RESULTS ARE PRELIMINARY
PLEASE DO NOT CITE

# The Wage Impact of Undocumented Workers


Julie L. Hotchkiss
Federal Reserve Bank of Atlanta
and Georgia State University
Research Department
Federal Reserve Bank of Atlanta
1000 Peachtree Street, NE
Atlanta, GA 30309-4470
Julie.L.Hotchkiss@atl.frb.org
404/498-8198

Myriam Quispe-Agnoli
Federal Reserve Bank of Atlanta
Research Department
Federal Reserve Bank of Atlanta
1000 Peachtree Street, NE
Atlanta, GA 30309-4470
Myriam.Quispe-Agnoli@atl.frb.org
404/498-8930

Fernando Rios-Avila
Georgia State University
friosavila1@gsu.edu
404/924-5176


April 29, 2013


Abstract

Using administrative, individual-level, longitudinal data from the state of Georgia, this paper finds that rising shares of undocumented workers results in higher earnings for documented workers, but by a small amount. A one-percentage point increase in the share of undocumented workers in a documented worker's county/industry results in an average wage boost of 0.44 percent. Within the firm, a one-percentage point increase in the percent of undocumented workers employed by the firm boosts wages by 0.09 percent (0.12, 0.11, and 0.04 in low, medium, and high skill firms, respectively). Potential explanations for a positive wage impact are discussed.



JEL Classification Codes:
J30 - Wages, Compensation, and Labor Costs; General
J15 - Economics of Minorities, Races, and Immigrants


---


The opinions expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of Atlanta, nor the Federal Reserve System. 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, University of Kentucky, University of Georgia Department of Public Administration and Policy, Claremont-McKenna College, and Univesity of California-Riverside. The research assistance of Gustavo Canavire, M. Laurel Graefe, Gustavo Uceda, and Eric Wang is much appreciated.

# 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, 2001), 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 Iskander and Lowe 2011, 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, 2013). 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 earn higher wages as the share of workers both in their firm and in the local labor market increases.

*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 Act (IIRIRA) of 1996. Both of these laws were passed in response to the growing population of

---

[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).

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 11.1 million unauthorized immigrants living in the U.S. as of March 2011 (Pew Research Center 2013). It is also estimated that about four percent of the total are located in Georgia. Between 2000 and 2010, Georgia experienced one of the largest percentage increases

[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>.

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 (2009), 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 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

[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.

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 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

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

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 sectors with the highest incidences. The concentration of workers in these sectors has also been

[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.

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. In spite of undercounting the actual levels of undocumented workers, we are confident that those we identify constitute a reliable representation of the presence of undocumented workers in Georgia. Evidence supporting that confidence is detailed in Appendix A.

---

[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).

*B. 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, presume the workers 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, forthcoming), 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, 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.

*C. Sample Means*

Table 1 presents sample means for workers classified as documented for the full sample of Georgia data, as well as sample means grouped by the skill classification of workers' firms. Because of the use of lagged and forward looking regressors, only data from 1995 through 2005 are used in the analysis. However, worker and firm longitudinal characteristics are calculated beginning in 1990 (the first year of available data). There are over 178 million observations even in this restricted time period, making estimation with high order fixed effects cumbersome, at best (see Abowd et. al 1999). The sample will be additionally restricted with the elimination of all multi-establishment firms. 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. The single-establishment restriction also has the advantage of giving us greater confidence in the identification of the (county) location of the employer. 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).

[Table 1 here]

Restricting the sample to single-establishment firms reduces the number of observations to about 73.5 million (5.7 million unique individual). Workers overall earn roughly $9,400 per quarter on average, with workers employed by firms in low skill industries earning roughly half of what workers earn if employed by firms in high skill industries. Not only are firms in low skill industries more likely to be located in counties/industries that employ undocumented workers, among firms that hire them, undocumented workers make up a greater share of these firms' workforces.

Older firms appear to be more concentrated in high skill industries and younger firms in lower skill industries. In addition, larger firms appear to be more concentrated in high skill industries. Table 1 also shows how firms in each broad sector (based on one-digit NAICS) are concentrated across industry skill classifications (base on two- and three-digit NAICS; see Appendix B for details). Some sectors, such as Other Services, have firms represented in each skill classification, while other, such as Construction and Wholesale Trade are fully concentrated at one skill level. This will be useful to remember when we turn to the analysis by industry skill classifications.

We also see in the sample means that the percent of workers being hired and separating (as well as churning, which is a function of hires and separations and total employment) decreases with a firm's skill classification. In addition, workers with the shortest amount of tenure are concentrated in the lowest skill classification. These sample means are consistent with Morales (1983) who finds a greater degree of undocumented worker employment among firms with greater churning, suggesting that employment of undocumented workers is a form of achieving greater flexibility in face of greater churning.

## 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), in a number of papers by George Borjas (alone and with co-authors; 2003, 2005, 2006), and by Orrenius and Zavodny (2007). The procedure makes use of large data sets and standard linear regression to identify a relationship between the density of immigrants on wages or employment across aggregated geographies (usually metropolitan statistical areas, MSAs), industries, or across specific type of demographic groups, such as labor market experience and education. These analyses typically exploit changes in the immigration density over time.

In these studies, and others, the issue of potential endogeneity has been a significant concern when estimating the impact of immigration on wages. For example, undocumented workers might naturally be attracted to locations or firms where higher wages can be found, so the share of undocumented workers employed becomes a function of the wage paid by the firm. Additionally, it might be the case that firms that employ undocumented workers are also intrinsically less productive and, thus, pay lower wages, which might translate into a negative correlation between the presence of undocumented workers and documented worker wages. Finally, there may be unobservable factors that both increase (decrease) a documented worker's wage and the probability that a firm hires undocumented workers (or hires more of them), this will result in a spurious positive (negative) correlation between a worker's observed wage and a firm's observed hiring behavior.

Previous work has applied various techniques to control for the potential endogeneity. Borjas (2003) simply performs his analysis at the national level to be free from concerns of geographically different wages affecting migration of immigrants, and Orrenius and Zavodny (2007) make use of instrumental variables along with inclusion of area-year fixed effects. Even though the data used here are not rich enough to allow for any attempts at instrumental variables estimation, the analysis using detailed individual level data in this paper has a primary advantage of being able to control for unobservable (time-invariant) firm and worker characteristics through the inclusion of fixed effects. The estimation also includes county, industry, and year-quarter fixed effects that should capture potential supply effects and aggregate firms' wage setting policies in the area. As a result, we should have a good chance of identifying the impact of the presence of undocumented workers geographically, within sector, and at the firm level on an individual documented worker's wage.

*A. Estimation at the County/Industry Level*

Wage equation models at three different levels of aggregation will be estimated in order to compare the results of undocumented workers to previous findings in the literature, and to provide evidence on the sensitivity of those estimations when data at a more disaggregated level are available. The first equation follows the specification similar to the one proposed Orrenius and Zavodny (2007), using aggregate information at the county (159 counties), industry (12 levels), and year and quarter information:

$$\ln \bar{w}_{c,k,t} = \alpha + \beta P_{c,k,t} + \gamma' \bar{X}_{c,k,t} + \varsigma_c + \iota_k + \tau_t + \varepsilon_{c,k,t} \quad , \tag{1}$$

where $\ln \bar{w}_{c,k,t}$ and $P_{c,k,t}$ are the log of the average wage among documented workers and percentage undocumented workers, respectively, in county *c*, industry *k*, and quarter *t*. The vector $X_{c,k,t}$ controls for average worker and firm characteristics for each county/industry cell in quarter *t*. We also control for county $\varsigma_c$, industry $\iota_k$, and quarter $\tau_t$ specific effects, to control for unobservable determinants of earnings. In the estimation, each county/industry cell in quarter *t* is weighted by the total number of workers represented by that cell. Firm characteristics, such as age and total employment (firm size) are categorized (as was seen in Table 1) so as to better capture the shape of the distribution of that characteristic at the county/industry level. This categorization (rather than using the continuous version of the characteristics) is retained for less aggregated analyses (at the firm and workers level) so that the estimation results can be as comparable as possible across levels of aggregation.

*B. Estimation at the Firm Level*

While the results of the specification in equation (1) are useful to compare with findings elsewhere in the literature, a geographically aggregated analysis does not allow us to control for firm level heterogeneity. Therefore, we estimate a regression similar to equation (1), but disaggregate information to the firm level:

$$\ln \bar{w}_{j,t} = \alpha + \beta_1 P_{j,t} + \beta_2 P_{c,k,t} + \gamma' \bar{X}_{j,t} + \varsigma_c + \iota_k + \tau_t + \varphi_j + \varepsilon_{j,t} \,, \tag{2}$$

where $\ln \bar{w}_{j,t}$ is the log of the average wage level of documented workers in the firm *j* at quarter *t*, $P_{j,t}$ is the percentage of undocumented workers in the firm, and $X_{j,t}$ is a vector of firm and worker characteristics, averaged at the firm level. To allow for an effect of the local potential supply of undocumented workers on average firm wages, $P_{c,k,t}$ is also included in the regression. In addition to the area, industry, and time specific fixed effects, equation (2) also includes firm specific fixed effects, $\varphi_j$. Each observation is weighted by the number of documented workers in each firm at time *t*.

*C. Estimation at the Individual Level*

Although the firm level estimation has advantages over the geographically aggregated estimation by being able to control for firm level heterogeneity in the determination of documented worker wages, the results may potentially suffer from composition bias. Undocumented workers tend to be lower skilled and be paid lower wages on average than documented workers (for example, see Pena, 2010 and Hotchkiss and Quispe-Agnoli, 2013). If firms replace their low-wage documented workers with low-wage undocumented workers, the average documented wage at the firm will rise -- not because firms employing undocumented workers pay their documented workers more, but because the bottom of the documented worker wage distribution has been chopped off. Use of individual level data eliminates this problem of composition bias by producing a within-worker estimate of the impact of the presence of undocumented workers.

In order to identify an unbiased, within-worker impact of undocumented workers on the wages of documented workers, we finally turn to a wage equation that is estimated at the individual worker level:

$$\ln w_{i,j,t} = \alpha + \beta_1 P_{j,t} + \beta_2 P_{c,k,t} + \gamma_1' X_{i,j,t} + \varsigma_c + \iota_k + \tau_t + \varphi_j + \theta_i + \varepsilon_{i,j,t} \ , \qquad (3)$$

where $\ln w_{i,j,t}$ is the log of the wage of individual *i*, working in firm *j*, in quarter *t*. Extending the specification of equation (2), equation (3) includes measures for the percentage of undocumented workers in the firm, $P_{j,t}$, and the local job market, $P_{c,k,t}$. It also includes worker and firm specific characteristics, but now worker characteristics are measured individually. Equation (3) also includes an individual fixed effect $\theta_i$, which will control for un-measured, time-invariant individual effects and produce a within-worker estimate of the impact of the presence of undocumented workers, at the firm and in the worker's county/industry. *At the present time, we are only able to include a nested firm*worker fixed effect, instead of the more appropriate approach of identifying individual and firm effects separately. Due to technological constraints including firm and worker fixed effects separately is too complex for available computing resources. We are investigating alternative estimation strategies.*

*D. Regressors*

Worker tenure is expected to positively influence wages through the presence of firm-specific human capital (Campbell 1993 and Altonji and Shakatko 1987). Individual general human capital (such as education, which is not available in the data) will be captured by the individual fixed effect. 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 will typically not have received a full quarter's worth of wages.

Firm size (categories) is included with the expectation that larger firms pay higher wages (Oi and Idson 1999). Firm age (categories) 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 (calculated 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).

Other characteristics of the firm that might be 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, forthcoming), but a full analysis of long-term labor market outcomes of potentially displaced documented workers will be the subject of a future investigation.

## IV. Estimation Results

### *A. Estimation at the County/Industry Level*

Table 2 contains the results from estimating equation (1) for two specifications -- one with an overall average impact and one that estimates a separate impact of undocumented workers by firms' skill classification. From the literature, we expect that the impact of the presence of undocumented workers will be felt most acutely by those workers who are most substitutable. The impact of average worker and firm characteristics on average wages earned by documented workers in each cell is as expected. The greater amount of worker churning is associated with lower average cell wages. Average wages increase with average firm size, but are highest where firms are youngest. Average wages are increasing in worker tenure, at a decreasing rate.[9] And, the average rate of separation/hire is not statistically significant, which may not be surprising if churning is picking up all the variation in hires and separations at the cell level. Employment growth in the county (entered as a proxy for worker demand) seems to be insignificantly related to average wages in the county/industry/time cell.

[Table 2 here]

Turning now to the regressor of interest, a greater percent of undocumented workers in the county/industry cell is associated with significantly lower wages. A one-percentage point increase in the share of undocumented workers in the local labor force is associated with a reduction in wages of 0.59 percent. Borjas (2003) finds a 0.3 percent reduction in wages and Onnenius and Zavodny (2007) report a 0.8 percent reduction for a similar one percent increase in immigrants. Altonji and Card (1991), using a different estimation approach, estimate between a 0.3 and 1.2 percent negative impact on native wages with a one-percentage point increase in the

[9] Results including average worker experience are similar to those discussed here, but omitted from the final regression due to a high degree of collinearity between average tenure and experience.

fraction of immigrants in an area. So the results in Table 2 are comparable to estimates found by others pertaining to the impact of immigrants as a whole.

Others (including those authors cited above) have found that the impact of the presence of immigrants differs by skill, education, or experience of workers, so specification (2) fully interacts the presence of undocumented workers in the county/industry cell with the share of firms in the cell in each skill classification. Consistent with what others have found, average wages decline even further as the share of firms that are low and middle skill increases. There is a mitigation of the negative impact of the presence of undocumented workers as the share of firms that are high skill increases. This is consistent with a complementarity or scale effect that might result from the arrival of an additional inexpensive production input; those workers employed by firms in high skill industries would be the least substitutable with the newly arriving cheaper labor.

*B. Estimation at the Firm Level*

Although the evidence from Table 2 suggests that the presence of undocumented workers has an important negative impact on wages of documented workers, the problem of endogeneity looms and we turn now to equation (2) which is estimated at the firm level and allows for control of firm level heterogeneity and fixed effects. Table 3 presents estimates for the analogous specification in Table 2, at the firm level and both with and without firm fixed effects in order to get a sense how much controlling for firm specific, time invariant characteristics matters.

[Table 3 here]

The results in column (1) of Table 3 suggest that the bulk of the negative impact on wages occurs from a rise in the share of undocumented workers in the firm versus a rise in the share of undocumented workers in the county/industry. A one-percentage point increase in the share of undocumented workers in the firm lowers average documented worker wages by one percent, whereas a one-percentage point increase in the share of undocumented workers in the county/industry lowers (marginally) documented worker wages by 0.2 percent. Of course, because the total number of workers is larger, a one-percentage point increase in a county represents a much larger number of workers than a one-percentage point increase in the share of workers at the firm level. The same pattern of a greater negative impact being felt among county/industry cells with a higher share of low skill firms is seen in column (2). However, at the firm level the greatest negative impact is felt among workers in firms classified as middle skill (recall, this is not a classification of worker skill, but the classification of the industry in which the firm operates).

The source of our concern that these results suffer from endogeneity, however, appears when considering the estimates that include firm fixed effects. These results are contained in columns (3) and (4) of Table 3. At both the firm and at the county/industry cell level, a larger share of undocumented workers is associated with *higher* wages among documented workers. A one-percentage point increase in the share of undocumented workers in the firm is associated with wages that are 0.14 percent higher. A one-percentage point increase in the share of undocumented workers in the local labor market (county/industry) is associated with wages that are 0.32 percent higher.

The inclusion of fixed effects influences the estimated relationship between other firm and averaged worker characteristics, as well. The relationship between firm size and wages even flip signs -- after controlling for firm specific fixed effects, average wages in larger firms is lower than in smaller firms. Most of the rest of the parameters merely become smaller (in absolute value) with the inclusion of firm fixed effects. This suggests that the firm fixed effects are effectively capturing firm specific wage determining characteristics not captured by the observed regressors. Percentage growth in county employment remains an insignificant determinant of wages.

The estimated impact of undocumented workers on wages reported in Table 3 are within-firm estimates, which means that they may suffer what we called earlier "composition bias." In other words, as the firm replaces low-skill, low-paid workers with lower-paid undocumented workers, the average wage among the remaining documented workers rises. The overall positive impact at the county/industry level may also suffer from a similar composition bias if workers migrate in response to an influx of undocumented workers (see Frey 1996 and Borjas 2005, although Card and DiNardo 2000 and Card 2001 find no evidence of out-migration impacts). In addition, localities with a large influx of low-skilled workers may differ in their adoption of technology, which might impact wages among native workers.[10]

*C. Estimation at the Individual Level*

We now turn to estimation of the wage equation using individual worker level data with the expectation that a within-worker estimate of the impact of an increased presence of undocumented workers will be free of endogeneity and composition bias. (*At the present time we are only able to report a nested worker/firm fixed effect result due to technical limitations.*) Table 4 reports these estimation results with varying inclusion of fixed effects.

[Table 4 here]

Columns (1) and (2) report the results at the individual level with no controls for firm or individual fixed effects. We once again see a negative association between worker's wages and an increasing share of undocumented workers, both in the firm and in the county/industry where

[10] Card and Lewis (2005) find little evidence of geographic adjustments in technology adoption. Also see Dunne and Troske 2005.

the worker is employed. A one-percentage point increase in the share of undocumented workers in the firm (county/industry) lower's a worker's wage by 0.97 (0.76) percent. The largest negative impact is found among firms classified as middle skill and in counties increasing in the percent of firms classified as middle skill. We also now see a significantly *negative* impact of county employment growth on workers' wages.

The negative impact on wages of the presence of undocumented workers is retained at the individual level even with the inclusion of individual fixed effects only, although the impact is reduced considerably. These results are in columns (3) and (4). A one percentage point increase in the percent of undocumented workers in the firm (in the worker's county/industry) lowers individual wages on average by 0.4% (0.08%).

Columns (5) and (6) once again demonstrate how important controlling for firm fixed effects is in estimating the impact of undocumented workers on documented worker wages. It's clear that employment of undocumented workers and wages paid by the firm are highly correlated. In other words, it is not the presence of undocumented workers that leads to lower wages in columns (1), (2), (3), and (4), but, rather, being employed by a firm that is more likely to hire undocumented workers.

Columns (7) and (8) combine the firm and individual fixed effect as a nested effect, so each worker/firm combination gets a time-invariant control -- not yet being able to include these fixed effects separately means that we lose information about the individual when he/she changes employers. These estimates are to be interpreted as within worker/within firm estimates. Knowing that the biggest effect comes from controlling for firm fixed effects, however, does suggest that these nested effects will be close to those that use individual information *across* firms.

The positive relationship between the presence of undocumented workers, at both the firm and in the local labor market seen in in columns (5) and (6) persist even with the addition of individual fixed effects, although the coefficients are of smaller magnitude (especially with increases of undocumented workers within the firm). A one-percentage point increase in the share of workers in the firm (in the worker's county/industry) that is undocumented increases workers wages by 0.09 (0.44) percent. The largest boost to wages accrues to workers in firms classified as low skill and as the share of low skill workers in the county/industry increases. This is the first specification (with a positive coefficient on undocumented workers) where low skill is associated with the highest returns. This will be addressed below.

These individual level results (with any combination of fixed effects) is also the first estimations where greater employment growth in the worker's county produces a positive wage effect, as we might expect if greater demand for workers is associated with higher wages. In addition, higher churning at the firm is related to higher individual wages, as is being employed by a larger and older firm. Workers with greater tenure earn higher wages (at an increasing rate) and earn lower wages if they either separated from or were hired by their firm during the quarter.

*D. Individual Level Results by Employer's Skill Classification*

We now return to the earlier mentioned result that workers employed by firms classified as low skill earn the greatest boost as either their employer hires more undocumented workers or as the share of undocumented workers in the workers county/industry increases. These results are the focus of Table 5, where the individual level estimation is performed separately by the skill level classification of the firm in which the worker is employed. This table contains the nested firm*worker fixed effects.

[Table 5 here]

Workers in low skill firms get a 0.1 (1.13) percent boost in wages for every percentage point increase in the share of undocumented workers in the worker's firm (county/industry), with workers in medium skill firms getting a similar boost. The county/industry result is an order of magnitude higher than seen in medium or high skill firms. How can a larger presence of undocumented workers result in what might seem to be a counter-intuitive result of higher earnings for documented workers? The most appealing explanation is found in earlier work by Peri (2009) and Peri and Sparber (2009), who argue that the arrival of unskilled undocumented workers with limited English capabilities allows documented workers to exploit their comparative advantage in higher productivity tasks, especially those requiring good communication skills (also see Iskander and Lowe 2011, Toussaint-Comeau 2007, and Cobb-Clark et al. 1995). Not only may these tasks tap into the higher productivity of documented workers, but merely being able to specialize increases a documented worker's productivity, thus resulting in higher earnings. Peri and Sparber (2009) estimate that a net one percent increase in employment of immigrants increases income per native worker from 0.6 to 0.9 percent -- consistent with the county/industry results reported here for workers employed by firms classified as low skill (results are considerably smaller at the firm level).[11] It is among workers employed in the lowest skill county/industries (where the undocumented workers are most likely to be employed) where we would expect to see the greatest opportunities for specialization.

The finding that workers employed at medium and high skill firms also benefit (although by a smaller amount) from a rising share of undocumented workers suggests that there may be more than one mechanism at work. These workers may be benefiting from a straight-forward

[11] Another explanation is suggested by research from Brown et al. (forthcoming), who find that firms who employ undocumented workers enjoy a competitive advantage over firms who do not. If firms share these rents with their documented workers, it could result in higher wages. However, given the fairly small estimates from the firm share, this explanation doesn't hold much weight.

scale effect. As a less expensive factor input becomes available, firms increase production, increasing demand for all workers. It's also important to keep in mind the results in Table 5 are not based on the skill levels of the workers per se, but, rather, the skill classification of the employer (see Appendix B), so employers in each skill classification will employ workers of all skill levels. And even those low skill workers in high skill firms will benefit from the arrival of undocumented workers, although the average impact will be smaller as there are likely fewer low skill workers in high skill firms than would be found in low skill firms.

Other results across skill classification of firms are also notable. County growth appears to only positively benefit workers employed by firms classified as middle and high skill. Workers in low skill firms appear to get the biggest return to being employed by larger firms, however it is workers employed in high skill firms that get the greatest return to their firm tenure.

## V. Conclusions and Implications for Policy

Using individual worker level data, linked to employer characteristics, the analysis in this paper finds that documented worker wages rise as the share of undocumented workers at the firm and in the local labor market increases, and that this positive wage impact is largest among workers employed by firms classified as low skill.

As a possible explanation for these mostly positive wage effects, Peri (2009) and Peri and Sparber (2009) suggest that efficiency and productivity can benefit from the task specialization that is likely to result as low-skill immigrants are hired 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, such as communication skills. The positive wage effect that also accrues to workers employed by middle and high skill firms suggests the possibility of a scale effect, as well. The positive wage effect in firms classified as middle and high skill may also be partially reflecting the fact that these firms also employ some lower-skill workers who can also benefit from specializing if their firm employs undocumented workers.

In assessing the results in this paper, it is important to keep in mind that the estimated effects of increasing shares of undocumented workers is very small. The thought experiments of a one-percentage point increase in the share of undocumented workers at the firm or at the county/industry level represents would be considered a very large increase in (roughly a doubling of) the presence of undocumented workers. 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; Julie L. Hotchkiss; and Myriam Quispe-Agnoli. "Does Hiring Undocumented Workers Give Firms a Competitive Advantage?" *Journal of Regional Science* (forthcoming).

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. "The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts." *NBER Working Paper #11552* (August 2005).

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).

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).

Dunne, Timothy and Kenneth Troske. "Technology Adoption and the Skill Mix of U.S. Manufacturing Plants." *Scottish Journal of Political Economy* 52(3) (July 2005): 387-405.

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. 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.

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.

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. "The Expected Impact of State Immigration Legislation on Labor Market Outcomes." *Journal of Policy Analysis and Management* 32(1) (Winter 2013): 34-59.

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.

Kossoudji, 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.

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).

MacQueen, J.B. "Some Methods for Classification and Analysis of Multivariate Observations." *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley: Berkeley University Press, 1967): 281-97.

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.

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.

Pew Research Center. "Unauthorized Immigrants: How Pew Research Counts them and What We Know About Them." 17 April 2013. http://www.pewresearch.org/2013/04/17/unauthorized-immigrants-how-pew-research-counts-them-and-what-we-know-about-them/ (accessed 19 April 2013).

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.

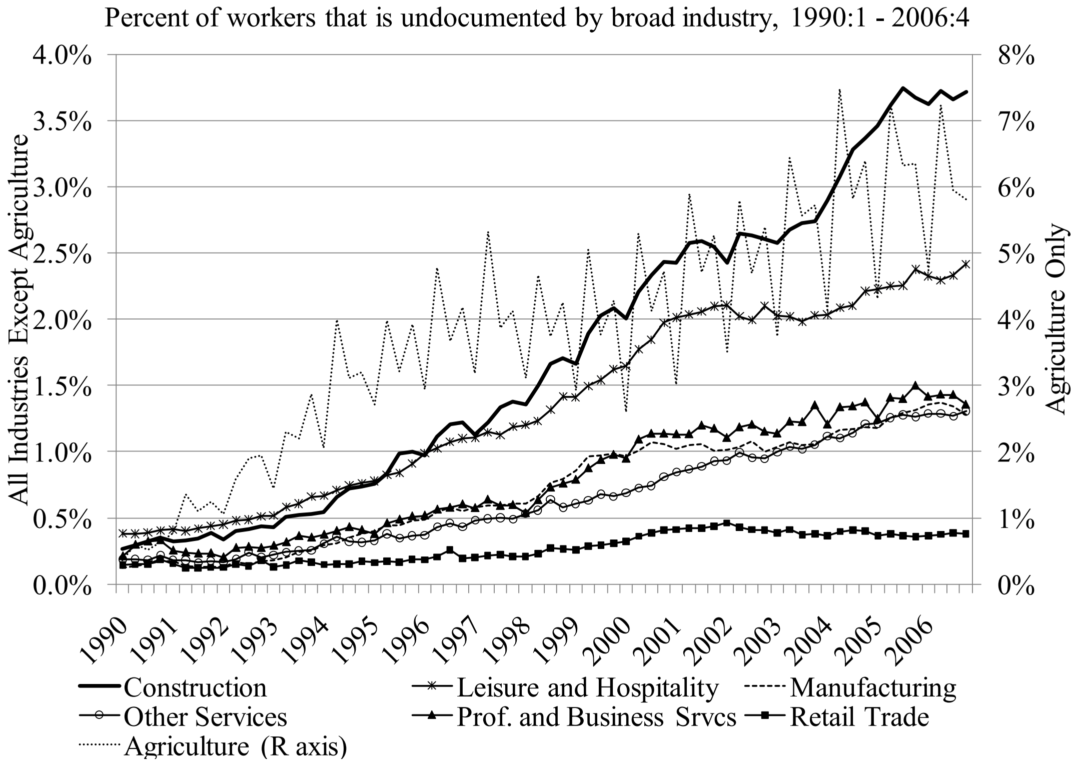

Percent of workers that is undocumented by broad industry, 1990:1 - 2006:4
All Industries Except Agriculture
Agriculture Only
4.0%
3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
8%
7%
6%
5%
4%
3%
2%
1%
0%
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
Construction
Leisure and Hospitality
Manufacturing
Other Services
Prof. and Business Srvcs
Retail Trade
Agriculture (R axis)

Table 1. Sample means, documented workers employed in single-establishment firms, 1995-2005, Georgia, by skill classification of employer.

| | Full | Low Skill | Middle Skill | High Skill |
|---|---|---|---|---|
| Wages (real qtrly earnings) | 9380.91 | 6663.01 | 8597.60 | 12422.37 |
| | (13204.96) | (9279.79) | (11494.39) | (16681.08) |
| % workers in firms that hire undoc workers | 25.82 | 35.72 | 26.01 | 18.02 |
| | (43.76) | (47.92) | (43.87) | (38.43) |
| % undocumented in the firm | 0.84 | 1.88 | 0.76 | 0.15 |
| | (2.89) | (4.26) | (2.65) | (1.09) |
| % undocumented in the firm's cnty/ind/qtr | 0.91 | 1.83 | 0.74 | 0.42 |
| | (1.31) | (1.98) | (0.84) | (0.61) |
| Y/Y % growth in county employment | 2.03% | 2.10% | 2.13% | 1.85% |
| | (0.097) | (0.109) | (0.099) | (0.084) |
| *Firm Age* | | | | |
| Firms age 1-4 quarters | 0.193 | 0.228 | 0.190 | 0.169 |
| | (0.394) | (0.420) | (0.392) | (0.375) |
| Firms age 5-12 quarters | 0.438 | 0.435 | 0.445 | 0.431 |
| | (0.496) | (0.496) | (0.497) | (0.495) |
| Firms age 13-24 quarters | 0.185 | 0.176 | 0.186 | 0.190 |
| | (0.388) | (0.381) | (0.389) | (0.392) |
| Firms age 25+ quarters | 0.185 | 0.161 | 0.179 | 0.210 |
| | (0.39) | (0.37) | (0.38) | (0.41) |
| *Firm Size* | | | | |
| Firms 1-9 workers | 0.142 | 0.162 | 0.124 | 0.148 |
| | (0.349) | (0.369) | (0.329) | (0.356) |
| Firms 10-49 workers | 0.287 | 0.353 | 0.287 | 0.237 |
| | (0.452) | (0.478) | (0.452) | (0.425) |
| Firms 50-249 workers | 0.291 | 0.314 | 0.324 | 0.232 |
| | (0.454) | (0.464) | (0.468) | (0.422) |
| Firms 250+ workers | 0.280 | 0.171 | 0.265 | 0.382 |
| | (0.449) | (0.376) | (0.441) | (0.486) |
| *Distribution of Firms, weighted by employment, across skill classification by sector* | | | | |
| Agriculture | | 87.43 | 12.57 | 0.00 |
| Construction | | 100.00 | 0.00 | 0.00 |
| Manufacturing | | 32.84 | 56.53 | 10.64 |
| Transport | | 0.00 | 76.55 | 23.45 |
| Whole sale | | 0.00 | 100.00 | 0.00 |
| Retail | | 18.40 | 77.68 | 3.92 |
| Financial Services | | 0.00 | 9.02 | 90.98 |
| Information | | 0.00 | 27.24 | 72.76 |
| Professional Services | | 0.00 | 49.58 | 50.42 |
| Education & Health | | 0.00 | 18.20 | 81.80 |
| Leisure & Hospitality | | 79.14 | 20.86 | 0.00 |
| Other Services | | 45.48 | 31.64 | 22.88 |
| *Total* | | *25.45* | *41.20* | *33.35* |
| Churning among doc Workers | 0.23 | 0.30 | 0.26 | 0.14 |
| | (0.25) | (0.26) | (0.27) | (0.16) |
| Percent of workers Hires or Separated | 22.914 | 28.381 | 25.033 | 16.125 |

| | Full | Low Skill | Middle Skill | High Skill |
|---|---|---|---|---|
| in the current quarter | (42.028) | (45.085) | (43.320) | (36.776) |
| *Worker Tenure within the firm (quarters)* | | | | |
| Workers Ten (1-4 qrts)=1 | 0.363 | 0.424 | 0.388 | 0.287 |
| | (0.481) | (0.494) | (0.487) | (0.452) |
| Workers Ten (5-8 qrts)=1 | 0.170 | 0.169 | 0.166 | 0.175 |
| | (0.375) | (0.375) | (0.372) | (0.380) |
| Workers Ten (9-16 qrts)=1 | 0.175 | 0.160 | 0.167 | 0.195 |
| | (0.380) | (0.367) | (0.373) | (0.396) |
| Workers Ten (17-28 qrts)=1 | 0.142 | 0.123 | 0.135 | 0.164 |
| | (0.349) | (0.328) | (0.342) | (0.370) |
| Workers Ten (29+ qrts)=1 | 0.151 | 0.124 | 0.144 | 0.180 |
| | (0.358) | (0.330) | (0.351) | (0.384) |
| N | 73,544,089 | 18,716,494 | 30,298,825 | 24,528,770 |
| Unique number of firms | 256,227 | 83,100 | 93,021 | 85,946 |
| Unique number of individuals | 5,720,423 | 2,437,501 | 3,376,102 | 2,237,834 |

Notes: Wages are real quarterly earnings, deflated by the chained price index for personal consumption expenditure $2006Q4. Individual sample means are across workers. Worker tenure and firm age are calculated using all of data available (starting in 1990), although the sample only contains observations from the years 1995-2005. Standard errors are in parentheses. A skill classification is assigned to a firm based on the average education of workers in the firm's three or two-digit NAICS (details described in Appendix B.

Table 2. Estimation at the county/industry level, 1995-2005.

| Dependent variable = log average quarterly earnings of documented workers in county/industry/quarter cell | (1) | (2) |
|---|---|---|
| % undocumented in the county/industry | -0.0059* | |
| | (0.0010) | |
| ... x % of firms that are low skill | | -4.10E-05* |
| | | (1.14E-05) |
| ... x % of firms that are medium skill | | -1.76E-04* |
| | | (2.08E-05) |
| ... x % of firms that are high skill | | 9.61E-05+ |
| | | (4.65E-05) |
| Y/Y % growth in county employment | 0.0004 | 0.0003 |
| | (0.0081) | (0.0080) |
| *Averaged Firm Characteristics* | | |
| Lag Churning among documented wrkrs | -0.6365* | -0.6316* |
| | (0.0271) | (0.0268) |
| % firms 10-49 workers | 0.2579* | 0.2565* |
| | (0.0124) | (0.0124) |
| % firms 50-249 workers | 0.4307* | 0.4298* |
| | (0.0117) | (0.0116) |
| % firms 250+ workers | 0.4364* | 0.4356* |
| | (0.0121) | (0.0120) |
| % firms age 5-12 quarters | -0.6078* | -0.6072* |
| | (0.0500) | (0.0499) |
| % firms age 13-24 quarters | -0.0679* | -0.0682* |
| | (0.0135) | (0.0135) |
| % firms age 25+ quarters | -0.1036* | -0.1018* |
| | (0.0160) | (0.0160) |
| *Averaged Worker Characteristics* | | |
| % workers Ten (5-8 qrts) | 0.8560* | 0.8510* |
| | (0.0494) | (0.0492) |
| % workers Ten (9-16 qrts) | 0.6687* | 0.6630* |
| | (0.0463) | (0.0462) |
| % workers Ten (17-28 qrts) | 0.5795* | 0.5743* |
| | (0.0473) | (0.0473) |
| % workers Ten (29+ qrts) | 0.5610* | 0.5576* |
| | (0.0413) | (0.0412) |
| % of workers hired or separated | -0.0098 | -0.0087 |
| | (0.0168) | (0.0168) |
| _cons | 7.9627* | 7.9641* |

| | | |
|---|---|---|
| | (0.0375) | (0.0374) |
| N (159 counties, 12 industries, 40 quarters) | 74157 | 74157 |
| R-Squared | 0.916 | 0.916 |

Notes: Regressors measured as percent range from 0 to 100. Standard errors in parentheses. *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. Estimation also includes county, industry, and quarter fixed effects. In the calculation of the share of firms in each cell that are low, middle, and high skill, each firm's classification is weighted by the employment in that firm (see Appendix B for details of how skill classification for each firm is determined). Lagged Churning is measured as the difference between worker flows and job flows divided by the average employment four quarters ago. 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).

Table 3. Estimation at the firm level, with and without firm fixed effects.

| Dependent variable = log average quarterly earnings of documented workers in the firm | No firm FE | | With Firm FE | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| % undoc employed in the firm | -0.0100* | | 0.0014* | |
| | (0.0002) | | (0.0001) | |
| ... x firm classified as low skill=1 | | -3.66E-05* | | 1.59E-05* |
| | | (1.99E-06) | | (1.41E-06) |
| ... x firm classified as medium skill=1 | | -2.00E-04* | | 1.45E-05* |
| | | (3.13E-06) | | (1.94E-06) |
| ... x firm classified as high skill=1 | | -6.28E-05* | | 8.71E-06+ |
| | | (5.84E-06) | | (3.41E-06) |
| % undoc in the county/industry (c/i) | -0.0019+ | | 0.0032* | |
| | (0.0008) | | (0.0005) | |
| ... x % of firms in c/i that are low skill | | -6.13E-05* | | 3.33E-05* |
| | | (8.81E-06) | | (6.16E-06) |
| ... x % of firms in c/i that are mid skill | | -5.91E-05* | | -3.18E-05* |
| | | (1.69E-05) | | (9.73E-06) |
| ... x % of firms in c/i that are high skill | | 2.02E-04* | | 1.82E-04* |
| | | (3.89E-05) | | (1.72E-05) |
| Y/Y % growth in county employment | 0.0025 | 0.0029 | -0.0017 | -0.0017 |
| | (0.0066) | (0.0066) | (0.0025) | (0.0025) |
| *Averaged Firm Characteristics* | | | | |
| Lag Churning among documented wrkrs | -0.6606* | -0.6562* | -0.0048* | -0.0048* |
| | (0.0060) | (0.0059) | (0.0016) | (0.0016) |
| % firms 10-49 workers | 0.2868* | 0.2862* | -0.0183* | -0.0184* |
| | (0.0011) | (0.0011) | (0.0008) | (0.0008) |
| % firms 50-249 workers | 0.3854* | 0.3843* | -0.0514* | -0.0515* |
| | (0.0017) | (0.0017) | (0.0014) | (0.0014) |
| % firms 250+ workers | 0.3344* | 0.3341* | -0.0823* | -0.0824* |
| | (0.0030) | (0.0030) | (0.0030) | (0.0030) |
| % firms age 5-12 quarters | -0.8508* | -0.8487* | -0.4152* | -0.4152* |
| | (0.0079) | (0.0079) | (0.0037) | (0.0037) |
| % firms age 13-24 quarters | -0.0133* | -0.0130* | -0.0007 | -0.0008 |
| | (0.0031) | (0.0031) | (0.0015) | (0.0015) |
| % firms age 25+ quarters | -0.0198* | -0.0193* | 0.0046^ | 0.0046^ |
| | (0.0035) | (0.0035) | (0.0024) | (0.0024) |
| *Averaged Worker Characteristics* | | | | |
| % workers Ten (5-8 qrts) | 0.4856* | 0.4831* | 0.1325* | 0.1325* |
| | (0.0068) | (0.0068) | (0.0027) | (0.0027) |

| Dependent variable = log average quarterly earnings of documented workers in the firm | No firm FE | | With Firm FE | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| % workers Ten (9-16 qrts) | 0.4593* | 0.4575* | 0.1821* | 0.1819* |
| | (0.0065) | (0.0065) | (0.0027) | (0.0027) |
| % workers Ten (17-28 qrts) | 0.4726* | 0.4714* | 0.2360* | 0.2357* |
| | (0.0077) | (0.0077) | (0.0031) | (0.0031) |
| % workers Ten (29+ qrts) | 0.4647* | 0.4638* | 0.2420* | 0.2419* |
| | (0.0072) | (0.0072) | (0.0039) | (0.0039) |
| % of workers hired or separated | 0.0105* | 0.0109* | 0.0126* | 0.0127* |
| | (0.0038) | (0.0038) | (0.0032) | (0.0032) |
| _cons | 7.9298* | 7.9277* | 8.6921* | 8.6926* |
| | (0.0119) | (0.0119) | (0.0263) | (0.0263) |
| N (256,227 unique firms) | 4652603 | 4652603 | 4652603 | 4652603 |
| R2 | 0.526 | 0.527 | 0.919 | 0.919 |

Notes: See notes to Table 2.

Table 4. Estimation at the individual level, with and without various fixed effects.

| Dependent variable = log average quarterly earnings of documented workers in the firm | No FE | | Individual FE only | | Firm FE only | | Nested FE | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| % undoc employed in the firm | -0.0097* | | -0.0040* | | 0.0022* | | 0.0009* | |
| | (0.0000) | | (0.0000) | | (0.0001) | | (0.0001) | |
| ... x firm classified as low skill=1 | | -1.39E-05* | | -5.05E-06* | | 2.57E-05* | | 1.17E-05* |
| | | (5.79E-07) | | (6.22E-07) | | (8.84E-07) | | (7.05E-07) |
| ... x firm classified as medium skill=1 | | -2.20E-04* | | -9.38E-05* | | 2.13E-05* | | 7.23E-06* |
| | | (7.02E-07) | | (7.32E-07) | | (1.07E-06) | | (8.57E-07) |
| ... x firm classified as high skill=1 | | -1.04E-04* | | -1.57E-05* | | 1.17E-06 | | 2.25E-06* |
| | | (1.86E-06) | | (1.61E-06) | | (2.16E-06) | | (0.0000) |
| % undoc in the county/industry (c/i) | -0.0076* | | -0.0008* | | 0.0050* | | 0.0044* | |
| | (0.0001) | | (0.0001) | | (0.0002) | | (0.0002) | |
| ... x % of firms in c/i that are low skill | | -1.22E-04* | | -2.25E-05* | | 5.15E-05* | | 6.65E-05* |
| | | (1.65E-06) | | (1.74E-06) | | (2.65E-06) | | (2.12E-06) |
| ... x % of firms in c/i that are mid skill | | -1.82E-04* | | -8.12E-05* | | 8.88E-06^ | | -1.24E-05* |
| | | (3.62E-06) | | (3.43E-06) | | (4.88E-06) | | (3.75E-06) |
| ... x % of firms in c/i that are high skill | | 2.70E-04* | | 2.40E-04* | | 1.38E-04* | | 5.36E-05* |
| | | (7.06E-06) | | (6.23E-06) | | (8.49E-06) | | (6.43E-06) |
| Y/Y % growth in county employment | -0.0064* | -0.0057* | 0.0037* | 0.0039* | 0.0030* | 0.0030* | 0.0023* | 0.0024* |
| | (0.0013) | (0.0013) | (0.0009) | (0.0009) | (0.0011) | (0.0011) | (0.0008) | (0.0008) |
| *Firm Characteristics* | | | | | | | | |
| Lag Churning among documented wrkrs | -0.9157* | -0.9085* | -0.3340* | -0.3325* | 0.0265* | 0.0265* | 0.0741* | 0.0741* |
| | (0.0006) | (0.0006) | (0.0005) | (0.0005) | (0.0007) | (0.0007) | (0.0005) | (0.0005) |
| firm 10-49 workers = 1 | 0.2109* | 0.2102* | 0.0694* | 0.0693* | -0.0398* | -0.0398* | 0.0458* | 0.0458* |
| | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0007) | (0.0007) | (0.0005) | (0.0005) |
| firm 50-249 workers = 1 | 0.3132* | 0.3119* | 0.1147* | 0.1147* | -0.0553* | -0.0553* | 0.1191* | 0.1190* |
| | (0.0004) | (0.0004) | (0.0005) | (0.0005) | (0.0009) | (0.0009) | (0.0007) | (0.0007) |

| Dependent variable = log average quarterly earnings of documented workers in the firm | No FE | | Individual FE only | | Firm FE only | | Nested FE | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| firm 250+ workers = 1 | 0.2462* | 0.2457* | 0.0953* | 0.0961* | -0.0498* | -0.0499* | 0.2000* | 0.1999* |
| | (0.0004) | (0.0004) | (0.0005) | (0.0005) | (0.0012) | (0.0012) | (0.0009) | (0.0009) |
| firm age 5-12 quarters = 1 | -0.0908* | -0.0905* | -0.0125* | -0.0126* | -0.0132* | -0.0132* | -0.0059* | -0.0059* |
| | (0.0003) | (0.0003) | (0.0003) | (0.0003) | (0.0006) | (0.0006) | (0.0005) | (0.0005) |
| firm age 13-24 quarters = 1 | -0.1414* | -0.1408* | -0.0342* | -0.0343* | 0.0082* | 0.0082* | 0.0034* | 0.0034* |
| | (0.0005) | (0.0005) | (0.0004) | (0.0004) | (0.0010) | (0.0010) | (0.0007) | (0.0007) |
| firm age 25+ quarters = 1 | -0.1394* | -0.1388* | -0.0489* | -0.0489* | 0.0442* | 0.0442* | 0.0232* | 0.0232* |
| | (0.0005) | (0.0005) | (0.0005) | (0.0005) | (0.0013) | (0.0013) | (0.0009) | (0.0009) |
| *Worker Characteristics* | | | | | | | | |
| worker Ten (5-8 qrts) = 1 | 0.2299* | 0.2295* | 0.0686* | 0.0685* | 0.1827* | 0.1827* | 0.0236* | 0.0236* |
| | (0.0004) | (0.0004) | (0.0003) | (0.0003) | (0.0003) | (0.0003) | (0.0002) | (0.0002) |
| worker Ten (9-16 qrts) = 1 | 0.3582* | 0.3579* | 0.1045* | 0.1044* | 0.3278* | 0.3278* | 0.0736* | 0.0736* |
| | (0.0004) | (0.0004) | (0.0003) | (0.0003) | (0.0003) | (0.0003) | (0.0003) | (0.0003) |
| worker Ten (17-28 qrts) = 1 | 0.5103* | 0.5100* | 0.1454* | 0.1453* | 0.4988* | 0.4988* | 0.1292* | 0.1292* |
| | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) | (0.0004) |
| worker Ten (29+ qrts) = 1 | 0.6673* | 0.6672* | 0.1209* | 0.1207* | 0.6920* | 0.6920* | 0.1539* | 0.1540* |
| | (0.0004) | (0.0004) | (0.0005) | (0.0005) | (0.0004) | (0.0004) | (0.0006) | (0.0006) |
| Hired or separated = 1 | -1.0536* | -1.0531* | -0.8198* | -0.8197* | -0.9873* | -0.9873* | -0.7001* | -0.7001* |
| | (0.0003) | (0.0003) | (0.0002) | (0.0002) | (0.0003) | (0.0003) | (0.0002) | (0.0002) |
| _cons | 7.9821* | 7.9826* | | | | | | |
| | (0.0044) | (0.0044) | | | | | | |
| N = 73,544,089<br>(5,720,423 unique workers; 256,227 unique firms)<br>R2 | 0.3895 | 0.3900 | 0.7454 | 0.7454 | 0.5417 | 0.5417 | 0.8570 | 0.8570 |

Notes: See notes to Table 2.

Table 5. Estimation at the individual level, nested fixed effects by skill classification of employer.

| Dependent variable = log average quarterly earnings of documented workers in the firm | Low | Medium | High |
|---|---|---|---|
| % undoc employed in the firm | 0.0012* | 0.0011* | 0.0004* |
| | (0.0001) | (0.0001) | (0.0001) |
| % undoc in the county/industry/quarter | 0.0113* | 0.0021* | 0.0027* |
| | (0.0002) | (0.0003) | (0.0004) |
| Y/Y % growth in county employment | -0.0002* | 0.0029* | 0.0064* |
| | (0.0015) | (0.0012) | (0.0014) |
| *Firm Characteristics* | | | |
| Lag Churning among documented wrkrs | 0.0695* | 0.0801* | 0.0508* |
| | (0.0009) | (0.0008) | (0.0010) |
| firm 10-49 workers = 1 | 0.0535* | 0.0402* | 0.0380* |
| | (0.0010) | (0.0009) | (0.0009) |
| firm 50-249 workers = 1 | 0.1457* | 0.1049* | 0.0912* |
| | (0.0014) | (0.0012) | (0.0013) |
| firm 250+ workers = 1 | 0.2540* | 0.1824* | 0.1501* |
| | (0.0019) | (0.0015) | (0.0016) |
| firm age 5-12 quarters = 1 | -0.0005* | -0.0052* | -0.0084* |
| | (0.0010) | (0.0007) | (0.0007) |
| firm age 13-24 quarters = 1 | 0.0072* | 0.0012* | 0.0026* |
| | (0.0015) | (0.0012) | (0.0011) |
| firm age 25+ quarters = 1 | 0.0165* | 0.0198* | 0.0261* |
| | (0.0020) | (0.0015) | (0.0015) |
| *Worker Characteristics* | | | |
| worker Ten (5-8 qrts) = 1 | 0.0146* | 0.0186* | 0.0388* |
| | (0.0005) | (0.0004) | (0.0004) |
| worker Ten (9-16 qrts) = 1 | 0.0608* | 0.0684* | 0.0897* |
| | (0.0007) | (0.0005) | (0.0005) |
| worker Ten (17-28 qrts) = 1 | 0.1098* | 0.1227* | 0.1484* |
| | (0.0009) | (0.0007) | (0.0006) |
| worker Ten (29+ qrts) = 1 | 0.1242* | 0.1472* | 0.1755* |
| | (0.0014) | (0.0010) | (0.0010) |
| Hired or separated = 1 | -0.7119* | -0.7231* | -0.6560* |
| | (0.0005) | (0.0003) | (0.0004) |
| N | 18,716,494 | 30,298,825 | 24,528,770 |
| Number of unique individuals | 2,437,501 | 3,376,102 | 2,237,834 |
| Number of unique firms | 83,100 | 93,021 | 85,946 |

| Dependent variable = log average quarterly earnings of documented workers in the firm | Low | Medium | High |
|---|---|---|---|
| R2 | 0.8353 | 0.8597 | 0.8536 |

Notes: See notes to Table 2.

# 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.[12] 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.[13] 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.

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

[13] 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."[14] 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.[15] 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

[14] "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>.
[15] 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).[16] 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.[17] The high degree of churning observed among workers with invalid SSNs for these other reasons is

[16] 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).

[17] 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.

*A.3. 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 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 A2 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 A2 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.[18] 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 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 A1 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

[18] 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).

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.

Figure A1.

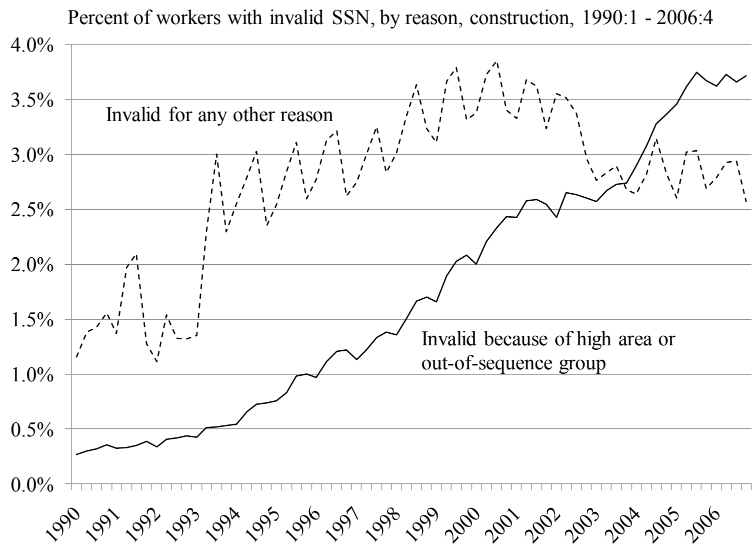
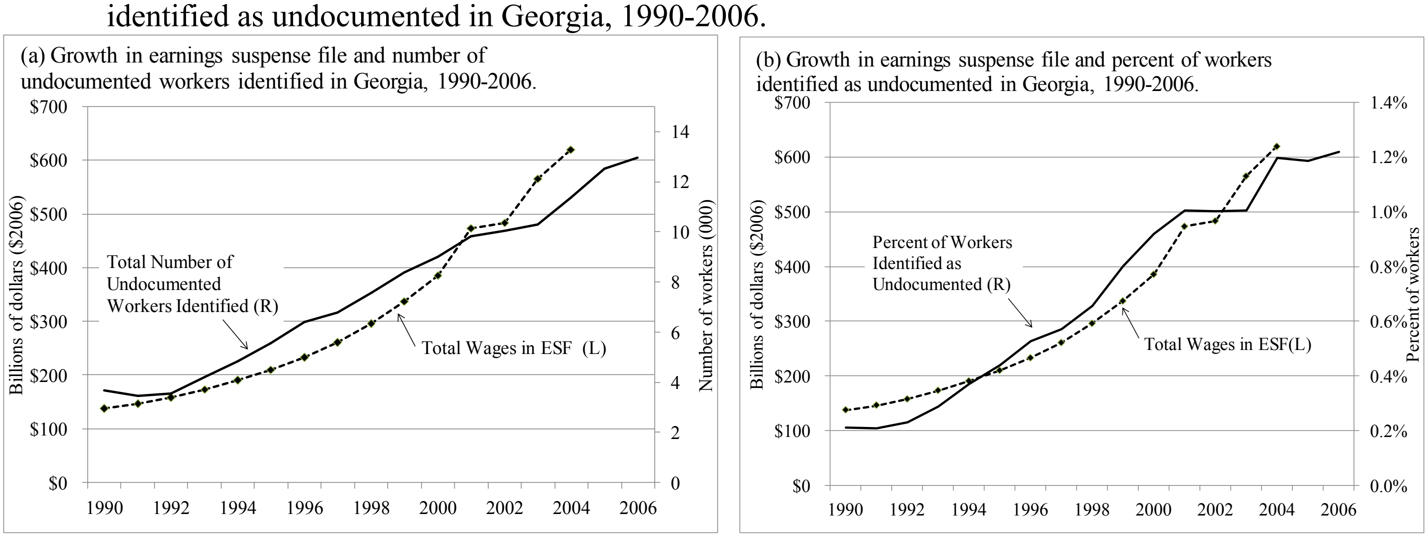



Figure A2. Growth in the earnings suspense file and the total number and percent of workers identified as undocumented in Georgia, 1990-2006.




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.

Table A1. 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.

## Appendix B: Determining a Firm's Skill Classification

The classification of firms by skill level has multiple steps. First, the March Current Population Survey for each year between 1995 and 2005 is used to calculate the average education level of workers by industry. The level of aggregation chosen depended on having enough observations (at least 1000) and on make industries as comparable as possible across the industry re-classification that took place in 2003. The share of workers with less than a high school degree, a high school degree, some college, college graduate, and graduate degree were calculated for each industry.

The second step in the process was to smooth each education share series by industry in order to better identify any trend. If the 2003 CPS reclassification of industries appears to break the series significantly, the trend between 1995 and 2002 was extended through the end of the series. Most series did not exhibit any significant trends over time, with some exceptions. For example, Figure B.1 shows the trend in share of workers in four industries exhibiting an increasing share of workers with less than a high school degree; these are also industries in which undocumented worker employment is relatively large.

(Figure B.1 about here)

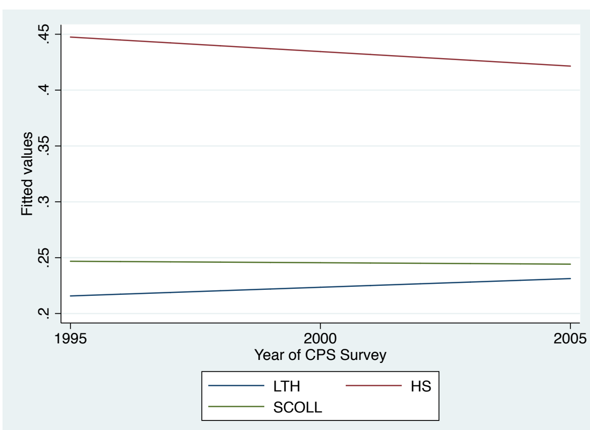
The third step involved grouping industries into three skill classifications based on information about how workers in the industry are distributed across education levels. We use k-means clustering analysis (with k=3) to perform this grouping (MacQueen 1967). Figure B.2 plots the distribution of industries across the share of workers in each education level based on the skill classification of the industry. As can be seen from the Figure, there is clear separation at the extreme education levels. There is significant overlap between the distributions of the share of high school graduates in low and middle skill industries, and those with some college in middle and high skill industries.

(Figure B.2 about here)

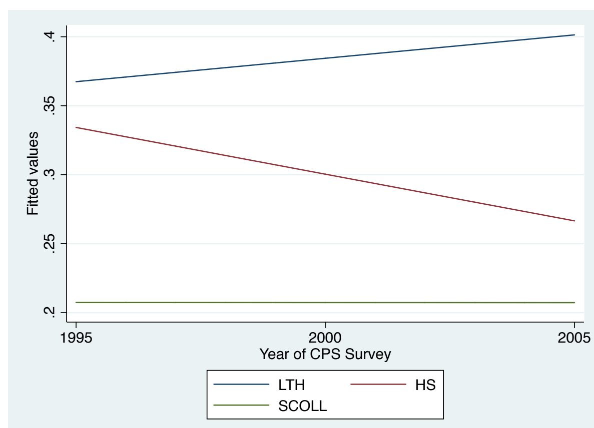
Each firm is assigned a skill level based on the industry in which the firm is located. This skill level can change over time as the educational attainment of workers in the industry changes over time.

Figure B.1. The share of workers in two industries with less than high school, high school, and some college.

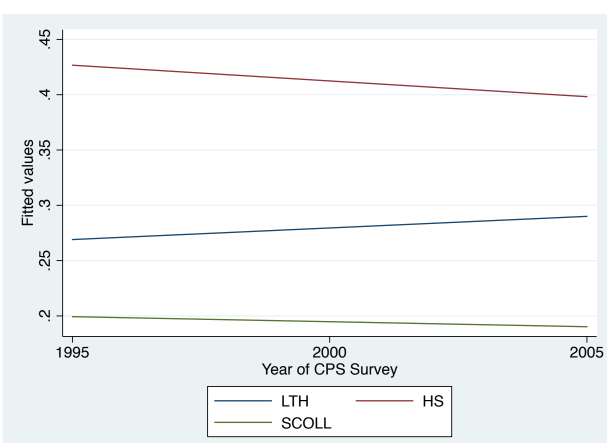
(a) Construction




(b) Support Activities for Ag and Forestry




(c) Food Manufacturing




(d) Apparel Manufacturing

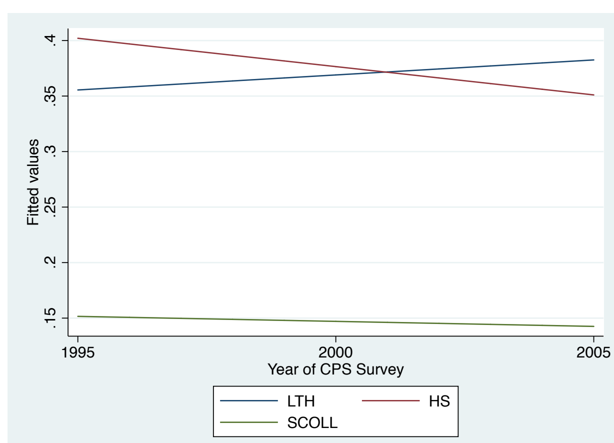
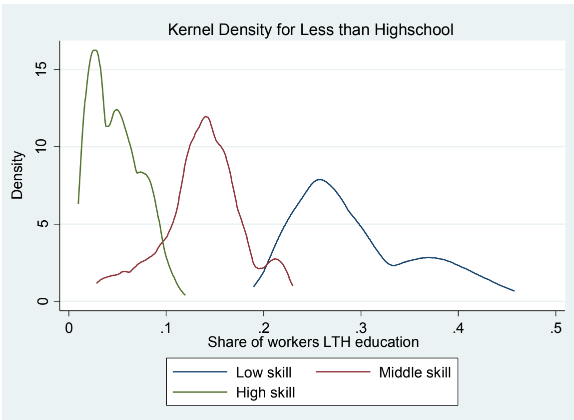
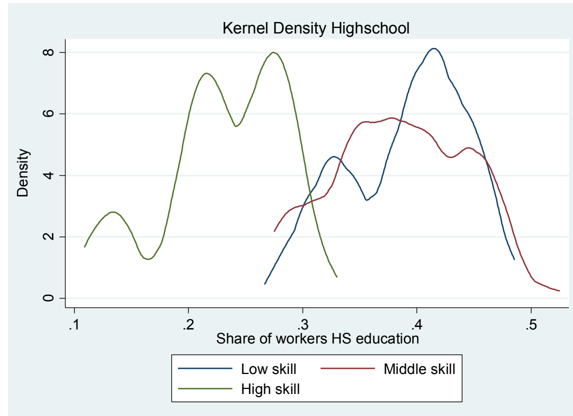


Figure B.2. Distribution of industries classified by skill level (as a result of clustering) by education level.

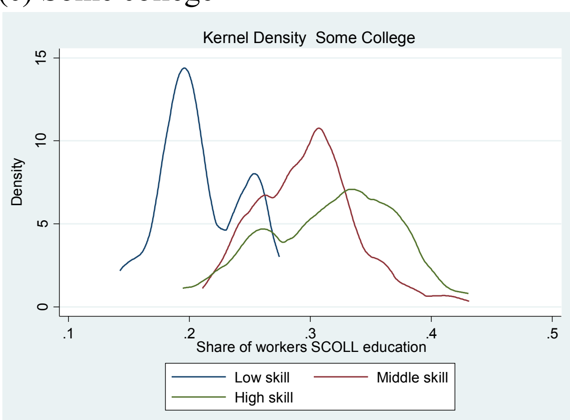
(a) Less than high school



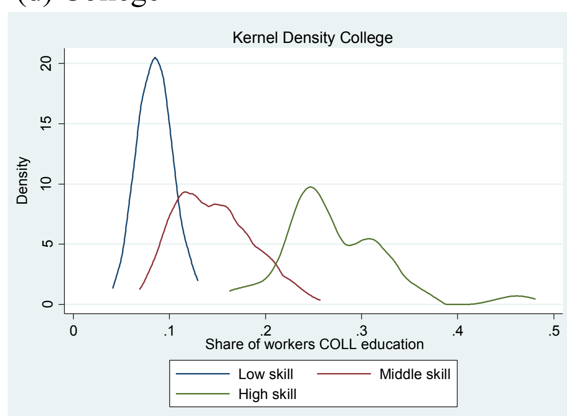

(b) High school




(c) Some college




(d) College




(e) Graduate