Hotchkiss, Julie L.; Quispe-Agnoli, Myriam

Working Paper

The labor market experience and impact of undocumented workers

Working Paper, Federal Reserve Bank of Atlanta, No. 2008-7c

Provided in Cooperation with:
Federal Reserve Bank of Atlanta

Suggested Citation: Hotchkiss, Julie L.; Quispe-Agnoli, Myriam (2008) : The labor market experience and impact of undocumented workers, Working Paper, Federal Reserve Bank of Atlanta, No. 2008-7c, Federal Reserve Bank of Atlanta, Atlanta, GA

This Version is available at:
http://hdl.handle.net/10419/70640

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
The Labor Market Experience and Impact of Undocumented Workers

Julie L. Hotchkiss and Myriam Quispe-Agnoli

Working Paper 2008-7c
June 2008
The Labor Market Experience and Impact of Undocumented Workers

Julie L. Hotchkiss and Myriam Quispe-Agnoli

Working Paper 2008-07c
June 2008

Abstract: Using administrative data from the state of Georgia, the authors find that average wages among documented workers are lower in industries that employ undocumented workers and that a greater share of undocumented workers in those industries further lowers wages. In addition, undocumented workers have significantly lower labor supply elasticity, likely as a result of their limited employment and grievance opportunities. Furthermore, the inflow of undocumented workers does more to displace earlier hired undocumented workers than it does to displace documented workers.

JEL classification: J61, J31, J42, F22

Key words: illegal immigration, wage impact, worker displacement
I. Introduction and Background

The United States has a long history of immigration debate. Through the last century and into this one, immigration policy has been subjected to changing economic needs, fears, and political whims. Positive contributions of immigration have been identified by Neal and Uselding (1972) who estimate that the flow of immigrants into the United States between 1790 and 1912 resulted in a 13 to 42 percent higher level of capital stock than would have prevailed in the absence of immigration during these years (also see Barro and Sala-i-Martin 1995 and Chiswick et al. 1997).\footnote{In contrast, Morley (2006) finds empirical evidence of economic growth leading to increased immigration, but not vice versa.} Immigration has also been more recently explored in various countries as a mechanism for replacing retiring baby-boom workers (e.g., Hamada and Kato 2007, Hotchkiss 2005, Denton and Spencer 1997).

The 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 literature presents a wide range of estimates of the effects of immigration on wages and employment of native workers. The consensus settles on a one to four percent decrease in native wages resulting from a 10 percent increase in the population share of immigrants (for example, see Friedberg and Hunt 1995 and Borjas et al. 2006). The measured impact of immigration on the displacement of workers is less clear. Card (1990), Wright et al. (1997), Butcher and Card (1991), and Card and DiNardo (2000) find no evidence of immigrant inflows affecting native migration patterns or employment outcomes. Whereas, Frey (1996) and Borjas (2005) identify a significant relationship between immigrant inflows and either native
outflows or lower net native in-migration, and Card (2001) finds lower rates of employment within cities with high immigrant arrivals.

While firmly rooted in the empirical literature measuring the impact of legal immigration on the labor market experiences of natives, this paper deviates slightly by exploring the impact of undocumented workers (assumed to be immigrants) on wages of documented workers, some of which are likely immigrants themselves. The analysis makes use of administrative data from the state of Georgia to investigate how the proportion of undocumented workers affects the wages of documented workers, the displacement of documented workers, and what role different labor supply elasticities might play in the observed wage gap between the two groups of workers. The results presented in this paper are of particular relevance as the immigration debate has recently narrowed its focus on undocumented immigrants.

\textit{A. The Impact of Immigration}

The impact of immigration on native worker wages essentially comes down to how complementary or how substitutable the immigrants are with native workers, and how responsive native migration patterns are to the influx of additional workers. Goldin (1994) documents significant wage effects of the large European migration to the United States in the early 20th century. Measured effects of more recent waves of immigrants have been more modest.\footnote{For example, Goldin's (1994) wage effect estimate is about 10 times larger than that of Altonji and Card (1991).} Friedberg and Hunt's (1995) review of the literature summarizes the consensus view through the mid-1990s that a 10 percent increase in the fraction of immigrants reduces native wages by (at most) one percent, and that immigration has no effect of practical significance on native employment.\footnote{Deviations from this consensus are nicely summarized in Fix and Passel (1994).} Friedberg and Hunt also point out, however, that the measured effect varies across skill groups. For example, Borjas et al. (1997) conclude that while immigration of less-
skilled workers (and trade influences) between 1980-1995 might account for half of the relative wage declines of high school dropouts over the period, the effect did not contribute significantly to the widening wage gap between skilled and low-skilled workers. A more recent analysis by Orrenius and Zavodny (2007) finds that, over the period between 1994 and 2000, wages of native manual laborers was about 0.8 percent lower than they would have been in the absence of legal immigrant inflows over the period. Furthermore, the impact on workers in professional occupations is found to be insignificant.

Other recent estimates put an upper bound on the negative impact of immigration on native wages, at most, at a four percent loss per 10 percent increase in immigrant share (see Borjas 2003, 2005; and Borjas et al. 2006). In addition, Borjas (2005) estimates that the measured wage effect is mitigated by out-migration of natives from areas experiencing significant immigrant in-flows. On the other end of the spectrum is the paper by Ottaviano and Peri (2006) who estimate an overall positive influence of immigration on native wages. This result is achieved by allowing immigrants and natives to be less than perfect substitutes and by allowing yearly adjustment of physical capital in response to immigration flows. Ottaviano and Peri, however, do find that wages of earlier immigrants suffer significantly with the arrival of more recent immigrants.4

The potential of displacement of native workers by the arrival of immigrants can result in a number of ways. If the arrival of immigrants depresses wages in a particular labor market, native workers, enjoying greater mobility, might migrate to a geographic location less inundated with immigrants or to a different industry/occupation all together. In addition, if native workers view the arrival of immigrants as "writing on the wall," they may choose to seek alternative employment (geographically or sectorally) before being replaced. These two forms of

4 This finding is consistent with earlier evidence provided in Lalonde and Topel (1991).
displacement are technically voluntary in nature, although precipitated by an inflow of immigrants, and have been the primary type of displacement documented in the U.S. Frey (1996) reports that traditional immigrant ports-of-entry metropolitan areas experienced significant and consistent net out-migration compared to other parts of the U.S. In addition, Borjas (2005) finds that states experience a reduction in native net in-migration (number moving in minus number moving out) of two people for every 10 new immigrants. Since immigrants tend to locate in cities, the impact is larger when measured at the metropolitan level.

Involuntary displacement can be said to occur if a native worker loses his/her job as a direct result of being replaced by an immigrant worker. Rosenfeld and Tienda (1999) find evidence that Mexican immigrants displaced or succeeded low-skilled African American natives in several industries in Los Angeles, Chicago, and Atlanta (also see Ong and Valenzuela 1996). Further, while finding no link between immigrant inflows and native migration, Card (1997) does find that between 1985 and 1990 the employment rate among low-skilled native men and earlier immigrants declined by a greater amount in metropolitan areas that were experiencing significant inflows of immigrants, such as Miami or Los Angeles. In addition, a case study of manicurists in California by Federman et al. (2006) found that for every five Vietnamese manicurists entering the market, two non-Vietnamese manicurists were displaced.

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. Second, undocumented workers are likely to be even more limited in their opportunities and therefore have lower elasticities of labor supply. 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.

B. Immigration Policy in the U.S.

Immigration legislation dates from the founding of the nation. The two most recent comprehensive 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. The IRCA responded to the growing population of undocumented immigrants by creating two amnesty programs for unauthorized immigrants and a new classification for seasonal agricultural workers. The Seasonal Agricultural Worker amnesty program allowed immigrants who had worked for at least 90 days in certain agricultural jobs to apply for permanent residence. The Legally Authorized Workers amnesty program allowed current undocumented immigrants who could prove residence in the U.S. since January 1, 1982 to legalize their status. Under the two amnesty programs, roughly 2.7 million undocumented people residing in the United States became lawfully permanent residents. At the same time, this reform established sanctions for employers who knowingly would hire or recruit undocumented workers. In addition, the legislation mandated that states use the Systematic Alien Verification for Entitlement System, an automated verification system to track the immigration status of applicants for welfare.

According to Fix and Passel (1994), the amnesty programs were very successful in legitimizing undocumented residents; however employer sanctions have, "largely failed to

---

5 For historical details, see CBO (2006) and FAIR (2007).
control illegal immigration in the 1990s. Employer sanctions have proven difficult to enforce because of the increased prevalence of fraudulent documents and the limited resources thus far dedicated to enforcement by the Immigration and Naturalization Services (INS)” (p.16).

Addressing the concerns of a growing population of unauthorized immigrants, the 1996 Illegal Immigration Reform and Immigrant Responsibility Act (IIRIRA) set new guidelines for border enforcement and eligibility verification for work or social services. The IIRIRA increased the number of border patrol agents and introduced new border control measures. In addition, it reduced government benefits available to immigrants, and established a pilot program in which employers and social service agencies could check the eligibility of applicants. The employment verification program is voluntary and the Government Accounting Office (GAO) has found that document fraud (use of counterfeit documents) and identity fraud (fraudulent use of valid documents belonging to others) made it very difficult for employers to comply with this verification process (GAO 2005).

According to the GAO study, between 1999 and 2003, the number of man hours that Immigration and Customs Enforcement (ICE) agents devoted to worksite inspections declined from 480,000 inspections (9 percent of total INS agent hours) to 180,000 hours (or 4 percent of total ICE agent hours). Therefore, this low worksite enforcement implies that fewer employers hiring undocumented workers are detected or prosecuted. Since September 11, 2001, ICE concentrated efforts on sites that could represent national security vulnerability, consistent with its mission to combat terrorism. Finally, the 2005 GAO study concludes that, “under the former INS and now under ICE, worksite enforcement has been a relatively low priority.”

---

7 The Homeland Security Act of 2002 created the Department of Homeland Security (DHS). The Immigration and Naturalization Service (INS), in charged of immigration services, border enforcement and border inspections, was restructured in three bureaus within DHS. Immigration and naturalization services are provided by the Bureau of
C. Undocumented Immigrant Legislation in Georgia

On July 1, 2007, the Georgia Security and Immigration Compliance Act (SB529) went into law. This first immigrant legislation in Georgia covers employment verification, eligibility for public benefits, human trafficking, tax withholdings, state enforcement of federal immigration laws, and ethics standards for the provision of immigration services. All employers who do business with the State of Georgia with more than 500 employees, along with all of their subcontractors (of more than 500 employees), are required to enroll in the federal government Basic Pilot Verification Program. By July 2009, coverage will be extended to all public employers. In addition, any company or individual (whether doing business with the state or not) is prohibited from deducting as a business expense on their state income taxes more than $600 per year paid to a person who has not been verified to legally work in the U.S. In addition, employers are required to withhold six percent of state income tax of an individual using a non-U.S. resident taxpayer identification number.

Moreover, SB529 authorizes the training of law enforcement officers to enforce federal immigration laws and it requires that an arresting officer determines the legal status of any person arrested for a felony or DUI and to notify ICE if the individual does not have documentation. The law also requires the verification of U.S. citizenship or residence status of any person, 18 years or older, applying for state or local benefits.

The intended effect of the legislation is to increase the penalty imposed on employers for hiring undocumented workers. The penalty imposed on undocumented workers also increases as employment opportunities will likely be further curtailed with increased enforcement. In Citizenship and Immigration Services; border enforcement functions are performed by two bureaus, the Bureau of Customs and Border Enforcement (ICE), and the Bureau of Customs and Border Protection.

8 Much of the information contained in this section was obtained from Kuck (2007) and Bess (2007). By November 2007, 244 immigration laws have been enacted by 46 state legislatures (NCSL 2007).
addition, driver’s licenses in Georgia are only obtainable with proof of citizenship or legal residence, which further restricts mobility of undocumented workers. The legislation also creates severe penalties for the offense of human trafficking and establishes ethics standards for the provision of immigration services.

D. Measuring Undocumented Immigration

The first step in determining the impact of the presence of undocumented immigrants is identifying who is undocumented. The most common method used to estimate the number of unauthorized immigrants is the residual approach, or merely calculating the difference between the total measured foreign-born population and the legal immigrant population.9 Table 1 shows the estimates from various sources of the number of unauthorized immigrants in the U.S. using the residual approach. The legal immigrant population includes lawful permanent residents (LPR), asylees, refugees, and non-immigrants whose information is obtained from the office of Immigration Statistics of the Department of Homeland Security (DHS). The foreign-born population is estimated from data collected by the U.S. Census Bureau, either through the American Community Survey (ACS) or the Current Population Survey (CPS). According to the latest figures, there are 12.4 million unauthorized immigrants living in the U.S. as of March 2007 (Camarota 2007; this figure accounts for the probable undercounting by the CPS). It is also estimated that about four percent of the total (554,000 persons) are located in Georgia. Between 2000 and 2006, the greatest percentage increase of unauthorized immigrants in the U.S. occurred in Georgia--a 123 percent increase, equivalent to an average annual increase of 45 thousand unauthorized immigrants (Hoefer et al. 2007).

[Table 1 here]

---

9 See Hanson (2006) for a review of different sources and estimates of undocumented immigrants.
Estimates using this residual approach suggest that stocks of undocumented immigrants have risen sharply over time. However, there is considerable variability in the estimates, associated with differing assumptions about the magnitude of errors in enumerating legal and unauthorized immigrants in official data sources. One of the primary sources of error in these estimates is the variability of sample sizes in the ACS across years, which leads to difficulties in assessing the degree to which the foreign-born population may be undercounted or may have emigrated (Hoefer et al. 2007). In addition, the ACS estimates assume that the current residence of a legal immigrant is the same one when they obtained LPR status; the estimates do not take into account internal migration. Other concerns exist regarding the validity and reliability of Census survey data on the year of entry. Further, errors on admission counts, length of visit, and changes in status might result in double counting of non-immigrants and persons adjusting to LPR status.

A second data source on unauthorized migration is information on border apprehensions from the U.S. Border Patrol. Estimating the level of unauthorized immigration using apprehension data is problematic, primarily because it is not only a function of the number of attempts to cross the border (which have been shown to vary with expected relative U.S./Mexico economic conditions), but also a function of the enforcement efforts of border patrol and a function of the number of attempts (see Hanson and Spilimbergo 1999, and GAO 2006). Evaluating apprehension data between 1977 and 1988, Espenshade (1995) estimates that unauthorized immigration exceeds the level of apprehensions by an order of 2.2. Even if that factor were cut in half (as a result of greater resources being devoted to border patrol efforts since 1988), the 1.1 million apprehensions along the Southwest border of the U.S. in 2004 would
mean that over a million undocumented migrants made it across the border.¹⁰

According to DHS estimates for January 2006, 57 percent of unauthorized immigrants come from Mexico, not a considerable change from 55 percent in January 2000. Therefore it is not surprising that surveys from Mexico constitute a third source of data on unauthorized immigrants. The Mexican Migration Project (MMP) is a household survey that has been conducted annually since 1982. The survey is conducted during the winter months when seasonal migrants return to Mexico. The Legalized Persons Survey (LPS) is another survey including undocumented immigrants who were granted permanent legal residence in the U.S. under the amnesty provision of the Immigration and Control Act of 1986. The LPS consisted of an initial survey in 1989 and a follow-up in 1992 (Hanson 2006: 884). In general, the MMP and LPS have been found to be more useful in characterizing undocumented immigrants than actually counting them. Orrenius and Zavodny (forthcoming), using the MMP, report that over the period between 1980 and 2004, approximately 62 percent of migrants from Mexico were undocumented.

Among the newest sources of data of information about immigrants is the New Immigrant Survey (NIS). The sample includes new legal permanent residents in the U.S., admitted in 1996, and over-samples employment based immigrants. The immigrants are administered three surveys over a 12 month period and are asked a host of questions about their original entry into the U.S. and about their experiences since arriving. Based on this survey, combined with administrative micro-data from the Department of Homeland Security, Jasso et al. (2008) estimate that 32 percent of new adult immigrants granted legal permanent residence in the U.S. in 1996 had originally arrived in the U.S. illegally, overstayed visa expirations, or were employed illegally.

¹⁰ Hanson (2006) estimates that number is closer to 300,000 per year since 2000.
In general, survey data sources on unauthorized immigrants are subject to sample-selection problems. For example, the MMP survey includes seasonal migrants, mostly in agriculture, and the LPS specifically excludes seasonal migrant workers. This paper differs in the way in which unauthorized individuals are identified. Most importantly, it does not rely on survey results. State administrative data are used to identify invalid social security numbers used by employers in reporting worker earnings. It is a common misconception that undocumented workers are all working "off the books." There is considerable evidence that employers do report, either knowingly or unknowingly, and pay taxes on the wages paid to undocumented workers. Unlike most other studies, the measure used here does not capture the supply of undocumented workers, but, rather, the demand, as the workers are identified through employment records. The advantage of this data source is that it is not subject to sample selection issues plaguing survey results. The disadvantage is that it does not capture undocumented workers that are not reported on employers' payrolls.

II. Data

The data used for the analyses in this paper are the Employer File and the Individual Wage File, compiled by the Georgia Department of Labor for the purposes of administering the state's Unemployment Insurance (UI) program. These data are highly confidential and strictly limited in their distribution. The Employer File provides an almost complete census of firms in non-farm sectors, covering approximately 97 percent of non-farm workers. The establishment level information includes the number of employees, the total wage bill and the NAICS classification of each establishment. The Individual Wage File contains quarterly earnings

---

11 White et al. (1990) provide an extensive discussion about the use of these employment data, commonly referred to as the Quarterly Census of Employment and Wages (QCEW), or ES-202 data.
information for all workers employed by these establishments. Regrettably, the data set contains no information about the worker's demographics (e.g., education, gender, race, etc.). There is no specific information about the worker's job (e.g., hours of work, weeks of work, or occupation). One important implication of this is that the worker's part-time/full-time status is unknown, so if undocumented workers are disproportionately part-time employed and documented workers are mostly full-time employed, any wage gap will be over-estimated.

Because of the lack of individual characteristics of workers (besides earnings), some of the analyses are performed at the 6-digit NAICS level. Workers are more homogeneous in skill level in some industries, such as construction, than in others, such as professional and business services. One seeming disadvantage to using UI records data to identify undocumented workers is the lack of coverage in the agriculture industry where one might expect to find a significant number of undocumented workers. (Less than half of agricultural workers are covered by UI.) However, Card and Lewis (2007) report that between 1990 and 2000, among Mexican migrants who have been in the U.S. 0-5 years, the share working in agriculture fell from 23 percent to 15 percent among men and from 13 percent to seven percent among women.

The data are available from the first quarter of 1990 through the fourth quarter of 2006. In each quarter, and within each 6-digit industry, the total number of workers, the total number of firms, the share of firms that were born or shut-down, the number of undocumented workers,

---

12 Included in earnings are pay for vacation and other paid leave, bonuses, stock options, tips, the cash value of meals and lodging, and in some states, contributions to deferred compensation plans (such as 401(k) plans). Covered employer contributions for old-age, survivors, and disability insurance (OASDI), health insurance, unemployment insurance, workers' compensation, and private pension and welfare funds are not reported as wages. Employee contributions for the same purposes, however, as well as money withheld for income taxes, union dues, and so forth, are reported even though they are deducted from the worker's gross pay. Because the Individual Wage file contains a firm rather than establishment identifier, a choice of which NAICS code to assign to each worker who was employed by a multi-establishment firm is required. Following the Department of Labor convention, a 6-digit NAICS code is assigned based on the largest share of the firm's total employment.

13 Restricting the sample to workers earning at least $1,000 (real) per quarter, in an effort to mitigate the potential problem of hours differences, produced qualitatively similar results.
and the average quarterly earnings of documented and undocumented workers are calculated. There is no identifier for whether a worker is an immigrant or not. Therefore, it is very likely that immigrants are included among the documented workers.

**A. Identifying Undocumented Workers using 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 Fair Labor Standards Act (FLSA).\(^4\) Each worker on this report is identified by his/her social security number (SSN). There are a number of ways in which one can establish that the reported social security number as invalid. The Social Security Administration provides a service by which an employer can upload a file of SSNs for checking, but one must register as an employer to obtain this service.\(^5\) Alternatively, there are several known limitations on what can be considered a valid social security number and so a simple algorithm is used to check each number to make sure it conforms to the valid parameters.

There are three pieces to the SSN.\(^6\) The first three numbers are referred to as the Area Number. This number is assigned based on the mailing address (the state) stated on the SSN application; 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 (see <http://www.ssa.gov/employer/ssnvhighgroup.htm>). Any SSN with an Area Number equal to 000 or greater than 772, or shows up before the officially assigned

---

\(^4\) Information about which workers are covered, see U.S. Department of Labor (2007).
date, will be considered invalid.

The second piece of the SSN consists of the two-digit Group Number. The lowest group number is 01 and they are assigned in non-consecutive order based on whether the Area Number is odd or even. The Social Security Administration publishes the maximum Group Number issued for every Area Number as of certain dates. No hits were found in checking for the presence of these invalid Group Numbers for one quarter of the data (several million observations). Given the time consuming nature of this particular search, all 18 years of data were not checked for this type of invalid SSN. Any SSN with a Group Number equal to 00 will be considered invalid.

In 1996 the Internal Revenue Service (IRS) introduced the Individual Tax Identification Number (ITIN). These numbers were introduced by the IRS to allow individuals who had income from the U.S. to file a tax return. It is simply a "tax processing number," and does not authorize an individual to work in the U.S. Employers are instructed by the IRS to, "not accept an ITIN in place of a SSN for employee identification for work. An ITIN is only available to resident and nonresident aliens who are not eligible for U.S. employment and need identification for other tax purposes." The importance of ITINs for undocumented immigrants, for example, to open bank accounts, is clear in noting the percent of workers identified with SSNs that match the number scheme of ITINs. These 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 number scheme will be identified as having an invalid Area Number, as they are not authorized to be working. The percent of SSN matching the ITIN number scheme, in construction for example, has risen from less than one percent in 1996 to over 40 percent in 2006.

---

The last four digits of the 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 will be considered invalid.

There were a series of SSNs that were de-commissioned by the Social Security Administration because they had been put on fake Social Security Cards used as props to sell wallets.\(^{18}\) Apparently, some people who purchased the wallets thought the fake Social Security Cards were real and started using them as their own. If any of these 21 "pocketbook" SSNs appear in the data, they are considered invalid. In addition, there are a number of SSNs that are exactly equal to the employer identification number. These are considered invalid. In any instance where a SSN is used for more than one person on a firm's UI wage report, that SSN will be considered invalid.\(^{19}\) Lastly, a SSN that does not have the required number of digits (including zeros) will also be considered invalid. Table 2 lists the reasons why a SSN is classified as invalid for the purposes of this paper.

[Table 2 here]

The means in Table 2, which reports the incidence of invalid SSNs over the entire time period, mask an important dynamic that will show up in the next section. That is, the incidence of some of these reasons for being invalid have a very strong cyclical component and some have grown remarkably over the time period in some industries. In addition, these reasons for being invalid are not mutually exclusive. For example, a SSN may be invalid because it has a high area number, but it also may be duplicated within the firm. Nonetheless, over the period the single largest reason a SSN is considered invalid is because of duplication on a UI report. 1.53


\(^{19}\) Since the same invalid SSN could be used by different people at different employers, only the undocumented worker who is earning the highest quarterly wage in any quarter is retained. The implication is that any calculated wage differential between documented and undocumented workers will be an underestimate of the actual difference. It will also result in an undercount of the number of undocumented workers in any quarter.
percent of all records over this time period are duplicates within a firm in the same quarter. The next largest reasons are zeros in any of the pieces and a SSN that is equal to the employer ID; each of these reasons appears for about 0.3% of the sample. The incidence of pocketbook SSNs is very small. SSNs with an Area Number that is too high shows up in 0.07% of the observations. Overall, about 0.39% of observations have an invalid SSN for any one of the reasons listed in Table 2. This amounts to just over one million of the workers in the full sample.

B. Growth in Undocumented Workers in Georgia 1990-2006

The means in Table 2 average the incidence of invalid SSNs across the whole time period and across all industries. There is reason to expect that invalid SSNs are more likely to be concentrated in certain industries and to have been growing over time. Figure 1 plots the percent of workers with invalid SSNs in six broadly defined NAICS industries (to keep the Figure from being too cluttered, only those industries with the greatest shares of undocumented workers are presented). Consistent with expectations, the industries in which the largest percent of workers with invalid SSNs are construction and leisure and hospitality. Manufacturing is in the middle of the pack at the bottom of the Figure. In addition to the significant growth in invalid SSNs seen in the construction industry beginning in about the year 2000, there appears to be a strong cyclical component to the presence of invalid SSNs. The peaks occur in the third quarter of every year and may have something to do with the timing of firms' cleaning up their records.

[Figure 1 here]

To explore the nature of the cyclicalality of the occurrence of invalid SSNs, Figure 2 plots the percent of workers with invalid SSNs in the construction industry only, by reason of invalid classification (excluding the pocketbook reason since it was so small). The striking story of this figure is that the third quarter cyclicalality is unique to SSNs with all zeros in one of the pieces and
those that are equal to the employer ID. Duplicate SSNs appear to have their own irregular
cyclicality with spikes happening often in the first quarter. Notably, the incidence of invalid
SSNs for these three reasons (as a percent of all workers) seems to be on the decline.
Remarkably, however, the incidence of SSNs with larger-than-valid Area Numbers exhibits no
cyclical behavior, and it has seen a fairly significant growth in the construction industry since the
late 1990s.20

[Figure 2 here]

Given the apparent administrative cyclical ity for all but the High Area Number invalid
SSN reason, only this reason is used as a conservative measure of the percent of undocumented
workers in an industry. This will clearly undercount the actual number of undocumented
workers, so that any effect identified in the analysis will also likely under-estimate the true effect
of the presence of undocumented workers on documented worker outcomes.

In order to have the "cleanest" group of documented workers possible, any worker with
an invalid SSN for a reason other than a High Area Number is deleted from the analysis. This
will ensure that an undocumented worker is not classified as documented. It is possible, of
course, that an undocumented worker fraudulently provides a valid SSN to his/her employer.
This person will only be in the sample if no one else reported this SSN or if the true owner of the
SSN is earning less than the fraudulent reporter.

Figure 3 plots the incidence of High Area Number by the top six broadly defined NAICS
industries. The growth of the percent of workers with this type of invalid SSN in construction is

---

20 One might suspect that the cyclicality associated with the incidence of the three bad SSN reasons other than bad
high area is related to employment variation in hiring and laying off of undocumented workers. However, the
 cyclicality of the share of workers that is undocumented is more correlated (positively) with the cyclicality of
documented worker employment than with undocumented worker employment, suggesting that the administrative
clean-up has more to do with the firms' overall employment record keeping than with any records or hiring practices
associated with undocumented workers in particular.
striking. In contrast to Figure 1, most industries see a growth in undocumented workers over this time period when only looking at the Area Number (the omitted industries also show a growth, but at a lower level). And it is here where the concentration of undocumented workers identified by Fortuny et al. (2007) is seen more clearly: construction, leisure and hospitality, professional and business services, wholesale and retail trade, and manufacturing. The growth in the share of undocumented workers seen in Figure 3 is also consistent with Fortuny et al. who estimate that 72 percent of unauthorized immigrants in Georgia arrived in the last 10 years.

[Figure 3 here]

Table 3 presents some sample means, across the entire time period and at each end of the time period. The percent of firms in an industry employing any undocumented workers increased over the time period from 37 percent to 54 percent. The experience of the construction and leisure and hospitality illustrates the variation observed across large industry groups. An average of 94 percent of firms within construction's 6-digit industries were employing undocumented workers by 2005, while an average of only 55 percent of firms in leisure and hospitality were doing so. There is not a great deal of average industry dynamics across industries, but these regressors are expected to vary more across business.

[Table 3 here]

Within industries that hired undocumented workers, the average percent of workers that were undocumented, across all industries (p), went from 0.12% in 1991 to 0.41% in 2005. Camarota (2007) estimated that seven percent of Georgia's workforce was undocumented in 2007. Of course, the method used in this paper is expected to under-count the number of undocumented workers, since it does not capture workers not reported on UI wage reports. However, the pattern of growth exhibited in Figure 3 is consistent with the numbers reported by
Fortuny et al. (2007). They report a 900 percent increase in the number of undocumented labor force participants in Georgia between 1990 and 2004; the data used for this analysis reflect a growth of about 800 percent in undocumented workers over the same time period. In addition, the number of undocumented workers identified with these data is roughly two percent of the number reported by Fortuny et al., and that relative measure is consistent across time.

The means also indicate that the average industry wage penalty (or percentage difference in earnings between documented and undocumented workers), across industries that employ undocumented workers, has increased over time. For the most part, real industry average quarterly earnings rose among documented workers but declined, for the most part, among undocumented workers.

III. Empirical Analysis

A. The Wage Impact of Undocumented Workers

A number of different approaches have been taken to quantify the impact of immigration on native worker wages and employment. The most common strategy is used by Altonji and Card (1991) and in a number of papers by George Borjas (alone and with co-authors; 2003, 2005, 2006). The procedure makes use of decennial census data and standard linear regression to identify a relationship between the difference in the density of immigrants and wages or employment across geographic areas (usually metropolitan statistical areas, MSAs). Various techniques (e.g., instrumental variables and fixed-effects through differencing) are employed to control for the endogeneity problem of immigrants selecting their geographic destination based on observed wages in those locations.
Two estimations are performed to determine the impact of the presence of undocumented workers on wages. The first specification estimates the impact of the presence of any undocumented workers on wages of documented workers in the same industry. Then, conditional undocumented workers being present in the industry, the second estimation includes the share of undocumented workers as a regressor determining wages of documented workers; the impact on undocumented worker wages is also explored. The second estimation is the one that follows Borjas more closely.

The first estimating equation is specified as:

\[
\ln w_{jt}^k = \alpha_0 Udum_{jt-1} + \alpha_1 x_{jt-1} + \tau_{jt}
\]

where \( \ln w_{jt}^k \) is the log of the average quarterly earnings of documented workers \((k=d)\) in industry \(j\) and time \(t\); \(Udum_{jt-1}\) is equal to one if any undocumented workers were identified working in industry \(j\) in the previous year, zero otherwise; \(x_{jt-1}\) are other (lagged) regressors expected to influence the observed base wage level; and \(\tau_{jt}\) is the random error. This equation is not estimated for undocumented workers since the average wage in industries that do not employ undocumented workers is zero.

The second specification replaces the undocumented dummy variable with \(p_{jt-1}\), which is the share of workers in industry \(j\) in the previous year that is undocumented:\(^{21}\)

\[
\ln w_{jt}^k = \beta_0^k p_{jt-1} + \beta_1^k x_{jt-1} + \epsilon_{jt}^k
\]

Equation (2) will be estimated conditional on \(p_{jt-1}>0\) and estimated separately for documented workers \((k=d)\) and undocumented workers \((k=u)\)

---

\(^{21}\) \(p_{jt-1} = 100 \times [N_{jt-1}^u / (N_{jt-1}^u + N_{jt-1}^d)],\) where \(N\) is the number of undocumented \((u)\) and documented \((d)\) workers in industry \(j\).
Both equations are estimated as fixed-effects models with first-order autocorrelation. Lagged values of the share of undocumented workers and other regressors are used in order to address potential issues of endogeneity. The data are not rich enough to allow for any attempts at instrumental variables estimation. The strict assumption of independence of regressors and the error term required for random effect estimation was rejected by the data, although results obtained via random effects are qualitatively the same. Serial independence of the error terms was rejected based on the evaluation of the Bhargava et al. (1982) modified Durbin-Watson test statistic, although fixed-effects results not correcting for autocorrelation were also qualitatively the same as those presented here (a test suggested by Wooldrige 2002: 282-3 and expanded on by Drukker 2003 also soundly reject serially independent errors).

The estimations also include a continuous year-quarter time trend, which will capture underlying quarterly cyclical variation and potential long-term trend influences. The number of firms and total employment in industry \( j \) at time \( t \) are also included, as well as the percent of firms in each industry that were born (newly open) or died (shut down). A firm is considered to have been born in time period \( t \) if it has positive employment in \( t \) but zero employment in the previous four quarters. A firm is considered to die in period \( t \) if it has positive employment in \( t \) but zero employment in the following four quarters. These regressors are expected to capture important industry-specific dynamics. The fixed-effects specification is designed to control for any industry-specific, time-invariant influences on wages.

Results for both specifications are presented in Table 4. Results in panel (a) indicate that average wages of documented workers are lower in those industries that employ undocumented workers. Wages are six percent lower in Construction (about $1,700 annually), 1.5 percent

\(^{22}\) Specifications including quarterly dummies and a measure of annual gross state product did not yield qualitatively different results.
lower overall (about $500 annually), but not statistically significantly lower in Leisure and Hospitality. After controlling for overall economic conditions, workers in an industry with a larger share of new firms earn lower wages, on average, but earn higher wages if a greater share of firms is going out of business. These regressors could be proxying for age of the industry, with older industries being more volatile and paying higher wages.\textsuperscript{23} In addition, wages appear to be decreasing at an increasing rate in the number of firms in an industry and in total industry employment. Greater number of workers could be pushing wages down as competition for jobs increases, or a greater number of workers could be reflecting lower worker productivity. It's unclear why a greater number of firms in an industry would result in lower average wages.

[Table 4 here]

Panel (b) of Table 4 presents the results from estimating equation (2), where the share of undocumented workers (in the previous year) is included as a regressor, and the sample consists only of those industry that employ undocumented workers.\textsuperscript{24} Among industries that employ undocumented workers, a larger share negatively impacts wages of documented workers but does not significantly impact wages of undocumented workers. The coefficient on $p_{t-1}$ indicates that a one percentage point increase in the share of undocumented workers can be expected to reduce earnings of documented workers by six percent across all sectors, on average; by 26 percent in construction; and by four percent in leisure and hospitality.

Based on the most recent estimates of the growth of undocumented workers between 2000 and 2007, the share of undocumented workers in Georgia has increased from four to seven percent (calculated using statistics reported in Hoefer et al. 2007 and in Camarota 2007). Our

\textsuperscript{23} Others have shown opposite results when analyzing micro-level data (Hotchkiss et al. 2004), and that the age of an individual manufacturing firm actually increases its chances of survival (Cefis and Marsili 2005). The difference in these results might be the consequence of performing the analysis on averages at the industry level.

\textsuperscript{24} Results by other broad industry classifications are reported in the Appendix and reflect a negative impact on documented worker wages within other broad industry groups.
analysis indicates that as a result of this three percentage point growth in the share of undocumented workers over seven years, average annual earnings of documented worker in Georgia were about 2.5 percent lower overall, 11 percent lower in Construction, and almost two percent lower in Leisure and Hospitality.

Recall that the bulk of the literature suggests that a one percent increase in the population share of immigrants results in a 0.1 to 0.4 percent decrease in native wages. The results in Table 4 indicate that the impact of a one percent increase in the share of undocumented workers can be expected to be considerably larger than the impact of the same increase in immigrants as a whole.\textsuperscript{25} As will be discussed below, this larger statistical impact is what one might expect given the more limited grievance and employment opportunities of undocumented workers, compared to that of immigrants generally.

Unlike the findings of Ottaviano and Peri (2006) and Lalonde and Topel (1991), who find that immigrants often experience an even greater negative impact from an increase in immigration, the results in Table 4 (and the Appendix) indicate this is not the case for wages among undocumented workers. However, these results are conditional on employment in an industry that hires undocumented workers in the first place; we do not know what wages for undocumented workers would look like in industries that do not employ undocumented workers.

\textit{B. Labor Supply Elasticities and the Undocumented Worker Wage Penalty}

As seen in Table 3, the average wage gap between documented an undocumented workers is sizeable. One of the ways for employers to successfully pay undocumented workers less than their documented co-workers is by exploiting a possible difference in the labor supply elasticities across the two workers. A commonly accepted source of differential elasticities of

\textsuperscript{25} Given that undocumented workers are such a small share of the total work force, the parameter estimate itself is a reasonable approximation of the elasticity.
labor supply across workers with different characteristics (usually gender or race) is the presence of constraints. It is argued that, historically, the employment opportunities for blacks and women are less than those for white men, that women are geographically constrained by their husband's employment choices (for example, see Raphael and Riker 1999 and Ofek and Merrill 1997), and these constraints at least partially contribute to observed wage differentials across race and gender.

Analogously, if undocumented workers are constrained in their employment opportunities, or at least in their grievance opportunities, their labor supply elasticity should be lower on average than that of a documented worker. A technique introduced by Ransom and Oaxaca (2008) is used to estimate the labor supply elasticity of documented and undocumented workers. The validity of this estimation relies on a number of assumptions. First, the technique requires that recruitment equals separation; that one employer's separation is another employer's recruitment. This means that it would not necessarily be valid in circumstances of very weak labor markets. Because of this, the separation equations will be estimated only during the period of expansion, 1994-2000, inclusive. A second assumption is required as a result of how undocumented workers are defined. Since undocumented workers are defined as those using invalid SSNs, it is not reasonable to expect that an undocumented worker would use the same SSN when moving from one employer to another. However, it must be assumed that an undocumented worker uses the same SSN while employed by the same employer, so that if the undocumented worker's SSN disappears from the employer's records for a period of time, it is assumed that the worker has separated. Of course, there is nothing to prevent a different worker from using the same invalid SSN later on with the same employer.
Given these considerations, the labor supply elasticity can be estimated as the negative of two times the separation elasticity (see Ransom and Oaxaca 2008, p. 4):

\[ \varepsilon_{nw} = -2\varepsilon_{sw} \]  
(3)

The separation elasticity is estimated by first estimating the following separation equation separately for documented workers \((k=d)\) and for undocumented workers \((k=u)\):

\[ s_{it}^k = \Phi \left[ \gamma_0^k + \gamma_1^k \ln(w_{it}^k) + \gamma_2^k h_{jt}^k + \gamma_3^k Y_{it}^k \right] = \Phi \left[ \Gamma_{ijt}^k \right], \]  
(4)

where \( s_{it} \) is the probability that worker \( i \) separates from his/her employer in industry \( j \) in quarter \( t; \ \Phi[\cdot] \) is the normal cumulative distribution function; \( w_{it} \) is the real quarterly wage observed for worker \( i \) in quarter \( t; \ h_{jt} \) is the percent of new hires in industry \( j \) that are undocumented;\(^{26}\) and \( Y_{it} \) are other characteristics of the worker, firm, or labor market that might affect the rate of separation. In order to control for the possibility that undocumented workers are drawn to industries experiencing a rising relative demand for their skills or to industries that have a history of hiring undocumented workers (see Card and DiNardo 2000), the share of workers in the industry that are undocumented is also included as a regressor, as well as industry fixed effects. Again, the state administrative wage files provide only limited information about the workers’ firms and industries and no information about the worker demographics. A worker is considered separated if the worker’s SSN disappears from the employer’s files for at least four quarters.

Given the estimation results from equation (4), obtained via maximum likelihood probit, the separation elasticity with respect to wages for workers of type \( k (= d, u) \) can be calculated as follows:

\[ \varepsilon_{sw}^k = \frac{\partial s}{\partial w s} = \frac{\hat{\gamma}_2^k \varphi(\hat{I}_i^k)}{w s} = \hat{\gamma}_2^k \left[ \frac{\varphi(\hat{I}_i^k)}{\Phi(\hat{I}_i^k)} \right], \]  
(5)

\(^{26}\) The coefficient on this regressor will be used to investigate potential displacement in the next section.
where $\varphi(.)$ is the standard normal density function (Ransom and Oaxaca 2008, p. 12). Table 5 reports the separation and labor supply elasticities for all workers and for workers in the construction and leisure and hospitality industries (estimates for other broad industries are presented in the Appendix).

Table 5 here

The first thing that the estimates in Table 5 indicate is that during this period of rapid economic expansion, all workers have fairly high labor supply elasticities; among documented workers in all industries, a one percent increase in the wage increases the quantity of labor supplied by 1.1 percent.\(^{27}\) The second point of interest from Table 5 is that the labor supply elasticities for undocumented workers are lower than that estimated for documented workers. Of course, the implication is that undocumented workers are less sensitive to wage changes than documented workers, which is what would be expected if undocumented workers are more restricted in their employment or grievance opportunities, giving employers more monopsony power over the terms of their employment.

This estimated lower labor supply elasticity among undocumented workers, and the ability of employers to identify and take advantage of the lower supply elasticity, likely explains why they are observed to be receiving lower pay. A competing hypothesis for observing a pay gap between two groups of workers is discrimination. A main prediction from the discrimination literature is that a larger supply of the disadvantaged group leads to a larger pay differential between the advantaged and the disadvantaged (see Becker 1971). The means in Table 3 suggest that this has occurred. However, the parameter estimates in Table 4 (and in the Appendix),

\(^{27}\) Another reason the labor supply estimate might be relatively high is that the data do not allow us to control for individual demographic characteristics, such as gender or age. Other's estimates of cross-sectional labor supply elasticities include 0.06 for men and 0.14 for women (Costa 1998), and 0.24 for real estate brokers (Benjamin et al. 2007). However, Ransom and Oaxaca (2007), in their single-firm study, estimate an even larger elasticity than reported here for both men and women which is close to 2.0.
which control for other characteristics of the industries, suggest something else; for a given increase in the share of undocumented workers, the documented worker wage is depressed by a greater amount than the undocumented worker wage, suggesting a mechanism other than discrimination at work in determining the pay gap.

C. Worker Displacement

To the extent that the arrival of undocumented workers depresses wages in a labor market or results in employers substituting documented workers with undocumented workers, an outflow of documented workers is expected. This potential outflow could not only affect estimates of the wage impact, but could also have considerable social welfare impacts if documented workers were flowing into unemployment (rather than to merely another job).

In order to investigate the impact of undocumented worker inflow on separation behavior, a measure of the share of new hires that is undocumented is included in equation (4). The percent of new hires in industry $j$ at time $t$ that are undocumented is calculated as:

$$h_{jt} = 100 \times \frac{H_{jt}^u}{(H_{jt}^u + H_{jt}^d)},$$ (6)

where $H_k^j$ is the number of undocumented ($k=u$) and documented ($k=d$) workers hired by the industry during the previous four quarters. The separation elasticity with respect to the share of new hires that are undocumented is calculated as:

$$\epsilon_{sh}^k = \frac{\partial s}{\partial h} = \hat{\gamma}_2^k \frac{\varphi(I_i^k)}{s} = \hat{\gamma}_1^k \left[ \frac{\varphi(I_i^k)}{\Phi(I_i^k)} \right],$$ (7)

The relevant parameter estimates and resulting separation elasticities are also reported in Table 5 (see the Appendix for estimates across other broad industries).

For all industries combined, a one percent increase in the share of new hires that are undocumented results in a 0.03 percent decrease in the separation probability of documented
workers, and a 0.24 percent increase in the separation probability of undocumented workers. These results are consistent with the positive (but only marginally significant) impact found by Card and DiNardo (2000) of immigration on population growth rates of natives in the same skill group. They suggest that local labor markets adjust to the arrival of immigrants through changes in skill composition and structure, rather than through population adjustments, or a flight of natives from the market. The idea is that the arrival of a less expensive version of one factor input produces a scale effect from which the demand for all input factors (e.g., documented workers) increases. The substitution effect appears to dominate with regard to the separation of undocumented workers, however, as they are (typically) more likely to separate with the arrival of new undocumented workers. Ong and Mar (2007) also find that newly arriving immigrants have a complementary effect on native hours of work and Ottaviano and Peri (2006) and Lalonde and Topel (1991) find that the arrival of immigrants negatively impacts the labor market outcomes of earlier arrivals more than those of natives.

IV. Conclusions and Implications for Policy

The analysis in this paper determined that documented workers in industries that employ undocumented workers earn less than in industries not employing undocumented workers. In addition, among those industries that employ undocumented workers, a higher proportion of undocumented workers statistically significantly reduces the wages of documented workers among all industries overall and within most broadly defined industries. In addition, the impact is expected to be larger than that of immigrants as a whole. It also appears as though the arrival of undocumented workers only displaces earlier arriving undocumented, and that much of the impact on documented worker wages is likely related to the lower labor supply elasticity.
estimated for undocumented workers. Given the limited employment and grievance opportunities of undocumented workers, employers likely enjoy some monopsony wage-setting power, which is expected to put extra downward pressure on wages in labor markets that employ undocumented workers.

If the policy goal is to forestall the negative impact of a growing number of undocumented workers, the results in this paper suggest three possible policy options. One option would be to completely stem the inflow of undocumented immigrants. The success of this effort seems improbable. Another option would be to remove employers' monopsony power by changing the legal status of undocumented workers. Extending to all immigrants the same employee rights afforded documented workers would eliminate a primary source of the measured downward wage pressure in industries that hire undocumented workers. This policy is not likely to garner wide support, however.

A third policy option, which would also have the effect of undermining employer monopsony power, would be to create a permeable border (assuming most of the concern is the flow of undocumented workers from Mexico). The flow of workers would be dictated by demand by employers in the U.S. Workers would be legitimized by the U.S. government and, therefore, would be able to seek redress for grievances, severely limiting an employer's monopsony power. Facilitating an employer's ability to draw workers from a larger pool when needed, would also likely have to be accompanied by strictly enforced penalties for hiring undocumented workers.

Of course, policy makers may have other goals in mind, such as ensuring a sufficient supply of labor to accommodate a desired economic growth rate, or low-prices for consumption goods. If this is the case, the implications for immigration policy would look very different.
References


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


Figure 1. Percent of workers with invalid SSNs for any reason by broad NAICS classification, 1994:1 - 2006:4
Figure 2. Percent of workers with invalid SSN by reason, construction, 1994:1 - 2006:4
Figure 3. Percent of workers with invalid area numbers by broad industry, 1990:1 - 2006:4
Table 1. Estimates of the U.S. unauthorized immigrant population 1990-2007 (millions).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>3.500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.500</td>
</tr>
<tr>
<td>1991</td>
<td>4.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>4.204</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>4.492</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>4.750</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>5.146</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>5.581</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>5.862</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>6.098</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>6.488</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>7.000</td>
<td>8.500</td>
<td>10.242</td>
<td></td>
<td>8.380</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.751</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.300</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.300</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.500</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.600</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.400</td>
</tr>
</tbody>
</table>

Note: Estimates are all made using the residual method as described in the text.
Table 2. Reasons for classifying a SSN as invalid, sample means for Georgia 1990-2006.

<table>
<thead>
<tr>
<th>Invalid Reason</th>
<th>Percent of Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area Number = 000 or Group Number = 00 or Serial Number = 0000, or not enough digits</td>
<td>0.32%</td>
</tr>
<tr>
<td>Area Number invalid</td>
<td>0.10%</td>
</tr>
<tr>
<td>Pocketbook SSN</td>
<td>0.000013%</td>
</tr>
<tr>
<td>SSN equal employer ID</td>
<td>0.30%</td>
</tr>
<tr>
<td>Duplicate SSN within firm</td>
<td>1.53%</td>
</tr>
</tbody>
</table>

Note: Total number of workers (number of person quarters between 1990:1 and 2006:4) is 277,183,148. These reasons are not mutually exclusive; one can have an invalid SSN for multiple reasons.

Table 3. Industry average sample means.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Industries</th>
<th>Construction Only</th>
<th>Leisure &amp; Hospitality Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w^d$</td>
<td>$7829$ (3588)</td>
<td>$6978$ (2997)</td>
<td>$8608$ (3906)</td>
</tr>
<tr>
<td>$Udum = 1$</td>
<td>0.43 (0.50)</td>
<td>0.37 (0.48)</td>
<td>0.54 (0.50)</td>
</tr>
<tr>
<td>$p$</td>
<td>0.09% (0.33)</td>
<td>0.04% (0.17)</td>
<td>0.22% (0.61)</td>
</tr>
<tr>
<td>No. of Firms</td>
<td>143 (392)</td>
<td>121 (319)</td>
<td>169 (516)</td>
</tr>
<tr>
<td>No. of Workers</td>
<td>3588 (11178)</td>
<td>2897 (7732)</td>
<td>3985 (13010)</td>
</tr>
<tr>
<td>Share of firms born</td>
<td>0.03 (0.13)</td>
<td>0.01 (0.04)</td>
<td>0.01 (0.05)</td>
</tr>
<tr>
<td>Share of firms died</td>
<td>0.01 (0.04)</td>
<td>0.01 (0.04)</td>
<td>0.01 (0.03)</td>
</tr>
</tbody>
</table>

Conditional on $p > 0$

| $w^d$                     | $7222$ (3315)  | $6651$ (2997)     | $7800$ (3352)             | $6433$ (1541)  | $5467$ (1303)     | $7283$ (1470)             | $3566$ (1803)  | $3369$ (2132)     | $3543$ (1185)             |
| $w^w$                     | $5511$ (7176)  | $6984$ (10033)    | $4928$ (4335)             | $4910$ (3603)  | $4559$ (4728)     | $5650$ (1312)             | $2979$ (4591)  | $3299$ (2128)     | $2716$ (1506)             |
| $p$                       | 0.22% (0.49)   | 0.12% (0.27)      | 0.41% (0.78)              | 0.48% (0.71)   | 0.12% (0.10)      | 1.17% (0.87)              | 0.23% (0.53)   | 0.15% (0.24)      | 0.27% (0.20)              |
| Wage penalty              | -23.2% (83.9)  | 0.42% (128.1)     | -32.5% (48.0)             | -20.9% (75.5)  | -13.4% (106.2)    | -20.2% (21.9)             | -15.9% (76.0)  | 11.35% (81.8)     | -20.5% (36.8)             |

Notes: Wages are real quarterly earnings, deflated by the chained price index for personal consumption expenditure $2006Q4. Means are averages across industries. Shares in this table do not match those in Figure 3 since means in this table correspond to averages across 6-digit industry shares. Standard errors are in parentheses.
Table 4. Industry fixed-effects regression results with first-order autocorrelation, impact on log wages.

(a) Specification with dummy variable indicating presence of undocumented workers in industry.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Industries</th>
<th>Construction</th>
<th>Leisure &amp; Hospitality</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Ud_{t-1} )</td>
<td>-0.015***</td>
<td>-0.062***</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.013)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>( N_{firms_{t-1}}/10^4 )</td>
<td>0.0002</td>
<td>-1.838***</td>
<td>-1.131***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.145)</td>
<td>(0.285)</td>
</tr>
<tr>
<td>( N_{firms_{t-1}}/10^6 )</td>
<td>8.6x10-6 (2.6x10-5)</td>
<td>0.004***</td>
<td>0.001***</td>
</tr>
<tr>
<td>( N_{wrkrs_{t-1}}/10^3 )</td>
<td>-0.048***</td>
<td>-0.605***</td>
<td>-0.106**</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.090)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>( N_{wrkrs_{t-1}}/10^4 )</td>
<td>1.5x10-5*** (3.4x10-6)</td>
<td>0.001***</td>
<td>3.3x10-5*</td>
</tr>
<tr>
<td>( Y.Q )</td>
<td>0.052**</td>
<td>0.088***</td>
<td>0.080***</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td>(0.0006)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>( Born_t )</td>
<td>-0.169***</td>
<td>-0.315***</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.120)</td>
<td>(0.117)</td>
</tr>
<tr>
<td>( Died_t )</td>
<td>0.018</td>
<td>0.846***</td>
<td>0.308***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.121)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>Intercept</td>
<td>3.73</td>
<td>0.9085***</td>
<td>0.792***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.031)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Within R²</td>
<td>0.42</td>
<td>0.91</td>
<td>0.94</td>
</tr>
<tr>
<td>No. of Observations</td>
<td>57,561</td>
<td>2,444</td>
<td>1,089</td>
</tr>
</tbody>
</table>

(b) Specification with share of undocumented workers in industry, conditional on \( p_{t-1}>0 \).

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Industries</th>
<th>Construction</th>
<th>Leisure &amp; Hospitality</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_{t-1} )</td>
<td>-0.064***</td>
<td>-0.024</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.020)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>( N_{firms_{t-1}}/10^4 )</td>
<td>-0.125***</td>
<td>-0.385**</td>
<td>-1.150***</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.175)</td>
<td>(0.596)</td>
</tr>
<tr>
<td>( N_{firms_{t-1}}/10^6 )</td>
<td>-7.3x10-5* (4.3x10-5)</td>
<td>-1.090***</td>
<td>-1.501***</td>
</tr>
<tr>
<td>( N_{wrkrs_{t-1}}/10^3 )</td>
<td>-0.178***</td>
<td>-0.030</td>
<td>-0.560***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.053)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>( N_{wrkrs_{t-1}}/10^4 )</td>
<td>5.0x10-5*** (4.9x10-6)</td>
<td>0.00002</td>
<td>0.00008</td>
</tr>
<tr>
<td>( Y.Q )</td>
<td>0.086***</td>
<td>0.073***</td>
<td>0.087***</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0007)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>( Born_t )</td>
<td>-0.246***</td>
<td>-0.340*</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.206)</td>
<td>(0.566)</td>
</tr>
<tr>
<td>( Died_t )</td>
<td>0.214***</td>
<td>0.379**</td>
<td>0.888*</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.151)</td>
<td>(0.505)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.446***</td>
<td>1.156***</td>
<td>1.624***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.029)</td>
<td>(0.110)</td>
</tr>
<tr>
<td>Within R²</td>
<td>0.95</td>
<td>0.39</td>
<td>0.92</td>
</tr>
<tr>
<td>No. of Observations</td>
<td>18,027</td>
<td>18,027</td>
<td>2,238</td>
</tr>
</tbody>
</table>

- 41 -
Notes: All regressions include a quarter-year trend regressor, a 6-digit NAIC industry fixed effect, and a first-order autoregressive error term. Standard errors are in parentheses. $Y_Q$ is a continuous variable constructed as $Y \in [90,105]$ and quarter $Q \in [00,75]$. *** $\Rightarrow$ statistical significance at the 99 percent confidence level; ** $\Rightarrow$ statistical significance at the 95 percent confidence level * $\Rightarrow$ statistical significance at the 90 percent confidence level.
Table 5. Maximum likelihood probit estimates of separation equation.

<table>
<thead>
<tr>
<th></th>
<th>All Industries</th>
<th>Construction</th>
<th>Leisure &amp; Hospitality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Documented</td>
<td>Undocumented</td>
<td>Documented</td>
</tr>
<tr>
<td>ln((\omega_{it}))</td>
<td>-0.467***</td>
<td>-0.337***</td>
<td>-0.522***</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.005)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>(h_{it})</td>
<td>-0.209***</td>
<td>0.915***</td>
<td>-0.085***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.055)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>(\phi(\hat{I}))</td>
<td>0.2094</td>
<td>0.3445</td>
<td>0.2548</td>
</tr>
<tr>
<td></td>
<td>(0.1000)</td>
<td>(0.0536)</td>
<td>(0.0846)</td>
</tr>
<tr>
<td>(\Phi(\hat{I}))</td>
<td>0.1761</td>
<td>0.4175</td>
<td>0.2270</td>
</tr>
<tr>
<td></td>
<td>(0.1703)</td>
<td>(0.1859)</td>
<td>(0.1739)</td>
</tr>
<tr>
<td>(\varepsilon_{sw})</td>
<td>-0.56</td>
<td>-0.28</td>
<td>-0.59</td>
</tr>
<tr>
<td>(\varepsilon_{nw})</td>
<td>1.12</td>
<td>0.56</td>
<td>1.18</td>
</tr>
<tr>
<td>(\varepsilon_{sh})</td>
<td>-0.03</td>
<td>0.24</td>
<td>-0.03</td>
</tr>
<tr>
<td>N</td>
<td>65,348,517</td>
<td>38,609</td>
<td>3,760,898</td>
</tr>
</tbody>
</table>

Notes: Analysis includes workers employed in Georgia 1994-2000 inclusive. Additional regressors include the number of firms and total employment in a worker's industry, the industry average proportion of undocumented workers, quarterly dummy variables, a measure of annual gross state product, and 3-digit NAIC industry fixed effect. Samples are too large to include 6-digit fixed effects. Only broad industry fixed effects are included in the estimation for all industries because of technical limitations resulting from the size of the documented sample. Based on a standard Z-test for differences in means, all estimates between documented and undocumented workers are significantly different from one another at the 99 percent confidence level. *** ⇒ statistical significance at the 99 percent confidence level; ** ⇒ statistical significance at the 95 percent confidence level; * ⇒ statistical significance at the 90 percent confidence level.
Appendix: Additional results.

Table A1. Estimates of coefficient on $p_{t-1}$ within broad industry classifications.

<table>
<thead>
<tr>
<th>Industry Classification</th>
<th>$\ln w_t$</th>
<th>$\ln w_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Sample</td>
<td>-0.064***</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Construction 3-digit NAICS = 236-238</td>
<td>-0.259***</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Leisure and Hospitality 3-digit NAICS = 711-722</td>
<td>-0.044**</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>Manufacturing 3-digit NAICS = 311-339</td>
<td>-0.174***</td>
<td>-0.115</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.074)</td>
</tr>
<tr>
<td>Professional &amp; Bus Srvcs 3-digit NAICS = 541-562</td>
<td>-0.334***</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.182)</td>
</tr>
<tr>
<td>Retail Trade 3-digit NAICS = 441-454</td>
<td>-0.005</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.106)</td>
</tr>
<tr>
<td>Wholesale Trade 3-digit NAICS = 423-425</td>
<td>-0.042</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.204)</td>
</tr>
</tbody>
</table>

*** ⇒ statistical significance at the 99 percent confidence level; ** ⇒ statistical significance at the 95 percent confidence level * ⇒ statistical significance at the 90 percent confidence level.

Table A2. Labor supply and new hire separation elasticities within broad industry classifications.

<table>
<thead>
<tr>
<th>Industry Classification</th>
<th>$\varepsilon_{nw}$</th>
<th>$\varepsilon_{sh}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Documented</td>
<td>Undocumented</td>
</tr>
<tr>
<td>Full Sample</td>
<td>1.12***</td>
<td>0.56***</td>
</tr>
<tr>
<td>Construction 3-digit NAICS = 236-238</td>
<td>1.18***</td>
<td>0.80***</td>
</tr>
<tr>
<td>Leisure and Hospitality 3-digit NAICS = 711-722</td>
<td>0.86***</td>
<td>0.68***</td>
</tr>
<tr>
<td>Manufacturing 3-digit NAICS = 311-339</td>
<td>1.72***</td>
<td>0.28***</td>
</tr>
<tr>
<td>Professional &amp; Bus Srvcs 3-digit NAICS = 541-562</td>
<td>0.81***</td>
<td>0.43***</td>
</tr>
<tr>
<td>Retail Trade 3-digit NAICS = 441-454</td>
<td>1.15***</td>
<td>0.53***</td>
</tr>
<tr>
<td>Wholesale Trade 3-digit NAICS = 423-425</td>
<td>1.38***</td>
<td>0.62***</td>
</tr>
</tbody>
</table>

Notes: Significance inferred from statistical significance of parameter estimates used to construct elasticities. *** ⇒ statistical significance at the 99 percent confidence level; ** ⇒ statistical significance at the 95 percent confidence level * ⇒ statistical significance at the 90 percent confidence level.