MEASURING IMMIGRANT ASSIMILATION IN THE UNITED STATES

Second Edition

Jacob L. Vigdor
Professor of Public Policy and Economics, Duke University
Faculty Research Fellow, National Bureau of Economic Research

The year 2007 marked an economic turning point in the United States. According to the National Bureau of Economic Research, the nation's economic output peaked late in the year and then began to contract. This development affected immigration in two important ways: immigrants began arriving in fewer numbers than they have since the 1960s; and those immigrants who not only arrived but stayed fell further behind the native-born population economically. Economic assimilation declined even among immigrants who arrived more than a decade ago, indicating that differences between that cohort and the native-born population widened.

This report, the second in an ongoing series, takes advantage of newly released U.S. Census Bureau data from 2007 to measure changes in an index describing the state of economic, civic, and cultural assimilation of immigrants to the United States. It also explores in detail two of the factors used to compute the index: immigrants' English-language ability and naturalization rates, both of which have been affected by the reduced inflow and increased outflow of recent immigrants. Because legal adult immigrants who have been here less than five years cannot become citizens and are unlikely to have mastered English in so short a period, the economic downturn is having an effect on all three assimilation indexes: economic, of course; but also cultural assimilation, of which English skills are an important component; and civic assimilation, of which citizenship is an important component.

Ironically, the effect of the reduction in the numbers of immigrants arriving and staying has been to offset the impact on the assimilation index of gradually declining levels of English skills upon arrival and afterward as well as lower rates of naturalization. The reason for this is that recent arrivals differ most from natives, and thus their absence raises the collective assimilation index values of immigrants who have been here longer.

The Manhattan Institute introduced its first summary measures of immigrant assimilation in the United States in 2008. Civic Report No. 53, "Measuring Immigrant Assimilation in the United States," presented a series of index measures describing the degree of similarity between foreign- and native-born residents of the United States between 1900 and 2006. The index rises only when the foreign-born population becomes less distinct from the native-born. In net terms, there has been no change in the assimilation index between 2006 and 2007. The composite measure, which considers all three categories of indicators—economic, cultural, and civic—remained at the same level. None of those three separately showed any variation from 2006 to 2007.

Analysis of English-language skills among immigrants between 1900 and 2007 reveals several important patterns. The key findings are:

- **The proportion of non-English-speaking immigrants peaked in 1910.** In that year, nearly a third of all immigrants could not speak English (once again, excluding those born in English-speaking nations). Only 10 percent of immigrants fall in that category today, but another 20 percent report that their English skills, while existent, are poor—a category not included in the early Census enumerations.

- **About half of all immigrants report speaking English "very well," and this proportion has not changed much since 1980.** Between 25 percent and 30 percent of all immigrants report either that they do not speak English or that their English skills are poor. These statistics exclude immigrants from English-speaking nations.

- **Immigrant children acquire English skills much more rapidly than their parents.** More than 80 percent of immigrants who arrive in the United States by the age of seven speak English very well as adults, while only 30 percent of those who arrive as adults are eventually able to do so.

- **Recent immigrants—especially Mexican-Americans—are acquiring English-language skills more slowly than their predecessors.** In the early twentieth century, roughly 75 percent of immigrants who arrived without knowing English learned the language within twenty years. In more recent years, this proportion has moved closer to 60 percent. While the English skills of Mexican immigrants are lower than those of other immigrant groups, their apparent rate of progress is higher. But this finding might simply reflect the higher rates of return migration of those who would not be expected to have learned English.

- **As the immigrant population grows, the English skills of newly arrived immigrants tend to decline.** Such a pattern is evident in the early twentieth century and in more recent data. Specifically, the proportion of foreign-born residents who cannot speak English increased from 9 percent in 2000 to a peak of 11 percent in 2006; the proportion with poor English skills increased from 19 percent to 20 percent in that same period. Immigrants don't have as great a need to learn English when they have an extensive network of fellow immigrants on whom they can rely.

Analysis of naturalization rates between 1900 and 2007 reveals several important patterns:

- **English-language requirements do not deter would-be citizens.** In the past and the present, virtually all immigrants interested in pursuing citizenship have acquired sufficient command of the English language to meet the official standard.

- **Immigrants continue to value citizenship highly.** The evidence for this is the rates at which immigrants became citizens, which were approximately the same in the late nineteenth and late twentieth centuries. These figures were stable in spite of significant changes in naturalization law, many of which imposed longer waiting periods before a legal immigrant could become a citizen.

- **There is some evidence that naturalization rates of the most recent immigrant cohorts are slowing.** While this may reflect the high number of recent immigrants who are illegal, and thus ineligible for citizenship, it could also reflect the cumulative impact of longer waiting periods.

- **Mexican immigrants become citizens at a lower rate than other immigrant groups.**

- **The 1986 immigration amnesty appears to have had a moderately positive effect on naturalization rates but resulted in very little improvement in language skills.**

Updated information on the assimilation of immigrants in the United States, along with the detailed analyses of language acquisition and citizenship, can help distinguish the success stories in American immigration from the failures. The challenge of any attempt at immigration reform will be to preserve success while remediating failure.

CONTENTS

# Measuring Immigrant Assimilation in the United States

Jacob L. Vigdor

CHAPTER I: THE WHY AND HOW OF MEASURING ASSIMILATION

Is modern American society as good at integrating immigrants into the economic, cultural, and civic mainstream as it was in past generations? Are today's immigrants equally able and willing to partake in the process? These basic questions motivated our initial effort to measure immigrant assimilation in the United States, using a uniform method applied to data spanning more than a century, from 1900 to 2006. This chapter briefly reviews the methods used to measure the assimilation of immigrants in the United States. It also reviews the conclusions of the initial report on assimilation issued in May 2008.

The central motivation for studying assimilation in the United States is the massive increase in the immigrant population between 1970 and 2007, shown in Figure 1. Over this time period, the number of foreign-born residents of the United States roughly quadrupled, from under 10 million to nearly 40 million. This growth can be attributed in part to the relaxation of government restrictions on legal immigration in 1965. As Figure 1 shows, however, the expansion of the immigrant population accelerated over the past two decades, as the period of robust economic growth that lasted through much of the 1980s, the 1990s, and the first portion of the present decade brought migrants in search of opportunity.

Figure I. Foreign-Born Population of the United States, by Region of Birth, 1960-2007
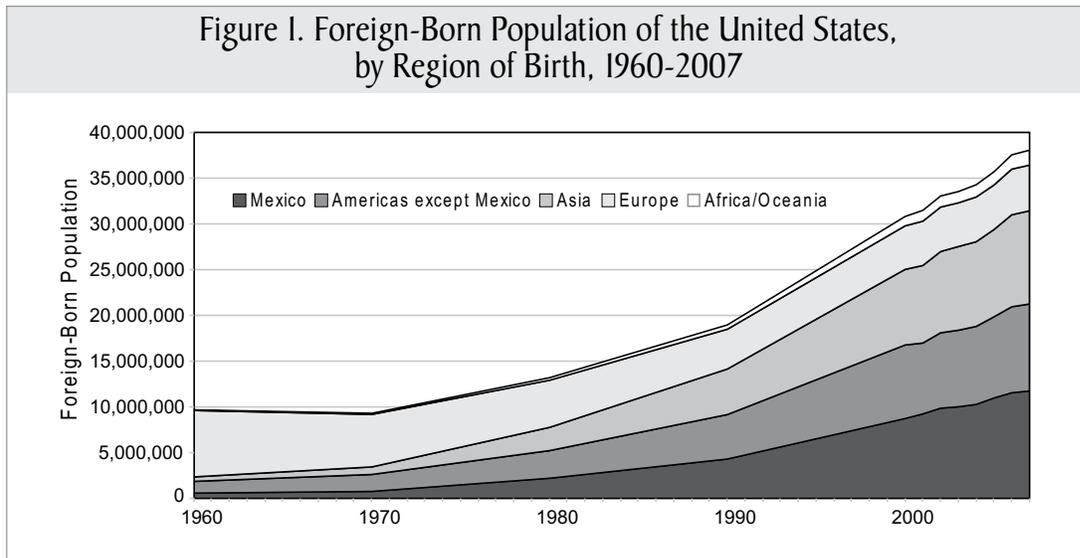
The very end of the time series plotted in Figure 1 shows evidence of a slowdown in the growth rate of the immigrant population. The average annual growth rate for the foreign-born population was in the range of 3 to 4 percent between 1970 and 2005; the Census Bureau-estimated growth rate from 2006 to 2007, by contrast, was only 1.4 percent. The net increase in the foreign-born population, according to Census Bureau estimates, was about 1.5 million in 2005 and 2.1 million in 2006; in 2007, the net increase was only 500,000. The Census Bureau intends to count only the number of residents in the United States and makes no effort to ascertain who and how many are residing in the nation legally.

The assimilation index tracks the integration of foreign-born individuals into the economic, cultural, and civic mainstream by measuring the ease with which nativity can be inferred on the basis of a variety of individual characteristics measured by Census Bureau surveys. A more complete description of the statistical procedure used to make these inferences can be found in Chapter 5 of this report. In brief, the 2007 index begins with a sample of more than a half million adults—evenly divided between native- and foreign-born—who participated in the American Community Survey (ACS).[1] The ACS is designed to be representative of the entire population of residents of the United States. While there are frequently voiced concerns that the ACS fails to sample certain segments of the population

adequately, including immigrants who are living or working in the country without legal documentation, the Census Bureau provides a set of statistical tools to address this concern.

We then use information collected in the ACS questionnaire to determine how easy it is to distinguish native- from foreign-born adults. This information can be divided into three sets: economic, civic, and cultural indicators of assimilation. The *composite* assimilation index uses all three sets to compare the native- and foreign-born. This report, like its predecessor, also discusses the *component* indices of economic, civic, and cultural assimilation.

Economic indicators include educational attainment, earnings, occupational prestige, employment status, and labor-force participation rates. In the case of the last four indicators, males and females are considered separately, since important and well-documented differences exist between them.

Civic indicators include citizenship and veteran status. In the case of the second indicator, males and females are considered separately, since military service is more common among males.

Cultural indicators include marriage to a foreign-born spouse, the number of children in an adult's household, the ability to speak English, and marital status.

Complete or "perfect" assimilation of the immigrant population would be deemed to have occurred when knowledge of each of these factors provides no information about whether an individual was born in the United States or abroad. In such a case, the assimilation index would take on a value of 100. In the opposite scenario, when this information is sufficient to correctly distinguish the native-born from the foreign-born in every case, the assimilation index would take on a value of zero.

In reality, the assimilation index tends to avoid these extremes. At all points in American history, the immigrant population has counted a mixture of individuals who blend in seamlessly as well as others who bear obvious marks of distinction from the native-born population.

The assimilation index can be computed for subsets of the immigrant population as well as for the immigrant population overall. For example, the index can be computed for immigrants born in a particular country, who arrived in the United States in a particular year, or who reside in a certain city or metropolitan area. When computed for a subset of the immigrant population, the index is a function of the proportion of individuals within that set correctly identified as foreign-born through the use of the basic algorithm described above.

A final alternative version of the index has been computed for members of "Generation 1.5," those persons who were born abroad but became residents of the United States by the time they were five years old. The index for Generation 1.5 is computed by drawing on information on adolescents and young adults, aged twelve to twenty-four, and relies on a more telling set of outcomes for that age group, including school attendance, English-speaking ability, residence with parents, parenthood, labor-force participation, and residence in a correctional facility or other nonacademic institutional setting.

The original assimilation-index report, released in May 2008, tracked the assimilation of immigrant groups in the United States between 1900 and 2006, using a combination of decennial Census Bureau data and the ACS. Among other things, the report noted the following patterns:

- By historical standards, the assimilation of immigrants in the United States in the early twenty-first century is low. The index fell during the 1980s and has remained at a persistently low level since 1990.

- Newly arrived immigrants are the least assimilated. Rapid growth in the immigrant population implies that the proportion of new arrivals is high. The low assimilation of immigrants in the early twenty-first century in part reflects this rapid growth. Assimilation progresses as time spent in the United States lengthens, and there is evidence that the assimilation rate overall is higher now than it was a century ago, during the last major wave of immigration to the United States.

- Assimilation varies substantially across national origin groups. Many immigrants born in developed nations are culturally and economically indistinguishable from the native-born. By contrast, immigrants from Mexico and nearby countries in Central America are, in general, quite distinct economically, culturally, and civically.

- One form of assimilation does not necessarily indicate another. Canadian immigrants are fully assimilated along cultural and economic dimensions, but their civic assimilation is not pronouced. Immigrants from Vietnam have very high levels of civic and economic assimilation but retain cultural distinctiveness. Immigrants from Mexico show low levels of economic and civic assimilation, quite possibly because a substantial proportion lack the legal right to live and work in the United States, but show cultural-assimilation levels similar to those of other groups.

The remainder of this report has three purposes. The first is to update the time series on immigrant assimilation in the United States with information from the 2007 American Community Survey. The second is to expand on the study of cultural assimilation by examining one characteristic used to compute it—the ability to speak English—in greater detail. As noted in the original index report, the Mexican and Vietnamese immigrant populations are very distinct from one another in terms of economic and civic assimilation, but receive similar

scores on the cultural scale. This could indicate that the English-language acquisition process is similar in both groups. Alternatively, large differences in the language facility of these groups may exist, but may be offset by other cultural patterns, such as the contrasting proportions of American-born spouses, since the cultural assimilation index uses information on both language and intermarriage. Evidence presented below will show the latter explanation to be more consistent with the data.

The final section of this report provides a detailed study of naturalization rates over the past century. The path to citizenship today is more difficult than it was a century ago. Before the early 1920s, white immigrants of any nationality could arrive in the United States and immediately join a five-year queue for citizenship. Although the twentieth century saw the eventual elimination of racial restrictions, the introduction of legal permanent residency as a way station on the road to citizenship increased the expected waiting period for most aliens.

In spite of this additional hurdle, naturalization rates for modern immigrants are in fact quite similar to those evidenced by their historical predecessors. Important differences in the numbers of applications for citizenship exist among nationalities; these reflect a combination of U.S. policy preferences and the motivation of the immigrants themselves.

Finally, English-language requirements for citizenship were established in 1906 and weakened later in the century. There is little or no evidence that these requirements have been of any consequence. Virtually every immigrant interested in becoming a citizen learns English, even if doing so is not required.

CHAPTER 2: ASSIMILATION IN 2007

The deceleration in growth in the immigrant population, documented in Figure 1, coupled with the fact that newly arrived immigrants tend to be the least assimilated, might suggest that the assimilation index should have increased between 2006 and 2007. Instead, as shown in Figure 2, the composite index and its three components did not change between 2006 and

2007. The composite index continues at its prior level of 28, cultural assimilation at 62, civic assimilation at 41, and economic assimilation at 87. The composite index has not changed since 2001, and the civic and economic indices have not changed since 2004 and 2003, respectively. Only the cultural index has shown evidence of a trend over the past few years, having increased since 2002, when its value was 60.

The absence of change in the assimilation index between 2006 and 2007 is not altogether a surprise. Figure 3, which shows the entire time path of the assimilation index from 1900 to 2007, indicates that there has been little change overall in the index since 1990. Even in a period of noteworthy change in the assimilation index, between 1980 and 1990, the annualized average change in the index came to less than


Figure 2. The Assimilation Index: 2007
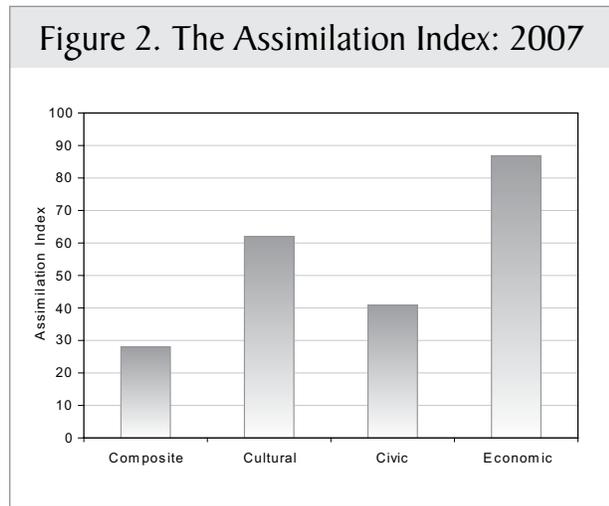
Figure 3. The Assimilation Index: 1900-2007

The Census Bureau couldn't collect sufficient data to compute the index for the years 1930-1970.
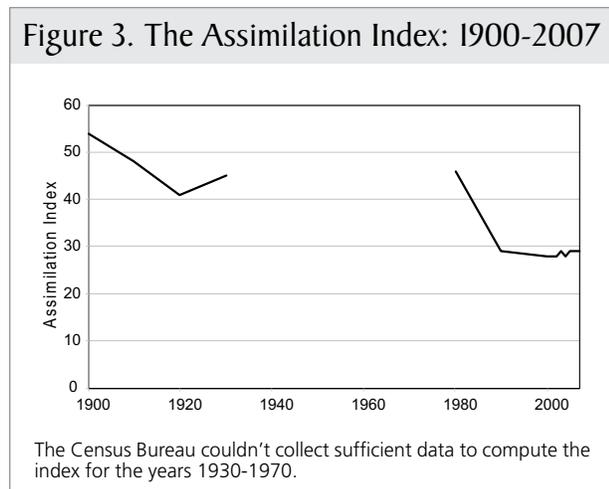
two points. A countervailing factor helps explain the failure of decreased growth in the immigrant population to increase assimilation: the economic slowdown that was at least partly responsible for the decline in the rate of immigration hurt immigrants more than the native-born population.

The original assimilation index found dramatic differences in assimilation among country-of-origin groups in 2006. By and large, these differences persisted in 2007. Before discussing groups' changes over time, it is worth discussing the methodological issues associated with inferring trends from data sets such as the ACS. The ACS interviews a small fraction of the U.S. population in any one year and draws an entirely new sample of respondents each year. Thus, neither the assimilation index nor any other summary measure perfectly reflects the experiences of a distinct set of individuals over time. Differences in the index from one year to the next may reflect the fact that a different set of persons was interviewed, and not that the experiences of any one set of individuals changed over time. This possibility is most troubling in the case of those groups with the fewest members. Each group discussed in this section was represented by at least 4,000 adults in the 2007 ACS. Table 3 in Chapter 5 reports index values for groups represented by as few as 100 adults in the 2007 ACS. Increasing degrees of caution must be used in interpreting trends in the assimilation index as the groups being examined diminish in size.

Figure 4 shows assimilation-index values for the ten countries of origin with the greatest numbers of representatives in the U.S. population in 2007. Of these ten groups, immigrants from Canada are the most assimilated, with an index value of 54 in 2007, up slightly from 53 in 2006. Immigrants from Mexico, the largest single country-of-origin group, post a 2007 index value of 13, identical to the 2006 value. Of the eight other large country-of-origin groups, four show no change in assimilation between 2006 and 2007. Immigrants from China and Vietnam show modest increases in assimilation; those from South Korea and Cuba show modest decreases. Just as there is little overall change in the index between 2006 and 2007, there is little evidence of uniform movement in one direction or the other by the largest country-of-origin groups.

Figure 4. Assimilation by Country of Origin: 2007
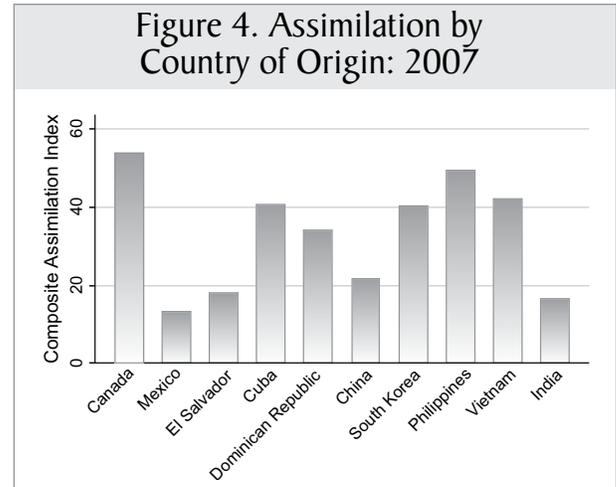
Figures 5, 6, and 7 examine the component indexes of cultural, economic, and civic assimilation for the largest country-of-origin groups in 2007. Once again, there is not much evidence of significant change between 2006 and 2007. The component indexes are entirely unchanged for immigrants from Mexico, the largest country-of-origin group.

Four of the largest country-of-origin groups experienced a decline in cultural assimilation, two showed an increase, and the remaining four showed no change. Because cultural assimilation is a slow process, the recent drop in immigrant arrival rates has not lifted its index value. It is conceivable that economic uncertainty reduces immigrants' incentive to assimilate culturally; if the likelihood of moving away from the United States increases, the potential gain from learning English or marrying a native-born spouse declines.

The majority of large groups registered no increase in economic assimilation, with only two of the ten increasing, two decreasing, and six remaining the same. The small number of increases reflects the fact that increases were impossible for the five groups exhibiting perfect or nearly perfect assimilation in 2006.

Civic assimilation shows an increase from 2006 to 2007 for six of the ten largest groups; only Cuban immigrants show evidence of a decline in civic assimilation. The gain is consistent with the overall trend toward higher civic assimilation witnessed over the past few years and may also reflect the slowdown in the arrival rate

## Figure 5. Cultural Assimilation by Country of Origin: 2007
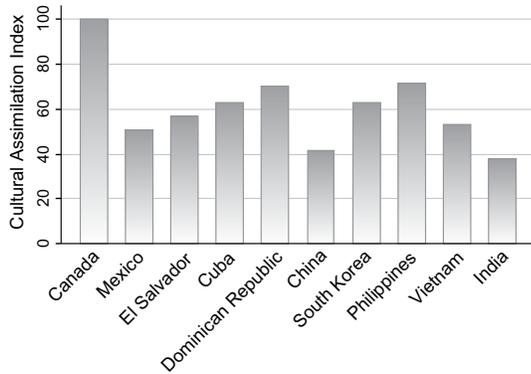
Figure 6. Economic Assimilation by Country of Origin: 2007
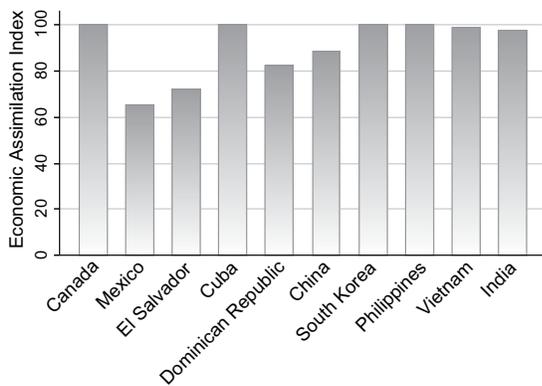
Figure 7. Civic Assimilation by Country of Origin: 2007
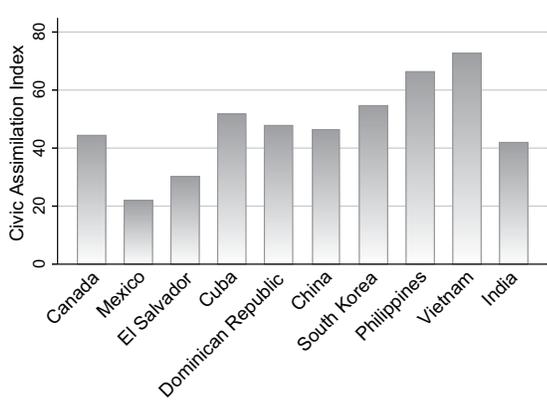
of new immigrants. Newly arrived immigrants are generally ineligible for immediate naturalization, the primary indicator of civic assimilation.

Overall, from 2006 to 2007, individual groups show little evidence of meaningful change in any of the assimilation categories. Interpretation of any individual movement should be undertaken with caution, as it may reflect changes in the set of individuals participating in the ACS rather than true changes in social conditions.

A similar caveat applies to the examination of trends in assimilation within destination metropolitan areas between 2006 and 2007. Figure 8 shows composite index values for the ten metro areas with the largest number of adult immigrants participating in the 2007 ACS. While 2007 assimilation-index values in these areas are similar to those from 2006, changes in them have a more recognizable pattern. The assimilation index declined in the four largest immigrant destination areas: Los Angeles, New York, Chicago, and Washington, D.C. Of the six remaining areas in Figure 8, four represent smaller California cities. A fifth California city, San Diego, was displaced from the list of the ten largest immigrant destinations. In four of these five smaller California cities, the assimilation index increased from 2006 to 2007.

Assimilation also increased in two other centers of Latin American and Caribbean immigration: Miami and Houston. The tendency of assimilation to decline in larger multiethnic centers of immigration, while increasing elsewhere, may reflect differential reactions to

Figure 8. Assimilation by Metro Area: 2007
Ten largest immigrant destinations in rank order

Figure 9. Assimilation by Years in the U.S.: 2007
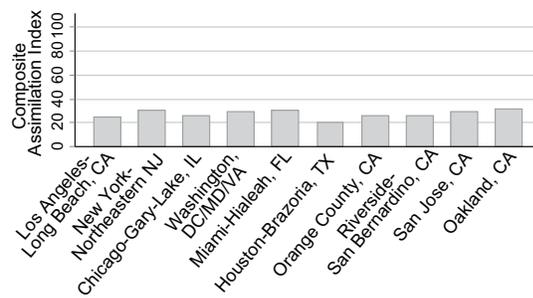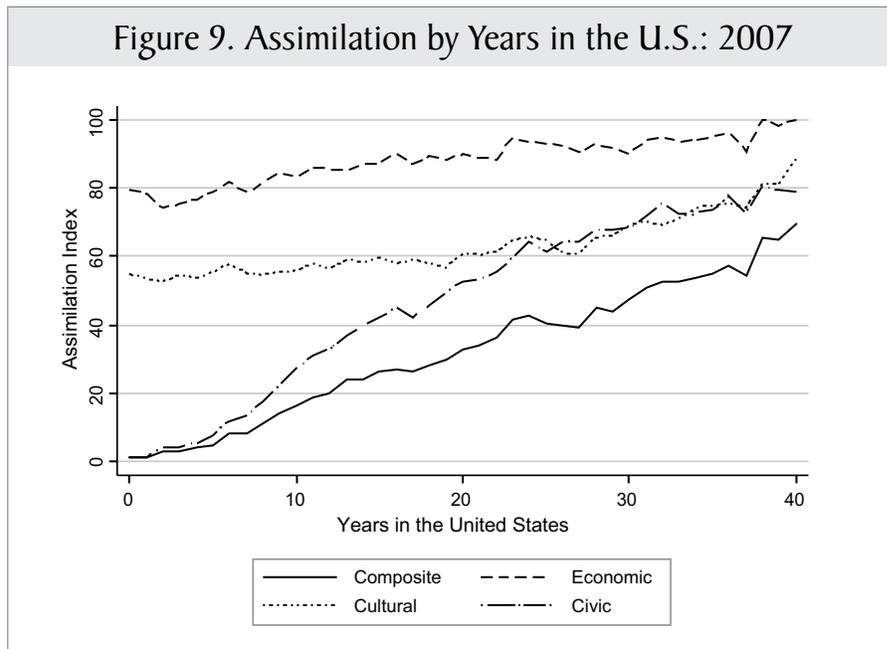
the nascent economic slowdown in 2007. Immigrants from nearby nations may have responded by returning to their country of origin; the flow of potential migrants to these areas may have declined as well. Immigrants with less feasible return options, in turn, may have remained and thus borne the full brunt of the downturn. Metropolitan areas on the fringes of the Los Angeles and San Francisco Bay areas, where immigrants have settled in large numbers, were among those most profoundly affected by the downturn in the housing market in 2007.

The assimilation index provides a snapshot of the degree of similarity between this country's native- and foreign-born populations at a single point in time. Assimilation itself is a process that takes place over time. Figure 9 sheds some light on the nature of this process, showing the composite and component assimilation indexes for immigrants in 2007 to be a function of the number of years since their arrival in the United States. The civic assimilation index for newly arrived immigrants is close to zero, largely because most foreign-born residents of the United States are not instantly eligible to become naturalized citizens. Immigrants with a longer history of residence in the United States have substantially higher civic-assimilation index values. There is also some evidence of assimilation along

the economic and cultural dimensions. Economic assimilation appears to occur slowly and steadily over time, to the point where adult immigrants with over three decades' residence in the United States appear economically indistinguishable from natives.

The cultural assimilation index is virtually identical for newcomers and even those immigrants with as many as twenty years' residence in the United States; those here beyond twenty years show some evidence of higher assimilation-index values. In many cases, adult immigrants with more than twenty years' residence in the United States arrived in the country as young children. These Generation 1.5 immigrants are generally difficult to distinguish from the native-born population.

Figure 10 compares the relationship between assimilation and years in the United States for 2007 and 2006. While there has been very little change from one year to the next, a pattern that is consistent with the stability of the overall index value, note that where the two lines diverge, the 2006 series is almost always higher than the 2007 series. Older cohorts of immigrants, in particular, appear to have made weak progress toward assimilation from 2006 to 2007. This pattern will be confirmed in the following analysis of the progress of individual cohorts over time.

Figure 10. Composite Assimilation by Years in the U.S.

Figure 11. The Progress of Individual Cohorts: Composite
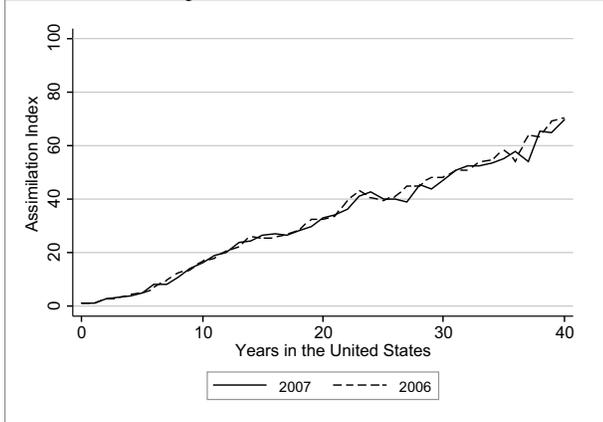
Immigrants who have spent more time in the United States may appear more assimilated for two reasons. First, assimilation is a process that takes time. Second, immigrants of a generation ago may have always been different from the newly arrived immigrants of today. Figures 9 and 10 thus cannot prove that immigrants assimilate over time. Fortunately, the availability of Census Bureau and ACS data at multiple points in time permits a more direct analysis of the progress of individual cohorts as their time in the United States lengthens. However, it should not be forgotten that changes in the index for a cohort of immigrants may occur either because those immigrants change relative to the native-born population or because a select group of that cohort elects to leave the country.

Figure 11 shows the evolution of the composite assimilation index over time for groups of foreign-born individuals arriving during four different five-year intervals. For each cohort, the index has tended to increase over time as a result of some combination of selective onward migration and actual changes in the circumstances of the foreign-born. Immigrants arriving in the late 1970s, for example, had a collective assimilation-index value of 5 in 1980. This value had increased to the low 20s by 1990, and to 40 by 2000. Cohorts arriving in later periods of time have tended to start with lower index values—this is one explanation for the decline in the overall index after 1980. The progress of these later cohorts over time resembles that of the late-1970s cohorts, however. Each cohort

Figure 12. The Progress of Individual Cohorts: Economic
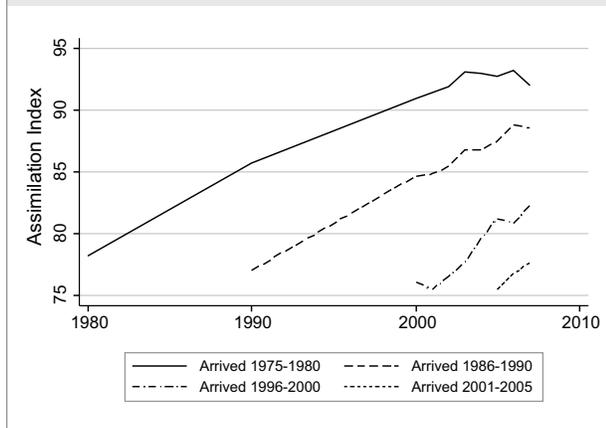
posts, or appears on track to post, increases on the order of 15 points in its first decade.

Figure 11 also shows, however, that increases in assimilation for the oldest cohort of immigrants have stalled in the past few years. There is similarly some evidence that the progress of cohorts arriving in the late 1980s and early 2000s is tepid when compared to the progress of other cohorts at other points in time. Figures 12 through 14 expose this pattern in greater depth by presenting similar charts for the three component indexes.

Figure 12 begins by charting the economic assimilation index for the four entry cohorts in Figure 11. While a

Figure 13. The Progress of Individual Cohorts: Cultural

Arrived 1975-1980 ——— Arrived 1986-1990 - - - - -
Arrived 1996-2000 -·-·-·- Arrived 2001-2005 ··········

Figure 14. The Progress of Individual Cohorts: Civic

Arrived 1975-1980 ——— Arrived 1986-1990 - - - - -
Arrived 1996-2000 -·-·-·- Arrived 2001-2005 ··········
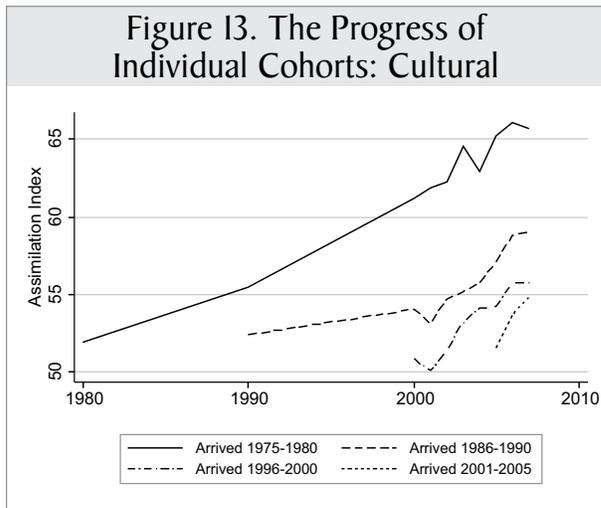
1970s were significantly more likely to be working a full fifty-two weeks per year than either the native-born or immigrants arriving in the early 2000s: about half of the older immigrants worked a full year, but only 33 percent of recent immigrants and 31 percent of natives did (a substantial number of workers report working fewer than fifty-two weeks per year). In general, immigrants are more likely to participate in the labor force than natives. At the time of their 2007 ACS interviews, however, the labor-force advantage of the older immigrants had disappeared—33 percent reported being unemployed or to be no longer seeking work at the time of their interview, the same proportion as in the native population. Younger immigrants continued to show a small advantage, with only 30 percent of them unemployed or gone from the labor force.

Why have the older cohorts of immigrants suffered to a greater extent? Although they have spent over a decade in the United States, these individuals may find themselves in economically marginal positions. Of course, younger immigrants might be expected to fare even worse, but, by virtue of their shorter stays in the United States, most will have probably put down fewer roots. Consequently, they may be more likely to go to another city in search of work. Younger immigrants may also be more likely to work in less cyclical sectors of the economy, such as education and health services.

pattern of strong increases in the early years of residence appears for all cohorts, there is striking evidence of a decline in the fortunes of older immigrants from 2006 to 2007. The decline is largest for the cohort arriving in the late 1970s but is also apparent among those arriving in the late 1980s. Younger cohorts, by contrast, continue a pattern of improvement from 2006 to 2007. Since the analysis excludes individuals over age sixty-five, retirement at the typical age cannot explain why older cohorts have suffered economically more than the younger.

A more detailed analysis of the factors underlying economic assimilation corroborates the impression that the fortunes of older immigrants declined from 2006 to 2007. In 2006, immigrants arriving in the late

Figure 13 shows that older cohorts experienced a decline in cultural assimilation as well. A similar break in trend can be observed in cohorts arriving as recently as the late 1990s; the most recently arriving cohorts, previously noted to be on a much more rapid trajectory than their predecessors, show continued evidence of progress through 2007. At first, it may seem peculiar that an economic downturn would have an impact on the indicators of cultural assimilation. In many ways, however, cultural assimilation is a form of investment, with up-front costs and returns that accrue only over time. Taking an English course, for example, is both costly and time-consuming. When immigrants fear that poor economic conditions will cause them to return home, they have little incentive to bear the cost, as they do not expect to have time to reap the benefit.

There is comparable evidence of a drop in the cultural assimilation of younger immigrant cohorts during the 2001 recession. Once again, the experiences of immigrants in 2008 will provide additional valuable information about the impact of economic conditions on cultural patterns.

If overall assimilation did not change for older cohorts from 2006 to 2007, while both cultural and economic assimilation declined, it stands to reason that civic assimilation improved. Figure 14 confirms the soundness of this inference. The civic-assimilation index for each of the four depicted cohorts is higher in 2007 than 2006. Once again, however, there is evidence that upward trends have moderated for the oldest cohorts. Just as cultural assimilation can be viewed as reflecting a recognition of the economic benefits of learning English, for example, the decision to become a naturalized citizen presumably reflects a recognition of the economic benefits of acquiring a permanent right to live and work in the United States. Immigrants may have expected these latter benefits as well to decline in the face of an economic downturn.

Historically, the foreign-born children of immigrants have assimilated more rapidly than their parents. Figure 15 shows the time path of an alternative assimilation-index measure, calculated for foreign-born individuals who arrived in the United States by the age of five. The assimilation index for this group, commonly called Generation 1.5, reflects the experiences and behaviors of individuals between the ages of twelve and twenty-four. Except for the fact that these individuals do not automatically become U.S. citizens at the moment of birth, they are very difficult to distinguish from the native-born. An assimilation index computed without citizenship information has been consistently above 90 in every year from 1900 to 2007. There is some evidence, however, of a decline in assimilation among the members of this group, perhaps attributable to the same set of economic forces acting on their parents. This downturn is also observable in the alternative assimilation index, which incorporates citizenship.

In summary, the economic downturn that began in 2007 affected the experience of immigrants, both first and second generation, in two ways. First, it slowed their net flow into the United States by discouraging some from arriving and impelling some already in the U.S. to depart. Second, there is evidence that the average immigrant began to suffer the ill effects of the downturn before the native-born citizen and that this economic suffering retarded the cultural and civic progress of immigrants as well.

On net, these two effects combined to produce little change in the assimilation index, as the losses experienced by long-term immigrants were offset by the reduced presence of less assimilated new arrivals.

Figure 15. The Assimilation of Generation 1.5
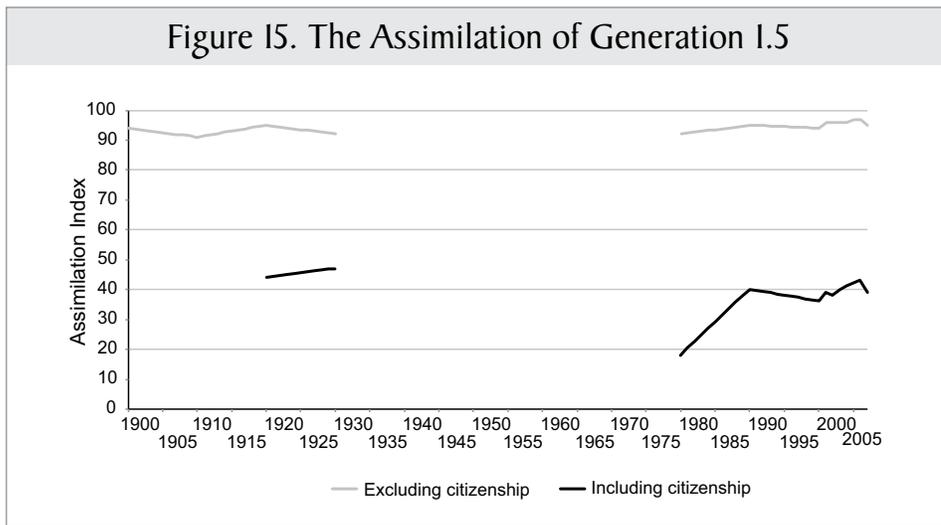
CHAPTER 3: SPOTLIGHT ON ENGLISH-LANGUAGE ACQUISITION

The index of cultural assimilation incorporates information on English-language ability, marital status and intermarriage, and childbearing patterns. As shown in Figures 9 and 13 above, cultural assimilation appears to occur less rapidly than civic and economic assimilation. Over their first decade in the United States, past cohorts of immigrants have posted 30 point increases in civic assimilation and 10 point increases in economic assimilation, but gains of only 2 to 5 points in cultural assimilation. Immigrants arriving within the past decade appear to be on a different trajectory.

Language plays a central role in current debates over immigration policy. While there are some examples of successful multilingual societies around the world, economic theory suggests that language barriers are costly, and economists have found considerable evidence to support this view.[2] Populist efforts to make English the official language of government activity have taken root in many parts of the country. Have the collective English-language skills of immigrants declined noticeably over time? If so, is it because immigrants don't acquire English as quickly as they had, or because today's foreign-born population contains such a high proportion of recently arrived immigrants? If not, are there any other warning signs that might justify popular concern?

Census Bureau questionnaires, including the ACS, have collected information about English-language ability since 1900. Through the first decades of the twentieth century, when the census was conducted by enumerators in face-to-face interviews, a single yes-or-no answer recorded whether an individual spoke English. This determination was ultimately the census enumerator's. In more recent years, most individuals have filled out the census questionnaire on their own and returned it by mail, requiring the Census Bureau to trust their own assessment of their English skills. The measurement of these skills, however, has become more informative over time, with individuals distinguished by whether they report that they speak English at home, and if they say they do not, whether in their judgment they speak English "very well," "well," "not well," or "not at all."

Figure 16 presents information on the English-language ability of immigrants over the period 1980-2007. Immigrants from nations where the predominant spoken language is English are excluded from the information in this figure and all figures in this chapter as well as the accompanying analysis.[3] Figure 16 shows evidence of a decline in English skills between 1980 and 1990, a period when the assimilation index itself also declines, and relative stability through the 1990s, with slight evidence of worsening since 2000. Between 2000 and 2006, the proportion of immigrants speaking no English rose from 9 to 11 percent, and the proportion speaking

Figure 16. Immigrant English Ability, 1980-2007

Legend:
- Speaks English at Home
- Speaks English Very Well
- Speaks English Well
- Speaks English, but Not Well
- Does Not Speak English

Excludes immigrants born in predominantly English-speaking nations
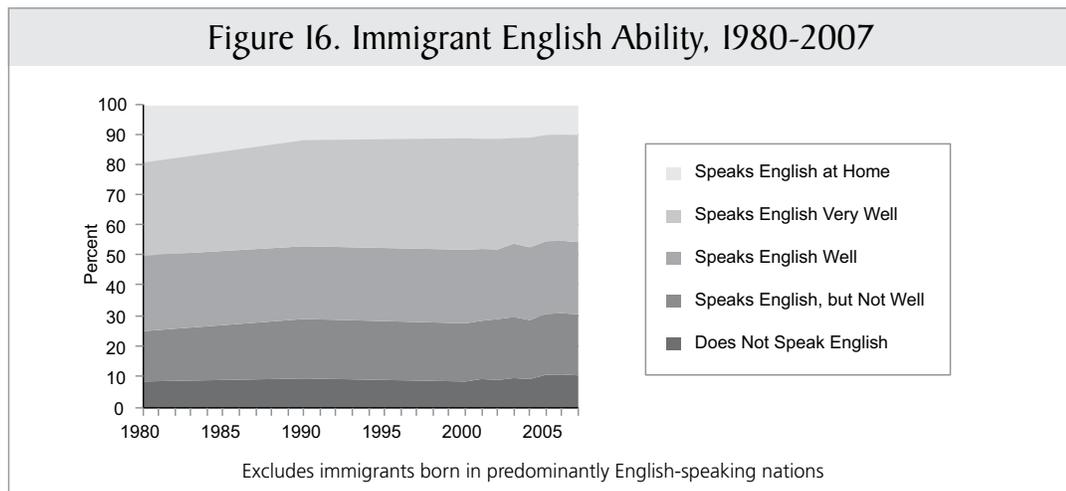
Figure 17. Immigrant English Ability, 1900-1930

Excludes immigrants born in predominantly
English-speaking nations
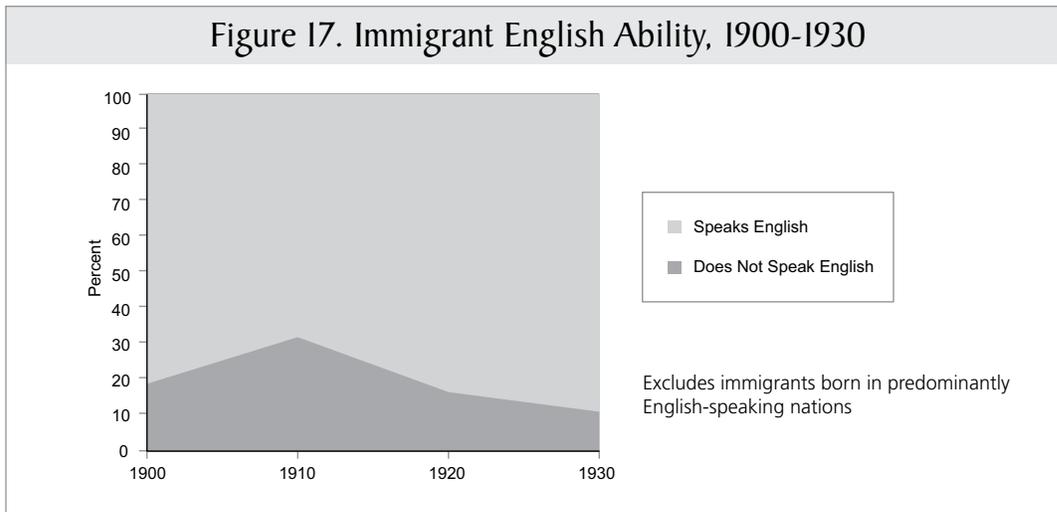
English poorly rose from 19 to 20 percent. In spite of these slight trends, the clearest picture in Figure 16 is one of stability. During a period when the immigrant population tripled, the proportion of immigrants from non-English-speaking nations who either speak English at home or speak English "very well" has held steady, at around 50 percent.

How do these patterns compare with those of the early twentieth century? Figure 17 provides the best information available for that period, tracking census enumerators' reports of whether the immigrants they interviewed could speak English. The proportion of non-English-speaking immigrants appears to have been larger in this earlier era, with a pronounced spike in 1910, when nearly a third of the immigrant population lacked the ability to speak English. After World War I and legal restrictions cut off the flow of new immigrants into the country, the linguistic skills of immigrants improved; by 1930, the proportion of non-English-speaking immigrants has approached its modern value of around 10 percent.

It is difficult to compare these two sets of information, collected by census enumerators using very different methods. At face value, the English-speaking ability of the non-Anglophone immigrant population appears to be considerably better than it was a century ago. It is possible, however, that many of the immigrants who report themselves to be poor English-speakers would have been labeled nonspeakers by census enumera-

tors. Nonetheless, we have uncovered no indication that the English-language skills of the immigrant population have deteriorated rapidly.

While the aggregate statistics show little cause for alarm, it is entirely possible that patterns vary considerably across cohorts of immigrants, or among individual immigrant groups. One possibility is that speakers of relatively uncommon languages have made rapid progress in learning English out of economic necessity, while immigrants from Spanish-speaking nations face less pressure to add to their linguistic skills. Another is that the aggregate statistics mask a combination of rapid progress among older immigrants and the poor English skills of new immigrants.

Figure 18 presents information on the progress in English-language acquisition of four cohorts of immigrants—the same four used to produce Figures 11–14 above. Each of the four panels in Figure 18 takes a single cohort and tracks the changing proportion of that cohort in each of the five English-ability categories of the census over time. Linguistic progress, to the extent that it occurs, would lead the lower categories on the graph to take up less of the vertical space over time.

Each cohort of immigrants, whether arriving in the late 1970s, early 2000s, or at any point between, shows some evidence of language acquisition over time. For the older cohorts, progress is most evident in the first

Figure 18

English Ability of 1975-1980 Arrivals Over Time

English Ability of 1986-1990 Arrivals Over Time

English Ability of 1996-2000 Arrivals Over Time

English Ability of 2001-2005 Arrivals Over Time

Speaks English at Home Speaks English Very Well
Speaks English Well Speaks English, but Not Well
Does Not Speak English

Excludes immigrants born in predominantly English-speaking nations

decade, when the proportion of immigrants with no or limited English skills declines and the proportion with very good English ability increases. Few immigrants, even after multiple decades of residence in the United States, adopt English as the language that they speak at home.

There is some evidence in Figure 18 that the most recent cohorts of immigrants arrive with poorer English skills than their predecessors. The proportion of immigrants arriving without the ability to speak English has increased over time, from 17 percent in the late-1970s cohort to 22 percent in the early-2000s cohort. In this most recent cohort, nearly 50 percent of all immigrants from non-Anglophone nations arrive with, at best, poor English skills, an increase of five percentage points over the late-1970s cohort.

Figure 19 repeats the strategy of Figure 18, following the linguistic progress of immigrant cohorts over time but focusing on the immigrants of a century ago. Like Figure 17, Figure 19 is restricted to analyzing only

whether census enumerators coded individuals as speakers or nonspeakers.

In this earlier era, the English-language skills of newly arrived immigrants appear to be much worse than they have been over the past few decades. Less than 50 percent of non-Anglophone immigrants arriving between 1896 and 1900 spoke English in 1900, and less than 40 percent of such immigrants arriving between 1906 and 1910 spoke English in 1910. Even allowing for the possibility that many immigrants who now report that they speak English poorly would have been classified as nonspeakers if interviewed by an enumerator, these figures are substantially worse than those depicting more recent immigrants.

Equally striking, however, is the rate at which members of these cohorts made progress as the time they spent in the United States lengthened. In 1920, 80 percent of the 1906-10 arrival cohort and 85 percent of the 1896-1900 cohort were coded as speaking English. To be precise, this process may not reflect immigrant prog-

Figure 19

English Ability of 1896-1900 Arrivals Over Time

English Ability of 1906-1910 Arrivals Over Time

Speaks English   Does Not Speak English

Excludes immigrants born in predominantly English-speaking nations
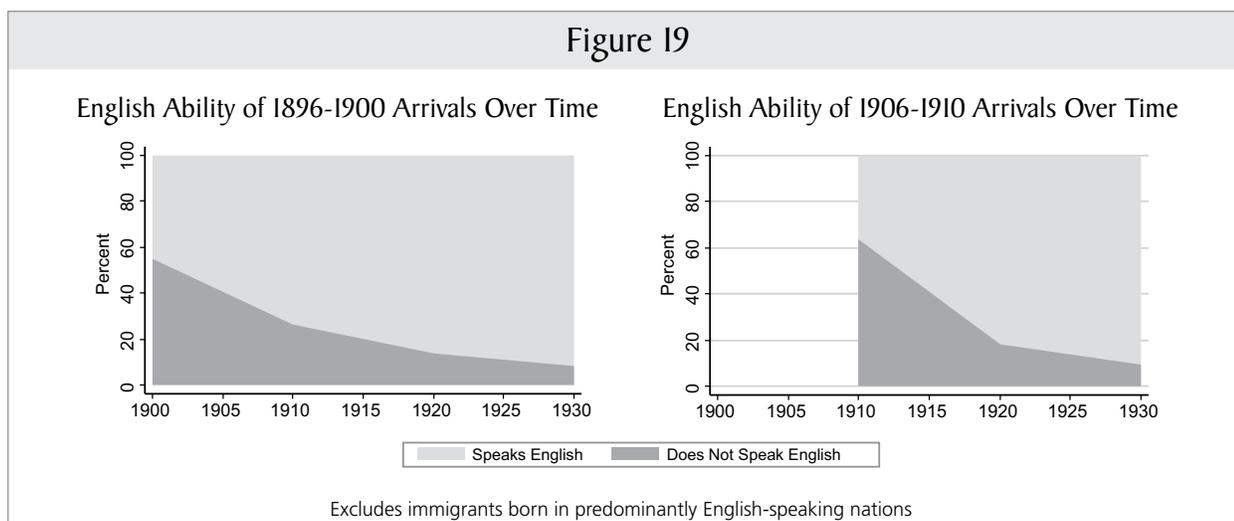
ress entirely—non-English-speakers may have been more likely to return to their respective homelands before 1920. Nonetheless, the rapid transformation of these cohorts is remarkable. If we put aside the possible impact of return migration for a moment, these statistics indicate that about 75 percent of the non-English-speakers in the 1896-1900 cohort acquired English skills within twenty years. By comparison, the proportion of non-English-speaking members of the 1975-80 cohort who acquired the ability to speak English over the following twenty years (once again, setting return-migration concerns aside for a moment) was only 58 percent.

How would return migration change this picture? Return migration was much costlier for the immigrants from a century ago, who came largely from overseas and, in some cases, knew that they would face war or persecution upon their return home. The option of returning is thus much more available in the present era to those migrants who have difficulty assimilating. It follows that the present cohort of immigrants is more likely to lose those members whose lack of English ability would have dragged cultural-assimilation values down if they had stayed. Altogether, then, there is substantial evidence that immigrants of a century ago, in spite of arriving with poorer English skills, acquired them at a significantly more rapid rate.[4]

As a final note, Figure 19 replicates the pattern observed among more recent immigrants: that those

arriving later in a wave of immigration have poorer English skills than their immediate predecessors. This is entirely logical: the first immigrants from a particular country of origin are those who anticipate fitting in rapidly with the native majority. Because later cohorts rely on the ethnic group-specific networks that their predecessors set up, they have less incentive to acquire English-language skills. As immigration from one nation or linguistic group accelerates, the average English-language skills of the group will tend to deteriorate. Continued growth in the immigrant population also reduces the pressure on the early waves of immigrants to learn English, as linguistic enclaves develop around them. The more rapid growth in the immigrant population in recent decades than occurred in the early twentieth century might explain why the rate of English-language acquisition appears lower now than the rate then.

Do the conclusions about the relative English skills of newly arrived immigrants, and the rate of language acquisition over time, hold equally for all immigrant groups? Figure 20 presents information similar to that in Figure 18 but focuses specifically on the English-language skills of immigrants born in Mexico. Recent cohorts of Mexican immigrants have arrived in the United States with poorer English skills than the rest of the non-Anglophone immigrant population. In the earliest cohort of immigrants, arriving in the late 1970s, more than a third did not speak English, and another third spoke English poorly. This group reported some

Figure 20

English Ability of 1975-1980 Arrivals from Mexico

English Ability of 1986-1990 Arrivals from Mexico

English Ability of 1996-2000 Arrivals from Mexico

English Ability of 2001-2005 Arrivals from Mexico

Speaks English at Home · Speaks English Very Well · Speaks English Well · Speaks English, but Not Well · Does Not Speak English

Excludes immigrants born in predominantly English-speaking nations

progress over time, to the point where more than half now report speaking English "well" or "very well." As is the case with the immigrant population as a whole, most of this progress occurred in the first decade of residence in the United States.

The cohorts of Mexican immigrants arriving in the late 1980s and late 1990s display trajectories similar to that of the first cohort, with roughly two-thirds of new arrivals having, at best, poor English skills and showing some evidence of progress over the first decade. The English skills of the most recent cohort, arriving in 2001 or later, are worse than those of earlier cohorts. Some 40 percent of this cohort arrived without the ability to speak English, and another 33 percent reported poor English skills. Between 2005 and 2007, this group showed evidence of progress at a rate comparable with that of earlier cohorts. The popular perception that the English-language skills of the nation's largest foreign country-of-origin group have declined is thus supported by the data.

While the English skills of Mexican immigrants are lower than those of other immigrant groups, their apparent rate of progress is higher. If we ignore the issue of return migration, then 63 percent of non-English-speakers in the late-1970s birth cohort acquired the ability to speak English by 2000. For Mexican immigrants, though, the phenomenon of return migration is particularly important, given their nation's proximity to the United States and the porousness of the border between the countries. Thus, the higher apparent rate of English-language acquisition for Mexican immigrants could be a function of rates of return migration.

Figure 21 examines a second country-of-origin group singled out in the initial assimilation-index report as having an assimilation experience quite different from that of Mexican-born immigrants: The group from Vietnam, who can be distinguished from Mexicans along a number of dimensions. Whereas the primary motivation for Mexican immigration is economic advancement, many Vietnamese arrived as political

---

Figure 21

English Ability of 1975-1980 Arrivals from Vietnam
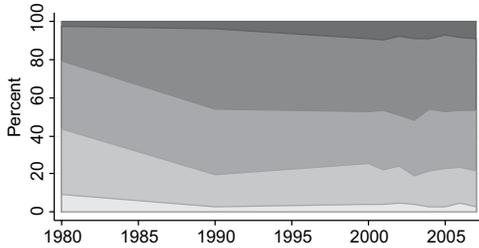
English Ability of 1986-1990 Arrivals from Vietnam
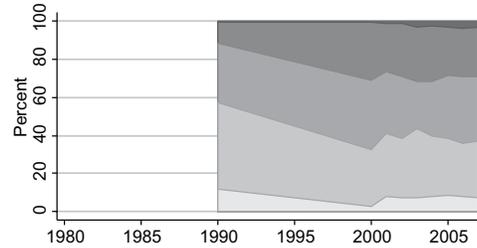
English Ability of 1996-2000 Arrivals from Vietnam
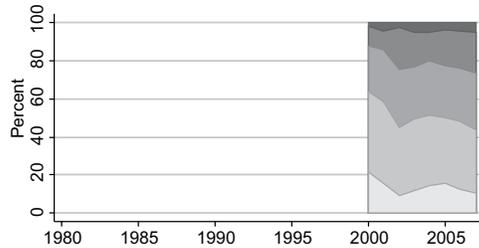
English Ability of 2001-2005 Arrivals from Vietnam
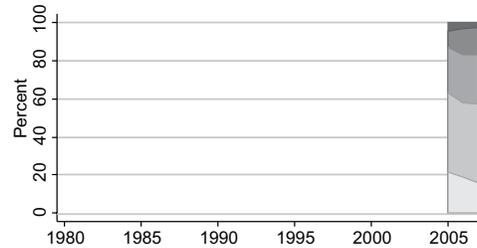
■ Speaks English at Home	■ Speaks English Very Well	
■ Speaks English Well	■ Speaks English, but Not Well	
■ Does Not Speak English		

Excludes immigrants born in predominantly English-speaking nations

---

refugees. Mexico is adjacent to the United States; Vietnam is almost as far away as a country of origin can be. Mexico is the source of the single largest country-of-origin group and shares a language with several other groups present in substantial numbers in the United States. Although Vietnam is one of the ten largest sources of immigrants, there are more than ten Mexican immigrants for every U.S. resident born in Vietnam. Moreover, while the Vietnamese language overlaps with Cantonese to a slight extent, it belongs to the relatively obscure Austro-Asiatic family of languages, of which it is the most commonly spoken. Each of these factors implies that immigrants from Vietnam have had stronger incentives to learn English and otherwise integrate themselves into the American mainstream, and the original index report found substantial evidence that these differences in incentives translated into differences in assimilation.

Among the more striking differences between Mexican and Vietnamese immigrants, which has been present in every cohort, is the latter's higher likelihood of speaking English upon arrival in the United States. Although virtually no immigrants born in Vietnam speak English in their household, the proportion with no knowledge of English upon arrival is never higher than 25 percent and was actually less than 10 percent in the earliest cohort. Like the members of other immigrant groups, individuals born in Vietnam show evidence of significant learning in their first decade in the United States. The proportion of late-1970s arrivals speaking English very well increased from 21 percent in 1980 to 47 percent in 2007. It makes little sense to track the progress of non-English-speakers in this cohort, since there were so few of them. Relative to Mexican immigrants, those from Vietnam arrived with a linguistic advantage and maintained that advantage over time.

It is also clear, however, that the English skills of more recent cohorts are poorer than those of their predecessors. Whereas less than 10 percent of the late-1970s cohort arrived without knowing English, 21 percent of post-2000 arrivals could not speak English in 2005. As was the case in the sample of immigrants

Figure 22. English Fluency and Age at Immigration for Foreign-born Adults in 2007

Excludes immigrants born in predominantly English-speaking nations
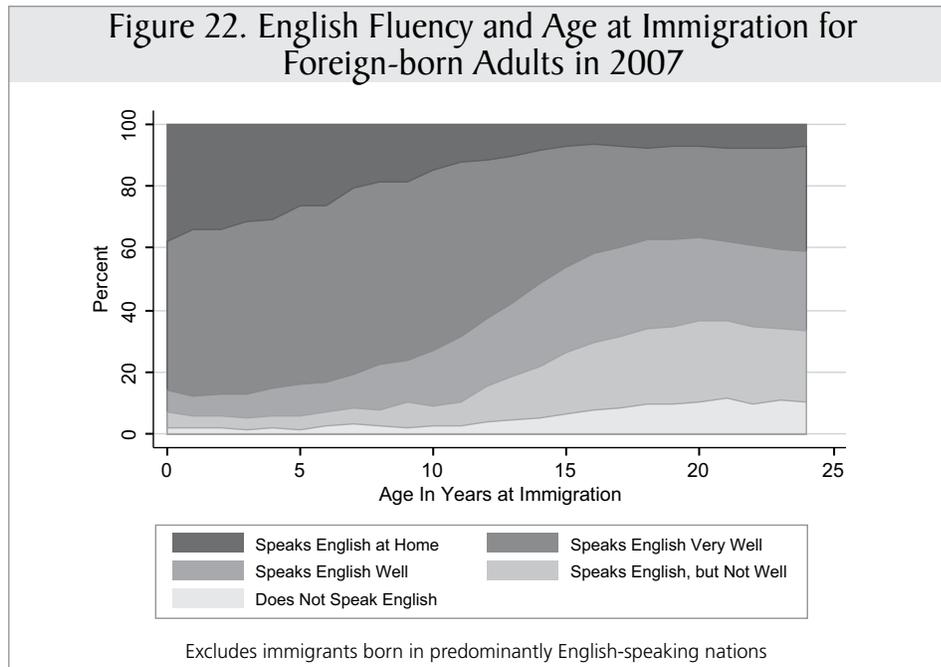
from Mexico, there is evidence that this more recent cohort is making progress at a pace similar to that of their predecessors.

As a final piece of evidence, Figure 22 presents information on the English-language skills of adult immigrants as a function of their age upon arrival into the United States. The figure considers only those immigrants at least twenty-five years of age in 2007. The vast majority of immigrants arriving as young children (Generation 1.5) report speaking English very well as adults. Over 80 percent of immigrants arriving at age seven or younger fall into this category. Over 25 percent of immigrants arriving at age six or younger speak English at home as adults.

While the differences in English skills between immigrants who arrive as newborns and as seven-year-olds are minor, the differences between those who arrive as seven-year-olds and as fourteen-year-olds are stark. Barely 50 percent of this latter group speak English very well, and over 20 percent speak either no or very little English. While consistent with the notion that younger children can acquire second-language skills more easily, the difference might also reflect the fact that teenage immigrants to the United States are able to drop out of school before they learn much

English. A considerable academic debate persists on the question of whether there is a "critical period" in second-language adoption.[5] In the figure, English-speaking skills are lowest among immigrants arriving in their late teens or early twenties; more than a third of immigrants arriving as twenty-one-year-olds, for example, speak either no or little English. Some portion of this trend may reflect the fact that adult immigrants who arrived as younger adults have had fewer years to learn English than those who arrived as children. Immigrants who arrive as older adults, although not depicted in this figure, have even poorer skills—nearly half of those arriving at age twenty-five or older speak little or no English. Once again, this deficiency may reflect their more recent arrival.

This brief study of English-language acquisition by immigrants has identified points of similarity and difference in the experiences of various groups over the past century. While changes in the measurement of English skills by the Census Bureau make long-term comparisons difficult, the contemporary immigrant population appears stronger, primarily because its members are more likely to speak English upon arrival in the United States. The superior English skills of newly arrived immigrants may reflect improvements in the education systems of foreign countries over the

last century or a change in the types of individuals who choose to emigrate.

In spite of this initial advantage, the rate of English-language learning among immigrants who arrive without the ability to speak English appears to be lower than it was a century ago. As discussed previously, the more rapid rate of growth in the immigrant population may have contributed to this trend. It is also possible that in an age of mass media and communications, a network of foreign-language resources makes it easier to get by.

Finally, the English skills of Mexican immigrants are worse, on average, than those of the immigrant population as a whole, though there is some evidence that their rate of English-language acquisition is higher.

Does the United States face a crisis resulting from the reduced English-speaking ability of first-generation immigrants? There is more than one way to read the evidence, and ultimately, the answer must be a subjective one. For the time being, the analysis in the preceding chapter suggests that any deterioration in the English skills of the immigrant population will slow for the foreseeable future, as the current economic downturn reduces the flow of new, less assimilated immigrants to the United States.

CHAPTER 4: SPOTLIGHT ON NATURALIZATION

There is almost certainly no stronger indication of a desire to join American society permanently than applying for citizenship. At the same time, naturalization measures more than merely an immigrant's desire for permanent membership in a particular society. It also measures that society's willingness to accept new members. Throughout American history, government policy has decided which immigrants were eligible for citizenship, and under what circumstances.

Citizenship status is the primary factor determining the index of civic assimilation; the conclusions we draw here regarding the overall trend in naturalization be-tween 1980 and 2007 will thus closely track those drawn from the civic-assimilation index overall. This chapter will expand on this previous evidence by presenting naturalization rates between 1900 and 1930, an era whose data limitations make it impossible to calculate the complete civic assimilation index. This long-run persepective will help highlight how both federal policy and immigrant motivation influence the rate of naturalization. We will pay especially close attention to whether the necessity of demonstrating some degree of ability in English has reduced naturalization rates.

Before we get to the evidence, we need to review the basics of naturalization policy in the United States and the basic patterns of naturalization in Census Bureau and ACS data.

A Brief History of Naturalization Policy

Article I, Section 8 of the Constitution grants Congress the power to enact laws governing the naturalization of immigrants. The first federal law governing citizenship was passed in 1790, and elements of this legislation have remained in place ever since. Specifically, throughout American history, foreign nationals interested in becoming citizens have been subjected to some form of waiting period and have been expected to exhibit "good moral character" and to take an oath of loyalty to the United States.

Over the years, naturalization policy has been restricted along some dimensions and liberalized along others. Racial restrictions on naturalization were imposed in at least some form for a period of over 160 years, and they didn't end until passage of the McCarran-Walter Act in 1952. There have been two significant restrictions over time. A requirement that immigrants be able to speak English before being naturalized was imposed in 1906. While the English-speaking requirement has since been relaxed for older immigrants who have spent a significant number of years in the United States, it is a condition binding about three-quarters of the current immigrant population.

The second major restriction concerned the exclusion of foreigners according to their country of origin. It

began with the Chinese Exclusion Act of 1882. This was followed by the immigration quotas of the 1920s. While the Hart-Celler Act of 1965 eliminated restrictions on access to the United States based on country of origin, it imposed legal permanent residence for some period of time before citizenship could be obtained. Since 1965, possession of a "green card" has been a prerequisite for citizenship. In practice, the limited availability of green cards has extended the waiting period for many would-be citizens, while eliminating any prospect of U.S. citizenship for certain types of legal but temporary residents.

The effects of these changes in policy over time are varied. In 1900, a non-English-speaking, poorly educated European immigrant could become a citizen, but a highly educated Chinese-born alien could not. By 2007, the converse was true: a highly skilled foreign national could gain a place on an employer-sponsored track to citizenship, but for a less educated immigrant there was no path to citizenship except on the basis of family ties to a U.S. citizen or legal permanent residency, and usually after a waiting period that could last a decade.

Basic Evidence on Naturalization

Figure 23 follows the progress toward naturalization of six immigrant cohorts over thirty years. The earliest cohort consists of immigrants who arrived in the last five years of the nineteenth century; the most recent consists of those who arrived between 2001 and 2005. The most obvious point of similarity among all these arrival cohorts is their low naturalization rates in the first few years of their residence in the United States. This pattern strongly reflects the impact of the minimum five-year waiting period, which has always been in force, notwithstanding the handful of exemptions extended—most notably, to those who served in the U.S. military.

It is also true of immigrants arriving before 1980 that naturalization rates, somewhat surprisingly, consistently stabilize at around 70 percent. Immigrants entering this country in the late nineteenth century faced few barriers to naturalization; but by the 1970s, the intermediate hurdle of legal permanent residency had been imposed. Latter-day immigrants also live in a world of cheaper transportation and easier communication, in which opportunities for return or onward migration are greater. The persistence of long-term naturalization rates around 70 percent could indicate either that lenient as well as restrictive policies have little effect, or that the value that immigrants place on obtaining citizenship went up at the same time that official obstacles did.

The civic-assimilation index, plotted for four of these cohorts in Figure 14 above, suggests that the progress exhibited by the most recent cohorts is comparable with that shown by those arriving in the late 1970s. Figure 23 casts at least some doubt on this conclusion. The late-1970s cohort posted a naturalization rate of over 40 percent in 1990; the late-1980s cohort, by contrast, did not hit the 40 percent mark by 2000. It is too early

Figure 23

Naturalization of Immigrant Cohorts 1900-1930

Naturalization of Immigrant Cohorts 1980-2007
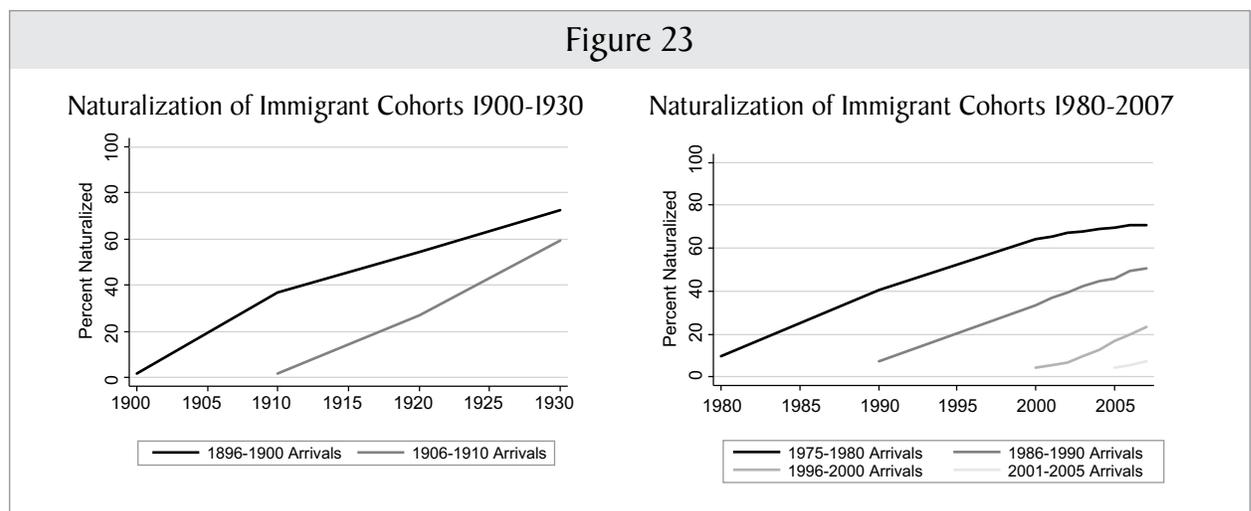
Figure 24. Naturalization of Mexican Immigrant Cohorts 1980-2007

Figure 25. Naturalization of Vietnamese Immigrant Cohorts 1980-2007
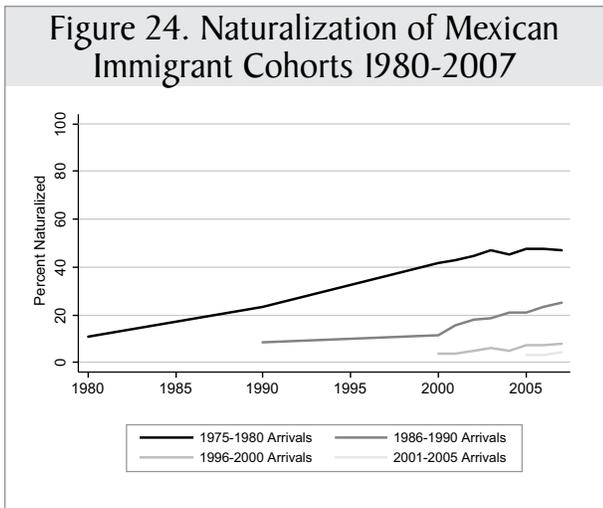
to tell what the late-1990s arrivals will accomplish by 2010, or the early-2000s arrivals by 2015, but there are indications that these cohorts will post naturalization rates that are below historical averages.

There are several possible explanations for the recent slowdown, and there is likely some truth to each of them. First, a higher proportion of post-1980 immigrants may be ineligible for citizenship, either because they hold only a temporary visa or no legal visa at all. Second, fewer eligible immigrants may be seeking citizenship. Third, the effective duration of the waiting period preceding naturalization may be lengthening, along with the queues for legal permanent residency. The number of green cards issued each year is limited by official policy; the number of immigrants eligible for green cards, however, is not directly limited, because of family preferences written into American immigration law. When the eligible population grows faster than the rate of green-card issuance, the wait time for new aliens entering the queue increases. In all likelihood, the first and third explanations count for more than the second.

Just as immigrants vary in their English-language ability and acquisition rates, so do they vary in their propensity to become citizens. Figures 24 and 25 plot naturalization rates for immigrant cohorts born in Mexico and Vietnam, respectively. Relative to rates for the immigrant population as a whole, naturalization rates for immigrants born in Mexico are low. The cohort of late-1970s arrivals took twenty-five years to

Figure 26. Naturalization of Italian Immigrant Cohorts 1980-2007
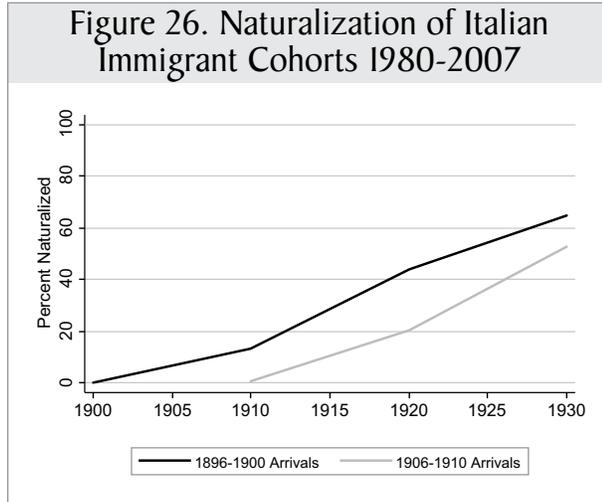
reach a naturalization rate of just under 50 percent; more recent cohorts appear to be on even weaker trajectories. The story is very different for immigrants from Vietnam: by 2007, the cohort of late-1970s arrivals had posted a naturalization rate above 90 percent; late-1980s arrivals had crossed the 80 percent level, and even late-1990s arrivals neared 60 percent.

The strong differences between Mexican and Vietnamese immigrants can be explained by a combination of policy and motivation. For refugees and asylum seekers, the official path to citizenship is easier than it is for other immigrants. They also face stronger incentives to naturalize, so long as a hostile regime retains power in their country of origin. At the other end of the spectrum, illegal immigrants have no path to citizenship,

and even legal migrants from Mexico face long queues for legal permanent residence that are pursued on the basis of family ties or employer sponsorship. The option of returning to one's home country can reduce an immigrant's incentive to naturalize as well.

Figure 26 presents uniformly derived information regarding the naturalization rates of immigrants from Italy in the early twentieth century. The experience of Italian immigrants was largely typical of the broader foreign-born population of the era; naturalization rates were higher than those of present-day Mexican immigrants but lower than those of present-day Vietnamese immigrants.

Language and Naturalization

Should the loopholes that permit some immigrants to become citizens without learning English be abolished? Should the English-language requirement be dropped so as to encourage more immigrants to pursue citizenship? Several independent pieces of evidence point to the conclusion that neither of these controversial policy proposals, if adopted, would have much impact. Those immigrants who aspire to citizenship already have a strong motivation to learn English, whether it is an official requirement or not.

The first piece of evidence is drawn from the early twentieth century. Even at the time that federal legislation imposed the English-language requirement, in 1906, immigrants had to wait a minimum of five years before becoming citizens. So immigrants arriving in 1901 or earlier could have become citizens before facing the language requirement, but those arriving in 1902 or later could not. It is therefore worth asking whether, on account of the onset of the language requirement, the citizenship rates of immigrants on either side of this cutoff point, which was, after all, determined several years later, differed starkly. Figure 27 shows the naturalization rates of Italian immigrants as of the 1920 census, by year of arrival. There is a broad tendency among immigrants who have more recently arrived to become citizens at lower rates, a pattern echoed in every figure in this chapter. However, we see no clear evidence that the English-language requirement by itself lowered naturalization rates among immigrants arriving after the cutoff point.

Two additional pieces of evidence can be drawn from more recent data. Currently, immigrants aged fifty and older are eligible to complete the naturalization examination in a foreign language, provided that they have spent at least twenty years in the United States. Thus, in 2007, we might expect to see two related patterns. Among those immigrants who arrived in

Figure 27. Naturalization of Italian Immigrants by Arrival Year: 1920
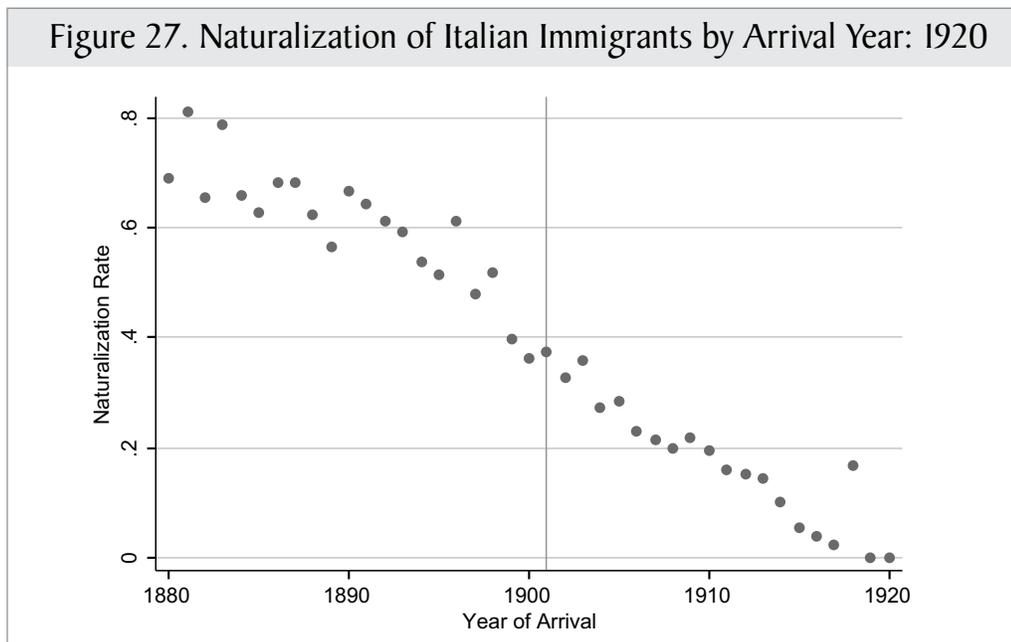
Figure 28. Age and Naturalization, 2007
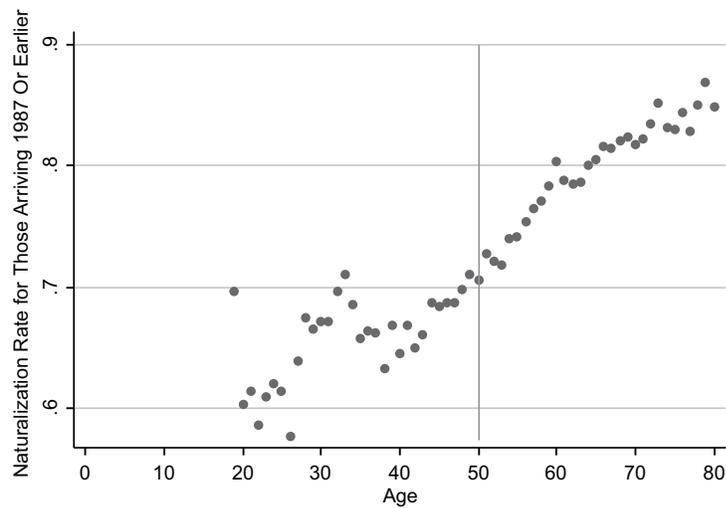
Figure 29. Naturalization of 50-54-Year-Old Immigrants, 2007
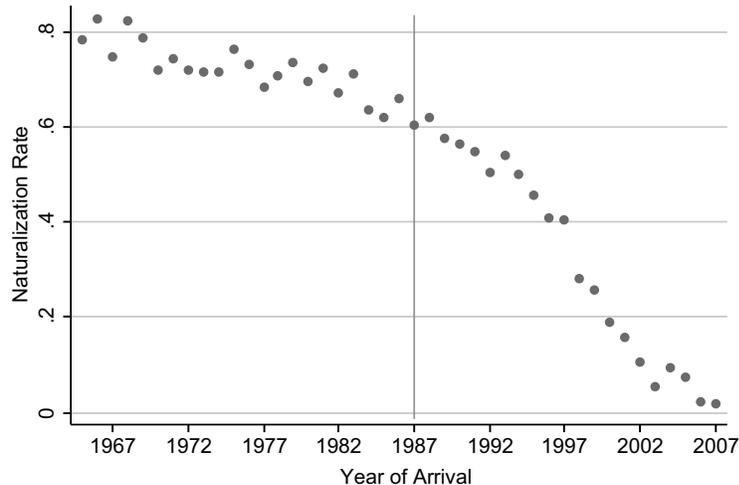
1987 or earlier, there should be a distinct jump in naturalization rates among those aged fifty and older. Similarly, we might expect to see immigrants aged fifty and older who arrived in 1987 become citizens at higher rates than those who arrived shortly thereafter. Figures 28 and 29 look for exactly these patterns in the 2007 ACS; neither one is apparent. While older immigrants tend to have higher naturalization rates (Figure 28), there is no notable break in this relationship at age fifty. And while those who immigrated sometime ago and are now old tend to have higher naturalization rates, there is little distinction between those of them who arrived just before 1987 and those who arrived just after.

The right of some older immigrants to become citizens without learning English may rankle many, but in practice, the loophole is of little consequence. Few, if any, immigrants longed to become citizens for decades but refused to learn sufficient English.

CHAPTER 5: METHODOLOGICAL APPENDIX

This section begins with a basic description of the procedure used to compute the assimilation index, which is followed by a more technical discussion of the statistical model used to distinguish the native-born from the foreign-born. The process used to generate the assimilation index can be divided into four steps.

Step 1: Build a Model That Predicts Immigrant Status

Imagine having access to a wide array of information on the social and economic characteristics of a group of people but no information on their place of birth. On the basis of social and economic information, it might be possible for a well-informed person to guess which individuals in the group were born in the United States and which were born abroad. Knowing that an individual has difficulty speaking English, for example, or that he or she works as an unskilled laborer, may be sufficient to infer that a person was born abroad.

The assimilation index is a measure of how easy it is to infer an individual's place of birth, whether domestic or abroad, on the basis of common social and economic data. The more difficult it is to tell immigrants and natives apart, the higher the index is. Computation of the index begins with data on a representative sample of the American population, evenly split between native- and foreign-born individuals who are at least twenty-five but no more than sixty-five years of age. The data source and exact set of variables used are described below.

The index is computed by guessing which individuals in the data set are native-born and which ones are foreign-born and seeing what proportion of the guesses is correct. The first step in the process is coming up with a method for making guesses. One could imagine many possible rules for guessing whether an individual is an immigrant on the basis of social and economic information; in practice, the index begins by employing a statistical procedure guaranteed to arrive at the most accurate guesses possible. This procedure, known as a probit regression, automatically identifies the personal characteristics most strongly associated with immigrant status, as well as those with little relevance. With this statistical procedure at the heart of the index, there is no need to subjectively assign varying weights to particular characteristics, such as income or marital status. The use of this procedure distinguishes the index from many other popular measures, such as indexes used to rank colleges.

As discussed in Chapter 2, the statistical model underlying the assimilation index considers three sets of factors: economic, cultural, and civic. The model considering all three sets produces the composite assimilation index. In addition to the composite index, this report analyzes the three component assimilation indexes, which are derived from statistical models that analyze only one of the three sets of factors.

Step 2: Use the Model to Make Educated Guesses

Once the model is constructed, information on actual immigrant status is temporarily eliminated from the data set. Once this information has been removed, the model is used to make educated guesses, or predictions, regarding which individuals are, in fact, foreign-born. The predictions take the form of probabilities. A predicted value of zero indicates that there is virtually no chance that the individual in question is foreign-born. A predicted value close to 100 percent indicates that an individual is almost certainly foreign-born.[6]

Complete assimilation is defined as a scenario in which it is impossible to distinguish immigrants from natives; that is, when the two groups are, on average, identical along all the dimensions incorporated into the probit model. In such a scenario, the model will assign each individual in the sample a 50 percent chance of being an immigrant. The educated guess of which individuals are immigrants would be, in this case, no more accurate than a random coin flip. At the other extreme, when the model can predict perfectly which individuals are native-born and which foreign-born, immigrants will receive a predicted probability of 100 percent, and natives a predicted probability of zero.

Table I. Probability Calculations Based on the Probit Regression Model			
	Case 1	Case 2	Case 3
Individual is a U.S. citizen	No	Yes	Yes
Individual is married to a native-born American	No	No	Yes
Individual speaks English	No	Yes	Yes
Individual is a veteran of the U.S. military	No	No	Yes
Result: Probability that individual is foreign-born	100%	94%	8%

Table 1 presents educated guesses at the immigrant status of three hypothetical individuals.[7] While the sets of characteristics of each individual are contrived and the set of characteristics included in Table 1 is far smaller than the set of characteristics incorporated in the probit model, the predicted probabilities are authentic and computed with the use of the same formula used to determine the assimilation index in 2006.

Case 1 concerns an individual who is not a U.S. citizen, is not married to a native-born American, does not speak English, and has not served in the U.S. military. The algorithm derived from the probit regression is used to predict this individual's nativity. In this case, the model is able to predict with 100 percent certainty that the individual is foreign-born. Residents of the United States who are not citizens, are married to foreigners, do not speak English, and are not veterans of the U.S. military are always foreign-born. The algorithm derived from the probit model makes this guess about every individual with this particular set of characteristics.

Case 2 is a more ambiguous scenario. The individual in question is a U.S. citizen and speaks English. However, this individual has not served in the military and is not married to a native-born American, which might indicate that the individual is married to a foreign-born spouse or that the individual is not married at all. While many foreign-born naturalized citizens undoubtedly fit this description, a number of native-born citizens would as well. The prediction offered by the model indicates that this scenario is less ambiguous than it might at first appear. On the basis of comparisons with the nativity of other individuals with similar characteristics, the model offers a 94 percent probability that the individual is foreign-born. In a sample evenly split between native- and foreign-born residents, nearly nineteen of every twenty English-speaking citizens with neither military service

nor a native-born spouse are, in fact, immigrants. The best guess for this particular individual, then, is that he or she is an immigrant.

Case 3 concerns a person who is a U.S. citizen, married to a native-born American, fluent in English, and with past or present service in the U.S. military. While there are some foreign-born citizens who fit this description, the overwhelming majority of persons in this category are, in fact, native-born. The model thus indicates that the likelihood that such an individual is an immigrant is a relatively remote 8 percent. The best guess in this case is that the individual is native-born.

Step 3: Determine the Accuracy of the Guesses

Having built a model in Step 1 and having used that model to make educated guesses in Step 2, we next need to determine just how accurate the guesses are. For this step, the actual information on birthplace is returned to the data set and the actual information is compared with the educated guesses that relied on the algorithm derived from the probit regression model. If the guesses are correct 100 percent of the time, the model can perfectly distinguish immigrants from natives, and the assimilation index will be zero. If the guesses are right only half the time—that is, if the algorithm performed no better than random guessing—then it is impossible to distinguish immigrants from natives, and the assimilation index will be 100 percent.

The composite assimilation index will always make more accurate guesses than any of the component indexes by themselves—statistically, guesses made on the basis of more information are always more accurate. Thus the summary measure of accuracy for the composite index will always be superior to the measure of accuracy for the individual components.

One useful summary measure of the model's accuracy is the average predicted probability among all immigrants in the data set. For example, suppose that the sample contains 100 foreign-born individuals, each of whom has a predicted probability of 100 percent. In this case, the model is perfectly accurate, as reflected by the group's average predicted probability of 100 percent. The assimilation index will equal zero. As another example, suppose that there are 100 foreign-born individuals in the sample and that the model assigned a probability of 80 percent to half of them and 50 percent to the other. In this case, the model was not perfectly accurate, and the group's average predicted probability is 65 percent. The model still performed better than random guessing, however, so the assimilation index will be less than 100 percent.

The average predicted probability can be computed for all immigrants, or for subsets of the immigrant population divided along lines of country of birth, region of residence in the United States, number of years since immigration, or other factors. In theory, averages can also be computed for individual persons.

Step 4: Convert the Average Accuracy Measure into an Index

The final step in computing the assimilation index entails rescaling the average predictions so that high values indicate more assimilation and low values less. In the hypothetical example in which all foreign-born individuals are predicted to be immigrants, the assimilation index takes on a value of zero. Immigrants who can be perfectly identified as such are defined as completely unassimilated. Conversely, a group of immigrants who cannot be distinguished from natives is defined as completely assimilated. The point of no distinction occurs when the probability assigned by the model equals the probability obtained through a random coin flip, or 50 percent.

Data

For the years 2000-2007, the composite assimilation index and its three components are computed with data from the Census Bureau's American Community Survey (ACS). We compute the index for 1990, 1980, 1930, 1920, 1910, and 1900 using the University of Minnesota's Integrated Public Use Microdata Samples (IPUMS) of the decennial census. In doing so, we also analyze the characteristics of males and females between the ages of twenty-two and sixty-five.[8] The alternative index computed for Generation 1.5 includes males and females between the ages of twelve and twenty-four.

Characteristics are incorporated into the predictive model according to the following guidelines: they must measure a characteristic that potentially distinguishes immigrants from natives, that is commonly observed in the ACS and Census Bureau data, and that has inspired at least some interest in previous studies of immigration or current policy debates. This last criterion excludes certain indicators, such as the age of children in an immigrant's household. While this indicator could distinguish immigrants from natives, the literature has not turned to it for this purpose and no current policy debates hinge on it. Our division of indicators into the three categories we refer to as economic, cultural, and civic is largely intuitive; there are several indicators, such as home ownership, that could fall into multiple categories.

Not all these characteristics are available in census data from 1900, 1910, 1920, and 1930. As a consequence, the probit model's capacity to predict immigrant status is slightly lower in these years. When we compare assimilation in the 1980-2006 period with that of the 1900-1930 period, we exclude from the predictive model the set of characteristics available in the later period but not the earlier period. This exclusion has only a modest impact on the assimilation-index computations for the most recent years.

The Predictive Regression Model

A probit regression model is based on the following conceptual model:

$$\Pr(Y=1) = \Pr(X_1\beta_1 + X_2\beta_2 + \dots + X_n\beta_n > \varepsilon).$$

In this context, the variable Y is an indicator set equal to 1 if an individual is an immigrant, and zero

otherwise. The variables X_1 through X_n are measures included in the predictive model: intermarriage, ability to speak English, and so forth. The error term, ε, is presumed to be drawn from a standard normal distribution: mean zero, standard deviation 1. The regression coefficients β_1 through β_n are chosen in a manner that leads the model to make the most plausible predictions possible. For individuals who are immigrants, the goal is to make the sum $X_1\beta_1 + X_2\beta_2 + ... + X_n\beta_n$ as large as possible. For individuals who are not immigrants, the goal is to make this sum as small as possible. We estimate the probit models using the maximum-likelihood method.

Probit regression models are not the only statistical method appropriate for predicting a binary outcome such as whether an individual is an immigrant. The simplest technique is to use an ordinary least-squares regression model, much like what one would use to analyze income or other continuous variables. This sort of model, often referred to as a linear probability model, is inappropriate for this exercise since it relies heavily on predicted probabilities from the model. A primary drawback of linear-probability models is that they can produce predicted probabilities that are less than zero or greater than 100 percent. A second alternative technique, which lacks this unattractive feature, is the logit model. In practice, there is very little difference between assimilation indexes based on probit models and those based on logit models.

The sum $X_1\beta_1 + X_2\beta_2 + ... + X_n\beta_n$ can be translated into a probability if the well-known properties of standard normal distributions are used:

if $X_1\beta_1 + X_2\beta_2 + ... + X_n\beta_n = 0$, then P(individual is immigrant) = 50%

if $X_1\beta_1 + X_2\beta_2 + ... + X_n\beta_n = 1$, then P(individual is immigrant) = 84%

if $X_1\beta_1 + X_2\beta_2 + ... + X_n\beta_n = -1$, then P(individual is immigrant) = 16%

if $X_1\beta_1 + X_2\beta_2 + ... + X_n\beta_n = 2$, then P(individual is immigrant) = 98%

and so forth.

We estimate the probit models using individual-level data from the U.S. Census Bureau enumerations of 1900, 1910, 1920, 1930, 1980, and 1990, as well as the American Community Survey samples of 2000 through 2007. Each data set is made available by the IPUMS project. The data sets are intended to be representative of the entire population of the United States, regardless of nativity or immigration status. It is relatively well-known that the census suffers from an undercount problem, which is thought to be especially severe among minority populations and illegal immigrants. To counteract this problem, the IPUMS project makes a series of sampling weights available. The sampling weights enable researchers to attach greater importance to individuals in the sample who are likely to share characteristics with the sorts of individuals who are probably undercounted. We employ these weights when estimating the probit equations and when aggregating the predicted probabilities that they generate.

Table 2 presents the probit coefficients estimated in the predictive equations for 1910, 1980, and 2007. Separate probit models are estimated each year in order to capture the predictive power of certain characteristics as they change over time. For each year, separate coefficients are estimated for males and females in acknowledgment of the fact that female labor-force participation, military service, and marriage patterns may differ significantly from those of males. In each model, positive coefficients indicate variables positively associated with immigrant status, and vice versa. Across years, the results are generally quite comparable. For each year, the impact of noncitizen status cannot be directly estimated because knowledge that an individual is not a citizen automatically implies that the individual is foreign-born. The predicted likelihood of being an immigrant is set equal to 100 percent for those individuals who are not citizens.

Marriage to an immigrant spouse is highly predictive of immigrant status, with coefficients above 2 in all years. The inability to speak English is another strong predictor, with coefficients between 1.5 and 2. Home ownership is less common among immigrants, though the association has strengthened over time as the overall home-ownership rate has increased. Immigrants

Table 2. Probit Coefficients			
Predictor variable	1910 coefficient	1980 coefficient	2007 coefficient
Not a U.S. citizen	*	*	*
Spouse is an immigrant	2.19	2.2	2.38
Owns residence	-0.087	-0.241	-0.268
Number of own children living in same household	0.026	-0.002^{ns}	0.041
Does not speak English	1.77	1.51	1.86
Married, spouse absent male/female	1.19//0.784	1.12//0.944	1.54//1.18
Separated male/female	---	0.507//0.354	0.908//0.752
Divorced male/female	0.567//0.476	0.313//0.242	0.455//0.406
Widowed male/female	0.775//0.919	0.494//0.355	0.557//0.541
Never married male/female	0.819//0.738	0.441//0.300	0.703//0.508
Occupation score male/female	0.021//0.008	0.008//0.002	-0.0004^{ns} //-0.003
Veteran male/female	---	-0.558//-0.175	-0.770//-0.581
Earned income (thousands) male/female	---	0.002//0.001	-0.002//-0.001
Unemployed male/female	---	-0.023^{ns} /0.006^{ns}	-0.348//-0.226
Out of labor force male/female	0.190/0.046^{ns}	0.037^{ns} /-0.015^{ns}	-0.578//-0.316
1–4 years of education	---	**	0.502
5–8 years of education	---	**	0.206
9 years of education	---	-0.481	-0.266
10 years of education	---	-0.56	-0.71
11 years of education	---	-0.744	-0.872
12 years of education, but no high school diploma	---	-0.647	-0.377
High school graduate or GED	---	-0.64	-0.766
Some college, no degree	---	**	-0.837
Associate degree	---	**	-0.743
Bachelor's degree	---	**	-0.63
Master's degree	---	**	-0.548
Professional degree	---	**	-0.379
Doctorate	---	**	-0.211
Constant term male/female	-1.01//-0.979	0.139//0.306	0.152//0.149

Note: All reported coefficients are statistically significant at the 1 percent level except those marked ^{ns}.

*All native-born individuals are U.S. citizens. Thus, any non–U.S. citizens can be perfectly identified as immigrants and receive a predicted probability of 100 percent.

**The 1980 census used a more exhaustive set of educational-attainment variables than it did in 1990 and 2000 and in the ACS. Complete results are available upon request.

are associated with larger numbers of children in a household in 1910 and 2007, and with categories of marital status other than "married with spouse present." With categorical variables such as marital status, there is always one category omitted from the regression: this becomes the baseline category to which all other categories are compared.

Surprisingly, immigrants are associated with higher-paying occupations in 1910 and 1980; the association is very weak for males in 2007 and negative for females. In 1910, a male physician otherwise identical to a male farm laborer with a predicted immigrant probability of 50 percent would have a predicted immigrant probability of 93 percent. The erosion of

occupational differences between immigrants and natives is a testament to the changing economic position of immigrants in society.

While the probit coefficients suggest that immigrants on the whole have descended the economic ladder relative to natives, there is also evidence that their attachment to the labor force has strengthened over time. Immigrants were more likely to be out of the labor force in 1910 than in 2007. When one examines this evidence, it is important to note that labor-market outcomes are cyclical in nature. Immigrant-native differences may arise and fall with the business cycle.

Among the characteristics not available in the 1910 census is military service, which is negatively associated with immigrant status. The association between educational attainment and the probability of being an immigrant is both positive and negative. In a comparison between two nearly identical individuals, one with an eighth-grade education and the other with a high school diploma, the more educated individual is more likely to be native-born. In a comparison between an individual with a high school diploma and an otherwise identical individual with a doctorate, however, the less educated individual is more likely to be native-born. In other words, immigrants are most underrepresented at intermediate levels of education.

As a final note, when male and female coefficients are allowed to differ from each other, the female coefficients are almost always closer to zero: that is, females are consistently more assimilated than males. It is more difficult to distinguish foreign-born from native-born females than it is to distinguish foreign-born from native-born males.

These coefficients can be used to illustrate the computation of predicted probabilities at the individual level. Suppose that in 2007, we observe a male high school graduate earning $16,000 per year as a cashier. He has no military record, speaks English, has never been married, has no children, is a U.S. citizen, and rents a unit in an apartment building. What is the likelihood that such an individual is foreign-born? First, we use the coefficients in Table 2 to compute an index number for this individual:

   0.152 (constant term)
- 0.766 (high school graduate)
+ 0.703 (never married)
- 0.002*16 (coefficient on income in thousands * income in thousands)
- 0.0004*18 (coefficient on occupation score * occupation score for a cashier)
= 0.039

The probability that this individual is an immigrant is equal to the probability of observing a draw from a standard normal distribution that is below 0.039. This is equal to 51.6 percent. In a sample split evenly between immigrants and natives, about half of all individuals matching these characteristics are foreign-born.

Suppose we take another individual identical to the first, except that he is married to and lives with a foreign-born wife. All other characteristics remain the same. The index number becomes:

   0.152 (constant term)
- 0.766 (high school graduate)
+ 2.38 (spouse is foreign-born)
- 0.001*16 (coefficient on income in thousands * income in thousands)
- 0.0004*18 (coefficient on occupation score * occupation score for a cashier)
= 1.74

The probability of observing a draw from a standard normal distribution below 2.148 is 95.9 percent. In a sample evenly divided between immigrants and natives, we expect about twenty-four of every twenty-five individuals meeting this description to be foreign-born.

Suppose we observe a similar individual in 1910 rather than 2007. The index number calculation uses the 1910 coefficients instead of the 2007 coefficients and omits those variables that are unobserved in the 1910 census:

- 1.01 (constant term)
+ 2.19 (spouse is foreign-born)
+ 0.008*18 (coefficient on occupation score * occupation score for a cashier)
= 1.558

This index number translates into a 94 percent probability of being an immigrant. The lack of relevant data in 1910, coupled with patterns of differences between the native- and foreign-born in that earlier era that are not today's, leads us to be a bit less certain that the individual we have observed is an immigrant.

From Predictions to Index

The probit regression models are used to compute predicted probabilities for every individual in the sample. Samples generally consist of hundreds of thousands of individual observations. Computing the assimilation index for immigrants as a whole or for specific groups of immigrants begins by finding the average, or mean, predicted probability for sample individuals who belong to the group in question. To compute an index for all immigrants, the predicted values of all immigrants in the sample are averaged. To compute an index for Mexican immigrants who arrived in the United States within the last five years, for example, the predicted values of individuals who meet that description are averaged. We always weight the averages using sample weights made available by the IPUMS project.

The averages are then converted into an index value by placing them on a scale between: (a) the value that would be expected if the model could not distinguish immigrants from natives; and (b) the value that would be expected if the model could perfectly distinguish immigrants from natives. The conversion uses the following formula:

Assimilation index = 2 × (100 − mean probability)

When the mean predicted probability is 100 percent—that is, when all immigrants are identified as such in the probit model with a probability of 100 percent—the assimilation index equals zero. A probit model that was completely ineffective in associating personal characteristics with immigrant status would assign all individuals a predicted probability of being an immigrant equal to 50 percent, the proportion of immigrants in the sample. In such a scenario, the index will equal 2 × (100 − 50) = 100%.

There are occasions when the assimilation-index formula returns a value greater than 100 percent. This is most likely to occur when considering the economic assimilation of immigrant groups from developed nations. It occurs when individuals are overrepresented in the educational and occupational categories that are more commonly associated with natives rather than immigrants. In this type of scenario, the assimilation index is reset to its theoretical maximum of 100 percent.

Component Indexes

To compute the component indexes, we recompute the probit regressions, restricting the set of predictor variables to those associated with economic, civic, or cultural assimilation. Removing variables from the predictive model always has the impact of making the predictions less accurate. This is why the component assimilation indexes are always a larger number than the corresponding composite index. The civic-assimilation index, which is based on only two variables, tends to come closest to the composite index because citizenship and military service are very strongly associated with native-born status. The cultural-assimilation index includes a broader array of variables; but in many cases, these variables are weaker predictors of immigrant status than citizenship and military service. Only groups with very low intermarriage rates, or low rates of speaking English, will have civic-assimilation values higher than cultural-assimilation values. Economic assimilation relies on educational attainment, occupation score, income, home ownership, and labor-force participation. As is shown above, the relationship between these factors and immigrant status is weak in recent data, and the association between educational attainment and immigrant status is complex. This explains the tendency of economic assimilation to approach 100 percent in many cases.

Analysis of English-Language Ability

Chapter 3 presents basic information on the English-language ability of U.S. residents who were born in non-English-speaking foreign countries. A nation is defined as English-speaking if at least half the U.S. immigrants from that nation in the 2007 ACS spoke English at home. The set of excluded English-speaking nations includes: Antigua-Barbuda, Australia, the Bahamas, Barbados, Belize, Bermuda, Canada, Grenada, Guyana, Ireland, Jamaica, Liberia, New

Zealand, South Africa, St. Lucia, St. Kitts and Nevis, St. Vincent, Trinidad and Tobago, the United Kingdom, and Zimbabwe.

The analysis is based on all individuals whose English ability is reported in the census or ACS. The census reports English ability for individuals aged ten and older between 1900 and 1930, for individuals aged three and older in 1980, and aged five and older thereafter. As noted in Chapter 3, English ability was recorded by census enumerators in the course of in-person interviews between 1900 and 1930; in 1980 and later, English ability is, for the most part, self-reported by individuals completing mail questionnaires. We weight all reported statistics on immigrant English ability using sample weights that attempt to correct for undercount.

Analysis of Naturalization

Chapter 4 presents information on the citizenship status of foreign-born aliens in the United States. Citizenship is self-reported by census or ACS respondents in 1980 and later, and was recorded by census enumerators between 1900 and 1930. In the early part of the century, aliens were required to state an intent to naturalize, at which point they received "first papers" and became eligible to become citizens after a waiting period. The census records whether aliens received "first papers" as well as whether they are citizens. Analysis here focuses exclusively on whether immigrants had become naturalized citizens by the time of enumeration. All reported statistics employ sampling weights that attempt to correct for undercount.

Caveats

The assimilation index and its components rely on publicly released data from the U.S. Census Bureau, both to build the probit model and to provide a set of individuals for whom predicted probabilities can be computed. While census data sets provide clear advantages, including relatively large samples, relevant variables, and consistent measurement over a time span exceeding a century, there are important limitations to the data. The Census Bureau intends each data set to be representative of the population

of the United States, at least when proper statistical weighting techniques are employed, but there remain concerns that certain segments of the population are undercounted in each census, primarily because they refuse to cooperate with survey enumerators. It is reasonable to believe that the undercounted population includes a disproportionate number of immigrants, particularly those who fear that their participation in the survey will lead to some form of government reprisal. In reality, the Census Bureau is statutorily prohibited from sharing information with any other government agency. Moreover, the census does not inquire whether survey respondents are legal or illegal residents of the United States. However, it may be difficult to convince an illegal immigrant of these protections.

In part to address undercount concerns, the Census Bureau supplies "weights" with each survey. The weights attempt to correct any differences between the sample of individuals who complete the survey and the underlying population by attaching greater importance to the survey responses of members of groups with low and less importance to the responses of members with high response rates. If, for example, non-English-speaking Mexican natives living in Los Angeles are less likely to fill out a survey form, the Census Bureau will assign higher weights to those non-English-speaking Mexicans living in Los Angeles who did participate. In this analysis, Census Bureau weights are employed in the construction of the predictive probit model and the computation of average predicted probabilities for all immigrants and for groups of immigrants.

If undercounted immigrants are less assimilated than those who appear in census enumerations and if the Census Bureau's efforts to correct the undercount by supplying sample weights are insufficient, the "true" index of assimilation will be lower than the reported index. It is more difficult to assess the impact of undercounting on trends in assimilation. According to some reports, the Census Bureau has reduced the magnitude of undercounting over time.[9] If so, the trend in reported assimilation may appear too negative. While it is ultimately difficult to make definitive judgments regarding the impact of undercounting on the assimilation index, the problem is probably not sufficiently large to produce a significant effect. For

example, the Census Bureau estimated that 5 percent of the Hispanic population was undercounted in the 1990 census.[10] The reported downward trend in undercounting implies that the problem was less severe in 2000.

A second caveat relates to the statistical properties of the assimilation index. The index and its components are estimates based on a sample of the U.S. population and, as such, are subject to sampling error. This error will be relatively inconsequential when describing the entire population of foreign-born individuals in the United States but will be more important when describing smaller groups, such as the set of immigrants from a relatively small foreign country or from now living in a small metropolitan area. Small fluctuations over time, or small differences between groups, should not be regarded as having much significance.

Finally, it should be noted that the index and its components are based on information that individuals themselves report to the Census Bureau. The Census Bureau makes few, if any, efforts to verify the accuracy of this information. Respondents may falsely state, for example, that they are U.S. citizens or exaggerate their ability to speak English. The full extent of misreporting in the census is not clear. The index and its components are computed under the assumption that all information reported to the Census Bureau is truthful.

Table 3. Assimilation Index by Birthplace, 2007				
Birthplace	Composite	Economic	Cultural	Civic
Afghanistan	34	94	52	69
Albania	15	95	42	37
Algeria	47	94	80	53
Antigua-Barbuda	63	100	95	69
Argentina	38	100	79	41
Armenia	27	100	46	56
Australia	28	100	100	24
Austria	78	100	100	60
Azores	39	80	68	64
Bahamas	61	100	100	52
Bangladesh	17	93	35	53
Barbados	64	100	93	69
Belgium	52	100	100	48
Belize/British Honduras	45	100	84	51
Bolivia	34	100	68	44
Bosnia	25	100	40	50
Brazil	21	93	72	24
Bulgaria	25	100	58	41
Burma (Myanmar)	25	97	47	52
Byelorussia	31	100	47	61
Cambodia (Kampuchea)	34	88	56	64
Cameroon	16	98	64	23
Canada	54	100	100	44
Cape Verde	42	86	82	53
Chile	41	100	79	45

Birthplace	Composite	Economic	Cultural	Civic
China	22	89	41	47
Colombia	38	100	70	47
Costa Rica	37	93	76	42
Croatia	51	100	66	66
Cuba	41	100	63	52
Czech Republic	45	100	100	49
Czechoslovakia	77	100	99	83
Denmark	43	100	100	37
Dominica	46	89	73	56
Dominican Republic	34	82	70	48
Ecuador	29	87	64	41
Egypt/United Arab Republic	38	99	58	61
El Salvador	18	72	57	30
England	61	100	100	50
Eritrea	31	95	55	63
Ethiopia	27	97	70	40
Fiji	37	100	63	56
Finland	48	100	99	44
France	55	100	100	47
Germany	91	100	100	71
Ghana	34	97	76	47
Greece	61	99	80	75
Grenada	48	98	71	66
Guatemala	15	64	56	23
Guyana/British Guiana	43	100	65	66
Haiti	33	97	67	49
Honduras	16	69	61	22
Hong Kong	53	100	65	77
Hungary	66	100	91	70
India	16	98	38	42
Indonesia	34	100	74	39
Iran	50	100	67	72
Iraq	38	97	60	63
Ireland	55	100	100	53
Israel/Palestine	52	100	79	59
Italy	69	100	96	70
Jamaica	52	100	85	61
Japan	38	100	91	34
Jordan	38	100	58	66
Kenya	23	100	77	27
Korea	40	100	63	55
Kuwait	44	100	74	59

Birthplace	Composite	Economic	Cultural	Civic
Laos	31	91	52	60
Latvia	41	100	83	54
Lebanon	50	100	69	70
Liberia	20	100	78	28
Lithuania	30	100	78	40
Macedonia	29	98	53	57
Malaysia	31	100	74	40
Mexico	13	66	51	22
Moldavia	24	100	57	46
Morocco	35	95	81	43
Nepal	10	91	57	17
Netherlands	63	100	100	48
New Zealand	43	100	100	35
Nicaragua	31	92	66	41
Nigeria	31	100	67	51
Northern Ireland	52	100	100	43
Norway	49	100	100	40
Pakistan	22	97	39	56
Panama	71	100	100	68
Peru	34	100	73	41
Philippines	49	100	72	67
Poland	37	100	62	53
Portugal	40	86	63	63
Romania	37	100	60	60
Russia	35	100	65	55
Scotland	63	100	100	49
Senegal	32	90	90	34
Sierra Leone	34	93	67	51
Singapore	29	100	88	35
Slovakia	35	100	74	52
Somalia	15	67	62	27
South Africa (Union of)	38	100	91	40
Spain	45	100	95	44
Sri Lanka (Ceylon)	23	99	57	37
St. Lucia	32	98	80	46
St. Vincent	55	100	89	64
Sudan	21	91	66	34
Sweden	44	100	100	38
Switzerland	42	100	100	37
Syria	40	100	52	69
Taiwan	42	100	62	68
Tanzania	25	100	66	38

Birthplace	Composite	Economic	Cultural	Civic
Thailand	50	100	95	53
Trinidad and Tobago	47	100	84	56
Turkey	36	94	75	47
Uganda	29	100	63	47
Ukraine	28	100	50	56
Uruguay	23	91	62	29
Uzbekistan	24	91	46	50
Venezuela	35	100	79	35
Vietnam	42	99	53	73
Wales	40	100	100	37
Yemen Arab Republic (North)	25	74	48	61
Yugoslavia	40	99	63	58
Zimbabwe	37	100	85	38

Note: Only birthplace groups with 100 or more representatives in the 2007 American Community Survey sample used to compute the assimilation index are included in this table.

Table 4: Assimilation Index by Metropolitan Area, 2007				
Metropolitan Area	Composite	Economic	Cultural	Civic
Akron, OH	33	99	71	47
Albany-Schenectady-Troy, NY	38	100	72	49
Albuquerque, NM	26	80	66	32
Allentown-Bethlehem-Easton, PA/NJ*	50	100	75	56
Amarillo, TX	28	76	73	32
Anchorage, AK	45	100	84	57
Ann Arbor, MI	31	97	69	36
Athens, GA	18	84	68	18
Atlanta, GA	24	90	62	34
Atlantic City, NJ	27	93	58	49
Augusta-Aiken, GA/SC*	50	96	87	53
Austin, TX	23	75	62	28
Bakersfield, CA	21	74	52	30
Baltimore, MD	33	99	72	46
Baton Rouge, LA	28	89	69	40
Beaumont-Port Arthur-Orange, TX	17	79	64	28
Bellingham, WA	37	100	80	43
Bergen-Passaic, NJ	36	98	61	55
Birmingham, AL	16	81	63	32
Boise City, ID	43	89	76	41
Boston, MA	30	91	68	42
Boulder-Longmont, CO	22	83	64	27
Brazoria, TX	23	81	56	34

Metropolitan Area	Composite	Economic	Cultural	Civic
Bremerton, WA	54	100	97	61
Bridgeport, CT	28	98	69	38
Brockton, MA	41	100	62	59
Brownsville-Harlingen-San Benito, TX	18	67	57	26
Bryan-College Station, TX	15	76	63	22
Buffalo-Niagara Falls, NY	31	97	68	51
Champaign-Urbana-Rantoul, IL	17	90	68	17
Charleston-N. Charleston, SC	38	90	85	44
Charlotte-Gastonia-Rock Hill, NC/SC*	22	87	60	31
Chicago, IL	26	90	55	43
Chico, CA	37	74	65	46
Cincinnati OH/KY/IN*	32	97	73	39
Cleveland, OH	44	99	71	57
Colorado Springs, CO	50	89	92	43
Columbia, SC	26	88	73	30
Columbus, OH	29	93	73	36
Corpus Christi, TX	37	93	80	41
Dallas, TX	17	75	51	28
Danbury, CT	32	99	68	47
Dayton-Springfield, OH	35	100	85	48
Daytona Beach, FL	42	100	85	47
Denver, CO	24	82	61	30
Des Moines, IA	29	93	65	41
Detroit, MI	34	97	63	49
Dutchess Co., NY	42	98	74	48
El Paso, TX	29	81	61	40
Eugene-Springfield, OR	35	86	77	34
Fayetteville, NC	65	100	94	66
Fayetteville-Springdale, AR	16	75	54	21
Fort Lauderdale-Hollywood-Pompano Beach, FL	36	100	70	47
Fort Myers-Cape Coral, FL	26	95	65	33
Fort Pierce, FL	25	84	54	35
Fort Walton Beach, FL	63	97	100	53
Fort Wayne, IN	33	85	82	32
Fort Worth-Arlington, TX	22	82	56	32
Fresno, CA	20	66	51	32
Gainesville, FL	38	96	77	47
Galveston-Texas City, TX	27	82	64	33
Gary-Hammond-East Chicago, IN	34	88	65	49
Grand Rapids, MI	24	86	60	35
Greeley, CO	18	80	53	22
Greensboro-Winston Salem-High Point, NC	18	79	55	25

Metropolitan Area	Composite	Economic	Cultural	Civic
Greenville-Spartanburg-Anderson SC	25	85	61	31
Harrisburg-Lebanon-Carlisle, PA	31	97	69	43
Hartford-Bristol-Middleton-New Britain, CT	36	100	71	46
Honolulu, HI	45	100	75	60
Houston-Brazoria, TX	20	79	54	32
Huntsville, AL	53	98	90	50
Indianapolis, IN	27	92	69	33
Jacksonville, FL	44	100	75	53
Jersey City, NJ	28	90	60	45
Kalamazoo-Portage, MI	34	96	65	42
Kansas City, MO/KS*	28	86	70	34
Killeen-Temple, TX	62	91	92	58
Knoxville, TN	26	94	69	37
Lakeland-Winterhaven, FL	31	85	66	39
Lancaster, PA	45	90	72	40
Lansing-E. Lansing, MI	34	89	80	39
Laredo, TX	21	71	56	27
Las Cruces, NM	24	74	61	32
Las Vegas, NV	28	86	66	38
Lawrence-Haverhill, MA/NH*	31	89	58	47
Lexington-Fayette, KY	16	92	59	30
Little Rock-North Little Rock, AR	27	85	77	30
Longview-Marshall, TX	20	69	58	32
Los Angeles-Long Beach, CA	24	80	55	41
Louisville, KY/IN*	31	92	76	34
Lowell, MA/NH*	36	96	70	52
Madison, WI	34	89	77	38
McAllen-Edinburg-Pharr-Mission, TX	16	71	49	24
Melbourne-Titusville-Cocoa-Palm Bay, FL	56	100	95	53
Memphis, TN/AR/MS*	20	80	63	32
Merced, CA	19	56	46	31
Miami-Hialeah, FL	31	97	58	44
Middlesex-Somerset-Hunterdon, NJ	26	97	50	48
Milwaukee, WI	30	88	72	38
Minneapolis-St. Paul, MN	27	91	66	40
Mobile, AL	17	83	68	24
Modesto, CA	23	78	53	37
Monmouth-Ocean, NJ	36	95	67	47
Naples, FL	16	84	60	21
Nashville, TN	24	82	60	32
Nassau Co., NY	37	100	62	55
New Bedford, MA	40	77	70	63

Metropolitan Area	Composite	Economic	Cultural	Civic
New Haven-Meriden, CT	25	95	79	34
New Orleans, LA	26	87	69	40
New York, NY	30	86	63	48
Newark, NJ	33	93	63	49
Newburgh-Middletown, NY	32	95	71	43
Norfolk-VA Beach-Newport News, VA	55	100	88	62
Oakland, CA	32	93	57	50
Ocala, FL	37	100	68	50
Odessa, TX	20	67	53	33
Oklahoma City, OK	19	84	53	27
Olympia, WA	48	100	75	55
Omaha, NE/IA*	27	85	65	38
Orange County, CA	26	83	54	42
Orlando, FL	31	96	69	43
Pensacola, FL	76	100	100	62
Philadelphia, PA/NJ*	35	97	66	48
Phoenix, AZ	20	78	56	27
Pittsburgh-Beaver Valley, PA	40	98	74	51
Portland-Vancouver, OR	27	89	65	36
Providence-Fall River-Pawtucket, MA/RI*	30	83	63	45
Provo-Orem, UT	22	94	73	27
Raleigh-Durham, NC	20	82	62	26
Reading, PA	26	89	66	34
Reno, NV	29	78	59	44
Richland-Kennewick-Pasco, WA	19	62	50	26
Richmond-Petersburg, VA	32	93	65	41
Riverside-San Bernardino, CA	26	84	57	40
Rochester, NY	50	100	83	56
Rockford, IL	31	95	67	39
Sacramento, CA	29	91	59	45
Salem, OR	14	61	50	18
Salinas-Seaside-Monterey, CA	17	68	52	24
Salt Lake City-Ogden, UT	30	88	67	36
San Antonio, TX	36	88	73	40
San Diego, CA	33	87	66	45
San Francisco, CA	36	91	65	54
San Jose, CA	29	93	53	50
San Luis Obispo-Atascadero-Paso Robles, CA	39	84	81	42
Santa Barbara-Santa Maria-Lompoc, CA	26	72	56	33
Santa Cruz, CA	20	68	54	30
Santa Fe, NM	24	75	80	25
Santa Rosa-Petaluma, CA	20	75	56	28

Metropolitan Area	Composite	Economic	Cultural	Civic
Sarasota, FL	28	98	66	37
Scranton-Wilkes-Barre, PA	35	93	69	50
Seattle-Everett, WA	30	98	67	43
Spokane, WA	52	100	90	49
Springfield-Holyoke-Chicopee, MA	38	91	72	50
St. Louis, MO-IL	34	95	71	41
Stamford, CT	27	92	65	36
Stockton, CA	22	83	48	38
Syracuse, NY	42	89	79	48
Tacoma, WA	53	100	85	56
Tallahassee, FL	32	96	85	42
Tampa-St. Petersburg-Clearwater, FL	39	99	72	45
Trenton, NJ	27	94	60	39
Tucson, AZ	34	89	77	37
Tulsa, OK	18	83	59	26
Tyler, TX	18	74	65	19
Utica-Rome, NY	41	100	71	47
Vallejo-Fairfield-Napa, CA	29	86	54	45
Ventura-Oxnard-Simi Valley, CA	23	81	54	39
Vineland-Millville-Bridgeton, NJ	27	80	72	31
Visalia-Tulare-Porterville, CA	15	57	42	27
Waco, TX	21	65	42	31
Washington, DC/MD/VA*	29	94	64	43
Waterbury, CT	37	89	69	46
West Palm Beach-Boca Raton-Delray Beach, FL	30	94	72	39
Wichita, KS	29	86	69	39
Wilmington, DE/NJ/MD*	36	96	66	48
Worcester, MA	33	95	72	43
Yakima, WA	18	57	59	27
Yolo, CA	24	82	61	39
Yuba City, CA	25	79	57	41
Yuma, AZ	19	83	51	32

Note: Only metropolitan areas with 100 or more foreign-born representatives in the 2007 American Community Survey sample used to compute the assimilation index are included in this table.

* These metro areas span state boundaries.

1. U.S. citizens born abroad, including children born to U.S. military families stationed overseas, are excluded from the analysis.

2. For a theoretical model, see Edward Lazear, "Culture and Language," *Journal of Political Economy* v.107 pp.S95-129 (1999). For a survey of the evidence on the costs of linguistic diversity, see Albert Alesina and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature* v.43 pp.762-800 (2005).

3. Specifically, if more than half of the immigrants from a particular nation interviewed in the 2007 ACS reported speaking English at home, all immigrants from that nation were excluded. The list of excluded nations consists primarily of Commonwealth nations—most notably, the United Kingdom, Ireland, Canada, former British possessions in the Caribbean, South Africa, Australia, and New Zealand.

4. One additional potential skewing factor bears mentioning: English-language skills may increase for some cohorts because young children, who are more likely to acquire English skills than those arriving at older ages (see Figure 22), are not counted in the initial sample. In the 1900-1920 period, for example, English-language ability is not recorded for individuals under the age of ten. In 1900, about 7 percent of the 1896-1900 arrival cohort was under the age of ten. Presuming that 100 percent of these young immigrants knew English in 1900 but were simply not asked if they knew it, the proportion of English-speaking immigrants would have been 48 percent instead of the 45 percent actually recorded. The proportion of non-English-speakers in this cohort who learned the language over twenty years would be recorded as 71 percent instead of 74 percent. This issue is less of a concern in the period 1980-2007; English-language ability was recorded for children as young as three in 1980 and as young as five in later years. The general point that the likelihood of learning English after arrival was greater in later-arriving cohorts still holds after taking this concern into consideration.

5. For a discussion of theories of language acquisition, see Patsy. M. Lightbown and Nina Spada, *How Languages Are Learned*, 3rd ed. (Oxford: Oxford University Press, 2006).

6. In this report, probabilities will be expressed in percentile form, between zero and 100 percent. It is also possible to express probabilities as decimals ranging between zero and 1.

7. The predictions listed in Table 1 are actually averages over all individuals with the listed characteristics in the 2005 sample. There are 3,419 individuals with characteristics matching case 1; 26,798 individuals with characteristics matching case 2; and 29,143 individuals with characteristics matching case 3. The model includes data on 245,480 individuals overall.

8. The index can also be constructed from a data set that is restricted to males only or females only. As discussed in the original assimilation-index report, females tend to have higher assimilation-index values than males. Beyond this difference, the substantive conclusions of the original report and this update are not affected if the analysis is restricted by gender.

9. See Paul M. Ong and Doug Houston, "The 2000 Census Undercount in Los Angeles County," Ralph and Goldy Lewis Center for Regional Policy Studies, Working Paper no. 42, University of California–Los Angeles (2002).

10. See http://www.census.gov/dmd/www/pdf/underus.pdf.

NOTES

NOTES

The mandate of the **Center for Civic Innovation (CCI)** is to improve the quality of life in cities by shaping public policy and enriching public discourse. The Center sponsors studies and conferences on issues such as education reform, welfare reform, crime reduction, fiscal responsibility, immigration, counter-terrorism policy, housing and development, and prisoner reentry. CCI believes that, although good government is essential to civic health, cities thrive only when power and responsibility devolve to the people closest to any problem, whether they are concerned parents, community leaders, or local police.

www.manhattan-institute.org/cci