

The Effects of English Proficiency among Childhood Immigrants:

Are Hispanics Different?*

Mevlude Akbulut-Yuksel

Hoyt Bleakley

Aimee Chin

Abstract

We test whether the effect of English proficiency differs between Hispanic and non-Hispanic immigrants. Using 2000 U.S. Census microdata on immigrants who arrived before age 15, we relate labor market, education, marriage, fertility and location of residence variables to their age at arrival in the U.S., and in particular whether that age fell within the “critical period” of language acquisition. We interpret the observed difference in outcomes between childhood immigrants arriving during the critical period and those arriving later (adjusted for non-language-related age-at-arrival effects using childhood immigrants from English-speaking countries) as an effect of English-language skills and construct an instrumental variable for English-language skills. We find that both Hispanics and non-Hispanics exhibit lower English proficiency if they arrive after the critical period, but this drop in English proficiency is larger for Hispanics. The effect of English proficiency on earnings and education is nevertheless quite similar across groups, which suggests that English proficiency is a key variable in understanding differences across these groups. On the other hand, the effects of English proficiency on social outcomes, such as marriage, fertility and enclave residence, are attenuated for Hispanics.

* Akbulut-Yuksel: Ph.D. Candidate, Department of Economics, University of Houston, 204 McElhinney Hall, Houston, TX 77204-5019 (email: makbulut@uh.edu). Bleakley: Assistant Professor, Graduate School of Business, University of Chicago, 5807 S. Woodlawn Ave., Chicago, IL, 60637 (email: bleakley@chicagosb.edu); Chin: Associate Professor, Department of Economics, University of Houston, 204 McElhinney Hall, Houston, TX 77204-5019 (email: achin@uh.edu). We thank Chinhui Juhn, Stephen Trejo and participants in the IUPLR Conference in April 2007 for helpful comments and discussion. Financial support from the National Institute of Child Health and Human Development (R03HD051562) is gratefully acknowledged. The authors bear sole responsibility for the content of this paper.

I. Introduction

What is the impact of English-language skills on the labor market, educational, marriage, fertility and residential location outcomes of U.S. immigrants? Does the impact differ between Hispanic and non-Hispanic immigrants? These are the questions we address in this chapter.

The increase in immigration in recent decades has drawn attention to the process of immigrant assimilation in the U.S. In 1970, just 4.8% of the U.S. population was foreign-born, but by 2005 the figure had risen to 12.4%.¹ Increasingly, immigrants are coming from countries where English is not widely spoken, leading to a rise in the number of U.S. residents who are not fluent in English. In 2005, 23 million U.S. residents age 5 and over reported speaking English less than very well, which is 8.6% of this subpopulation. Among foreign-born U.S. residents age 5 and over, 52% spoke English less than very well. In this context, it is useful to understand the role that English proficiency has in the process of immigrant assimilation. This knowledge may enable us to form policies that facilitate adjustment to living in America for immigrants and their descendants. Such policies may be desirable because immigrants tend to be worse off educationally and economically compared to natives, and some of their disadvantages are passed on to their U.S.-born offspring. Immigrants are more concentrated in the lowest parts of the education and wage distributions. For example, in 2005, 20.3% of immigrants had completed less than nine years of schooling (compared to 3.7% for natives), and 17.1% of immigrants lived in poverty (compared to 12.8% of natives).

Hispanics accounted for 47% of the foreign-born population and 14.5% of the total population, making them the largest minority race/ethnicity in the U.S., and making Spanish the second most spoken language in the U.S. behind English. Given the large number of Hispanics in the U.S. and their geographic concentration—although the degree of concentration has been

¹ The 1970 figure is from the 1970 U.S. Census and the 2005 figure is from the 2005 American Community Survey. Here and in the rest of this paragraph, we use tabulations of the 2005 American Community Survey done by the Pew Hispanic Center (Fry and Hakimzadeh (2006a)).

decreasing in the past decade, with recent immigrant choosing new areas to settle—it is possible that English-language skills may have different effects for Hispanics than non-Hispanics. For example, it may be more feasible for Hispanics to live and work mostly within the ethnic community, in which case English proficiency may have a reduced role in determining their outcomes. In this chapter, we formally analyze whether Hispanic immigrants' outcomes are indeed less sensitive to English proficiency compared to non-Hispanic immigrants' outcomes.

There are considerable challenges to estimating the effect of an individual's English-language skills on his or her socioeconomic outcomes. Since language skills are correlated with many other variables that also affect these outcomes, such as individual ability, family background and cultural attitudes, it is difficult to separate out what is the causal effect of language skills from the effects of these other correlated variables. This problem is called omitted variable bias since by omitting relevant variables one is left with an estimated effect that is biased (i.e., not giving the true effect). Another source of bias is reverse causality. One can imagine high earnings enabling an immigrant to afford better instruction in English which in turn raises his or her English proficiency. Yet another source of bias is measurement error. It is not easy to accurately measure English-language skills, and not having accurate measures would tend to bias the estimated effect. In the Census data we use, individuals are asked to rate their own English-speaking ability, which may lead to some over-reporting or under-reporting relative to some unstated scale. In these situations—with omitted variable bias, reverse causality or measurement error—ordinary least squares (OLS) regressions of socioeconomic outcomes on language skills are unlikely to give the causal effects of language skills.

One way to obtain causal effects is to run an experiment in which people are randomly assigned different levels of English proficiency. For example, take a group of people who do not know English and randomly select some to have high levels of English proficiency and the

remainder to have low levels of English proficiency. Then the researcher can observe these people's outcomes. In this experimental setting, in which English-language skills are manipulated by the experimenter and not the result of choices made or constraints faced by the individual, then a simple comparison of people who have higher English proficiency and people who have lower English proficiency will give the causal effect of English proficiency. Of course for a variety of reasons, such an experiment is not feasible to do. For one thing, it seems unethical to assign anyone into the lower English proficiency group when there is such a strong prior that English proficiency improves economic well-being in the U.S. Additionally, we would have to follow the subjects for many years before some of the outcomes are realized, such as marriage and fertility. But waiting may not be desirable either because answers are needed now to guide policymaking, or because sample attrition will worsen over time which may offset the benefits of the original experimental design.

A more practical approach to obtaining the causal effect of English proficiency is to take advantage of experiments provided by nature. In our work, we use an identification strategy based on the *critical period of language acquisition*: younger children learn languages more easily than older children and adults. Earlier exposure to English, therefore, should improve the odds that an immigrant to the U.S. becomes proficient in the language. So it is as if nature is assigning each immigrant with a higher or lower cost of acquiring English-language skills based on his or her age at arrival in the U.S.

We apply this identification strategy to study the effect of English proficiency on a number of economic and social outcomes. Section II describes the related literature and the data. We document the strong relationship between age at arrival and English proficiency for immigrants from non-English-speaking countries and detail our identification strategy in Section III. In Section IV we present our findings on the effect of English proficiency on labor market,

education, marriage, fertility and location of residence variables. We discuss some policy implications of our findings in Section V.

II. Background and Data

A. Related literature

A large number of studies has examined the correlation between immigrants' English proficiency and their earnings (e.g., McManus, Gould and Welch (1983), Kossoudji (1988), Tanier (1988) and Chiswick (1991)), education (e.g., Portes and MacLeod (1999) and Glick and White (2003)), marital status (e.g., Stevens and Swicegood (1987), Davila and Mora (2001), Meng and Gregory (2005)), and fertility (e.g., Sorenson (1988), Swicegood, Bean, Stephen and Opitz (1988)). Some have focused on Hispanics in particular and others have examined all immigrants. However, English proficiency is likely to be endogenous for reasons mentioned in the introduction (e.g., omitted variable bias, reverse causality and measurement error), thus OLS estimates of the effect of English proficiency are likely to be biased.

Only a small handful of studies have attempted to address the problem of endogeneity in language skills when estimating the effect of destination-country-language proficiency on earnings (e.g., Chiswick and Miller (1995), Angrist and Lavy (1997), Dustmann and van Soest (2002), and Bleakley and Chin (2004)), education (Bleakley and Chin (2004)), and marital status and fertility (Bleakley and Chin (2007)). Chiswick and Miller (1995) and Bleakley and Chin (2004, 2007) study the U.S. context like we do here, but neither addresses whether the effect of English proficiency differs between Hispanic and non-Hispanic immigrants.

Bleakley and Chin (2008) estimate the causal effects of immigrants' English proficiency on various outcomes of their U.S.-born children, and allow the effects to vary by Hispanic origin. We find that children with limited-English-proficient parents have worse English-language skills

and are more likely to drop out of high school, be below their age-appropriate grade, and not attend preschool. These intergenerational effects do not differ between Hispanic and non-Hispanics. However Bleakley and Chin (2008) do not examine whether the effects of immigrants' English proficiency on their *own* outcomes varies by Hispanic origin.

The main contribution of this chapter is to test whether the effects of English proficiency on immigrants' own outcomes differ between Hispanic and non-Hispanic immigrants. This study uses the same identification strategy as Bleakley and Chin (2004, 2007, 2008). We are not aware of any previous studies that both address the endogeneity of English proficiency and allow the effects of English proficiency to vary by Hispanic origin.

B. Data

Our empirical analysis uses individual-level data from the 2000 U.S. Census of Population and Housing.² The 2000 Census contains a question on English-speaking ability, and we use the responses to this question to construct measures of English-language skills.³ The main measure of English-speaking ability that we use is one coded as follows: 0 = speaks English not at all, 1 = speaks English not well, 2 = speaks English well and 3 = speaks English very well or speaks only English. Thus, a higher value for this measure corresponds to a higher level of English proficiency. We may be concerned, given that this measure is based on individuals' self-reports of English-speaking ability rather than some objective test, whether this measure really captures English proficiency. Kominski (1989) finds that measures of English-speaking ability based on the Census question are highly correlated with both scores from tests designed to measure English-language skills and functional measures of English-language skills.

² Specifically, we combine the 1% and 5% samples from Integrated Public Use Microsample Series (IPUMS) (Ruggles et al. (2004)).

³ The Census question based on which the English-ability measure in this paper is constructed is: "How well does this person speak English?" with the four possible responses "very well," "well," "not well" and "not at all." This question is only asked of individuals responding affirmatively to "Does this person speak a language other than English at home?" We have coded immigrants who do not answer "Yes" to speaking another language as speaking English "very well."

Our analysis is conducted using childhood immigrants currently aged 25 to 55.⁴ We define a childhood immigrant as an immigrant who was under age 15 upon arrival in the U.S. For these immigrants, age at arrival is not a choice variable since they did not time their own immigration but merely come with their parents to the U.S.⁵ Given these age and age at arrival restrictions, individuals in our sample arrived in the U.S. between 1945 and 1989, with 86% of the sample arriving in 1980 or earlier. These individuals have been in the U.S. a minimum of 11 years and an average of 30 years. Given the relatively long spans in the U.S., it is reasonable to believe these individuals would have had the opportunity to learn English if they wished to and could.

The 2000 Census is a general-purpose survey, and enables us to look at a variety of outcomes. For all childhood immigrants, we analyze earnings, employment status, marital status, number of children and location of residence as outcomes. When we examine the effect on spouse's characteristics, we restrict the sample to those who are currently married with spouse present in the household.

We divide our sample into three mutually exclusive language categories: individuals from non-English-speaking countries of birth, countries of birth with English as an official language that have English as the predominant language, and other countries of birth with English as an official language.⁶ The first category is our "treatment" group, and is further

⁴ For the purposes of the empirical analysis immigrant is defined as someone born outside the fifty states and the District of Columbia. This means that a person born in Puerto Rico is considered an immigrant, although legally he/she is a U.S. citizen at birth.

⁵ According to the U.S. Citizenship and Immigration Services, immigrating parents may bring any unmarried children under age 21. We use a more restricted set of childhood immigrants: immigrants who were under 15 upon arrival (i.e., maximum age at arrival is 14). Using this lower age at arrival cutoff should mitigate the concern that many low-educated young men migrate on their own to the U.S. from Mexico and Central America to look for work, which makes age at arrival a choice variable and may make our identification strategy less plausible.

⁶ We used *The World Almanac and Book of Facts, 1999*, to determine whether English was an official language of each country. Recent adult immigrants from the 1980 Census were used to provide empirical evidence of the prevalence of English in countries with English as an official language. English-speaking countries are defined as those countries from which more than half the recent adult immigrants did not speak a language other than English at home. The remaining countries with English as an official language are excluded from the main analysis. We made two exceptions to this procedure. First, despite the fact that Great Britain was not listed as having an official

divided into Hispanic and non-Hispanic.⁷ The second category is our “control” group. The last category is omitted from the main analysis since we are not sure how much exposure to the English language immigrants from these countries would have had before immigrating. Appendix Table 1 reports the mean and standard deviation of the variables we use in our analysis for Hispanics in the treatment group, non-Hispanics in the treatment group, and the control group.

III. Age at Arrival and English Proficiency

There is a window of time in which it is easier to learn languages; this window is known in psychology as the “critical period of language acquisition.”⁸ This appears to be linked to physiological changes in the brain (Lenneberg (1967)): maturational changes starting just before puberty reduce a child’s ability to acquire second languages. If exposure to the language begins during the critical period, acquisition of the language up to native-level proficiency is almost certain. If first exposure commences afterward, the individual’s language proficiency is less assured.

U.S. immigrants from non-English-speaking countries generally do not receive their first exposure to English until they enter the U.S. Given the aforementioned biological constraints to new language acquisition, those who arrive at an early enough age (i.e., during the critical period) can be expected to develop native-level proficiency in English while those who arrive at a later age can be expected to attain a lower level of proficiency. Indeed this is what we observe

language, we included it in the list of English-speaking countries. Second, we classified Puerto Rico as non-English speaking even though English is an official language due to its colonial history.

⁷ We classify someone as Hispanic if they responded affirmatively to the Hispanic origin question (“Is this person Spanish/Hispanic/Latino?”). In theory, there could be endogeneity in this measure of Hispanic origin (the most well assimilated people may cease to call themselves Hispanic; see Duncan and Trejo (2007), however this is unlikely to be a concern in our analysis. Our sample consists of foreign-born, for whom self-reporting to be Hispanic is almost the same as being born in a Spanish-speaking country. When we perform our analysis using being born in a Spanish-speaking country (country of birth is exogenous) to define Hispanic origin, results are basically unchanged.

⁸ See Newport (2002) for an overview of the theory and empirical evidence related to the critical period.

in the data. Figure 1 plots for each age at arrival the mean English-speaking ability in adulthood. For immigrants from English-speaking countries (diamond-marker line), there is no relationship between age at arrival and English proficiency. This makes sense because their age at first exposure to English did not depend on age at migration—English surrounds them in both their country of birth and the U.S. For immigrants from non-English-speaking countries (square-marker line), though, there is a nonlinear relationship between age at arrival and English proficiency: earlier arrivers' English proficiency is similar to that of immigrants from English-speaking countries while later arrivers' English proficiency declines as a function of age at arrival. This pattern for immigrants from non-English-speaking countries is consistent with the critical period of language acquisition. For childhood immigrants arriving well within the critical period of language acquisition, a slightly later arrival does not depress English proficiency in the long run. On the other hand, those who arrived as their critical period was coming to a close attained significantly worse eventual English skills.

Moreover, among immigrants from non-English-speaking countries, the drop in proficiency is around 2-3 times greater for Hispanics than non-Hispanics. There should be no differences in biological constraints to new language acquisition by age between Hispanics and non-Hispanics, and the probable explanation for the steeper decline for Hispanic older arrivers is due to lower pre-U.S.-entry exposure to English for Hispanics.⁹ This seems reasonable since Hispanic immigrants to the U.S. in our sample tend to come from countries that are nearer to the U.S. The shorter distance facilitates emigration from these countries to the U.S., and on average emigrants from these countries will be more numerous and less selected. In contrast, on average

⁹ This explanation for the weaker relationship between age at arrival and English skill for non-Hispanics compared to Hispanics finds corroboration in two other observations. First, for immigrants born in countries with English as an official but non-dominant language, there is also a weaker relationship than for immigrants born in countries without English as an official language. Second, for non-Mexican Hispanic immigrants from non-English-speaking countries, there is also a weaker relationship compared to Mexican immigrants from non-English-speaking countries.

emigrants from countries that non-Hispanics tend to come from will be less numerous and more selected. For example, they may be from wealthier households, have attended private schools that teach in English or at least teach English as a subject, or have American expatriate parents.

For the purpose of the statistical analysis in the next section, the relationship between age at arrival and English proficiency shown in Figure 1 can be captured by an interaction between age at arrival and coming from a non-English-speaking country. When we estimate our models, we always control for a full set of dummies for age at arrival and a full set of dummies for country of birth, and are using only the *interaction* as the instrumental variable for English proficiency. This means that we are not attributing the entire difference in outcome between younger and older arrivers from non-English-speaking countries to English proficiency. Instead, we are attributing only the portion that is *over and above* the difference in outcome between younger and older arrivers from English-speaking countries to English proficiency.

A simple example illustrates the intuition behind our instrumental-variables strategy. Consider four immigrants, each brought to the U.S. as a child. Two are from Jamaica (an English-speaking country), one aged 5 at arrival and the other aged 14. The other two are from Mexico (a non-English-speaking country), with parallel ages of arrival. If we observe a difference in outcome between the two Jamaicans, we could attribute it to non-language age-at-arrival effects (e.g., younger arrivers are able to adjust better to American institutions). But all of these effects are also present in the case of the two Mexicans, in addition to the fact that the Mexicans had substantially less exposure to English before immigrating. As such, the Jamaicans can be used to control for the non-language age-at-arrival effects. Any difference in outcome between the Mexicans in excess of the difference between the Jamaicans can be attributed to language effects, because the Mexican child who immigrated younger has an age of first exposure to English within the critical period while the other Mexican child who immigrated

does not.

We estimate the relationship between English skill and age at arrival in the equation,

$$(1) \quad \text{ENG}_{ija} = \alpha_1 + \pi_1 k_{ija} + \delta_{1a} + \gamma_{1j} + \mathbf{w}_{ija}' \boldsymbol{\rho}_1 + \varepsilon_{1ija},$$

for individual i born in country j arriving in the U.S. at age a . ENG_{ija} is a measure of English proficiency, k_{ija} is an interaction between age at arrival and coming from a non-English-speaking country¹⁰, δ_{1a} is a set of age-at-arrival dummies, γ_{1j} is a set of country-of-birth dummies and \mathbf{w}_{ija} is a vector of exogenous explanatory variables (including age, sex, race and Hispanic origin). Because there are no endogenous variables on the right-hand side, Equation 1 can be consistently estimated using OLS.

Table 1, Column 1 displays the results of estimating Equation 1. There is a significant negative effect of k_{ija} on English skill: for each year past age at arrival 9, English skill declines by a tenth of a unit. In Column 2, we allow the effect of k_{ija} to vary by Hispanic origin. As was seen in Figure 1, the effect of age at arrival is stronger among Hispanic immigrants from non-English-speaking countries: for each year past age at arrival 9, English skill declines by 0.135 units for Hispanics but 0.048 units for non-Hispanics. But for both groups, the effect is significantly different from zero.

We wish to assess whether the effect of English differs between Hispanics and non-Hispanics. Therefore, our empirical work will divide the English skill measure into two categories: English skill of Hispanic immigrants and English skill of non-Hispanic immigrants. Not surprising based on Column 2 results, $k_{ija} \times \text{Hispanic}_{ija}$ and $k_{ija} \times (1 - \text{Hispanic}_{ija})$ are jointly significant predictors of English skill of Hispanics (Column 3) and English skill of non-

¹⁰ A parsimonious way to parameterize the relationship is as an interaction between a piece-wise linear function of age at arrival and a dummy variable for coming from a non-English-speaking country. The specific piece-wise linear function we use takes on the value zero up through age at arrival nine, and is linear thereafter: $k_{ija} = \max(0, a - 9) \times I(j \text{ is a non-English-speaking country})$. We have used other parameterizations of age at arrival and used full set of dummies for age at arrival to form the interaction and obtained similar results.

Hispanics (Column 4).¹¹

IV. Effects of English Proficiency

In this section, we show the results of estimating the effect of English proficiency on outcome y_{ija} using the following equation:

$$(2) \quad y_{ija} = \alpha + \beta_1 \text{ENG}_{ija} \times \text{Hispanic}_{ija} + \beta_2 \text{ENG}_{ija} \times (1 - \text{Hispanic}_{ija}) + \delta_a + \gamma_j + \mathbf{w}_{ija}' \boldsymbol{\rho} + \varepsilon_{ija}$$

for individual i born in country j arriving in the U.S. at age a . Hispanic_{ija} is a dummy for being Hispanic, $\text{ENG}_{ija} \times \text{Hispanic}_{ija}$ is the English proficiency of Hispanics, $\text{ENG}_{ija} \times (1 - \text{Hispanic}_{ija})$ is the English proficiency of all others, δ_a is a set of age-at-arrival dummies, γ_j is a set of country-of-birth dummies and \mathbf{w}_{ija} is a vector of exogenous explanatory variables (including age, sex, race and Hispanic_{ija}). Because English-language skills are endogenous, OLS estimates of Equation 2 will tend to be biased. To obtain consistent estimates of the parameters in Equation 2, we will use 2SLS estimation with $k_{ija} \times \text{Hispanic}_{ija}$ and $k_{ija} \times (1 - \text{Hispanic}_{ija})$ as the excluded instruments.¹² We test whether the effect of English skill for Hispanics significantly differs from the effect for non-Hispanics by performing an F-test where the null hypothesis is that $\beta_1 = \beta_2$.

A. Labor market outcomes

Immigrants tend to earn less than natives in the U.S. Since English is the language used in the American workplace, it is natural to ask the extent to which English proficiency can raise the earnings of immigrants. This is the question we addressed in Bleakley and Chin (2004). This paper introduced the identification strategy described above and applied it to 1990 Census data. In this subsection, we extend this paper by applying the same empirical methodology to

¹¹ Columns 3 and 4 are the first stage regressions that underlie the two-stage least squares regressions estimated in the next section.

¹² As we documented in Section III, these two variables are strong predictors of the endogenous regressors ($\text{ENG}_{ija} \times \text{Hispanic}_{ija}$ and $\text{ENG}_{ija} \times (1 - \text{Hispanic}_{ija})$). Moreover they are plausible exclusion restrictions since they arise from biological constraints to human language acquisition.

newer data, allowing the effects to vary by Hispanic origin, and examining additional labor market outcomes.

Figure 2 shows the mean log annual wages as a function of age at arrival for immigrants from non-English-speaking countries and for those from English-speaking countries.¹³ As in Figure 1, the lines corresponding to the means of the two groups are similar at earlier ages at arrival and diverge for later ages. Among the younger arrivers, whether they come from non-English-speaking countries makes no significant difference in their wages. Among the adolescent arrivers, however, wages tend to be lower for the immigrants from non-English-speaking countries. The line for immigrants from English-speaking countries is nearly flat, suggesting that the non-language effects of age at arrival are small.

It is striking how similar the patterns are in Figures 1 and 2. This makes it especially convincing that the lower earnings observed for older arrivers from non-English-speaking countries observed in Figure 2 are attributable to English proficiency; why else would there be a relationship between wages and age at arrival that is shaped in a way that is consistent with the critical period of language acquisition? Thus, one estimate of the effect of English proficiency on wages comes from estimating Equation 1 but with wages as the dependent variable—the coefficient for the interaction between age at arrival and coming from a non-English-speaking country gives the effect (where we keep in mind that arriving at a later age means lower English proficiency). Though such reduced-form estimates are interesting in and of themselves, sometimes we want to rescale them in order to obtain the effect of a one-unit increase in English proficiency. But this is exactly what we obtain by estimating Equation 2 using 2SLS using the interaction as the excluded instrument; the identifying assumption is that the interaction affects the outcome only through English proficiency, and Figures 1 and 2 provide visual evidence in

¹³ Our measure of wages is from the “wage and salary income” item in the IPUMS 2000 Census data, which is money received as an employee for the previous calendar year.

support of this.

In Row A of Table 2, we present the results of estimating Equation 1 with wages as the outcome. Here and below, we will focus our discussion on the 2SLS estimates in Columns 4-6 since they provide consistent parameter estimates whereas the OLS estimates are not necessarily consistent.¹⁴ English skill raises wages for both Hispanics and non-Hispanics, with the return to English being similar for the two groups. Thus the value of a unit of English skill (e.g., moving from speaks not well to speaks well, or speaks well to speaks very well) is the same across all immigrants. Specifically, a one-unit increase in English proficiency raises annual wages about 35%. The common effect of English proficiency on wages between the two groups is probably to be expected considering that skill prices are determined by the broader labor market; if a profit-maximizing firm wishes to have English-proficient workers, why would it be willing to pay a higher price for the same skill (English proficiency) to one group when it can acquire the same skill from another group at a lower price?

However, the effect of English on other labor market outcomes does appear to differ between Hispanics and non-Hispanics. First, it can be noted in Row B of Table 2 that English does raise the probability that a childhood immigrant works, and this effect is significantly stronger for non-Hispanics than Hispanics (the p-value in Column 6 is 0.0391, which means the difference is significant at the 3.91 level of significance). It turns out that all the effects on employment status reported in Row B derive from women (Row C); male employment is not sensitive to English skill (Row D). Hispanic women have both lower levels of employment (see Appendix Table 1) and lower sensitivity to English at the extensive margin of labor supply than non-Hispanic women. Second, self employment status of non-Hispanic immigrants is also more

¹⁴ As we have found in Bleakley and Chin (2004, 2007), the 2SLS estimates of the effect of English skill tend to be of greater magnitude than the OLS estimates. Analysis in Bleakley and Chin (2004) suggests that while there does appear to be upward bias due to omitted variables-type stories, the downward bias from classical measurement error more than offset it for an overall downward bias in the OLS estimate.

sensitive to English proficiency than that of Hispanic immigrants: there is a 10.2 percentage point reduction in being self-employed for non-Hispanics and no effect for Hispanics. Perhaps this lower sensitivity for Hispanics is due to the larger Hispanic community in America (which may make operating an ethnic business more attractive for Hispanics), or due to the smaller improvements in job quality as a result of raising English proficiency for Hispanics (Hispanic immigrants are more likely to be undocumented and not have a college education, both of which may limit access to the attractive jobs outside self-employment).

B. Educational Attainment

In Bleakley and Chin (2004), we found that English proficiency raised wages mainly in an indirect way, through raising educational attainment. In this subsection we quantify the effect of English proficiency on education and assess the role of education in the observed relationship between English proficiency and wages using 2000 Census data.

Figure 3 shows the relationship between years of schooling completed and age at arrival. The pattern of years of schooling completed by age at arrival bears remarkable resemblance to the pattern of English proficiency by age at arrival (in Figure 1), which supports the interpretation of English proficiency playing a causal role in the poorer educational outcomes of older arrivers from non-English-speaking countries. We proceed by estimating Equation 2 by 2SLS to quantify the causal effects of English proficiency on educational outcomes by Hispanic origin. Columns 4-6 of Table 3 show that a unit increase in English skill raises years of schooling by about three years for both Hispanics and non-Hispanics—there is no significant difference in the effect. However, it is interesting to note that the three years of extra schooling appear to be coming from different parts of the educational distribution for the two groups. In particular, Rows B-D suggests that there is more increase at the high school diploma margin for Hispanics compared to non-Hispanics, for whom the largest gain is registered at the college

level. Of course the only significant difference is in Row D, for attaining a Bachelor's degree or higher.

Assuming standard returns to a year of schooling (say, 8%), the educational channel accounts for over two-thirds of the total effect of English proficiency on wages. Thus, other channels—including the direct effect on productivity—play a much smaller role in determining the wages of childhood immigrants. This may sound surprising initially since we may have found the story of direct productivity effects of English proficiency appealing, e.g., since business is conducted in English in the U.S., English proficiency should facilitate communication with co-workers and customers. However, once we recognize that the return to education in the U.S. has grown dramatically since the 1970s, and that a high school diploma and college education are now prerequisites for most good jobs, then this result is much less surprising.

C. Family formation

English proficiency may impact not only the economic outcomes of immigrants, but also their social outcomes. For example, when one becomes proficiency in English, one's pool of potential spouses may expand since there are more people with whom one can communicate. Alternatively, English proficiency may bring about changes in preferences regarding marriage and family, such as from home-country norms to U.S. norms. Bleakley and Chin (2007) estimated the effect of English proficiency on marriage and fertility outcomes, and we extend this paper by testing whether these effects vary by Hispanic origin.

In Panel I of Table 4, we estimate the impact of English skills on current marital status. We find that English proficiency reduces the probability of being currently married for both Hispanics and non-Hispanics, with the effect being larger for non-Hispanics. This effect derives both from a higher likelihood of being divorced (a one-unit increase in English increases the

probability of being divorced 4.6% for Hispanics and 9% for non-Hispanics) and a lower likelihood of having ever married (a one-unit increase in English increases the probability of having ever married by 4.4% for Hispanics and 23% for non-Hispanics). The differences in the point estimates are large, but it must be pointed out that the estimates in Column 5 are somewhat imprecise and admit a wide range of values within the 95% confidence interval.

In Panel II of Table 4, we examine the effect on spousal attributes. This analysis is restricted to childhood immigrants who are currently married and living with their spouse. We find that increasing English skill of the childhood immigrant significantly increases his/her spouse's level of English proficiency and the probability that the spouse is U.S.-born. The flip side of being more likely to marry a native is the lower likelihood of marrying someone with the same country of birth; marrying immigrants from other countries seems to have a negligible role. It is interesting to note that although childhood immigrants with better English are less likely to marry someone with the same country of birth, there is a less dramatic decline in the propensity to marry someone of the same primary ancestry, at least for Hispanics.¹⁵ Indeed, Hispanic childhood immigrants are not significantly more likely to marry outside their primary ancestry though they are significantly more likely to marry a native, reflecting the fact that some are marrying natives of the same ancestry (e.g., a Mexican immigrant marrying a U.S.-born person of Mexican ancestry).

Another note is that the marriage outcomes in Panel I of Table 4 are significantly more sensitive to English proficiency for non-Hispanics than Hispanics. For example, in Row B, a one-unit increase in English raises the probability of marrying a U.S. native by 25% for Hispanics and 80% for non-Hispanics. One possible explanation for the differential effects on

¹⁵ Our measure of primary ancestry is from the "ancestry, first response" item (the 3-digit version) in the IPUMS Census data. Often, the ancestry reflects the country where a person or his ancestors were born, but some countries have more detailed categories. For example, there are several distinct codes for people of Mexican ancestry: Mexican, Mexican American, Chicano/Chicana, Nuevo Mexicano and Californio.

marriage outcomes is that though both Hispanics and non-Hispanics have similar preferences for assortative mating by ethnicity and maintaining “traditional” values at the outset, non-Hispanics may have a reduced opportunity to do so due to their smaller ethnic group size in the U.S. Due to the larger Hispanic community in America, there is greater potential to find a spouse who not only satisfies the usual criteria for a spouse (e.g., age, educational attainment, personality), but also is of the same cultural heritage. In contrast, a non-Hispanic childhood immigrant may have to marry outside his/her country of birth or ancestry in order to find someone with the usual criteria. Another possible explanation for the differential effects is that Hispanic and non-Hispanic childhood immigrants have different preferences for assortative mating by ethnicity and maintaining “traditional” family values. These two explanations are unlikely to be independent. For example, it is plausible that ethnic group size might itself shape individual preferences. Perhaps growing up in an ethnic enclave causes one to retain more traditional values even as one acquires English-language skills. Although all immigrant groups have their enclaves, enclaves may be more numerous and larger among Hispanics compared to non-Hispanics due to the former’s larger population size in the U.S.

In Table 5, we examine fertility outcomes. Our measure of fertility is the number of children currently living in the same household.¹⁶ We find that an increase in English proficiency reduces the number of children, especially for non-Hispanics (Row A). For Hispanics, there is no effect on the extensive margin of having any children (Row B), but there is an effect on the number of children conditional on having at least one. For non-Hispanics, there is a significant reduction both in the probability of having any children and the number of children. The effect on the extensive margin shown in Row B appears to come entirely from the effects of English skill on marital status. Restricting the sample only to married couples, we find

¹⁶ The 2000 Census does not contain information on the number of children ever born, a better measure of fertility, but we verified in Bleakley and Chin (2007) using 1990 Census data that there are similar estimated effects of English using either fertility measure.

there is no significant effect on having any children for both Hispanics and non-Hispanics (Rows C and D). Thus, it does not appear that couples are foregoing parenthood altogether, but there is a reduction in the number of children conditional on having any children or at least there is a postponement in completing one's family.

D. Location of residence

An additional marker of social integration of immigrants is the extent to which they live in ethnic enclaves. The public-use 2000 Census data that we use are not ideal for studying residential choice decisions because they do not provide detailed information about the neighborhood that a person resides to preserve respondent privacy. The lowest level of geographic aggregation that we can measure is something called a PUMA (public-use microdata area), which is an area containing at least 100,000 people. A more accurate characterization of one's neighborhood would contain fewer people, but given the data limitations, we form a couple of variables intended to capture the idea of an ethnic enclave. One set of measures that we use is based on the fraction of the population in one's PUMA that shares the same country of birth as the childhood immigrant.¹⁷ Of course there are people of the same background who are born in the U.S. (e.g., U.S.-born Mexicans may have many similarities to Mexicans born in Mexico), so a second set of measures used is based on the fraction of one's PUMA that shares the same primary ancestry. A larger fraction from the same country or ancestry may be associated with being in a larger ethnic community and greater likelihood of living in an ethnic enclave. In Table 6, we find that English proficiency significantly reduces the probability that childhood immigrants live in PUMAs with a high fraction from own country or with same ancestry for non-

¹⁷ Recognizing that an enclave is really where there is a big concentration of people from own group, we also tried nonlinear functions of the fraction. Here, we show results not only for when the fraction itself is the outcome, but also for when a dummy for whether an individual lives in a PUMA that has a fraction from same country of birth that is above the mean fraction for all immigrants from that same country of birth (this latter measure picks up whether for someone from your country, you tend to live in a neighborhood with an above-average number of fellow countrymen).

Hispanics, but the extent to which Hispanics live with their own group does not depend on English proficiency. For example, Row B shows that for a one-unit increase in English raises the probability of living in a PUMA with above-average fraction of people from same country of birth for non-Hispanics but has no impact for Hispanics' probability. This is consistent with non-Hispanic immigrants moving away from ethnic enclaves and integrating geographically with mainstream America as they gain English proficiency. Hispanic immigrants do not appear to do the same as their English improves.

E. Why might the effect of English proficiency differ between Hispanics and non-Hispanics?

We have shown that except for wages and years of schooling, English skill tends to have effects of greater magnitude for non-Hispanics than Hispanics. In this subsection, we explore two reasons that we *ex ante* thought could be responsible for the observed difference in effect of English between Hispanics and non-Hispanics.

One potential reason for the differential effects on marriage and fertility outcomes is that English proficiency affects the two groups at different parts of the education distribution. As discussed above, an improvement in English skill tends to raise years of schooling at the high school level for Hispanics and at the college level for non-Hispanics. Attending college could affect outcomes beyond raising years of schooling. People are more likely to leave their hometown or at least their childhood home in order to attend college. By changing the place of residence, attending college could change the pool of potential spouses and reduce the influence of the family and ethnic community. But to the extent that English skill is only raising schooling at the high school level (e.g., from dropping out at grade 9 or 10 to completing high school), then the three extra years of schooling associated with a unit increase in English proficiency is unlikely to change the potential pool of spouses and relevant community. To assess this education story for differential effects, we started with Equation 2 and added dummies for each

level of educational attainment measured by the 2000 Census.¹⁸ When we do this, we still find similar differences in effects of English proficiency between Hispanics and non-Hispanics as in the base specification. However, as far as the levels of the effects of English skill for Hispanics and non-Hispanics are concerned, it is notable that the effect on wages declines drastically, mirroring the Bleakley and Chin (2004) finding that much of the effect of English proficiency on wages is mediated through years of schooling. The effect decreases more for non-Hispanics than Hispanics, which is consistent with mechanisms besides educational attainment mattering more for Hispanics.¹⁹

A second potential reason for the differential effects on marriage and fertility outcomes is the large size of the Hispanic population in the U.S. Hispanic immigrants share a common language, Spanish, and can associate with each other even if they are born in different countries. Even if Hispanic immigrants did not associate with people born in other countries, there would be large ethnic communities by country of birth due to their relatively large population and geographic concentration. For example, Mexicans (both immigrants and U.S.-born) make up 9% of the U.S. population and are still relatively concentrated in the Southwest, although they are increasingly settling elsewhere. On average then, Hispanic immigrants are more likely to live in ethnic enclaves, and their ethnic social networks are larger than non-Hispanic immigrants. This may impact the role of English in economic and social assimilation. One can imagine that the existence of a large and diverse ethnic community may create opportunities that are not available to all immigrants, and even an immigrant who becomes English-proficient may have a desire to continue participating in this community. For example, this immigrant may participate in the

¹⁸ There are 14 educational categories altogether: no schooling, completing 1st-4th grade, completing 5th-8th grade, 9th grade, 10th grade, 11th grade, 12th grade without diploma, high school graduate or GED, some college without degree, associate degree in occupational program, Bachelor's degree, Master's degree, professional degree and doctorate degree.

¹⁹ The effect on wages controlling for the education dummies for non-Hispanics is about one-third of the original 2SLS estimate reported in Table 2 and is statistically insignificant, and for Hispanics is about one-half of the original estimate and still statistically significant. The difference in effect between the two groups remains insignificant.

broader society for school and work, but his or her social life is spent in the ethnic community. Thus, it is plausible that the muted effects of English proficiency on marriage and fertility outcomes for Hispanics could be the result of the larger Hispanic community.

To assess this second reason, we did the following. First, we added super-Public Use Microdata Area (PUMA) of residence fixed effects as an attempt to control for neighborhood characteristics. Some neighborhoods may be heavily concentrated ethnically and others more diverse, and adding super-PUMA fixed effects would control for any fixed features (such as ethnic diversity) of the super-PUMA. Considering a super-PUMA contains about 400,000 people, it is a rough measure of one's community, so perhaps not surprisingly results did not change after allowing for super-PUMA fixed effects. Second, we controlled for a variable measuring the fraction of one's PUMA's population with the same country of birth. Adding this fraction as a control did not change any of the results. Finally, we controlled for a variable measuring the fraction of one's PUMA's population with the same primary ancestry. Again, the differential effects between Hispanics and non-Hispanics remained as they were in the base results.

The three empirical tests described in the previous paragraph do not provide evidence in favor of the ethnic enclave story for why effects tend to be lower for Hispanics, but they are rather weak tests. On the one hand, we probably do not have a good measure of each individual's relevant community. PUMAs are still large areas, and two PUMAs with a similar fraction of own group may nonetheless differ in the degree of ethnic segregation across neighborhoods within the PUMA. On the other hand, perhaps there is a nonlinear effect of ethnic community size wherein the ethnic community does not have effects until it reaches a certain size, and currently we do not have precise enough measures of ethnic community size to account for this. Thus, we should not yet dismiss the ethnic enclave story.

It is apparent from this subsection that we do not know why the effects of English on marriage and fertility outcomes differ between Hispanics and non-Hispanics. Undoubtedly there are reasons besides the education and ethnic enclave explanations offered here. As well, better measures of one's relevant community would enable a better assessment of the ethnic enclave explanation than what we have done. The role of ethnic enclaves in mediating the effects of English-language skills merits more research, as do other potential explanations for the differential effects between Hispanics and non-Hispanics.

F. Robustness checks

Central to our ability to interpret the 2SLS estimates presented in this section as the causal effects of English-language skills is the assumption that non-language age-at-arrival effects are identical for immigrants from English-speaking countries and from non-English-speaking countries. Is this a plausible assumption? One potential concern is that immigrants from non-English-speaking countries are on average from poorer countries than immigrants from English-speaking countries, which may generate stronger age-at-arrival effects for the immigrants from non-English-speaking countries. That is, richer countries may have better education systems and more modern institutions, so the transition to living in the U.S. would be easier for immigrants from richer countries than poorer countries. A second concern is that differences in geographic or cultural distance to the U.S. among English-speaking countries, non-English-speaking countries from which Hispanics tend to come, and non-English-speaking countries from which non-Hispanics tend to come might cause differential non-language age-at-arrival effects among these three country groups for immigrants.

To address these concerns, we did a variety of robustness checks with different specifications and subsamples. It turns out that the estimated effects of English on Hispanics and non-Hispanic, and the difference in effects between the two groups do not change much when we

control for these alternative explanations, which supports the assumption and therefore the interpretation of the 2SLS estimates as being related to English proficiency. Since the results do not change much in these robustness checks, we do not report their results in tables.

V. Discussion

In this chapter, we document a strong relationship between age at arrival and English proficiency using 2000 U.S. Census microdata on childhood immigrants. This observed relationship is consistent with the critical period of language acquisition, with young arrivers from non-English-speaking countries developing a high level of proficiency and older arrivers—who entered the U.S. as their critical period was coming to a close—developing lower proficiency. The pattern of our outcomes by age at arrival bears remarkable resemblance to the pattern of English proficiency by age at arrival (to save space, we have only provided these graphs for two outcomes, wages and years of schooling), motivating us to use an instrumental-variables strategy based upon the critical period to identify the causal effects of English-language skills. We allow these causal effects to vary by Hispanic origin, which previous studies have not done. We find for both Hispanics and non-Hispanics, an increase in English proficiency significantly increases wages, educational attainment, intermarriage and divorce, and significantly decreases being currently married, being ever married, fertility and living in an ethnic enclave. The sensitivity of wages and years of schooling to English proficiency was the same for Hispanics and non-Hispanics, but English proficiency tended to have significantly larger-magnitude effects on marriage, fertility and location of residence outcomes for non-Hispanics than non-Hispanics.

If policymakers want to close the education and earnings gap between Hispanics and natives, it seems viable to direct resources at raising the English proficiency of Hispanics. Both

Hispanic and non-Hispanic immigrants' schooling and earnings exhibit the same responsiveness to a unit change in English proficiency, but Hispanics have on average lower levels of English proficiency. It seems logical, then, to raise the level of English proficiency for Hispanics. Had we found that Hispanics' schooling and wages were not sensitive to English proficiency, this would not be a viable way to achieve the policy objective.

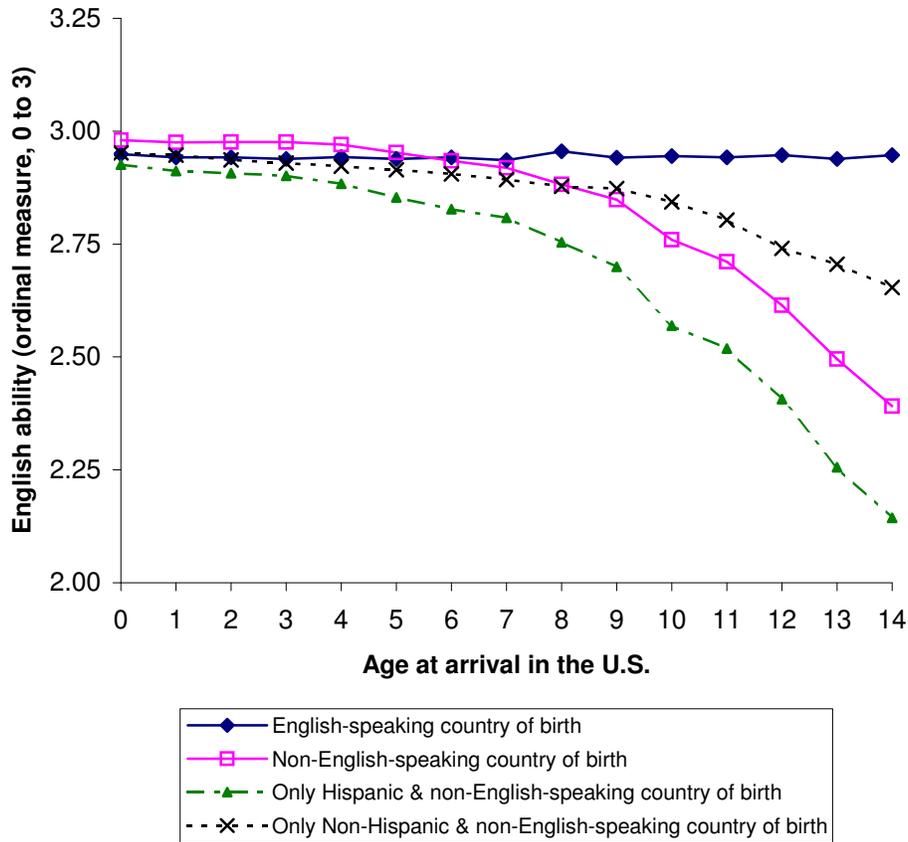
Some advocate English classes for immigrants not so much to help immigrants raise their schooling and wages but to assimilate them into the mainstream culture. A recent *Washington Post* article reports on the debate over whether the U.S. government should be more pro-active in promoting assimilation into the mainstream culture such as through English classes (Brulliard (2007)). Without taking a position on this debate, this chapter suggests that English-language proficiency does not automatically lead to social assimilation. That the effects of English-proficiency on marriage and fertility outcomes differ between Hispanics and non-Hispanics suggests that ethnicity-specific forces are at least in part moderating the role of English proficiency. It is conceivable that the two groups are taking different lengths of time to reach U.S. norms, or that one or both groups will never reach the U.S. norm and instead set their own cultural norm for living in the U.S.

References

- Angrist, Joshua D. and Victor Lavy, "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco," *Journal of Labor Economics* 15:1 (1997), S48-S76.
- Bleakley, Hoyt and Aimee Chin, "Language Skills and Earnings: Evidence from Childhood Immigrants," *Review of Economics and Statistics* 86:2 (2004), 481-496.
- Bleakley, Hoyt and Aimee Chin, "English Proficiency and Social Assimilation Among Immigrants: An Instrumental-Variables Approach," Center for Comparative Immigration Studies Working Paper 149, San Diego, CA: University of California at San Diego, April 2007.
- Bleakley, Hoyt and Aimee Chin, "What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital among Immigrants," *Journal of Human Resources* 43 (Spring 2008), 267-298.
- Brulliard, Karin, "At Odds over Immigrant Assimilation: Whether the U.S. Government Should Offer Encouragement is Debated," *Washington Post*, August 7, 2007, page A1.
- Chiswick, Barry R., "Speaking, Reading and Earnings among Low-Skilled Immigrants," *Journal of Labor Economics* 9:2 (1991), 149-170.
- Chiswick, Barry R. and Paul W. Miller, "The Endogeneity between Language and Earnings: International Analyses," *Journal of Labor Economics* 13:2 (1995), 246-288.
- Davila, Alberto and Marie T. Mora, "The Marital Status of Recent Immigrants in the United States in 1980 and 1990," *International Migration Review* 35:2 (2001), 506-524.
- Duncan, Brian and Stephen Trejo, "Ethnic Identification, Inter-marriage and Unmeasured Progress by Mexican Americans," in Borjas, George, ed., *Mexican Immigration*, Chicago: National Bureau of Economic Research and the University of Chicago Press, 2007.
- Dustmann, Christian and Arthur van Soest, "Language and the Earnings of Immigrants," *Industrial and Labor Relations Review* 55:3 (2002), 473-492.
- Fry, Richard and Shirin Hakimzadeh, "A Statistical Portrait of the Foreign-born Population at Mid-Decade," Pew Hispanic Center Report, Washington, DC: Pew Hispanic Center, October 2006a.
- Fry, Richard and Shirin Hakimzadeh, "A Statistical Portrait of Hispanics at Mid-Decade," Pew Hispanic Center Report, Washington, DC: Pew Hispanic Center, October 2006b.
- Glick, Jennifer E. and Michael J. White, "The Academic Trajectories of Immigrant Youths: Analysis Within and Across Cohorts," *Demography* 40:4 (2003), 759-783.
- Kominski, Robert, "How Good Is 'How Well'? An Examination of the Census English-Speaking Ability Question," presented at the 1989 annual meeting of the American Statistical Association, Social Sciences Section, Washington, DC: US Census Bureau, Population Division, 1989.
- Kossoudji, Sherrie A., "English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men," *Journal of Labor Economics* 6:2 (1988), 205-228.
- Lenneberg, Eric H., *Biological Foundation of Language*, New York: Wiley & Sons, 1967.
- McManus, Walter, William Gould and Finis Welch, "Earnings of Hispanic Men: The Role of

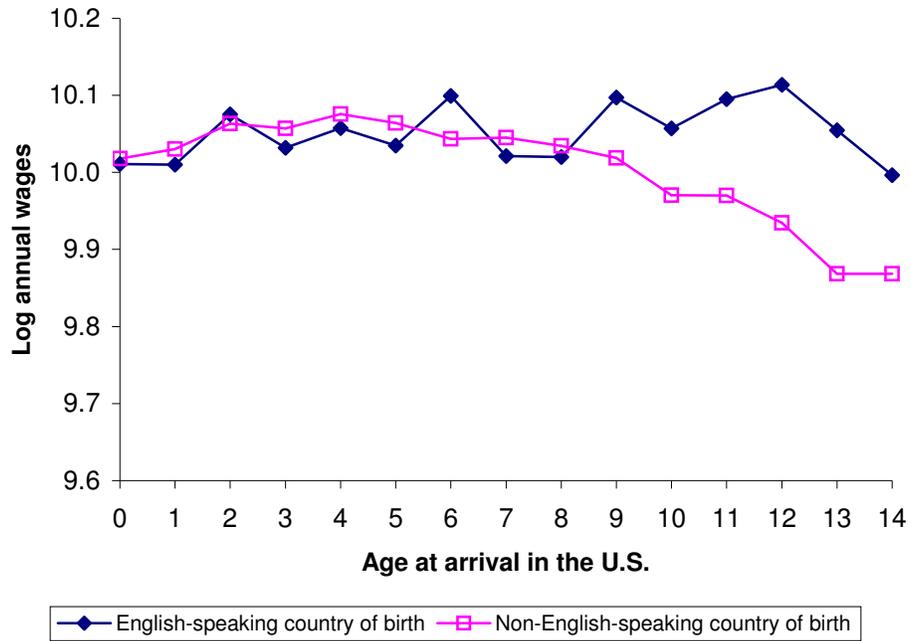
- English Language Proficiency,” *Journal of Labor Economics* 1:2 (1983), 101-130.
- Meng, Xin and Robert G. Gregory, “Intermarriage and the Economic Assimilation of Immigrants,” *Journal of Labor Economics* 23:1 (2005), 135-175.
- Newport, Elissa L., “Critical Periods in Language Development,” in L. Nadel (Ed.) *Encyclopedia of Cognitive Science*, London: Macmillan Publishers Ltd./ Nature Publishing Group, 2002.
- Portes, Alejandro and Dag MacLeod, “Educating the Second Generation: Determinants of Academic Achievement Among Children of Immigrants in the United States,” *Journal of Ethnic and Migration Studies* 25:3 (1999), 373-396.
- Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander, *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database], Minneapolis, MN: Minnesota Population Center [producer and distributor] (2004) (<http://www.ipums.org>).
- Sorenson, Ann Marie, “The Fertility and Language Characteristics of Mexican-American and Non-Hispanic Husbands and Wives,” *Sociological Quarterly* 29:1 (1988), 111-130.
- Stevens, Gillian and Gray Swicegood, “The Linguistic Context of Ethnic Endogamy,” *American Sociological Review* 52:1 (1987), 73-82.
- Swicegood, Gray, Frank D. Bean, Elizabeth Hervey Stephen and Wolfgang Opitz, “Language Usage and Fertility in the Mexican-Origin Population of the United States,” *Demography* 25:1 (1988), 17-33.
- Tanier, Evelina, “English Language Proficiency and Earnings among Foreign-born Men,” *Journal of Human Resources* 23:1 (1988), 108-122.
- Weil, David N., *Economic Growth*, U.S.: Pearson Addison Wesley (2005).
- World Almanac and Book of Facts, 1999*, New York: World Almanac Books (1999).

Figure 1. English-Speaking Ability by Age at Arrival



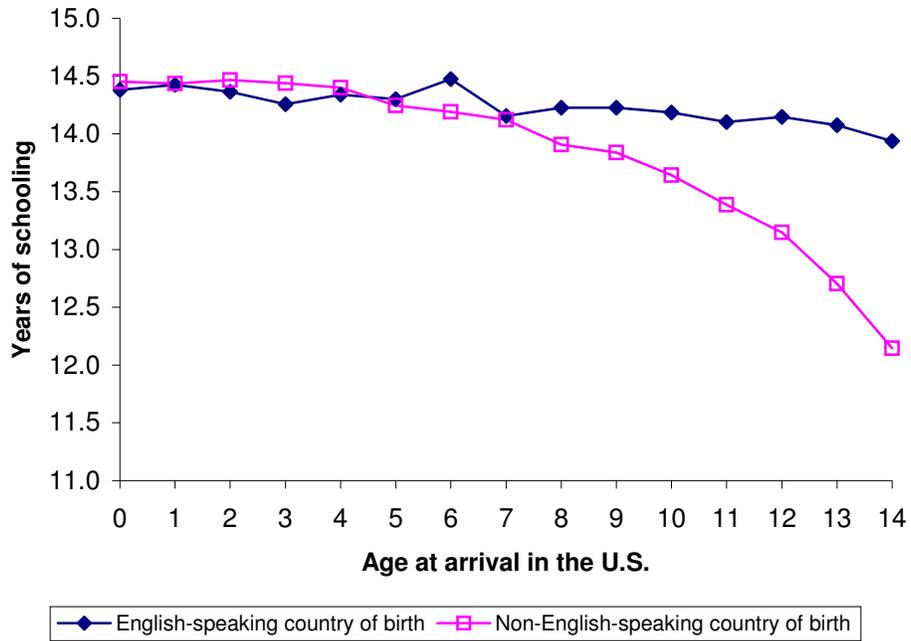
Notes: Data are from the 2000 IPUMS. Sample size is 191,534 (composed of people who immigrated to the U.S. before age 15 and are currently aged 25-55, and with nonmissing English variable). Displayed for each age at arrival is the mean English-speaking ability. Means are weighted by IPUMS weights, and regression-adjusted for age, race, Hispanic and sex dummies. The race categories used were White, Black, Asian & Pacific Islander, Multiracial and Other. The English ordinal measure is defined as: 0 = no English, 1 = not well, 2 = well and 3 = very well.

Figure 2. Wages by Age at Arrival



Notes: Data are from the 2000 IPUMS. Sample size is 127,616 (composed of people who immigrated to the U.S. before age 15 and are currently aged 25-55, and with nonmissing English and wage variables). Displayed for each age at arrival is the mean log wages in 1999. Means are weighted by IPUMS weights, and regression-adjusted for age, race, Hispanic and sex dummies. The race categories used were White, Black, Asian & Pacific Islander, Multiracial and Other.

Figure 3. Years of Schooling by Age at Arrival



Notes: Data are from the 2000 IPUMS. Sample size is 188,191 (composed of people who immigrated to the U.S. before age 15 and are currently aged 25-55, and with nonmissing English and educational attainment variables). Displayed for each age at arrival is the mean years of schooling. Means are weighted by IPUMS weights, and regression-adjusted for age, race, Hispanic and sex dummies. The race categories used were White, Black, Asian & Pacific Islander, Multiracial and Other.

Table 1. First Stage Regressions

	Endogenous regressor:			
	English ability (1)	English ability (2)	English ability× Hispanic (3)	English ability× non-Hispanic (4)
Identifying instruments:				
Max (0, age at arrival - 9)×non-English-speaking country of birth	-0.104 *** (0.029)			
Max (0, age at arrival - 9)×non-English-speaking country of birth×Hispanic		-0.135 *** (0.027)	-0.136 *** (0.027)	0.0010 (0.0010)
Max (0, age at arrival - 9)×non-English-speaking country of birth×non-Hispanic		-0.048 *** (0.009)	-0.001 (0.002)	-0.047 *** (0.008)
F-statistic associated with test of joint significance of identifying instruments (p-value)	13.14 0.0004	21.38 <0.0001	13.75 <0.0001	21.59 <0.0001

Notes: The sample is as described in the notes to Appendix Table 1. Each column is from a separate OLS regression that is weighted by IPUMS weights and contains dummies for age at arrival, country of birth, age, sex, race (White, Black, Asian, Multiracial and Other), Hispanic origin and Hispanic×non-English-speaking country of birth. The country-of-birth dummies are based on IPUMS detailed birthplace codes. Standard errors adjusted for country of birth clusters are shown in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01). English-speaking ability is measured on an ordinal scale as follows: 0 = no English, 1 = not well, 2 = well and 3 = very well.

Table 2. Effect of English Proficiency on Labor Market Outcomes

	OLS			2SLS		
	Effect of English for Hispanics (1)	Effect of English for non-Hispanics (2)	p-value of test of equality (3)	Effect of English for Hispanics (4)	Effect of English for non-Hispanics (5)	p-value of test of equality (6)
A. Log annual wages	0.249 *** (0.011)	0.204 *** (0.027)	0.1235	0.377 *** (0.063)	0.354 ** (0.170)	0.8571
B. Worked last year	0.077 *** (0.012)	0.066 *** (0.007)	0.4335	0.070 *** (0.014)	0.152 *** (0.044)	0.0391
C. Worked last year, females only	0.118 *** (0.006)	0.075 *** (0.008)	<0.0001	0.140 *** (0.013)	0.280 *** (0.062)	0.0120
D. Worked last year, males only	0.036 ** (0.015)	0.056 *** (0.009)	0.2712	-0.011 (0.019)	-0.030 (0.050)	0.6260
E. Is self-employed	0.004 *** (0.001)	-0.006 (0.005)	0.0462	-0.008 (0.009)	-0.102 *** (0.037)	0.0057

Notes: The sample is as described in the notes to Appendix Table 1. In each lettered row, Columns 1-3 report the coefficient for English ability for Hispanics, coefficient for English ability for non-Hispanics, and p-value of test of equality of the two aforementioned coefficients, respectively, from an equation estimated using OLS, and Columns 4-6 report the same from an equation estimated using 2SLS. All regressions are weighted by IPUMS weights and contain dummies for age at arrival, country of birth, age, sex, race, Hispanic origin and Hispanic×non-English-speaking country of birth. The identifying instruments when 2SLS is used are $\max(0, \text{age at arrival} - 9) \times \text{non-English-speaking country} \times \text{Hispanic}$ and $\max(0, \text{age at arrival} - 9) \times \text{non-English-speaking country} \times (1 - \text{Hispanic})$. Standard errors adjusted for country of birth clusters are shown in parentheses. Asterisks denote significance levels (*=.10, **=.05, ***=.01).

Table 3. Effect of English Proficiency on Educational Attainment

	OLS			2SLS		
	Effect of English for Hispanics (1)	Effect of English for non-Hispanics (2)	p-value of test of equality (3)	Effect of English for Hispanics (4)	Effect of English for non-Hispanics (5)	p-value of test of equality (6)
A. Years of schooling	1.971 *** (0.036)	1.366 *** (0.128)	<0.0001	3.374 *** (0.192)	3.089 *** (0.772)	0.6836
B. Has high school diploma	0.234 *** (0.008)	0.127 *** (0.012)	<0.0001	0.400 *** (0.018)	0.306 *** (0.091)	0.2727
C. Has any college or more	0.175 *** (0.016)	0.171 *** (0.020)	0.8663	0.289 *** (0.041)	0.383 *** (0.136)	0.4059
D. Has Bachelor's degree or more	0.052 *** (0.010)	0.146 *** (0.023)	0.0003	0.092 ** (0.042)	0.313 *** (0.111)	0.0087

Notes: See notes for Table 2.

Table 4. Effect of English Proficiency on Marriage Outcomes

	OLS			2SLS		
	Effect of English for Hispanics (1)	Effect of English for non-Hispanics (2)	p-value of test of equality (3)	Effect of English for Hispanics (4)	Effect of English for non-Hispanics (5)	p-value of test of equality (6)
Panel I: Marital Status						
A. Is currently married with spouse present	0.008 (0.012)	0.009 (0.009)	0.9361	-0.070 ** (0.027)	-0.318 *** (0.076)	<0.0001
B. Is currently divorced	0.013 *** (0.002)	-0.001 (0.003)	0.0006	0.046 *** (0.011)	0.090 ** (0.037)	0.1137
C. Has ever married	-0.001 (0.008)	-0.005 (0.007)	0.7258	-0.044 ** (0.017)	-0.227 *** (0.055)	0.0001
Panel II: Spouse's Ethnicity and Nativity (conditional on being married with spouse present)						
A. Spouse's English-speaking ability ordinal measure	0.524 *** (0.010)	0.463 *** (0.039)	0.1354	0.753 *** (0.032)	1.198 *** (0.099)	<0.0001
B. Spouse is US-born	0.097 *** (0.006)	0.138 *** (0.014)	0.0172	0.246 *** (0.082)	0.800 *** (0.263)	0.0073
C. Spouse has the same country of birth	-0.113 *** (0.005)	-0.150 *** (0.013)	0.0179	-0.272 *** (0.080)	-0.846 *** (0.264)	0.0048
D. Spouse has the same ancestry	-0.065 *** (0.007)	-0.126 *** (0.013)	0.0002	-0.095 (0.072)	-0.737 *** (0.230)	0.0004

Notes: See notes for Table 2.

Table 5. Effect of English Proficiency on Fertility

	OLS			2SLS		
	Effect of English for Hispanics (1)	Effect of English for non-Hispanics (2)	p-value of test of equality (3)	Effect of English for Hispanics (4)	Effect of English for non-Hispanics (5)	p-value of test of equality (6)
A. Number of children living in same household	-0.112 *** (0.022)	-0.078 ** (0.036)	0.4398	-0.374 *** (0.070)	-0.814 *** (0.201)	0.0129
B. Has a child living in same household	-0.003 (0.005)	-0.013 (0.009)	0.3405	-0.044 (0.027)	-0.236 *** (0.083)	0.0022
C. Number of children living in same household, only individuals married with spouse present	-0.178 *** (0.011)	-0.173 *** (0.051)	0.9288	-0.351 *** (0.052)	-0.765 *** (0.207)	0.0250
D. Has a child living in same household, only individuals married with spouse present	-0.015 *** (0.002)	-0.037 *** (0.007)	0.0020	0.008 (0.017)	-0.099 (0.067)	0.0622

Notes: See notes for Table 2.

Table 6. Effect of English Proficiency on Neighborhood of Residence

	OLS			2SLS		
	Effect of English on Hispanics (standard error) (1)	Effect of English on non-Hispanics (standard error) (2)	p-value of test of equality (3)	Effect of English on Hispanics (standard error) (4)	Effect of English on non-Hispanics (standard error) (5)	p-value of test of equality (6)
Panel I: Neighborhood Measures based on Fraction of Population in PUMA from Same Country of Birth						
A. Fraction from same country of birth	-0.011 *** (0.001)	-0.005 *** (0.001)	0.0004	-0.005 (0.008)	-0.017 ** (0.008)	0.1447
B. Fraction is above mean	-0.044 *** (0.007)	-0.078 *** (0.010)	0.0183	0.014 (0.050)	-0.296 *** (0.091)	<0.0001
Panel II: Neighborhood Measures based on Fraction of Population in PUMA from Same Ancestry						
A. Fraction from same ancestry	-0.014 *** (0.001)	-0.007 *** (0.002)	0.0005	-0.010 (0.013)	-0.062 *** (0.018)	0.0005
B. Fraction is above mean	-0.039 *** (0.007)	-0.061 *** (0.009)	0.0910	0.004 (0.050)	-0.214 ** (0.088)	0.0026

Notes: See notes for Table 2.

Appendix Table 1. Descriptive Statistics

	Born in non-English-speaking country		Born in English-speaking country
	Hispanics (1)	non-Hispanics (2)	(3)
Panel I: Regressors			
English-speaking ability (ordinal measure, 0 to 3)	2.567 (0.747)	2.883 (0.375)	2.980 (0.167)
Age	35.871 (7.987)	37.285 (8.476)	38.403 (8.367)
Female	0.496 (0.500)	0.504 (0.500)	0.528 (0.499)
White	0.477 (0.499)	0.638 (0.481)	0.666 (0.472)
Black	0.017 (0.130)	0.045 (0.206)	0.247 (0.432)
Asian/Pacific Islander	0.003 (0.052)	0.253 (0.435)	0.028 (0.164)
Other single race	0.455 (0.498)	0.007 (0.083)	0.019 (0.136)
Multiracial	0.049 (0.215)	0.058 (0.233)	0.040 (0.196)
Panel II: Labor Market Outcomes			
Log annual wages	9.993 (0.919)	10.281 (0.964)	10.288 (0.956)
Worked last year	0.829 (0.376)	0.878 (0.327)	0.891 (0.312)
Worked last year, females only	0.750 (0.433)	0.823 (0.381)	0.849 (0.358)
Worked last year, males only	0.908 (0.289)	0.934 (0.248)	0.938 (0.242)
Is self-employed	0.064 (0.244)	0.099 (0.299)	0.089 (0.284)
Panel III: Schooling Outcomes			
Years of schooling	11.882 (3.608)	14.485 (2.712)	14.527 (2.452)
Has high school diploma	0.640 (0.480)	0.922 (0.269)	0.940 (0.237)
Has any college or more	0.406 (0.491)	0.730 (0.444)	0.746 (0.435)
Has Bachelor's degree or more	0.133 (0.340)	0.403 (0.491)	0.375 (0.484)

Notes: The table continues on the next page.

Appendix Table 1. Descriptive Statistics (Continued)

	Born in non-English-speaking country		Born in English-speaking country
	Hispanics	non-Hispanics	
	(1)	(2)	(3)
Panel IV: Marital Status			
Is currently married with spouse present	0.605 (0.489)	0.602 (0.489)	0.561 (0.496)
Is currently divorced	0.094 (0.292)	0.099 (0.299)	0.120 (0.325)
Has ever married	0.786 (0.410)	0.747 (0.435)	0.736 (0.441)
Panel V: Spouse's Ethnicity and Nativity (conditional on being married with spouse present)			
Spouse English-speaking ability ordinal measure	2.356 (0.908)	2.839 (0.456)	2.979 (0.170)
Spouse is US-born	0.346 (0.476)	0.654 (0.476)	0.804 (0.397)
Spouse has the same country of birth	0.536 (0.499)	0.239 (0.426)	0.094 (0.292)
Spouse has the same ancestry	0.646 (0.478)	0.426 (0.494)	0.245 (0.430)
Panel VI: Fertility			
Number of children living in same household	1.468 (1.418)	1.005 (1.219)	0.974 (1.178)
Has a child living in same household	0.645 (0.478)	0.510 (0.500)	0.506 (0.500)
Number of children living in same household, only individuals married with spouse present	1.997 (1.325)	1.479 (1.243)	1.421 (1.208)
Has a child living in same household, only individuals married with spouse present	0.855 (0.352)	0.734 (0.442)	0.715 (0.451)
Panel VII: Neighborhood of Residence			
Fraction of population from PUMA from same country of birth	0.110 (0.122)	0.010 (0.026)	0.008 (0.019)
Fraction from same country of birth is above mean, within country of birth	0.415 (0.493)	0.333 (0.471)	0.410 (0.492)
Fraction of population from PUMA with same primary ancestry	0.165 (0.184)	0.058 (0.080)	0.044 (0.062)
Fraction with same ancestry is above mean, within ancestry	0.440 (0.496)	0.425 (0.494)	0.470 (0.499)

Notes: The sample consists of individuals from the 2000 1% and 5% PUMS files who are currently aged 25-55, immigrated to the U.S. before age 15 and have nonmissing own age, year of immigration, country of birth and English variables. Total number of observations is 191534 for the English variable, with Columns 1-3 containing 86387, 79241, and 25906 observations, respectively. Statistics are weighted by IPUMS weights. The English-speaking ability ordinal measure is defined as: 0 = no English, 1 = not well, 2 = well and 3 = very well.