FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

The Wage Impact of Undocumented Workers

Julie L. Hotchkiss, Myriam Quispe-Agnoli,
and Fernando Rios-Avila

Working Paper 2012-4
March 2012

Abstract: Using administrative, individual-level, longitudinal data from the state of Georgia, this paper finds that a documented worker employed by a firm that hires undocumented workers can expect to earn 0.15 percent less than if employed by a firm that does not hire undocumented workers. However, in sectors where there are opportunities for task specialization and benefits from communication skills, documented workers can expect to earn a wage premium of less than 1 percent from being employed at a firm that also hires undocumented workers.

JEL classification: J30, J15

Key words: illegal immigrants, immigration, unauthorized immigration

---

The paper has benefited from comments and suggestions from George Borjas, J. David Brown, Curtis Florence, Jerry Gonzalez, Gordon Hanson, Robert E. Moore, Pia Orrenius, Giovanni Peri, Debra Sabia, Eric Seiber, Chad Sparber, Madeline Zavodny, and seminar participants at the Departments of Economics at the Georgia Institute of Technology, the University of Kentucky, the University of Georgia Department of Public Administration and Policy, Claremont-McKenna College, and the University of California-Riverside. The research assistance of Gustavo Canavire, M. Laurel Graefe, Gustavo Uceda, and Eric Wang is much appreciated. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Julie L. Hotchkiss (contact author), Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-8198, julie.l.hotchkiss@atl.frb.org; Myriam Quispe-Agnoli, Community and Economic Development, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-8930, myriam.quispe-agnoli@atl.frb.org; or Fernando Rios-Avila, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-7995, fernando.rios@atl.frb.org.

The Wage Impact of Undocumented Workers

I. Introduction and Background

The United States has a long history of immigration debate. Through the last century and into this one, immigration policy has been subjected to changing economic needs, fears, and political whims. Positive contributions of immigration have been identified by Neal and Uselding (1972) who estimate that the flow of immigrants into the United States between 1790 and 1912 resulted in a 13 to 42 percent higher level of capital stock than would have prevailed in the absence of immigration during these years (also see Barro and Sala-i-Martin 1995 and Chiswick et al. 1992). Immigration has also been more recently explored in various countries as a mechanism for replacing retiring baby-boom workers (e.g., Hamada and Kato 2007, Hotchkiss 2005, Denton and Spencer 1997).

Concerns surrounding immigration are rooted in an expectation that the arrival of new workers into a labor market would displace native workers and/or put downward pressure on wages. The purpose of this paper is to investigate the impact on wages of the presence of a specific class of immigrants -- undocumented workers. The literature presents a wide range of estimates of the effects of immigration on wages and employment of native workers, but little is known about the impact of undocumented workers. The conventional wisdom has been that a 10 percent increase in the population share of immigrants results in a one to four percent decrease in native wages (for example, see Friedberg and Hunt 1995, Borjas et al. 2006, and Orrenius and Zavodny 2007). The measured impact of immigration on the displacement of workers is less clear. Card (1990), Wright et al. (1997), Butcher and Card (1991), and Card and DiNardo (2000) find no evidence of immigrant inflows affecting native migration patterns or employment outcomes. Whereas, Frey (1996) and Borjas (2005) identify a significant relationship between

immigrant inflows and either native outflows or lower net native in-migration, and Card (2001) finds lower rates of employment within cities with high immigrant arrivals.

More recent evidence from Peri (2009) and Peri and Sparber (2009) suggests that immigrants do not crowd-out employment of native born workers; there is no significant effect on hours worked of native born workers in the short run, but hours significantly increase in the long-run; and that there is no short-run impact on native worker income. However, over time, a net increase of immigrants equal to one percent of employment significantly *increases* income per worker from 0.6 to 0.9 percent. This positive impact on worker income derives from increased efficiency and productivity through task specialization, especially among low-skilled natives (also see Toussaint-Comeau 2007 and Cobb-Clark et al. 1995). In the short-run, capital intensity is decreased as additions to the workforce are from lower skilled workers, but over time businesses expand their capital as they increase production. These conclusions are consistent with those made in earlier work by Barro and Sala-i-Martin (1995) and Chiswick et al. (1992), linking higher levels of immigration to capital deepening and higher per capita consumption.

While estimates of the impact of immigration as a whole on the labor market outcomes of native workers abound, much less is known about the impact of undocumented workers. The reason is the dearth of information about the labor market presence or characteristics of undocumented workers. To a certain extent, the impact of undocumented workers can be expected to be similar to that of immigrants as a whole; however there are some important differences between the two groups of workers. First of all, the number of undocumented workers in any labor market is only a fraction of the total number of immigrants, suggesting the impact, in either direction, would be much weaker. Second, undocumented workers are likely to be even more limited in their opportunities and therefore have lower elasticities of labor supply

(see Hotchkiss and Quispe-Agnoli 2009). This would tend to make them an even less expensive factor substitute for native labor of similar skill. This lower elasticity of labor supply will also have implications for wage differentials between documented and undocumented workers. The more concentrated undocumented workers are in an industry the greater is the opportunity for firms to exercise monopsony power and keep wages of undocumented workers low. And, thirdly, certain skills, such as communication, are likely to be more lacking in undocumented workers (than in immigrants in general). And, according to Peri and Sparber's (2009) model, the presence of undocumented workers with limited communication skills would provide opportunities for even low-skilled native workers (or their employers) to shift the native skill contribution to production toward those that are more highly rewarded (specializing in tasks requiring greater communication skills).[1]

The analysis in this paper makes use of longitudinal, administrative, individual-level data from the state of Georgia to investigate how the presence of undocumented workers affects the wages of documented workers. Controlling for individual and firm level fixed effects, the results indicate that workers employed by single-establishment firms who hire undocumented workers can expect to earn wages about 0.15 percent lower than they would at a firm that does not employ undocumented workers, but the effect does vary across sectors, with evidence of a wage *premium* in low-skilled sectors in which communication skills could be valuable.

A. Immigration Policy

Immigration legislation dates from the founding of the nation.[2] The two most recent Federal efforts to address concerns of undocumented immigration are the Immigration and Control Act (IRCA) of 1986, and the Illegal Immigration Reform and Immigrant Responsibility

[1] The importance of communication skills in occupational mobility is highlighted by Kossoudji and Cobb-Clark (2000) who find that deficiency in English severely limits occupational mobility of undocumented workers.
[2] For historical details, see CBO (2006) and FAIR (2007).

Act (IIRIRA) of 1996. Both of these laws were passed in response to the growing population of unauthorized immigrants identified at the time, however they were quite different in their approaches to addressing the concerns spawned by that growth. Whereas the IRCA is best known for creating two amnesty programs for unauthorized immigrants, the focus of the IIRIRA was one of border enforcement (see Fix and Passel 1994 and Nuñez-Neto and Viña 2006 for more details).

Since the terrorist attacks of 2001, and in response to continued dramatic growth in unauthorized immigrants, there have been renewed calls for additional comprehensive immigration policy reform. The absence of forthcoming Federal legislation has been the likely motivation of many states to pass state-level laws targeted at unauthorized immigrants. The number of laws enacted has grown from 39 in 2005 to 208 in 2010 and 197 in 2011.[3] Fifteen additional bills were passed out of legislatures in 2011, but vetoed by governors. The first major immigration legislation in Georgia became law in July 2007 and the second in 2011. The analysis in this paper makes use of data through 2006, so the relatively recent change in the legal environment in Georgia will not confound the current analysis.

B. Identifying Unauthorized Immigrants

Identifying unauthorized immigrants is the greatest challenge in investigating their impact. The most common method used to estimate the number of unauthorized immigrants is the residual approach, or merely calculating the difference between the total measured foreign-born population and the legal immigrant population (see Hanson 2006). According to the latest figures, there are 11.2 million unauthorized immigrants living in the U.S. as of March 2010 (Passel and Cohn 2011). It is also estimated that about four percent of the total are located in

[3] See the National Conference of State Legislatures website, "Issues and Research: Immigration," <http://www.ncsl.org/issues-research/immig/state-laws-related-to-immigration-and-immigrants.aspx>.

Georgia. Between 2000 and 2010, Georgia experienced one of the largest percentage increases of unauthorized immigrants in the U.S. -- 70 percent (Passel and Cohn 2011).

A second data source that has been used to look at unauthorized immigration is information on border apprehensions from the U.S. Border Patrol. Estimating the level of unauthorized immigration using apprehension data is problematic, primarily because it is not only a function of the number of attempts to cross the border (which have been shown to vary with expected relative U.S./Mexico economic conditions), but also a function of the enforcement efforts of border patrol and a function of the number of attempts (see Hanson and Spilimbergo 1999, and GAO 2006).

According to DHS estimates for January 2009, 62 percent of unauthorized immigrants come from Mexico, as compared to 55 percent in January 2000. Therefore it is not surprising that surveys from Mexico constitute a third source of data on unauthorized immigrants. The Mexican Migration Project (MMP) is a household survey conducted during the winter months when seasonal migrants return to Mexico. The Legalized Persons Survey (LPS) is a survey of unauthorized immigrants who were granted permanent legal residence in the U.S. under the amnesty provision of the Immigration and Control Act of 1986. In general, the MMP and LPS have been found to be more useful in characterizing undocumented immigrants than actually counting them. Orrenius and Zavodny (forthcoming), using the MMP, report that over the period between 1980 and 2004, approximately 62 percent of migrants from Mexico were unauthorized.

Among the newest sources of data of information about immigrants is the New Immigrant Survey (NIS). The data set now includes two waves of new legal permanent residents in the U.S., admitted in 1996 and 2003, and over-samples employment based immigrants. The

immigrants are administered three surveys over a 12-month period and are asked a host of questions about their original entry into the U.S. and about their experiences since arriving. Jasso (2011) reports that roughly 40 percent of new legal immigrants in 2003 had some experience of being in the U.S. illegally at some time before attaining legal status.[4] The percentage whose spell of illegality is most likely to have more immediately preceded legalization is about 12 percent (Jasso 2011: Table 6). This does not mean that 12 (or even 40) percent of the undocumented workers eventually become documented, however, since those who obtain legal status are going to be a very select group of those who initially entered illegally (Jasso et al. 2000, p. 136).

This paper differs in the way in which unauthorized individuals are identified. In addition, it is not the goal of this paper to obtain an accurate count of unauthorized immigrants, but to identify a reasonable sample with which to perform statistical analyses of labor market outcomes. State administrative data are used to identify invalid social security numbers used by employers in reporting worker earnings. It is a common misconception that all undocumented workers are working "off the books." There is considerable evidence that many employers report, either knowingly or unknowingly, and pay taxes on the wages paid to undocumented workers.[5] Unlike most other studies, the measure used here does not capture the supply of undocumented workers, but, rather, the demand, as the workers are identified through employment records. The advantage of this data source is that it is not subject to sample selection issues plaguing survey results. The disadvantage is that it does not capture

[4] Jasso et al. (2008) estimate that 32 percent of new adult immigrants granted legal permanent residence in the U.S. in 1996 had originally arrived in the U.S. illegally

[5] The Social Security Administration keeps track of wages reported by employers but cannot be matched to a valid name or SSN. This repository of unmatched wages is referred to as the Earnings Suspense File (ESF). It is widely agreed that the exponential growth in the ESF is attributable to the growth in unauthorized immigrants. For tax years 2001 and 2002 alone, 1.8 billion dollars were placed into the ESF.

undocumented workers not reported on employers' payrolls. However, the result is a sample of undocumented workers that represents about 20 percent of all undocumented workers in the state of Georgia.

II. Data

The primary data used for the analyses in this paper are the Employer File and the Individual Wage File, compiled by the Georgia Department of Labor for the purposes of administering the state's Unemployment Insurance (UI) program. These data are highly confidential and strictly limited in their distribution. The data are available from the first quarter of 1990 through the fourth quarter of 2006. The Employer File provides an almost complete census of firms in the U.S., covering approximately 99.7 percent of all wage and salary workers (Committee on Ways and Means 2004).[6] The establishment-level information includes the number of employees, the total wage bill, and the NAICS classification of each establishment. The Individual Wage File, which links individual workers to their employer, is used to construct workforce characteristics at the firm level. We take advantage of the longitudinal nature of the data to calculate the firm's age, employment variability, turnover rates, and worker tenure. The data also contain a 6-digit NAICS industry code and the county of location, allowing us to construct or merge in various industry- and county-level indicators.

We restrict the analysis to single establishment firms for two reasons. First, workers are only linked to the firm in which they are employed. If a firm has multiple establishments, we do not know at which establishment the worker is employed; nor do we know exactly the physical location of the firm, as the address in the file could correspond to the firm headquarters, physical location, mailing address, etc.; nor do we know, if a firm employed undocumented workers, in

[6] Certain jobs in agriculture, domestic services, non-profit organizations are excluded from UI coverage; excluded workers are not represented in the data.

which establishment those workers are employed or who are those undocumented workers' documented colleagues. These problems of measurement error don't arise when we limit the analysis to single establishment firms. The second reason we restrict the analysis is because it is a clear way to reduce the number of observations without employing some sampling scheme. The full data sample has over 178 million observations, restricting to single establishment firms reduces the sample by about half. Conclusions are only generalizable to single establishment firms.

Regrettably, the data set contains no information about workers' demographics or, more importantly, immigration status. However, again making use of the longitudinal nature of the data, we estimate an individual fixed effects model, allowing us to control for individual characteristics that do not vary over time (e.g., innate human capital, immigration status).

A. Using SSNs to Identify Undocumented Workers

Details of how the SSN is used to identify undocumented workers are contained in Appendix A. The abbreviated version is that there are some easily identifiable ways in which a SSN is determined to be invalid. We conclude that some of those reasons are either errors or the result of incomplete record keeping by the firm. We restrict our identification of undocumented workers to invalid SSN that are more likely to have been generated by the workers -- numbers that look valid, but are not. Workers with invalid SSNs for any other reason are considered neither undocumented nor documented and, thus, are excluded from the analysis; this will clearly undercount the actual number of undocumented workers. However, all workers, regardless of SSN classification, are included in counts of aggregate firm employment.[7]

Figure 1 plots the prevalence of undocumented workers in the seven broadly defined

[7] The only other use of SSN to identify unauthorized immigrants we have found is by Maloney and Kontuly (2010) and Wen and Malony (2011), who identify Individual Taxpayer Identification Numbers (ITINs) in driver's license records to track changes in neighborhood living conditions as these individuals change residences.

sectors with the highest incidences. The concentration of workers in these sectors has also been identified nationally by Fortuny et al. (2007).[8] The pattern of growth is also consistent with Fortuny et al. who estimate that 72 percent of unauthorized immigrants in Georgia arrived in the last 10 years.

[Figure 1 here]

Fortuny et al. (2007) estimate that 4.5 percent of the workforce in Georgia was undocumented in 2004. In our sample 1.0 percent of workers are classified as undocumented in 2004, implying that the sample used for the analysis in this paper is capturing about 22 percent of all undocumented workers in the state of Georgia. This is a respectable representation, given that to be included in the sample all workers have been included on the firm's wage report in the first place, and we are being very conservative in the identification of workers as undocumented. Note that the identification process we use in this paper does not make any assumptions about whether the employer knows a worker is documented or undocumented. In addition, the goal of the conservative identification process is to end up with a sample in which we can have a high degree of confidence that the sample is representative of the undocumented workforce, not to actually count the number of undocumented workers in Georgia. The implication of undercounting the number of undocumented workers present in the labor force has the potential to undermine our ability of identifying a statistically significant systematic effect of their presence on documented worker wages, hence likely underestimating any measured effect.

B. Are Undocumented Workers Correctly Identified?

There are several reasons we are confident that the sample of undocumented workers is representative. First of all, the rate of growth seen in both the number and percent of

[8] Fortuny et al. (2007) estimate that nationally in 2004 the percent of workers in leisure and hospitality and construction that was undocumented was 10 percent each, nine percent of workers in agriculture, and six percent each in manufacturing, professional and business services, and other services. Also see Pena (2009).

undocumented workers identified in Georgia matches closely the rate of growth in the Social Security Administration's (SSA) earnings suspense file (ESF). The ESF is a repository of social security taxes paid by employers that cannot be matched to a valid name or SSN. It is widely believed that this growth in the ESF reflects growing incidence of unauthorized work in the U.S. (Bovbjerg 2006).

Figure 2 plots the number of workers (panel a) and the percent of workers (panel b) identified as undocumented along with the size of the ESF. This figure shows a remarkable consistency between the growth seen in workers identified as undocumented and the ESF.

[Figure 2 here]

Data from Census and Homeland Security suggest that between 40 and 60 percent of Mexicans in the U.S. are undocumented, and that 61 percent of unauthorized immigrants come from Mexico.[9] Clearly not all Hispanics are undocumented, or vice versa, however using weighted data from the Current Population Survey (CPS), we calculate the average annual growth in total workers and total number of foreign born, Hispanic workers in the U.S. and in Georgia in order to compare growth rates to those in our sample. These results are reported in Table 1. The workforce in GA grew faster over the period than the U.S. workforce (2.9 percent vs. 1.5 percent, respectively). In addition, the number of foreign born, Hispanic workers in the U.S. grew faster (eight percent per year) than the overall workforce; this phenomenon has been documented by others (Passel and Cohn 2009). But most importantly for our purposes is that the growth rate of foreign born, Hispanic workers in GA (roughly 27 percent per year), which is much larger than in the U.S. overall (also see Passel and Cohn 2009), is similar to the growth in

[9] The 2008 ACS estimates that 11.4 million people in the U.S. were born in Mexico (http://www.census.gov/population/www/socdemo/hispanic/cps2008.html). The DHS estimates that 7.03 million undocumented workers from Mexico were in the U.S. in 2008 (http://www.dhs.gov/xlibrary/assets/statistics/publications/ois_ill_pe_2008.pdf).

the number of workers in GA classified here as undocumented. We also observe a similarly large growth rate in the number of foreign born, Hispanic workers with less than a high school degree (21%), among which we might expect a larger share of undocumented workers than among foreign born, Hispanics in general.

[Table 1 here]

The close match in growth rates in the number of workers classified as undocumented with that of the SSA ESF and with the number of foreign born, Hispanic workers in Georgia as measured by the CPS, suggests that the mechanism employed in this paper to identify undocumented workers is accurate; it's clear that not all undocumented workers are being captured in the data, but likely represent the tip of the ice burg of hiring behavior of any firm. Any remaining mis-classifications will show up in the error term and limit the estimation in its ability to identify any systematic relationships between wages and the presence of undocumented workers.

C. What Do Firms Know and Does it Matter?

A natural question arises as to whether an employer knows when he/she is hiring an undocumented worker, and, more importantly, whether that knowledge has any implication for interpretation of the results in this paper. If the undocumented worker is perfectly indistinguishable from documented workers then the only expected impact on wages is what would result from the increase in the supply of a substitute factor input--wages will fall (Borjas 2009). However, if the employer is able to identify the new workers as undocumented, and, thus, have limited employment opportunities (e.g., Bohon, et al. 2008) and are likely to accept a wage lower than his productivity (e.g., Hotchkiss and Quispe 2009), then there is room for

overall productivity gains and rents to either be enjoyed by the employer or shared with documented workers.

There is reason to expect that employers have a fairly good idea when a worker is undocumented. Up to 60 percent of Mexicans in the U.S. are undocumented (see footnote 9), and, thus, ethnic Hispanic characteristics and limited English skills are features employers can use to identify which workers are likely undocumented; there is no need to carefully scrutinize the presented SSN to determine with a high degree of accuracy whether a worker is undocumented. A firm's willingness, then, to hire undocumented workers will be a function of the expected benefit from hiring versus the expected cost of breaking the law. These benefits and costs are likely to vary by industry and firm characteristics (such as firm size). On the whole, the expected costs are considered to be relatively negligible, especially for a non-border state. For example, CBO (2010) reports that 91 percent of all apprehensions of unauthorized immigrants occur at the border. In addition, prior to 2006, workforce enforcement did not figure very large in efforts to combat unauthorized immigration (CBO 2006, also see Jordan 2011).

A firm's decision to hire undocumented workers, then, would depend on the assessments of costs and benefits to their own economic outcome and, simply, the ethics of the person making the hiring decision. There is a possibility that firms that hire undocumented workers also have a higher propensity to break other laws; it's unclear how this propensity might be expected to affect wage determination policies.

D. Sample Means

Table 2 presents some sample means for workers classified as documented. While worker and firm longitudinal characteristics are calculated beginning in 1990 (the first year of available data), estimations are performed on the sample period 1995-2005 to focus more acutely

on the greater prevalence of undocumented workers during this time period, and because some geographic regressors are only available beginning in 1995. In addition, estimation is performed on single-establishment firms only. Given that there are over 178 million observations in the full sample, estimation with high order fixed effects is cumbersome, at best (see Abowd et. al 1999). Restricting the estimation to single-establishment firms allows us to reduce the structure of the sample in a more predictable way than a random sampling of workers; wage variation that is correlated with whether a firm is multi-establishment or not will be lost. However, since this paper is able to make use of the population of workers (at least the population employed in single-establishment firms), the estimates will not suffer from the attenuation bias highlighted in Aydemir and Borjas (2010). The first column of Table 2 contains means for the full sample, the second column contains means for single-establishment firms only, and the third and fourth columns contains means only for single-establishment firms that do and do not hire undocumented workers.

[Table 2 here]

There are some characteristic differences between single establishment firms and all firms; these differences should be kept in mind when interpreting the generalizability of the results. First of all, the worker level characteristics are all quite similar across both types of firms, although workers in single establishment firms have moderately less tenure and work experience. There is also somewhat greater variation in quarterly earnings among workers employed in single establishment firms.

Single establishment firms appear to be smaller and older and experience less churning than other firms. The biggest difference comes in the distribution of workers across broad industry sectors. There seems to be a greater share of workers employed in single establishment

firms found in construction, wholesale trade, professional and business services, and other services. There is a noticeably smaller share in retail trade and education and health services.

Turning to differences across firms that hire and don't hire undocumented workers, firms that hire tend to be larger, experience greater churning, and are likely to operate in industries that employ fewer workers with college degrees. The greater churning among firms that hire undocumented workers is consistent with earlier findings by Morales (1983) who suggests that these firms need greater workforce flexibility.[10] Workers in these firms also exhibit shorter tenure, lower wages, and are more likely to either be newly hired or separating from their employer.

Among firms that employ undocumented workers, workers are more concentrated in industries characterized as low-skill and less concentrated in high-skill industries, compared with workers employed by firms not hiring undocumented workers. The share of workers employed by firms who hire undocumented workers is much larger in agriculture, construction, manufacturing, and leisure and hospitality.

III. Empirical Specification

A number of different approaches have been taken to quantify the impact of immigration on native worker wages and employment. The most common strategy is used by Altonji and Card (1991) and in a number of papers by George Borjas (alone and with co-authors; 2003, 2005, 2006). The procedure makes use of decennial census data and standard linear regression to identify a relationship between the difference in the density of immigrants on wages or employment across geographic areas (usually metropolitan statistical areas, MSAs). Various

[10] Churning is measured as the difference between worker flows and job flows divided by the average employment during the period. Worker flows is the sum of hires and separations and job flows is net employment change. $CHURN_{jt} = \frac{[Hires+Separations]-[|N_{jt}-N_{jt-1}|]}{[(N_{jt}+N_{jt-1})/2]}$, N_t is the number of workers in time *t* (Burgess et al. 2001).

techniques (e.g., instrumental variables and fixed effects) have been employed to control for the endogeneity problem of immigrants selecting their geographic destination based on observed wages in those locations.

The issue of reverse causality through selection has been a significant concern of analyses estimating the impact of immigration on wages since, as it relates to the analysis in this paper, one can reasonably expect that firms paying higher wages will attract more undocumented workers, so the share of undocumented workers becomes a function of the wage paid by the firm. Regarding this concern, the analysis in this paper has two advantages. First, the estimation includes as a regressor the percent of K-12 students that is Hispanic in the county in which the firm is located in order to capture potential supply effects from the firm's wage setting policy. Second, the estimation is performed at the individual level, where individual worker wages are modeled as a function of firm behavior. It's not likely that a firm bases its decision on whether to hire undocumented workers, or how many to hire, based on the observed wage of any single documented worker.

Endogeneity through omitted variables has also plagued studies of immigration effects on wages. For example, if there are unobservable factors that both increase a documented worker's wage and the probability that a firm hires undocumented workers (or hires more of them), this will result in spurious positive correlations between a worker's observed wage and a firm's observed hiring behavior (and vice versa). Even though the data are not rich enough to allow for any attempts at instrumental variables estimation, all equations are estimated with as many other firm characteristics as possible that will proxy for unobservable firm characteristics. In addition, estimations will control for both individual fixed effects, which will account for time-invariant unobservable individual heterogeneity, and firm level fixed effects, which will control for time-

invariant correlation between a firm's wage setting policy and its tendency to hire undocumented workers.[11]

The primary advantage of the analysis in this paper is that it makes use of matching techniques in order to construct a synthetic control group for workers employed by firms that also employ undocumented workers. The goal is to obtain a control as similar to the treated worker as possible, and controlling for as much heterogeneity as possible. This alleviates concerns of selection by workers toward or away from firms with certain employment practices, allowing us to draw causal inferences on the implications for working in a firm that employs undocumented workers. The analysis also includes proxies to control for supply of undocumented workers, total labor supply, and local labor market conditions through the unemployment rate. The use of individual worker data and the fixed effects means the results are interpreted as within worker, within firm effects and won't suffer from composition bias, which may plague analyses at the broad geographic, industry, or firm level.

A. The Estimating Equation

The baseline estimating equation is specified as:

$$\ln w_{ijt} = \beta_0 U_{jt} + \beta_1 U_{jt} P_{jt} + \beta_2' x_{it} + \beta_3' y_{jt} + \delta_i + \theta_j + \tau_{ijt}, \quad (1)$$

where $\ln w_{ijt}$ is the log of the quarterly earnings of documented worker *i* at firm *j* at time *t*; U_{jt} is equal to one if any undocumented workers were identified working in firm *j* at time *t*, zero otherwise; P_{jt} is the percent of workers at firm *j* that is undocumented at time *t* (this is undefined if $U_{jt} = 0$, so it enters interactively); x_{it} are individual characteristics expected to influence the observed base wage level; y_{jt} are firm level characteristics expected to influence worker *i*'s observed wage; δ_i is an individual fixed-effect; θ_j is a firm fixed-effect; and τ_{ijt} is the random

[11] Malchow-Møller et al. (2007) also use individual level, panel data to investigate the impact of immigrants on native wages in Denmark.

error. The equation will also include a broad sector industry fixed-effect when estimated on the full sample. The model will also be estimated across categories of workers differentiated by sector.[12] Year and quarter fixed effects are included to control for cyclical and seasonal variation in wages. We also include interactions of sector and year fixed effects to control for the possibility of increasing saturation of undocumented workers across sectors over time.

Worker tenure and labor market experience is expected to positively influence wages (at decreasing rates) through the presence of firm- and general-specific human capital (Campbell 1993 and Altonji and Shakatko 1987). Indicators for whether the worker is newly hired at the firm (not employed by the firm in the four preceding quarters) or is separating (not employed by the firm in the following four quarters) are also included; these workers are likely to not have received a full quarter's worth of wages.

Firm size (measured by log employment) is included with the expectation that larger firms pay higher wages (Oi and Idson 1999). Firm age (and its square) is also included, but the relationship between firm age and wages paid is less straightforward (Brown and Medoff 2003). A firm level measure of worker churning (among documented workers only) is included as a measure of employment cost, which might suggest lower wages at firms with greater churning (Burgess et al. 2001, see footnote 10). By merging in additional data, we are able to assign rough measures of three or four digit industry level worker skill and labor intensity of the production process (see Appendix B). Firms who employ higher educated workers are expected to pay higher wages, and we would expect to observe lower wages for workers employed at firms with a more labor intensive production process, ceteris paribus. This is because capital is

[12] One might expect that the firm fixed-effect would capture the sector effect, as a firm's industry would not be expected to change over time. However, firms do change industry from time to time as their primary business activity might change over time. In addition, given the heterogeneity on characteristics and presence of undocumented workers in each industry, we might expect for the effect of undocumented workers to also differ across industries.

typically found to be complimentary with higher skilled labor (for example, see Krussel et al. 2000) suggesting that a production process that uses less capital (one that is labor intensive) will also employ lower skilled workers.

The unemployment rate in the county in which the firm is located and the population density of the county are included to capture overall worker demand and strength of consumer demand, along with potential alternative employment opportunities in the worker's geographic area. We expect to see higher wages being paid by firms facing higher product demand and tighter labor markets. And, as mentioned before, the percent of students in the firm's county that is Hispanic is also included to capture the potential supply of undocumented workers to the firm, although, clearly, not all Hispanics are undocumented workers, and vice versa.

With the specification estimated in this paper, the focus is on the workers' experiences within their employer; the analysis will have nothing to say about a more aggregated effect of undocumented workers at the industry or geographic level. Other factors that are expected to modify the impact of undocumented workers on wages, such as a firm's "hiring intensity" (or, how often the firm hires undocumented workers), are captured by the firm fixed-effect. Unfortunately, there is rarely a clearly defined pre and post time period in which firm hires undocumented workers, making a difference-type analysis fruitless.

One question that presents itself in considering the impact of undocumented workers is what happens to workers who might be displaced when his/her employer begins hiring undocumented workers. The analysis in this paper does not speak to this question. Other work has compared the separation behavior of undocumented workers with that of documented workers (see Hotchkiss and Quispe-Agnoli 2009), but a full analysis of long-term labor market

outcomes of potentially displaced documented workers will be the subject of a future investigation.

B. Flexible Functional Form

There is strong empirical evidence that relationship between wages and a firm's undocumented worker hiring behavior (U_{jt} and P_{jt}) might not be linear, as depicted in equation (1). Figure 3 plots the coefficients on an estimated version of equation (1) that regresses workers' log wage on a set of dummy variables reflecting their firm's workforce share of undocumented workers for the full sample, agriculture, construction, and leisure & hospitality.

[Figure 3 here]

Several things are apparent from this figure. First, the relationship between wages and a firm's undocumented workforce share is likely to vary across sector. Second, that relationship is also likely to vary across workforce share, with the relationship appearing to change somewhere between five and 10 percent. And, third, the relationship appears to be quite small, especially at small workforce shares. In order to allow as flexible a relationship as possible between wages and the hiring behavior/intensity of firms, equation (1) is estimated as, at most, a six-order polynomial with regard to the interaction term ($U_{jt}P_{jt}$); results are reported for the polynomial order in each sector (and the full sample) that best fits the data, based on the Akaike information criteria.[13]

Since the baseline model assumes that the presence of even one undocumented worker (or the presence of one invalid SSN) in any quarter means the firm employed undocumented workers during that quarter, it is possible that there may be a certain degree of measurement

[13] While it is typical to limit functional flexibility to three polynomial terms, we found a significant difference in results between the specification with three and six terms, however, no difference in conclusions beyond the inclusion of six terms (up to nine). In addition, full non-parametric estimation of this relationship is precluded by the desire to control for both firm and individual fixed effects in the estimation.

error in identifying a firm that employs undocumented workers if the reporting and recording of SSNs is error prone. We find that allowing for the greater flexibility in functional form reduces to negligible the sensitivity of results to more rigid definitions of undocumented worker employment (e.g., requiring a certain workforce percentage before declaring the firm employs undocumented workers).

C. Matching

The goal of the empirical analysis is to determine what the implications for a documented worker's wage is of working in a firm that employs undocumented workers compared to working in a firm that does not employ undocumented workers. To be able to draw causal inference the estimation needs to be free of any individual selection by workers toward or away from firms that employ undocumented workers and to control for as much worker and firm heterogeneity as possible. A technique that has been often used in the absence of a random experiment or appropriate instruments, is to construct a synthetic control group through matching techniques (see Imbens 2004).[14] The goal is to find a control observation for each worker who has been "treated" by working in a firm that employs undocumented workers. This control will be as similar to the treated worker in as many dimensions as possible, *except* the control will be employed in a firm that does *not* hire undocumented workers. Details of the matching methodology and results are contained in Appendix C.

IV. Estimation Results

A. Matching

Appendix C contains the estimation results from the propensity score matching exercise, as well as evidence that the matching strategy was successful in that it significantly reduced the

[14] Heyman et al. (2007) use matching techniques and a linked employer-employee data set to investigate the presence of a foreign ownership wage premium among workers in Sweden.

bias between the means of workers employed by firms who hire undocumented workers and by firms that do not. Appendix C also contains logit estimates corresponding to the propensity of a worker to be observed working in a firm the employs undocumented workers. Estimates for the full sample are presented; those for each sector separately are available upon request. Results are fairly consistent across sectors. A worker is more likely to be employed in a firm that hires undocumented workers if the worker has lower tenure; less labor market experience; is a new or separating worker; and is employed in a small, young firm with a high amount of churning. Being employed in a county with a larger share of Hispanic students in the school system, greater population density, and lower unemployment increase a worker's chances of being employed by a firm that also hires undocumented workers.

B. Wage Regression Results

Table 3 contains parameter estimates for equation (1) estimated for the full sample with the full set of firm and individual level fixed effects, as well as an estimation via OLS without fixed effects. Both OLS and fixed effect models are estimated using the same matched sample and functional form. The preferred model, based on the AIC, for the full sample is a sixth order polynomial. Parameter estimates for each sector separately are found in Appendix D.

[Table 3 here]

Focusing first on the fixed effects estimation results and the impact of the control variables, the impact of the worker and firm characteristics on wages earned by documented workers is as expected. Wages are increasing, at a decreasing rate, as tenure and overall work experience increases. Being either a new hire or one about to separate leads to much lower observed wages. The sizes of the estimated coefficients reflect the high probability that these workers are likely to be working less than a full quarter during their employment transition (thus

being observed with considerably lower earnings). Workers can expect to earn higher wages if employed at larger and younger firms, firms that employ workers with more education, and firms that are less labor intensive. Higher wages are also found among workers employed at firms with higher levels of churning. The churning result suggests that firms with high levels of churning may pay an efficiency wage to reduce churning, which can be costly.

Workers facing a higher unemployment rate in the county and quarter in which their employer is located can expect to earn a lower wage, but wages are higher in more densely populated counties. In addition, wages are significantly lower in counties with higher levels of Hispanic populations.

The importance of controlling for individual and firm level fixed effects is apparent even before considering the impact of the presence of undocumented workers on wages. While most of the point estimates across the two columns are similar in sign and magnitude, there are some exceptions. For example, the estimated impact of tenure is roughly four times larger when one does not control for individual fixed effects. In addition, the relationship between churning and wages is negative when firm level fixed effects aren't included, the OLS results would suggest that workers earn a higher wage working in a county with a higher share of Hispanics, and the role of the unemployment rate and population density is much larger in the absence of individual and firm level fixed effects.

Before turning to the overall marginal effect of working at a firm that employs undocumented workers, further evidence of the importance of controlling for fixed effects can be seen in the first row of Table 3. Without considering the interactions with the share of workers undocumented, the coefficient on whether or not the firm employs undocumented workers is a large negative number when estimated without fixed effects, and small and positive when fixed

effects are included; controlling for firm and individual fixed effects is clearly important.

Table 4 presents the average marginal effect on a documented workers' wages of being employed at a firm that also employs undocumented workers, relative to being employed at a firm that does *not* hire undocumented workers. Since the relationship between wages and the undocumented workforce share is highly non-linear in some sectors (and the full sample), these marginal effects are presented at different points of the distribution of workers across firms' undocumented workforce shares. The undocumented workforce share at each of the respective points of the distributions are in brackets below the marginal effect.

[Table 4]

For the full sample of workers across all sectors, a documented worker at a firm with an average undocumented workforce share (3.24%) suffers a 0.15 percent wage penalty compared to what that worker would earn if employed by a firm that did not hire any undocumented workers. Figure 4 illustrates the information contained in Tables 4 for the full sample and some select individual sectors. As can easily be seen from Figure 4, the distribution of workers across firms' undocumented workforce share is highly skewed to the right; less so in agriculture and construction. Except in agriculture, half of workers employed at firm who hire undocumented workers in all sectors are employed in firms whose undocumented workforces share is less than five percent. And only in construction (other than agriculture) does that share exceed three percent.

[Figure 4 here]

The marginal effects reported in Table 4 indicate that in most sectors, workers can expect to earn a statistically significantly higher wage if employed by a firm that does not also hire undocumented workers, compared to what he/she would earn at a firm that employs an average

share of undocumented workers. However, the practical magnitude of that difference is questionable. The largest penalty for working at a firm that hires undocumented workers is 1.7 percent in professional and business services. This amounts to about $192 per quarter (the average quarterly earnings of a documented worker in a professional and business services firm *not* hiring undocumented is $11,303). In addition, there are two notable sectors in which documented workers earn a *premium* for being employed at firms that hire undocumented workers. Workers employed at firms in retail trade and leisure and hospitality that hire an average share of undocumented workers will earn a 0.6 and 0.3 (respectively) percent wage premium (roughly $39 and $9, respectively, per quarter) compared to being employed in the same sectors at a firm that does not hire undocumented workers. However, these, too, while statistically significant, are of questionable importance. On the other hand, in retail trade, for example, the premium rises with the firm's undocumented workforce share so that a worker employed in a firm that is in the 90th percentile of workforce share (7.69 percent) will be earning a two percent wage premium ($131/quarter) relative to what he/she would earn in a firm in retail trade not employing undocumented workers (see Figure 4, as well for a visual illustration).

C. A Wage Premium?

For the most part, although statistically significantly different from zero, there is only a negligible impact of working at a firm that also employs undocumented workers. But in some sectors, that impact, although also negligible, is positive. This seems to be counter-intuitive. However, a positive expected wage impact is consistent with results of Peri (2006) and Peri and Sparber (2009), who conclude that increased immigration leads to increased efficiency and productivity through task specialization. At the firm level this might manifest itself in documented manual laborers being reassigned to tasks that require more communication skills

(or oversight of the new undocumented labor), while the undocumented workers specialize in the manual labor tasks. The net result is a more efficient production process with documented workers (especially low-skill documented workers) being more productive and earning higher wages (also see Orrenius and Zavodny 2007, and Hanson 2007).[15] In addition, Brown et. al (2012) find that firms who hire undocumented workers have a competitive advantage and employers may share rents from this advantage with their documented workforce. Further, the availability of lower cost, low-skill workers could result in a scale effect with firms producing more, using more capital, and hiring workers complementary to capital (documented, higher-skilled workers); see Ottaviano and Peri (2006) and Brown et. al (2010). It may also be the case that documented workers require a compensating differential for their discriminatory negative attitudes toward unauthorized immigrants or Latinos, in general (see Cowan et al. 1997).

Other studies have found fairly strong negative impact of the arrival of immigrants on wages of previous wages of immigrants (e.g., Ottaviano and Peri 2006 and Lalonde and Topel 1991). Since we cannot identify immigration status among our documented workers, there are surely immigrants among them who are likely to experience a negative consequence of the presence of undocumented workers. The implication is that our results are likely biased toward finding a negative wage impact, which means the very small negative impact estimated is likely even closer to zero than reported here, and the positive impact may be even be a larger positive impact.

V. Conclusions and Implications for Policy

Using individual-level data, linked to employer characteristics, the analysis in this paper finds that a documented worker employed at a firm that hires the average share of undocumented

[15] Iskander (2011) suggests that this productivity gain may derive from the transformation of immigrants' tacit knowledge in a new environment into innovation and improved production processes.

workers (among firms hiring undocumented workers) can expect to earn a 0.15 percent lower wage than if he/she worked at a firm that does not employ undocumented workers. This effect of working in a firm that employs undocumented workers varies across sectors, is even positive in two notable sectors (retail trade and leisure & hospitality), but is small in every sector. The largest negative impact is found among workers in agriculture employed in firms in the 90th percentile of firms' undocumented workforce share (20% or more of the firm's workforce is undocumented). Documented workers in these firms earn roughly three percent less than they would if employed in an agricultural firm that does not hire undocumented workers. This amounts to less than $300 per quarter.

The premium earned by documented workers employed by firms that hire undocumented workers in the retail trade and leisure and hospitality sectors, although, again, small, warrants some additional consideration since this result may have some implications for other results found in the literature. Peri (2009) and Peri and Sparber (2009) suggest that efficiency and productivity can benefit from the task specialization that is likely to result as firms hire low-skill immigrants to perform the tasks previously performed by natives. The natives are re-assigned to relatively higher-skilled tasks that make better use of their comparative advantage, say, communication. Both retail trade and leisure and hospitality, unlike other immigrant-rich sectors like agriculture, construction, and manufacturing, are sectors in which there is ample opportunity for customer contact and specialization in communication-intensive activities.

The bottom line from the analysis in this paper is that whether a worker experiences a premium or a penalty from working at a firm that also employs undocumented workers, that impact is expected to be very small. There is reason to believe that actual experience is even smaller than that estimated here, but one could argue that as the share of undocumented workers

increases in the labor market, these impacts could grow, as well. In addition, the analysis in this paper is a partial equilibrium analysis and does not consider the long-run implications for technology or capital usage by the firm from increasing employment of undocumented workers. Further, the analysis in this paper says nothing about the impact of the presence of undocumented workers on overall employment, prices, or economic growth.

References

Abowd, John M.; Francis Dramarz; and David N. Margolis. "High Wage Workers and High Wage Firms." *Econometrica* 67(2) (March 1999): 251-333.

Altonji, Joseph and David Card. "The Effects of Immigration on the Labor market Outcomes of Less-skilled Natives." In J. Abowd and R. Freeman, eds. *Immigration, Trade, and the Labor Market*, 201-34. Chicago: University of Chicago Press, 1991.

Altonji, Joseph and Robert A. Shakotko. "Do Wages Rise with Job Seniority?" *Review of Economic Studies* 54(3) (July 1987): 437-59.

Aydemir, Abdurrahman and George J. Borjas. "Attenuation Bias in Measuring the Wage Impact of Immigration." *Journal of Labor Economics* 29(1) (January 2011): 69-113

Barro, Robert and Xavier Sala-i-Martin. *Economic Growth*. New York: McGraw Hill, 1995.

Bohon, Stephanie A.; Katherine Stamps; and Jorge H. Atiles. "Transportation and Migrant Adjustments in Georgia." *Population Research Policy Review* 27 (2008): 273-91.

Borjas, George J. "The Analytics of the Wage Effect of Immigration." *Mimeo,* Harvard University (March 2009).

Borjas, George J. "Native Internal Migration and the Labor Market Impact of Immigration." *NBER Working Paper #11610* (September 2005).

Borjas, George J. "The Labor Demand Curve *Is* Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *The Quarterly Journal of Economics* (November 2003): 1335-74.

Borjas, George J.; Jeffrey Grogger; and Gordon H. Hanson. "Immigration and the Economic Status of African-American Men." *Economica* 33(306) (April 2010): 255-82..

Bovbjerg, Barbara D. 2006. *Social Security Numbers: Coordinated Approach to SSN Data Could Help Reduce Unauthorized Work*, Testimony before the Subcommittees on Social Security and on Oversight, Committee on Ways and Means, House of Representatives, GAO-06-458T, http://www.gao.gov/new.items/d06458t.pdf, (accessed 15 December 2008).

Brown, Charles and James L. Medoff. "Firm Age and Wages." *Journal of Labor Economics* 21(3) (July 2003): 677-98.

Brown, J. David; Serife Genc; Julie L. Hotchkiss; and Myriam Quispe-Agnoli. "Undocumented Workers' Employment Across the U.S. Business Cycle." *Mimeo, Federal Reserve Bank of Atlanta* (November 2010).

Brown, J. David; Julie L. Hotchkiss; and Myriam Quispe-Agnoli. "Does Hiring Undocumented Workers Give Firms a Competitive Advantage?" *FRBA Working Paper #2012-02* (February 2012).

Burgess, Simon; Julie Lane; and David Stevens. "Churning Dynamics: An Analysis of Hires and Separations at the Employer Level." *Labour Economics* 8 (2001): 1-14.

Butcher, Kristin F. and David Card. "Immigration and Wages: Evidence from the 1980s." *American Economic Review Papers and Proceedings* 81(2) (May 1991): 292-6.

Campbell, Carl M. III. "Do Firms Pay Efficiency Wages? Evidence with Date at the Firm Level." *Journal of Labor Economics* 11(3) (July 1993): 442-70.

Card, David. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19(1) (January 2001): 22-64.

Card, David. "The Impact of the Mariel Boatlift in the Miami Labor Market." *Industrial and Labor Relations Review* 43(2) (January 1990): 245-57.

Card, David and John DiNardo. "Do Immigrant Inflows Lead to Native Outflows?" *American Economic Review Papers and Proceedings* 90 (2) (May 2000), 360-7.

CBO. *Immigration Policy in the United States*, Pub. No. 2499. Washington, D.C.: Congressional Budget Office, February 2006.

CBO. *Immigration Policy in the United States: An Update*, Pub. No. 4160. Washington, D.C.: Congressional Budget Office, January 2010.

Chiswick, Carmel U.; Barry R. Chiswick; and Georgios Karras. "The Impact of Immigrants on the Macroeconomy." *Carnegie-Rochester Conference Series on Public Policy* 37 (December 1992): 279-316.

Cobb-Clark, Deborah A.; Clinton R. Shiells; and B. Lindsay Lowell. "Immigration Reform: The Effects of Employer Sanctions and Legalization on Wages." *Journal of Labor Economics* 13(3) (1995): 472-98

Committee on Ways and Means, House of Representatives. 2004. *Greenbook,* WMCP 108-6, Section 4, (April).

Cowan, Gloria; Livier Martinez; and Stephanie Mendiola. "Predictors of Attitudes Toward Illegal Latino Immigrants." *Hispanic Journal of Behavioral Sciences* 19(4) (November 1997): 403-15.

Dehejia, Rajeev H. and Sadek Wahba. "Propensity Score-matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84(1) (February 2002): 151-61.

Denton, Frank T. and Byron G. Spencer. "Population, Labour Force and Long-term Economic Growth." McMaster University, Quantitative Studies in Economics and Population Research Reports (1997).

FAIR. *U.S. Immigration History*. <http://www.fairus.org/site/PageServer?pagename=research_research9c29>. Access 26 October 2007.

Fix, Michael and Jeffrey S. Passel. *Immigration and Immigrants: Setting the Record Straight*. Washington, D.C.: The Urban Institute, May 1994.

Fortuny, Karina; Randy Capps; and Jeffrey S. Passel. "The Characteristics of Unauthorized Immigrants in California, Los Angeles County, and the United States." *Mimeo*. Washington, D.C.: The Urban Institute (March 2007).

Frey, William H. "Immigration, Domestic Migration, and Demographic Balkanization in America: New Evidence for the 1990s." *Population and Development Review* 22(4) (December 1996): 741-63.

Friedberg, Rachel M. and Jennifer Hunt. "The Impact of Immigrants on Host Country Wages, Employment and Growth." *Journal of Economic Perspectives* 9 (Spring 1995): 23-44.

GAO. "Illegal Immigration: Border-Crossing Deaths Have Doubled Since 1995; Border Patrol's Efforts to Prevent Deaths Have not Been Fully Evaluates," GAO-06-770. Report to the Honorable Bill Frist, Majority Leader, U.S. Senate (August 2006).

Hamada, Koichi and Hiromi Kato, eds. *Ageing and the Labor Market in Japan: Problems and Policies*. ESRI Studies Series on Ageing. Cheltenham, U.K. and Northampton, Mass: Elgar, 2007.

Hanson, Gordon H. "Illegal Immigration from Mexico to the United States." *Journal of Economic Literature* 44 (December 2006): 869-924.

Hanson, Gordon H. "Logic of Illegal Immigration." *Council on Foreign Relations, CSR No. 26* (April 2007).

Hanson, Gordon H. and Antonio Spilimbergo. "Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border." *The American Economic Review* 89 (Dec 1999): 1337-57.

Heyman, Fredrik; Fredrik Sjoholm; Patrik Gustavsson Tingvall. "Is There Really a Foreign Ownership Wage Premium? Evidence from Matched Employer-Employee Data." *Journal of International Economics* 73(2) (November 2007): 355-76.

Hotchkiss, Julie L. "Employment Growth and Labor Force Participation: How Many Jobs are Enough?" *Federal Reserve Bank of Atlanta Economic Review* 90 (1) (Quarter 1, 2005): 1-13.

Hotchkiss, Julie L. and Myriam Quispe-Agnoli. "Employer Monopsony Power in the Labor Market for Undocumented Workers." *FRBA Working Paper 2009-14c* (December 2009).

Huse, James G. Jr. "Status of the Social Security Administration's Earnings Suspense File" *Congressional Response Report #A-03-03-23038* (November 2002).

Hotchkiss, Julie L. and Myriam Quispe-Agnoli. "Employer Monopsony Power in the Labor Market for Undocumented Workers." *FRBA Working Paper 2009-14c* (December 2009).

Imbens, Guido W. 2004. "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review." *Review of Economics and Statistics* 86(1) (February): 4-29.

Iskander, Natasha and Nichola Lowe. "The Transformers: Immigration and Tacit Knowledge Development." *NYU Wagner Research Paper No. 2011-01* (January 2011).

Jasso, Guillermina. "Migration and Stratification." *Social Science Research* 40 (2011): 1292–1336.

Jasso, Guillermina; Douglas S. Massey; Mark R. Rosenzweig; and James P. Smith, J.P. (2000). "The New Immigrant Survey Pilot (NIS:P): Overview and New Findings about U.S. Legal Immigrants at Admission." *Demography* 37 (2000): 127–138.

Jasso, Guillermina; Douglas S. Massey; Mark R. Rosenzweig; and James P. Smith. "From Illegal to Legal: Estimating Previous Illegal Experience Among New Legal Immigrants to the United States." *IZA Discussion Paper #3441* (April 2008).

Johnson, Mary. "The Growing Cost of Illegal Immigration to Social Security: Unprecedented Growth in Social Security's 'Earnings Suspense file'." *Mimeo*, The Senior Citizen's League (http://www.tscl.org/NewContent/102880.asp) (March 2007, 2/10/10).

Jordan, Miriam. "Crackdown on Illegal Workers Grows." *WSJ.com* (20 January 2011) <http://online.wsj.com/article/SB10001424052748703951704576092381196958362.html >, accessed 1 April 2011.

Kossoudiji, Sherrie A. and Deborah a. Cobb-Clark. "IRCA's Impact on the Occupational Concentration and Mobility of Newly-legalized Mexican Men." *Journal of Population Economics* (13) (2000): 81-98.

Krusell, Per; Ohanian, Lee E.; Jose-Victor Rios-Rull; and Giovanni L. Violante. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," Econometrica, Econometric Society, vol. 68(5), pages 1029-1054, (September 2000).

Lalonde, Robert and Robert Topel. "Labor Market Adjustments to Increased Immigration." In J. Abowd and R. Freeman, eds. *Immigration, Trade, and the Labor Market*, 167-200. Chicago: University of Chicago Press, 1991.

Malchow-Møller, Nikolaj; Jakob Roland Munch; and Jan Rose Skaksen. "Do Immigrants Affect Firm-specific Wages?" *IZA Discussion Paper No. 3264* (December 2007).

Maloney, Thomas N. and Thomas Kontuly. "Legal Status and Economic Mobility Among Immigrants in the Early Twenty-first Century: Evidence from the 'New Gateway' of Utah." In Maloney and Korinek, eds., *Migration in the 21st Century: Rights, Outcomes, and Policy.* New York: Routledge (2010).

Morales, Rebecca. "Transitional Labor: Undocumented Workers in the Los Angeles Automobile Industry," *International Migration Review*, 17(4), (1983): 570-596.

Neal, Larry and Paul Uselding. "Immigration, a Neglected Source of American Economic Growth: 1790 to 1912." *Oxford Economic Papers,* New Series 24 (Mary 1972): 68-88.

Nuñez-Neto, Blas and Stephen R. Viña. "Border Security: Fences along the U.S. International Border." *CRS Report for Congress* (order Code RS22026) (11 January 2006).

Oi, Walter Y. and Todd L. Idson. "Firm Size and Wages." In Orley Ashenfelter and David Card, eds. *Handbook of Labor Economics*, pp. 2165-214. Amsterdam: Elsevier, 1999.

Orrenius, Pia M. and Madeline Zavodny. "Does Immigration Affect Wages? A Look at Occupation-level Evidence." *Labour Economics* 14 (2007): 757-73.

Orrenius, Pia M. and Madeline Zavodny. "Why Stop There? Mexican Migration to the U.S. Border Region." In Marie Mora and Alberto Davila, eds. *Labor Market Issues along the U.S.-Mexican Border: Economic and Demographic Analyses*. Forthcoming.

Ottaviano, Gianmarco I.P. and Giovanni Peri. "Rethinking the Effects of Immigration on Wages." *NBER Working Paper No. 12497* (August 2006).

Passel, Jeffrey S. and D'Vera Cohn. "A Portrait of Unauthorized Immigrants in the United States." *Pew Hispanic Center Report* (14 April 2009).

Passel, Jeffrey S. and D'Vera Cohn. "Unauthorized Immigrant Population: National and State Trends, 2010." *Pew Hispanic Center Report* (1 February 2011).

Pena, Anita Alves. "Legalization and Immigrants in U.S. Agriculture." *The B.E. Journal of Economic Analysis & Policy* 10(1) (2010): article 7.

Peri, Giovanni. "The Effect of Immigration on Productivity: Evidence from U.S. States." *NBER Working Paper No. 15507* (November 2009).

Peri, Giovanni and Chad Sparber. "Task Specialization, Immigration, and Wages." *American Economic Journal: Applied Economics* 1(3) (2009): 135-69.

Toussaint-Comeau, Maude. "The Impact of mexican Immigrants on U.S. Wage Structure." *Federal Reserve Bank of Chicago Working Paper* (2007).

Wen, Ming and Thomas N. Maloney. "Latino Residential Isolation and the Risk of Obesity in Utah: The Role of Neighborhood Socioeconomic, Built-Environmental, and Subcultural Contect." *Journal of Immigrant and Minority Health* 13(6) (December 2011): 1134-41.

Wright, Richard A.; Mark Ellis; and Michael Reibel. "The Linkage between Immigration and Internal Migration in Large Metropolitan Areas in the United States." *Economic Geography* 73(2) (April 1997): 234-53.

Figure 1.

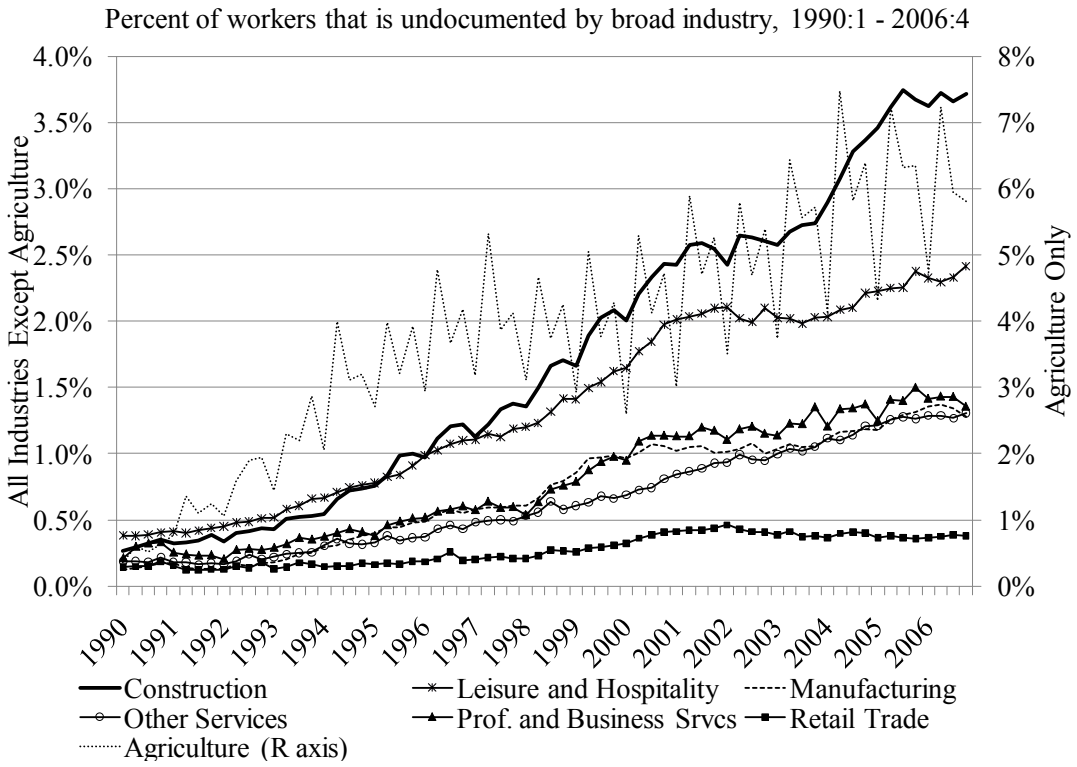
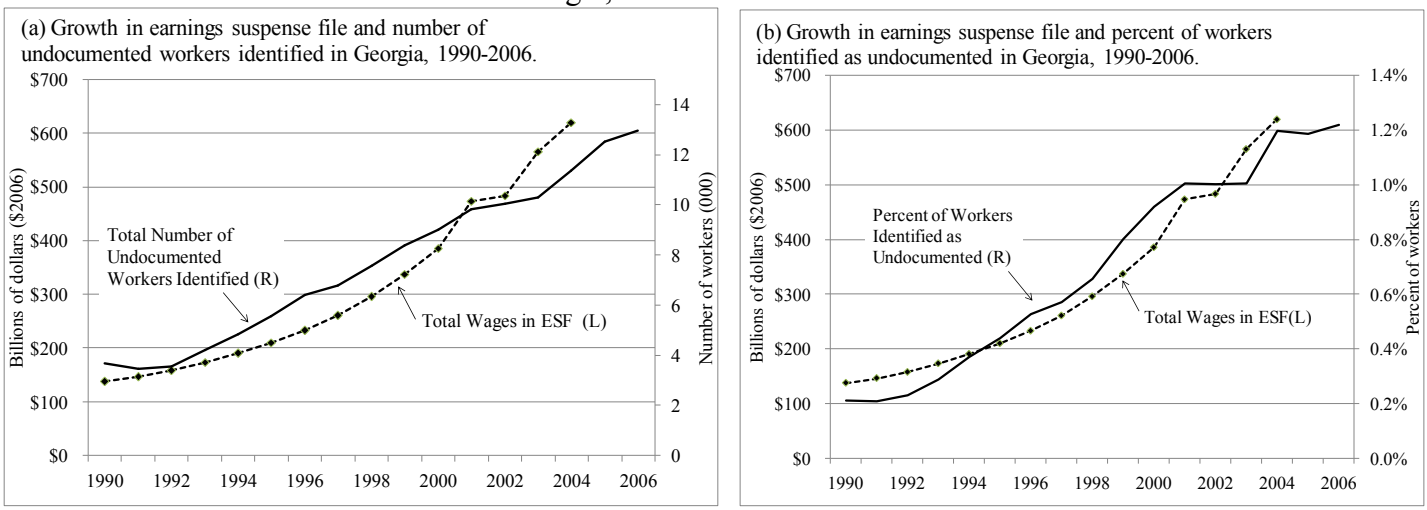
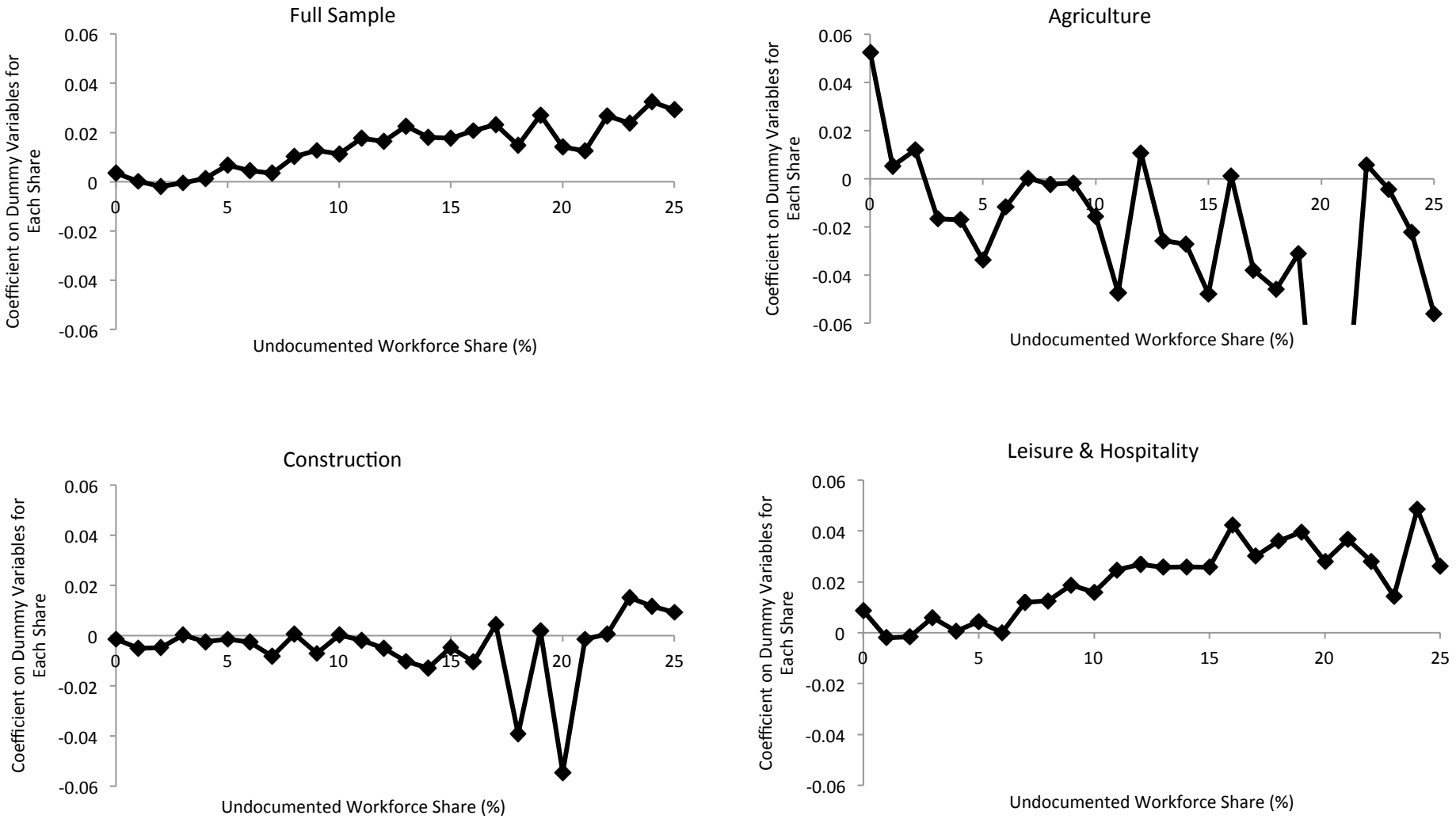


Figure 2. Growth in the earnings suspense file and the total number and percent of workers identified as undocumented in Georgia, 1990-2006.



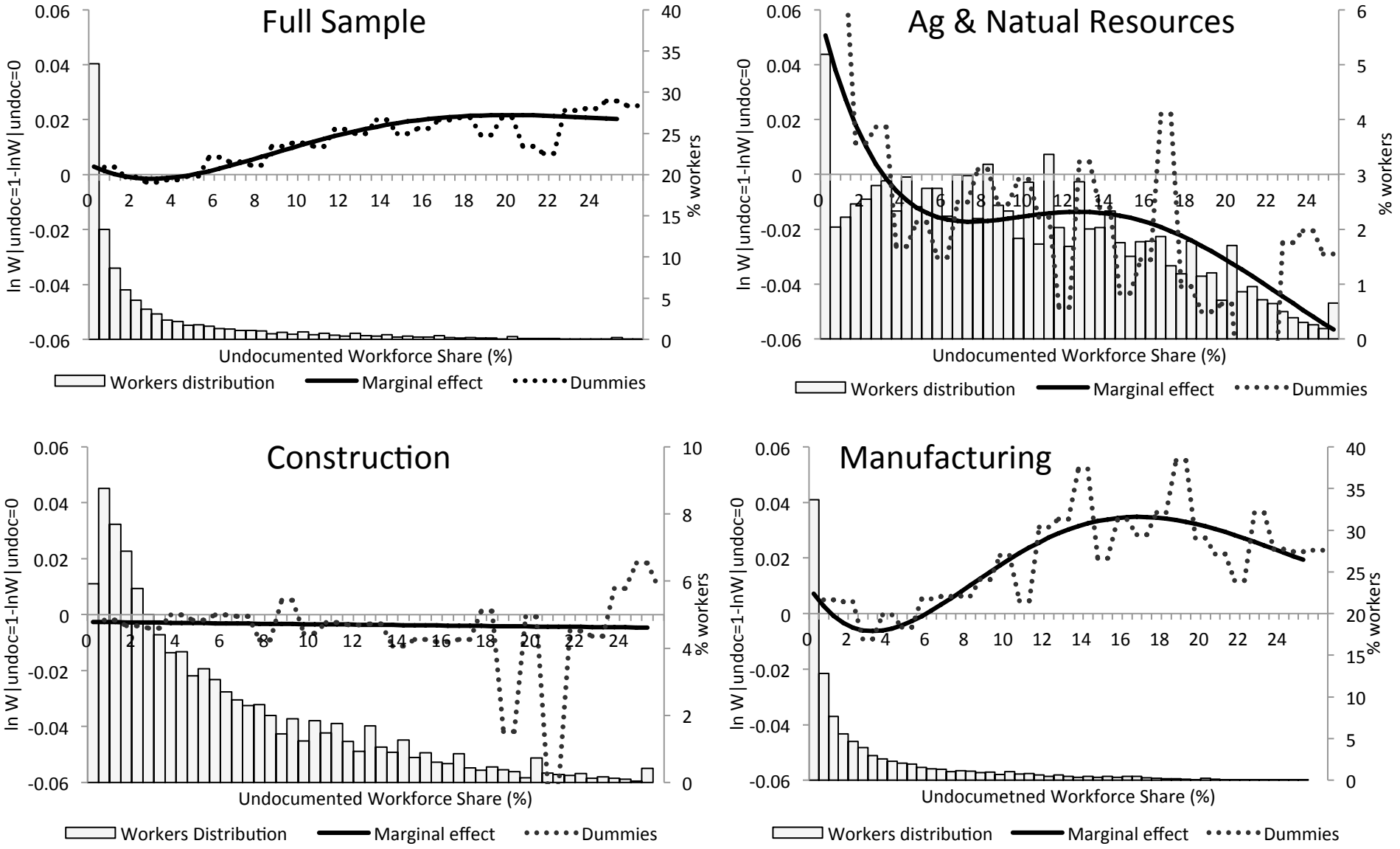
Source: Huse (2002) for estimates 1990-2000, Johnson (2007) for estimates 2001-2004, and authors' calculations. Dollar estimates reflect 2006 values, using the PCE chain-weighted deflator.

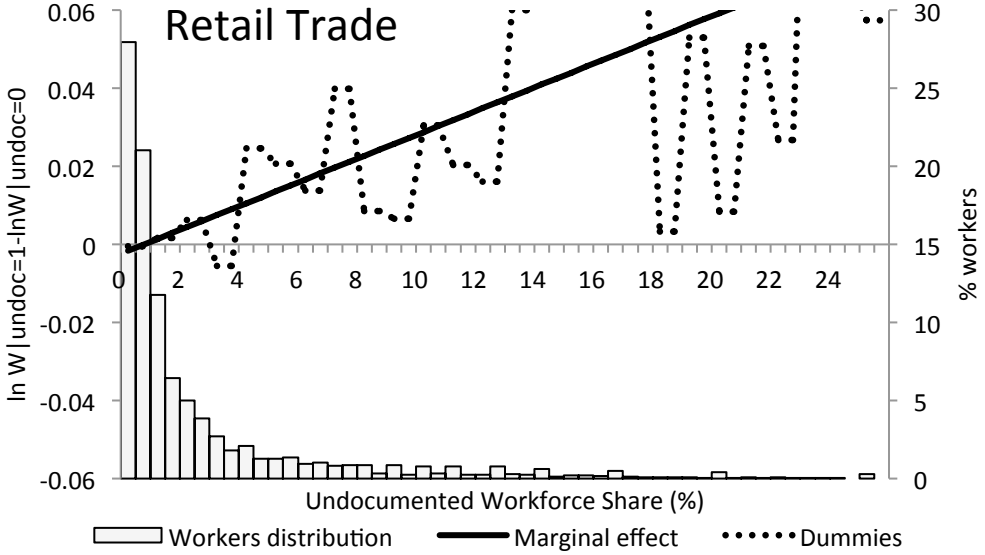
Figure 3. Linear probability model estimated with dummy variables for each share of undocumented workers on firms' payrolls.

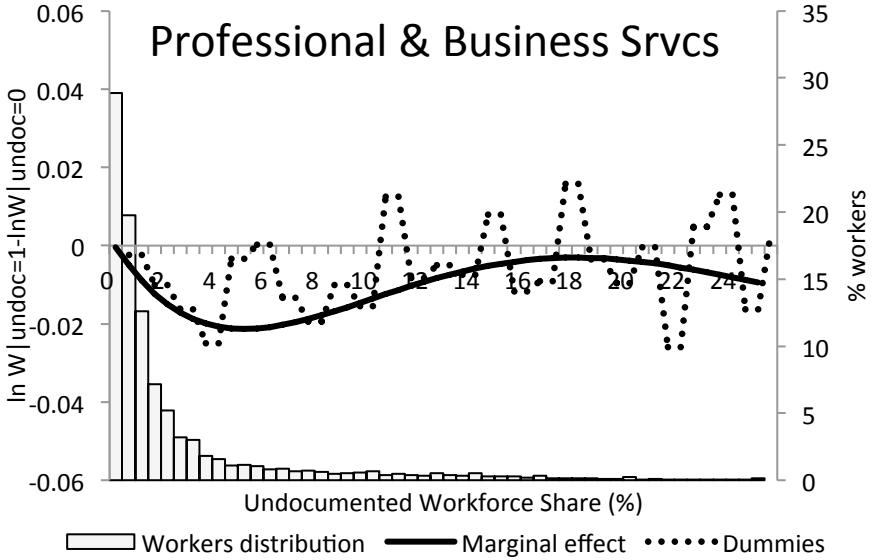


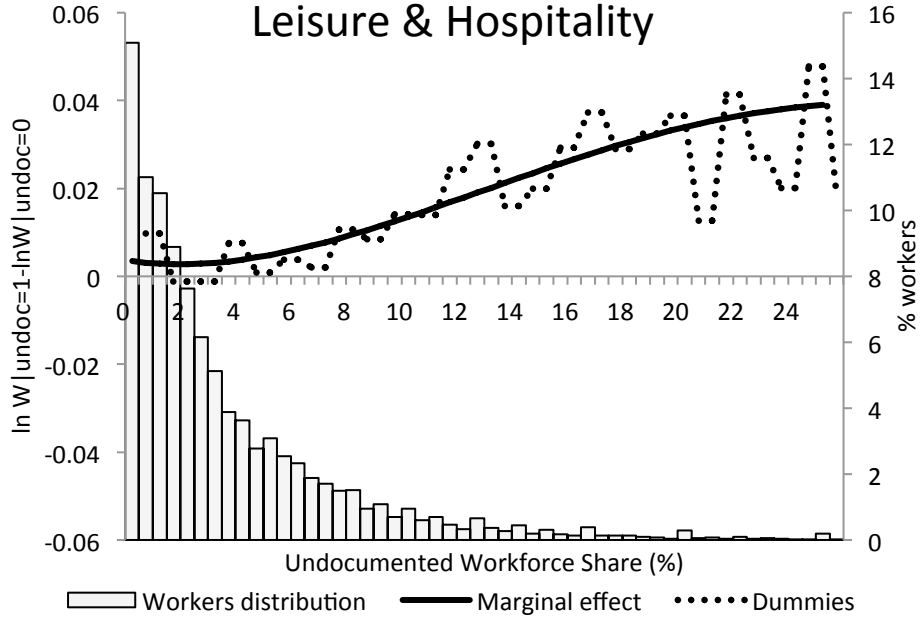
Notes: Estimation using final matched, trimmed samples.

Figure 4. Estimated marginal effects of being employed by a firm that hires undocumented workers, with the distribution of workers across undocumented workforce share.




Retail Trade
ln W|undoc=1-lnW|undoc=0
% workers
Undocumented Workforce Share (%)
Workers distribution
Marginal effect
Dummies


Professional & Business Srvcs
ln W|undoc=1-lnW|undoc=0
% workers
Undocumented Workforce Share (%)
Workers distribution
Marginal effect
Dummies


Leisure & Hospitality
ln W|undoc=1-lnW|undoc=0
% workers
Undocumented Workforce Share (%)
Workers distribution
Marginal effect
Dummies

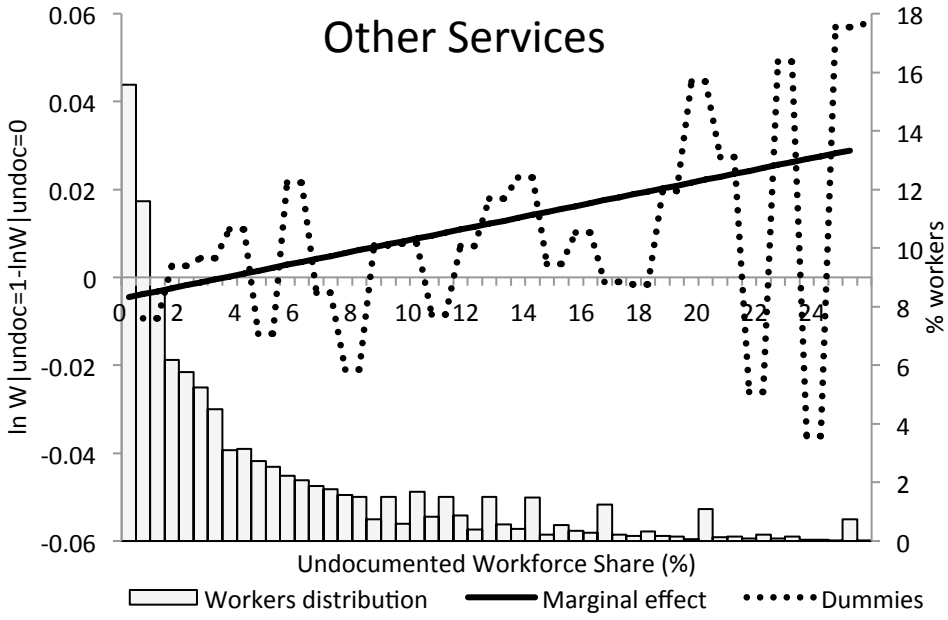

Other Services
ln W|undoc=1-lnW|undoc=0
% workers
Undocumented Workforce Share (%)
Workers distribution
Marginal effect
Dummies

Table 1. Average annual growth, 1994-2006, in US and GA employment, Hispanic workers, and workers identified as undocumented.

Average Annual Growth Rate of:	
Total number of workers in the U.S.	1.48%
Total number of foreign born, Hispanic workers in the U.S.	8.03%
Total number of foreign born, Hispanic workers with less than a high school degree in Georgia	7.28%
Total number of workers in Georgia	2.92%
Total number of foreign born, Hispanic workers in Georgia	26.82%
Total number of foreign born, Hispanic workers with less than a high school degree in Georgia	21.48%
Total number of workers in GA identified as undocumented	*25.29%*

Source: Current Population Survey, Basic Survey (March), 1994-2006; and authors' calculations.
Note: 1994 is used as starting year since is the first year the Current Population Survey has a reliable indicator of Hispanic ethnicity.

Table 2. Sample means, documented workers, 1995-2005.

			Single Estab. Firms	
	Full Sample	Single Establish-ment Firms Only	Workers employed at firms with at least one undocumented worker ($U_{jt} = 1$)	Workers employed at firms with no undocumented workers ($U_{jt} = 0$)
Worker Characteristics				
Quarterly wages	$8,277 (11,522)	$8,547 (12,741)	$6,520 (9,711)	$9,272 (13,589)
Worker tenure (quarters)	13 (14)	11 (13)	9 (12)	12 (13)
Worker experience (quarters)	26 (16)	25 (16)	22 (16)	26 (16)
Worker is a new hire = 1	0.17	0.18	0.26	0.16
Worker is separating = 1	0.17	0.18	0.25	0.16
Firm Characteristics (firm-level means)				
Percent of firms hiring undoc (U_{jt})		5.0%	--	--
Percent of workers at firm that is undoc. (P_{jt})	0.56% (2.79)	0.54% (3.4)	11% (10.9)	--
Firm size (number of workers)	21 (207)	15 (87)	83 (287)	12 (58)
Age of firm (quarters)	16 (14)	25 (17)	28 (18)	25 (17)
Churning at worker's firm (doc. workers only)	0.24 (0.38)	0.14 (0.29)	0.34 (0.33)	0.13 (0.29)
Share of workers with college degree or higher (determined at 3-digit NAIC level)[a]		24% (0.16)	15% (0.12)	25% (0.16)
Labor share in production process (determined at 3- or 4-digit NAIC level)[a]		40% (0.15)	38% (0.12)	40% (0.15)
Geographic Controls				
Population density (population/sq mi)	1,032 (770)	981 (770)	1,082 (757)	945 (772)
County unemployment rate	4.49 (1.30)	4.49 (1.35)	4.42 (1.26)	4.51 (1.39)
Percent of county school enrolment that is Hispanic	5% (5)	5% (5)	5% (5)	5% (5)
NAICS Sector Employment Shares				
Natural Resources and Agriculture	0.9%	1.4%	2.4%	1.0%
Construction	5.1%	8.9%	10.3%	8.4%

Manufacturing	13.5%	13.5%	16.2%	12.5%
Transportation and Utilities	4.4%	3.8%	2.5%	4.3%
Wholesale Trade	4.8%	7.0%	3.4%	8.3%
Retail Trade	13.4%	9.1%	5.1%	10.5%
Financial Activities	5.3%	5.8%	2.1%	7.1%
Information	3.4%	3.0%	2.7%	3.1%
Professional and Business Services (includes temporary services)	14.8%	17.8%	22.9%	16.0%
Education and Health Services	17.2%	15.7%	14.4%	16.2%
Leisure and Hospitality	10.4%	10.0%	16.0%	7.8%
Other Services (includes private household, laundry, and repair and maintenance services)	2.6%	4.0%	1.9%	4.7%
No. of observations	178m	99,868,841	26,313,397	73,555,444

Notes: Wages are real quarterly earnings, deflated by the chained price index for personal consumption expenditure $2006Q4. Individual sample means are across workers. All averages reflect four quarters of data for each year. Standard errors are in parentheses. Note in comparing values across years that 1990 is the first year in which any firm or worker is observed in the data. See Appendix B for details related to the construction of high education, labor intensity, and broad sector classifications.

[a] In order to conserve space, the variables are not merged into the full data set.

Table 3. Linear, fixed effects estimation of log wages.

Dependent variable = Log(wage)	Includes individual and firm level F.E. (1)	OLS estimation (no F.E.) (2)
Firm hires undocumented workers (U_{jt})	0.0012***	-0.0489***
	(0.0004)	(0.0006)
Coefficients for Percent undocumented interaction are available upon request.		
Worker Characteristics		
Worker tenure (in quarters)	0.0078***	0.035***
	(0.0001)	(0.0001)
Worker tenure squared / 1000	-0.1079***	-0.4396***
	(0.0011)	(0.0012)
Worker experience (in quarters)	0.0096***	0.0089***
	(0.0002)	(0.0000)
Worker experience squared / 1000	-0.0731***	-0.0515***
	(0.0011)	(0.0008)
Worker is new hire = 1	-0.7021***	-0.7542***
	(0.0005)	(0.0006)
Worker is separating = 1	-0.8931***	-1.1024***
	(0.0005)	(0.0005)
Firm Characteristics		
Age of firm (in quarters)	-0.0132***	-0.0100***
	(0.0005)	(0.0001)
Age of firm squared / 100	0.0001***	0.0001***
	(0.0000)	(0.0000)
Log total employment (firm size)	0.1539***	0.0351***
	(0.0007)	(0.0002)
Churning at worker's firm (documented wrkrs only)	0.1461***	-0.6065***
	(0.0012)	(0.0009)
Share of workers with higher education	-0.01690	0.6171***
	(0.0110)	(0.0019)
Labor intensity	-0.0564***	-0.3263***
	(0.0148)	(0.0022)
Geographic Controls		
County unemployment rate	-0.0057***	-0.0279***
	0.00023	0.00017
Population density /1000	0.00001***	0.00014***
	0.00000	0.00000
Percent of students that is Hispanic	-0.0828***	0.2664***
	0.01108	0.00409
Intercept		8.1743***
		(0.0085)
R-squared	0.8657	0.4420

No. of Observations	32,085,002	32,085,002

Note: Both models presented here reflect the preferred specification based on criteria described in text (see equation 1; a 6th order polynomial is the preferred specification for the full sample). Standard errors in parentheses. Analysis includes documented workers employed in Georgia 1995-2005 inclusive. Also included as regressors are broad sector industry fixed effects (see Appendix B) and year and quarter fixed effects, as well as sector/year interactions. P_{jt} is measured from 0 to 100. *** statistically significantly different from zero at the 99% confidence level, ** statistically significantly different from zero at the 95% confidence level, and * statistically significantly different from zero at the 90% confidence level.

Table 4. Marginal effects on documented worker wages, by industry, across the distribution of undocumented workforce shares; undocumented workforce shares (%) across the distribution are in brackets.

Industry	Average expected wage difference $\frac{1}{N}\sum_{i}^{N}\left\{\widehat{\ln w_{iJ}}\big\|_{U_j=1} - \widehat{\ln w_{iJ}}\big\|_{U_j=0}\right\}$						Average Quarterly Earnings Among Documented Workers at Firms (no Undoc Wrkrs)
	at Mean	P10	P25	P50	P75	P90	
Full Sample	**-0.0015** [3.24]**	**0.0035*** [0.08]**	**0.0027*** [0.29]**	**0.0003 [1.16]**	**-0.0012* [3.92]**	**0.0095*** [9.59]**	$9,272
Ag and Natural Resources	-0.0152** [10.12]	0.0204*** [1.61]	-0.0103** [4.39]	-0.0163*** [9.09]	-0.0151** [14.75]	-0.0298*** [19.66]	$7,121
Construction	-0.0031** [6.21]	-0.0027* [0.72]	-0.0027** [1.69]	-0.0029** [4.17]	-0.0033* [8.97]	-0.0038 [14.29]	$8209
Manufacturing	-0.0063*** [3.24]	0.0088*** [0.10]	0.0068*** [0.29]	-0.0005 [1.20]	-0.0054*** [4.24]	0.0174*** [9.76]	$9,766
Transportation and Utilities	-0.0063*** [1.89]	0.0088*** [0.07]	0.0068*** [0.22]	-0.0005 [0.70]	-0.0054*** [1.90]	0.0174*** [4.88]	$9,349
Wholesale Trade	-0.00002 [4.57]	0.0138*** [0.27]	0.0103*** [0.70]	0.0022 [2.34]	0.0018 [6.44]	0.007 [12.34]	$13,307
Retail Trade	0.0062*** [2.82]	-0.0017 [0.22]	-0.001 [0.45]	0.0007 [1.03]	0.0063*** [2.86]	0.021*** [7.69]	$6,552
Financial Activities	-0.0024 [1.94]	0.0022 [0.09]	0.0019 [0.22]	0.0009 [0.58]	-0.002 [1.75]	-0.0067 [5.00]	$12,258

Information	0.0086*** [0.59]	0.0241*** [0.04]	0.0234*** [0.06]	0.0207*** [0.14]	0.0136*** [0.38]	-0.0033 [1.25]	$14,267
Professional and Business Srvcs	-0.0173*** [2.81]	0.0004 [0.01]	-0.0021 [0.42]	-0.0074*** [1.03]	-0.0169*** [2.70]	-0.0179*** [8.21]	$11,303
Education and Health Services	-0.0021*** [0.50]	-0.003*** [0.01]	-0.0029*** [0.03]	-0.0028*** [0.12]	-0.0024*** [0.33]	-0.0012 [0.93]	$8,855
Leisure and Hospitality	0.0034** [3.79]	0.0034* [0.33]	0.003* [0.95]	0.0028* [2.27]	0.0046** [5.00]	0.0108*** [8.89]	$3,056
Other Services	0.0021 [5.18]	-0.0043 [0.39]	-0.0036 [0.88]	-0.0012 [2.70]	0.0045 [6.98]	0.013* [13.33]	$6,162

Note: *** statistically significantly different from zero at the 99% confidence level, ** statistically significantly different from zero at the 95% confidence level, and * statistically significantly different from zero at the 90% confidence level. Earnings are real quarterly earnings, deflated by the chained price index for personal consumption expenditure $2006Q4.

Appendix A: Using SSNs to Identify Undocumented Workers

A.1. Identifying Invalid Social Security Numbers

Every quarter employers must file a report with their state's Department of Labor detailing all wages paid to workers who are covered under the Social Security Act of 1935. Each worker on this report is identified by his/her social security number (SSN). There are a number of ways in which one can establish that a reported social security number is invalid. The Social Security Administration provides a service by which an employer can upload a file of SSNs for checking, but one must register as an employer to obtain this service.[i] In addition, there are several known limitations on what can be considered a valid social security number, so a simple algorithm is used to check whether each number conforms to the valid parameters.

There are three pieces to a SSN.[ii] The first three numbers are referred to as the Area Number. This number is assigned based on the state in which the application for a SSN was made; it does not necessarily reflect the state of residence. The lowest Area Number possible is 001 and the highest Area Number ever issued, as of December 2006, is 772. Using information provided by the SSA, the dates at which area numbers between 691 and 772 are first assigned can be determined. Any SSN with an Area Number equal to 000, greater than 772, or which shows up before the officially assigned date, will be considered invalid.

The second piece of a SSN consists of the two-digit Group Number. The lowest group number is 01, and they are assigned in non-consecutive order. Any SSN with a Group Number equal to 00 or with a Group Number that appears in the data out of sequence with the Area Number will be considered invalid.

---

[i] See Social Security Number Verification Service <http://www.ssa.gov/employer/ssnv.htm>.

[ii] Historical information and information about valid SSNs can be found at the Social Security Administration's web sites: <http://www.ssa.gov/history/ssn/geocard.html>, <http://www.xocialsecurity.gov/employer/stateweb.htm>, and <http://www.socialsecurity.gov/employer/ssnvhighgroup.htm >.

The last four digits of a SSN are referred to as the Serial Number. These are assigned consecutively from 0001 to 9999. Any SSN with a Serial Number equal to 0000 is invalid.

In 1996 the Internal Revenue Service (IRS) introduced the Individual Tax Identification Number (ITIN) to allow individuals who had income from the U.S. to file a tax return (the first ITIN was issued in 1997). It is simply a "tax processing number," and does not authorize an individual to work in the U.S. Employers are instructed by the IRS to "not accept an ITIN in place of a SSN for employee identification for work. An ITIN is only available to resident and nonresident aliens who are not eligible for U.S. employment and need identification for other tax purposes."[iii] ITIN numbers have a "9" in the first digit of the Area Number and a "7" or "8" in the first digit of the Group Number. Anyone with this numbering scheme will be identified as having an invalid Area Number; the percent of SSNs with high area numbers that also match the ITIN numbering scheme has risen from about one percent in 1997 to over 60 percent by the end of 2006.

A series of SSNs were de-commissioned by the Social Security Administration because they had been put on fake Social Security Cards used as props to sell wallets.[iv] Apparently, some people who purchased the wallets thought the fake Social Security Cards were real and started using them as their own. If any of these 21 "pocketbook" SSNs appear in the data, they are considered invalid, although their frequency is so low as to be inconsequential. In addition, a number of SSNs are exactly equal to the employer identification number. These are invalid, primarily because they have too few digits. In any instance where a SSN is used for more than one person on a firm's UI wage report or does not have the required number of digits (including

[iii] "Hiring Employees," <http://www.irs.gov/businesses/small/article/0,,id=98164,00.html>. Also see, "Individual Taxpayer Identification Number (ITIN)," <http://www.irs.gov/individuals/article/0,,id=96287,00.html>.
[iv] See "Disclosure and Verification of Social Security Numbers (SSNs) for the Section 235 Program" (9 November 1990), <http://www.hud.gov/offices/adm/hudclips/letters/mortgagee/files/90-39ml.txt> (accessed 8 February 2011).

zeros), the SSN is considered invalid.

The possibility that someone fraudulently uses a valid SSN assigned to someone else poses a special problem. First of all, the SSN will show up multiple times across firms in one quarter for workers with different surnames (the wage report includes the first three characters of the workers' surnames). With this information alone, it is not possible to know which worker is using the SSN fraudulently and who the valid owner of the number is. If one of the SSN/surname pairs shows up in the data initially in a quarter by itself, this is the pair that is considered valid and all other duplicates (with different surnames) are considered invalid.

A.2. Does "Invalid" mean "Undocumented?"

Not all invalid SSN are classified as undocumented workers; examining the patterns of incidence of different types of invalid SSNs suggests that some types are firm generated rather than worker generated. Figure 1 illustrates the incidence patterns across types of invalid SSNs in construction. The percent of workers with SSNs having a high area number or out-of-sequence group number displays the expected growth in undocumented workers, whereas the incidence of SSNs for other reasons exhibits a flat to declining, highly seasonal pattern (this seasonality appears in all other sectors, as well).[v] The strong seasonal nature of the other invalid reasons suggests that firms are temporarily assigning invalid SSN numbers to workers before having time to gather the information for the purpose of record keeping/reporting. Or, firms may decide to not bother obtaining a SSN for workers who will only be employed a very short time.[vi] The high degree of churning observed among workers with invalid SSNs for these other reasons is

[v] Documentation of growth in undocumented workers can be found in Michael Hoefer, Nancy Rytina, and Christopher Campbell, "Estimates of the Unauthorized Population Residing in the United States: January 2006," *Population Estimates* (Washington D.C.: US Department of Homeland Security, Office of Immigration Statistics, February 2009).

[vi] Indeed, a worker has 90 days to resolve a discrepancy that results in the receipt of a "no-match" letter from the Social Security Administration. The employee may be long gone before such a letter is even received.

consistent with either of these practices.

[Figure A1 here]

Since there is no way to know whether a temporary assignment by the firm of an invalid SSN is to merely cover for temporary employment of an undocumented worker or to allow the firm to file its wage report before having had a chance to record the worker's valid SSN, the analysis below takes the conservative tack by considering as undocumented only those workers whose SSNs are classified as invalid because the area number is too high or the group number is assigned out of sequence; workers with invalid SSNs for any other reason are considered neither undocumented nor documented and, thus, are excluded from the analysis. This will clearly undercount the actual number of undocumented workers. However, all workers, regardless of SSN classification, are included in counts of aggregate firm employment.

Figure A1.

Appendix B: Definitions of broad sectors, industry skill, and industry labor intensity.

B.1. Definitions of Broad Sectors

Throughout this paper, regressors are measured at different levels of industry aggregation. The process of matching is performed at the broad sector level, which are defined based on two-digit NAICS classifications. These classifications are designed to match as closely as possible the former SIC classifications and are reported in Table A1.

Table B1: Definitions of broad sectors based on 2-digit NAICS classifications.

Sector	Included 2-digit NAICS
Agriculture and Natural Resources	11, 21
Construction	23
Manufacturing	31-33
Transportation and Utilities	22, 48-49
Wholesale Trade	42
Retail Trade	44-45
Financial Activities	52-53
Information	51
Professional and Business Services (includes temporary services)	54-56
Education and Health Services	61-62
Leisure and Hospitality	71-72
Other Services (includes private household, laundry, and repair and maintenance services)	81

B.2. Construction of the Measure of Industry Labor Intensity

Labor intensity for each industry is based on coefficients from the U.S. Input-Output (I-O) Benchmark Tables 2002 (http://www.bea.gov/industry/index.htm#benchmark_io). The labor intensity coefficient is defined as the share of compensation of employees (wage bill) in total industry output. Compensation of employees includes wages and salaries and their supplements. Total industry output is the sum of the products consumed by the industry, compensation of employees, taxes on production and imports less subsidies, and gross operating surplus.

B.3. Construction of Industry Skill

Each industry is assigned a skill intensity based on the weighted average of educational attainment of workers in that industry, using the Current Population Survey for 1994. This year was chosen since this is the first year in which the nativity (place of birth) of respondents is reported. For each industry, the percent of workers with less than a high school education (LTHS), a high school education (HS), some college (SCOLL), college degree (COLL), and graduate education (GRAD) is calculated. The regressor *HigherEducation* is the share of workers in the firm's industry with a college or graduate education.

Appendix C: Matching strategy and first stage matching results.

A combination of exact- and propensity score one-to-one matching without replacement is used. For each worker employed in a firm that hires undocumented workers, a set of potential controls is constructed from workers in the same year, quarter, and sector who are employed by firms not hiring undocumented worker. Among these candidate controls, the treated worker's match is chosen as that control with the closest propensity score. The propensity score for all workers is constructed by estimating a logit model with the dependent variable equal to one for workers employed by firms who hire undocumented workers. Determinants for this outcome include both worker and firm characteristics, most of which are also included in the second stage wage equation. Dehejia and Wahba (2002) show that without-replacement matching can reduce the quality of the match when few control candidates are available. However, when there are a large number of control candidates, as is the case here, one-to-one matching improves the precision of the estimates.

Preserving the longitudinal nature of the data makes the matching process more complicated. It is conceivable that a match is found for a treated worker in only one quarter of their observations; this worker (along with its control) will be rejected because of a lack of multiple observations with which to control for fixed effects. An iterative process takes place after the matching exercise to ensure that the number of remaining single observations of workers and firms is minimized for the final wage estimation.[22]

Table C1 presents evidence of the quality of the matching through a comparison of the bias that exists between the controls (workers employed at firms that do not hire undocumented workers) and the treated (workers employed at firms that do hire undocumented). For this effect

[22] The final sample will have only a very small number of workers with only one observation. The goal was to minimize the occurrence of singles in order to have confidence in the standard errors of the estimates; too many singles will produce inappropriately small standard errors.

two statistics are provided: the standardized difference test (%bias) and a test of equality of means (t-test).

The t-tests reject equality of the variable means, for both the unmatched and matched sample. This result is not unexpected considering that we have over 16 million observations, and that the standard deviation of the mean tends to zero with such sample size. Still, there is a considerable reduction on the t-statistics for the matched sample.

Considering the sample size, the standardized bias statistic is perhaps more suitable to test the quality of the matching, since it does not depend on the number of observations in the sample. Before matching it is evident that there are considerable differences in the controls between treatment and comparison group. With a bias of ranging from 2 to 108 percent, only 4 of the covariates showed a difference that could be considered "small" compared to the threshold suggested by Rosenbaum and Rubin (1985) of 20 percent. After matching, all the differences are considerably reduced with a median bias reduction of 91 percent (ranging in absolute value from a 54 percent reduction to a 99.5 percent reduction). Further, all covariates present less than eight percent bias (ranging from 0.4 to 8 percent), supporting the quality of the matching.[23]

[23] The full set of bias statistics is available upon request (for the full sample and for each sector separately).

Table C1. T-Test and Bias statistic before and after Matching/trimming

Full Sample	Unmatched				Matched			
	Pd=0	Pd=1	%bias	t-test	Pd=0	Pd=1	%bias	t-test
Share of Wkrs with higher education	0.25 (0.16)	0.20 (0.15)	**29.74**	**1328.29**	0.20 (0.15)	0.21 (0.16)	**2.81**	**17.41**
Labor intensity	0.39 (0.13)	0.43 (0.18)	**28.19**	**1158.23**	0.39 (0.14)	0.39 (0.14)	**1.33**	**37.67**
Worker tenure	12.08 (13.26)	9.20 (11.76)	**22.97**	**1041.09**	9.78 (11.38)	10.66 (12.39)	**7.43**	**210.37**
Worker tenure^2	0.32 (0.63)	0.22 (0.52)	**17.13**	**791.62**	0.23 (0.49)	0.27 (0.56)	**7.96**	**225.40**
Worker experience	26.17 (15.92)	21.70 (15.74)	**28.23**	**1246.35**	23.82 (15.52)	24.21 (15.76)	**2.50**	**70.92**
Worker experience^2	0.94 (0.95)	0.72 (0.86)	**24.26**	**1094.45**	0.81 (0.89)	0.83 (0.91)	**2.93**	**83.04**
New Hire	0.16 (0.37)	0.26 (0.44)	**23.99**	**1015.18**	0.19 (0.39)	0.18 (0.39)	**2.21**	**62.60**
Separating	0.16 (0.36)	0.25 (0.43)	**22.89**	**968.94**	0.19 (0.40)	0.18 (0.39)	**2.64**	**74.78**
Age of firm	31.07 (16.67)	30.75 (16.24)	**1.98**	**87.89**	32.69 (16.25)	32.55 (16.33)	**0.86**	**24.32**
Age^2	1243.56 (1094.11)	1209.06 (1055.74)	**3.21**	**142.48**	1332.71 (1091.79)	1326.16 (1094.43)	**0.60**	**16.97**
Log total employment	3.83 (1.85)	5.73 (1.67)	**107.78**	**4861.28**	5.16 (1.56)	5.21 (1.48)	**3.28**	**92.99**
Churning	0.22 (0.26)	0.41 (0.34)	**65.64**	**2720.49**	0.32 (0.30)	0.31 (0.25)	**3.83**	**108.48**
Cnty unemployment	4.51 (1.39)	4.43 (1.26)	**6.26**	**281.73**	4.43 (1.27)	4.44 (1.31)	**0.74**	**20.82**
Pop. Density	945.08 (771.59)	1081.94 (756.76)	**17.91**	**792.04**	1003.15 (775.08)	1003.74 (776.89)	**0.08**	**2.15**
Share of Hispanic students	0.05 (0.05)	0.05 (0.05)	**10.92**	**476.10**	0.05 (0.05)	0.05 (0.05)	**0.38**	**10.81**
No of observations	73,555,444	26,313,397			16,042,501	16,042,501		

Table C2. Logit estimation of the probability of being employed by a firm that hires undocumented workers.

	Parameter Estimates (st. errors)	Marginal Effect
Worker Characteristics		
Worker tenure (in quarters)	-0.003***	0.000
	(0.000)	
Worker tenure squared / 1000	0.002	0.000
	(0.002)	
Worker experience (in quarters)	-0.021***	-0.003
	(0.000)	
Worker experience squared / 1000	0.148***	0.021
	(0.001)	
Worker is new hire = 1	-0.088***	-0.012
	(0.001)	
Worker is separating = 1	-0.025***	-0.003
	(0.001)	
Firm Characteristics		
Age of firm (in quarters)	0.012***	0.002
	(0.000)	
Age of firm squared / 100	0.000***	0.000
	(0.000)	
Log total employment (firm size)	0.726***	0.103
	(0.000)	
Churning at worker's firm (documented wrkrs only)	1.890***	0.267
	(0.001)	
Share of workers with higher education	-2.252***	-0.318
	(0.003)	
Labor intensity	-0.241***	-0.034
	(0.003)	
Geographic Controls		
County unemployment rate	-0.039***	-0.006
	(0.000)	
Population density /1000	0.000***	0.000
	(0.000)	
Percent of students that is Hispanic	3.072***	0.434
	(0.006)	
Intercept	-3.459***	
	(0.012)	
No. of Observations		

Note: Analysis includes documented workers employed in Georgia 1995-2005 inclusive. Also included as regressors are broad sector industry fixed effects (see Appendix B) and year and quarter fixed effects, as well as sector/year interactions. *** statistically significantly different from zero at the 99% confidence level, ** statistically significantly different from zero at the 95% confidence level, and * statistically significantly different from zero at the 90% confidence level..

Appendix D: Sector specific wage results.

Table D1. Linear, fixed-effects estimation of log wages, by sector.

Dependent variable = Ln wage	Agriculture	Construction	Manufacturing	Trans, Comm &Util
Firm hires undocumented wrkrs:	0.0577***	-0.0026*	0.0099***	0.046***
	(0.0103)	(0.0014)	(0.0009)	(0.0030)
Coefficients on percent undocumented interaction are available upon request.				
Order of preferred Specification K=	6	1	6	6
Worker Characteristics				
Worker tenure (in quarters)	0.0084***	0.0044***	0.0059***	0.0069***
	(0.0007)	(0.0002)	(0.0001)	(0.0006)
Worker tenure squared / 1000	-0.1181***	-0.0822***	-0.1155***	-0.1316***
	(0.0111)	(0.0033)	(0.0017)	(0.0080)
Worker experience (in quarters)	0.0036**	0.0174***	0.0123***	0.000141289
	(0.0016)	(0.0006)	(0.0004)	(0.0014)
Worker experience squared / 1000	-0.0704***	-0.1277***	-0.0561***	0.0256***
	(0.0111)	(0.0033)	(0.0018)	(0.0065)
Worker is new hire = 1	-0.5236***	-0.6787***	-0.7446***	-0.6918***
	(0.0049)	(0.0014)	(0.0009)	(0.0029)
Worker is separating = 1	-0.6911***	-0.8643***	-0.9279***	-0.9267***
	(0.0046)	(0.0014)	(0.0009)	(0.0029)
Firm Characteristics				
Age of firm (in quarters)	-0.0027	-0.0011	-0.005***	-0.005**
	(0.0025)	(0.0013)	(0.0009)	(0.0025)
Age of firm squared / 100	0.00004***	0.0001***	0.0001***	0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log total employment (firm size)	0.2455***	0.19***	0.1332***	0.163***
	(0.0046)	(0.0018)	(0.0013)	(0.0042)
Churning at worker's firm (among documented workers only)	0.2368***	0.107***	0.1213***	0.1627***
	(0.0082)	(0.0030)	(0.0025)	(0.0076)
Share of workers with higher education	-2.823**		0.03176	-0.08251
	(1.2616)		(0.0289)	(0.2191)
Labor intensity	-0.3536	-0.0486	-0.2308***	0.1337
	(0.2177)	(0.1159)	(0.0426)	(0.1313)
Geographic Controls				
County unemployment rate	-0.0048***	0.0023***	-0.0032***	-0.0326***
	(0.0014)	(0.0008)	(0.0003)	(0.0016)
Population density /1000	-0.00009***	0.000001	0.00004***	-0.0001***
	(0.00002)	(0.000003)	(0.000002)	(0.00001)
Percent of students that is Hispanic	-0.10857	-0.2191***	0.0841***	-1.0788***
	(0.0976)	(0.0322)	(0.0153)	(0.0825)
R-squared	0.7931	0.7850	0.8101	0.8131
No. of Observations	384,138	3,358,162	6,339,192	900,136

Table D1, cont.

Dependent variable = Ln wage	Wholesale Trade	Retail Trade	Financial Services	Information
Firm hires undocumented wrkrs:	0.0163***	-0.0024	0.0025	0.0253***
	(0.0027)	(0.0018)	(0.0023)	(0.0041)
Coefficients on percent undocumented interaction are available upon request.				
Order of preferred Specification K=	6	1	2	6
Worker Characteristics				
Worker tenure (in quarters)	0.0066***	0.0071***	-0.0004	0.0119***
	(0.0004)	(0.0004)	(0.0006)	(0.0009)
Worker tenure squared / 1000	-0.0983***	-0.0847***	-0.0576***	-0.1295***
	(0.0047)	(0.0057)	(0.0078)	(0.0109)
Worker experience (in quarters)	0.0052***	0.0075***	0.0366***	0.0028
	(0.0011)	(0.0012)	(0.0021)	(0.0028)
Worker experience squared / 1000	-0.0585***	-0.0993***	-0.0996***	-0.0461***
	(0.0045)	(0.0054)	(0.0075)	(0.0103)
Worker is new hire = 1	-0.7167***	-0.773***	-0.6995***	-0.6416***
	(0.0021)	(0.0021)	(0.0030)	(0.0043)
Worker is separating = 1	-0.8964***	-1.0242***	-0.954***	-0.7043***
	(0.0021)	(0.0021)	(0.0030)	(0.0041)
Firm Characteristics				
Age of firm (in quarters)	-0.0046**	-0.0128***	0.0198***	-0.0992***
	(0.0022)	(0.0026)	(0.0048)	(0.0064)
Age of firm squared / 100	0.00005***	0.0001***	0.0001***	0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log total employment (firm size)	0.1188***	0.0862***	0.0834***	0.0922***
	(0.0032)	(0.0038)	(0.0045)	(0.0058)
Churning at worker's firm (among documented workers only)	0.1052***	0.1733***	0.1553***	0.2039***
	(0.0054)	(0.0057)	(0.0098)	(0.0145)
Share of workers with higher education			0.00150	3.5482***
			(0.1698)	(1.3180)
Labor intensity			0.1508	0.8861***
			(0.1467)	(0.2736)
Geographic Controls				
County unemployment rate	-0.00689***	0.0048***	-0.0052**	0.011***
	(0.0011)	(0.0014)	(0.0022)	(0.0028)
Population density /1000	0.000004	-0.00003***	0.0001***	0.0001***
	(0.000004)	(0.00001)	(0.00001)	(0.00001)
Percent of students that is Hispanic	0.2644***	0.02538	-0.16630	-1.1929***
	(0.0372)	(0.0571)	(0.1027)	(0.1433)
R-squared	0.8416	0.8522	0.8380	0.8772
No. of Observations	1,242,614	1,567,908	653,832	320,958

Table D1, cont.

Dependent variable = Ln wage	Prof & Bus Services	Education & Health	Leisure & Hospitality	Other Services
Firm hires undocumented wrkrs:	0.0022	-0.003***	0.0037*	-0.0048
	(0.0017)	(0.0008)	(0.0021)	(0.0036)
Coefficients on percent undocumented interaction are available upon request.				
Order of preferred Specification K=	6	2	4	1
Worker Characteristics				
Worker tenure (in quarters)	0.0059***	0.0143***	0.0062***	0.0078***
	(0.0003)	(0.0002)	(0.0003)	(0.0008)
Worker tenure squared / 1000	-0.0656***	-0.1608***	-0.0414***	-0.1063***
	(0.0053)	(0.0021)	(0.0055)	(0.0113)
Worker experience (in quarters)	0.0102***	0.0149***	0.0053***	0.0034*
	(0.0008)	(0.0005)	(0.0008)	(0.0019)
Worker experience squared / 1000	-0.0637***	-0.0662***	-0.1042***	-0.0508***
	(0.0041)	(0.0024)	(0.0045)	(0.0098)
Worker is new hire = 1	-0.6697***	-0.6762***	-0.7234***	-0.6732***
	(0.0014)	(0.0012)	(0.0014)	(0.0037)
Worker is separating = 1	-0.8558***	-0.8565***	-0.9065***	-0.8469***
	(0.0014)	(0.0012)	(0.0014)	(0.0036)
Firm Characteristics				
Age of firm (in quarters)	-0.0195***	-0.0164***	-0.0114***	-0.0081**
	(0.0016)	(0.0020)	(0.0012)	(0.0038)
Age of firm squared / 100	0.0001***	0.0002***	0.00004***	0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log total employment (firm size)	0.1505***	0.161***	0.183***	0.0927***
	(0.0020)	(0.0025)	(0.0028)	(0.0058)
Churning at worker's firm (among documented workers only)	0.1379***	0.1268***	0.1503***	0.1104***
	(0.0039)	(0.0048)	(0.0036)	(0.0085)
Share of workers with higher education	0.0818***	0.1425**	0.10800	
	(0.0236)	(0.0638)	(0.1556)	
Labor intensity	-0.0123	0.5269***	0.3502	0.2264
	(0.0358)	(0.1548)	(0.2815)	(0.3267)
Geographic Controls				
County unemployment rate	-0.0106***	-0.0075***	-0.0046***	-0.0104***
	(0.0012)	(0.0005)	(0.0011)	(0.0028)
Population density /1000	0.000004	-0.00002***	0.00001*	0.00001
	(0.000004)	(0.00001)	(0.00001)	(0.00001)
Percent of students that is Hispanic	-0.02190	-0.1525***	-0.0946*	0.214*
	(0.0506)	(0.0301)	(0.0521)	(0.1286)
R-squared	0.8747	0.8304	0.8182	0.8498
No. of Observations	3,972,320	5,112,496	3,676,432	521,206

Note: See Appendix B for sector definitions. Standard errors in parentheses. Analysis includes documented workers employed in Georgia 1995-2005 inclusive. Also included as regressors are year and quarter fixed effects, as well as sector/year interactions. P_{jt} is measured from 0 to 100. *** statistically significantly different from zero at the 99% confidence level, ** statistically significantly different from zero at the 95% confidence level, and * statistically significantly different from zero at the 90% confidence level.