

Estimating unauthorized immigrant populations in sub-state areas: California as a test case

California is home to approximately 2.6 million unauthorized immigrants (Passel and Cohn, 2010 and Hoefer, Rytina, and Baker, 2010). Careful estimates of the population's size at the state and national level are produced regularly by the Pew Hispanic Center (PHC) and the Department of Homeland Security (DHS), but cannot provide information about the distribution of the population within the state. We propose to develop the first-ever estimates of the unauthorized immigrant population for California's counties and some sub-county areas. We will use several data sets to develop these estimates, relying primarily on newly generated and recently available data derived from IRS tax filings.

Getting estimates of the unauthorized immigrant population into the hands of local service providers, foundations, county governments, and policy makers will be useful to them in a number of ways. For example, it will help non-profits and counties targeting outreach for service provision and benefits, such as access to public programs for which children of unauthorized immigrants are eligible (e.g. SNAP, Medi-Cal, CalWORKs). It will help non-profits and funders target outreach to increase civic engagement among immigrants, and to spread knowledge related to new immigration/enforcement policy. The numbers will help health care providers of last resort to the uninsured plan for enactment of health care reform. Finally, these data could guide planning for increased demand for legal services connected to comprehensive immigration reform.

Our research will provide local area estimates of unauthorized immigrants within California. The specific objectives are to (1) produce estimates of unauthorized immigrants for California's counties (and sub-county areas in Los Angeles County), and (2) assess the feasibility of using ITIN data as a proxy for the unauthorized immigrant population for sub-state areas in California in the future, as well as their use for estimating populations at the sub-state level in other states.

In order to develop the estimates, we will use several data sets, relying primarily on newly generated and recently available data derived from IRS tax filings. Since 1996, unauthorized immigrants have been able to file tax returns using Individual Taxpayer Identification Numbers (ITINs) on their 1040 forms. The IRS has made counts of ITIN filers for each zip code publically available for tax years 2000-2007.

In addition, we will use estimates of the unauthorized population produced for the states by the Pew Hispanic Center (PHC) and the Office of Immigration Statistics at the Department of Homeland Security (DHS). These estimates of the unauthorized population all are derived using a residual method with counts of the foreign-born in population surveys. In this method, naturalized citizens are subtracted from the population of the foreign-born, as are estimates of legal permanent residents and temporary visa holders. The remainder are considered to be unauthorized immigrants (see Passel and Cohn 2010 and Hoefer, Rytina, and Baker 2010 for more detailed descriptions).

Below, we describe the ITIN data and our estimation methods.

Data: The Individual Taxpayer Identification Number (ITIN)

Despite having worked without authorization, unauthorized immigrants are obligated to file tax returns. Estimates suggest a large percentage do so. Fifty percent is the most commonly used estimate, but the Social Security Administration's chief actuary assumed closer to three-

quarters of unauthorized immigrants have payroll taxes withheld (Porter 2005). A recent PPIC study found that over 85% of working unauthorized immigrants reported having filed federal taxes in the year prior to earning legal permanent residence (Hill, Lofstrom, and Hayes 2010).

The IRS has facilitated the process for unauthorized immigrants to file taxes (and possibly earn refunds) by allowing unauthorized immigrants to apply for ITIN numbers. To apply for an ITIN, unauthorized immigrants may attach an ITIN application form (W-7) to their tax return and leave the SSN blank. An ITIN functions like a social security number (SSN) for the purpose of filing taxes or opening bank accounts for those without work authorization, but cannot be used to secure employment. Once obtained, ITIN numbers are used in the place of SSNs on an unauthorized immigrant's 1040 form. Immigrants working in jobs that provide W-2 forms to their employees would attach their W-2 form to the 1040, but the W-2 would have a false or fraudulent SSN. The IRS then processes the return and issues any refund the person listed on the ITIN, not the owner of the SSN on the W-2. One can file a tax return without a W-2 if a W-2 was not provided by the employer.

Unauthorized immigrants have many incentives to file tax returns. First, filing tax returns will establish a paper trail that indicates employment and paying taxes, both of which are likely to increase an unauthorized immigrant's chance of earning legal status if comprehensive immigration reform passes. Second, some unauthorized immigrant workers will be eligible for tax refunds and might file with ITIN in order to claim them. Unauthorized immigrant workers can receive child tax credits, however they cannot receive the Earned Income Tax Credit even if their children are US citizens. Third, after becoming legal, formerly unauthorized immigrants can link payroll taxes withheld under a false or fraudulent SSN through their ITIN. Unauthorized immigrants may legitimately be concerned that filing tax returns with an ITIN would identify them to the federal government and increase the chances that they would be punished or even deported. However, the IRS has taken the position that they are required to make it possible for all who are required to file taxes to do so and that they will not identify taxpayers using ITINs to other branches of the government.

Not all tax returns filed with ITINs belong to unauthorized immigrants. Some authorized immigrants are required to file U.S. tax returns but are not authorized to work, and therefore do not have SSNs. For example, some international students must pay taxes on grants or stipends, but are not authorized to work, and thus do not have social security numbers. Other types of ITIN filers who are not unauthorized immigrants include international investors or foreign gambling winners. Those that do not have work authorization, or in many cases, do not even live in the country, must still file taxes, and would not have an SSN. They would use an ITIN, and would likely file a 1040NR (non-resident) tax return. Approximately 7% of ITIN filers in 2007 filed 1040NRs.

Anyone who is authorized to work in the United States, such as those on temporary work visas (e.g. H1-B, or foreign students with work authorization) are required to apply for and file federal taxes with a SSN. The vast majority of ITIN filers appear to be unauthorized immigrants. Because most tax returns filed with ITINs also include wages (85% in 2007), we assume most ITIN filers are unauthorized workers.

Since 1996, the number of tax returns filed with ITINs has increased dramatically. By December 2008, the IRS had issued over 14 million ITINs (Treasure Inspector General for Tax Administration, 2009). A substantial number of all unauthorized immigrants in California filed income taxes using an ITIN. The use of ITINs has been increasing over time, and in California, nearly 6% of filers used an ITIN in 2007, up from 2% in 2000. It appears that a substantial

number of all unauthorized immigrants filed income taxes: in 2007, the number of ITIN filers was 36% of the number of unauthorized immigrants estimated to be in California.

Method

An important aspect of this research will be to validate the ITIN data as a proxy for unauthorized immigrants, and adjust ITIN numbers upwards sensibly, to account for difference in ITIN usage across the states.

Our first task is to assess how well the distribution of ITIN filers across the states matches estimates from PHC and DHS, and with American Community Survey data. Preliminary results for 2007 indicate a high degree of correlation between these estimates and the ITIN data in 2007. We find a correlations ranging from 0.96 to 0.98.

Table 1. 2007 ITIN correlations for the 50 states and DC

2007 ITIN with:	Correlation
Pew Hispanic Center estimates	
2006-2008	0.961
2009	0.974
Average 2006-2008 and 2009	0.970

Note: DHS estimates of the unauthorized population are produced only for the 10 states with the largest populations of unauthorized immigrants, and correlations are not included here

ITIN counts are completely independent of the ACS, but are still highly correlated with them. The main advantage of the ITIN counts is that they are available at the zip code level, while the PHC and DHS estimates are available only at the state level.

Table 2. Correlations between ITIN tax filers and ACS,
50 States and DC

	ITIN 2007
2006-2008 ACS	
Total population	0.871
Foreign-born	0.950
Foreign-born, non citizen	0.966
Foreign-born, male 18+, non citizen	0.967

Note: DHS estimates of the unauthorized population are produced only for the 10 states with the largest populations of unauthorized immigrants, and correlations are not included here.

We will also examine the correlations for the years 2000-2006, but we expect earlier years will not exhibit as high a degree of correlation. It appears unauthorized residents in some states, such as California, were early adopters of the ITIN, while residents in other states were late adopters.

Second, we will adjust state ITIN totals to match estimates of the unauthorized immigrant population for 2008. Before doing so, however, we will attempt to understand how relationship varies from state to state, and over time. For example, our early work suggests that Illinois has larger numbers of ITIN filers and Florida has fewer than we would expect based on the 2008 estimates from PHC and DHS.

Differences in ITIN usage may be related to industry mix, participation in the formal economy, prevalence of self-employment, prevalence of paid tax preparers, presence of children in the household, and infrastructure to support immigrant communities, among other things. We will estimate a model to account for state variation in the number of ITIN filers relative to the number of unauthorized immigrants estimated to be residing in the state by the PHC. Using the ACS, we will estimate the following simple equation:

$$\left(\frac{PHC}{ITIN}\right)_{state} = (\% \text{ self employed})_{state} + (\% \text{ NILF})_{state} + (\% \text{ Mexican} - \text{born})_{state} + (\% \text{ with own kids in hh} < 18)_{state} + (population)_{state} + \varepsilon$$

Finally, we will adjust the zip code ITIN counts to match PHC state totals for California. We will use the coefficient estimates from the above equation to adjust ITIN estimates for regions within California. Then we will apply those adjustments to the zip codes within the regions in the state. Where there are estimates of the unauthorized population, such as for Los Angeles County (from the 2000 Los Angeles Family and Neighborhood Survey), we will examine the match.

These three steps will allow us to produce estimates of unauthorized immigrants for counties in California for 2007, and may be able to produce estimates for prior years as well. This paper will also assess the feasibility of making substate estimates of the unauthorized immigrant population for other states.