Southern Economic Journal 2015, 81(4), 874–906
DOI: 10.1002/soej.12020

Symposium: Economic Impact of Unauthorized Workers

The Wage Impact of Undocumented Workers: Evidence from Administrative Data

Julie L. Hotchkiss,*† Myriam Quispe-Agnoli,‡ and Fernando Rios-Avila§

Using administrative, individual level, longitudinal data from the state of Georgia, this article 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%. Within the firm, a one percentage point increase in the percent of undocumented workers employed by the firm boosts wages by 0.09% (0.11, 0.12, and 0.04 in low, medium, and high skill firms, respectively). Potential explanations for a positive wage impact are discussed.

JEL Classification: J30, J15

1. 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 13–42% higher level of capital stock than would have prevailed in the absence of immigration during these years (also see Chiswick, Chiswick, and Karras 1992; Barro and Sala-i-Martin 1995). Immigration has also been more recently explored in various countries as a mechanism for replacing retiring baby-boom workers (e.g., Denton and Spencer 1997; Hotchkiss 2005; Hamada and Kato 2007).

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 article 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% increase in the population share of immigrants results in a 1–4% decrease in native wages (e.g., see Friedberg and

* Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, NE, Atlanta, GA, 30309-4470, USA; E-mail: Julie.L.Hotchkiss@atl.frb.org
† Department of Economics, Georgia State University
‡ Department of Economics, University of Georgia, Athens, GA 30602; E-mail: mquispe@uga.edu
§ Levy Economics Institute, Annandale-on-Hudson, NY, 12504-5000, USA; E-mail: friosavi@levy.org

Hunt 1995; Orrenius and Zavodny 2007; Borjas, Grogger, and Hanson 2010). The measured impact of immigration on the displacement of workers is less clear. Card (1990, 2001), Butcher and Card (1991), Wright, Ellis, and Reibel (1997), and Card and DiNardo (2000) find no evidence of immigrant inflows affecting native migration patterns or employment outcomes. Whereas Frey (1996) and Borjas (2006) identify a significant relationship between immigrant inflows and either native outflows or lower net native in-migration, Card (2001) finds lower rates of employment within cities with high immigrant arrivals.

More recent evidence from Peri (2012) 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. In addition, over time, a net increase of immigrants equal to 1% of employment significantly increases income per worker by 0.5%. This positive impact on worker income derives from increased efficiency and productivity through task specialization, especially among low-skilled natives (also see Cobb-Clark, Shiells, and Lowell 1995; Toussaint-Comeau 2007; Peri and Sparber 2009; Iskander and Lowe 2011). 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 Chiswick, Chiswick, and Karras (1992) and Barro and Sala-i-Martin (1995), 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 (e.g., specializing in tasks requiring greater communication skills).[1]

The analysis in this article 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

[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.

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.

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 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; Nuñez-Neto and Viña 2005 for more details).

Since the terrorist attacks of 2001, and in response to continued dramatic growth in unauthorized immigrants, there have been renewed calls in the U.S. 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 156 in 2012.[3] The first major immigration legislation in Georgia became law in July 2007 and the second in 2011. The analysis in this article makes use of data through 2006, so the relatively recent change in the legal environment in Georgia will not confound the current analysis.

Identifying Unauthorized Immigrants

Identifying unauthorized immigrants is the greatest challenge in investigating their impact. A common strategy has been to use a combination of U.S. Census Bureau surveys and administrative data from the Department of Homeland Security (e.g., Passel 2013). Other sources of information include apprehension data from the U.S. Border Patrol (e.g., Hanson and Spilimbergo 1999; GAO 2006), the Mexican Migration Project (Orrenius, Zavodny, and Lukens 2008), the Legalized Persons Survey, and the new Immigrant Survey (see Jasso et al. 2000; Jasso 2011). This article differs in the way in which unauthorized individuals are identified. The purpose of our strategy is not to obtain a count of unauthorized immigrants but to identify a reasonable sample that captures the trends of undocumented workers in the labor market in Georgia with which to perform statistical analyses of labor market outcomes.

State administrative data are used to identify invalid social security numbers (SSN) 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.[4] Unlike most other studies, the measure used here does not capture the supply of undocumented workers, but, rather, the demand, as the workers are

[2] For historical details, see CBO (2006) and FAIR (2007).

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

[4] 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.

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% of all undocumented workers in the state of Georgia,based on totals estimated by Fortuny, Capps, and Passel (2007).

2. Data

The primary data used for the analyses in this article 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 United States, covering approximately 99.7% of all wage and salary workers (Committee on Ways and Means 2004).[5] 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 various levels of aggregation. 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 six-digit NAICS industry code and the county of location, allowing us to construct or merge in various industry- and county-level indicators. It is estimated that in 2010 about 4% of the total number of undocumented workers in the United States were located in Georgia, and that between 2000 and 2010 Georgia experienced a 70% increase in the number of unauthorized immigrants, one of the largest percentage increases in the United States (Passel and Cohn 2011).

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, and so forth; 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 do not arise when we limit the analysis to single establishment firms. The second reason we restrict the analysis is because it is a clear way to reduce the number of observations without employing some sampling scheme. The full data sample has over 178 million observations (repeated worker observations between 1995 and 2005), restricting to single establishment firms reduces the sample by about half. Conclusions are only generalizable to single establishment firms. About 98.5% of all firms in Georgia during this time period are 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

---

[5] Certain jobs in agriculture, domestic services, and nonprofit organizations are excluded from UI coverage; excluded workers are not represented in the data.

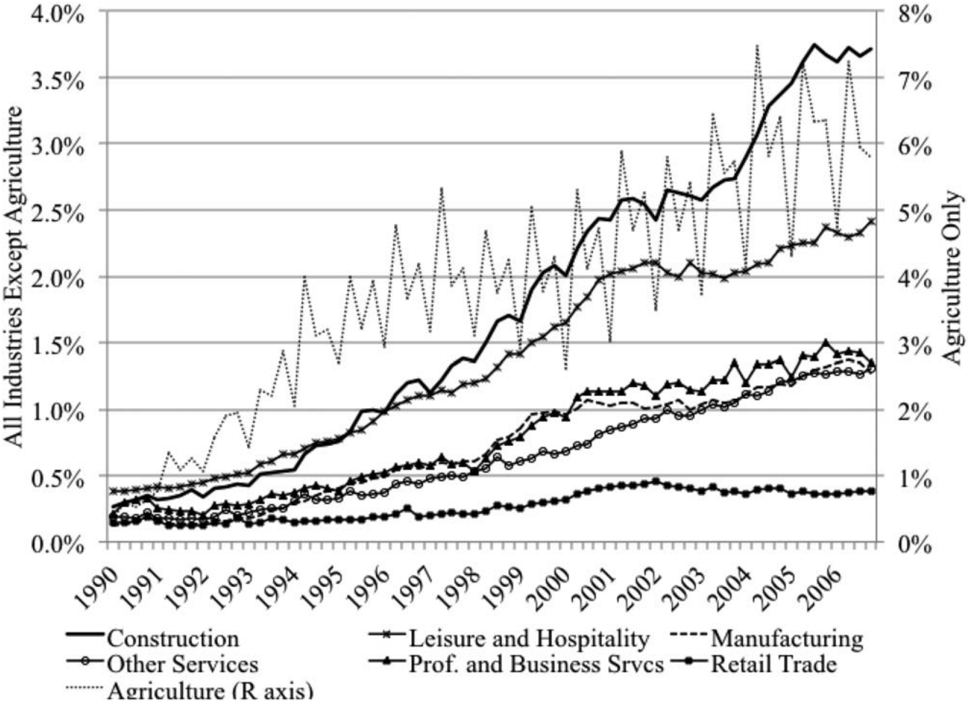

Figure 1. Percent of Workers That is Undocumented by Broad Industry, 1990:1–2006: 4

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).

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.[6] 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 SSNs 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 representation of undocumented workers in each industry matches that reported by Fortuny, Capps, and Passel (2007), who estimate that nationally in 2004 the percent of workers in leisure and hospitality and construction that was undocumented was

---

[6] The assignment of SSNs was changed to a randomized system on June 25, 2011.

[7] The only other use of SSNs 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.

10% each, 9% of workers in agriculture, and 6% each in manufacturing, professional and business services, and other services (also see Pena 2010). The pattern of growth is also consistent with Fortuny, Capps, and Passel (2007), who estimate that 72% of unauthorized immigrants in Georgia arrived since, roughly, 2000.

Fortuny, Capps, and Passel (2007) estimate that 4.5% of the workforce in Georgia was undocumented in 2004. In our sample 1.0% of workers are classified as undocumented in 2004, implying that the sample used for the analysis in this article is capturing about 22% 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 conservative in the identification of workers as undocumented. Note that the identification process we use in this article 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. Evidence supporting that confidence is detailed in Appendix A.

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 article. 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, Stamps, and Atiles 2008) and are likely to accept a wage lower than his/her productivity (e.g., Hotchkiss and Quispe-Agnoli 2013), 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% of Mexicans in the United States are undocumented (Hoefer, Rytina, and Baker 2012), 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 nonborder state. For example, CBO (2010) reports that 91% 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.

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 present in the analysis. However, worker and firm longitudinal characteristics, like tenure and age, are calculated beginning in 1990 (the first year of available data). There are over 178 million observations in this restricted time period, making estimation with high order fixed effects cumbersome, at best (see Abowd, Dramarz, and Margolis 1999). As mentioned earlier, the sample will be additionally restricted with the elimination of all multiestablishment firms. Since this article 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 2011).

The final estimating sample contains 73.5 million observations (5.7 million unique individuals). Workers overall earn roughly \$9400 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 (based on two- and three-digit NAICS; see Appendix B for details). 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. Some sectors, such as Other Services, have firms represented in each skill classification, while others, 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.

3. 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 articles by George Borjas (alone and with co-workers 2003, 2005, 2010),

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 quarterly earnings)	$9,380.91	$6,663.01	$8,597.60	$12,422.37
	(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 (all firms)	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 qtrs	0.193	0.228	0.190	0.169
	(0.394)	(0.420)	(0.392)	(0.375)
Firms age 5–12 qtrs	0.438	0.435	0.445	0.431
	(0.496)	(0.496)	(0.497)	(0.495)
Firms age 13–24 qtrs	0.185	0.176	0.186	0.190
	(0.388)	(0.381)	(0.389)	(0.392)
Firms age 25+qtrs	0.185	0.161	0.179	0.210
	(0.39)	(0.37)	(0.38)	(0.41)
Firm Size				
Firms 1–9 wrkrs	0.142	0.162	0.124	0.148
	(0.349)	(0.369)	(0.329)	(0.356)
Firms 10–49 wrkrs	0.287	0.353	0.287	0.237
	(0.452)	(0.478)	(0.452)	(0.425)
Firms 50–249 wrkrs	0.291	0.314	0.324	0.232
	(0.454)	(0.464)	(0.468)	(0.422)
Firms 250+ wrkrs	0.280	0.171	0.265	0.382
	(0.449)	(0.376)	(0.441)	(0.486)
Distribution of Firms	Across sectors (%)	Across skill classification, within sector (%)		
Agriculture	1.23	87.43	12.57	0.00
Construction	12.33	100.00	0.00	0.00
Manufacturing	4.58	32.84	56.53	10.64
Transport	3.09	0.00	76.55	23.45
Wholesale	8.63	0.00	100.00	0.00
Retail	12.45	18.40	77.68	3.92
Financial srvcs	8.96	0.00	9.02	90.98
Information	1.60	0.00	27.24	72.76
Professional srvcs	19.47	0.00	49.58	50.42
Education & hlth	8.89	0.00	18.20	81.80
Leisure & hosp.	7.70	79.14	20.86	0.00
Other services	11.06	45.48	31.64	22.88
Total	100.00	25.45	41.20	33.35
Churning among documented workers	0.23	0.30	0.26	0.14
	(0.25)	(0.26)	(0.27)	(0.16)
Percent of workers hired or separated in the current quarter	22.914	28.381	25.033	16.125
	(42.028)	(45.085)	(43.320)	(36.776)
Worker Tenure within the Firm (Quarters)				
Workers Tenure (1–4 qrts) = 1	0.363	0.424	0.388	0.287
	(0.481)	(0.494)	(0.487)	(0.452)
Workers Tenure (5–8 qrts) = 1	0.170	0.169	0.166	0.175
	(0.375)	(0.375)	(0.372)	(0.380)

Table 1. (*Continued*)

	Full	Low Skill	Middle Skill	High Skill
Workers Tenure (9–16 qrts) = 1	0.175	0.160	0.167	0.195
	(0.380)	(0.367)	(0.373)	(0.396)
Workers Tenure (17–28 qrts) = 1	0.142	0.123	0.135	0.164
	(0.349)	(0.328)	(0.342)	(0.370)
Workers Tenure (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

Note: 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).

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), 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 article 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, across time, and at the firm level on an individual documented worker's wage.

It's important to point out the assumptions under which these fixed effects results will be interpreted as causal. The ability to control for firm and individual fixed effects is what differentiates these results from previously estimated impacts of immigration on wages. The success of these fixed effects in sweeping away coefficient-biasing endogenous characteristics of the firm and worker depends on these effects remaining constant over time. If the arrival of additional unauthorized immigrants changes the behavior of firms (e.g., a firm's inclination to break the law) or changes the job search behavior of workers (e.g., a worker's distaste for working with undocumented workers), then the ability of these results to be interpreted as causal is compromised.

Wage equation models at three different levels of aggregation will be estimated 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.

Estimation at the County/Industry Level

The first estimation follows the specification similar to the one proposed by Orrenius and Zavodny (2007), using aggregate information at the county (159 counties), industry (12 levels), and year and quarter information:

$$\ln \bar{w}_{ckt} = \alpha + \beta P_{ckt} + \gamma' \bar{X}_{ckt} + \varsigma_c + \iota_k + \tau_t + \varepsilon_{ckt}, \tag{1}$$

where $\ln \bar{w}_{ckt}$ and P_{ckt} are the log of the average quarterly earnings among documented workers and percentage undocumented workers, respectively, in county c, industry k, and quarter t. The vector $\bar{X}_{ckt}$ controls for average worker and firm characteristics for each county/industry cell in quarter t. We also control for county ς_c, industry l_k, and quarter τ_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 documented workers represented by that cell. This is done for two reasons; first, to maintain consistency of the specification with less aggregated models, and second, because our dependent variable, together with other variables of interest, is the average among documented workers.

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.

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:

$$\ln \bar{w}_{jt} = \alpha + \beta_1 P_{jt} + \beta_2 P_{ckt} + \gamma' X_{jt} + \varsigma_c + \iota_k + \tau_t + \varphi_j + \varepsilon_{jt}, \tag{2}$$

where $\ln \bar{w}_{j,}$ is the log of the average quarterly earnings of documented workers in the firm j at quarter t, P_{jt} is the percentage of undocumented workers in the firm, and X_{jt} is a vector of firm and (average) worker characteristics. To allow for an effect of the local potential supply of undocumented workers on average firm wages, P_{ckt} 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, φ_j. Similar to the county/industry level analysis, each observation is weighted by the number of documented workers in each firm at time t, as average quarterly earnings and other variables of interest are constructed with respect to documented workers. This weighting scheme also aims to maintain consistency with the individual-level specification.

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 (e.g., see Pena 2010; 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.

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:

$$\ln w_{ijt} = \alpha + \beta_1 P_{jt} + \beta_2 P_{ckt} + \gamma_1' X_{ijt} + \varsigma_c + \iota_k + \tau_t + \varphi_j + \theta_i + \varepsilon_{ijt}, \tag{3}$$

where $\ln w_{ijt}$ is the log of the quarterly earnings 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_{jt}, and the local job market, P_{ckt}. It also includes worker and firm specific characteristics, but now worker characteristics are measured individually. Equation 3 also includes an individual fixed effect θ_i, which will control for unmeasured, 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.

The difficulty of controlling for more than one high-order fixed effect, however, is notorious (e.g., see Abowd, Dramarz, and Margolis 1999). To control for both firm level and individual worker level fixed effects (as well as county, industry, and quarter fixed effects), we apply the methodology developed in Rios-Avila (2013), which is described in Appendix C. Unlike Abowd, Dramarz, and Margolis (1999), we are not interested in identifying the fixed effects themselves, merely removing their biasing influence on the other coefficients. The procedure is analogous to a differencing approach used to remove a single fixed effect.

Regressors

Worker tenure is expected to positively influence wages through the presence of firm-specific human capital (Altonji and Shakatko 1987; Campbell 1993). 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, Lane, and Stevens 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 article 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 2013), but a full analysis of long-term labor market outcomes of potentially displaced documented workers will be the subject of a future investigation.

4. Estimation Results

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.[8] 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.

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%. Borjas (2003) finds a 0.3% reduction in wages and Orrenius and Zavodny (2007) report a 0.8% reduction for a similar 1% increase in immigrants. Altonji and Card (1991), using a

[8] 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.

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 mid 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 Tenure (5–8 qtrs)	0.8560*	0.8510*
	(0.0494)	(0.0492)
% Workers Tenure (9–16 qtrs)	0.6687*	0.6630*
	(0.0463)	(0.0462)
% Workers Tenure (17–28 qtrs)	0.5795*	0.5743*
	(0.0473)	(0.0473)
% Workers Tenure (29+ qtrs)	0.5610*	0.5576*
	(0.0413)	(0.0412)
% of workers hired or separated	−0.0098	−0.0087
	(0.0168)	(0.0168)
Intercept	7.9627*	7.9641*
	(0.0375)	(0.0374)
N = 74,157 (159 counties, 12 industries, 40 quarters)		
R-Squared	0.916	0.916

Note: 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).

different estimation approach, estimate between a 0.3% and 1.2% negative impact on native wages with a one percentage point increase in the 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.

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 to get a sense how much controlling for firm specific, time invariant characteristics matters.

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 1%, whereas a one percentage point increase in the share of undocumented workers in the county/industry lowers (marginally) documented worker wages by 0.2%. 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% 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% 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

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)
% Undocumented 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 mid 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 workers	−0.6606*	−0.6562*	−0.0048*	−0.0048*
	(0.0060)	(0.0059)	(0.0016)	(0.0016)
% Firms 10–49 wrkrs	0.2868*	0.2862*	−0.0183*	−0.0184*
	(0.0011)	(0.0011)	(0.0008)	(0.0008)
% Firms 50–249 wrkrs	0.3854*	0.3843*	−0.0514*	−0.0515*
	(0.0017)	(0.0017)	(0.0014)	(0.0014)
% Firms 250+ wrkrs	0.3344*	0.3341*	−0.0823*	−0.0824*
	(0.0030)	(0.0030)	(0.0030)	(0.0030)
% Firms age 5–12 qtrs	−0.8508*	−0.8487*	−0.4152*	−0.4152*
	(0.0079)	(0.0079)	(0.0037)	(0.0037)
% Firms age 13–24 qtrs	−0.0133*	−0.0130*	−0.0007	−0.0008
	(0.0031)	(0.0031)	(0.0015)	(0.0015)
% Firms age 25+ qtrs	−0.0198*	−0.0193*	0.0046^	0.0046^
	(0.0035)	(0.0035)	(0.0024)	(0.0024)
Averaged Worker Characteristics				
% Workers Tenure (5–8 qtrs)	0.4856*	0.4831*	0.1325*	0.1325*
	(0.0068)	(0.0068)	(0.0027)	(0.0027)
% Workers Tenure (9–16 qtrs)	0.4593*	0.4575*	0.1821*	0.1819*
	(0.0065)	(0.0065)	(0.0027)	(0.0027)
% Workers Tenure (17–28 qtrs)	0.4726*	0.4714*	0.2360*	0.2357*
	(0.0077)	(0.0077)	(0.0031)	(0.0031)
% Workers Tenure (29+ qtrs)	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)
Intercept	7.9298*	7.9277*	8.6921*	8.6926*
	(0.0119)	(0.0119)	(0.0263)	(0.0263)
N = 4,652,603 (256,227 unique firms)				
R-Squared	0.526	0.527	0.919	0.919

Note: See notes to Table 2.

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 fact that the coefficient on the share of undocumented workers switches from negative to positive suggests that firm and county/sector labor markets that employ undocumented workers also pay lower wages. However, since these estimates are measuring within-firm effects, 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; however 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.[9]

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. Table 4 reports estimation results with varying inclusions of fixed effects.

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 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%). 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, it is being employed by a firm that is more likely to hire undocumented workers that leads to lower wages.

---

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

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		Individual and Firm FE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Undocumented 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*		1.28e-05*
		(5.79E-07)		(6.22E-07)		(8.84E-07)		(8.17E-07)
x firm classified as mid skill = 1		−2.20E-04*		−9.38E-05*		2.13E-05*		7.00e-06*
		(7.02E-07)		(7.32E-07)		(1.07E-06)		(9.78E-07)
x firm classified as high skill = 1		−1.04E-04*		−1.57E-05*		1.17E-06		2.28E-06
		(1.86E-06)		(1.61E-06)		(2.16E-06)		(1.78E-06)
% Undocumented in the county/industry (c/i)	−0.0076*		−0.0008*		0.0050*		0.0047*	
	(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*		6.2E-05*
		(1.65E-06)		(1.74E-06)		(2.65E-06)		(2.50E-06)
x % of firms in c/i that are mid skill		−1.82E-04*		−8.12E-05*		8.88E-06^		−1.9E-05*
		(3.62E-06)		(3.43E-06)		(4.88E-06)		(4.10E-06)
x % of firms in c/i that are high skill		2.70E-04*		2.40E-04*		1.38E-04*		1.1E-04*
		(7.06E-06)		(6.23E-06)		(8.49E-06)		(6.70E-06)
Y/Y % growth in county employment	−0.0064*	−0.0057*	0.0037*	0.0039*	0.0030*	0.0030*	0.0034*	0.0035*
	(0.0013)	(0.0013)	(0.0009)	(0.0009)	(0.0011)	(0.0011)	(0.0009)	(0.0009)
Firm Characteristics								
Lag Churning among documented workers	−0.9157*	−0.9085*	−0.3340*	−0.3325*	0.0265*	0.0265*	0.0525*	0.0524*
	(0.0006)	(0.0006)	(0.0005)	(0.0005)	(0.0007)	(0.0007)	(0.0006)	(0.0006)
Firm 10–49 workers = 1	0.2109*	0.2102*	0.0694*	0.0693*	−0.0398*	−0.0398*	0.0234*	0.0234*
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0007)	(0.0007)	(0.0006)	(0.0006)
Firm 50–249 workers = 1	0.3132*	0.3119*	0.1147*	0.1147*	−0.0553*	−0.0553*	0.0696*	0.0695*
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0009)	(0.0009)	(0.0008)	(0.0008)
Firm 250+ workers = 1	0.2462*	0.2457*	0.0953*	0.0961*	−0.0498*	−0.0499*	0.1326*	0.1325*
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0012)	(0.0012)	(0.0011)	(0.0011)
Firm age 5–12 quarters = 1	−0.0908*	−0.0905*	−0.0125*	−0.0126*	−0.0132*	−0.0132*	0.0019*	0.0019*
	(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*	0.0025*	0.0026*
	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0010)	(0.0010)	(0.0008)	(0.0008)

Table 4. (*Continued*)

Dependent Variable = Log Average Quarterly Earnings of Documented Workers in the Firm	No FE		Individual FE only		Firm FE only		Individual and Firm FE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Firm age 25+ quarters = 1	−0.1394*	−0.1388*	−0.0489*	−0.0489*	0.0442*	0.0442*	0.0081*	0.0081*
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0013)	(0.0013)	(0.0011)	(0.0011)
Worker Characteristics								
Worker Tenure (5–8 qtrs) = 1	0.2299*	0.2295*	0.0686*	0.0685*	0.1827*	0.1827*	0.0563*	0.0563*
	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Worker Tenure (9–16 qtrs) = 1	0.3582*	0.3579*	0.1045*	0.1044*	0.3278*	0.3278*	0.1023*	0.1023*
	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Worker Tenure (17–28 qtrs) = 1	0.5103*	0.5100*	0.1454*	0.1453*	0.4988*	0.4988*	0.1459*	0.1459*
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Worker Tenure (29+ qtrs) = 1	0.6673*	0.6672*	0.1209*	0.1207*	0.6920*	0.6920*	0.1323*	0.1323*
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0005)	(0.0005)
Hired or separated = 1	−1.0536*	−1.0531*	−0.8198*	−0.8197*	−0.9873*	−0.9873*	−0.7954*	−0.7954*
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Intercept	7.9821*	7.9826*						
	(0.0044)	(0.0044)						
N = 73,544,089 (5,720,423 unique workers; 256,227 unique firms)								
R-Squared	0.3895	0.3900	0.7454	0.7454	0.5417	0.5417	0.8570	0.8570

Note: See notes to Table 2.

Columns 7 and 8 contain the results controlling for both firm and individual fixed effects. The estimates are interpreted as within-worker effects, controlling for the time-invariant influence of working for a firm that hires (or does not) undocumented workers. Although expectedly smaller in magnitude, the positive relationship between wages and the presence of undocumented workers at both the firm and in the local labor market indicates that the positive coefficients from the firm level analysis were not the result of composition bias. But, rather, firms that hire undocumented workers also pay lower wages, and that controlling for this firm effect is essential in identifying the positive impact of undocumented workers on the individual worker. 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% (0.47%). The largest boost to wages accrues to workers in *firms* classified as low skill and as the share of high skill workers in the *county/industry* increases. 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. It's also worth noting that this final specification indicates an increasing wage premium with firm size, as we would expect from a long literature relating firm size and wages (e.g., see Oi and Idson 1999). In addition, it is only the inclusion of both firm and individual fixed effects that produces an increasing wage premium with firm age, and the smallest impact of separating or being newly hired. And, as we would expect, controlling for individual fixed effects has the largest mitigating effect on the relationship between worker tenure and wages, although wages are still increasing with tenure.

Individual Level Results by Employer's Skill Classification

We now return to the earlier mentioned result that workers employed by firms classified as low (or medium) skill earn the greatest boost as either their employer hires more undocumented workers or as the share of undocumented workers in the worker's 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 reports results in which both firm and individual fixed effects are controlled.

Workers in low skill firms get a 0.11% (1.08%) 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 (within the firm). The county/industry result is an order of magnitude higher for workers employed by low skill firms than experienced by workers in medium or higher skill firms.

Other results across skill classification of firms are also notable. County growth appears to only significantly positively benefit workers employed by firms classified as middle and high skill (with the largest and most significant impact on workers with high skill firms). 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.

Table 5. Estimation at the Individual Level, Controlling for Both Individual and Firm Fixed Effects, by Skill Classification of Employer

Dependent Variable = Log Average Quarterly Earnings of Documented Workers in the Firm	Low	Medium	High
% Undocumented employed in the firm	0.0011*	0.0012*	0.0005*
	(0.0001)	(0.0001)	(0.0002)
% Undocumented in the county/industry/quarter	0.0108*	0.0025*	0.0037*
	(0.0003)	(0.0004)	(0.0004)
Y/Y % growth in county employment	0.00001	0.0033+	0.0062*
	(0.0017)	(0.0013)	(0.0016)
Firm Characteristics			
Lag Churning among documented workers	0.0582*	0.0567*	0.0480*
	(0.0011)	(0.0010)	(0.0012)
Firm 10–49 workers = 1	0.0370*	0.0238*	0.0282*
	(0.0011)	(0.0010)	(0.0011)
Firm 50–249 workers = 1	0.1044*	0.0679*	0.0775*
	(0.0015)	(0.0013)	(0.0015)
Firm 250+ workers = 1	0.2022*	0.1315*	0.1315*
	(0.0021)	(0.0017)	(0.0019)
Firm age 5–12 quarters = 1	0.0040*	0.0049*	−0.0056*
	(0.0010)	(0.0009)	(0.0008)
Firm age 13–24 quarters = 1	0.0043*	0.0050*	−0.0004
	(0.0016)	(0.0013)	(0.0013)
Firm age 25+ quarters = 1	−0.0012	0.0099*	0.0166*
	(0.0021)	(0.0017)	(0.0016)
Worker Characteristics			
Worker Tenure (5–8 qtrs) = 1	0.0371*	0.0422*	0.0481*
	(0.0005)	(0.0004)	(0.0004)
Worker Tenure (9–16 qtrs) = 1	0.0745*	0.0854*	0.0951*
	(0.0006)	(0.0005)	(0.0005)
Worker Tenure (17–28 qtrs) = 1	0.1053*	0.1255*	0.1446*
	(0.0008)	(0.0006)	(0.0007)
Worker Tenure (29+ qtrs) = 1	0.0829*	0.1163*	0.1435*
	(0.0012)	(0.0009)	(0.0010)
Hired or separated = 1	−0.7762*	−0.7953*	−0.6981*
	(0.0006)	(0.0005)	(0.0006)
N	18,716,494	30,298,825	24,528,757
Number of unique individuals	2,437,501	3,376,102	2,237,834
Number of unique firms	83,100	93,021	85,946
R-Squared	0.7760	0.7925	0.8111

Note: See notes to Table 2.

Explaining Positive Wage Effects

From a basic aggregate demand and supply labor market model, one would expect that the arrival of any additional labor would simply put downward pressure on the wage of existing workers who are substitutes for the new arrivals. So, 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? Simple theories, such as scale effects or rent sharing, would lead us to expect a differential ordering of, or uniformly higher wages across workers employed by firms of different skill levels. The scale effect would suggest that access to a cheaper factor input (undocumented workers) would lower costs, raising production, increasing demand (and wages) for all workers. However this sort of effect would increase demand, thus wages, among complementary (presumably higher skilled) workers more than demand among lower skilled workers. This is not what we find in Table 5, where we see the largest premiums among workers at the lowest skilled firms.

Alternatively, an explanation is suggested by research from Brown, Hotchkiss, and Quispe-Agnoli (2013), who find that firms that employ undocumented workers enjoy a competitive advantage over firms who do not. If firms share the rents accruing from this competitive advantage with their documented workers, it could result in higher wages. However, one might expect these rents to be more uniform across firm skill level classification.

The most appealing explanation for a positive wage effect from increasing shares of undocumented workers, especially among workers employed by low skill firms, is found in work by Peri and Sparber (2009) and Peri (2012), 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 English communication skills (also see Cobb-Clark, Shiells, and Lowell 1995; Baker 1999; Toussaint-Comeau 2007; Iskander and Lowe 2011). 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. We might expect the opportunity for specialization to be even greater in the presence of undocumented workers, relative to the presence of immigrants as a whole.[10]

Peri and Sparber (2009) present a theoretical general equilibrium static model in which an increase in immigrant share results in a series of events: (i) the ratio of the supply of communication-to-manual skills declines; (ii) this results in a decline in returns to manual skill and an increase in returns to communication skill; and (iii) tasks are reallocated to workers' comparative advantage, with the relative increase in communication skill compensation benefiting natives. This task specialization is possible because even low-skilled immigrants and natives are imperfect substitutes. The theory predicts that there will be an overall rise in the value of now relatively scarcer communication skills (hence the positive impact measured at the county/industry level), plus an additional productivity boost in those firms able to take advantage of specialization—those employing undocumented workers.

One would expect less of an impact of the opportunity to specialize among higher skilled workers, as they are already specializing in higher productivity tasks. This is likely why the positive wage effect is significantly smaller when estimated among workers employed by firms classified as high skill. Recall that 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), implying that 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.

How Do Positive Wage Effects Persist?

Peri and Sparber's (2009) theoretical model is developed in the context of a perfectly competitive product and perfectly competitive labor market. The key to their result is that natives and

---

[10] In addition, Autor and Handel (2013) illustrate that there is a great deal of flexibility in task specialization even within a single occupation.

immigrants are imperfect substitutes and that natives are able to specialize in the higher value task, which becomes relatively scarcer as more immigrants arrive. So a natural question arises: "Why aren't all the documented workers clamoring for jobs at firms that employ more undocumented workers so that they can earn higher wages by specializing in higher productivity tasks?" Such a reallocation of labor would eventually equalize wages across firms along the distribution of shares of undocumented workers. In other words, "How does a positive wage effect persist in firms that employ undocumented workers?"

The now relatively scarcer (communication) skill is rewarded broadly across the labor market; hence, the positive effect measured at the county/industry level. However, the ability of workers to specialize in that task (and, hence, be even more productive) is only realized in firms that employ undocumented workers. In addition, not all native workers are equally endowed with the now more valuable skill and those with the greatest endowment of the skill will sort to where that skill will reap the highest rewards—firms employing undocumented workers. The value of these workers is greater to the firm than those less endowed with the skill and, hence, firms will pay the more endowed workers just enough more than what they could earn at a firm not employing undocumented workers to keep them from leaving. It can be thought of as a firm-specific efficiency wage that exists only in firms employing undocumented workers.

Robustness Check—Sorting on Wages?

If undocumented workers, generally, sort to higher paying firms (or higher paying than documented workers are sorting to), then we would observe a spurious (noncausal) relationship between higher wages and higher shares of undocumented workers. To see how much worker mobility decisions based on wages (generally, not based on communication skill) is contributing to the identification of the positive wage effect of higher shares of undocumented workers, we re-estimated Equation 3 with a single worker cross-firm fixed effect (not shown here). The results are nearly identical to those reported in Table 4, column 7, indicating that the positive effect of the presence of undocumented workers on documented workers' wages is primarily taking place within the firm, rather than across firms. This does not negate the potential importance of workers with higher skill endowment sorting to firms employing undocumented workers, but, rather, eliminates concern that the positive correlation between wages and undocumented workers only occur as a result of documented and undocumented workers sorting differentially based on firm wages.

Individual Level Results in Perspective

The estimated parameters have been discussed above in terms of their implication for wages of a one percentage point increase in the share of undocumented workers in the worker's firm and county/industry. Over the sample period, the share of undocumented workers in firms that employed undocumented workers grew an average of 0.36 percentage points per year.[11] If we apply this percentage point growth to both the county/industry sector and firm exposure to undocumented workers, the parameter estimates in Table 5 suggest that workers experienced an annual

[11] Passel and Cohn (2009) estimate that the entire population of unauthorized immigrants in the United States grew an average of 0.25 percentage points per year between 1996 and 2006. Since Georgia experienced one of largest influxes of unauthorized immigrants over this period, the 0.36 percentage point growth in our sample seems to be a reasonable estimate.

average wage boost of 0.20% (0.43% in low skill firms and 0.15% in high skill firms) between 1995 and 2005. Using the average earnings reported in Table 1, this amounts to roughly $75/year ($115/year for workers employed by low-skill firms and $75/year among workers employed by high-skill firms). While potentially a noticeable amount among low-skill workers, these wage effects are arguably small, and are considerably smaller than some others' estimates of positive immigrant wage impacts.[12] The smaller effects are not unexpected because of our ability to control for individual fixed effects.

In addition, if a labor market were to become saturated with low-skill undocumented workers to the extent that all opportunities for specialization by documented workers is exhausted, we would expect the positive wage effect to be reduced. However, this dynamic would merely place a lower bound of zero on the wage effect, as newer arriving undocumented workers would simply substitute for earlier arriving undocumented workers (see Hotchkiss and Quispe-Agnoli 2013).

5. Conclusions and Implications for Policy

Using administrative individual worker level data, linked to employer characteristics, this article 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 positive wage effects, Peri and Sparber (2009) and Peri (2012) 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 reassigned to relatively higher-skilled tasks that make better use of their comparative advantage, such as communication skills. The positive wage effect persists because not all natives are endowed with the same level of the now relatively scarce (e.g., communication) skill, and only firms who hire undocumented workers are able to benefit from the task reallocation/specialization.

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 hires undocumented workers.

In the context of the on-going debate in the United States regarding immigration reform, the results in this article have a couple of implications. First, one of the goals of most proposals for immigration reform is to increase the availability of low-skill workers to U.S. employers. As the number of low-skill workers increases, the labor market may reach a point of saturation where the opportunities for natives to specialize in higher skill tasks become eroded. As this happens, the premium associated with working for a firm that employs (now legitimized) low skill immigrants would likely disappear, or even be reversed. In addition, results presented by Hotchkiss and Quispe-Agnoli (2013) suggest that employers enjoy a certain degree of monopsony power over employment of undocumented workers, keeping their wages lower than if those workers enjoyed greater labor market flexibility. If any implemented immigration reform resulted in greater labor market flexibility—for example, being able to move more freely across employers than is allowed by current H-2A or H-2B programs—then the ability of firms to exploit monopsony power over low skill immigrants is reduced and wages to low skill immigrants will rise. This would erode any

[12] For example, Baker (1999) estimates a $400 per year wage boost to African American and non-Hispanic white women for every percentage point increase in the share of female Mexican immigrants in the labor force.

boost to wages that (mostly higher skill) documented workers might be deriving from scale effects that might be explaining some of the positive wage effects in this article.

In assessing the results in this article, it is important to keep in mind that the analysis assumes that firm and individual fixed effects adequately control for coefficient-biasing endogenous characteristics. In addition, the estimated effects of increasing shares of undocumented workers is very small, and the effects may get even smaller as the opportunities for specialization become exhausted as the share of undocumented workers grows. Further, the analysis in this article 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.

Appendix A: Using SSNs to Identify Undocumented Workers

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 SSN. There are a number of ways in which one can establish that a reported SSN is invalid.[13] The Social Security Administration (SSA) 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.[14] In addition, there are several known limitations on what can be considered a valid SSN, so a simple algorithm is used to check whether each number conforms to the valid parameters.

There are three pieces to a SSN.[15] 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 nonconsecutive 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.

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 United States 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 United States. 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."[16] 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 1% in 1997 to over 60% by the end of 2006.

A series of SSNs were decommissioned by the SSA because they had been put on fake Social Security Cards used as props to sell wallets.[17] 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 appears in the data, they

---

[13] The assignment of Social Security Numbers was changed to a randomized system on June 25, 2011.

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

[15] 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>.

[16] "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>.

[17] See "Disclosure and Verification of Social Security Numbers (SSNs) for the Section 235 Program" (November 9, 1990), <http://www.hud.gov/offices/adm/hudclips/letters/mortgagee/files/90-39ml.txt> (Accessed 8 February 2011).

Table A1. Average annual growth, 1994–2006, in U.S. and GA employment, Hispanic workers, and workers identified as undocumented

Average Annual Growth Rate of:	
Total number of workers in the United States	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.

Notes: 1994 is used as starting year since is the first year the Current Population Survey has a reliable indicator of Hispanic ethnicity.

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 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, then that pair is considered valid and all other duplicates (with different surnames) are considered invalid.

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 A1 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).[18] 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.[19] The high degree of churning observed among workers with invalid SSNs for these other reasons is consistent with either of these practices.

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.

[18] 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).

[19] 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.

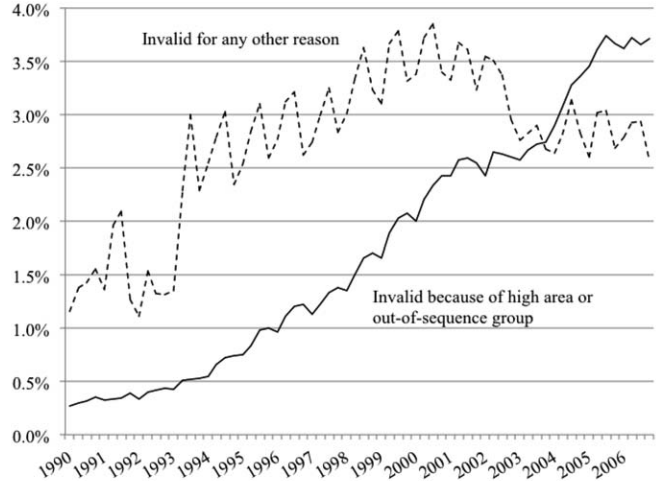

Figure A1. Percent of Workers with Invalid SSN, by Reason, Construction, 1990:1–2006:4.

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 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 United States (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.

Data from Census and Homeland Security suggest that between 40% and 60% of Mexicans in the United States are undocumented, and that 61% of unauthorized immigrants come from Mexico.[20] 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 United States and in Georgia to compare growth rates to those in our sample. These results are reported in Table A1. The workforce in GA grew faster over the period than the U.S. workforce (2.9% vs. 1.5%, respectively). In addition, the number of foreign born, Hispanic workers in the U.S. grew faster (8% 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% 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.

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 article to identify undocumented workers is accurate; it's clear that not all undocumented workers are being captured in the data, but likely represent the tip of the ice burg of hiring behavior of any firm. Any remaining misclassifications 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.

[20] The 2008 ACS estimates that 11.4 million people in the United States 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 United States in 2008 (http://www.dhs.gov/xlibrary/assets/statistics/publications/ois_ill_pe_2008.pdf).

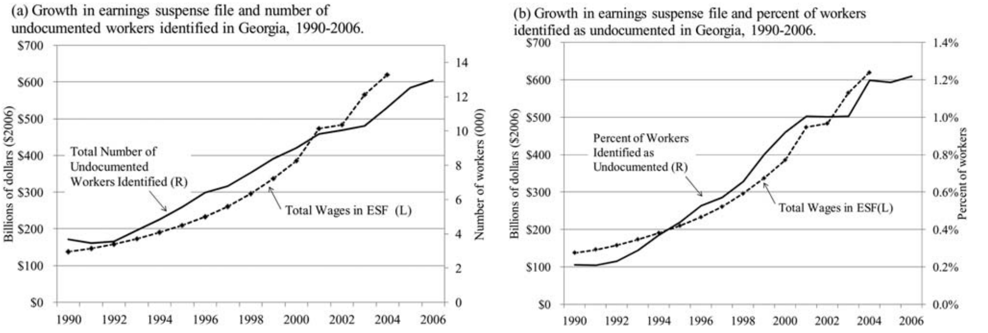

Figure A2. Growth in the Earnings Suspense File and the Total Number and Percent of Workers Identified as Undocumented in Georgia, 1990–2006. Source: Huse (2002) for Estimates 1990–2000, Johnson (2007) for Estimates 2001–2004, and Authors' Calculations. Dollar Estimates Reflect 2006 Values, Using the PCE Chain-Weighted Deflator.

Appendix B: Determining a Firm's Skill Classification

The classification of firms by skill level has multiple steps. First, the March CPS 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 to make industries as comparable as possible across the industry reclassification 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 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 B1 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.

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 B2 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.

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.

Appendix C: Strategy for Controlling for Multiple High-Order Fixed Effects

The availability of longitudinal data allows one to control for group-specific, as well as time-specific fixed effects that would other wise be absorbed by the error term, and, if correlated with the other regressors, result in biased coefficients. When the number of unique observations within group are of reasonable size, controlling for group-specific heterogeneity can easily be accomplished by adding sets of dummy variables to the regression that absorb the group-specific fixed effects. When the number of observations within a defined group is large, estimating a model that includes group-specific dummy variables can be difficult or even impossible due to limitations on either the software or hardware capacity to manage large matrixes of estimated parameters.

There are many options available within the statistical software Stata to estimate normal linear models with two fixed effects (two group categories, such as firm and worker fixed effects as seen in this article) (see McCaffrey et al. 2012 for a review of the different potential strategies). Nevertheless, the memory requirements of these strategies makes estimation of the individual level analysis in this article infeasible; this model contains roughly 250,000 unique firms and 5.7 million unique workers, in addition to 159 counties, 12 industries, and 40 time periods.

To provide feasible estimates of the parameters of interest, we implement a demeaning strategy that allows us to remove all (firm, worker, county, industry, time) of the fixed effect variation from the estimated relationship between the dependent and independent variables of interest. This strategy can be described as follows (see Rios-Avila 2013 for full details).

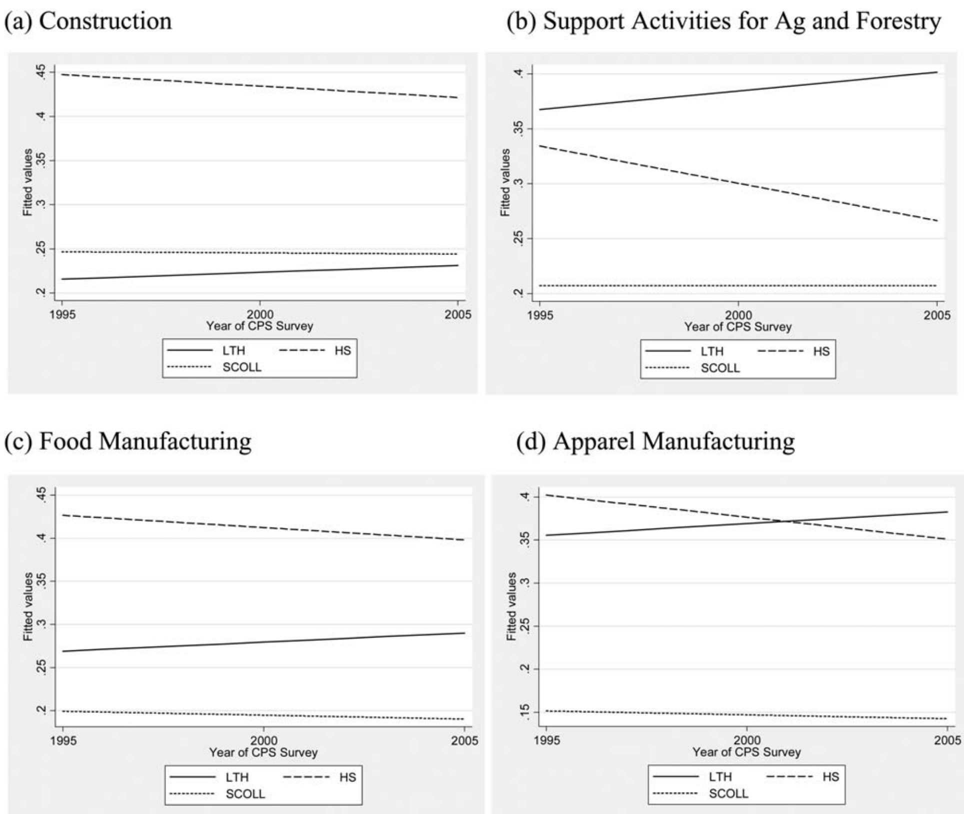

Figure B1. The Share of Workers in Four Industries with Less than High School (LTH), High School (HS), and Some College (SCOLL).

Assume the outcome y (log wages) can be expressed as a linear function of individual and firm characteristics X (variables of interest), and two fixed effects including individual (α_i) and firm (β_j) fixed effects.[21] It is also a function of a random error (ε) that is uncorrelated with the independent variables and fixed effects.

$$y = \alpha_i + \beta_j + X'\delta + \varepsilon. \tag{C1}$$

Estimating the individual means of this model, we obtain:

$$i_y = \alpha_i + i_{\beta_j} + i_X{}'\delta, \tag{C2}$$

where i_z is the average of Z (i.e., variables y or X, or parameter β) over individual i. Subtracting (C2) from (C1), we obtain:

$$y - i_y = \beta_j - i_{\beta_j} + (X - i_X)'\delta + \varepsilon, \text{ or } \quad \tilde{y} = \tilde{\beta}_j + \tilde{X}'\delta + \varepsilon. \tag{C3}$$

Note that while Equation C3 is not a function of the individual fixed effect (α_i) anymore, the transformation of the firm and county fixed effect ($\tilde{\beta}_j$) now varies with respect to each individual. Similar to before, we can obtain the mean model with respect to each firm:

[21] While the model in the article also controls for county and year quarter fixed effect, we present this model using a two fixed effects for simplicity, but the strategy can be extended to estimate any number of fixed effects (see Rios-Avila 2013 for details).

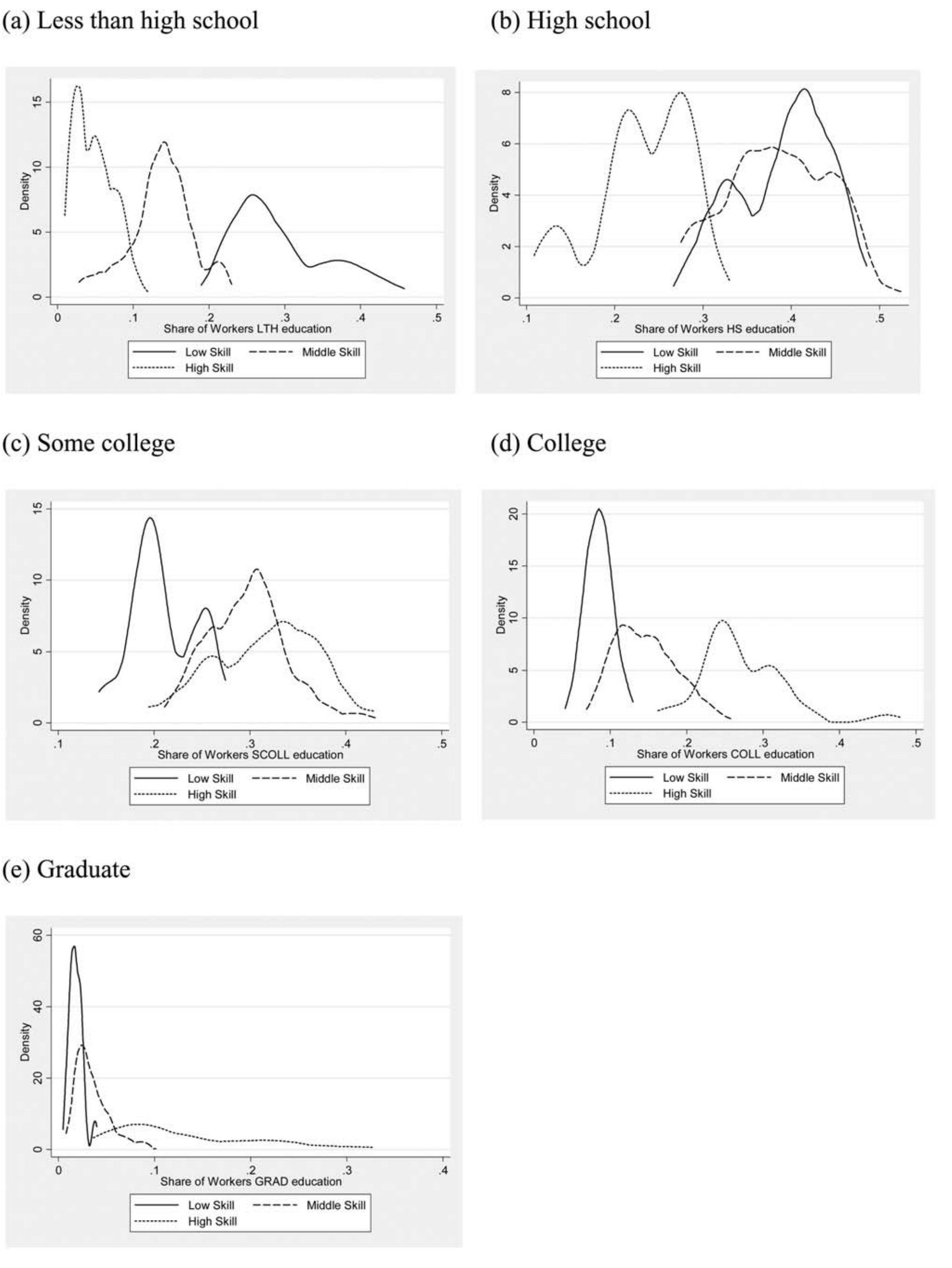

Figure B2. Distribution of Industries Classified by Skill Level (As a Result of Clustering) by Education Level.

$$j_{\tilde{y}} = j_{\tilde{\beta}_j} + j_{\tilde{X}}\delta. \tag{C4}$$

Once again, subtracting Equation C4 from Equation C3, we obtain:

$$\tilde{y} - j_{\tilde{y}} = \tilde{\beta}_j - j_{\tilde{\beta}_j} + (\tilde{X} - j_{\tilde{X}})'\delta + \varepsilon, \text{ or } \tilde{\tilde{y}} = \tilde{\tilde{\beta}}_j + \tilde{\tilde{X}}'\delta + \varepsilon. \tag{C5}$$

Note that in this case, $j_{\tilde{\beta}_j}$ is not equal to $\tilde{\beta}_j$, and the transformed coefficient $\tilde{\tilde{\beta}}_j \neq 0$. While $\tilde{\tilde{\beta}}_j$ still varies across individuals and firms, its variance is lower than the variance of β_j.[22] To eliminate the remaining variance contributed by the presence of fixed effects, the algorithm requires repeating the iterative demeaning process until the variance of $\ddot{\tilde{\tilde{\beta}}}_j$ is sufficiently small. Conversely, since the firm fixed effect is unobservable, the iterative demeaning process should be repeated until the variance of $i_{\tilde{\tilde{y}}}$ and $j_{\tilde{\tilde{y}}}$ is sufficiently small, and should not affect the estimates of the parameters. The final transformed variables ($\ddot{\tilde{\tilde{y}}}$ and $\ddot{\tilde{\tilde{X}}}$) are orthogonal to the individual and firm fixed effects, and can be used to estimate the parameters of interest. Using these transformed variables, the following equation is estimated in place of model (A3.1):

$$\ddot{\tilde{\tilde{y}}} = \ddot{\tilde{\tilde{X}}}'\delta + \varepsilon. \tag{C6}$$

Given a sufficient number of iterations to produce the transformed variables, $\ddot{\tilde{\tilde{y}}}$ and $\ddot{\tilde{\tilde{X}}}$, ordinary least squares (OLS) estimates of $\hat{\delta}$ will converge to the true δ parameter in model Equation C1.

1. Standard Errors and $R-$Squared

Under the assumption that the error term is well behaved, ε is the same in models Equation C1 and Equation C6, and the corresponding variance $Var(\varepsilon)$ can be used to estimate the variance-covariance matrix of the parameter of interest $\hat{\delta}$ While model Equation C6 can be estimated by OLS, the reported parameter standard errors need to be corrected to take into account the differences in degrees of freedom. Assuming full identification of the fixed effects, the original model Equation C1 required the estimation of $K + 1$ parameters for the X variables and the constant, $N_I - 1$ individual fixed effects, and $N_J - 1$ firm fixed effects. In contrast model Equation C6 requires the estimation of only K parameters. Taking this into account, the variance-covariance matrix of the parameters $\hat{\delta}$ should be estimated as:

$$cov(\hat{\delta}) = \frac{\sum \hat{\varepsilon}^2}{T - K - N_I - N_J + 1} (\ddot{\tilde{\tilde{X}}}' \ddot{\tilde{\tilde{X}}})^{-1} \tag{C7}$$

where T is the total number of observations. Similarly, the corresponding R-squared can be directly estimated as:

$$R^2 = 1 - \frac{\sum \hat{\varepsilon}^2}{\sum y^2} \tag{C8}$$

2. Extension to N-Fixed effects

In cases when one requires estimates beyond two fixed effect, as in this article, the strategy described above can be modified to allow the estimation of linear models with N-way fixed effects. First, instead of implementing the demeaning process back and forth between the first and second fixed effect group, the demeaning algorithm rotates between the different fixed effect groups in turn. Second, for the estimation of the variance-covariance matrix $cov(\hat{\delta})$, the number of original degrees of freedom is calculated as:

$$d.f. = T - \left(1 + K + (N_I - 1) + (N_J - 1) + (N_T - 1) + \ldots\right) \tag{C9}$$

While not all fixed effects can be estimated, as there might not be enough observations to fully identify all parameters, the expression in Equation C9 provides a lower bound approximation of the number of degrees of freedom in

[22] $Var(\beta_j) = E(\beta_j - \bar{\beta}_j)^2 = E(\beta_j - \tilde{\tilde{\beta}}_j + \tilde{\tilde{\beta}}_j - \bar{\beta}_j)^2 = E(\beta_j - \tilde{\tilde{\beta}}_j)^2 + Var(\tilde{\tilde{\beta}}_j) \Rightarrow Var(\beta_j) > Var(\tilde{\tilde{\beta}}_j)$

the original model Equation 1. This will provide a conservative estimation of the significance levels of the parameters of interest.

Acknowledgments

The opinions expressed in this article do not necessarily reflect those of the Federal Reserve Bank of Atlanta, nor the Federal Reserve System. The article has benefited from comments and suggestions from George Borjas, J. David Brown, Tim Conley, Curtis Florence, Jerry Gonzalez, Gordon Hanson, Robert E. Moore, Tom Mroz, Pia Orrenius, Giovanni Peri, M. Melinda Pitts, 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.

References

Abowd, John M., Francis Dramarz, and David N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67: 251–333.

Altonji, Joseph, and David Card. 1991. The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, edited by John M. Abowd and Richard B. Freeman. Chicago: University of Chicago Press, pp. 201–34.

Altonji, Joseph, and Robert A. Shakotko. 1987. Do wages rise with job seniority? *Review of Economic Studies* 54: 437–59.

Autor, David H., and Michael J. Handel. 2013. Putting tasks to the test: Human ges. *Journal of Labor Economics* 31: S59–96.

Aydemir, Abdurrahman, and George J. Borjas. 2011. Attenuation bias in measuring the wage impact of immigration. *Journal of Labor Economics* 29: 69–113.

Baker, Susan Gonzalez. 1999. Mexican-origin women in southwestern labor markets. In *Latinas and African American women at work: Race, gender, and economic inequality*, edited by Irene Browne. New York: Russell Sage Foundation, pp. 244–69.

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

Bohon, Stephanie A., Katherine Stamps, and Jorge H. Atiles. 2008. Transportation and migrant adjustments in Georgia. *Population Research Policy Review* 27: 273–91.

Borjas, George J. 2003. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics* 118: 1335–74.

Borjas, George J. 2006. Native internal migration and the labor market impact of immigration. *Journal of Human Resources* 41: 221–58.

Borjas, George J. 2009. The analytics of the wage effect of immigration. Unpublished paper, Harvard University.

Borjas, George J., Jeffrey Grogger, and Gordon H. Hanson. 2010. Immigration and the economic status of African-American Men. *Economica* 33: 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. Accessed December 2008. Available at http://www.gao.gov/new.items/d06458t.pdf.

Brown, Charles, and James L. Medoff. 2003. Firm age and wages. *Journal of Labor Economics* 21: 677–98.

Brown, David J., Julie L. Hotchkiss, and Myriam Quispe-Agnoli. 2013. Does hiring undocumented workers give firms a competitive advantage? *Journal of Regional Science* 53: 158–70.

Burgess, Simon, Julie Lane, and David Stevens. 2001. Churning dynamics: An analysis of hires and separations at the employer level. *Labour Economics* 8: 1–14.

Butcher, Kristin F., and David Card. 1991. Immigration and wages: Evidence from the 1980s. *American Economic Review Papers and Proceedings* 81: 292–6.

Campbell, Carl M., III. 1993. Do firms pay efficiency wages? Evidence with date at the firm level. *Journal of Labor Economics* 11: 442–70.

Card, David. 1990. The impact of the Mariel Boatlift in the Miami labor market. *Industrial and Labor Relations Review* 43: 245–57.

Card, David. 2001. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics* 19: 22–64.

Card, David, and Ethan G. Lewis. 2005. The diffusion of Mexican immigrants during the 1990s: Explanations and impacts. NBER Working Paper No. 11552.

Card, David, and John DiNardo. 2000. Do immigrant inflows lead to native outflows? *American Economic Review Papers and Proceedings* 90: 360–7.

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

CBO. 2010. *Immigration Policy in the United States: An update*, Pub. No. 4160. Washington, DC: Congressional Budget Office.

Chiswick, Carmel U., Barry R. Chiswick, and Georgios Karras. 1992. The impact of immigrants on the macroeconomy. *Carnegie-Rochester Conference Series on Public Policy* 37: 279–316.

Cobb-Clark, Deborah A., Clinton R. Shiells, and B. Lindsay Lowell. 1995. Immigration reform: The effects of employer sanctions and legalization on wages. *Journal of Labor Economics* 13: 472–98.

Committee on Ways and Means, House of Representatives. 2004. "Greenbook". WMCP 108-6, Section 4. Accessed March 2004. Available at http://waysandmeans.house.gov/media/pdf/greenbook2003/108transmittal.pdf.

Denton, Frank T., and Byron G. Spencer. 1997. Population, labour force and long-term economic growth. Unpublished paper, McMaster University, Quantitative Studies in Economics and Population Research Reports.

Dunne, Timothy, and Kenneth Troske. 2005. Technology adoption and the skill mix of U.S. manufacturing plants. *Scottish Journal of Political Economy* 52: 387–405.

FAIR. "U.S. Immigration History". Accessed October 2007. Available at http://www.fairus.org/site/PageServer?pagename=research_research9c29.

Fix, Michael, and Jeffrey S. Passel. 1994. *Immigration and immigrants: Setting the record straight*. Washington, DC: The Urban Institute.

Fortuny, Karina, Randy Capps, and Jeffrey S. Passel. 2007. *The characteristics of unauthorized immigrants in California, Los Angeles County, and the United States.* Washington, D.C.: The Urban Institute.

Frey, William H. 1996. Immigration, domestic migration, and demographic balkanization in America: New evidence for the 1990s. *Population and Development Review* 22: 741–63.

Friedberg, Rachel M., and Jennifer Hunt. 1995. The impact of immigrants on host country wages, employment and growth. *Journal of Economic Perspectives* 9: 23–44.

GAO. 2006. "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. Accessed August 2006. Available at http://www.gao.gov/new.items/d06770.pdf

Hamada, Koichi, and Hiromi Kato, eds. 2007. *Ageing and the labor market in Japan: Problems and policies.* ESRI Studies Series on Ageing. Cheltenham, UK and Northampton, MA: Elgar.

Hanson, Gordon H., and Antonio Spilimbergo. 1999. Illegal immigration, border enforcement, and relative wages: Evidence from apprehensions at the U.S.-Mexico Border. *The American Economic Review* 89: 1337–57.

Hoefer, Michael, Nancy Rytina, and Bryan Baker. 2012. "Estimates of the unauthorized immigrant population residing in the United States: January 2011." *Population Estimates* (March). Available at http://www.dhs.gov/xlibrary/assets/statistics/publications/ois_ill_pe_2011.pdf.

Hotchkiss, Julie L. 2005. Employment growth and labor force participation: How many jobs are enough? *Federal Reserve Bank of Atlanta Economic Review* 90: 1–13.

Hotchkiss, Julie L., and Myriam Quispe-Agnoli. 2013. The expected impact of state immigration legislation on labor market outcomes. *Journal of Policy Analysis and Management* 32: 34–59.

Huse, James G., Jr. 2002. "Status of the Social Security Administration's Earnings Suspense File". *Congressional Response Report* #A-03-03-23038. Accessed November 2002. Available at http://oig.ssa.gov/sites/default/files/audit/full/pdf/A-03-03-23038.pdf.

Iskander, Natasha, and Nichola Lowe. 2011. The transformers: Immigration and tacit knowledge development. NYU Wagner Research Paper No. 2011-01.

Jasso, Guillermina. 2011. Migration and stratification. *Social Science Research* 40: 1292–336.

Jasso, Guillermina, Douglas S. Massey, Mark R. Rosenzweig, and James P. Smith. 2000. The New Immigrant Survey Pilot (NIS: P): Overview and new findings about U.S. legal immigrants at admission. *Demography* 37: 127–38.

Johnson, Mary. 2007. "The growing cost of illegal immigration to social security: Unprecedented growth in social security's 'Earnings Suspense file'". *The Senior Citizen's League.* Accessed March 2007. Available at http://www.tscl.org/NewContent/102880.asp.

Jordan, Miriam. 2011. "Crackdown on Illegal Workers Grows." *WSJ.com.* Accessed April 2011. Available at http://online.wsj.com/article/SB10001424052748703951704576092381196958362.html.

Kossoudji, Sherrie A., and Deborah A. Cobb-Clark. 2000. IRCA's impact on the occupational concentration and mobility of newly-legalized Mexican men. *Journal of Population Economics* 13: 81–98.

MacQueen, James B. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability.* Berkeley, CA: Berkeley University Press, pp. 281–97.

Maloney, Thomas N., and Thomas Kontuly. 2010. Legal status and economic mobility among immigrants in the early twenty-first century: Evidence from the 'New Gateway' of Utah. In *Migration in the 21st century: Rights, outcomes, and policy*, edited by Thomas N. Maloney and Kim Korinek. New York: Routledge, pp. 226–46.

McCaffrey, Daniel F., John R. Lockwood, Kata Mihaly, and Tim R. Sass. 2012. A review of Stata commands for fixed-effects estimation in normal linear models. *Stata Journal* 12: 406–32.

Morales, Rebecca. 1983. Transitional labor: Undocumented workers in the Los Angeles automobile industry. *International Migration Review* 17: 570–96.

Neal, Larry, and Paul Uselding. 1972. Immigration, a neglected source of American economic growth: 1790 to 1912. *Oxford Economic Papers, New Series* 24: 68–88.

Nuñez-Neto, Blas, and Stephen R. Viña. 2005. "Border security: Fences along the U.S. International Border". *CRS Report for Congress*, order Code RS22026. Accessed January 2005. Available at http://www.fas.org/sgp/crs/RS22026.pdf.

Oi, Walter Y., and Todd L. Idson. 1999. Firm size and wages. In *Handbook of labor economics*, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier, pp. 2165–214.

Orrenius, Pia M., and Madeline Zavodny. 2007. Does immigration affect wages? A look at occupation-level evidence. *Labour Economics* 14: 757–73.

Orrenius, Pia M., Madeline Zavodny, and Leslie Lukens. 2008. Why stop there? Mexican migration to the U.S. Border Region. Federal Reserve Bank of Dallas WP #803.

Passel, Jeffrey S. 2013. *Unauthorized immigrants: How Pew Research counts them and what we know about them*. Washington, DC: Pew Hispanic Center.

Passel, Jeffrey S., and D'Vera Cohn. 2009. *A portrait of unauthorized immigrants in the United States*. Washington, DC: Pew Hispanic Center.

Passel, Jeffrey S., and D'Vera Cohn. 2011. *Unauthorized immigrant population: National and state trends, 2010*. Washington, DC: Pew Hispanic Center.

Pena, Anita Alves. 2010. Legalization and immigrants in U.S. agriculture. *The B.E. Journal of Economic Analysis & Policy* 10:article 7.

Peri, Giovanni. 2012. The effect of immigration on productivity: Evidence from U.S. States. *Review of Economics and Statistics* 94: 348–58.

Peri, Giovanni, and Chad Sparber. 2009. Task specialization, immigration, and wages. *American Economic Journal: Applied Economics* 1: 135–69.

Rios-Avila, Fernando. 2013. Feasible estimation of linear models with N fixed effects. Working paper #782, Levy Economics Institute of Bard College.

Toussaint-Comeau, Maude. 2007. The impact of Mexican immigrants on U.S. wage structure. Federal Reserve Bank of Chicago Working Paper 07-24.

Wen, Ming, and Thomas N. Maloney. 2011. Latino residential isolation and the risk of obesity in Utah: The role of neighborhood socioeconomic, built-environmental, and subcultural context. *Journal of Immigrant and Minority Health* 13: 1134–41.

Wright, Richard A., Mark Ellis, and Michael Reibel. 1997. The linkage between immigration and internal migration in large metropolitan areas in the United States. *Economic Geography* 73: 234–53.