

# Profiling the Eligible to Naturalize

*By*

Manuel Pastor, Patrick Oakford, and Jared Sanchez

**Center for the Study of Immigrant Integration &  
Center for American Progress**

Research Commissioned by the

**National Partnership for New Americans**

Draft: November 24, 2014

**Please do not quote or cite without permission.**



## Introduction

Recent research has suggested that the high fees to naturalize may serve as an impediment to Legal Permanent Residents (LPRs) seeking citizenship (Pastor et al. 2013). That research has been based on observing responses over time to increases in the fee (particularly to the differential between the costs to naturalize and the costs to renew a “green card”) and seeing how that affects both the level and composition of those choosing to naturalize in any given year. In general, those results square with the notion of price sensitivity: the share of LPRs with less education (and likely less income) tends to fall dramatically when the fees go up as does the share of Mexican-origin LPRs choosing to make the transition to citizenship.

To understand what fee changes might mean going forward, it would be useful to have a more detailed profile of those eligible to naturalize, including by income, education and English language ability. Unfortunately, this information is not readily available in the data maintained by the Office of Immigration Statistics which essentially relies on broad flows to calculate the number of Legal permanent residents and the share of those eligible to naturalize. This memo seeks to address that gap by attempting to simulate the LPR population with data in the 2012 American Community Survey (ACS) that can offer a more detailed picture.

## Methods and Data

The methods we use to identify individuals in the ACS are described in the appendix. In general, the approach is to use well-known techniques to estimate the undocumented, call the remaining foreign-born resident LPRs (with a cut-out made to exclude those foreign-born who are likely to be here on student visas), and then to apply the standards utilized by Citizenship and Immigration Services (CIS) to tag individuals as eligible to naturalize or not. We wind up with over 12.1 million LPRs, of which roughly 8.8 million are adults eligible to naturalize.

These aggregate numbers, particularly on the eligible to naturalize, are very similar to those in the latest report from the Office of Immigrant Statistics on LPRs; we are underestimating LPRs by about one million but we are very close to the estimated number of those eligible to naturalize which is the crucial measure here (see Table 1 in Rytina (2013). And while our composition of those immigrants eligible to naturalize differs to some degree in terms of countries of origin, period in which status was adjusted, and state (the first two of these breakdowns are shown in Tables 1 and 2), the deviations are not large.

For example, the OIS estimates that Mexican-origin immigrants comprise 30.6 percent of those eligible to naturalize and our estimated share is 29.0 percent of adults. While our figures are higher for Indians, this is likely because included in our LPR numbers (unavoidably given this data exercise) are holders of H-1B visas (Indians are the largest group of these) As for the period of adjustment, we proxy that with year of entry – and unsurprisingly our numbers for the 1980s are much higher since many of the immigrants who entered in those years adjusted their status later due to the Immigration Reform and Control Act of 1986. In any case, while there is some noise in this data, the use of the ACS allows us to give a general estimate of some of the broad income and other characteristics of the eligible to naturalize population that cannot be made without this simulation.

**Table 1. Country of Birth of LPR Eligible to Naturalize Adult Population, 2012**

Year	Number	Percent
Total	8,825,670	
Pre-1960	166,445	1.9%
1960-69	390,352	4.4%
1970-79	905,553	10.3%
1980-89	2,114,181	24.0%
1990-99	2,660,320	30.1%
2000-2009	2,588,819	29.3%
2010-2012	-	-

**Table 2. Year LPR Status Obtained for the Eligible to Naturalize Adult Population, 2012**

Country of Birth	Number	Percent
Total	8,825,670	
Mexico	2,555,338	29.0%
China	361,291	4.1%
Philippines	249,765	2.8%
India	442,959	5.0%
Dominican Republic	279,797	3.2%
Cuba	293,079	3.3%
Vietnam	109,070	1.2%
El Salvador	244,065	2.8%
Canada	283,233	3.2%
United Kingdom	256,822	2.9%

## Characteristics of the Eligible to Naturalize

Given our general confidence that these numbers likely offer a somewhat representative view, we now turn to a breakdown to those eligible to naturalize by various factors. Of particular concern is income. It is already the case that immigrants eligible to naturalize who fall below 150 percent of the poverty line can apply for a fee waiver. One key question is whether extending some form of waiver to those in a slightly higher bracket – presumed to be the working poor and so price-sensitive – might encourage naturalization and be feasible in terms of the costs incurred by USCIS in processing those applicants. We cannot address the cost side of the equation but we do look at the income profile of those eligible to naturalize.

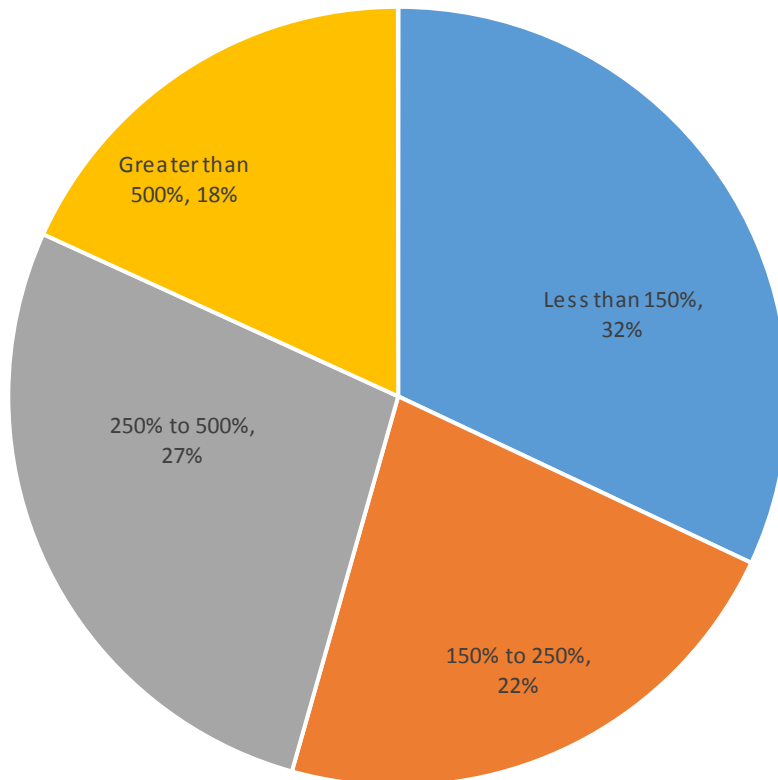
Table 3 takes up this topic by breaking the eligible to naturalize adult population into those living below 150 percent of the poverty line, those living between 150 to 250 percent of the poverty line, and those living above that income marker. We should note that the raw numbers fall in this exercise since the poverty line is a family-ascribed characteristic and is not available in the data for boarders, foster children, and other non-relatives.

In any case, note that there are nearly two million adult eligible to naturalize LPRs within the 150% to 250% poverty band – representing 22 percent of all adults eligible to naturalize LPRs for whom we have poverty data, a lower share of the total than the group (those living below 150% of the poverty level) for whom the waiver is already generally available. We offer a further breakdown (separating out very high income immigrants) in Figure 1.

**Table 3. Adult LPRs Eligible to Naturalize by Poverty Bands, 2012**

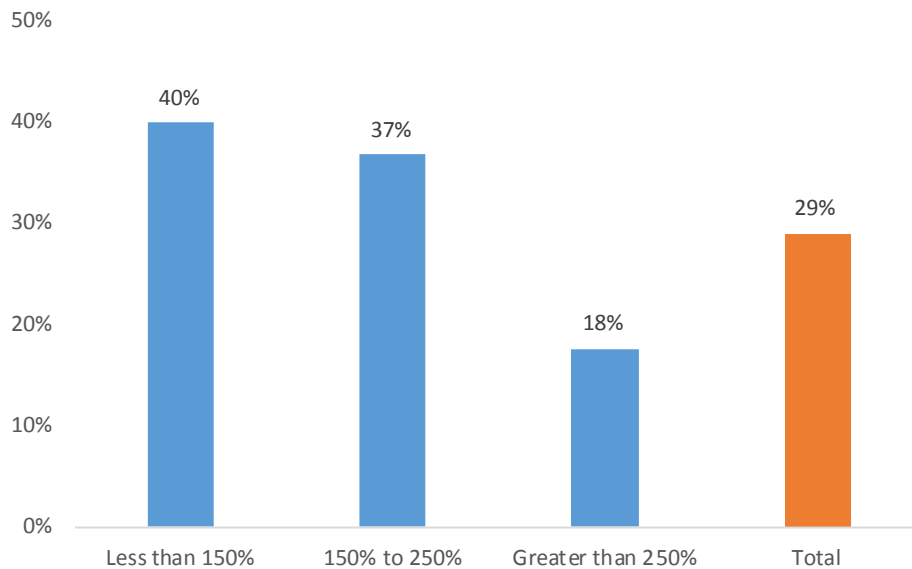
Percent of Poverty Level	Number	Percent
Less than 150%	2,790,404	32%
150% to 250%	1,949,149	22%
Greater than 250%	3,974,413	46%
Total	8,713,966	100%

**Figure 1. Adult LPRs Eligible to Naturalize by Poverty Bands, 2012**



Granting some form of waiver to this group in the 150 to 250 percent band might have important impacts on the ethnic or national-origin composition of those who may naturalize. As reported in Pastor, et al., (2013), previous increases seem to have had a sharper effect on the Mexican-origin population. One can see why by looking at the numbers: while Mexican-origin immigrants comprise 29 percent of all adult eligible to naturalize, they comprise 40 percent of all adults eligible to naturalize in the band below 150% of the poverty level, 37 percent of all adults eligible to naturalize within the 150% to 250% poverty band, and only 18 percent of those eligible to naturalize in the higher income groups (see Figure 2). From another perspective, while 22 percent of all eligible to naturalize LPRs fall in the 150% to 250% poverty band, 28 percent of all eligible to naturalize Mexicans are within that band (see Table 4).

**Figure 2. Share that are Mexican-Origin of those Adults Eligible to Naturalize by Poverty Bands, 2012**

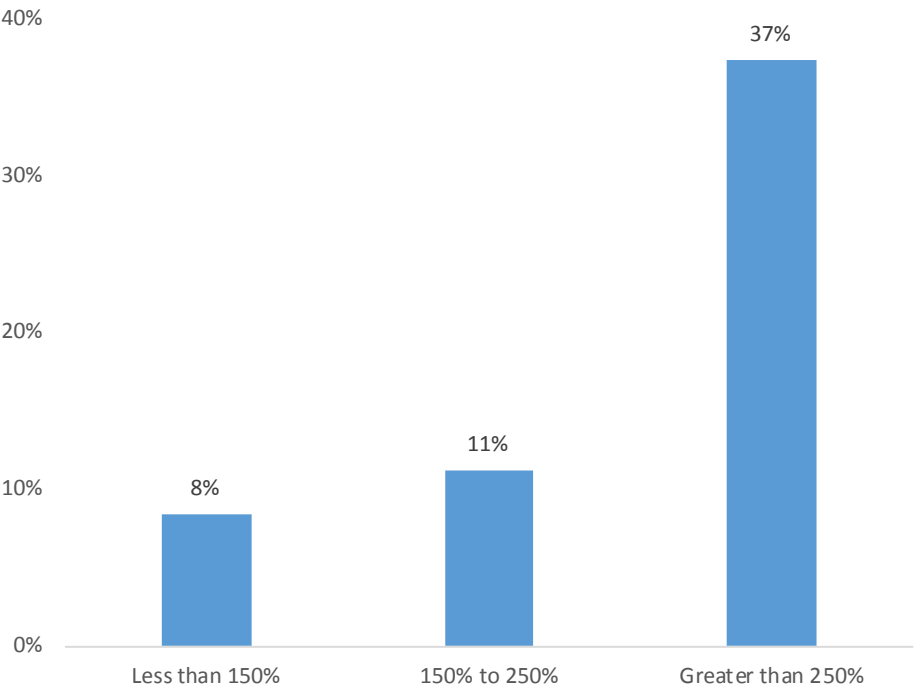


**Table 4. Adults Eligible to Naturalize By Country of Birth by Poverty Bands, 2012**

	Less than 150%	150% to 250%	Greater than 250%	Total
Mexico	44%	28%	28%	100%
China	29%	17%	54%	100%
Philippines	14%	17%	69%	100%
India	9%	9%	82%	100%
Dominican Republic	48%	26%	26%	100%
Cuba	42%	27%	30%	100%
Vietnam	36%	23%	41%	100%
El Salvador	36%	32%	31%	100%
Canada	14%	15%	71%	100%

While the share of eligible to naturalize LPRs in the 150 to 250% of poverty band is lower than the share in the band with incomes below 150% of the poverty level, one should recognize that the former group might have a higher take-up in the case of a reduced fee. Disaggregating the poverty/income band between three separate categories (less than 150%, 150% to 250%, and greater than 250% of the poverty level), we find that 47 percent, 57 percent, and 77 percent (respectively) of adults eligible to naturalize report speaking English well or better (or report that it is their only language). For the same three income groups but comparing educational attainment, 8 percent, 11 percent, and 37 percent have a Bachelor’s Degree or higher (see Figure 3). Since these are all characteristics also associated with lower rates of naturalization, a slightly higher take-up for this group than for the poorest immigrant may be likely.

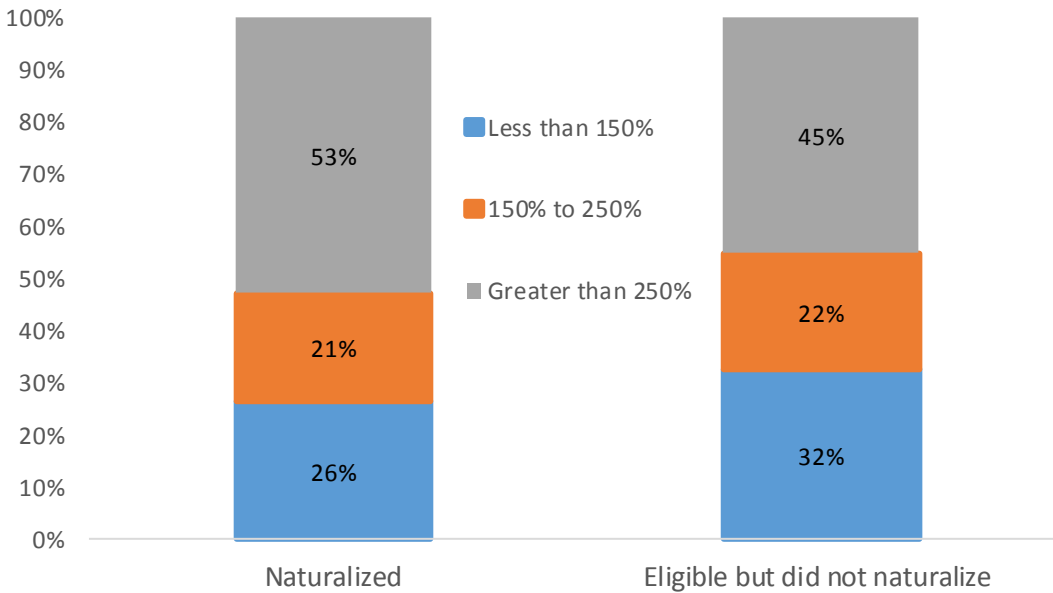
**Figure 3. Share with BA or Higher for those Eligible to Naturalize by Poverty Bands, 2012**



Would there be takeup? While the data is not immediately amenable to examining price sensitivity, we created a group of LPRs who would have been eligible to naturalize as of 2011 and compared the mix of their income levels to those LPRs who actually naturalized in 2011 or in 2012 (since 2008, the ACS has included the year of naturalization in the individual microdata). As it turns out, that number for 2011 is virtually identical to the number of naturalizations recorded for 2011 in the official OIS data – we included 2012 naturalizations in the determination since these individuals had an intention to naturalize very soon but those numbers do not match 2012 OIS number because the ACS is a rolling sample and so fewer 2012 naturalizations are recorded.

In Figure 4, we show that the income composition of those who naturalized relative to those who did not. As can be seen, there was a distinct bias toward higher income groups. Another way of looking at this is to estimate naturalization rates by dividing the number of those who naturalized by those who were eligible but did not (plus those who did). In this case, it makes sense to focus just on those who naturalized in 2011. We find low rates of naturalization that should worry any observer but also clear differences by income band: 6.0 percent for those below 150 percent of the poverty level, 7.1 percent for those in the 150-250 percent band, and 9.2 percent for those in the band above 250 percent of the poverty level. If, for example, the naturalization rate for the middle band had risen to the naturalization rate for the above 250 band, there would have been just under 40,000 additional naturalizations that year. In general, it would be useful to raise naturalization rates, particularly given the civic and economic benefits associated with citizenship, and so the likelihood of higher tax revenues in the future (Pastor and Scoggins 2012).

**Figure 4. Income Composition of those who naturalized in 2011 and 2012 compared to those eligible to naturalize in those periods who did not**



### Future Research Directions

Past increases in the naturalization fees seem to have discouraged less educated (and presumably less well-off) immigrants from becoming citizens. The impacts have been ethnically disproportionate with a larger share of Mexican-origin LPRs dropping off the path to naturalization. US CIS has sought to address part of the issue by providing waivers for those falling below 150% of the poverty line.

The analysis above suggests that there may also be a population whose income is slightly higher than that that might benefit from a fee waiver as well. Currently, a family of four living at the poverty level (so in the midst of the usual poverty threshold band) would have to give up three weeks of labor to have both parents pay the citizenship fees. For the population eligible to naturalize living at 150 to 250 percent of the poverty level, fees likely also pose a challenge for naturalizing, since the fees alone for the same family of four to naturalize both parents consume roughly one and half weeks of labor for a family whose income is in the middle of the range. Such families are not likely to have sizeable extra resources. Lowering the effective cost by, say, half would likely encourage citizenship and also alleviate some of the unintentional country-of-origin differentials that emerged as a result of the earlier fee increases.

Further research is needed both on price sensitivity and cost recovery, particularly whether any additional naturalizations could be covered under a partial waiver system for what are likely working poor LPRs. However, the initial analysis here suggests ways in which the use of the American Community Survey could facilitate a better view of what might happen should authorities make an even stronger commitment to facilitating citizenship on the part of Legal Permanent Residents, a goal on which there seems to be broad consensus.



## Technical Appendix

There is no existing profile on the income, educational, and language of Legal Permanent Residents (LPRs) who are eligible to naturalize, mostly because the OIS data is constructed through broad aggregate flows. To simulate this population, we worked with the 2012 American Community Survey (ACS). The ACS data does have information on who is foreign-born and non-citizen which is the building block for this exercise. Of course, that population also includes undocumented immigrants and the first step was to identify those who in the sample were likely to be documented (LPRs) and undocumented.

To do this, we assume that the aggregate total of undocumented adults in the U.S. in 2012 was similar to that reported in the most recent estimate from the Office of Immigration Statistics (Rytina 2013). We then take every non-citizen, non-Cuban foreign-born respondent in the ACS sample and assign to each of these respondents an initial documentation status based on certain characteristics. For example, we assumed that any non-citizen non-Cuban immigrant with military experience is an LPR. Other characteristics that pushed a respondent into LPR status included whether or not the respondent worked for the public sector; had an occupation (such as police officer) that required documents; received social security or disability payments; or received food stamps but did not have a child in the house (who could have been the legal source of the assistance). We also removed from the undocumented category those who immigrated as adults and were currently enrolled in higher education on the grounds that they were likely student visa holders (while they could be future permanent residents, we actually remove them later in the process since they are not eligible to naturalize).

We also placed respondents in the LPR category if they received Medicare (and were older than 65), Veterans Affairs Care, or Indian Health Services. We did not automatically make someone younger than 65 who reported Medicare as documented as they could have been reporting for another member in the household (as might be guessed, this affected a very small number of people and so was an appropriate but not a critical decision). We did not assume that reporting Medicaid was sufficient to designate one as documented because the question asked is ambiguous and could be interpreted as asking about any kind of public health assistance such as emergency care at the hospital.

That initial assignment left us with an undocumented population that was significantly larger than it should have been given OIS estimates. To assign the rest, we first determined their probability of being undocumented using a technique similar to that in Capps et al., (2013). Following the very clear directions kindly provided by those authors, we started with Wave 2 of the 2008 Survey of Income and Program Participation (SIPP) in which respondents offered answers with regard to whether they had LPR status upon arrival and whether they had ever achieved it later; those who answered no were considered to be undocumented. To move these responses over to the American Community Survey, Capps, et al. (2013) use a multiple imputation strategy to populate “missing” answers for the ACS (which are basically all the answers). We instead utilize a logistic regression strategy in which the probability of being undocumented is determined by an equation in which the right-hand side variables include gender, age, years since arrival, education levels, marital status (whether never married and if married, whether married to a U.S.-born citizen), several dummy variables for broad region, and several interactive terms to allow for coefficients to vary.

The combination of conditions and probabilities can actually get one very close to the correct number and composition of the undocumented, using the weighted average of the probability as the cut-off. To get a better fit, we then took advantage of the fact that the OIS offers a breakdown of the top 10

nations of origin of the undocumented (Rytina 2013). Adjusted for age to look at adults, we took all non-citizen non-Cuban adult residents who had not been assigned to documentation by the conditions and sorted them in order of the probability of being undocumented, using a random number to all respondents to break ties where a large group of respondents shared the same probability. We then sorted over to the LPR column till we arrived at the right number of undocumented resident (accounting for an undercount explained below).

For the remaining countries, we used a two-year average from 2009 and 2010 for the Brazilian undocumented (their number fell recently and so the count was not in the most recent OIS reports) since unauthorized Brazilians are a very large share of the non-citizen Brazilian immigrant population (Marcelli et al. 2009). For the rest of the unauthorized population, the easiest approach would be to assume that all nations of origin have exactly the same share of undocumented residents by comparing the remaining OIS numbers to the non-citizen non-Cuban immigrant numbers in the ACS. However, that is clearly not the case and we investigated the next 20 largest countries sending immigrants, taking advantage of several bits of knowledge in the field, including an estimate of undocumented Canadians that was generated by the Migration Policy Institute (MPI) in 2008 (see (Slovic 2008) and work by Marcelli as well as the Migration Policy Institute that suggests that the share of undocumented is quite low in the Dominican community (Grieco 2004; Marcelli et al. 2009). For other countries, we use available information on similar countries in their hemisphere (either from the overall data or from the information in the SIPP data) to target a percent undocumented and hence number undocumented. At the end of the targeting and assignment process, we have a total number of adult undocumented residents that is close to the OIS totals.

Better put, it is close given an assumed undercount. There because there is a widely shared assumption that the undocumented are undercounted by around 10% in the decennial census (see (Marcelli and Ong 2002)) and more in other samples. To account for this, we had initially set the targets below the target adult numbers (nation-by-nation) so that when we reweighted all of those observations up, we would arrive at a more likely number. (Warren and Warren 2013) contend, reasonably enough, that the undercount might be as high as 20% in recent years because the ACS is perceived as a more voluntary survey by respondents than is the Census. To implement this procedure, we stuck closer to the earlier research and set the undercount assumption of 12.5%.

While this approach is only a first approximation to a full identification of undocumented immigrants and LPRs in the ACS, it is similar to approaches used by other researchers such as the propensity approaches developed by James Bachmeier and Jennifer Van Hook in which they identify undocumented immigrants in the 2008 Survey of Income and Program Participation by respondents' answers to a variety of migration related questions and use that information as well as certain screens similar to the ones we deploy (Bachmeier, Van Hook, and Bean 2014; Batalova, Hooker, and Capps 2014). While we have likely identified some undocumented residents as documented (and the other way around), this overall approach is likely sufficient for the broad approximations in this memo.

To the extent that there is a bias in our numbers, we think that we may be overestimated the number – but not the composition – of those eligible to naturalize. That is, our numbers are basically identical to the 8.8 million reported in the most recent OIS report but would be higher if we were to include potential derivative citizenship numbers (that is, children who could naturalize if the parent could naturalized). We do not tackle those derivative numbers in this memo but could in a future iteration of this work.

## References

- Bachmeier, James D., Jennifer Van Hook, and Frank D. Bean. 2014. "Can We Measure Immigrants' Legal Status? Lessons from Two U.S. Surveys." *International Migration Review* 48(2):538–66.
- Batalova, Jeanne, Sarah Hooker, and Randy Capps. 2014. *DACA at the Two-Year Mark: A National and State Profile of Youth Eligible and Applying for Deferred Action*. Washington, D.C.: Migration Policy Institute. Retrieved November 11, 2014 (<http://www.migrationpolicy.org/research/daca-two-year-mark-national-and-state-profile-youth-eligible-and-applying-deferred-action>).
- Capps, Randy, James D. Bachmeier, Michael Fix, and Jennifer Van Hook. 2013. *A Demographic, Socioeconomic, and Health Coverage Profile of Unauthorized Immigrants in the United States*. Washington, D.C.: Migration Policy Institute. Retrieved November 25, 2014 (<http://www.migrationpolicy.org/research/demographic-socioeconomic-and-health-coverage-profile-unauthorized-immigrants-united-states>).
- Grieco, Elizabeth M. 2004. *The Dominican Population in the United States: Growth and Distribution*. Washington, DC: Migration Policy Institute. Retrieved ([http://www.migrationpolicy.org/pubs/mpi\\_report\\_dominican\\_pop\\_us.pdf](http://www.migrationpolicy.org/pubs/mpi_report_dominican_pop_us.pdf)).
- Marcelli, Enrico A. et al. 2009. *(In)Visible (Im)Migrants: The Health and Socioeconomic Integration of Brazilians in Metropolitan Boston*. San Diego, CA: Center for Behavioral and Community Health Studies, San Diego State University. Retrieved ([http://boston.com/bonzai-fba/Third\\_Party\\_PDF/2009/10/17/Marcelli\\_et\\_al\\_BACH\\_2009\\_Brazilian\\_\\_1255753970\\_2565.pdf](http://boston.com/bonzai-fba/Third_Party_PDF/2009/10/17/Marcelli_et_al_BACH_2009_Brazilian__1255753970_2565.pdf)).
- Marcelli, Enrico A. and Paul Ong. 2002. "Estimating the Sources of the 2000 Census Undercount among Foreign-Born Mexicans in Los Angeles County."
- Pastor, Manuel, Jared Sanchez, Rhonda Ortiz, and Justin Scoggins. 2013. *Nuturing Naturalization: Could Lowering the Fee Help?*. Los Angeles, CA: Center for the Study of Immigrant Integration, University of Southern California.
- Pastor, Manuel and Justin Scoggins. 2012. *Citizen Gain: The Economic Benefits of Naturalization for Immigrants and the Economy*. Los Angeles, CA: Center for the Study of Immigrant Integration, University of Southern California. Retrieved (<http://csii.usc.edu/CitizenGain.html>).
- Rytina, Nancy. 2013. *Estimates of the Legal Permanent Resident Population in 2012*. U.S. Department of Homeland Security; Office of Immigration Statistics.
- Slovic, Beth. 2008. "He's An... Illegal Eh-Lien." *Willamette Week*, February 20. Retrieved ([http://www.wweek.com/portland/article-8470-herss\\_an\\_illegal\\_eh\\_lien.html](http://www.wweek.com/portland/article-8470-herss_an_illegal_eh_lien.html)).
- Warren, Robert and John Robert Warren. 2013. "Unauthorized Immigration to the United States: Annual Estimates and Components of Change, by State, 1990 to 2010." *International Migration Review*. Retrieved April 15, 2013 (<http://doi.wiley.com/10.1111/imre.12022>).