

Unauthorized Immigrants in California

Estimates for Counties

July 2011

Laura E. Hill and Hans P. Johnson

with research support from David Ezekiel and Joseph M. Hayes

Supported with funding from the Silicon Valley Community Foundation

Summary

California has more unauthorized immigrants than any other state, about 2.6 million of the nation's 11 million; they make up 7 percent of the total California population and 9 percent of the state's labor force. For decades, unauthorized immigrants have been a part of California: in many industries in the economy and in rural and urban communities.

But recent and comprehensive information about the numbers and location of this population *within* California—at the county and sub-county level—does not exist. That this wide information gap exists is doubly surprising given the amount of energy spent and attention paid to this issue by policymakers and the public over those same decades.

This report is the first to use a new source of administrative data at the local level to produce comprehensive and systematic sub-state estimates of the unauthorized immigrant population in California.

We find that unauthorized immigrants live in every county in the state, primarily but not only in highly agricultural or highly urban areas. As in the country as a whole, unauthorized workers here reside not just in traditional immigrant communities, but have found homes throughout all regions of the state.

Contents

Summary	2
Tables	4
Figures	5
Introduction	6
Counting California’s Unauthorized Immigrants	7
Obstacles to Counting the Unauthorized	8
Combining New Administrative Data and Residual Method Data	9
Methodology	12
Where in California Do Unauthorized Immigrants Live?	16
Unauthorized Immigrant Population Time Trends	17
Unauthorized Immigrants Live in Zip Codes throughout the State	21
Conclusion	28
References	29
About the Authors	30
Acknowledgments	30

A technical appendix to this paper is available on the PPIC website:
http://www.ppic.org/content/pubs/other/711LHR_appendix.pdf

Tables

Table 1	Estimates of unauthorized immigrants in California and the United States (millions)	9
Table 2	Stepwise regression estimates	14
Table 3	Estimates of California county unauthorized immigrant populations (2008)	16
Table 4	Estimates of county unauthorized immigrant populations: preferred model; implied by ITIN tax filings; implied by new noncitizens	19

Figures

Figure 1	Unauthorized immigrants in California and all other states, selected years	7
Figure 2	Tax returns filed with ITINs	11
Figure 3	Correlations between ITIN tax filers and unauthorized immigrant estimates, all states	12
Figure 4a	Estimates of unauthorized immigrants in California, by zip code	22
Figure 4b	Estimates of unauthorized immigrants in California, percent of population, by zip code	23
Figure 5a	Estimates of unauthorized immigrants, Los Angeles County zip codes	24
Figure 5b	Estimates of unauthorized immigrants, percent of population, Los Angeles County zip codes	25
Figure 6a	Estimates of unauthorized immigrants, San Francisco Bay Area zip codes	26
Figure 6b	Estimates of unauthorized immigrants, percent of population, San Francisco Bay Area zip codes	27

Introduction

The most basic measures of any population group, its size and location, remain elusive when it comes to unauthorized immigrants in California. Although information about the size of the unauthorized immigrant population in the state—about 2.6 million in 2010¹— is easy to find, no current, comprehensive estimates of this population exist for smaller areas such as counties or cities. Nor are they available for any other state. This information gap creates a problem for local, state, and even federal authorities as they try to evaluate this population and create policies affecting them. It is especially problematic in California, where the unauthorized immigrant population is so large and its location so geographically diverse.

In this work, we seek to close this knowledge gap at the sub-state level for California.

Producing sub-state estimates of unauthorized immigrants is challenging because immigrants cannot be counted directly. In this work, we take a unique set of administrative data, in this case IRS tax return data for unauthorized immigrants, and then model the data's relationship to estimates of unauthorized immigrants for the states using regression analysis. Because the IRS tax data is available at the zip code level, we can use the observed relationship between IRS tax data and state populations of unauthorized immigrants to estimate unauthorized immigrant populations for counties and sub-county areas.

Our methodology is one common for demographers estimating population growth; both the Census Bureau and the California Department of Finance, for example, use administrative data on births and housing to estimate size and change for large population sets such as age and ethnic groups (although not unauthorized immigrants). A further advantage to our method is that IRS data are released annually. Thus, estimates can be both updated and replicated across other states that have large unauthorized immigrant populations.

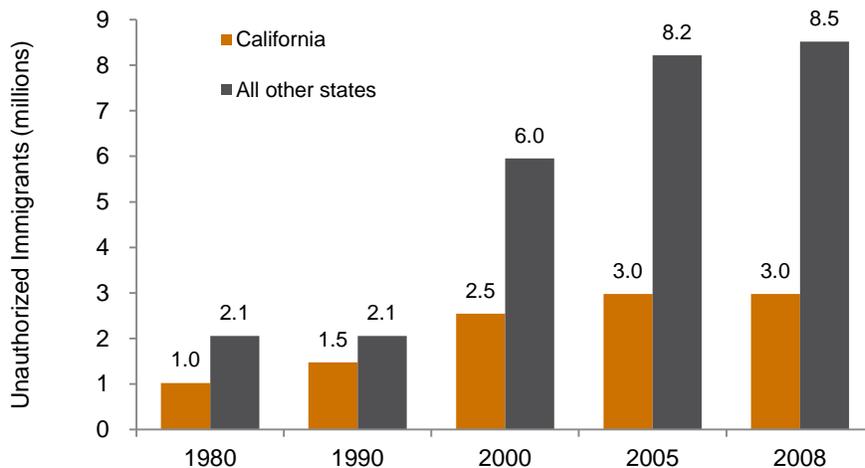
We first explain our data and methodology in detail, then present our results.

¹ Passel and Cohn (2010); Hoefler, Rytina, and Baker (2010).

Counting California's Unauthorized Immigrants

The Pew Hispanic Center (PHC), the Department of Homeland Security (DHS), and Warren (2011) produce careful estimates of the size of the unauthorized immigrant population at the state and national level. Because there are no national or state level surveys that obtain the documentation status of the foreign-born residing in the United States, these three sources provide our best *indirect* estimates of the number and distribution of unauthorized immigrants. This population is in the midst of a major shift. After many years of increases, the number of California's unauthorized immigrants has remained stable or even declined slightly recently.² At the same time, the number living in other states has increased substantially compared to California (Figure 1). In 1980, approximately half the nation's unauthorized immigrants lived in the state, but that share had fallen substantially, to about 26 percent, by 2008.

FIGURE 1
Unauthorized immigrants in California and all other states, selected years



SOURCES: Passel and Woodward (1984); Warren (2011).

NOTE: All years except 1980 are from Warren (2011).

There are other signs of change in national settlement patterns: states with the highest rates of growth of unauthorized immigrants in recent years are not the traditional ones: Mississippi, Alabama, and South Carolina (Warren 2011). Further, the greatest numerical gains in recent years have been in Texas, Florida, and North Carolina. More recently, estimates suggest some states are losing unauthorized immigrant populations—including California, which had 250,000 fewer in 2009 than in 2008, according to DHS

² Passel and Cohn (2011); Warren (2011); Department of Homeland Security (2011).

estimates and 100,000 fewer in 2009 than in 2007, according to the PHC. (The PHC estimated difference is not statistically significant.) Within California, settlement patterns may be changing as well but until now, there was no way of gauging these.

Obstacles to Counting the Unauthorized

Because the Census and national population surveys place their primary focus on generating full participation, they do not ask foreign-born participants to reveal their immigration status for fear that they will not participate. The Census, the American Community Survey (ACS), and the Current Population Survey (CPS) do ask respondents who identify themselves as foreign-born to state their country of birth, their date of arrival, and whether or not they are naturalized U.S. citizens. For those not naturalized, however, no further survey questions provide the detail necessary to determine if they are legally resident (either permanently or temporarily) or unauthorized.

A few other surveys come closer to providing counts of the unauthorized than do the census, CPS, or ACS, but none can do the job completely. The National Agricultural Worker Survey asks foreign-born participants to state their official immigration status, but their respondents are all agricultural workers, and therefore not representative of the full population of unauthorized immigrants. (Although more than half of all agricultural workers are estimated to be unauthorized, only 4 percent of the nation's total unauthorized population is employed in agriculture, Passel (2009) finds). The California Health Interview Survey asks respondents to indicate if they are citizens or legal permanent residents, but does not differentiate between temporary visa holders and unauthorized immigrants. The New Immigrant Survey provides retrospective information on prior legal status of immigrants who eventually gain legal permanent residency, but the sample size is too small to pinpoint locations and these unauthorized immigrants may not be representative of all unauthorized immigrants in all ways.³

Thus, estimates of unauthorized immigrants at the national and state level are produced indirectly, using what is commonly referred to as a residual technique. We rely on three sources for California data on the unauthorized that use this technique—PHC, DHS, and Warren (2011). Each is computed using slightly different data and slightly different variants on a residual method approach. To count the foreign-born population, the PHC estimates use Current Population Survey (CPS) data while the DHS and Warren estimates use the American Community Survey (ACS). Next, each subtracts estimates of the legal foreign-born residents from the counts of the foreign-born in the CPS or ACS. The remaining, or residual, foreign-born comprise the estimates of unauthorized immigrants. The three estimates vary somewhat in how they determine the legally resident population of the foreign-born, but all three combine administrative counts of legal admissions with standard demographic techniques.⁴ (For more detailed descriptions, see Passel and Cohn 2010, Passel 2007, Hoefler, Rytina, and Baker 2011, and Warren 2011.) The estimates computed by these three are generally in close agreement.

³ Hill, Lofstrom, and Hayes (2010) demonstrated many similarities.

⁴ Warren does not estimate the legal population fully each year. Instead, he estimates legal foreign-born arrivals each year and uses demographic techniques to estimate changes in the legal population from one year to the next.

TABLE 1
Estimates of unauthorized immigrants in California and the United States (millions)

	2000	2005	2006	2007	2008	2009	2010
California							
Warren	2.6	2.9	2.9	2.9	2.9		
PHC	2.3	2.7		2.8	2.7	2.6	2.6
DHS	2.5	2.8	2.8	2.8	2.9	2.6	2.6
Total U.S.							
Warren	8.6	10.9	11.1	11.2	11.2		
PHC	8.4	11.1		12.0	11.6	11.1	11.2
DHS	8.5	10.5	11.3	11.8	11.6	10.8	10.8

SOURCES: Warren (2011); Passel and Cohn (2010, 2011); DHS (2010, 2011).

In California, credible efforts to estimate some sub-state unauthorized populations have been undertaken, but none are as current as our data and none derive from annually updated, independent, administrative data. Pastor and Marcelli in 2004 published estimates of Mexican unauthorized immigrants in California Public Use Microdata Areas (PUMAs) for 1990 and 2000.⁵ These were based on two surveys of Mexican immigrants in Los Angeles County, with the results used to determine characteristics of Mexican unauthorized immigrants. These characteristics were in turn used to assign probabilities of unauthorized status among Mexican immigrants in the population data. Heer and Passel (1987) demonstrated that estimates similar in type to those of Pastor and Marcelli (based on survey data that are applied to population data) matched well with estimates derived by interpolating from estimates based on the residual method.

The Census Bureau did estimate counts of unauthorized immigrants for California counties in the 1980s, but to our knowledge, these estimates were not published, with the exception of Los Angeles County (Heer and Passel 1987), and were later discontinued. More recently, Fortuny, Capps, and Passel (2007) published estimates for five metropolitan areas in California (and 25 such areas nationwide) for the years 2000 and 2003–2004. In addition, Paral and Associates publishes on its web page estimates of unauthorized immigrants by U.S. congressional districts for 2000 and 2005. They apportion the Passel estimates to congressional districts based on demographic correlates.

Combining New Administrative Data and Residual Method Data

As described above, residual method estimates for unauthorized immigrants at the state level are in wide agreement and are generally believed to be the best source of information about unauthorized immigrants in the United States. We use these reliable estimates as the basis from which to derive new estimates for regions, counties, and smaller geographic areas within California, combining them with new administrative data to do so. Because the administrative data we use are not in wide use (despite having been collected since 1996), and have never before been used for this purpose, we devote this section to explaining them and their usefulness for our estimates.

⁵ PUMAs are larger than zip codes and census tracts. In sparsely populated regions, PUMAs can span many counties, but within dense counties, such as Los Angeles, there may be several.

Since 1996, unauthorized immigrants have been permitted to file federal tax returns using a unique identifier, the Individual Taxpayer Identification Number, or ITIN. Immigrants and the native-born with the right to work use social security numbers (SSN) when filing tax returns, but those without work authorization do not have valid SSNs and should use ITINs instead.⁶ The IRS has made counts of ITIN filers by zip code publicly available for tax years 2000–2007 (which correspond to calendar years 2001–2008). As we show, these counts of ITIN filers are a basis from which to estimate unauthorized immigrants in counties and sub-count areas in California.

Even if they have worked in the United States without proper authorization, unauthorized immigrants are nevertheless required under federal law to file tax returns. Some estimates suggest that about half do so (Immigration Policy Center 2011; Pastor et al. 2010; Hinojosa-Ojeda 2010), but others show the share is much higher: a recent PPIC report found that over 80 percent of unauthorized immigrants reported having filed federal income taxes in the year prior to earning legal permanent residence (Hill, Lofstrom, and Hayes 2010). The Social Security Administration’s chief actuary estimated that about 75 percent of unauthorized immigrants have payroll taxes withheld (Porter 2005).

Unauthorized immigrants have many incentives to file tax returns. First, some who have had taxes withheld by their employers throughout the year would be eligible for tax refunds and might use an ITIN to claim that money. Second, if an unauthorized immigrant ultimately does become a legal permanent resident and receives an SSN, he or she can link any social security earnings withheld under the ITIN to earnings under the new SSN and have them counted toward later benefits. Third, even unauthorized immigrants who have not had taxes withheld (being self-employed or paid in cash, for example) might file tax returns to establish a positive paper trail for the future: should comprehensive immigration reform ever become a reality, a clear record of employment and tax payments are factors likely to increase an unauthorized immigrant’s chances of attaining legal status.

The majority of ITIN users are unauthorized immigrants, as we explain below. If a former ITIN filer is ever granted the legal ability to work, he or she should begin using his or her assigned SSN immediately. Further, anyone who is authorized to work in the United States, such as those on temporary work visas (e.g. H-1B or foreign students with work authorizations) are required to apply for and file federal taxes with an SSN.

By December 2008, the IRS had issued more than 14 million ITINs. Not all of these ITINs are used on tax returns filed domestically; some are filed from abroad. Many may have been retired after an immigrant was issued a valid SSN. Some may never have been used to file taxes (they may have been obtained to open a bank account, for example). Some may no longer be in use because the ITIN holder no longer files taxes, either by choice or because he or she no longer is required to do so (having insufficient income or no longer living in United States).

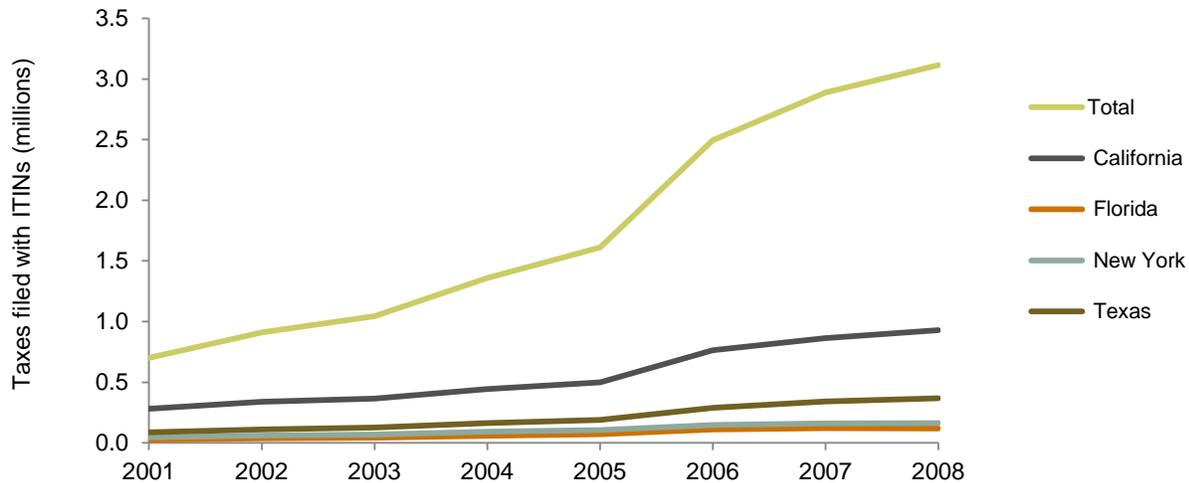
The number of tax returns filed with ITINs and that use U.S. addresses has increased dramatically since 1996.⁷ The years for which we have data—2001 to 2008—reveal that many tax filers in California were early users of the ITIN, making up 40 percent of ITIN filers nationally in 2001, although this fell to 30 percent of filers in by 2008 (Figure 2). Other states have seen similar but smaller changes in their share of ITIN tax filers.

⁶ ITINs are not a valid proof of employment eligibility; ITIN tax filers could not receive federal stimulus tax rebates, nor are they eligible for the Earned Income Tax Credit (EITC), which requires a filer, spouse, and child to have valid SSNs.

⁷ Anomalous increases in ITIN filings were investigated and in some cases, those records were deleted from the totals reported here. Texas and Georgia both had single-year increases in a few zip codes that were larger than the entire zip code population, so we kept the growth in ITIN numbers for those four zip codes at the level recorded in the previous year.

For example, Illinois had 4.7 percent of all filers in 2001 and 6.1 percent in 2008. Like California, ITIN tax filers in New York also declined in share of ITIN filers (from 6.5 percent to 5.2 percent). North Carolina’s share rose (from 1.8 percent to 3.7 percent over the same period). These changes are largely consistent with the residual estimates of unauthorized immigrants in these states. For example, California’s share of the nation’s unauthorized population fell from 43 percent in 2000 to 35 percent in 2008, according to the Warren estimates.

FIGURE 2
Tax returns filed with ITINs



SOURCE: Authors’ calculations of ITIN numbers.

NOTE: Beginning in 2006, dependents with ITINs were also recorded.

In California, nearly 6 percent of tax filers used an ITIN in 2008 (929,000), up from 2 percent in 2001. These include primary filers, spouses, or dependents. Although not all ITIN filers were unauthorized, the number of ITIN filers was 36 percent of the number of estimated unauthorized immigrants.

Although ITIN tax filing data may serve as a good proxy for unauthorized immigrants, they cannot provide a precise count, for two reasons. First, not all unauthorized immigrants file taxes, and among those that do, not all use the ITIN. Some may file instead using a false, fraudulent SSN, or an SSN issued many years ago that did not permit work⁸ or one that no longer does. Second, not all ITIN filers are unauthorized immigrants. However, once we exclude tax filers from abroad, the vast majority of ITIN filers do appear to be unauthorized. When we examine ITIN tax filers with U.S. filing addresses, we find that 90 percent in 2008 include wages. Only unauthorized workers would file tax returns with wages and an ITIN; authorized workers with wages would file tax returns using SSNs. Therefore, it is safe to conclude that the vast majority of ITIN filers are unauthorized workers.

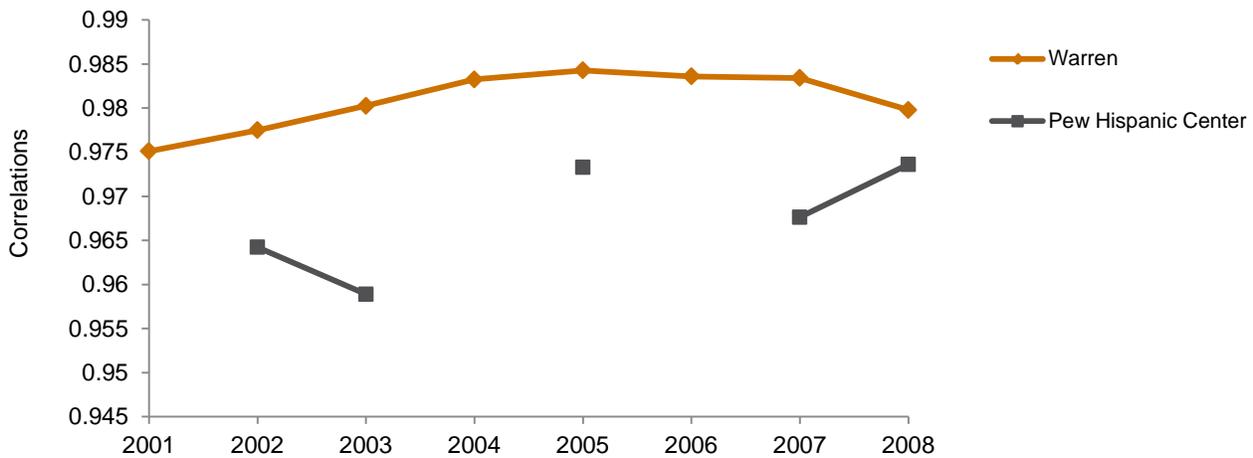
⁸ These “no work” SSNs are no longer issued.

Methodology

Using ITIN data to estimate the distribution of unauthorized immigrants within California (or any state) depends on a strong relationship between the ITIN numbers and existing estimates of unauthorized immigrants. ITIN numbers will not include all unauthorized immigrants, nor will they include only unauthorized immigrants. But to use ITINs as a basis from which to scale local area estimates to match state total estimates we need the correlation between ITIN counts and unauthorized estimates to be high.

And indeed, we find ITINs are highly correlated with independently derived state estimates of unauthorized immigrants. We calculated correlations for each year of available ITIN data with the Warren and PHC estimates (Figure 3). Each point in the figure represents that year's correlations for the 50 states and the District of Columbia.

FIGURE 3
Correlations between ITIN tax filers and unauthorized immigrant estimates, all states



SOURCES: Authors' calculations from ITIN counts; Warren (2011); Passel and Cohn (2011, 2010).

NOTE: 2008 ITIN correlations for Passel and Cohn use 2009 data; DHS estimates are available only for the 10 states with the largest unauthorized populations.

The Warren estimates have a correlation value of nearly 0.98 with the ITIN tax filing records (see [Appendix Table A1](#) for the full set of Warren estimates). The PHC estimates use ranges for states with small populations, which may partially explain their lower correlation, which is still very high: in excess of 0.955 for all years. In the PHC estimates, those of 32 states are derived using multiple years of data because of the very small CPS data sample of likely unauthorized immigrants in those states—fewer than 50 in each year (Passel and Cohn, 2011). Altogether, these extremely high correlations give us great confidence that ITIN filings are an excellent indicator of the unauthorized immigrant population.

We also examined the relationship between the estimates of unauthorized immigrants and ITIN data filed with wages attached (*W-2s*), ITIN taxing filings prepared by a paid tax preparer, and ITINs not filed as a nonresident. Each of these has the potential to be more highly correlated with unauthorized immigrants than the overall ITIN filing rates. ITIN returns filed with wages are almost definitely unauthorized immigrant tax filers (all other filers with the legal right to work should have SSNs), but not all ITIN unauthorized tax filers

will have wages attached (e.g. those who file but are paid cash or are self-employed). The vast majority of ITIN filers use paid preparers, and do so at much higher rate than tax filers in general. This may be because unauthorized immigrant tax filers may prefer to pay for help in filing correctly if their motivation is to document earnings and tax paying; or it may simply be that unauthorized tax filers who use paid preparers are more likely to learn about ITINs than if they do not.⁹ Unauthorized immigrants should also be more likely to file as residents (rather than using the nonresident 1040NR form) because nonresidents cannot claim child tax credits and because filing as a nonresident might do less to establish one’s intent to ultimately legalize.¹⁰ However, we found none of these other measures of ITIN tax filers were as consistently highly correlated as the correlations with the overall ITIN tax filings for the states.

Given the high correlation between ITIN filers and estimates of the unauthorized for the 50 states and D.C., we could use the simple ratio of unauthorized immigrants to ITIN filers as the factor to estimate local populations of unauthorized immigrants. However, we know this ratio varies across states and time, and so suspect that it might also vary within the state. We use regression analyses to account for this cross-state and intra-state variation. Differences in the rate of ITIN usage by a state’s unauthorized population may be related to variations in the characteristics of employment and earnings among unauthorized immigrants, their demographic characteristics, or infrastructure available to support immigrant tax filing, among others.

Using the Warren estimates, ITIN data, the ACS (for the 2006 and 2008 models), and the 2000 census (for the 2001 model), we use weighted least squares (WLS) to estimate the following regression model of the ITIN coverage rate for each state (*s*) and each year (*t*), with proportionate weight applied to the estimated size of the undocumented population. We restrict the population to foreign-born residents:

$$\left(\frac{ITIN}{Warren\ Estimate} \right)_{ts} = \mathbf{X}_{ts}\alpha + \mathbf{W}_{ts}\beta + \mathbf{Z}_{ts}\gamma + \varepsilon_{ts}$$

In the equation, X represents a matrix of demographic characteristics of the immigrant population in each state and each year. It includes age, proportion Latino, and proportion born in Central America; we do not include the proportion born in Mexico because that is so closely correlated with the proportions that are Latino and Central American-born. Employment characteristics for the immigrant population are represented by the matrix W and include the share employed in construction, the share employed in restaurants, proportion self-employed, and proportion not in the labor force. Tax filing characteristics are represented by the matrix Z and include filing as married, being a new tax filer, and filing using a paid preparer. Because we have only 51 observations (50 states and D.C.), we restricted our possible covariates to just a few (Table 2).

We want these relationships to be able to vary, first because the increase in ITIN usage was so great during the interval we studied (Figure 2), and second, because we do not fully understand why some states seemed to have a higher percentage of unauthorized immigrants filing income taxes with ITINs. We therefore allowed the variables that entered into this equation to change across time. We used a backward elimination stepwise regression method, removing the least significant variable from the model one at a time until all variables met a predetermined threshold of significance. In this case our threshold was a *p*-value of 0.10. Regressions were run separately for each year.

⁹ Filers claiming the EITC also use paid preparers at a very high rate. In California, 76 percent of EITC filers used paid preparers in 2006 (Danielson 2010).

¹⁰ 1040NR tax forms can be filed from U.S. addresses by those who are only temporarily residing in the U.S.

In developing our model, we considered different population bases to use to estimate the ratio of ITIN tax filers to Warren estimates, and settled on the foreign-born as the base population of each state.¹¹ We discarded some of the variables that were determined to be significant based on the state-level models because the range of values for those variables for counties in California was far outside the range of values for states. For example, the share of the foreign-born population working in agriculture ranges from 1 percent to 25 percent in the 50 states and D.C., but in California counties, the high end is much greater: Tulare County has 61 percent of its foreign-born labor force working in agriculture, and many other California counties were above 40 percent.

We considered models in which we predicted the ITIN numbers, with the Warren estimates on the right hand side of the equation. We also predicted the logged ITIN numbers. Both of these models resulted in state-level computed estimates that had a poorer fit than the ratio models—that is, our predictions for states were not as close to the actual Warren estimates as in our final model. Further, we estimated our models using the PHC estimates as robustness check. Because correlations are lower (Figure 3) and the estimates are not available for all years, we prefer the Warren estimates. (The models estimated with the PHC data are available on request.)

Our final model varied for each year. We report those for 2001, 2006, and 2008 below; Table 2 reports all of the variables that we allowed to enter into the stepwise regression. As noted above, only those variables which were estimated with a *p* value of 0.10 were ultimately included in the regression. We have the most confidence in the county estimates for 2008.

TABLE 2
Stepwise regression estimates

	2001		2006		2008	
R squared	0.701		0.593		0.618	
hettest	0.012		0.501		0.692	
	Coeff	p-value	Coeff	p-value	Coeff	p-value
Age 0–17						
Age 35–54	0.215721	0.038			–	
Age 55+	0.186604	0.006				
Proportion Latino						
Born in Central America	-0.1601	0.000	-0.41705	0.000	-0.5378	0.000
Construction						
Restaurants						
Self-employed						
Not in the labor force						
Filed taxes as married	-0.11164	0.000	-0.39599	0.000	-0.48947	0.000
New tax filer						
Filed using paid preparer	0.065578	0.064				
Constant	0.015552	0.718	0.489259	0.000	0.598122	0.000

SOURCES: Authors' calculations using ITIN, ACS, 2000 census, and Warren and Associates data.

¹¹ We also considered models using recently arrived foreign-born noncitizens as the base population. Those models had so much more variation within California counties than across the 50 states and strained the ability of regression to make useful estimates. Expanding the sample more widely to include the entire population resulted in a poorer fitting model as well.

In each of the years 2001, 2006, and 2008, the share of the foreign-born population from Central America and the share of ITIN tax filers that filed tax returns as married was very important in scaling the ITIN returns to match the state-level estimates of unauthorized immigrants. Also important for 2001 was the share of the foreign-born population between ages 35 and 54, and 55 or older, as was the share of ITIN tax filers who used a paid tax preparer.

Applying these coefficient estimates for the states to ITIN tax-filer and ACS data for the counties, we computed a county- (or regional-) level count of unauthorized immigrants.¹² These were then totaled and scaled to match the estimate of unauthorized immigrants for the state in that year. Our final step was to scale these local estimates back down to the zip code level, using the distribution of ITIN filings filed by zip code within that county. (We could not use our model and ACS data in the same way because ACS data are not available for that small geography.)

¹² We totaled zip codes to the county or regional level. Many zip codes span two (or more) counties, and are allocated to counties based on 2000 census block populations (Kneebone 2008).

Where in California Do Unauthorized Immigrants Live?

For 2008, we find unauthorized immigrants residing in all counties (or county groups) throughout the state.¹³ Unauthorized immigrants are found in major urban areas, agricultural regions, and places in between. According to our estimates, their raw numbers range from just over 1,000 in the Del Norte/Siskiyou/Modoc/ Lassen county grouping to just under 1 million in Los Angeles County (Table 3).

In general, unauthorized immigrants make up small but notable shares of county populations. In only four counties or county groupings do they make up more than 10 percent of the total population. In 22 counties, mostly rural and mountainous, but also including Sacramento, unauthorized immigrants make up less than 5 percent of the population. Not surprisingly, the most populous counties have the largest populations of unauthorized immigrants.

TABLE 3
Estimates of California county unauthorized immigrant populations (2008)

County	Total population (2008 ACS estimate)	Unauthorized immigrants	
		Population estimate	% of county total population
Alameda	1,475,000	124,000	8.4%
Amador, Calaveras, Tuolumne, Mariposa, Alpine, Mono, Inyo	191,000	2,500	1.4%
Butte	220,000	4,000	1.8%
Colusa, Glenn, Tehema, Trinity	124,000	10,000	8.3%
Contra Costa	1,029,000	79,000	7.7%
Del Norte, Siskiyou, Modoc, Lassen	118,000	1,000	1.0%
El Dorado	176,000	4,000	2.2%
Fresno	909,000	49,000	5.3%
Humboldt	129,000	2,000	1.6%
Imperial	164,000	21,000	12.8%
Kern	801,000	46,000	5.7%
Kings	150,000	9,000	5.8%
Los Angeles	9,860,000	916,000	9.3%
Madera	149,000	12,000	7.7%
Marin	249,000	14,000	5.6%
Mendocino, Lake	151,000	8,000	5.0%
Merced	246,000	22,000	9.1%
Monterey, San Benito	463,000	62,000	13.5%
Napa	134,000	16,000	12.0%

¹³ Because the ACS is a sample, not all counties have large enough populations to be reported separately in it. Because of these ACS sample size restrictions, we report unauthorized immigrant estimates for 34 counties and 7 county groups rather than 58 counties for both 2001 and 2008.

TABLE 3 (continued)

County	Total population (2008 ACS estimate)	Unauthorized immigrants	
		Population estimate	% of county total population
Orange	3,010,000	289,000	9.6%
Placer	342,000	8,000	2.3%
Plumas, Sierra, Nevada	120,000	2,000	1.5%
Riverside	2,101,000	146,000	7.0%
Sacramento	1,394,000	65,000	4.6%
San Bernardino	2,015,000	150,000	7.5%
San Diego	3,002,000	198,000	6.6%
San Francisco	809,000	30,000	3.7%
San Joaquin	673,000	54,000	8.0%
San Luis Obispo	266,000	9,000	3.5%
San Mateo	712,000	55,000	7.8%
Santa Barbara	405,000	37,000	9.0%
Santa Clara	1,764,000	180,000	10.2%
Santa Cruz	253,000	21,000	8.2%
Shasta	180,000	1,000	0.6%
Solano	407,000	24,000	6.0%
Sonoma	467,000	41,000	8.8%
Stanislaus	511,000	39,000	7.6%
Sutter, Yuba	165,000	9,000	5.6%
Tulare	426,000	29,000	6.8%
Ventura	798,000	74,000	9.3%
Yolo	198,000	12,000	6.2%
Total	36,756,000	2,876,000	7.8%

SOURCES: Authors' calculations; ACS.

Unauthorized Immigrant Population Time Trends

In this section, we present our estimates for 2001 and 2008 together. We expected that our method and models would be more reliable for years when ITIN numbers were more commonly used than the years when ITINs were new and less likely to be used by unauthorized immigrants. However, the fit for our model that predicted the ratio of ITIN filings to the Warren unauthorized estimates was actually slightly better in 2001 than 2008 (Table 2), despite the fact that, as Figure 2 illustrated, there was a dramatic uptick in use of ITINs among tax filers. We also found that from 2001 to 2008 the number of zip codes with ITIN tax filers increased.

Taken individually, the estimates for the single years seem reasonable. Although our 2001 model fits well, we are still cautious about our results from years before 2008 mainly because a smaller share of unauthorized immigrants was filing ITIN returns in the earlier years; when we examine the change from 2001 to 2008, we have less certainty about the prior years. We find that for many of the small county and small county groupings, the growth in unauthorized immigrants that is implied from our estimates is perhaps

implausible. Therefore, the 2001 estimates we present are only for those counties with 2008 populations of 200,000 or greater and those in which our 2001 estimate of unauthorized immigrants was greater than 10,000.

We offer two benchmarks, for comparison only. The first is the simple distribution of the estimated number of unauthorized immigrants using the ITIN counts for counties from the administrative data. Comparing our model estimates to the ITIN results gives a sense of how our model may be an improvement over simply scaling the administrative tax data. The second benchmark is the distribution of the estimate of unauthorized immigrants using the distribution of the state's new noncitizens (arrived within the last 20 years) to counties. This is one way to allocate the reputable state estimates to sub-state areas (but not a method employed by any of those who compute those residual methods). Our method, because its underlying data are available every year, for all zip codes nationwide, and because it does not rely on any other allocation or estimation (with the exception of state-level estimates), is the best methodology available given the current data constraints.

TABLE 4

Estimates of county unauthorized immigrant populations: preferred model; implied by ITIN tax filings; implied by new noncitizens*

County	County population (2008 ACS)	2001			2008			Change 2001 to 2008		
		Model	ITIN filers	New non-citizens	Model	ITIN filers	New non-citizens	Model	ITIN filers	New non-citizens
Alameda	1,475,000	163,000	144,000	121,000	124,000	116,000	116,000	-39,000	-28,000	-5,000
Amador, Calaveras, Tuolumne, Mariposa, Alpine, Mono, Inyo	191,000				2,500	3,000	2,000			
Butte	220,000				4,000	4,000	6,000			
Colusa, Glenn, Tehama, Trinity	124,000				10,000	10,000	6,000			
Contra Costa	1,029,000	63,000	63,000	50,000	79,000	71,000	71,000	16,000	8,000	21,000
Del Norte, Siskiyou, Modoc, Lassen	118,000				1,000	1,000	1,500			
El Dorado	176,000				4,000	4,000	4,000			
Fresno	909,000	30,000	25,000	63,000	49,000	51,000	72,000	19,000	26,000	9,000
Humboldt	129,000				2,000	2,500	1,000			
Imperial	164,000				21,000	18,000	14,000			
Kern	801,000	21,000	18,000	36,000	46,000	46,000	53,000	25,000	28,000	17,000
Kings	150,000				9,000	9,000	12,000			
Los Angeles	9,860,000	924,000	948,000	1,069,000	916,000	894,000	987,000	-8,000	-54,000	-82,000
Madera	149,000				12,000	12,000	12,000			
Marin	249,000	16,000	19,000	12,000	14,000	14,000	12,000	-2,000	-5,000	0
Mendocino, Lake	151,000				8,000	8,000	5,000			
Merced	246,000	15,000	12,000	16,000	22,000	24,000	24,000	7,000	12,000	8,000
Monterey, San Benito	463,000	39,000	37,000	43,000	62,000	73,000	48,000	23,000	36,000	5,000
Napa	134,000				16,000	15,000	12,000			
Orange	3,010,000	349,000	387,000	273,000	289,000	323,000	267,000	-60,000	-64,000	-6,000
Placer	342,000				8,000	7,000	10,000			
Plumas, Sierra, Nevada	120,000				2,000	2,000	500			

TABLE 4 (continued)

County	County population (2008 ACS)	2001			2008			Change 2001 to 2008		
		Model	ITIN filers	New non-citizens	Model	ITIN filers	New non-citizens	Model	ITIN filers	New non-citizens
Riverside	2,101,000	78,000	75,000	85,000	146,000	145,000	145,000	68,000	70,000	60,000
Sacramento	1,394,000	42,000	36,000	60,000	65,000	64,000	80,000	23,000	28,000	20,000
San Bernardino	2,015,000	100,000	99,000	95,000	150,000	141,000	125,000	50,000	42,000	30,000
San Diego	3,002,000	189,000	165,000	175,000	198,000	184,000	177,000	9,000	19,000	2,000
San Francisco	809,000	42,000	51,000	64,000	30,000	33,000	65,000	-12,000	-18,000	1,000
San Joaquin	673,000	31,000	27,000	35,000	54,000	51,000	43,000	23,000	24,000	8,000
San Luis Obispo	266,000				9,000	10,000	9,000			
San Mateo	712,000	64,000	71,000	60,000	55,000	55,000	65,000	-9,000	-16,000	5,000
Santa Barbara	405,000	37,000	36,000	29,000	37,000	39,000	35,000	0	3,000	6,000
Santa Clara	1,764,000	241,000	246,000	182,000	180,000	185,000	190,000	-61,000	-61,000	8,000
Santa Cruz	253,000	15,000	17,000	16,000	21,000	24,000	17,000	6,000	7,000	1,000
Shasta	180,000				1,000	1,000	2,000			
Solano	407,000	16,000	15,000	16,000	24,000	23,000	22,000	8,000	8,000	6,000
Sonoma	467,000	42,000	43,000	22,000	41,000	43,000	26,000	-1,000	0	4,000
Stanislaus	511,000	23,000	22,000	24,000	39,000	38,000	28,000	16,000	16,000	4,000
Sutter, Yuba	165,000				9,000	9,000	10,000			
Tulare	426,000	33,000	33,000	29,000	29,000	32,000	33,000	-4,000	-1,000	4,000
Ventura	798,000	48,000	49,000	45,000	74,000	83,000	54,000	26,000	34,000	9,000
Yolo	198,000				12,000	12,000	12,000			
Total	36,757,000	2,711,000	2,711,000	2,711,000	2,876,000	2,876,000	2,876,000	165,000	165,000	165,000

SOURCES: Authors' calculations; 2008 ACS.

*New noncitizens are non-naturalized foreign-born who arrived in the previous 20 years (ACS 2008).

According to our estimates, the population of unauthorized immigrants in Los Angeles County was 924,000 in 2001 and declined very slightly to 916,000 by 2008. Two other sources have also estimated this population for similar years; the Los Angeles Family and Neighborhood Survey of 2000 found that there were about 664,000 unauthorized adults in Los Angeles County. Using the residual method, Fortuny, Capps, and Passel (2007) estimated about 937,000 in the same year, 2000.¹⁴ Our estimates are much closer to those that use the residual approach. Similarly, changes over time in large counties appear to be smaller proportionally, according to our model, than unauthorized immigrant population changes in the smaller counties.

Unauthorized Immigrants Live in Zip Codes throughout the State

Within counties, we are able to match, approximately, the unauthorized immigrant population to the zip codes of residence. As we described above, however, we cannot use our model to directly estimate unauthorized immigrant residents in zip codes. Instead, we take our county-level estimates (derived from our models), and allocate the unauthorized immigrants based on the distribution of ITIN tax return filers by zip code within the county.

There are three potential problems with this approach. First, counts of fewer than 10 ITIN tax filers at the zip code level were suppressed by the IRS, so county totals of unauthorized immigrants cannot be allocated to that zip code. However, the number of zip codes with 10 or more ITIN tax filers has risen rapidly in our data years. In 2001, 54 percent of zip codes with any tax filers had ITIN tax filers; by 2008, that share had risen over two thirds, 67 percent.

Second, some zip codes are actually points, such as post office boxes, or office buildings. These are not mapped, but the data from them are included in our county estimates (Tables 3 and 4).

Third, because tax filers may use a work address or an address other than a residence, we find that in a few zip codes we predict higher numbers of unauthorized residents than there are total residents. These are few: in 2008, we found nine, defined as zip codes where the total population was fewer than 1,000 and the percentage of unauthorized residents was greater than 35 percent. In our maps that display the percentage of zip code residents that are unauthorized immigrants, we do not show levels over 15 percent, and so do not expect that these nine zip codes dramatically alter the visual presentation of our results.

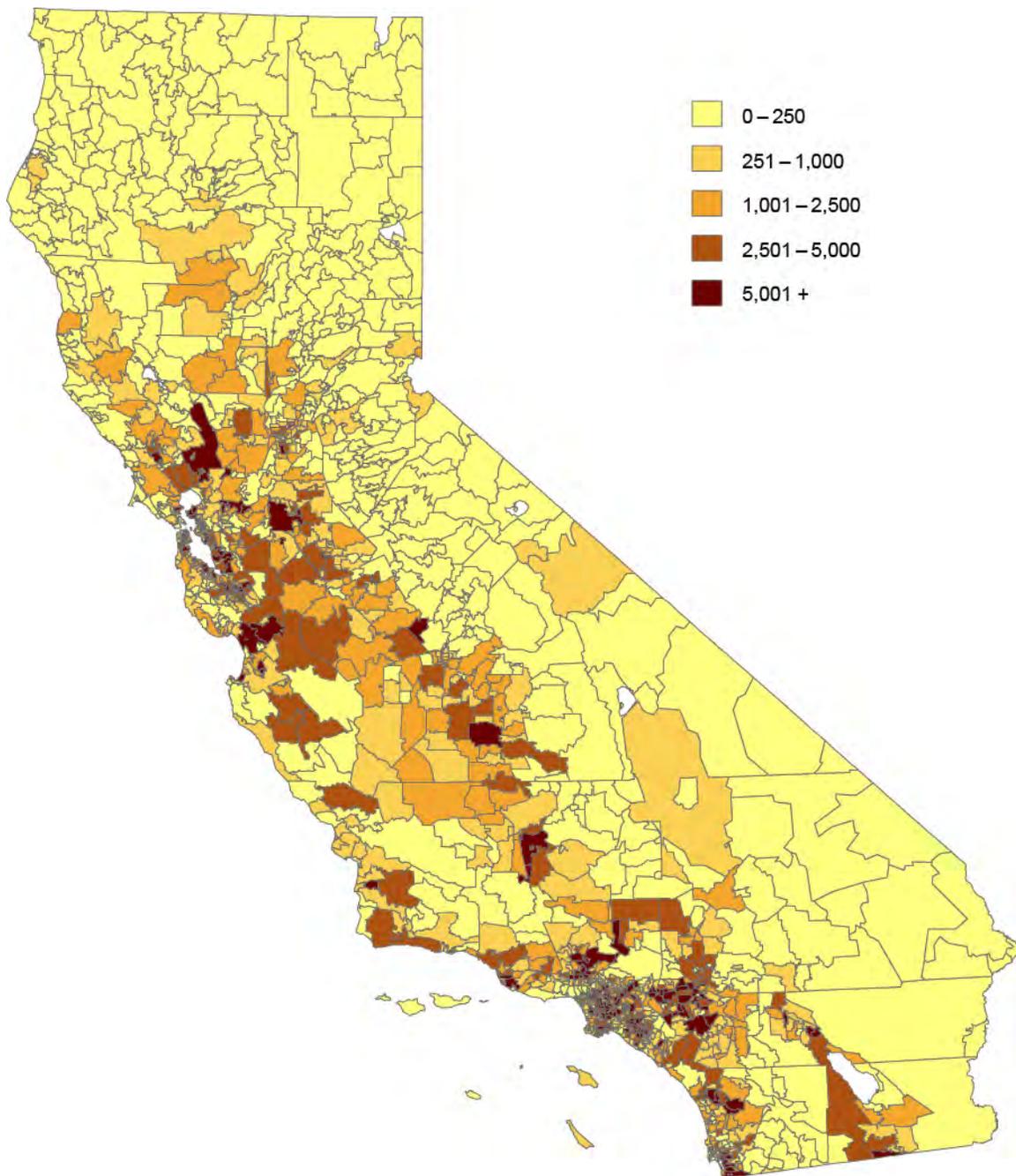
Our methods clearly cannot predict unauthorized immigrants residing in the state's zip codes with exact precision. For that reason, we present our zip code results in ranges, rather in specific number or tabular form. In addition, we do not separate zip codes with zero unauthorized immigrants from zip codes with just a few unauthorized immigrants because of the IRS data suppression issue.

Maps of the state by zip code reveal unauthorized immigrants residing in some very highly concentrated pockets throughout the state, but also located in some places of relative isolation. Throughout the state, we find zip codes with more than 5,000 unauthorized immigrant residents well outside highly urbanized areas (Figure 4a). When we consider the unauthorized as a percentage of the population, we find many zip codes where 15 percent of the population is unauthorized across even more diffuse and diverse geographies (Figure 4b).

Maps for Los Angeles County (Figures 5a and 5b) and for the San Francisco Bay area (Figures 6a and 6b) are provided to illustrate the patterns that emerge from estimating sub-county distributions. All maps reflect 2008 data.

¹⁴ Fortuny et al. (2007) also provide estimates for Los Angeles County in 2003-04 (1,000,000), Orange County PMSA (245,000 in 2000), and Riverside/San Bernardino PMSA (175,000 in 2000).

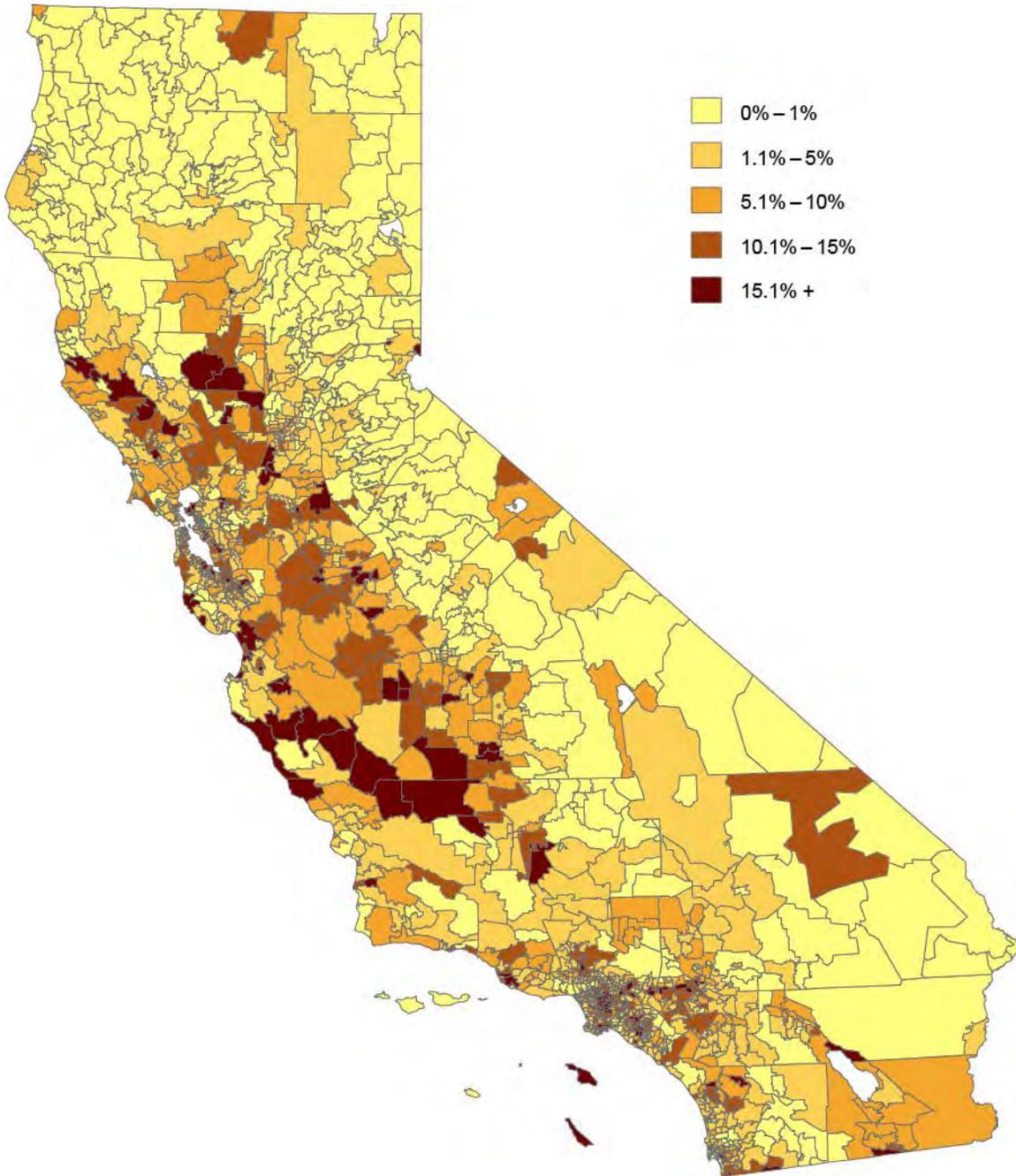
FIGURE 4A
Estimates of unauthorized immigrants in California, by zip code



SOURCE: Authors' calculations using ITIN and ACS data.

NOTE: Areas in white indicate no population.

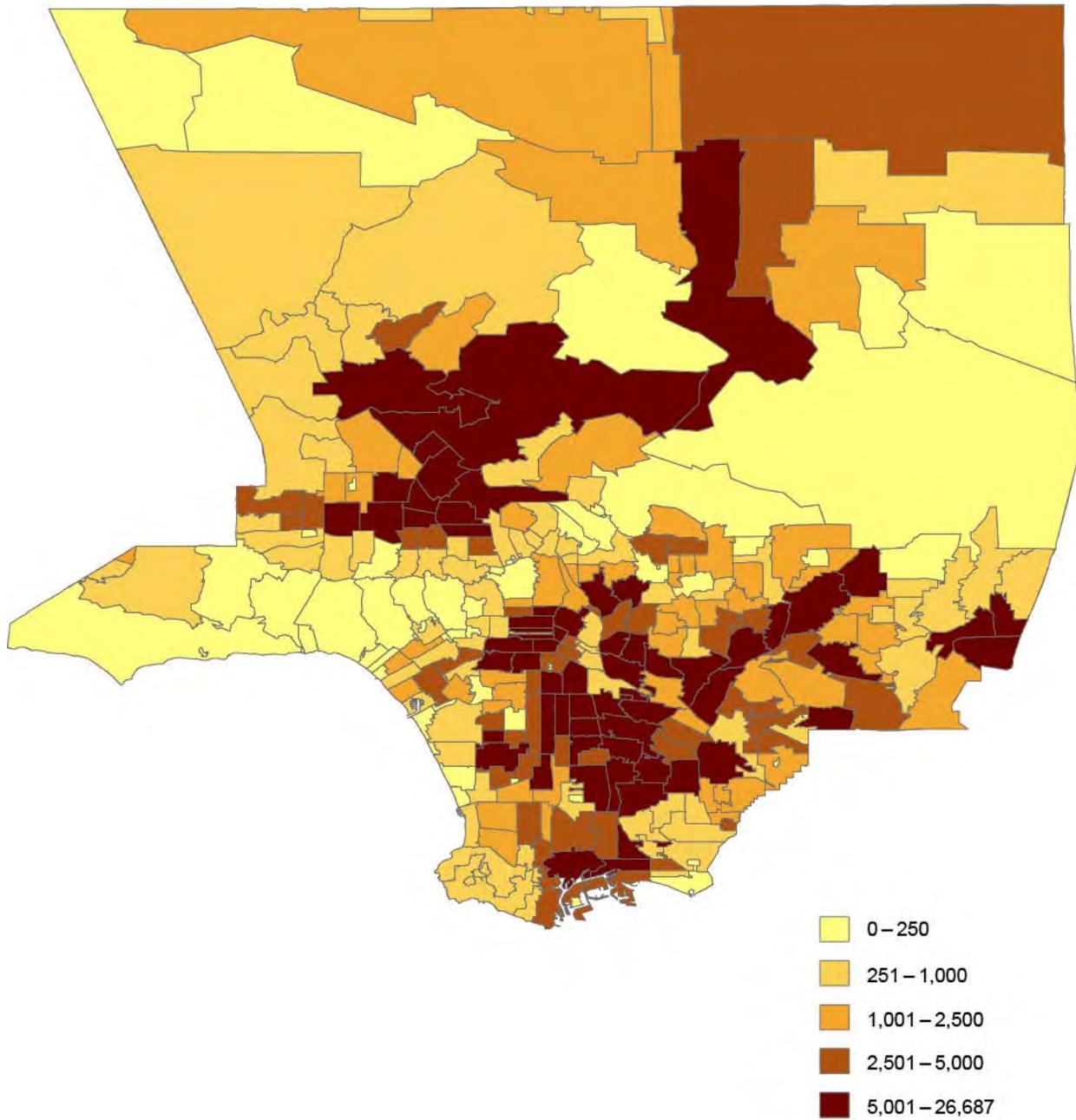
FIGURE 4B
Estimates of unauthorized immigrants in California, percent of population, by zip code



SOURCE: Authors' calculations using ITIN and ACS data.

NOTE: Areas in white indicate no population.

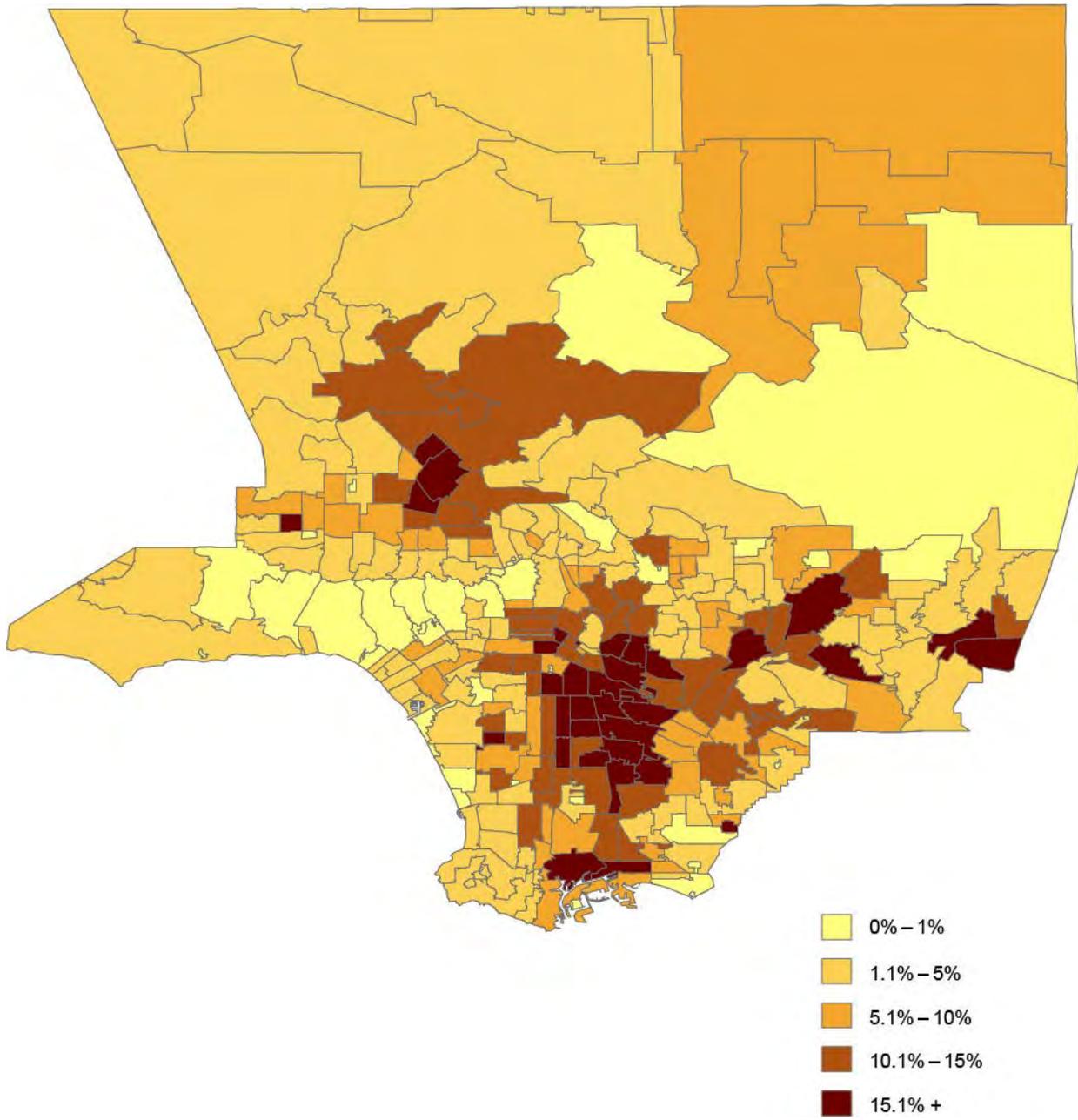
FIGURE 5A
Estimates of unauthorized immigrants, Los Angeles County zip codes



SOURCE: Authors' calculations using ITIN and ACS data.

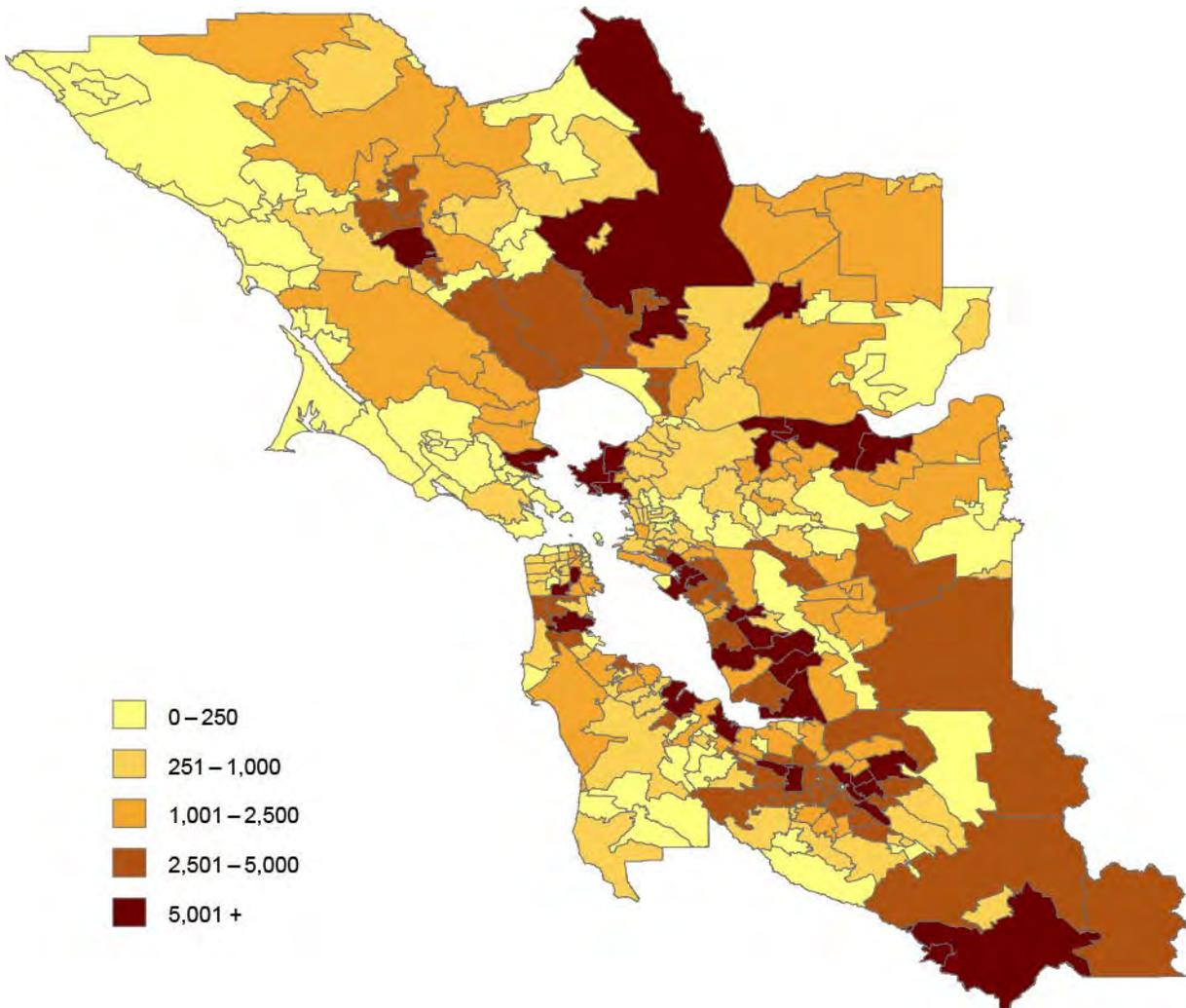
NOTE: Areas in white indicate no population.

FIGURE 5B
Estimates of unauthorized immigrants, percent of population, Los Angeles County zip codes



SOURCE: Authors' calculations using ITIN and ACS data.
NOTE: Areas in white indicate no population.

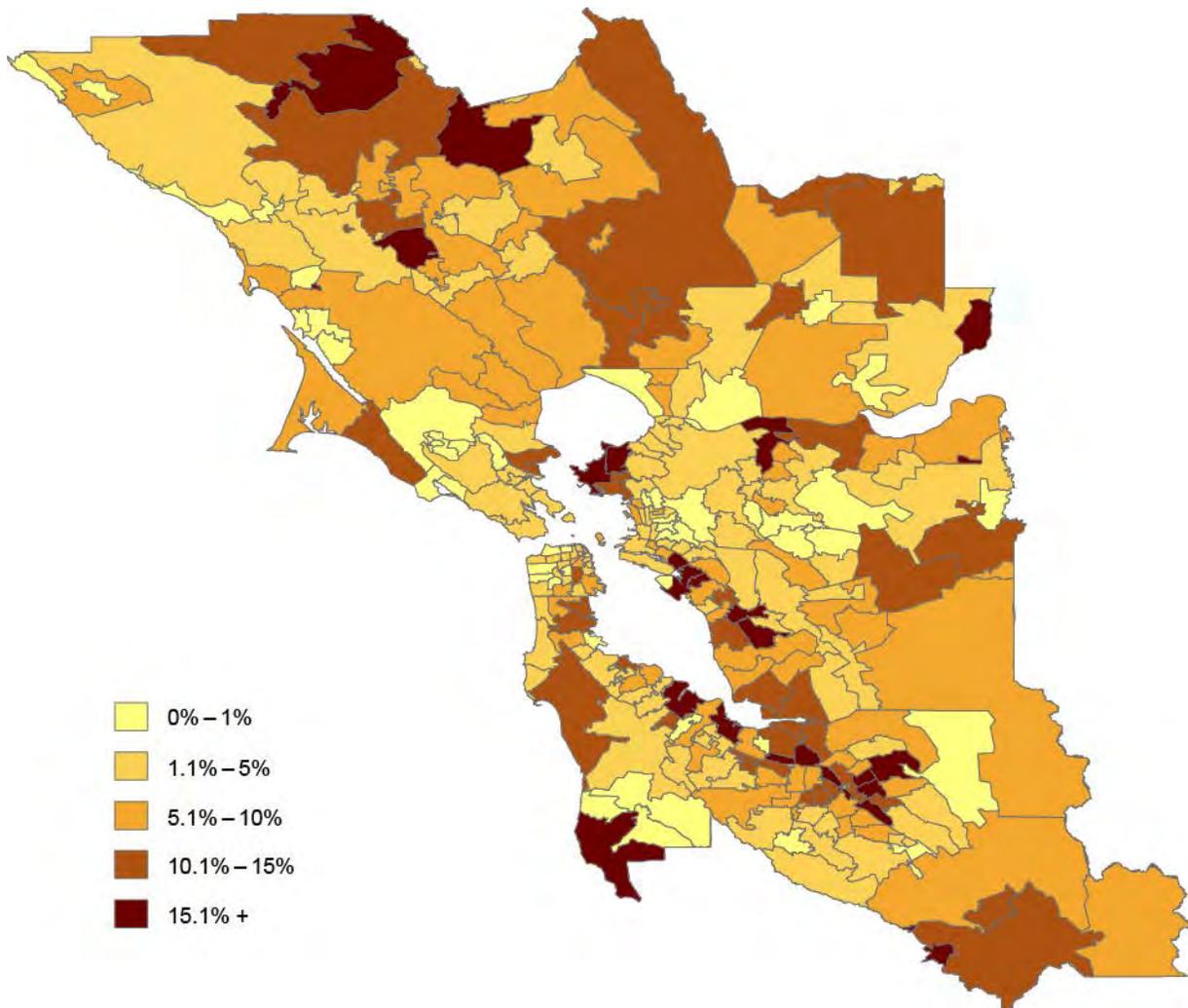
FIGURE 6A
Estimates of unauthorized immigrants, San Francisco Bay Area zip codes



SOURCE: Authors' calculations using ITIN and ACS data.

NOTE: Areas in white indicate no population.

FIGURE 6B
Estimates of unauthorized immigrants, percent of population, San Francisco Bay Area zip codes



SOURCE: Authors' calculations using ITIN and ACS data.

NOTE: Areas in white indicate no population.

Conclusion

We have developed comprehensive sub-state estimates of unauthorized immigrants for California—perhaps the most important state to understand when considering immigration issues. Our estimates are based on administrative data— income tax returns by unauthorized immigrants—available for local areas. Prior to their availability, the best estimates about where in the state unauthorized immigrants reside were limited to larger levels of geography and are now either outdated or available only for subsets of this population.

As with any estimates of unauthorized immigrants, these numbers are subject to uncertainty. However, the administrative data we rely on is highly correlated with independently developed residual estimates of state unauthorized immigrant populations. We take further comfort in our results for Los Angeles County, which are consistent with other estimates derived from the residual method. We expect that as the percentage of unauthorized immigrants using ITIN numbers on their tax returns increases across the country, our method could be used to compute similar sub-state estimates in other locations. Currently, it is reasonable to attempt to do so with states with large populations of unauthorized immigrants who appear to be using ITINs in large numbers, such as Texas, New York, Illinois, and Arizona.

References

- Danielson, Caroline. 2010. "The Earned Income Tax Credit in California." Just the Facts. Public Policy Institute of California. Available at www.ppic.org/main/publication.asp?i=925.
- Fortuny, Karina, Randy Capps, and Jeffrey S. Passel. 2007. "The Characteristics of Unauthorized Immigrants in California, Los Angeles County, and the United States." Washington DC: Urban Institute.
- Goldman, Dana P., James P. Smith, and Neeraj Sood. 2006. "Immigrants and the Cost of Medical Care." *Health Affairs* 25 (6).
- Heer, David M., and Jeffrey S. Passel. 1987. "Comparison of Two Methods for Estimating the Number of Undocumented Mexican Adults in Los Angeles County." *International Migration Review* 21 (4).
- Hill, Laura E., Magnus Lofstrom, and Joseph M. Hayes. 2010. *Immigrant Legalization: Assessing the Labor Market Effects*. San Francisco: Public Policy Institute of California. Available at www.ppic.org/main/publication.asp?i=869.
- Hinojosa-Ojeda, Raul. 2010. "Raising the Floor for American Workers: The Economic Benefits of Comprehensive Immigration Reform." Washington DC: Center for American Progress and Immigration Policy Center.
- Hoefer, Michael, Nancy Rytina, and Bryan C. Baker. 2011. "Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2010." *Population Estimates*, Office of Immigration Statistics, Department of Homeland Security.
- Immigration Policy Center. 2011. "Unauthorized Immigrants Pay Taxes, Too: Estimates of the State and Local Taxes Paid by Unauthorized Immigrant Households."
- Internal Revenue Service. SPEC Tax Data. Available through Brookings EITC interactive (www.brookings.edu).
- Kneebone, Elizabeth. 2008. "Bridging the Gap: Refundable Tax Credits in Metropolitan and Rural America." *Earned Income Tax Credit Series*, Metropolitan Policy Program at Brookings. Washington, DC: Brookings Institute.
- Passel, Jeffrey. 2007. "Unauthorized Immigrants in the United States: Estimates, Methods, and Characteristics." OECD Social, Employment and Migration Working Paper.
- Passel, Jeffrey S., and Karen A. Woodrow. 1984. "Geographic Distribution of Undocumented Immigrants: Estimates of Undocumented Aliens Counted in the 1980 Census by State." *International Migration Review* 18 (3).
- Passel, Jeffrey S., and D'Vera Cohn. 2011. "Unauthorized Immigrant Population: National and State Trends, 2010." Washington DC: Pew Hispanic Center.
- Passel, Jeffrey S., and D'Vera Cohn. 2010. "U.S. Unauthorized Immigration Flows Are Down Sharply since Mid-Decade." 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.
- Pastor, Manuel, Justin Scoggins, Jennifer Tran, and Rhonda Ortiz. 2010. "The Economic Benefits of Immigrant Authorization in California." Los Angeles: Center for the Study of Immigrant Integration, University of Southern California.
- Pastor Jr., Manuel, and Enrico A. Marcelli. 2004. "Somewhere Over the Rainbow? African Americans, Unauthorized Mexican Immigration, and Coalition Building." In *The Impact of Immigration on African Americans*, ed. Steven Schulman (Piscataway, NJ: Transaction Publishers).
- Porter, Eduardo. 2005. "Illegal Immigrants Are Bolstering Social Security with Billions." *New York Times*, April 5.
- Rob Paral and Associates. "Undocumented Immigrants in Congressional Districts." Available at www.robparal.com/MapPage.html?map=14&type=G. Accessed May 5, 2011 (Google Earth plug-in required).
- Warren, Robert. 2011. "Annual Estimates of the Unauthorized Immigrant Population in the United States, by State: 1990 to 2008." Working paper, Public Policy Institute of California. Available at <http://www.ppic.org/main/publication.asp?i=992>.

About the Authors

Laura E. Hill is a policy fellow at the Public Policy Institute of California. Her research interests include immigrants, immigration, race and ethnicity, and youth. She has been a research associate at The SPHERE Institute and a National Institute of Aging postdoctoral fellow. She holds a Ph.D. in demography from the University of California, Berkeley.

Hans Johnson is a senior policy fellow at the Public Policy Institute of California. His research focuses on the dynamics of population change in California and policy implications of the state's changing demography. At PPIC, he has conducted research on international and domestic migration, population projections, housing, and higher education. Before joining PPIC as a research fellow, he was senior demographer at the California Research Bureau, where he conducted research on population issues for the state legislature and the governor's office. He has also worked as a demographer at the California Department of Finance, specializing in population projections. He holds a Ph.D. in demography from the University of California, Berkeley.

Acknowledgments

The authors wish to thank Robert Warren and Elizabeth Kneebone and for their help with the data. Technical reviews of the method and paper by Jeffrey Passel and Manuel Pastor are much appreciated. We also thank Helen Lee, Kim Belshé, Robert Gleeson, and Abby Cook for multiple reviews and conversations, and Richard Greene for his skillful wordsmithing.

PUBLIC POLICY INSTITUTE OF CALIFORNIA

Board of Directors

John E. Bryson, *Chair*

Retired Chairman and CEO
Edison International

Mark Baldassare

President and CEO
Public Policy Institute of California

Ruben Barrales

President and CEO
San Diego Regional Chamber of Commerce

María Blanco

Vice President, Civic Engagement
California Community Foundation

Gary K. Hart

Former State Senator and Secretary
of Education
State of California

Robert M. Hertzberg

Partner
Mayer Brown LLP

Walter B. Hewlett

Director
Center for Computer Assisted Research
in the Humanities

Donna Lucas

Chief Executive Officer
Lucas Public Affairs

David Mas Masumoto

Author and farmer

Steven A. Merksamer

Senior Partner
Nielsen, Merksamer, Parrinello,
Gross & Leoni, LLP

Constance L. Rice

Co-Director
The Advancement Project

Thomas C. Sutton

Retired Chairman and CEO
Pacific Life Insurance Company



PPIC

PUBLIC POLICY
INSTITUTE OF CALIFORNIA

The Public Policy Institute of California is dedicated to informing and improving public policy in California through independent, objective, nonpartisan research on major economic, social, and political issues. The institute's goal is to raise public awareness and to give elected representatives and other decisionmakers a more informed basis for developing policies and programs.

The institute's research focuses on the underlying forces shaping California's future, cutting across a wide range of public policy concerns, including economic development, education, environment and resources, governance, population, public finance, and social and health policy.

PPIC is a private operating foundation. It does not take or support positions on any ballot measures or on any local, state, or federal legislation, nor does it endorse, support, or oppose any political parties or candidates for public office. PPIC was established in 1994 with an endowment from William R. Hewlett.

Mark Baldassare is President and Chief Executive Officer of PPIC.
John E. Bryson is Chair of the Board of Directors.

Short sections of text, not to exceed three paragraphs, may be quoted without written permission provided that full attribution is given to the source and the above copyright notice is included.

Research publications reflect the views of the authors and do not necessarily reflect the views of the staff, officers, or Board of Directors of the Public Policy Institute of California.

© 2011 Public Policy Institute of California
All rights reserved.
San Francisco, CA

PUBLIC POLICY INSTITUTE OF CALIFORNIA
500 Washington Street, Suite 600
San Francisco, California 94111
phone: 415.291.4400
fax: 415.291.4401
www.ppic.org

PPIC SACRAMENTO CENTER
Senator Office Building
1121 L Street, Suite 801
Sacramento, California 95814
phone: 916.440.1120
fax: 916.440.1121