logit model that incorporates both characteristics of the individual and the alternatives.

Table (2.11) shows the change in log odds comparing two alternatives. The share of white coworkers significantly increases in the log odds of a white being hired. We also show the predicted hiring probabilities for each owner type (Table 2.12) computed at the means of all firms and dummy variables. The change in log odds between hiring a white worker versus hiring a Hispanic or an Asian decreases when the firm is immigrant-owned. Immigrant owners are 3 percentage points more likely to hire Asians and Hispanics than native firms. These results support the analysis in the previous section.

Table 2.11: Multinomial Logit Model: Effects of Owner Type and Coworkers on Type of New Hires

	Change in log odds comparing alternative 1 to alternative 2					
Cow. Share	White to Black	White to Asian	White to Hispanic	Black to Hispanic	Black to Asian	Asian to Hispanic
White	**1.97*** **	**2.32*** **	**3.53*** **	1.44*	0.53	0.92
	0.646	*0.761*	*0.421*	*0.71*	*0.723*	*1.017*
Black	**-5.352*** **	1.186	2.145***	**7.456*** **	**5.456*** **	1.014
	0.892	*1.086*	*0.661*	*0.957*	*0.968*	*1.131*
Asian	-1.391	**-7.243*** **	0.041	1.433	-5.682***	**7.126*** **
	0.951	*1.001*	*0.591*	*1.108*	*0.946*	*1.143*
Hispanic	-0.236	*-0.086*	**-3.675*** **	**-3.127*** **	0.15	**-3.654*** **
	1.03	*1.102*	*0.527*	*1.09*	*0.952*	*1.361*

Note: Other controls include log of employment, percentage of immigrant workers in the surrounding counties, population in the county, legal form of organization, Msa location, 2-digit industry, interaction 2-digit industry and English speaker dummy, state and year dummies. Results from race/ethnicity 'others' are not shown. Number of observation 135,583 workers, and 2,662 firms. Robust standard errors in italic allow for arbitrary correlation within the same firm.
* significant at 10%,** significant at 5%, *** significant at 1%.

Table 2.12: Multinomial Logit Model: Predicted Probability of Covariates

	Workers			
Owner	White	Black	Asian	Hispanic
Native	0.740	0.120	0.031	0.100
Immigrant	0.710	0.102	0.060	0.126
Mix	0.690	0.119	0.052	0.134

Note: Based on multinomial logit predictions of the race of new hires from previous table.

2.6.2.2 Worker and Owner Races

After looking at the effect of owner birthplace on the probability of being hired for each particular worker's race, the natural question is whether we can detect similar effects when we separate owner types by race. As explained in Section 2.4.1, owner's race is obtained from the Small Minority Owner Business Employers Survey(SMOBE). For multiple-owned firms, the median race is used; in the case of ties, the hours worked in the firm are also considered to determine the predominant race of the firm. The race categories are: white, black, Asian and Hispanic.

The likelihood of a new worker being Hispanic or Asian significantly decreases when the employer is White. This result holds even after including a exhaustive list of controls(Tables 2.13 and 2.14). The direct impact of owner type is reduced, however, once we control for the share of Hispanic coworkers. For instance, an increment of one percentage point in the share of Hispanics as current employees in the firm increases the probability that a new worker is Hispanic (by up to 0.95 in Hispanic owned firms).

Table 2.13: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Hispanic

			Hispanic		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Owner Black	-0.2215***	-0.1514***	-0.024**	-0.0284**	-0.0165**
	0.0271	0.0085	0.0112	0.015	0.0013
Owner Asian	-0.135***	-0.1264***	-0.0257***	0.0868	0.1429
	0.0229	0.0043	0.0076	0.7279	0.1309
Owner White	-0.2358***	-0.176***	-0.0318***	-0.0231***	-0.0158***
	0.0154	0.0034	0.0064	0.0015	0.0045
Hispanic Cowkrs			0.9512***		
			0.0176		
Asian Cowkrs				-0.918*	-0.9114***
				0.0669	0.0707

Continued on next page.

Table 2.13: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Hispanic (continued)

	Hispanic				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
White Cowkrs				-0.9086***	-0.6494***
				0.0231	0.0269
Black Cowkrs				-0.8322***	-0.6335***
				0.0317	0.0349
Other Cowkrs				-0.6857***	-0.5814***
				0.0434	0.0478
Owner Black*			-0.0335**		
Hispanic Cowkrs			0.0145		
Owner Black*				0.0367	0.0922
Asian Cowkrs				0.2128	0.2218
Owner Black*				0.0132	0.0024
White Cowkrs				0.0678	0.0748
Owner Black*				-0.0382**	0.0063
Black Cowkrs				0.0109	0.0783
Owner Black*				0.0048	0.092
Other Cowkrs				0.1592	0.1826
Owner Asian*			-0.0427*		
Hispanic Cowkrs			0.0252		
Owner Asian*				-0.0531	-0.1636**
Asian Cowkrs				0.0711	0.0756
Owner Asian*				-0.0869**	-0.1622***
White Cowkrs				0.0326	0.037
Owner Asian*				-0.1764***	-0.144***
Black Cowkrs				0.0493	0.0533
Owner Asian*				-0.2705***	-0.33***
Other Cowkrs				0.0627	0.0676
Owner White*			-0.09***		
Hispanic Cowkrs			0.0194		
Owner White*				-0.2077***	-0.1057*
Asian Cowkrs				0.0703	0.0748
Owner White*				-0.1218***	-0.0811**
White Cowkrs				0.0254	0.0296
Owner White*				-0.2015***	-0.1071***
Black Cowkrs				0.0342	0.0385
Owner White*				-0.2358***	-0.184***
Other Cowkrs				0.0486	0.054
Share of workers					0.0016***
with HSD (firm)					0.0005
Share of workers					0.0022***
with HSG					0.0003
Share of workers					0.0032***
with SOG					0.0009
Log employment					0.0011*
					0.0009
Work. Pop.					0.0422***
Total					0.024
Work. Pop.					0.0002**

Continued on next page.

Table 2.13: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Hispanic (continued)

	Hispanic				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
% Hisp.					0
Constant	0.3286***	0.0436*	0.0211	0.8779	0.6446
	0.0136	0.0175	0.0828	0.0864	0.1755
Year dumies	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes
State dummies	-	yes	yes	yes	yes
Other Controls	-	-	-	yes	yes
p-value	0.01	0.01	0.01	0.01	0.01
R-Square	0.24	0.29	0.31	0.33	0.45

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

The impact of Hispanic coworkers is smaller for other types of firms. The results for other characteristics of the firm and its location are similar to previous sections. The results show that a higher share of low-educated workers in the firm increases the probability that the new worker is Hispanic. I also include the shares of coworkers in each racial group. Black and White owned firms are 2 to 3 percentage points less likely to hire a Hispanic worker compared to Hispanic and Asian owned firms, holding constant the worker race distribution.

Looking at Asian new hires (Table 2.14), we find that white employers are less likely to hire Asian workers. White owners are mostly natives. The inclusion of additional controls reduces the difference in probability of hiring an Asian between Asian and white owned firms. Another interesting result is that Asians are less likely to be hired in firms with bigger proportion of workers with educational attainment below the high school level.

Table 2.14: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Asian

	Asian				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Owner Black	-0.181***	-0.1771***	-0.06***	-0.2035*	-0.1628***
	0.0067	0.0054	0.006	0.1164	0.0268
Owner Hispanic	-0.1977**	-0.1842***	-0.013	-0.0333***	0.0503
	0.0215	0.0027	0.0034	0.00419	0.0462
Owner White	-0.187***	-0.1775***	-0.049***	-0.1257***	-0.1749***
	0.0049	0.0021	0.0026	0.0149	0.0163
Asian Cowkrs			0.98***		
			0.0074		
Hispanic Cowkrs				-0.987***	-0.8883***
				0.0149	0.017
White Cowkrs				-0.920***	-0.7414***
				0.0101	0.0118
Black Cowkrs				-0.978***	-0.7986***
				0.0202	0.022
Other Cowkrs				-0.9928***	-0.8541***
				0.0237	0.0257
Owner Black*			-0.2765***		
Asian Cowkrs			0.1041		
Owner Black*				0.1801	0.1481
Hispanic Cowkrs				0.1258	0.1367
Owner Black*				0.1915	0.1471
White Cowkrs				0.1186	0.1305
Owner Black*				0.2172	0.2031
Black Cowkrs				0.1184	0.1299
Owner Black*				0.3027	0.2474
Other Cowkrs				0.1514	0.17
Owner Hispanic*			-0.0374		
Asian Cowkrs			0.0339		
Owner Hispanic*				0.0019	-0.0373
Hispanic Cowkrs				0.044	0.0488
Owner Hispanic*				-0.0421**	-0.0302**
White Cowkrs				0.0235	0.0129
Owner Hispanic*				-0.0032	-0.0009
Black Cowkrs				0.0506	0.0559
Owner Hispanic*				-0.0684	-0.0754
Other Cowkrs				0.0552	0.0608
Owner White*			-0.2475***		
Asian Cowkrs			0.0131		
Owner White*				0.1633***	0.2095
Hispanic Cowkrs				0.0203	0.0229
Owner White*				-0.1158***	-0.1092***
White Cowkrs				0.0158	0.0177
Owner White*				-0.144***	-0.1205***
Black Cowkrs				0.0242	0.0263
Owner White*				-0.1223***	-0.108***
Other Cowkrs				0.0287	0.0316
Share of workers					-0.0017***

Continued on next page.

Table 2.14: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Asian (continued)

	Asian				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
with HSD (firm)					0.0004
Share of workers					-0.0001
with HSG					0.0002
Share of workers					-0.0021***
with SOG					0.0005
Log employment					0.0033***
					0.0006
Work. Pop.					0.0643***
Total					0.0154
Work. Pop.					0.0002***
% Asian					0
Constant	0.2431***	0.2131***	0.0373	0.0275	0.1242
	0.0052	0.0557	0.0519	0.0528	0.1129
Year dumies	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes
State dummies	-	yes	yes	yes	yes
Other Controls	-	-	-	yes	yes
p-value	0.01	0.01	0.01	0.01	0.01
R-Square	0.28	0.29	0.35	0.38	0.43

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

Whites are more likely to be hired by white owned firms (see Table F.4). However, the probability that a new hire is white depends postively on the share of whites in the firm at the time of the recruitment process. The owner's race effect is lower for those owners from a different racial group.

Multinomial analysis is also applied to the combination of worker and owner races. The results for this regression are shown in Tables (2.15) and (2.16). Table (2.15) shows the change in log odds comparing two alternatives. The change in log odds between hiring a white worker versus hiring a Hispanic or an Asian decreases when firms are Hispanic or Asian owned. A higher share of white coworkers significantly increases the log odds of a white being hired, and a similar result holds for other races. We also show the predicted hiring probabilities for each owner race (Table 2.12) computed at the means of all firms and dummy variables. Hispanic owners are 3 percentage points more likely to hire Asians and Hispanics than White and Black owners. These results support the analysis in previous sections.

In sum, Hispanic and Asian workers are generally more likely to be hired by Hispanic or Asian owned firms. In this detailed presentation, it seems that Asian owned firms tend to employ Asian and Hispanic workers more readily than black and white workers. Almost 70% of Asian and Hispanic owners are immigrants. We would also like to analyse the impact of immigrant/native oner effects after controllinf for owner race, including owner birthpalce and race simultaneously. However, the variation across the sample is not enough to identify whether birthplace or owner race is more important. Most of the native owners are either white or black, with a large proportion of them being white. While, our sample has a small representation of black immigrant owners.

Given the structure of our sample, white and black owners are mainly natives, while Asian and Hispanic owners are immigrants, and after looking at our by racial groups results, we can see that our previous result. Immigrant owners tend to hire immigrant workers, while native owners tend to hire native workers.

Table 2.15: Multinomial Logit Model: Effects of Owner's Race and Coworkers on Type of New Hires

Covariates	**Change in log odds comparing alternative 1 to alternative 2**					
	White to Black	White to Asian	White to Hispanic	Black to Hispanic	Black to Asian	Asian to Hispanic
Owner Black	-0.379***	-0.313*	-0.234	0.145***	0.067	-0.079
	0.03	0.15	0.2	0.04	0.1	0.1
Owner Hispanic	-0.09*	-0.186***	-0.052**	-0.039**	-0.096***	-0.134***
	0.06	0.03	0.01	0.01	0.02	0.03
Owner Asian	0.11	-0.229***	-0.113	-0.340**	-0.223**	0.116***
	0.09	0.03	0.03	0.04	0.03	0.02
Cow. Share white	2.522***	1.7560**	6.631***	4.111***	0.766	1.875
	0.892	0.413	1.121	1.153	0.651	1.034
Cow. Share black	-7.130***	-0.823	-0.385	6.745***	6.308***	0.438
	1.203	0.723	0.241	1.412	1.324	0.56
Cow. Share Asian	-3.532***	-7.742***	-1.832*	1.700	-4.210***	5.910***
	0.731	1.202	0.891	1.342	1.154	1.265
Cow. Share Hispanic	-1.522*	-1.756***	-6.632***	-4.210**	0.667	-4.875***
	0.641	0.952	1.678	1.023	0.801	1.123

Note: Other controls include log of employment, percentage of immigrant workers in the surrounding counties, population in the county, legal form of organization, Msa location, 2-digit industry, interaction 2-digit industry and English speaker dummy, state and year dummies. Results from race/ethnicity 'others' are not shown. Number of observation 135,583 workers, and 2,662 firms. Robust standard errors in italic allow for arbitrary correlation within the same firm.
* significant at 10%,** significant at 5%, *** significant at 1%.

Table 2.16: Multinomial Logit Model: Predicted Probability of Covariates Owner and Worker Races (%)

	Workers			
Owner	White	Black	Asian	Hispanic
White	74.11	11.99	3.05	10.64
Black	65.66	15.33	3.90	11.97
Asian	60.92	10.91	6.84	12.50
Hispanic	61.00	12.75	6.71	13.87

Note: Based on multinomial logit predictions of the race of new hires from previous table.

2.6.3 Workers' earnings and analysis of results

We estimate the effects of owner type and coworker shares on workers' compensation using a human capital approach. The dependent variable is the natural logarithm of workers' real annual wages.[50] The regression includes dummy variables for owner type, the share of similar coworkers, worker type, and other firm characteristics. Using wage estimates at the individual level, we can evaluate the impact of owners' characteristics on wage differentials by using equation (2.4).

$$
\begin{aligned}
ln(w_{kjt}) = \quad & c + \beta_1 * I_k + X_k^{'} * B_2 + O_j^{'} * B_3 + I_k * O_j^{'} * B_4 \\
& +COW_{kj}^{'} * B_5 + I_k * COW_{kj} * B_6 \\
& +O_j^{'} * COW_{kj}^{'} * B_7 + I_k * O_j^{'} * COW_{kj}^{'} * B_8 \\
& +F_j^{'} * \Phi + Z_{kj}^{'} * Z + T * M_t + \mu kjt
\end{aligned} \tag{2.4}
$$

[50] In order to approximate the individual's full-year annual wage rate and thus reduce the importance of within-year labor supply decision, we include the additional information of whether the worker is a full quarter employee. That is, full quarter worker is an individual with positive earnings during all the quarters of the year. Controlling for full quarter workers allows us to make UI's annual earnings comparable to CPS salary and wages. Abowd et al. [2002] have a discussion on the comparison between LEHD and CPS annualized wages. After controlling for dominant employer and full-time status, CPS and LEHD earnings data are more comparable. LEHD annualized wages are slightly higher than CPS' annualized wages. However, when looking at our analysis we should keep in mind that an individual's labor supply depends on both the duration and the average number of hours worked at the job.

where k identifies information on the worker and j refers to information on the firm. w_{kj} stands for worker k's log real annual earnings at firm j. I_k is a dummy variable for whether the worker is an immigrant. In an effort to establish how much the immigrant earnings differential is due to differences in predetermined personal characteristics, we add a vector X_k of employee characteristics including age, age squared, education, sex, and race. O_j is a vector of dummy variables for owner type (birthplace). COW_{kj} stands for the proportion of immigrant coworkers in the firm (explained in section 2.4.5). The expected sign for β_1 is negative, assuming that immigrants earn lower wages, and its significance would indicates whether there is substantial wage variation across the different worker types. With the inclusion of owner type dummies, the estimate of β_1 will represent the difference in wages between immigrants and natives in native owned firms. The sum of β_1 and the B_3 and B_4 coefficients corresponding to an immigrant owned firm will be positive if immigrant workers earn higher wages when working for immigrant-owned businesses than native workers in an immigrant firm. The coworker share accounts for the potential impact on wages of having better connections to similar types of workers in the firm. The interaction between COW_{kj} and the vector of owner types is included to assess whether the effect of coworkers differs according to the type of employer that is hiring the employee. We explore a 3-way interaction among owner type, worker type, and the immigrant coworker share. In equation (2.4), B_2X_k absorbs the effects of variations in personal characteristics. We would expect estimates of β_1 and the vector B_3 to change after including workers' characteristics.

We should be aware of the potential presence of omitted variable bias. Unobservable characteristics could bias estimated coefficients in equation (2.4). Ignoring these unobservables could causes us to overestimate the impact of owner type and immigrant coworkers on individual earnings. High ability workers of type k should look for firms that pay higher earnings. If native-owned firms offer higher wages and employ these high ability workers, the estimated model would not be capturing the effect of owner type on workers' earnings; rather it would be capturing individuals' ability to find better jobs. Also, worker preferences and comparative advantage can influence the results. Variations in preferences for particular job characteristics across different workers could provide an alternative explanation for both earnings differentials and sorting. To account for some of this variation, we include the fraction of workers in the firm with education lower than high school, equal to high school, higher than high school with some college, and equal to college or higher. The omitted category is college graduate.

Characteristics of firms (F_j) and of the local community (Z_j) are also included. These controls include the population share of each group in the local community, population density, firm's size (log of reported employment), and legal form of organization. M_t are year dummies.

The first column of Table (2.17) shows results from a baseline model including immigrant status, individual age, education, and part-time status, but excluding other variables of interest. The table reports the betas estimated by equation (2.4). To make the analysis easier to interpret, we transform these unstandardized β coefficients with the usual formula $[(e^{\beta}-1)*100]$, so that we can analyze the percentage change in wages associated with a 1-unit change in a continuous independent predictor variable. In the case of a dichotomous independent variable, we interpret the percentage wage difference in the target category compared to the reference category. After controlling for typical human capital variables, full-time immigrant workers earn about 8% less than native workers (3,293 dollars less each year). In the Table (2.17), we progressively include covariates that control for firms and coworker shares. Column 1 shows the typical human capital analysis. Column 2 includes owner dummies and their interaction with worker type. Then, in column 3, immigrant coworker share is included and its interaction with worker type to see whether the effect of immigrant coworker share differs across worker types. Then, the interaction of immigrant coworker share and owner types are added to the regression (column 5). Finally, we include a 3-way interaction among immigrant coworker share, worker type, and owner type.

Evaluating the variables at their means and sample distribution, we find that, when working for native employers, the difference between native and immigrant wages increases to 11%. Meanwhile, immigrant workers earn 10% more than native workers in immigrant owned firms (4,398 dollar more each year).

The human capital results in Table (2.17) are consistent with the literature. Age positively affects wages but at a decreasing rate. Education is significant and positive. Part-time workers earn less than full-time workers. The inclusion of additional independent variables does not modify these patterns. After controlling for individual characteristics, immigrant workers are paid less than native workers in native firms, but they receive a significantly higher wage than native workers when working for immigrant firms. The inclusion of the share of immigrant coworkers

Table 2.17: OLS Results: Effect of Owner Type and Coworker Share on Log Real Annual Wages

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Immigrant	-0.08***	-0.1503***	-0.1205***	-0.1171***	-0.1017***
	0.007	*0.0069*	*0.0073*	*0.0025*	*0.0008*
Age	0.0806***	0.080***	0.0803***	0.0802***	0.0749***
	0.0007	*0.0007*	*0.0007*	*0.0006*	*0.0008*
Age square (’)	-0.080***	-0.080***	-0.080***	-0.080***	-0.080***
	0.000	*0.000*	*0.000*	*0.000*	*0.000*
Education	0.506***	0.506***	0.511***	0.504***	0.484***
	0.0007	*0.0007*	*0.0007*	*0.0007*	*0.0007*
Partime	-2.1847***	-2.1805***	-2.1792***	-2.1783***	-2.1240***
	0.0044	*0.0044*	*0.0044*	*0.0036*	*0.0049*
Owner Mix		0.1808***	0.1615***	0.1407***	0.0936
		0.0128	*0.013*	*0.0481*	*0.0155*
Owner Immigrant		-0.1191***	-0.1443***	-0.1495***	-0.1087***
		0.0097	*0.0101*	*0.017*	*0.012*
Owner Mix*Immigrant		0.0054	-0.0007	0.1866	0.1368
		0.0224	*0.02*	*0.251*	*0.243*
Owner Immigrant*Immigrant		0.3205***	0.3030***	0.3174***	0.3131***
		0.0251	*0.0153*	*0.017*	*0.0252*
Imm.Coworker			-0.1398***	-0.2457***	-0.3797***
			0.0163	*0.0203*	*0.0276*
Imm.Coworker*Immigrant			0.09***	0.12***	0.1520***
			0.013	*0.011*	*0.02*
Imm.Coworker*Oimm				-0.0966***	-0.2925***
				0.036	*0.0734*
Imm.Coworker*Omix				-0.095**	-0.1359
				0.0582	*0.0513*
Imm.Coworker*Oimm*					0.6456***
Immigrant					*0.1285*
Imm.Coworker*Omix*					-0.3285
Immigrant					0.641
Constant	10.665***	10.7015***	10.686***	10.653***	10.615***
	0.26	*0.262*	*0.263*	*0.241*	*0.235*
Year dummies	yes	yes	yes	yes	yes
2-digit industry dummies	yes	yes	yes	yes	yes
Other controls	no	no	no	no	yes
R-Square Adjusted	0.25	0.27	0.28	0.31	0.35

Note: The number of observations includes 214,398 workers. Standard Errors are Huber-White robust standard errors, corrected by firm clustering. Reference group are full time native workers in native firms. (+) Neighborhood is defined as the contiguous counties to the county where the firm is located. Population in 100,000’s. (’) $Age * 10^2$.
***significant at 1%, ** significant at 5%, * significant at 10%.

produces interesting results. Immigrants earn more when working for immigrant employers and when the immigrant coworker share increases. The opposite is true for native workers. In general, a native worker receives higher wages if he or she works for a native firm with a low share of immigrant workers.

These results are striking in two senses. One, the ability to look at individual wages and identify the types of firm owners is only possible with this database. We have individual earnings for each firm. Although immigrants are paid less on average, they find themselves in a better position when working for immigrant firms. Second, we can look at the entire workforce and identify each individuals' types of coworkers in the firm. This allows us to make inference on the impact of social ties on worker wages.

2.7 Conclusions

This paper takes advantage of unique employee and employer matched micro-data from the U.S. Census Bureau to examine the effect of owner types and coworker types on firms' hiring patterns and workers' earnings. Particular attention was paid to the birthplace of employers and to the share of similar coworkers (by birthplace and ethnicity) at firms when new workers are hired. We examined the effect of those variables on hiring rates and on the wage differential between immigrants and natives.

In general, employees' wages are affected by the type of owner of the firm. For native employees working for immigrant owners the effect is very interesting. Natives are paid lower when working for immigrant employers, and in these firms natives have lower average earnings than immigrants. One explanation for these findings is that immigrant bosses have a better understanding of and networking with the immigrant community, and therefore can find and contract immigrant workers more easily than native-owned firms. Why can't native-owned firms quickly adjust and find this cheaper labor? Lack of language knowledge and lack of networking make it harder for native bosses to find immigrant workers. These findings justify further analysis of differences in contracting ability across employers. The evidence that the type of owner matters for wage differentials among workers also implies an important role for owner type on personnel policy.

In addition to examining the effect of owners and coworkers on differences between immigrants and natives, we evaluate the effect of owner and coworker types on ethnically(racially) different groups. An individual's race is an important source of variation across workers and owners. The evidence suggests that employers tend to hire workers from the same ethnic group. A significant impact of similar coworkers in the hiring process is observed across all types of owners, even after controlling for firm fixed effects. Immigrant owners tend to hire more Hispanics and Asians, while native owners hire more blacks and whites.

By shedding light on the ways workers and employers interact in the labor market to affect job and wage outcomes, this research makes a contribution to the sociology, labor economics, and demography literatures. It also opens up numerous avenues for future research. On the microeconomic side, we can further evaluate job flows and wage profiles of workers inside different types of firms. The analysis of assimilation can also take advantage of the results presented here, to further our understanding of the adjustment process of new immigrant workers. The empirical analysis in this paper makes some progress toward mitigating biases of skill sorting. This paper controls for a broad number of observable characteristics that try to capture other explanations for segregation. However, if owner unobservable characteristics are correlated with worker characteristics, the results of the analysis would be biased. Different empirical approaches such as instrumental variables or owner fixed-effects could be good options in future research, although this would demand a more exhaustive matched database that follows workers after leaving the firm and firms after ownership changes. Narrowing the scope of the analysis by looking at one industry could also provide information on the costs and benefits of firm recruitment processes. For instance, we could examine with more detail the effect of worker type concentration on firms' labor productivity. On an aggregate view, we can evaluate the effect of large flows of immigrants on the economy with the combined analysis of push and pull factors. Immigrant firms and immigrant workers seem to match quickly in the labor market. The analysis of the impact of immigration on unemployment and aggregate vacancies in the labor market can be extended to incorporate the findings in this paper.

Chapter 3

Workplace Concentration of Immigrants

3.1 Introduction[1]

Over the last several decades, labor markets in many U.S. cities have absorbed large inflows of new immigrants. The size of these flows has generated intense interest in their effects on the employment and wages of natives, as well as in the extent to which new immigrants have assimilated into the U.S. economy. New immigrants find employment and accumulate location-specific skills and work experience, gradually becoming integrated into local economies and potentially changing them in substantial ways. While outcomes of this process have been the subject of much research, less is known about the process itself. Which businesses hire immigrants? To what extent do immigrants work with natives? How does these patterns change as immigrants accumulate U.S. specific skills? Do the characteristics of different immigrant groups and different geographic labor markets affect the way in which assimilation plays out?

A lack of suitable data has limited economists' ability to address these questions. Our contribution is to bring to bear a rich set of matched employer-employee data that allows us to identify immigrants, their coworkers, and their employers. Our unique data permit quantifying the extent of and covariates of the workplace

[1]This chapter draws heavily on a joint paper with John Haltiwanger, Kristin McCue, Seth Sanders and Fredrik Andersson with the same title.

concentration of immigrants. The paper has two broad objectives. The first is primarily descriptive. The descriptive findings show that immigrants are much more likely to have immigrant coworkers than are natives. This pattern is driven partly by the geographic concentration of immigrants, but the patterns hold true even within local labor markets. At the same time, most immigrants do have native coworkers: only a small share work in immigrant-only workplaces. The concentration of immigrants is higher for recent immigrants and, conditional on recent arrival, for older immigrants: part of the assimilation process is a movement towards more interaction with natives in the workplace, and younger immigrants are more likely to work with natives. We find large differences associated with firm size: concentration is much higher in smaller firms, but is far from zero even in the largest firms. We also find substantial variation in the extent of immigrant concentration across industries even after controlling for a detailed set of location, employer and employee characteristics.

Second, our finding that the allocation of immigrants across workplaces is far from random raises the question: what does drive this workplace concentration? Both the existing literature and our descriptive findings suggest that it is important to consider how businesses hire their employees and the choices that businesses make about the skill mix of their workforce. One relevant issue here is the role that language skills play in governing interactions among employees and between employees and customers. A second issue is the role of social networks in the process that matches workers and firms. A third issue is human capital - the sorting and concentration of immigrants in the workplace may reflect sorting by skills. In the second part of the paper, we explore the role of these factors. We find evidence that immigrants with primarily immigrant coworkers are likely to have coworkers who live in the same residential tract. This pattern is robust to the inclusion of controls for other closely related factors such as residential segregation. We also find evidence that immigrant workers with poor English speaking ability and low education are more likely to work with immigrant coworkers.

The paper proceeds as follows. Section 3.2 provides an overview of the relevant theoretical and empirical literature that helps guide our empirical analysis. Section 3.3 describes the measurement of immigrant concentration, the matched employer-employee data we use in our analysis and the methods we use to explore the correlates of immigrant concentration. In section 3.4 we present our main results quantifying the extent and nature of immigrant concentration across workers and businesses. Section 3.5 analyzes the impact of factors such as social networks, language skills and human capital on the patterns of immigrant concentration. Most of the analysis focuses on native born, recent immigrants and established immigrants without specific reference to country of origin. Section 3.6 extends the analysis in terms of the basic patterns of concentration by country of origin. Concluding remarks are provided in section 3.7.

3.2 Background

3.2.1 Literature on earnings differences

Work examining earnings differences between whites and other groups in the U.S. has largely focused on netting out differences in skill (often captured by education and labor market experience) and geography (often using place of residence and urban residence) to assess the potential role of discrimination in labor market outcomes. This assumes that earnings differences are generated either by differing worker characteristics or differing returns to those characteristics. By extension, closing gaps in earnings requires equalizing worker characteristics and their return across groups. Differences in returns to characteristics are assumed to reflect unobserved ways in which the wage generating process differs and is typically viewed as an upper bound on the potential for discrimination to play a role in explaining wage disparities. A huge number of papers use this approach; some classic examples that examine earnings differences relative to white men are Smith and Welch [1977] for African American men, Borjas [1982] for Hispanic men, Chiswick [1983] for Asian men, and Corcoran et al. [1983] for women.

There is also a large literature assessing the sources of earnings differences between immigrants and native born workers (for example, Chiswick [1978], or Butcher and DiNardo [2002]). These papers generally augment the basic human capital framework used in the studies above by allowing for skill differences that are specifically relevant to immigrants. These include potential differences in the value of education and work experience accumulated outside the U.S., and the importance of differences in English language skills. Immigrant assimilation into the U.S. labor market is viewed as occurring through a narrowing of the earnings gap, resulting largely from increased U.S.-specific skills with time spent in the U.S. While there is debate over the speed at which the earnings gap between immigrant and native born workers closes, most studies find a substantial narrowing with time spent in the U.S. (see Chiswick [1978] and Borjas [1985]).

An older literature in sociology and economics stresses that earnings differences between groups may be driven by the characteristics of the firms that employ the majority and minority groups, rather than solely by human capital characteristics. Usually termed 'dual labor market theory,' this idea gained considerable attention in the late 1960s and early 1970s (see for example Averitt [1968] or Galbraith [1971]). According to this theory, many firms (especially industrial firms) are not governed by competitive processes. Instead, these firms enjoy market power. They insulate themselves and stabilize their workforce through job training and promotional ladders (Edwards [1972]). Firms that are constrained by competition do not invest in work skills and are characterized by low wages and high turnover, with low returns to human capital including job tenure.

The existence of 'good jobs' and 'bad jobs' by itself would not imply an earnings disadvantage to minority workers. Sociologists typically rely on a form of employer discrimination to explain why dual labor markets lead to minority disadvantage. Queuing theory suggests that good jobs always have an excess supply of applicants and firms then order workers by preferences and hire down the queue until vacancies are filled. If race or ethnicity plays a role in this ordering, a higher fraction of minority workers will be employed in the secondary market and have relatively low wages and wage growth.

While dual labor market theory per se has largely fallen out of the mainstream literature in economics and sociology, a newer literature that similarly argues that firm characteristics may be partially responsible for the level and growth in earnings of workers has gained growing acceptance. Wages appear to be positively correlated with firm productivity and firm size (Abowd et al. [2005]). While more controversial, there is some evidence that firm-level technological adoption also affects workers' wages (Dunne et al. [2004]). Lengermann [2002] finds that coworker characteristics, in addition to firm characteristics, may affect wages. Specifically, he finds that having more skilled coworkers independently raises a worker's wages. If firm characteristics play a major role in wage setting, then understanding how race and ethnicity affect the matching of workers to firms becomes important for understanding wage disparities across groups. Lengermann et al. [2004] explore the issues of sorting of immigrants across firms and find that sorting matters for wage differences between native born and immigrant workers.[2] We now turn to theories of worker segregation with special attention to how immigrants sort into firms.

3.2.2 Literature on segregation

Four broad overlapping theories explain segregation of workers into firms. These theories focus on sorting based on (a) productive characteristics, (b) pref-

[2]Some of our basic findings on immigrant concentration are also found in Lengermann et al. [2004]. Using the same data infrastructure that we use in this paper, they find for example differences in immigrant concentration by industry and employer size.

erences of workers or employers, (c) information available to workers or employers, or (d) cost of commuting to jobs. Some, but not all, of these theories imply that segregation results in a disadvantage for one group of workers relative to another.

There is substantial evidence of segregation by skill. For example, Kremer and Maskin [1996] look at the sorting of high and low skilled workers into firms over time and across three countries, the U.S., Britain and France. They find a high and rising correlation between worker skill levels in firms over the 1970s and 1980s. This may occur either because a firm demands a particular type of worker (for example skilled workers) or because coordination within a firm demands that workers share a common characteristic such as a common language. Cabrales et al. [2008] emphasize a different skill-based mechanism: if a worker's utility is a function of both absolute wages and their wages relative to those of coworkers, and if movement of workers across firms is costless, complete segregation of workers by skill is optimal. A mixed-skill workforce generates wage inequality within a firm, reducing worker utility. All workers are made better off by grouping workers with similar skills and avoiding these reference group costs. Regardless of the mechanism, segregation by skill will cause immigrant-native differences in the distribution of skill to contribute to segregation. For example, immigrants are both much more likely than natives to have an 8th grade education or less (23% vs. 5.2% for natives in the 2000 census), and also more likely to have an advanced degree (10.3% vs. 8.6% for natives). Therefore, firms that specialize in hiring exclusively low-skilled or exclusively high-skilled workers will tend to have a workforce that has a higher fraction of immigrants than the fraction in the population.

Language differences provide another productivity-based motivation for segregation. If working with someone who does not speak the same language generates transaction costs, employers may increase productivity by hiring only workers who share a common language. In this case, immigrants from non-English speaking countries may be particularly likely to be segregated, and may also be particularly likely to work with their compatriots rather than other immigrants. Lang [1986] develops a formal model of wage differences arising because of the costs to firms of having to pay a premium for bilingual workers who can bridge the language barrier. One of the results of this model is that complete segregation would occur if both capital and labor were owned by each language group. Hellerstein and Neumark [2003] find evidence that Hispanics with poor English-language skills are particularly likely to work with other Hispanics. Their data do not allow them to examine how much of this is due to Hispanic workers working for Hispanic-owned firms as in the Lang model.

Becker [1957] is the classic model of preference-based segregation. In this model, segregation of workers by race occurs as the result of discriminatory preferences on the part of co-workers. White workers would demand a premium to work with black workers. In response, firms segregate workers into separate facilities, avoiding the need to pay a wage premium to discriminating white workers. Depending on conditions including the relative size of the minority and majority group, the number of firms, and returns to scale in production, segregation may be extreme but with limited disadvantage in wages to the minority group. Dual labor market theory, described above, also generates wage differences across groups if discriminating employers put minority job candidates lower down the queue. In this case, higher wages in the primary sector ensure that a higher fraction of the majority group works in the primary sector and hence gives a wage advantage to the majority group.

Information-based theories concentrate on the mechanisms that workers use to find jobs. For example, firm use of employee referrals to fill jobs may contribute to workplace segregation. For workers, use of personal contacts to search for jobs is inexpensive and has relatively high rates of success (Holzer [1988]). For employers, employee referrals provide both a low cost recruitment strategy and, on average, new hires with higher productivity and lower turnover rates (Holzer [1987]; Montgomery [1991]). If workers tend to refer others who have similar characteristics, use of referrals can increase rates of segregation. Elliot [2001] finds that recent Latino immigrants are more likely than blacks or Latino natives to use personal contacts to find jobs. Weak English skills explain much of this difference. A greater reliance on referrals in small workplaces in combination with a concentration of recent immigrants in small firms also contributes to the difference.

Information flows may combine with residential segregation to contribute to workplace segregation. Neighborhoods play an important role in who you know and hence may provide important job contacts and references. Several papers have established that workers in the same firm are disproportionately from the same neighborhoods. Using data from Boston, Bayer et al. [2008] find that a worker is about one-third more likely to work with someone who lives in the same census block as to work with someone who lives in other blocks in their block group (typically eight or so contiguous blocks). This comparison of blocks to block groups provides important evidence that having coworkers who are neighbors does not stem from unobserved factors such as transportation routes or distance that make a place of employment a natural place to work for those living in a particular location. Many of these unobserved factors would be similar for a block group and block of residence, and so should have similar effects on the likelihood of working with more or less immediate neighbors. This paper is limited in that the exact establishment can not be observed, while sample sizes as well as the ethnic make-up of Boston restrict the authors' investigation to black-white differences.

Hellerstein et al. [2008a] also present evidence of neighborhood network effects. Using matched employer-employee data, they compare how likely an individual is to work in the same establishment as his neighbor, relative to the likelihood that this would result if their employer hired workers randomly from the geographic areas of residence of all individuals who work in the employer's census tract. Their dataset is large enough to disaggregate the analysis for whites, blacks and Hispanics. They find that another worker living in the same census tract has twice the probability of working in your firm than what one would expect from randomness. They do not investigate the importance of other mechanisms for sorting workers into firms.

A final theory of the sorting of workers into firms also works through residential segregation but focuses on the fact that not all jobs are equally accessible from different places of residence. Kain [1968] investigated employment patterns of blacks and whites in Chicago and Detroit. He found that blacks were unlikely to be employed in areas that were predominantly white, that blacks would have higher employment rates if housing segregation was lower, and that the movement of jobs from central cities to suburban areas depressed the employment prospects of blacks. A number of other studies followed that compared employment differences between central city and suburban residents within an urban area. These tests often found employment prospects lower for central city residents, but controlling for unmeasured skill differences between residents of different locations remained an issue in inference. A recent study by Hellerstein et al. [2008b] questions the interpretation that a lack of jobs near where blacks live is a major source of racial employment differences. They find that the employment prospects of black residents are positively correlated with the number of nearby jobs in which blacks work, but not with the number of nearby jobs in which whites work. This indicates that even within close geographic proximity, job markets are racially segregated. They conclude that spatial mismatch has little effect on employment prospects of blacks but that what they term racial mismatch—few nearby jobs that employ blacks—has a large effect.

Clearly, residential segregation could contribute to workplace segregation of immigrants. There is ample evidence that immigrants' places of residence are spatially concentrated. Iceland [2009] describes the high level of residential segregation in the U.S. among immigrant groups but also shows that immigrants migrate to neighborhoods that are more ethnically integrated as they spend more time in the U.S. However, Porter and Wilson [1980] argue that, unlike for black Americans, residential segregation may aid immigrants—especially new immigrants–while also leading to segregation of workers in firms. Studying the post-Castro immigration from Cuba to Miami, Portes and Wilson show that not only do Cubans in the U.S.

work together, many work in firms owned by other Cubans. Moreover, Cuban employees of Cuban-owned firms tended to display the same patterns of wage growth and returns to human capital as workers in firms classified as in the 'primary sector,'defined as firms with a promotion ladder, over 1000 workers, and high average wages. While an impressive source of employment, it is not clear that the example of Cubans generalizes to other foreign-born groups. Capital owners specifically were forced to leave Cuba, which may have led to higher levels of capital with which to start businesses and more experience with small businesses among Cubans than among other foreign born groups. Having said this, Wilson and Portes report that much of the capital used to start these businesses was accumulated in the U.S. and not transferred from Cuban concerns.

3.3 Methodology and Data

3.3.1 Measuring immigrant concentration

We follow several recent papers that study workplace segregation (Hellerstein and Neumark [2007]; Aslund and Skans [2005a], Aslund and Skans [2005b]—henceforth HN and AS) by using the share of coworkers in a particular group as a measure of exposure. That is, we exclude the worker himself when measuring the concentration of immigrants in the business he works in. For worker i, employed by business j which has s_j employees, the share of immigrants among coworkers is:

$$C_{ij} = \frac{1}{s_j - 1} \sum_{k \neq i}^{s_j} I_k \tag{3.1}$$

where I_k is an indicator for whether or not worker k is an immigrant. For the sake of brevity, we will refer to this simply as the coworker share. As pointed out by these authors, excluding the worker's own characteristic in calculating concentration ensures that in large samples the coworker share for both immigrants and natives should on average equal the share of immigrants in the workforce in the absence of any systematic concentration. Based on this, we use the difference between the mean coworker share for immigrants and natives as a measure of immigrant concentration. A positive value indicates that immigrants are more concentrated than would be expected based on random allocation. At the extreme, if immigrants worked only with immigrants and natives with natives, the difference in coworker means would equal one. A negative value for this difference would indicate that immigrants were more likely to work with natives than would be expected based on random allocation—a pattern that could arise where the two groups provide different but complementary skills.

We depart from the approach of these authors in two ways: in the way in which we condition on observable characteristics, and in choice of a normalization to gauge whether the concentration we find is large relative to some alternative. There are two types of questions that can be addressed by conditioning on observable characteristics in studying segregation: to what extent can segregation be explained by differences in the characteristics of the two groups, and which characteristics are most associated with segregation. HN and AS both focus more on the first issue, while we explore some aspects of both questions. As an example to provide some context, the immigrant and native education distributions differ, and particular employers may hire primarily from one part of the education distribution, leading to concentration of immigrants because of differences in skill. HN and AS both use the difference between measured concentration and the amount of concentration that would be generated solely by the way in which education is distributed across employers as their conditional measure of concentration. In contrast, we condition on a worker's own characteristics and on the characteristics of his or her employer (e.g. employer size and industry), but do not directly condition on coworker characteristics. Our measure of concentration is the mean difference between immigrants and natives with the same characteristics.

We take a different approach in part because the worker characteristics in our data that vary within employer (age, gender) do not differ dramatically between immigrants and natives, and they also turn out not to have a strong correlation with immigrant concentration. Controlling for a worker's own characteristics should remove the effects of age and gender from the measured difference in coworker mean, and the estimated coefficients allow us to examine the characteristics of immigrants and natives who work in heavily immigrant workplaces.

Both HN and AS normalize their measures of concentration, though they choose different references for the normalization. While both of their normalizations have intuitive appeal, we take a different approach. We use the immigrant-native difference in coworker shares as our measure of concentration, but in most cases also present information on the coworker share for natives as a point of reference. Our regression approach makes doing so straightforward, and also allows us to more directly illustrate patterns of concentration. For example, using the regressions to predict means for a given set of covariates allows us to illustrate the strong positive relationship between immigrant concentration and immigrant share of the workforce, when looking across groups defined by characteristics such as area of residence and employer size. In addition, the regression approach using our coworker index at the person level as the dependent variable permits us to normalize our measure of concentration effectively along a number of dimensions. For example, HN normalize to control for between MSA differences in various groups (e.g., differences in the distribution of blacks and whites across MSAs). We control for such differences directly in our regression approach by, for example, including controls for MSAs.

3.3.2 Data

We use the data from the Longitudinal Employer - Household Dynamics (LEHD) database, which draws much of its data from complete sets of unemployment insurance (UI) earnings records for a subset of U.S. states. The database includes records for 1990 to 2004, though some states only have data for a subset of those years. The workers' earnings records have also been matched to characteristics of their employer gathered in quarterly administrative reports and through Census Bureau business censuses and surveys. Basic demographic data are also available for workers, including place of birth. For those born outside the U.S. (and its territories), we treat the year in which they first applied for a Social Security Number (SSN) as the date of their arrival. While this may not precisely date arrival, preliminary results based on a sample of immigrants for whom both LEHD and decennial population census data are available suggest that the year the individual first applied for a social security number proxies the reported year of arrival fairly well.[3] In the current analysis, we use data from selected metropolitan areas in 11 states. While we do not use a large number of states, our sample does include five of the six states that had immigrant populations of 1 million or more.

These data give us two unique advantages. First, we have earnings for a group large enough to include millions of immigrants. Second, we can observe the firms in which workers are employed, allowing us to measure both employer characteristics and the characteristics of coworkers. These data have other advantages that we do not exploit here but plan to in future work: for example, the data can be used to generate a panle on both employers and employees that would allow us to track earnings of immigrants over time in the U.S. as well as to observe contemporaneous changes for native-born workers. The main disadvantage of these data for studying immigration is that they include only on-the-books employees and so do not cover the self-employed or those working in the informal sector. Thus they likely have poor coverage of undocumented immigrants. Coverage of employment in agriculture is incomplete in the LEHD data, so we exclude employers in that sector.

Calculating the share of coworkers who are immigrants requires at least one

[3] Here we use year of arrival only to split immigrants between those arriving very recently (within the last 5 years) and other immigrants. Comparing our classification based on date of SSN application to one based on responses in the 2000 census, 92% of immigrants are classified in the same way according to both sources. Among those for whom the classification differs, the most common pattern is that 4% of Mexicans are considered new immigrants in Decennial Census versus 10% in LEHD. The lag in the registration process by immigrants, specially in the case of Mexicans, explains these differences. The patterns by age are very similar between LEHD and Decennial Census, however younger immigrant workers are also reporting a small lag in their application for social security.

coworker, so we restrict our sample to businesses with at least two employees.[4] We measure concentration using a cross-section of data based primarily on the second quarter of 2000, but we use LEHD data for the 1995-2000 period to define business age. In computing the coworker share, we use all coworkers, whether or not they hold other jobs. However, the set of observations used in our regressions includes only one job for each individual: the job where they received their highest earnings in that quarter.

We draw data from employers in 31 MSAs. We include all MSAs that have substantial foreign-born populations and are in states for which we have the required data, but we also included several smaller MSAs that experienced very rapid growth in foreign-born residents between 1990 and 2000.[5] Even in the smallest of our MSAs we have data on more than 30,000 immigrant workers, so small sample sizes are never an issue.

Table 3.1 summarizes the across-MSA variation in immigrant shares for our sample of MSAs. In the average MSA in our dataset, 18.9% of workers are immigrants. In what follows, we are interested in deviations in workplace shares from the overall-average. Clearly the substantial variation in immigrant share across MSAs will contribute to finding immigrant concentration. The shares of both recent and

---

[4]Immigrants account for 27% of employment in single-employee businesses, and 16% of employment in businesses with more than one employee.

[5]More precisely, we started from the list of MSAs used in Singer [2004], which included all MSAs with at least 1 million residents in 2000, and meeting at least one of the following criterion: (i) at least 200,000 foreign-born residents, (ii) a foreign-born share higher than the 2000 national average (11.1%), (iii) 1990-2000 growth rate of the foreign-born population above the national growth rate (57.4%), or (iv) above national average percentage foreign-born in 1900-1930 (''former gateways"). We drop 14 of Singer's 45 MSAs because we do not currently have access to all of the data we need from the relevant states.

Table 3.1: Variation in Immigrant Share of Workforce across Sample MSAs

	Percent Immigrant		
	Total	**Recent**	**Established**
Mean	18.86	3.40	15.46
Standard Deviation	10.27	1.85	8.57
P25	10.57	1.94	8.52
Median	16.26	2.92	13.54
P75	26.60	4.37	22.82
P90	32.58	6.03	27.23

Source: Authors calculations based on LEHD UI-ES202 database.
Note: Unit of observation is an MSA. Immigrant shares are measured as of the second quarter of 2000, and recent immigrants are those arriving between 1995 and 2000. The table presents fuzzed percentile values.

established immigrants vary substantially across MSAs as well.

For roughly 10% of workers in our sample, we match in additional information on educational attainment and English language skills from the long form of the 2000 population census. Using propensity score models, we develop weights for the matched sample that allow us to closely replicate our results based on the overall sample.[6] We then use weighted estimation with the matched sample to examine the relationship between these measures of skill and immigrant concentration.

3.3.3 Regression specifications

Our primary empirical approach is to run a series of regressions with the coworker share as the dependent variable, and individual workers on their primary job as the unit of analysis. As a rough way to capture the way in which immigrant

[6]The variables used in the propensity score procedure were: worker age, sex, country of origin (11 groups=Mexico, China, Cuba, El Salvador, India, Korea, Japan, Vietnam, Phillipines, other country of origin groups, and natives), log earnings, worker status, industry (4 digits), Msa indicator variables and population density, plant age and size, and firm's # of establishments.

concentration changes with time spent in the U.S., we include indicators for whether an individual is a recent immigrant (RI, defined as arriving in the last 5 years), or a more established immigrant (EI, arriving more than 5 years ago). Since we use a cross-section of data, the differences between recent and more established immigrants confound the effects of time spent in the U.S. with changes in labor markets and in immigrant and native characteristics over time. We would need to exploit the panel aspect of our database to seriously address the affects of assimilation, but believe this is useful as a starting point that provides suggestive evidence on whether assimilation effects on concentration are likely to be important.

Our initial regression specification is:

$$C_{ij} = \gamma_N + \gamma_{EI} EI_i + \gamma_{RI} RI_i + \beta x_{ij} + \epsilon_{ij} \tag{3.2}$$

where (again) i denotes an individual and j denotes a workplace. Here, the constant term (γ_N) represents the mean coworker share for the omitted category, which in our simplest specification consists simply of natives.

Coefficients γ_{EI} and γ_{RI} give us estimates of the differences between immigrants and natives in how likely they are to have immigrant coworkers. We use controls for MSA and for various worker and employer characteristics to examine the extent to which immigrant concentration can be accounted for by differences between natives and immigrants in their geographic distribution and in worker and job characteristics. In section 3.6, we define coworker shares for specific countries of origin and look at which immigrants are most likely to work together.

Specification (3.2) assumes that the effects of covariates are the same for immigrants and natives. To examine whether this in fact holds, we use an alternative specification that includes interactions between our immigrant dummy variables and other covariates:

$$C_{ij} = \gamma_N + \gamma_{EI} EI_i + \gamma_{RI} RI_i + \beta x_{ij} + \phi_{EI} EI_i * x_{ij} + \phi_{RI} RI_i * x_{ij} + \epsilon \tag{3.3}$$

Once we add interaction terms, the intercept rarely identifies effects for a group of particular interest. To illustrate the effects of a particular covariate in specifications of form 3.3, we present predicted means for immigrants and natives, by which we evaluate differences between immigrants and natives based on the pooled distribution of the variables in x.

To ease computations with our 36 million records, we use linear regression models rather than adopting an approach that accounts for the limited range of the dependent variable. In this draft, we also ignore the effect of clustering within employer in estimating the standard errors. For most of our specifications, the dependent variable mean is not close to either 0 or 1, which mitigates some of the problems inherent in the linear model. The strong positive correlation in the coworker share among employees of the same business will lead to a downward bias in our estimated standard errors in all worker-level regressions. Given the huge size of our sample, the results we present would generally remain significant at standard levels even if the corrected standard errors were 100 times larger. The few exceptions (in Table 3.6) are estimates that are too small to be meaningfully different from zero anyway. [7]

3.3.4 Descriptive statistics

Table 3.2 presents summary statistics for immigrant and native workers in our full sample. The first row gives coworker shares for the three groups. For the average native, about 15% of coworkers are immigrants, while 42% of the coworkers of recent immigrants are fellow immigrants, and 36% of the coworkers of established immigrants are immigrants. The immigrant-native difference in coworker means—our measure of concentration—is .272 for recent immigrants and .214 for more established immigrants, indicating substantial concentration.

Table 3.2: Characteristics of Immigrant and Native Workers, Full Sample

	Immigrants		
	Recent	**Established**	**Natives**
Coworker share	42.1	36.3	14.9
Age			
Age<30	43.6	19.7	29.3
30<Age<40	35.6	33.2	30.0
Age>40	20.8	47.0	40.7
Male	56.8	56.4	51.7
Age at arrival (*)			
<12	1.1	14.7	.
12-25	36.2	49.6	.
26-35	37.0	24.8	.
35+	25.7	10.9	.
Establishment size			
2-9	8.5	9.0	8.0
10-49	23.6	22.6	23.5

Continued on next page.

[7] So far, using statistical software to handle clustering does not seem feasible. However, we could put an upper bound on the standard errors by summarizing data at the establishment level for immigrants and natives, and then running our regressions weighted by employment and clustering on establishment. This reduces the number of records to less than two times our number of establishments and cluster size to at most 2.

Table 3.2: Characteristics of Immigrant and Native Workers, Full Sample(continued)

	Immigrants		
	Recent	**Established**	**Natives**
50-99	14.4	13.3	13.6
100-499	31.4	30.9	29.6
500 or more	22.2	24.2	25.3
Firm has multiple establishments	31.6	34.3	41.4
Establishment age			
0-1	12.3	10.7	10.8
2-4	25.6	22.0	23.6
Age 5 or more	62.2	67.3	65.7
Sector			
Construction	4.7	5.5	5.9
Manufacturing	18.9	21.3	12.8
Transportation & utilities	3.7	5.2	6.5
Wholesale	6.8	7.0	6.5
Retail	22.7	18.3	21.4
FIRE	3.2	5.6	7.2
Services	40.0	37.1	39.7
Log quarterly earnings	8.20	8.52	8.39
Consecutive quarters on 2000-Q2 job			
Quarter before AND after	58.7	70.9	65.6
Quarter before OR after (not both)	32.9	23.4	26.6
Neither quarter before NOR after	8.5	5.7	7.8
Immig. share of wkrs in residence tract	36.5	35.2	15.2
Neighborhood network index	2.10	1.78	1.70
Shared commute index:			
Immigrant co-commuters	0.14	0.11	0.05
Native co-commuters	0.24	0.23	0.48

Source: LEHD database and author calculations.
(*) Year of application for a SSN is used as a proxy for time of arrival in the U.S.
Note: The unit of observation is a worker. Employer characteristics and earnings are for the first quarter 2000 job with the highest earnings. All figures except for log earnings represent percentages. There are 35,966,450 workers in total for our 31 MSAs.

The following rows give demographic information for each group. Recent immigrants are substantially younger than natives while earlier immigrants are older. Combining the two, immigrants are slightly older than natives in our sample. Men substantially outnumber women among both recent and established immigrants, while among natives men are more narrowly in the majority. Differences between immigrant and native women in rates of labor force participation likely contribute to these gaps. Overall, more established immigrants are also more likely to have entered the country before reaching prime working age. This is in part by definition, as immigrants who are old enough to work and arrived in the country within the last five years could not have arrived when they were young children.

Comparing information on employer characteristics, immigrants are more concentrated in manufacturing than are natives, but generally the differences by sector are not particularly large.[8] The size distributions look surprisingly similar. Immigrants are more likely to work in the smallest firms, and less likely to work in the largest, but overall the differences are small. Our restriction to employers and on-the-books employment may affect these statistics as well. Similarly, differences in distribution of employment by plant age are also small. However, immigrants are less likely than natives to work for multi-unit firms.

The statistics on earnings show rather large differences, with recent immigrants having lower earnings than natives, while established immigrants have higher earnings. Differences in job tenure likely explain some of this pattern, as established immigrants are least likely to be in a job that lasts only one quarter, and recent immigrants are most likely to be in such a transitory job.

We construct three additional measures from our data base using information on worker tract of employment and tract of residence, which we use in section 3.5 to

[8] Comparing our estimates to published 2000 population census estimates is inexact for several reasons: our analysis includes only a subset of MSAs; our sectors are defined based on SIC codes while the 2000 industry codes are NAICS based; and we exclude the self-employed and those working off the books, both of which may be included in household estimates of employment. But for reference purposes, in the 2000 decennial census 17% of immigrants and 14% of natives worked in manufacturing, while 8% of immigrants worked in construction compared to 7% of natives Bureau [2005].

explore the relationship between workplace concentration and social networks. The first of these is simply the share of immigrants among people in a worker's tract of residence which we use to capture the degree of residential segregation.[9] Because we only have data on those who work, we use the share of immigrants among resident workers in a particular tract rather than the share among all residents. As can be seen in Table 3.2, immigrants are more likely to live with other immigrants than are natives, but there is little difference in residential segregation between recent and earlier immigrants. On average, both groups live in tracts that are majority non-immigrant.

As a proxy for social networks, we calculate for each worker the fraction of their coworkers who also reside in the worker's tract of residence. This proxy, which we refer to as a neighborhood network index, may reflect many factors. For example, as discussed in Section 3.2, referrals by current employees may be an important recruitment source, and many referrals may come about through contacts with neighbors. If neighborhood referrals are important we would expect to find people who work together also living close together. Our network variable will, in principle, capture such effects but should more generally be viewed as capturing the extent to which residential location and employment location are correlated. It is for this reason that when we explore the role of this variable in the regression analysis that follows, we include both the residential segregation variable above and the shared commute

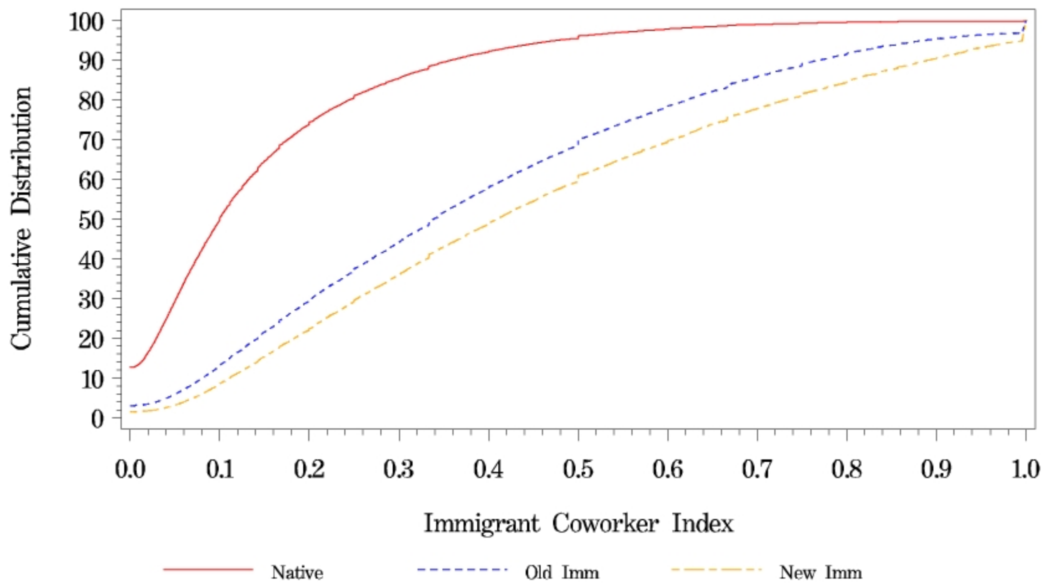
[9]Census tracts are small geographical areas with a population between 1,500 and 8,000 individuals. They are designed to be relatively homogeneous with respect to socio-economic characteristics. As such, they are arguably well-suited to serve as a proxy for the geographical reach of a social network, i.e. the limited distance between residents of a census tract, both in terms of geography and socio-economic factors, suggests that the within-area likelihood of interactions between members is high relative to the between areas.

variable described below as controls. Not surprisingly the mean of the network index is small: for the average worker the fraction of coworkers who live in the same tract is about 1.7 percent. This fraction is substantially higher for small businesses, and it falls systematically with employer size. It is also instructive to observe that while the average is small there is considerable variation across workers and it is the latter variation we exploit in our analysis.

A related but distinct relationship between workplace and residence is that proximity or transportation infrastructure may mean that employees working at a particular location are likely to live in a particular set of neighborhoods. To distinguish this effect from the network measure, we construct an additional variable for each worker, namely the share of employees at other businesses located close to his employer (defined as other employers in the same tract) who also live in the worker's residential tract (as in Hellerstein et al. [2008a]). This is a rich measure of the propensity for workplace and residence to be connected that will control for factors such as commuting patterns and even the extent to which the connection between workplace and residence might reflect sorting across workplace and residence locations by skill. We refer to this as a shared commute index, which we split between immigrant co-commuters and native co-commuters. That is, the denominator for both components of the shared-commute index is the number of employees working for other employers in a worker's tract of employment.[10] The numerator for the immigrant co-commuter variable is the number among that group who are

[10] In our sample, there are on average 49.19 employers per tract (excluding tracts that are strictly residential). 7% of tracts with employment have only one employer, and for those tracts, the shared commute variables are zero. Only 9% of workers in our sample work in single-employer tracts.

Figure 3.1: Cumulative Distribution of Coworker Share by Worker Type



Source= LEHD database year 2000 second quarter

immigrants and who live in the worker's residential tract. These shares are quite small, but differ between immigrants and natives.

Figure 3.1 provides some basic information on the distribution of our dependent variable. The three lines plot the cumulative distribution of immigrant coworker share for natives and for recent and more established immigrants. About 13% of natives work in native-only workplaces (having coworker immigrant share=0) in our sample of immigrant-rich MSAs, but the share of immigrant employment in immigrant-only businesses is surprisingly small (2.8% of immigrants). In this set of MSAs, about 10% of the median native's coworkers are immigrants, while for established immigrants the share at the median is about 34%, and for recent immigrants, the share is about 41%.

3.4 Accounting for immigrant concentration

We carry out two sets of exercises to examine the degree and nature of immigrant concentration. First, we address the extent to which observable factors can account for immigrant concentration using a series of regressions with the coworker share as the dependent variable based on specification (3.2). Second we apply specification (3.3) in which we add interactions between the immigrant dummy variables and our explanatory variables. In doing so, we let the difference between coworker shares for immigrants and natives vary with observable characteristics, which allows us to determine in what sort of workplaces and for which kinds of workers we see the most concentration.

Table 3.3 presents estimates of the key parameters from the first set of regressions. The first two columns present estimates of the coefficients on the dummy variables identifying our two immigrant groups—recent immigrants, defined as those who arrived between 1995 and 2000, and more established immigrants who entered before 1995.

In the first row of Table 3.3, we report results from the base specification without any controls. This simply reproduces the differences in means one finds from the first row of 3.2. The subsequent rows of Table 3.3 show the effects of adding each of the sets of controls. We include MSA dummies in all but the first row, but add the other controls one set at a time. Note that in doing this we are allowing the immigrant share of employment to vary with the controls, but assuming that within-cell immigrant concentration is the same for all control categories. Our intent here

Table 3.3: Contribution of Covariates to Immigrant Concentration (Full Sample)

Covariates	Recent immigrant	Established immigrant	R-squared
Full sample			
No covariates	0.272	0.214	0.198
MSA dummies	0.224	0.156	0.379
MSA+:			
Worker age	0.225	0.155	0.379
Worker sex	0.224	0.156	0.379
Employer size	0.224	0.156	0.380
Employer age	0.225	0.156	0.379
Employer age * Multi-unit	0.221	0.154	0.387
Industry detail	0.195	0.133	0.460
Size and industry	0.195	0.135	0.461
Log earnings and full-quarter controls	0.223	0.115	0.379
Neighborhood network index	0.221	0.155	0.381
Shared commute index variables	0.222	0.156	0.379
Immigrant share in residential tract	0.181	0.132	0.471
All of the above	0.143	0.089	0.495

Notes: Figures in the first two columns give the predicted difference in mean coworker share between the immigrant group and natives. As a point of reference, the mean coworker share for natives in the first line is .149 (as in Table 3.2). It is also .149 for all other specifications if evaluated at the native mean for all included covariates, but somewhat higher if evaluated at the pooled sample mean. The unit of observation is a worker. N=35,966,450 for the full sample. The variables are as described in Table 3.2, except that we use 185 detailed industry categories in place of sector. All standard errors are less than 0.0001.

is to determine whether any of the employer or worker characteristics available to us identify cells with a large share of immigrant employment, but within which the immigrant-native differences are significantly smaller than the overall difference. For example, if immigrants were mostly employed in a few industries, but were randomly distributed across workplaces within industry, industry controls would reduce the concentration coefficients to zero because there is no concentration within industry. If that were the case, then explaining immigrant concentration would boil down to explaining why immigrants worked in different industries than natives.

In broad terms, Table 3.3 shows that our measures of employer and worker characteristics account for a substantial amount of the observed concentration of immigrants in the workplace, but about half of the concentration remains unexplained. Differences in the immigrant share of employment across MSAs and detailed industries account for roughly one-third of total concentration. The other control with substantial explanatory power is the share of immigrants in a worker's residential tract. Living in an immigrant-rich residential tract is positively correlated with the share of coworkers who are immigrants, so the difference between immigrants and natives in the likelihood of working with immigrants is substantially smaller for those living in neighborhoods with similar immigrant shares than for immigrants and natives overall. In the last row, we include all of our controls but still find that, compared to natives, the difference in the share of coworkers who are immigrant is 14% for recent immigrants and 9% for established immigrants.

In the subsections that follow, we discuss the results of adding particular controls in Table 3.3 along with results from our second exercise based on (3.3), in which we add interactions with the immigrant dummy variables. Because the patterns identified by the interaction terms are easier to grasp visually, we present the findings from this exercise through graphs of predicted coworker shares. We do note that the R-squared for this full model with interactions is 0.567.

3.4.1 Location

Covariates have the greatest potential to account for differences in coworker means when their distribution differs substantially between immigrants and natives. Geography is one dimension along which there are substantial differences. In this section we look only at differences across metropolitan areas, but in section 3.5 we explore how differences in location within MSAs may also contribute to immigrant concentration.

Immigrants are much more likely than natives to live in the largest metropolitan areas in the U.S. For example, in 2000 55% of immigrants lived in the 9 metro areas having a population of at least 5 million, compared to 27% of natives. While 21% of natives lived in nonmetropolitan areas, only 3% of immigrants did (Schmidley [2001]). Even if immigrants were randomly sorted into jobs within their local labor markets, the fact that many natives live in areas with few immigrants would lead us to find substantial concentration in coworker means for the nation as a whole. By restricting our sample to urban areas that have many immigrants, we increase the overall share of immigrant coworkers above the national average. At the same time, we reduce the difference between immigrants and natives by excluding areas where natives work with few immigrants. But substantial variation in the immigrant share of employment remains across our sample MSAs, as illustrated in Table 3.1 above.

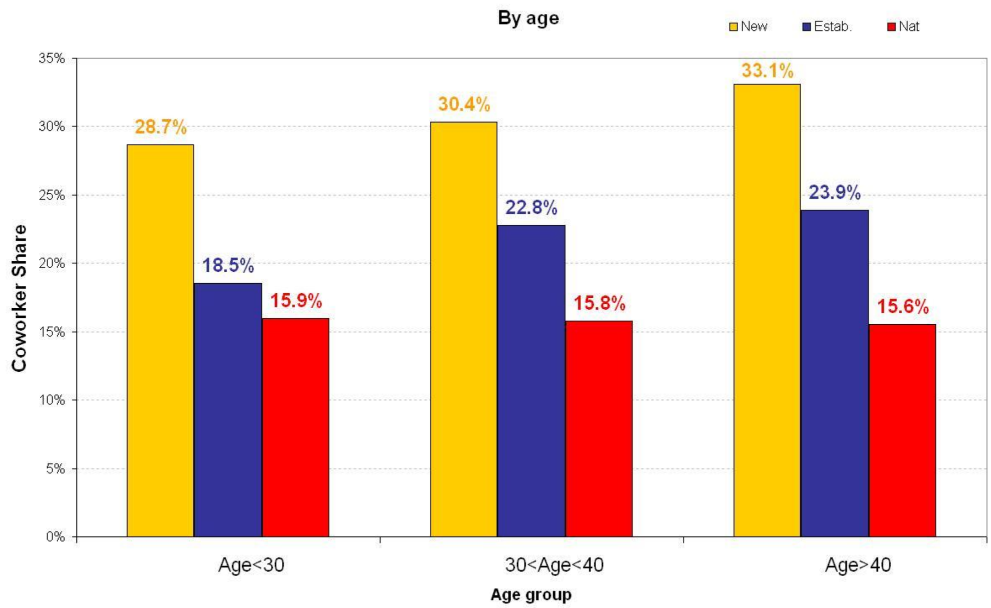
In Table 3.3, including MSA dummies almost doubles the R-squared. The reduction in the recent immigrant coefficient between row (1) and row (2) indicates that roughly one-fifth of the overall difference between recent immigrants and natives simply reflects differences in their geographic distribution: unsurprisingly, the metropolitan areas in which immigrants work have higher immigrant shares than the areas in which the average native works. Similarly, about one-quarter of the native/immigrant differential for established immigrants is due to differences in which cities they live in. When we allow immigrant concentration to differ across MSAs (results not reported), we find that immigrant concentration rises very consistently with the overall immigrant share in an MSA, and that concentration is consistently higher for recent immigrants than for more established immigrants.

3.4.2 Worker Demographics

We have limited data on the characteristics of workers—basically age and gender, in addition to knowing the country in which a worker was born. As the third and fourth rows of Table 3.3 illustrate, adding age and gender to the specification with MSA dummy variables has essentially no effect. Allowing the effects of these variables to differ between natives and immigrants shows that age does have a weak association with immigrant concentration, though gender does not. As Figure 3.2 illustrates, older immigrants are somewhat more likely to work with other immigrants than are younger immigrants, but there is little difference by age for natives. Note that because recent immigrants have by definition arrived within the preceding 5 years, age and age at arrival are necessarily highly correlated for that group, so what we observe are the combined effects. We would need to move beyond the cross section we are using here to disentangle their separate effects for recent immigrants.

We use similar methods for most of the following bar charts, so it is useful to clarify how the estimates were constructed for the first chart. The coworker shares here are based on regression estimates from specifications that include all of the sets of variables listed in Table 3.3. We constructed the estimates presented in the figures using the pooled mean values of all controls except for those used in defining the categories for the bars. So for Figure 3.2 the pooled mean values for all variables except age are used to get predicted values for each age and immigrant status group. The age dummy values are set according to the labels on each of the three clusters of bars. The differences between bars for a given age group are determined by the coefficients on the immigrant group dummy variables and by the product of the interaction effects for the group with the pooled mean the other controls. The age interaction terms determine how much the bars vary across age categories for a given group (i.e. natives, recent, or established immigrants) while the pooled means and coefficients for other variables determine the average level of the bars for a group.

Figure 3.2: Coworker share by age of employee



Note: Size, Sector, plant age, sex, units and MSA groups use total population distribution. Using full two-way interactions with individual status.

3.4.3 Employer characteristics

We have a rich set of employer characteristics in our data. Most of the measures we use are defined for an establishment (or business location). The measures include establishment size (measured by employment), detailed industry and detailed location.[11]

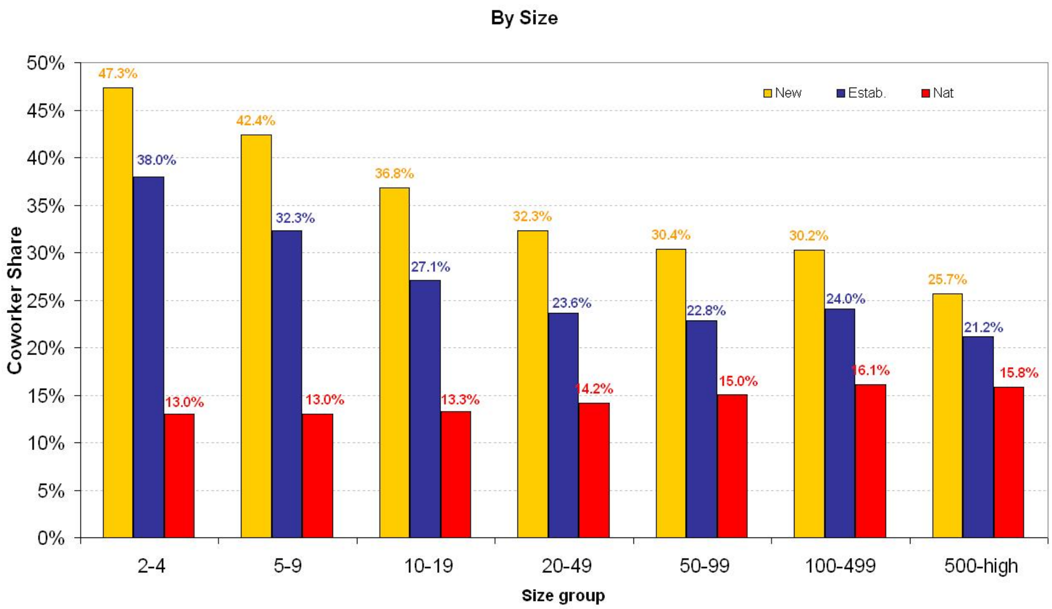
[11]There are some technical issues in assigning workers to multi-unit establishments in the LEHD data. The UI wage records at the person-level include state-specific employer identifiers which identify the firm that a worker is employed by. The UI wage records link to ES-202 records filed by the firm that provide employment, payroll, industry, and location information for each of the firm's establishments in that state. LEHD has developed algorithms for assigning workers to multi-unit establishments which multiply impute an establishment identifier to affected workers based on the worker's place of residence; the locations, sizes, and ages of the employing firm's establishments; and the timing of the worker's employment. Once a worker is assigned to a specific establishment while working for a given employer, the worker remains with that establishment as long as the worker remains employed with that employer. We weight each implicate based on the estimated probability of being employed at that establishment. More details are available in Abowd et al. [2006].

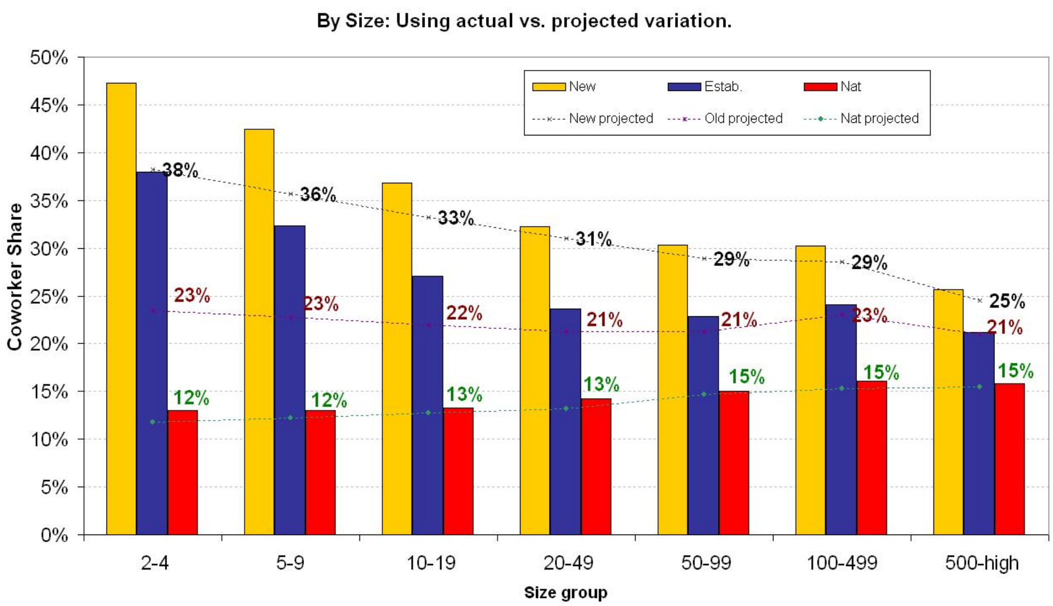
3.4.3.1 Employer size

We classify employer size into the size bins depicted in Figure 3.3. Recall that we are excluding establishments with only one worker since the coworker index is by construction not defined for a worker who has no coworkers, so the bins begin with size 2. When we constrain the size effects to be the same for immigrants and natives, adjusting for immigrant/native differences in employer size has virtually no effect on the difference in average coworker share (see Table 3.3). This is unsurprising given that the distribution of employment across employer size classes (given in Table 3.2) is similar for immigrants and natives; the immigrant share varies little across these classes.

Despite this similarity in distributions, when we allow the size effect on coworker share to differ between immigrants and natives we find large size effects. That is, while the share of immigrants is relatively constant across size classes, the concentration of immigrants in the workplace falls substantially with size, as illustrated in Figure 3.3. Natives are slightly more likely to work with immigrants in larger firms than in smaller firms, while immigrants are much less likely to work with other immigrants in larger firms. For example, 37 percent of coworkers are immigrants for recent immigrants who work at establishments with 10-19 workers, while for recent immigrants at establishments with 500 or more employees, that figure is 26 percent. It is striking that these large size effects hold even after controlling for many other factors including detailed industry. We also find that the difference in immigrant coworker shares between recent and more established immigrants falls with size.

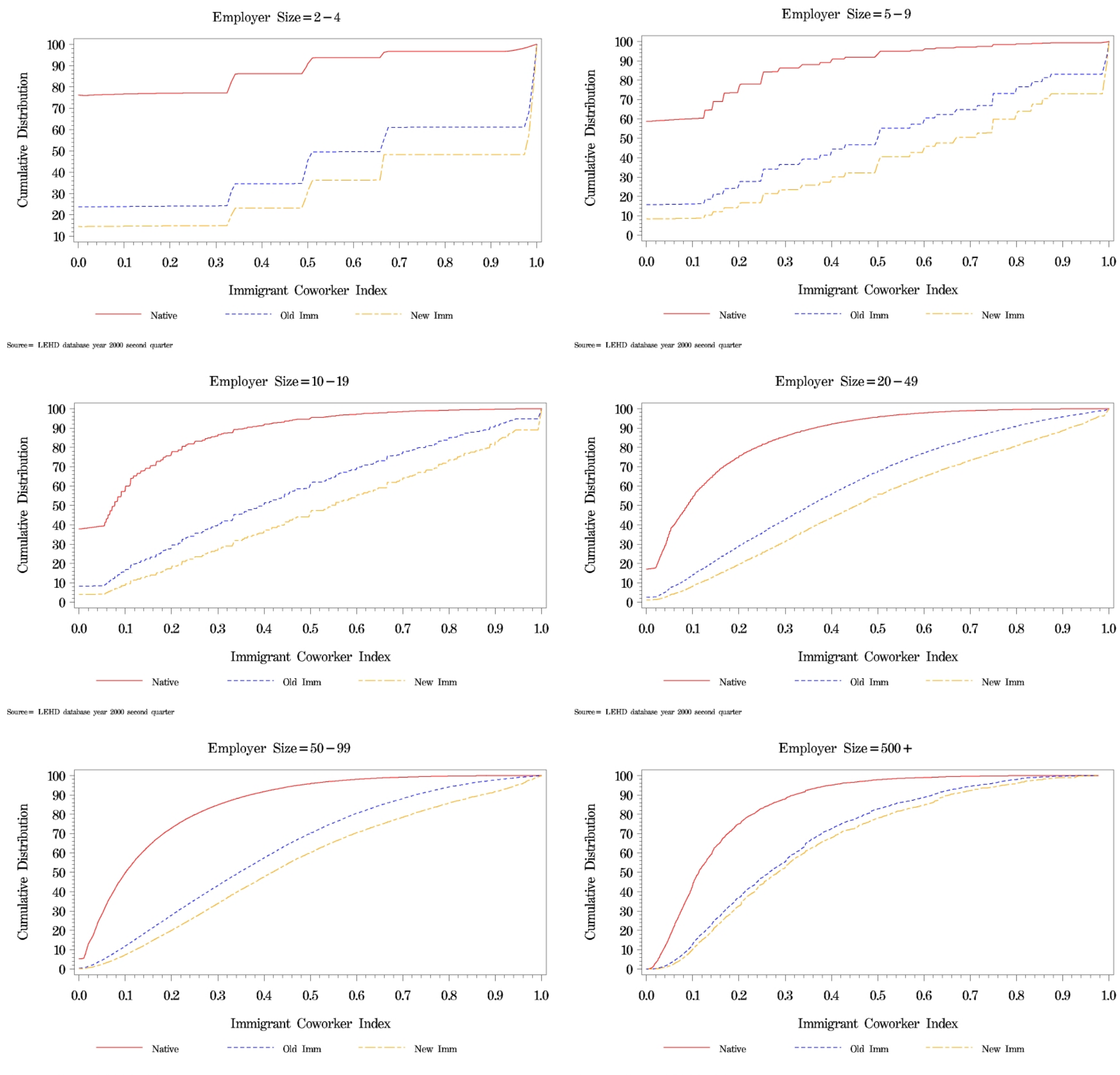
Figure 3.3: Coworker share by employer size





Notes: Evaluated at pooled mean for other control variables–MSA, sector, immigrant demographics, establishment age interacted with multi-unit status. Sector, individual's age, plant age, sex, units and MSA groups use total population distribution. Using full two-way interactions with individual status.

Figure 3.4: Cumulative Distribution of Coworker Share by Worker Type and Employer Size



Source=LEHD database. Year 2000 second quarter.

To illustrate more concretely how these differences arise, Figure 3.4 gives cumulative distributions of employment across coworker shares for different employer size classes. The size of the gap between the native and immigrant cumulative distributions represents the size of the differences in means, or the amount of concentration. For the smallest firms, much of the concentration comes from segregated workplaces—those with only immigrant or only native employees. About three-quarters of natives in this size class work only with other natives, while roughly half of recent immigrants and two out of five established immigrants work only with other immigrants. Looking across the different size classes, the share of employment accounted for by all-immigrant and all-native workplaces falls quickly as firm size increases.

We think two mechanisms drive this pattern. One is a size effect, which we find interesting—a greater tendency for immigrants to work with natives in larger firms. The second is a statistical artifact that arises from the fact that the variance across employers in the coworker share falls with employer size. For a given size-neutral tendency to group like workers together, the difference in mean coworker share will tend to fall as the variance of the mean falls—that is, with employer size.

To see this, consider 2-employee firms. The only possible outcomes are complete segregation (2 natives or 2 immigrants), or integration (1 native, 1 immigrant). If workers are randomly allocated to employers, the expected values of mean coworker shares for immigrants and natives will both equal the overall immigrant share of the (employer size=2) workforce—a difference of 0. But given some tendency to group like workers together, moving some of the weight of the distribution towards segregated workplaces has a relatively large effect on the mean difference because it moves immigrants towards workforces with coworker share=1, and natives towards coworker shares of 0. As employer size increases, extreme values become less likely. If we think of some process shifting weight away from integrated workforces to those with more segregation, with larger firms this has a smaller effect on the mean difference because less of the weight ends up at extreme values. Appendix G shows that this is true for a particular statistical model, but we think this point holds more generally.

This statistical effect is not particularly interesting but we need some way to gauge how much of the size effect it accounts for. Because the change in variance with sample size falls off quite quickly as size increases, we think the statistical effect is unlikely to account for size effects among firms with more than 20 employees. Thus it might be reasonable to think of size effects based on the portion of our sample with at least 20 employees as representing the economic size relationship, while in smaller firms the size effect combines the economic and statistical relationships. Based on this assumption, we fit a flexible functional form to the size effect for the portion of our sample with at least 20 employees, and then use the fitted model to predict the size effect for smaller firms.[12] The lower panel of Figure 3.3 superimposes this estimated/extrapolated relationship on the actual size-specific means.

For each of our three groups, we separately fit the relationship between mean coworker share and firm size over the range of firm sizes above 20. The points

[12]We use linear, quadratic and cubic functional forms to predict the size effect for smaller firms. The quadratic and cubic specifications gave very similar results. We show the quadratic results here.

marked on each line represent the mean predicted coworker share for that employer size grouping. For example, in the lower graph, the 23% marked on the established immigrant line for the 500+ size group is the mean predicted value for established immigrants in this size range—a bit lower than the actual 27% share, which is labeled in the upper graph. For groups 2-4, 5-9, and 10-19, the actual coworker share does not influence the fit of the model. The model projection fits the native means closely, which is unsurprising given that the native mean varies little with size. For immigrants, the projections under-predict the coworker means, with a particularly large gap for recent immigrants in the smallest firm size classes. If we take the projection as tracing out the real size effect, the evidence is consistent with a modest underlying size effect. Given that interpretation, the gap between the actual and projected mean then represents the purely statistical effect of size. Consistent with the statistical model in Appendix G, this effect is large for very small firms, but rapidly decreases with size.

We think that the size effects, especially after controlling for the statistical aggregation effects, potentially reflect a number of factors that influence concentration as described in section 3.2. One reason that size may matter is that the production process (even within industrial sectors) varies across establishments of different sizes. Job tasks and division of labor are likely less formal in small establishments, with all workers more likely to interact with coworkers and customers. As such, more concentrated workplaces permit immigrant workers to work alongside with immigrants,potentially overcoming language and related barriers. Still, Figure 3.4 shows that except for the smallest workplaces (where concentration is by construction high at the individual establishment level) we find that not much mass is concentrated in completely segregated workplaces. This suggests that even small businesses can find ways to organize their production activities to permit native and immigrant workers to work side-by-side.

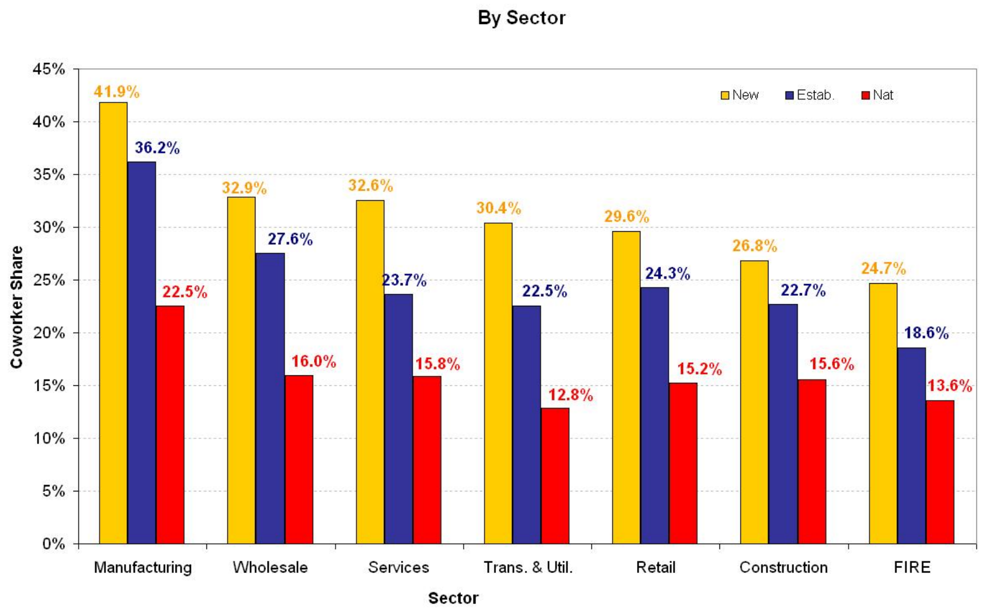
A related argument is that the hiring process is likely to be more informal for small businesses. Moreover, the number of open vacancies is likely to much smaller for a given small business (even if the rate of vacancies is as high or higher than larger businesses). Both of these effects might make social networks more important in the hiring process for small businesses. At this point of the analysis we cannot distinguish between these or alternative channels for our findings on the role of employer size. For now, we highlight the importance of employer size but we also explore some of these channels in our analysis below.

3.4.3.2 Industry

Industry differences in immigrant concentration are of particular interest to us because, given the available data sources, industry provides the best way of grouping employers that face similar constraints in choosing the skill mix of their workforce. Significant variation in immigrant concentration by industry would be consistent with technological differences playing an important role in determining how employers combine employment of natives and immigrants.

Controlling for detailed industry reduces our measure of concentration by about 13% for recent immigrants and 15% for established immigrants, while sub-

Figure 3.5: Coworker share by employer sector



Note: Size, individual's age, plant age, sex, units and MSA groups use total population distribution. Using full two-way interactions with individual status.

stantially increasing the explanatory power of the regression (as illustrated in Table 3.3). Whether we control for employer size or not has little effect on this conclusion. It is impractical to illustrate differences across 185 detailed industries, but Figure 3.5 illustrates differences by broad sector to give a sense of where immigrants are most concentrated. The figure orders sectors according to coworker shares for recent immigrants. Manufacturing is the most immigrant intensive sector in our data; even among natives, immigrants account for more than one out of five coworkers. The concentration of immigrants is also highest in manufacturing: despite the large coworker share for natives, the share for immigrants is about double the native share. The other sectors also show substantial levels of immigrant concentration at least for recent immigrants, with even the least concentrated sector (finance, insurance, and real estate) having an 18% percentage point higher coworker share for recent immigrants than for natives. Note that using the coworker share for established immigrants (or for natives) to order the sectors would change the ranking of sectors—there is less consistency across groups in ranks by sector than we found when looking at variables such as size.

3.5 Exploring social networks, language skills, and human capital as possible explanations for concentration

Section 3.4 has three main findings. First, there is substantial concentration of immigrants at the workplace. Second, even after accounting for many employer and worker characteristics including employer location, industry and size, concentration remains substantial within employer and worker characteristic groups. Third, the differences in coworker means between immigrant and native workers vary substantially with employer and worker characteristics. The most interesting interaction effects we find are by employer size and industry. These effects are especially intriguing because they arguably reflect differences in how businesses organize their workplaces. As discussed above in section 3.2, there are number of potential channels for immigrant concentration depend on the type of technology (broadly defined), organizational structure and recruiting methods of a business. In this section, we present results of analyses that look more directly at possible channels. In particular, we explore the role of network effects, English language skills and human capital.

To explore these issues, we add variables intended to capture the effects of social networks, language skills and human capital variables to the regression specifications in section 4. As a proxy for social networks, we use the neighborhood network variable defined and discussed in section 3.4. As discussed above, this network measure is likely to be correlated with a variety of factors that connect the workplace to the place of residence. Accordingly, as controls we include both the residential segregation measure and the shared commute variables. We also include education variables as well as variables measuring English speaking ability (these variables are discussed in more detail below). The education and language skill variables are of interest in their own right since the concentration of immigrants may reflect sorting by these characteristics. In addition, social network effects are also likely related to language skills so it is of interest to include both the proxy of social networks as well as language skills in the specification.

Table 3.4: Characteristics of Matched Sample Workers (Unweighted)

	Immigrants		
	Recent	**Established**	**Natives**
Coworker share	41.3	36.1	13.6
Age			
Age<30	43.6	19.0	31.0
30<Age<40	35.8	32.7	26.0
Age>40	20.6	48.3	43.0
Male	55.9	55.2	50.8
Age at arrival (*)			
<12	0.8	11.0	.
12-25	35.4	58.4	.
26-35	38.5	21.5	.
35+	25.3	9.2	.
Establishment Size			
2-9	7.8	8.21	9.2
10-49	22.7	24.2	22.3
50-99	13.4	14.2	13.2

Continued on next page.

Table 3.4: Characteristics of Immigrant and Native Workers, Full Sample(continued)

	Immigrants		
	Recent	**Established**	**Natives**
100-499	29.5	31.8	30.81
500 or more	26.7	21.6	24.6
Firm has multiple establishments	33.6	44.6	44.8
Establishment age			
0-1	12.5	10.0	11.4
2-4	25.9	19.2	14.3
Age 5 or more	61.6	70.8	74.3
Sector			
Construction	5.2	4.9	5.9
Manufacturing	17.9	17.7	12.4
Transportation & utilities	2.6	17.8	6.3
Wholesale	6.3	5.7	6.0
Retail	24.5	16.5	22.9
FIRE	2.9	4.3	6.4
Services	40.6	33.1	40.2
Log quarterly earnings	8.21	8.54	8.34
Consecutive quarters on 2000-Q2 job:			
Quarter before AND after	56.0	70.1	64.1
Quarter before OR after (not both)	32.9	21.6	27.3
Neither quarter before NOR after	11.1	8.3	8.6
Immig. share of wkrs in residence tract	35.8	34.6	13.5
Neighborhood network index	0.97	0.88	0.95
Shared commute index:			
Immigrant co-commuters	0.07	0.06	0.03
Native co-commuters	0.13	0.14	0.32

Source: LEHD database and author calculations.
(*) Year of application for a SSN is used as a proxy for time of arrival in the U.S.
Note: The unit of observation is a worker. Employer characteristics and earnings are for the first quarter 2000 job with the highest earnings. All figures except for log earnings represent percentages. There are 3,549,111 matched workers in total for our group of MSAs.

For this analysis, we focus on our matched sample since it is for this sample that we have education and English language skill measures. Table 3.4 shows unweighted summary statistics for our matched sample. Comparing Table 3.4 to Table 3.2 illustrates that most of the differences between the matched and full samples are modest. Matched natives have a somewhat lower coworker share than in the full sample, but the shares for immigrants are little changed. There seems to be a gen-

Table 3.5: Characteristics of Immigrant and Native Workers, Matched Sample (weighted)

Weighted	Immigrants Recent	Immigrants Established	Native
Education categories			
High school drop-out	34.90	28.03	17.08
High school graduate	19.38	15.78	25.38
Some college	13.51	15.01	26.02
Bachelor's degree	20.80	32.98	23.94
Advanced degree	11.46	8.20	7.58
Does not speak English well	31.35	16.89	1.14

Source: LEHD database and author calculations.
(*) Year of application for a SSN is used as a proxy for time of arrival in the U.S.
Note: The unit of observation is a worker. All figures represent percentages. There are 3,549,111 matched workers in total for our group of MSAs.

eral tendency for workers at multi-unit firms to be overrepresented in the matched sample, but the difference is large only for established immigrants. Workers with very transitory jobs also tend to be somewhat overrepresented. The most dramatic differences are in the mean values for the shared commute and neighborhood network index ,which are much smaller in the matched sample for all three groups. Table G.1 gives weighted statistics for the matched sample to illustrate that the weights we construct bring us reasonably close to replicating the observable characteristics of the full sample.

Table 3.5 presents summary statistics for the additional variables that we can construct using the matched data (using the propensity score weights to make the sample representative). Immigrants are much more likely to be high school drop-outs than are natives, particularly very recent immigrants. But immigrants are also overrepresented among those with advanced degrees. The English language measure we use is based on a sequence of questions asked on the census long form questionnaire. All respondents are asked whether they speak English or another language at home. Those who report that they speak another language at home are then asked to categorize how well they speak English–not at all, poorly, well, or very well. We categorize all those responding either not at all or poorly as not speaking English well. Note that this includes a small fraction of natives, presumably primarily second generation immigrants. Unsurprisingly, recent immigrants are more likely than others to report not speaking English well.

Our analysis of the contribution of these additional variables is in two steps. First, we estimate the model without any interactions to examine the direct role of these variables in accounting for immigrant concentration. Table 3.6 is the extension of Table 3.3 with these additional variables. The first two rows correspond to the first two rows in Table 3.3, but estimated on the matched sample using our weights. The third row corresponds to the "All of the above" row in Table 3.3. While estimated concentration is somewhat lower for established immigrants using the matched sample, the estimates using all of the controls from the full sample match Table 3.3's results very closely. Adding the additional controls without interactions modestly increases the explanatory of the model, with the English language measure having a more substantial effect than the education measure.

In Table 3.7 we show results for the specification that includes a full set of interactions so that the effects of variables can differ between immigrants and natives. To simplify the model, we categorize all immigrants together rather than distinguishing between recent and more established immigrants. In Table 3.7 we

Table 3.6: Contribution of Covariates to Immigrant Concentration (Matched Sample)

Covariates	Recent immigrant	Established immigrant	R-squared
Matched sample			
No covariates	0.269	0.186	0.175
MSA dummies	0.218	0.139	0.369
Full sample specification	0.148	0.090	0.498
Full sample specification +:			
Education controls	0.146	0.087	0.500
English language controls	0.129	0.080	0.504
Education and English controls	0.128	0.078	0.506

Notes: Figures in the first two columns give the predicted difference in mean coworker share between the immigrant group and natives. As a point of reference, the mean coworker share for natives in the first line is .145 (as in Table G.1). It is also .145 for all other specifications if evaluated at the native mean for all included covariates, but somewhat higher if evaluated at the pooled sample mean. The unit of observation is a worker and N=3,549,111. The categories for worker age, employer size and employer age are the same as in Table G.1. There are 185 detailed industry categories. All standard errors are less than 0.0001.

only report the interaction coefficients for added variables of interest but note that the patterns by employer and employee characteristics discussed in section 4 are still present. It is also of interest to observe that the fully interacted model now accounts for a substantially larger share of variation – the R-squared in Table 3.7 is 0.605.

The results in Table 3.7 support the hypothesis that social networks play an important role in workplace concentration. The network index variable is positively associated with concentration: natives who work with their neighbors have fewer immigrant coworkers, while immigrants who work with their neighbors have more. It is important to emphasize that this pattern holds controlling for a rich set of employer and employee characteristics and for shared commute patterns, residential location, education and language skills.

Turning to the other effects of interest, the coefficients on education and the

Table 3.7: Network Effects from Coworker Share Regressions

Covariates	**(1)**
Neighborhood network index	-0.081
Network index * Immigrant	0.443
Native shared commute index	0.033
Native shared commute * Immigrant	-0.754
Immigrant shared commute index	-0.774
Immigrant shared commute * Immigrant	0.351
Immigrant share in residential tract	0.075
Immigrant residential share * Immigrant	0.053
High school drop-out	0.016
High school graduate	0.002
College graduate	0.002
Graduate degree	0.004
High school drop-out * Immigrant	-0.004
High school graduate * Immigrant	0.009
College graduate * Immigrant	0.009
Graduate school degree * Immigrant	0.009
Does not speak English well	0.216
Does not spreak English well* Immigrant	0.035
R-Squared	0.605

Note: All standard errors are below 0.001. Controls in all columns include MSA, detailed industry, employer age and size, worker age and sex, along with interactions with immigrant for each, in addition to the variables listed in the table. Education and English language controls and their interactions with immigrant are added as well. The reference group is the group natives who speak English well. The unit of observation is a worker. N=3,549,111. These are the workers with complete information on their residential location.

English language variable indicate that high school drop-outs and those with poor English language skills are more likely to work with immigrants. While significant, even the largest education effect–the difference between high school drop-outs and those with some college–is quite small. The effect of having limited ability to speak English is substantial–immigrants who do not speak English well have about 25% more immigrant coworkers.

The control variables are also of interest. The residential segregation index has the expected impact. A higher share of immigrants in the residential tract is associated with a higher share of coworkers who are immigrants for both natives and immigrants. The shared commute pattern has a somewhat anomalous effect, although these coefficients are highly sensitive to the inclusion of the residential segregation index. Thus, it appears that these two controls are capturing related effects that may be difficult to identify separately. We note that the main effects of interest are robust to including only one of these controls (i.e., either residential segregation or shared commute).

To explore the role of education and English speaking ability further, we also consider a richer version of the specification presented in Table 3.7. That is, returning to the specifications considered in section 4, we estimate a specification with the new variables and also distinguishing between new and established immigrants. The pattern results are quite similar to those reported in Table 3.7 but enriched by the additional degree of variation.

3.6 Country of origin differences

In the analysis above, we distinguish between natives, recent immigrants and established immigrants. Our data also permit exploring how the patterns of immigrant concentration vary by country of origin. That is, instead of only asking how likely it is for immigrants to have co-workers that are immigrants, we can ask how likely it is for an immigrant from say, Mexico, to have co-workers who are Mexicans. Examining such patterns could potentially shed further light on the relative merit of various language- and cultural-based explanations for immigrant workplace concentration. In particular, we would expect social networks to have much stronger effects within country-of-origin groups, while language-based explanations would imply similar effects for immigrants from different countries that share the same language. We plan to investigate these issues in the next draft of this paper.

3.7 Concluding remarks

Using matched employer-employee data that extensively cover employment in our sample of MSAs, we find that immigrants are much more likely to work with each other—and hence less likely to work with natives—than would be expected given random allocation of workers. This is in part driven by the distribution of immigrants across MSAs, but within an MSA, substantial concentration remains. We find evidence that immigrant assimilation into the U.S. workforce generated a tendency to have more native coworkers as more time is spent in the U.S. We also document that immigrant concentration is greatest in small firms, and varies substantially across industries.

After presenting descriptive results, we examine possible underlying causes of this concentration. We find evidence that supports the hypothesis that social networks, language skills and education are important factors in accounting for workplace concentration. Our results indicate that natives who live near some of their coworkers are more likely to work with others who are native born. The effect for immigrants is similar—they are more likely to work with immigrants if they live near coworkers—but larger. These findings hold controlling for a variety of other factors (e.g., residential segregation) that could lead to a correlation between residential and employment location. We also find that immigrant workers with poor English speaking skills and less educated workers are more likely to work with other immigrants. These effects are of interest in their own right, since they suggest that some of the workplace concentration we observe is associated with sorting by skill and language, but these effects also act as controls for the other variables of interest. The results for the full specification indicate that our results on social networks are robust to inclusion of language and education controls.

Chapter 5

Conclusion

This thesis took advantage of unique employee and employer matched microdata from the U.S. Census Bureau to examine immigrant segregation at the workplace, the effect of owner types, and coworker types on firms' hiring patterns and workers' earnings. In Chapter 2 particular attention was paid to the nativity of employers and to the share of similar coworkers (by nativity and ethnicity) at firms when new workers are hired. We examined the effect of those variables on hiring rates and on the wage differential between immigrants and natives. In 3 a detailed geographic residential location was used to describe networks by links between residential neighbors.

We found that immigrants are much more likely to work with each other, and hence less likely to work with natives, than would be expected given random allocation of workers. This is in part driven by the distribution of immigrants across MSAs, but within an MSA, substantial concentration remains. We found evidence that suggests that immigrant assimilation into the U.S. workforce includes a tendency to have more native coworkers with more time in the U.S. We also documented that immigrant concentration is greatest in small firms, and varies substantially across industries.

In general, employees' wages are affected by the type of owner of the firm.

For native employees the effect on wages is higher when working for immigrant employers. Natives are paid lower when working for immigrant employers, and in these firms natives have lower average earnings than immigrants. One explanation for these findings is that immigrant bosses have a better understanding of and networking with the immigrant community, and, therefore, can better find and contract immigrant workers than native-owned firms. Why can't native-owned firms quickly adjust and find this cheaper labor? Lack of language knowledge and lack of networking make it harder for native bosses to find immigrant workers. These findings justify further analysis of differences in contracting ability across employers. The evidence that the type of owner matters for wage differentials among workers also implies an important role for personnel policy.

After presenting descriptive results, we examine possible underlying causes of this concentration. We find evidence that supports the hypothesis that social networks, language skills and education are important factors in accounting for workplace concentration. Our results indicate that natives who live near some of the coworkers are more likely to work with others who are native born. The effect for immigrants is similar—they are more likely to work with immigrants if they live near coworkers—but larger. These findings hold controlling for a variety of other factors (e.g., residential segregation) that could lead to a correlation between residential and employment location. We also find that immigrant workers with poor english speaking skills and low educated workers are more likely to work with other immigrants. These effects are of interest in their own right since they suggest some of the workplace concentration we observe is associated with sorting by skill and language but these effects also act as controls for the other variables of interest. Namely, that the impact of social networks we estimate is robust to inclusion of language and education controls. These findings are consistent with an important role for social networks, though there are other mechanisms that could lead to a correlation between residential and employment location that we have not yet investigated. We posited that social network effects should be more important in smaller firms that are less likely to have formal human resources practices. We also found that poor English language skills and low education levels are highly correlated to immigrant workplace concentration, though our education measures are not high in levels in our regressions once the language measure is included.

By shedding light on the ways workers and employers interact in the labor market to affect job and wage outcomes, this research makes a contribution to the sociology, labor economics, and demography literatures. It also opens up numerous avenues for future research. On the microeconomic side, we can further evaluate job flows and wage profiles of workers inside different types of firms. The analysis of assimilation can also take advantage of the results presented here, to further our understanding of the adjustment process of new immigrant workers. The empirical analysis in this paper makes some progress toward mitigating biases of skill sorting and controls for a broad number of observable characteristics that try to capture other explanations for segregation and groups differentials.

Appendix A

Matching Rate

To have an idea of the groups of firms included in both database, I include a short discussion on firms matching rate. Mostly, matches between CBO and SSEL in 1992 are employer firms (See Table(A.1)). However, there is a small portion of non-employers that match with SSEL[1]. The matching rate, although very high, is not 100%. There is a group of industries that are not included in SSEL such as Private Households (88), and Direct Sellers (5963) that are included in the CBO. The matching rate increases with the number of owners of the firm. Number of owners in the firm and size of the firm are also relatively proportional.

Table A.1: Matching and Non-matching rate of firms in CBO and SSEL(single-unit)

	CBO: One Owner	
	Non-Employer	**Employer**
Non-matches	99.28	10.79
Matches	0.72	89.21

	CBO: Two Owners	
	Non-Employer	**Employer**
Non-matches	79.23	2.55
Matches	20.77	97.45

	CBO: Multiple Owners	
	Non-Employer	**Employer**
Non-matches	88.31	3.92
Matches	11.69	96.08

Source: Authors calculation based on CBO(1992) and SSEL(1992).

[1]The SSEL in 1992 seems to include those nonemployer firms that were subject to Federal Income Tax[2].

Appendix B

Definitions

Standard Statistical Establishment List (SSEL).

- Sales: Following Spletzer [1998], I consider the variable sales as to the sum over ACSR1, ACSR2, ACSR3. We can include ACSR4 for Corporations and Partnerships. Spletzer [1998] gives a more detailed information on the sale data in the SSEL. For single establishment file has sales data. This variable contains data for the Value of Shipments, Sales, Receipts, or Revenue. It may include ''total revenues, gross income, operating receipts (gross receipts or sales less returns and allowances), interest income, and gross rents". In 1992, 80.1 percent of SSEL single establishments have current year sales data. When matched with CBO we obtain sales data for all firms. Spletzer [1998] compared the values for Employment, Payroll, and Sales between SSEL and the Economic Census. He found that, for 1992 and single establishments in Maryland, the number for the two first match well. However, the numbers for sales and sales per worker shows a difference above 8%. He speculates that the difference is coming from distint definitions, specifically for commissions for wholesalers ''that sell as an agent for another company (Type of Operation code=43 or 46)."

- For the employment variable I use the sum of ACEMP and AC943E.

Appendix C

Unknown-Owned Firms

Before continuing with the analysis, it is worthy of attention to mention that there exists a group of undefined firms for which owners' nativity is unknown. Mostly, this group corresponds to firms that did not answer the survey (CBO). 95% of the owners of these firms did not answer the survey. I obtained information on their distribution, size, sales and payroll using SSEL and EC in 1992. At a glance, from Table(2.1) the distribution of this type of firms across size and sectors is similar to the rest of the group. For further analysis, I compute a t-test of equality of productivity proxies and earnings per employees means between the unknown-owned firms and a weighted average value of the productivity proxies and earnings of the other three groups. A t-test cannot reject the null hypothesis that the mean of labor productivity (t-test=-0.89), and logarithm of earnings per employee (t-test=2.2) of unknown-owned firms and the rest of the groups are the same at the 90% level. The average number of owners and the average share immigrant workers at the firm is also similar when we compare unknown-owned firms with the average of immigrant, mix and native-owned firms.

Additionally, a chi-square test over unknown-owner firms' and the rest of the group's distribution across categories (size and sector) cannot reject the hypothesis that unknown-owned firms and the total of firms excluding unknowns are similarly distributed across the categories size and sector presented in Table(2.1) at the 90% level. The Pearson chi-square for the size distribution is 9.02 (Pr=0.11). For the distribution across sectors, the Pearson chi-square with 6 degree of freedom is equal to 16.02 (Pr=0.08). Like the other firms types, unknown-owned firms are highly represented in Wholesale, Manufacturing, Services and Retail. At the same time, more than 70% of unknown-nativity employers are businesses with less than 20 employees.

It is very relevant for us to know whether the owner (or owners) of the firm was (were) born in the USA or otherwise. We cannot identify this profile for the group unknown. Given that the characteristics of the unknown group are very similar, in average, to rest of the sample, I decide to drop these observations. For the rest of the paper, we only consider those groups for which nativity is obtained (three groups of firms: native-owned, mix-owned, and immigrant-owned).

Appendix D

Weights and Selection

According to Heckman(1979), I obtain the probit estimate from the probit selection equation in order to estimate the inverse mills ratio. The probability of being a matched firm in the sample is estimated as a function of the characteristics of the firm: size, industry, legal form of organization, geographic location, owner type, and continuing or exiting firm.

$$\lambda(z) = \frac{\phi(z)}{\Phi(z)} \tag{D.1}$$

where $\Phi(.)$ is the standard normal cfd, $\phi(.)$ is the standard normal density and z is $x^{'}\beta/\sigma$. The covariates x are the ones discussed above and the coefficients are estimates of the probit model.

Table D.1: Descriptive Statistics - CBO(1992) and Sample/Matched Firms

	Matched Sample(CBO-LEHD) unweighted				
Distribution/Type of firm	**Mix**	**Imm**	**Nat**	**Unk**	**ALL**
Size (%)					
2-4	28.93	29.50	25.32	20.40	24.53
5-9	14.93	21.40	19.78	28.30	21.35
10-19	15.98	25.56	21.67	17.89	21.40
20-49	19.56	17.15	18.37	19.42	18.79
50-99	10.62	4.01	9.40	9.42	8.89
100+	9.98	2.38	5.46	4.57	5.04
Sector(%)					
Construction	4.35	3.20	13.40	10.20	11.34
Manufacturing	27.50	17.02	20.50	14.50	18.25
Transp. & Utility	6.40	2.48	8.50	6.55	7.20
FIRE	17.89	15.32	16.70	21.50	19.34
Retail	15.40	38.40	18.40	22.35	19.45
Wholesale	10.50	4.01	5.45	4.56	4.79
Services	17.96	19.57	17.05	20.34	19.63
Legal Form(%)					
Sole Propietorship	-	55.70	53.40	15.43	49.30
Partnership	24.30	17.50	15.30	30.27	25.43
Corporation (*)	75.70	26.80	31.30	54.30	25.27
l(sales/employment) (1)	13.45	13.39	12.49	15.30	13.55
	(1.08)	*(1.19)*	*(1.05)*	*(1.18)*	*(1.13)*
Average Number of Owners	4.01	1.60	2.00	1.96	1.94
Race/Ethnicity					
Hispanic	10.45	20.45	5.00	4.21	6.40

Continued on next page.

Table D.1: Descriptive Statistics - CBO(1992) and Sample/Matched Firms (continued)

	Matched Sample(CBO-LEHD) unweighted				
Distribution/Type of firm	**Mix**	**Imm**	**Nat**	**Unk**	**ALL**
Asian	11.50	45.50	2.43	3.19	9.29
Black	0.95	0.50	2.59	2.30	1.54
White	77.10	33.55	89.98	90.30	82.77

Note: Statistics based on unweighted outcomes.
(1) Single-unit firms that matched with SSEL.
(2)Only S- Corporation .
(3)Source SSEL: Sales (total receipts/sales), and employment (Employment March12th). Numbers in parenthesis are standard deviations.

Appendix E

IPUMS 1990: Descriptive Statistics

Table E.1: Descriptive Statistics - Characteristics of Workers

	Individual		
	IM	**US**	**ALL**
Age	36.28	36.74	36.68
	(11.49)	(11.93)	(11.74)
Education	14.45	16.9	16.5
	(2.41)	(2.53)	(2.89)
Log(annual earnings)	9.75	9.96	9.94
	(0.96)	(0.97)	(0.98)
DISTRIBUTION (%)			
AGE			
Under 25	15.68	16.58	16.47
25-39	48.76	45.01	45.49
40+	35.56	38.40	38.04
EDUCATION			
High School Dropout	35.21	14.98	17.54
High School Graduate	18.15	28.55	27.24
Some College Education	15.66	23.48	22.50
College Graduate	30.98	32.98	32.73
SECTOR			
Construction	10.55	9.87	9.95
Manufacturing	26.36	25.01	25.18
Transportation and Utilities	7.62	11.72	11.21
Wholesale	5.66	6.81	6.66
Retail	19.01	16.23	16.58
FIRE	5.38	6.08	5.99
Services	25.41	24.27	24.41
RACE			
White	28.07	83.51	76.50
Asian	21.74	0.81	3.46
Black	6.69	9.90	9.49
Hispanic	43.08	5.28	10.06
Other	0.42	0.50	0.49
All	**12.64**	**87.36**	**100.00**

Continued on next page.

Table E.1: Descriptive Statistics - Characteristics of Workers (continued)

	Individual		
	IM	US	ALL

Note: The sample includes only male workers older than 16 years. Sectors Agriculture, Mining and Public Administration are not included. Self-employed workers are not considered. Statistics based on weighted outcomes. Standard Deviations in parenthesis.

Appendix F

Linear Probability Estimates

Table F.1: Linear Probability Estimates of the Effect of Owner Type on the Probability that a New Hire is Black

	Black					
	(1)	**2)**	**(3)**	**(4)**	**(5)**	**FE**
Owner Native	0.0231***	0.0312***	0.0235***	0.0345***	0.0345***	
	0.002	*0.002*	*0.003*	*0.013*	*0.013*	
Hispanic Cowkrs				-0.8262***	-0.8321***	-0.7143***
				0.0299	*0.0337*	*0.034*
Asian Cowkrs				-0.8516***	-0.8363***	-0.6523***
				0.0269	*0.0302*	*0.032*
White Cowkrs				-0.9126***	-0.9073***	-0.2834***
				0.0252	*0.0284*	*0.054*
Black Cowkrs			0.9513***			
			0.0345			
Hispanic Cowkrs *				0.0009	0.0048	
Owner Native				*0.03*	*0.03*	
Asian Cowkrs *				0.0898	0.0707	
Owner Native				*0.103*	*0.0384*	
White Cowkrs *				0.0386	0.0584**	
Owner Native				*0.026*	*0.0293*	
Black Cowkrs*			0.1298***			
Owner Native			*0.023*			
log(employment)					0.0082***	
					0.00	
Share of workers					0.0011***	
with HSD (firm)					*0.0005*	
Share of workers					-0.0001	
with HSG (firm)					*0.0001*	
Share of workers					0.0005	
with SCG (firm)					*0.0008*	
Working Pop.					-0.0362*	
% immigrant(+)					*0.01*	

Continued on next page.

Table F.1: Linear Probability Estimates of the Effect of Owner Type on the Probability that a New Hire is Black (continued)

	Black					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**FE**
Working Pop. % white(+)						
Working Pop. % black (+)					0.095***	
					0.01	
Constant	0.0689***	0.0269***	0.0867***	0.4116***	0.4116***	0.3145
	0.0026	*0.0077*	*0.01*	*0.1612*	*0.1612*	*0.201*
year dummies	yes	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes	-
State dummies	-	-	yes	yes	yes	-
Other controls (+)	-	-	-	-	yes	-
p-value	*0.0001*	*0.002*	*0.0001*	*0.003*	*0.003*	*0.01*
R-Square	*0.21*	*0.23*	*0.31*	*0.31*	*0.33*	*0.34*

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. Mix-owned firms outcomes are not reported in the table but they are used in the regressions. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

Table F.2: Linear Probability Estimates of the Effect of Owner Type on the Probability that a New Hire is White

	White					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**FE**
Owner Mix	0.0197***	0.0518***	0.0077	-0.0007	-0.012	
	0.0074	*0.0073*	*0.005*	*0.02*	*0.02*	
Owner Native	0.2633***	0.2307***	0.0412***	0.0391***	0.0233***	
	0.004	*0.004*	*0.003*	*0.009*	*0.010*	
Hispanic Cowkrs				-0.8325***	-0.8197***	-0.147***
				0.0255	*0.0257*	*0.0276*
Asian Cowkrs				-0.6869***	-0.6904***	-0.5821***
				0.0211	*0.0213*	*0.0234*
White Cowkrs			0.8861***			
			0.0279			
Black Cowkrs				-0.8936***	-0.8654***	-0.6634***
				0.0375	*0.0378*	*0.465*
Hispanic Cowkrs * Owner Mix				-0.1247***	-0.1252***	
				0.05	*0.0566*	
Asian Cowkrs * Owner Mix				-0.2546	-0.2212	
				0.0507	*0.0518*	
White Cowkrs* Owner Mix			0.1015***			
			0.0454			
Black Cowkrs * Owner Mix				-0.047	-0.0236	
				0.1047	*0.1048*	
Hispanic Cowkrs* Owner Native				0.0351*	0.0390*	
				0.017	*0.017*	

Continued on next page.

Table F.2: Linear Probability: Effect of Owners types on the Probability that a New Hire is White (continued)

	White					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**FE**
Asian Cowkrs *				0.0996**	0.1154**	
Owner Native				*0.0363*	*0.0364*	
White Cowkrs *			0.077***			
Owner Native			*0.0354*			
Black Cowkrs *				0.0194	0.0257	
Owner Native				*0.0386*	*0.0388*	
log(employment)					0.0061**	
					0.00	
Share of workers					-0.0018***	
with HSD (firm)					*0.0008*	
Share of workers					-0.0024***	
with HSG (firm)					*0.0005*	
Share of workers					-0.0046***	
with SCG (firm)					*0.0012*	
Working Pop.					-0.1471***	
% immigrant(+)					*0.036*	
Working Pop.					*0.0754****	
% white (+)					*0.023*	
Working Pop.						
% black (+)						
Constant	0.4407***	0.4112***	0.9702***	0.8747***	0.915***	0.71*
	0.004	*0.127*	*0.08*	*0.2327*	*0.237*	*0.33*
year dummies	yes	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes	-
State dummies	-	-	yes	yes	yes	-
Other controls (+)	-	-	-	-	yes	-
p-value	*0.0001*	*0.002*	*0.0001*	*0.003*	*0.003*	*0.01*
R-Square	*0.27*	*0.31*	*0.33*	*0.35*	*0.39*	*0.41*

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

Table F.3: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Black

	Black				
	(1)	**2)**	**(3)**	**(4)**	**(5)**
Owner Asian	-0.4114***	-0.3728***	-0.0373**	-0.057***	-0.2238
	0.0076	0.0074	0.0134	0.0316	0.0369
Owner White	-0.3928***	-0.3697***	-0.0407***	-0.1313***	-0.1594
	0.0072	0.007	0.013	0.0178	0.0226
Black Cowkrs			0.9512***		
			0.0249		

Continued on next page.

Table F.3: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Black (continued)

	Black				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Asian Cowkrs				-0.6087***	-0.62
				0.1743	0.1981
White Cowkrs				-0.9677***	-0.956
				0.0292	0.037
Hispanic Cowkrs				-0.8834***	-0.85
				0.0586	0.0701
Other Cowkrs				-0.7862***	-0.8954
				0.1046	0.1312
Owner Asian*			-0.115***		
Black Cowkrs			0.0375		
Owner Asian*				-0.2324*	-0.0597
Asian Cowkrs				0.1767	0.2008
Owner Asian*				-0.1286***	0.193
White Cowkrs				0.0427	0.0514
Owner Asian*				-0.1383***	0.2152
Hispanic Cowkrs				0.0679	0.0803
Owner Asian*				0.1728	0.157
Other Cowkrs				0.1115	0.138
Owner White*			-0.0862***		
Black Cowkrs			0.0262		
Owner White*				-0.135	-0.0504
Asian Cowkrs				0.1749	0.1989
Owner White*				-0.0993**	0.1262
White Cowkrs				0.0253	0.0381
Owner White*				0.0493	0.0758
Hispanic Cowkrs				0.0597	0.0712
Owner White*				0.2033**	0.1182
Other Cowkrs				0.1057	0.1323
Share of workers					0.0013
with HSD (firm)					0.0005
Share of workers					-0.0003
with HSG					0.0003
Share of workers					0.0011
with SCG					0.0008
Log employment					0.0082
					0.0009
Work. Pop.					0.1346
Total					0.0239
Work. Pop.					-0.0001
% Black					0.0001
Constant	0.4784***	0.4082***	0.0869	0.0717	0.2882
	0.0072	0.0766	0.0784	0.0787	0.1743
Year dumies	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes
State dummies	-	yes	yes	yes	yes
Other Controls	-	-	-	yes	yes
p-value	0.01	0.02	0.02	0.02	0.03

Continued on next page.

Table F.3: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is Black (continued)

	Black				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
R-Square	0.22	0.24	0.28	0.28	0.31

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373. Standard Errors are Huber-White robust standard errors, corrected for firm clustering. Outcomes from Hispanic-owned firms are used in the regression but are not reported because of disclosure revision. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

Blacks new workers are more likely to be hired by Black owned firms(See Tables(F.3). The effect of having other black coworkers is stronger for black owned firms and then white owned firms. As for Hispanic workers, black employees are more likely to be hired by firms with higher concentration of high school drop outs compared to college graduates.

Table F.4: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is White

	White				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Owner Black	-0.3989***	-0.3946***	-0.0909***	-0.0202	-0.0198***
	0.0324	0.0114	0.0178	0.0334	0.0085
Owner Asian	-0.3681***	-0.3019***	-0.0808***	-0.0393**	-0.0219***
	0.0366	0.0046	0.0077	0.0134	0.0031
Owner Hispanic	-0.3496***	-0.2735***	-0.0608***	-0.0331**	-0.0137**
	0.0272	0.0048	0.0086	0.0143	0.0071
White Cowkrs			0.8745***		
			0.0065		
Black Cowkrs				-0.8774***	-0.5805***
				0.0292	0.0789
Asian Cowkrs				-0.8707***	-0.804***
				0.0168	0.0627
Hispanic Cowkrs				-0.7089***	-0.7722***
				0.0139	0.1131
Other Cowkrs				-0.619***	-0.6094***
				0.0192	0.1152
Owner Black*			-0.0796*		
White Cowkrs			0.042		
Owner Black*				-0.0776	-0.1903

Continued on next page.

Table F.4: Linear Probability Estimates of the Effect of Owner Race on the Probability that a New Hire is White (continued)

	White				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Asian Cowkrs				0.2722	0.298
Owner Black*				0.0715	-0.0166
Hispanic Cowkrs				0.0917	0.1139
Owner Black*				-0.1588***	-0.1637*
Black Cowkrs				0.0466	0.1012
Owner Black*				0.0173	0.0725
Other Cowkrs				0.1679	0.2905
Owner Asian*			-0.0819***		
White Cowkrs			0.0157		
Owner Asian*				-0.1210***	-0.1311**
Asian Cowkrs				0.0362	0.0852
Owner Asian*				-0.1219***	-0.0696
Hispanic Cowkrs				0.036	0.0826
Owner Asian*				0.052	-0.0481
Black Cowkrs				0.0498	0.1291
Owner Asian*				-0.1573***	-0.1257
Other Cowkrs				0.0506	0.1408
Owner Hispanic*			-0.0551***		
White Cowkrs			0.0187		
Owner Hispanic*				0.105	0.0516
Asian Cowkrs				0.1008	0.1285
Owner Hispanic*				-0.184***	-0.1561**
Hispanic Cowkrs				0.0361	0.0795
Owner Hispanic*				-0.1482***	-0.2106***
Black Cowkrs				0.0392	0.1209
Owner Hispanic*				-0.2512***	-0.192
Other Cowkrs				0.046	0.1294
Share of workers					-0.0025***
with HSD (firm)					0.0006
Share of workers					-0.0019
with HSG					0.0016
Share of workers					-0.0055*
with SOG					0.0035
Log employment					0.0133***
					0.0048
Work. Pop.					0.1475
Total					0.1113
Work. Pop.					0.0002
% White					0.0003
Constant	0.7209***	0.8591***	0.0411	0.9194***	0.1583*
	0.0152	0.126	0.1174	0.1195	0.0946
Year dumies	yes	yes	yes	yes	yes
Industry dummies	-	yes	yes	yes	yes
State dummies	-	yes	yes	yes	yes
Other Controls	-	-	-	yes	yes
p-value	0.01	0.01	0.01	0.01	0.01
R-Square	0.24	0.25	0.33	0.35	0.40

Note: Reference group is native firms.Reference Sector is Services. The number of observations is 147,373.

Standard Errors are Huber-White robust standard errors, corrected for firm clustering. (+) Other controls include: location in a MSA dummy, legal form of organization, population in thousands in the neighborhood, interaction between 2-digit industry dummy and English speaker dummy. Neighborhood is defined as the adjacent counties to the county where the firm is located. Population in 100,000's. ***significant at 1%, ** significant at 5%, * significant at 10%.

Appendix G

Simulations of employer size effects in a statistical model with segregation

If immigrants and natives are randomly allocated to jobs in proportion to their presence in the working population, the expected difference between immigrants and natives in the share of coworkers who are immigrant is zero regardless of employer size. However, we find that the distribution of immigrants across workplaces is clearly not consistent with random allocation, and that concentration is particularly high in small businesses. This raises the question of whether we should expect some general tendency to segregate to have the same effects on measured concentration in small and large businesses. The following sets up a statistical model that incorporates some segregation, then uses the model to simulate how employer size might affect differences in coworker immigrant share.

Suppose that an employer of given size s draws its workforce randomly from the population, but that some fraction of initial draws that involve an integrated workforce (i.e. some natives and some immigrants) are rejected and replaced with a new draw.

Assume that the outcome of each draw can be described using the binomial probability mass function:

$$b(i,s) = \begin{pmatrix} i \\ \\ s \end{pmatrix} p_D^i (1 - p_D)^{s-i} \tag{G.1}$$

where i represents the number of immigrants in the workforce draw, s represents employer size, and p_D represents the fraction of workers who are immigrants in the group being sampled in draw D. For the initial draw, the parameter p_0 will equal the overall share of immigrants in the workforce.

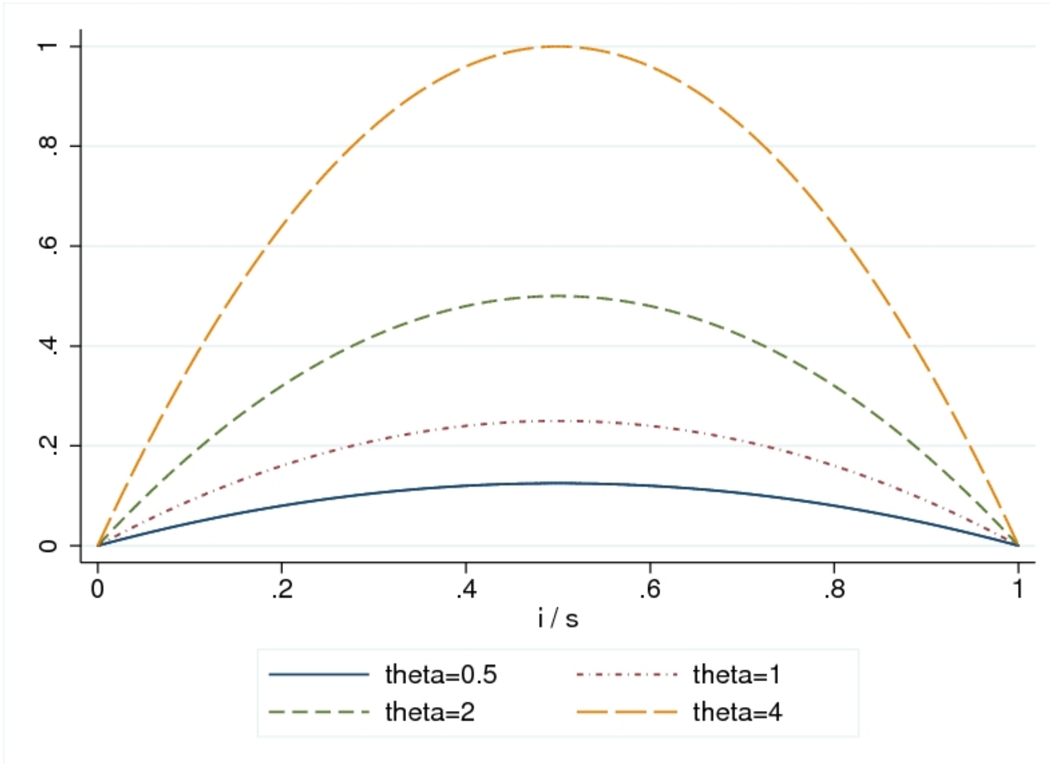
Suppose that employers discard a draw with probability d which depends on the composition of the workforce, and a parameter θ which indexes the tendency to segregate ($0 \leq \theta \leq 4$).

$$d(i; s, \theta) = \frac{i}{s} \left(\frac{s - i}{s} \right) \theta \tag{G.2}$$

If an employer draws only immigrants or only natives, then $d = 0$—the original draw is kept. If there are some of both types of employees, then the workforce is redrawn with probability d. This shifts some of the probability mass from more integrated towards more segregated types of employee mixes. Figure G.1 illustrates the shape of $d()$ for various values of θ.

For $\theta = 4$, all draws with immigrants making up exactly half the workforce ($i/s = .5$) are discarded in the first round. However, even with $s = 2$, the final distribution includes some workforces with i/s = .5 because 1 immigrant and 1 native can be drawn in the second round.

Figure G.1: Shape of function d



If immigrants account for a small share of the population, they are more likely than natives to be included in integrated workforces in the first draw. Because of this, the population that the second draw is taken from has a somewhat higher share of immigrants than the initial population. (e.g. with $s = 2$, immigrants are always half of the workers in discarded first round draws.)

Thus the second draw is also binomial, but the immigrant share is given by:

$$p_1 = \frac{\sum_{j=1}^{s} b(j, s|p_0) * d(j; s, \theta) * j}{\sum_{j=1}^{s} b(j, s|p_0) * d(j; s, \theta) * s} \tag{G.3}$$

and:

$$Pr(i, s|p_0, \theta) = b(i, s|p_0) * (1 - d(i; s, \theta)) + b(i, s|p_1) * \left(\sum_{j=0}^{s} b(j, s|p_0) * d(j; s, \theta) \right) \tag{G.4}$$

For the simple case $s = 2$ and $\theta = 4$ (so $d = 1$ for the only integrated workforces—those with 1 immigrant, 1 native), $p_1 = .5$, and the probability of observing a workforce with 1 immigrant and 1 native in the final distribution simplifies to $p_0(1 - p_0)$ (half the binomial probability). Figure G.2 illustrates the difference between the distribution of the coworker mean with segregation and without for employers of varying size. It uses parameter values $\theta = 4$ and $p_0 = .25$. Smaller values of θ would reduce the shift in the distribution, while smaller values of p_0 shift the weight of both distributions to the left.

For immigrants, mean share of coworkers who are immigrant for employer size s is:

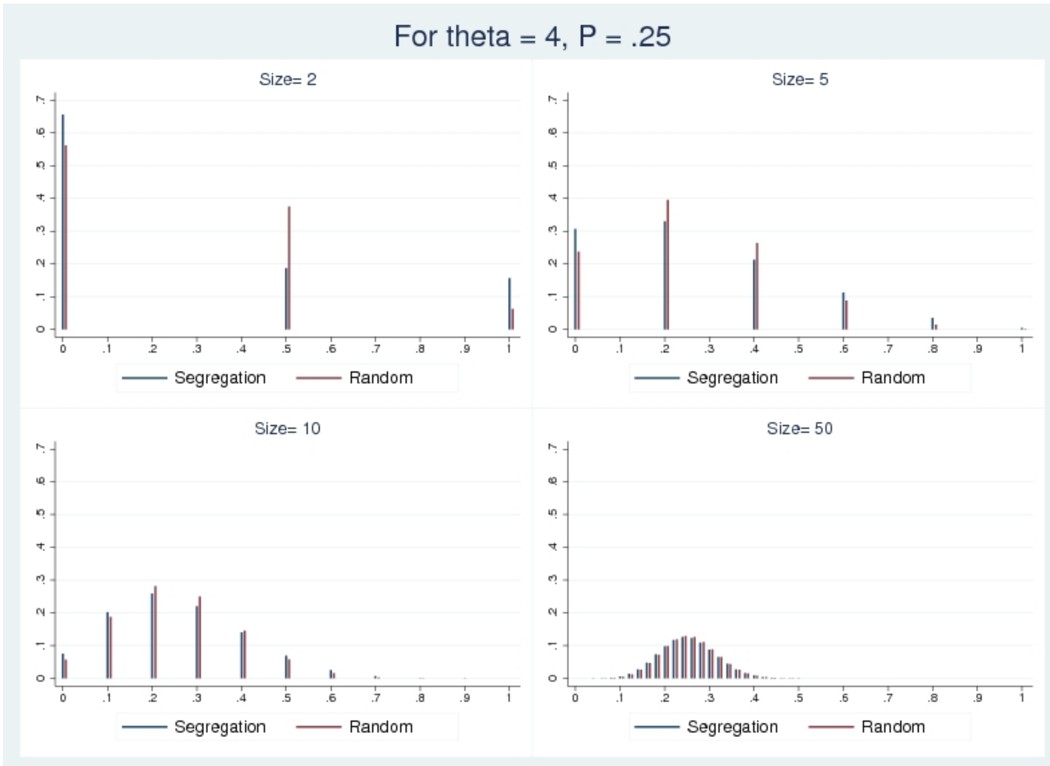
$$E(cw_I|s) = \sum_{i=0}^{s} \left(Pr(i, s|p_0, \theta) * i * \frac{i-1}{s-1} \right) \tag{G.5}$$

and for natives,

$$E(cw_N|s) = \sum_{i=0}^{s} \left(Pr(i, s|p_0, \theta) * (s - i) * \frac{i}{s-1} \right) \tag{G.6}$$

The difference is then:

Figure G.2: Immigrant share distribution with and without segregation



$$E(cw_N - cw_I|s) = \sum_{i=0}^{s} \left(Pr(i, s|p_0, \theta) * \frac{i(i-1)-(s-1)i}{s-1} \right) \qquad \text{(G.7)}$$

Figures G.3 to G.5 plot out the relationship between employer size and coworker means for various values of the immigrant share of the overall workforce p(different colored lines in each graph), using segregation parameter $\theta = 4$. Figure G.3 graph gives the mean by firm size for immigrants, figure G.4 is for natives, and figure G.5 gives the difference between them. Figure G.6 repeats figure G.5, except that it is parameterized to represent a lower level of segregation ($\theta = 1$).

Examination of these figures makes a couple of patterns clear:

- For very small employers (< 10 employees), the model can generate a large

Figure G.3: Immigrant coworker mean and employer size ($\theta = 4$)

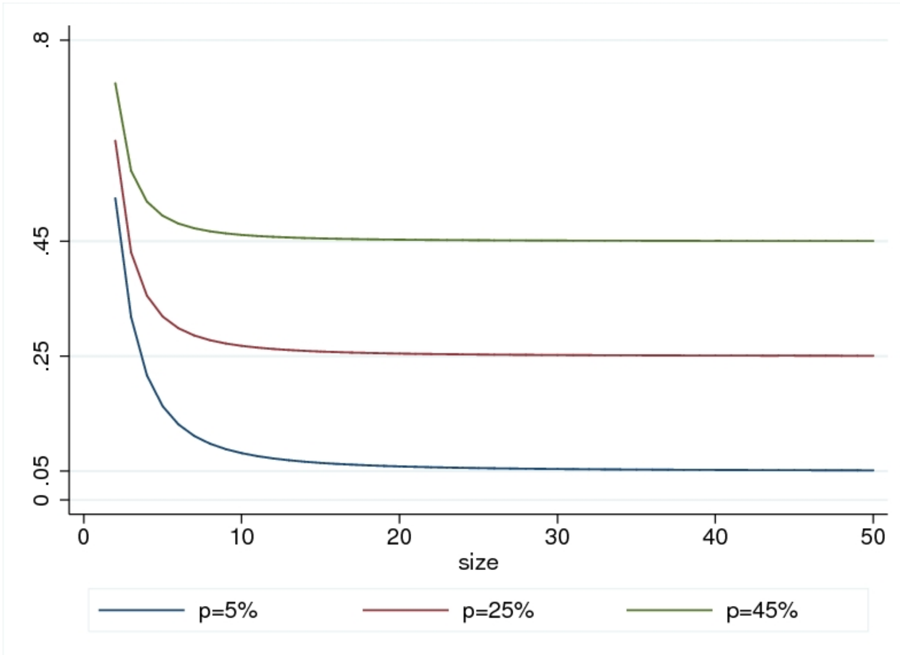


Figure G.4: Native coworker mean and employer size ($\theta = 4$)

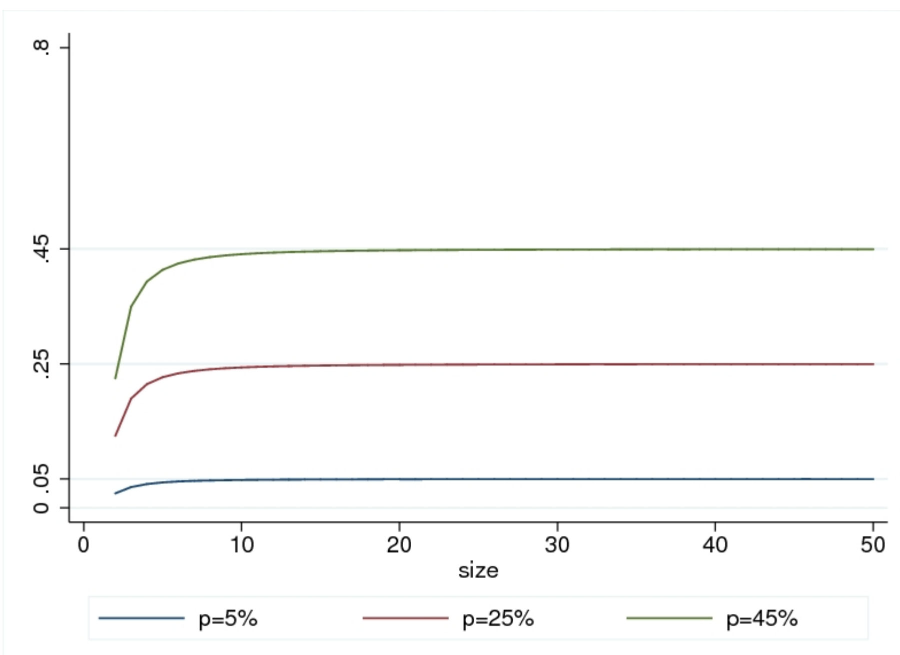


Figure G.5: Immigrant-native difference in coworker mean and employer size ($\theta = 4$)

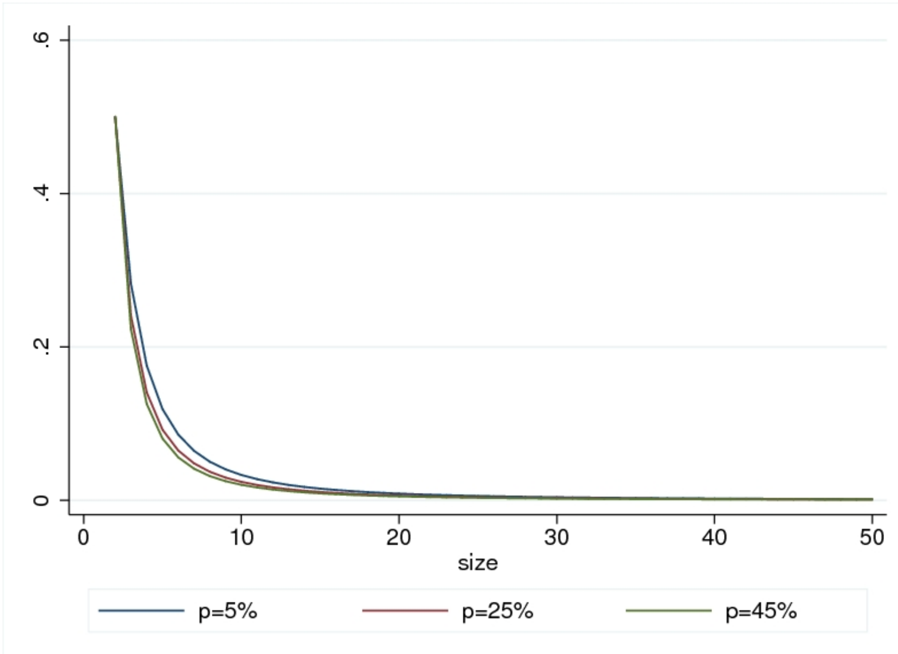
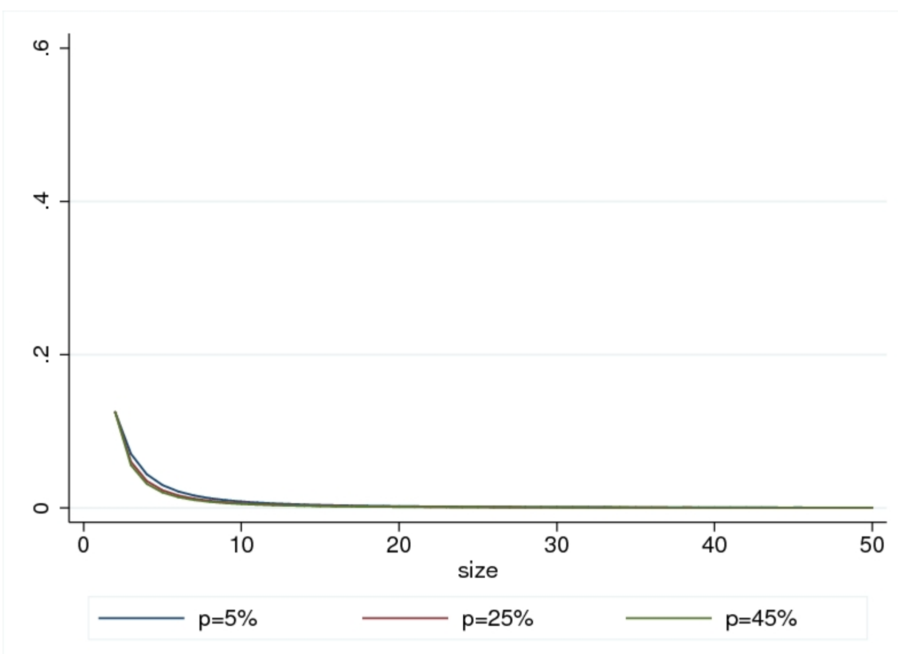


Figure G.6: Immigrant-native difference in coworker mean and employer size ($\theta = 1$)



difference in coworker means, even with a relatively mild tendency to segregate.

- Even for large theta, this model generates essentially no segregation in large firms.

Table G.1: Characteristics of Weighted Matched Sample

Weighted	Immigrants Recent	Immigrants Established	Native
Coworker Share	42.0	36.1	14.3
Age			
Age<30	44.6	19.8	30.9
30<Age<40	35.7	34.3	28.6
Age>40	19.7	45.0	40.0
Male	56.9	62.4	51.9
Age at arrival (*)			
<12	0.9	12.3	.
12-25	36.3	51.5	.
26-35	38.4	25.4	.
35+	24.5	10.8	.
Establishment size			
2-9	8.4	9.2	7.8
10-49	23.6	23.2	22.9
50-99	13.4	13.4	13.4
100-499	29.1	30.6	29.1
500 or more	25.0	23.1	26.8
Firm has multiple establishments	34.6	36.3	42.6
Establishment age			
0-1 years	13.0	11.3	11.9
2-4 years	26.2	22.2	24.6
Age 5 or more	60.8	66.6	63.6
Sector			
Construction	4.85	5.47	5.71
Manufacturing	18.75	21.70	13.43
Transportation & utilities	2.60	3.77	5.26
Wholesale	6.46	6.73	6.18
Retail	23.99	18.86	22.46
FIRE	3.01	5.14	6.71
Services	40.34	38.34	40.26
Log quarterly earnings	8.15	8.54	8.34
Consecutive quarters on 2000-Q2 job:			
Quarter before AND after	57.5	70.4	64.2
Quarter before OR after (not both)	34.7	24.0	28.0
Neither quarter before NOR after	7.8	5.6	7.8
Immigrant share of workers in residence tract	36.5	36.3	14.2
Neighborhood network index	2.00	1.50	1.78
Shared commute index:			
Immigrant co-commuters	0.13	0.09	0.04
Native co-commuters	0.21	0.17	0.47

Source: LEHD database and author calculations.
(*) Year of application for a SSN is used as a proxy for time of arrival in the U.S.
Note: The unit of observation is a worker. All figures represent percentages. There are 3,549,111 matched workers in total for our group of MSAs.

Table G.2: Linear Regression of Full Specification

Parameter	Estimate
Constant	0.0843
New Immigrant	0.1283
Established Immigrant	0.0784
Neighborhood network index	0.0410
Immigrant share commute index	1.9923
Native share commute index	-0.5932
Log earn	0.0020
Work Quarter 1 and 2 of 2000	0.0036
Work Quarter 2 and 3 of 2000	0.0016
Full quarter worker	0.0016
College Graduate	0.0013
Graduate Schl.	0.0075
High Sch. Drop	0.0188
High Sch. Grad	0.0024
Poor English	0.0759
Immigrant residential tract	0.1859
Plant age 0-1	-0.0023
Plant age 2-4	0.0039
Number of plants= Multi	-0.0315
30<Age<40	-0.0019
Age<30	-0.0088
Female	0.0023
Plant size 2-4	0.0234
Plant size 5-9	0.0054
Plant size 10-19	-0.0035
Plant size 20-49	-0.0068
Plant size 50-99	-0.0032
Plant size 100-499	0.0045
plant age*Number of plants= 0-1 Multi	0.0009
plant age*Number of plants= 2-4 Multi	-0.0001

Note: All standard errors are below 0.001. Controls in all columns include MSA, detailed industry, employer age and size, worker age and sex, in addition to the variables listed in the table. The reference group is the group who speak English well. The unit of observation is a worker. N=3,549,111. These are the workers with complete information on their residential location.

Bibliography

John M. Abowd, Francis Kramarz, and David N. Margolis. High wage workers and high wage firms. *Econometrica*, 67(2):251–333, Mar. 1999.

John M. Abowd, Paul Lengermann, and Kevin McKinney. The measurement of human capital in the u.s. economy. Lehd technical report, August 2002.

John M. Abowd, Francis Kramarz, and Paul Lengermann. Are Good Workers Employed by Good Firms? A Test of a Simple Assortative Matching Model for France and the United States. Mimeo, January 2004.

John M. Abowd, Julia I. Lane, Kristin Mccue, John C. Haltiwanger, Ron S. Jarmin, and Paul Lengermann. The relation among human capital, productivity and market value: Building up from micro evidence. In *Measuring Capital in the New Economy*, chapter 5, pages 153–198. University of Chicago Press for the NBER, Chicago, 2005.

John M. Abowd, Lars Vilhuber, Kevin McKinney, Kristin Sandusky, Bryce Stephens, Fredrik Andersson, Mark Roemer, and Simon Woodcock. The lehd infrastructure files and the creation of the quarterly workforce indicators. Lehd techincal working paper, US Census Bureau, TP 2006.01 2006.

John M. Abowd, Julia I. Lane, Kevin McKinney, Kristin Sandusky, and John C. Haltiwanger. Technology and skill: An analisis of within and between firm differences. Lehd techincal working paper, US Census Bureau, February 2007.

Fredrik Andersson, John C. Haltiwanger, and Seth Sanders. The role of firms in immigrant assimilation and labor market adjustment. Grant application to the national institute of child health and human development, July 2007.

Fredrik Andersson, Mónica Gárcia-Pérez, John C. Haltiwanger, Kristin Mccue, and Seth Sanders. Workplace concentration of immigrants. mimeo, October 2008.

Pernilla Andersson and Eskil Wadensjó. The employees of native and immigrant self-employed. Iza dp no.3147, IZA, November 2001.

Olof Aslund and Oskar Skans. Measuring conditional segregation: Methods and empirical examples. working paper, Institute for Labor Marker Policy Evaluation (IFAU), 2005a.

Olof Aslund and Oskar Skans. Will i see you at work? ethnic workplace segregation in sweden 1985-2002'[;. working paper, Institute for Labor Marker Policy Evaluation (IFAU), 2005b.

Robert T Averitt. *The Dual Economy: The Dynamics of American Industrial Structure*. Norton, New York, 1968.

Abigail Barr and Abena Oduro. Ethnicity and wage determination in ghana. working paper WPS/2000-9, 2000.

Patrick Bayer, S. Ross, and G. Topa. Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economics*, forthcoming 2008.

Paul Beaudry, Mark Doms, and Ethan Lewis. Endogenous skill bias in technology adoption: City-level evidence from the it revolution. Working paper 12521, NBER, September 2006.

Gary S. Becker. *The Economics of Discrimination*. University of Chicago Press, Chicago, 1957.

George J. Borjas. The labor demand curve is downward slopping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics*, pages 1335–1374, November 2003.

George J. Borjas. Native internal migration and the labor market impact of immigration. *The Journal of Human Resources*, pages 221–258, August 2005.

George J. Borjas. The earnings of male hispanic immigrants in the united states. *Industrial and Labor Relations Review*, 35(3):343–353, 1982.

George J. Borjas. Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics*, 3(4):463–489, October 1985.

George J. Borjas. The economics of immigration. *Journal of Labor Economics*, 32 (4):1667–1717, 1994.

U.S Census Bureau. Census 2000 special tabulations (stp-159). Special tabulation, U.S. Census Bureau, 2005. `http://www.census.gov/population/cen2000/stp-159/foreignborn.pdf` and `http://www.census.gov/population/cen2000/stp-159/native.pdf`.

C Butcher and John E. DiNardo. The immigrant and native-born wage distributions: Evidence from united states censuses. *Industrial and Labor Relations Review*, 56 (1):97–121, October 2002.

Antonio Cabrales, Antoni Calvo-Armengol, and N. Pavoni. Social preferences, skill segregation, and wage dynamics. *Review of Economic Studies*, 75(1):65–98, 2008.

Gilberto Cardenas and Niles Hansen. Immigrant and native ethnic enterprises in mexican american neighborhoods: Differing perceptions of mexican immigrant workers. *International Migration Review*, 22(226242), 1988.

William J. Carrington and Kenneth Troske. Gender segregation in small firms. *Journal of Human Resources*, 30(3):503–533, Summer 1995.

William J. Carrington and Kenneth Troske. Sex segregation in u.s. manufacturing. working paper CES 96-4, June 1996.

Barry. R. Chiswick. The effect of americanization on the earnings of foreign-born men. *Journal of Political Economics*, 86(5):897–921, October 1978.

Barry. R. Chiswick. An analysis of the earnings and employment of asian-american men. *Journal of Labor Economics*, 1(2):197–214, April 1983.

Mary Corcoran, Greg Duncan, and Michael Ponza. A longitudinal analysis of white women's wages. *Journal of Human Resources*, 16(4):497–520, Autumn 1983.

T. Dunne, Lucia Foster, and John C. Haltiwanger. Wage and productivity dispersion in us manufacturing: The role of computer investments. *Journal of Labor Economics*, 22(2):397–429, 2004.

Richard C. Edwards. The social relations in the firm and labor market structure. In M. Reich R.C. Edwards and D. M Gordon, editors, *Labor Segmentation*, pages 3–26. Mass: Heath, Lexington, 1972.

James R. Elliot. Referral hiring and ethnically homogeneous jobs: How prevalent is the connections and for whom? *Social Science Research*, 30(1):401–425, 2001.

John K. Galbraith. *The New Industrial State*. Mentor, New York, 1971.

Laura Giuliano and MIchael Ransom. Manager ethnicity and employment segregation. Mimeo, June 2008.

Laura Giuliano, David Levine, and Jonathan Leonard. An analysis of quits, dismissals, and promotions at a large retail firm. Mimeo, July 2006.

Mark Granovetter. *Getting a Job: A Study of contacts and Careers*. The University of Chicago Press, Chicago, 1995. 251 pages.

Erica L. Groshen. How are wages determined? *Federal Reserve Bank of Cleveland*, February 1990.

Erica L. Groshen. Five reasons why wages vary among employers. *Industrial Relations*, 30(3), Fall 1991a.

Erica L. Groshen. Sources of intra-industry wage dispersion: How much do employers matter? *Quarterly Journal of Economics*, 106(3):869–84, August 1991b.

John C. Haltiwanger. Entrepreneurship and job growth. Mimeo, October 2006.

John C. Haltiwanger, Julia I. Lane, and James R. Spletzer. Wages, Productivity, and the Dynamic Interaction of Businesses and Workers. Mimeo, June 2005.

Brian Headd. The characteristics of business owners database, 1992. Ces 99-8, Census Bureau. U.S. Small Business Administration, May 1999.

Judith Hellerstein and David Neumark. Workplace segregation in the united states: Race, ethnicity and skill. Ces 07-02, Center of Economic Studies Discussion Paper, January 2007.

Judith Hellerstein, David Neumark, and McInerney Melissa. Measuring the importance of labor market networks. Working paper, NBER, 2008a.

Judith K. Hellerstein and David Neumark. Ethnicity, Language, and Workplace Segregation: Evidence from a New Matched Employer-Employee Data Set. *Annales D'Économie et de Statistique*, pages 19–78, 2003.

Judith K. Hellerstein, David Neumark, and McInerney Melissa. Spatial mismatch or racial mismatch. *Journal of Urban Economics*, 64(2):464–479, 2008b.

Harry Holzer. Search method use by unemployed youth. *Journal of Economic Literature*, 6(1):1–20, 1988.

Harry Holzer. Hiring procedures in the firm:their economic determinants and outcomes. Nber working paper series. wp-2185, NBER, March 1987.

John Iceland. *Where we Live Now: Immigration and Race in the United States*. University of California Press, 2009.

Yannis M. Ioannides and Linda Datcher Loury. Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, XLII(1):1056–1093, December 2004.

Ron S. Jarmin and Javier Miranda. The longitudinal business database. Report, Census Bureau. U.S. Center for Economic Studies, July 2002.

John Kain. Housing segregation, negro employment and metropolitan decentralization. *Quarterly Journal of Economics*, 82(2):175–95, May 1968.

Michael Kremer and Eric Maskin. Wage inequality and segregation by skill. Working Paper 5718, NBER, August 1996.

Kevin Lang. A language theory of discrimination. *Quarterly Journal of Economics*, 2(101):363–381, 1986.

Paul Lengermann. Is it who you are, where you work, or with whom you work? reassessing the relationship between skill segregation and wage inequality. Technical paper, Census-LEHD, October 2002.

Paul Lengermann, Kevin McKinney, and Robert Pedace. New evidence on immigration and labor market outcomes using linked longitudinal employer-employee data. mimeo, Census-LEHD, April 2004.

Ethan Lewis and David Card. The diffusion of mexican immigrants during the 1990s: Explanations and impacts. Working paper 11552, NBER, August 2005.

Ivan Light. *Deflecting Immigration*. Sage Publications, New York, 2006. 246 pages.

Ivan Light and Steven J. Gold. *Ethnic Economies*. American Press, California, 2000. 302 pages.

Douglas Massey. *Why Does Immigration Occur? A Theoretical Synthesis.* In the Handbook of International Migration: The American Experience, edited by Charles Hirschman, Philip Kasinitz, and Josh DeWind., New York: Russell Sage Foundation, 1999.

Douglas Massey, Alarcon Rafael, Durand Jorge, and Gonzalez Humberto. *Return to Aztlan: The Social Process of International Migration from Western Mexico*. University of California Press, Berkley, CA, 1987.

P.G. Min. Some positive functions of ethnic business for an immigrant community: Korean immigrants in los angeles. *Final Report Submitted to the National Science Foundation*, 1989.

Suzanne Model. Ethnic economy and industry in mid-twentieth century. *Social Problems*, 4:445–463, November 1997.

James Montgomery. Social networks and labor market outcomes: Toward an economic analysis. *American Economic Review*, 81(5):1408–18, 1991.

U.S. Bureau of the Census. 1992 characteristics of business owners. *Government Printing Office*, 1997.

Robert P. Parker and James R. Spletzer. The business establishment list -standard statistical establishment list comparison project. Prepared for the federal economic statistics advisory committee, U.S. Census Bureau, June 2000.

Krishna Patel and Francis Vella. Immigrant networks and their implications for occupational choice and wages. Working Paper 3217, IZA, December 2007.

Alejandro Porter and Kenneth L. Wilson. Immigrant enclaves: An analysis of the labor market experiences of cubans in miami. *The American Journal of Sociology*, 86(2):295–319, 1980.

Seth Sanders and Victor Nee. Limits of ethnic solidarity in the enclave economy. *American Sociological Review*, 52(6):745–773, Dec. 1987.

A. Dianne Schmidley. Profile of the foreign-born population in the United States: 2000. Current population reports, series P23–206, U.S. Census Bureau, 2001. `http://www.census.gov/prod/2002pubs/p23-206.pdf`.

Curtis J. Simon and John T. Warner. Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure. *Journal of Labor Economics*, 10(3):306–330, Jul. 1992.

Audrey Singer. *The Rise of New Immigrant Gateways*. Brookings, Washington, 2004.

James P. Smith and Finis Welch. Black-white male wage ratios: 1960-70. *American Economic Review*, 67(3):323–338, June 1977.

Sandra Smith. Mobilizing social resources: Race, ethnic,and gender difference in social capital and persisting wage inequalities. *Sociological Quartely*, 41(4):509–537, 2000.

James R. Spletzer. Sales data in the ssel and the economic census. Technical report, U.S. Census Bureau, June 1998.

Michael A. Stoll, Steven Raphael, and Harry Holzer. Black job applicants and the hiring officer's race. *Industrial and Labor Relations Review*, 57:267–287, 2004.

Stephen J. Trejo. Why do Mexican Americans Earn Low Wages? *Journal of Political Economics*, 105(6):1235–1268, 1997.

Kenneth Troske. Evidence on the Employer Size-Wage Premium from Worker-Establishment matched data. *Review of Economics and Statistics*, 81(1):15–26, Feb. 1999.

Thierry Volery. Ethnic entrepreneurship: a theoretical framework. *Handbook of research on ethnic minority entrepreneurship*, 2005.

R. Waldinger. Black/immigrant competition re-assessed: New evidence from los angeles. *Sociological Perspectives*, 40:365–386, 1997.

Ed Walker. The census bureau's business register: Basic features and quality issues). Report presented at the joint statistical meetings, American Statistical Association, August 1997.