The Effects of English Proficiency on Economic and Social Outcomes: Summary of Evidence from Childhood Immigrants in the U.S. Census

Hoyt Bleakley
Graduate School of Business
University of Chicago

Aimee Chin
Department of Economics
University of Houston

September 2007

ABSTRACT

In this chapter, we quantify the effect of English-language skills on various socioeconomic outcomes of U.S. immigrants, including earnings, marital status, fertility, location of residence and intergenerational progress. Because conventional methods of estimating the effect of English proficiency may lead to biased estimates—due to omitted variables, reverse causality or measurement error—we take advantage of a "natural experiment" that provides plausibly exogenous variation in English proficiency. In particular, we rely on the psychobiological phenomenon that younger children learn new languages more easily than older children and adults to identify the causal effect of English proficiency. We implement our research design using data on childhood immigrants in the 2000 U.S. Census. We find that English proficiency has an important role in determining the economic and social outcomes of immigrants, with the effects of English proficiency tending to be larger using our approach relative to the conventional approach. Additionally, we find that although both Hispanic and non-Hispanic immigrants' outcomes depend on English proficiency, Hispanic immigrants' marriage, fertility and location of residence outcomes tend exhibit less sensitivity to English proficiency relative to non-Hispanic immigrants'.

* Bleakley: Assistant Professor, Graduate School of Business, University of Chicago, 5807 S. Woodlawn Ave., Chicago, IL, 60637 (email: bleakley@chicagogsb.edu); Chin: Associate Professor, Department of Economics, University of Houston, 204 McElhinney Hall, Houston, TX 77204-5019 (email: achin@uh.edu). This chapter is a summary and extension of our four previous papers on the effects of English-language skills on childhood immigrants in the U.S. 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 chapter.

I. Introduction

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 age 5 and over, 52% spoke English less than very well. In this context, it is useful to understand the role that English proficiency plays in the process of immigrant assimilation. This knowledge may enable us to form policies that facilitate the adjustment to living in America for immigrants and their descendents. Such policies may be desirable because immigrants tend to be worse off educationally and economically compared to natives, and their disadvantages may be 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 have completed less than nine years of schooling (compared to 3.7% for natives), and 17.1% of immigrants live in poverty (compared to 12.8% of natives).

There are considerable challenges to estimating the effect of an individual's English-language skills on his or her outcomes. Examples of outcomes we will examine below are the individual's earnings, educational attainment, location of residence, propensity to intermarry, fertility, and children's English proficiency and educational progress. Since English-language skills are correlated with many other variables that also affect these outcomes, such as individual ability and cultural attitudes, it is difficult to separate out what is the causal effect of English proficiency from the effects of these other correlated variables. This problem is called omitted

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

variable bias by economists, 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 raises his or her English proficiency, or children's English proficiency affecting parent's English proficiency. So then, what is the cause and what is the effect? 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 may not yield the causal effects of language skills.

One way to obtain causal effects is to run an experiment in which people are randomly assigned levels of English proficiency. For example, take a group of people who do not know English and randomly select half 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. Additionally, we would have to follow the subjects for many years before some of the outcomes are realized, such as fertility and children's

education. 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 construct a research design based around 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 elaborate on this research design below.

We have applied this research design to study the effect of English proficiency on a number of economic and social outcomes. Section II describes the data used. We document the strong relationship between age at arrival and English proficiency for immigrants from non-English-speaking countries using individual-level data from the 2000 U.S. Census in Section III. In Section IV, we summarize our findings on the effect of English proficiency on earnings, family formation (marital status, spousal characteristics and fertility), location of residence, and children's outcomes (children's English proficiency and educational progress). We discuss some policy implications of our findings in Section V.

II. Data

For most of the empirical analysis we report on below, we use 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

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

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.

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.

[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."

[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 research design less plausible.

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, location of residence and number of children as outcomes. When we examine the effect on spouse's characteristics, we will restrict the sample to those who are currently married with spouse present in the household. When we look at children's outcomes, we will be using all U.S.-born children up to age 17 who are living with at least one parent who is in the main sample (i.e., a childhood immigrant who is aged 25 to 55).[6]

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.[7] The first category is our "treatment" group and the second 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 shows the decomposition of the sample by country of birth, and also presents our classification of countries by English-speaking status.

III. Age at Arrival and English Proficiency

According to the critical period of language acquisition, there is a window of time during

[6] In the Census, we are able to match child to parent only when they reside in the same household. Since children aged 18 and over are likely to move out of their parents' home, it is not possible to analyze the impact of parental English on adult children's outcomes. Separately, in instances where both parents are childhood immigrants aged 25 to 55, the mother's characteristics are assigned. In Bleakley and Chin (2007), we find that results do not change when we assign father's characteristics instead.

[7] 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 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.

which it is easier to learn a new language and after which it is more difficult (see Newport (2002) for a brief review). This appears to be linked to physiological changes in the brain (Lenneberg (1967)). Immigrants coming from non-English-speaking countries tend not to be exposed to English until they enter the U.S. Then immigrants who arrive at an early enough age—during this window—will develop native proficiency in English, and those arriving after the window will not be assured of this.

We see in Figure 1 that there is indeed a strong association between immigrants' age at arrival and their adult English-language skills. The lower proficiency among later arrivers is quite pronounced for immigrants from non-English-speaking countries (see diamond-marker line), and consistent with psycholinguistic evidence on the critical period. On the other hand, for immigrants from English-speaking countries, the connection between age of arrival and age at first exposure to English is not so tight, and therefore differences in proficiency by age at arrival are barely perceptible for this group (square-marker line). (Age at arrival probably affects immigrants through channels other than language, such as through better knowledge of American culture and institutions. Throughout the studies cited here, we use immigrants from English-speaking countries to control for non-language-related age-at-arrival effects.)

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.[8] 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 the interaction as the instrumental variable for English proficiency. This means that we are not attributing the entire difference in outcome

[8] 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. 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.

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.

To clarify this research design, we offer this hypothetical example: 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 15. The other two are from Mexico (a non-English-speaking country), with parallel ages of arrival. If we observe a difference between the wages of the two Jamaicans, we could attribute it to secular age-at-arrival effects. 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 the English language before immigrating. As such, the Jamaicans can be used to control for the non-language age-at-arrival effects. Any differences between the Mexicans in excess of the differences between the Jamaicans can be attributed to language effects, because the Mexican child who immigrated younger has an earlier age of first exposure to English.

Thus, when we say that we are using age at arrival interacted with coming from a non-English-speaking country as an instrumental variable for English proficiency, we mean that we are only using the part of the variation in English proficiency that is attributable to the biological constraints to human language acquisition that vary by age. We are not using the rest of the variation because this may be correlated with omitted variables or may be affected by the outcome itself, which would lead to biased estimates of the effects of English proficiency. Thus, we are keeping only the "good" variation in English proficiency—good in the sense that we can consider it exogenous conditional on our control variables—and leaving out the "bad" variation when we use our instrumental variables method. The specific instrumental variables estimator we use is the two-stage least squares (2SLS) estimator.

We are interested in the effect of the English on a variety of economic and social outcomes. For each outcome, we will be estimating a regression model of the following form:

$$(1) \qquad y_{ija} = \alpha + \beta\, ENG_{ija} + \delta_a + \gamma_j + \mathbf{w}_{ija}'\rho + \varepsilon_{ija}$$

for individual i born in country j arriving in the U.S. at age a. y_{ija} is the outcome, ENG_{ija} is the measure of English proficiency, δ_a is a full set of age-at-arrival dummies, γ_j is a full set of country-of-birth dummies and $\mathbf{w}_{ija}$ is a vector of exogenous explanatory variables (including age, sex, race and Hispanic origin). The parameter of primary interest is β, the effect of English proficiency on the outcome. Because English-language skills are endogenous due to reasons discussed in Section I (e.g., omitted variables, reverse causality, measurement error), OLS estimates of Equation 1 will tend to be biased. To obtain consistent estimates of the parameters in Equation 1, 2SLS estimation, with the aforementioned interaction between age at arrival and coming from a non-English-speaking country as the excluded instrument for the English proficiency of the childhood immigrant.

IV. Effect of English Proficiency on Economic and Social Outcomes

A. Earnings and Education

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 research design described above and applied it to 1990 Census data. In this subsection, we extend this paper by repeating the empirical analysis using the 2000 Census sample described in Section II. We also describe the main findings in this paper.

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. 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? From this, we proceed to estimate Equation 1 by 2SLS. We find that English proficiency has a significant positive effect on wages. In particular, raising English proficiency by one unit (e.g., from speaking English well to very well, or from speaking English not well to well) raises wages by 37%. This 2SLS estimate is larger than OLS estimate (which is 24%), albeit less precise. We explored this at length in Bleakley and Chin (2004); the main conclusion is that both classical measurement error and omitted ability appear to be biasing the OLS estimate, with the downward bias caused by the former more than offsetting the upward bias of the latter.

Much of the effect of English proficiency on wages is mediated through education. 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 earnings by age at arrival (in Figure 2), which is suggestive that the effect of English proficiency on educational attainment may in large part be responsible for the effect of English proficiency on wages. In other words, English proficiency raises wages mainly in an indirect

way, through raising educational attainment. 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. We find that a one-unit increase in English proficiency raises completed schooling by over three years, with much of the effect coming from moving childhood immigrants away from dropping out of high school to completing high school and attending college. This is a large effect. Assuming standard returns to a year of schooling, the educational channel accounts for about 85% 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 (collectively accounting for the remaining 15%).

In Akbulut, Bleakley and Chin (2007), we test whether the effects of English proficiency are the same for Hispanic immigrants and non-Hispanic immigrants. We find that for wages and years of schooling, there is no statistically significant difference in the effect of English proficiency for Hispanics compared to non-Hispanics. The common effect of English proficiency on wages 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?

The results summarized in this subsection point to English-language skills being tremendously important for labor market success in the U.S. Much of this labor market success seems attributable to the impact of English proficiency on a pre-market factor, investment in

education.

B. 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. In addition, limited English proficiency may cause people to self-segregate (e.g., live in ethnic enclaves) because they cannot interact easily with the mainstream population. We apply the research design described above to marriage and fertility outcomes in Bleakley and Chin (2007b), and to location of residence in Akbulut, Bleakley and Chin (2007).

With regard to current marital status of childhood immigrants, we find that English proficiency significantly reduces the probability of being married with spouse present. This negative effect is attributable primarily to higher divorce rates among women and a lower propensity to ever marry among men. Thus, English proficiency, which would expand the set of people with whom one can communicate, does not raise marriage rates. English proficiency appears to be increasing acceptance and practice of divorce, and at least among men delaying marriage, both which mark a convergence to U.S. norms since U.S. divorce rates and age at first marriage tend to be higher than in the average source country.

Not only is the likelihood of being married with spouse present lower, but spousal characteristics change as well with changes in English proficiency. We found that intermarriage—which we alternately define as marrying someone with a different country of birth or a different ancestry—was significantly less common when English proficiency increases.

That is, English proficiency makes childhood immigrants less likely to marry someone of the same nativity and ethnicity, with much of the gap made up for by marrying U.S.-born natives. Some researchers use intermarriage as a measure of social assimilation, so this means that English proficiency expedites social assimilation.

Another finding in relation to spousal characteristics is that when a childhood immigrant's English proficiency increases, his or her spouse's English proficiency, years of schooling and earnings are higher too. This suggests that there is strong assortative matching along these dimensions, i.e., people tend to marry people who are similar to themselves. This is an important finding because it means that marriage decisions magnify existing differences across individuals along linguistic lines. A child who has one limited-English-proficient parent probably has both parents who are limited-English-proficient (and will thus tend to have parents who have less education and income). Thus, given how people select spouses, if we wish to help the second generation, it might be efficient to target scarce resources to those children with limited-English-proficient parents; children in other immigrant households already have access to early exposure to English.

We estimated the effect of English proficiency on the number of children living in the household.[9] We find that English proficiency reduces the number of children. Some of this effect comes from the lower likelihood of men to ever marry; fertility rates are higher among married people than unmarried people. But most of the effect comes from a reduction in family size. While among married childhood immigrants there is no impact on whether one has at least one child, there is a decrease in having larger family sizes. 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

[9] This is the only measure of total children born afforded by the 2000 Census. However, we have repeated the analysis with 1990 Census data, which asked about children ever born, and obtained similar results.

family.

An additional marker of social integration of immigrants is the extent to which to 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 measure that we use is that fraction of the population in one's PUMA that shares the same country of birth as the childhood immigrant. 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 measure used is the fraction of one's PUMA that shares the same ancestry. We find that English proficiency significantly reduces the probability that a childhood immigrant lives in PUMAs with a high fraction from own country or with same ancestry. This is consistent with more English proficient immigrants moving away from ethnic enclaves, to integrating geographically with mainstream America.

In Akbulut, Bleakley and Chin (2007), we found that English proficiency tended to have significantly larger size effects on the marriage, fertility and location of residence outcomes for non-Hispanics than non-Hispanics. This is an interesting result because it suggests either the two groups are converging to U.S. norms at different rates (e.g, Hispanic immigrants may take longer), or that one or both groups are converging to something other than the U.S. norms (e.g., instead of all its immigrants joining the melting pot, America will be a multi-cultural society).

The results summarized in this subsection show that English-language skills play an

important role in determining outcomes in the private sphere of childhood immigrants. Of course we could have predicted given English proficiency's effects on education and wages that there would be effects on neighborhood of residence, fertility and marriage outcomes. However, even when we control for education and income in our regression models, we still find a significant effect of English proficiency on these social outcomes. Thus, there are additional mechanisms at work.

C. Children's English Proficiency and Education

The children of immigrants lag behind the children of natives in educational and economic achievement. The gap between the second generation and other natives appears to be linked to English-language proficiency. Over 8% of students enrolled in U.S. public schools are classified as limited English proficient, of which three-quarters are Hispanic. Surprisingly, only half of these limited-English-proficient students are actually foreign-born. The rest are U.S.-born, mostly the children of immigrants. Evidently, growing up and attending school in the U.S. are not sufficient for developing English-language proficiency. In this setting, parental English-language skills might have an important role in the educational (and therefore economic) progress of their children.

In Bleakley and Chin (2007a), we estimate the effect of immigrant parents' (first generation's) English-language skills on their U.S.-born children's (second generation's) English-language skills and educational outcomes. As mentioned in Section II, we are using data on all the U.S.-born children of the childhood immigrant sample we have been using for the preceding analyses who are aged 17 or younger and still residing at home. Before this paper, there had been no work on the intergenerational impact of immigrant parents' destination-country skills.

We know from Figure 1 that the critical period leaves its "footprint" on English proficiency for the first generation, but does this pattern carry forward to the second generation? We see in Figure 4 that it does. For young children, the shape of the relationship between their English skill and their parent's age at arrival (the diamond-marker line) corresponds approximately to that seen for the first generation in Figure 1. We take this as evidence that immigrant parents' English-speaking proficiency has a significant effect on U.S.-born children's English-speaking proficiency while the children are young. On the other hand, as these children grow, they are exposed more to their peers and teachers, and parental English matters less and less. By the time the child reaches middle school, parental English ceases to matter; children with older-arriving parents no longer have worse English-language skills at least as measured by the rough measure of English-speaking skills afforded by the Census. The square-marker line in Figure 4 shows that there is no relationship between parents' age at arrival and children's English proficiency using Census data for children aged 16-17; the flat relationship holds beginning with child age 12.

Because children with limited-English-proficient parents enter school with poorer English-language skills on average, a logical question to ask is whether this early deficiency has longer-term consequences. Using the same research design (described in Section III), we estimate the effect of parental English on children's educational outcomes, and find that children with limited-English-proficient parents are more likely to drop out of high school, be below their age-appropriate grade and not attend preschool. Relatedly, we examine standardized scores in the NLSY (whose respondents took the ASVAB test) as a function of whether English was the principal language spoken in the childhood home. The estimated probability distribution of scores for vocabulary (ASVAB test #3), so decomposed, is seen in Figure 5. (Raw scores are shown in Panel A, while Panel B shows the scores controlling for parental education.) The

distribution of scores for those who spoke a foreign language at home (the blue line) is shifted to the left relative to those who spoke English at home (the red line), and this is particularly evident in the tail of the distribution. Similar results were found for paragraph comprehension and arithmetic (the latter being learned in the early years by rote, linguistic memorization), but not for "coding speed", which should be least affected by language deficits.

Bleakley and Chin (2007a) did not find any differences in effect of parental English on children's educational outcomes between Hispanic and non-Hispanic childhood immigrants. Additionally, much of the intergeneration effects remain even if one controls for parental education and family income, suggesting that much of the way parental English benefits children is through direct channels, such as parents helping children with schoolwork or maneuvering their children into the right schools and classes.

The results summarized in this subsection point to parental English skill being an important input into their children's human-capital formation.

V. Discussion

Childhood immigrants from non-English-speaking countries who arrive in the U.S. after the critical period look different from earlier arrivers along a number of dimensions. Most notably, they are less proficient in English, they earn less, and both they and their children obtain less education. The coincidence of these shifts with closing of the critical period points to important role played by English proficiency in their adaptation to the U.S.

While English skill appears linked to numerous important outcomes, further research is needed before we can make definitive policy recommendations based on our findings. Policymakers might consider favoring the migration of families with children who are still in the critical period (in effect swapping a person that looks like a first-generation immigrant for one

that resembles a second-generation immigrant), although it would be difficult to implement such a policy without creating perverse incentives. On the other hand, our research indicates that poor English derails educational progress for the childhood immigrants and their children. For the childhood immigrants themselves, entering the school system with limited English is no doubt a barrier to comprehension and achievement, which suggests a transition period that includes some instruction in the native language. On the other hand, this crowds out learning English, which is best begun before the critical period has completely ended. This is no doubt a hard balance to strike, and might not be the same for all children in any case. For their second-generation children, it appears that having the children start school with limited English is an impediment, but educators face exactly the same tradeoff as for their first-generation parents. One additional option is available, however, because they are already in the country prior to starting school. It might be productive, therefore, to increase exposure to English through preschool or family-literacy programs, although this remains more conjecture than fact at this point.

References

Akbulut, Mevlude, Hoyt Bleakley and Aimee Chin, "The Effects of English Proficiency Among Childhood Immigrants: Are Hispanics Different?," University of Houston Mimeo, August 2007.

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, "What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital Among Immigrants," Forthcoming, *Journal of Human Resources* (2007a).

Bleakley, Hoyt and Aimee Chin, "English Proficiency and Social Assimilation Among Immigrants: An Instrumental-Variables Approach," University of Houston Mimeo, March 2007b.

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 2006.

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.

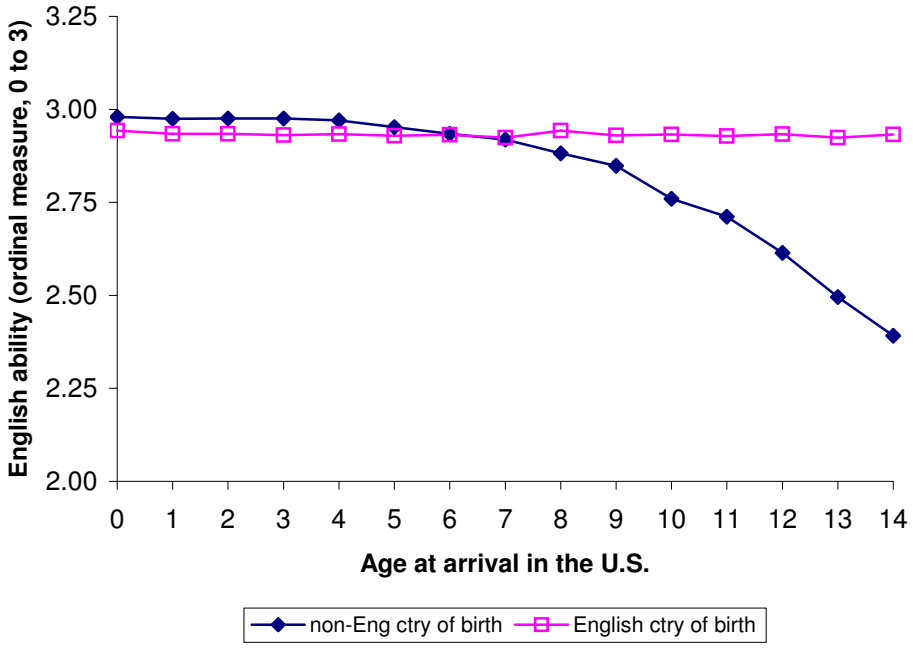
Lenneberg, Eric H., *Biological Foundation of Language*, New York: Wiley & Sons, 1967.

Newport, Elissa L., "Critical Periods in Language Development," in L. Nadel (Ed.) *Encyclopedia of Cognitive Science*, London: Macmillan Publishers Ltd./ Nature Publishing Group, 2002.

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).

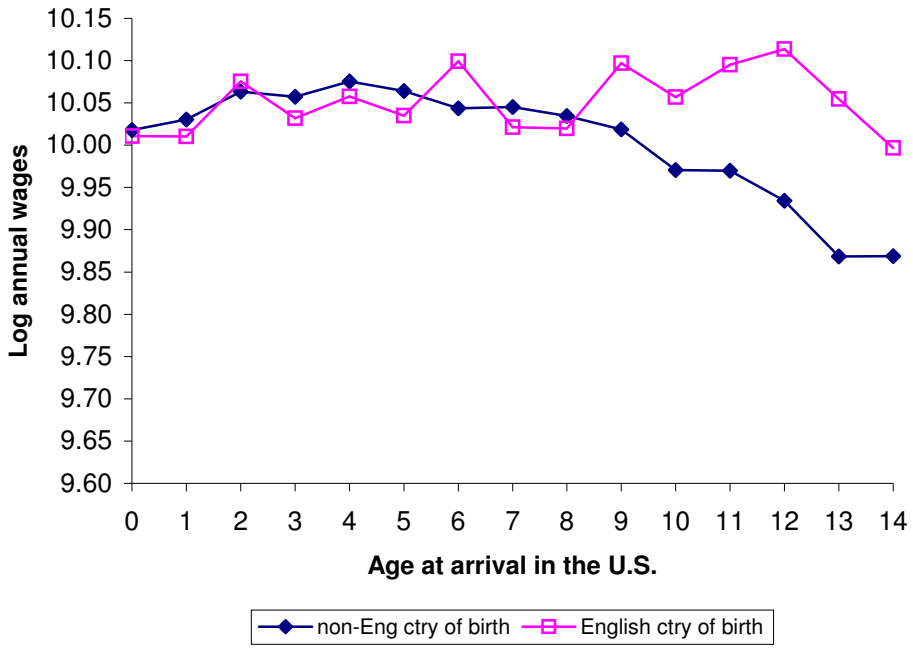
World Almanac and Book of Facts, 1999, New York: World Almanac Books (1999).

Figure 1. English-Speaking Ability by Age at Arrival



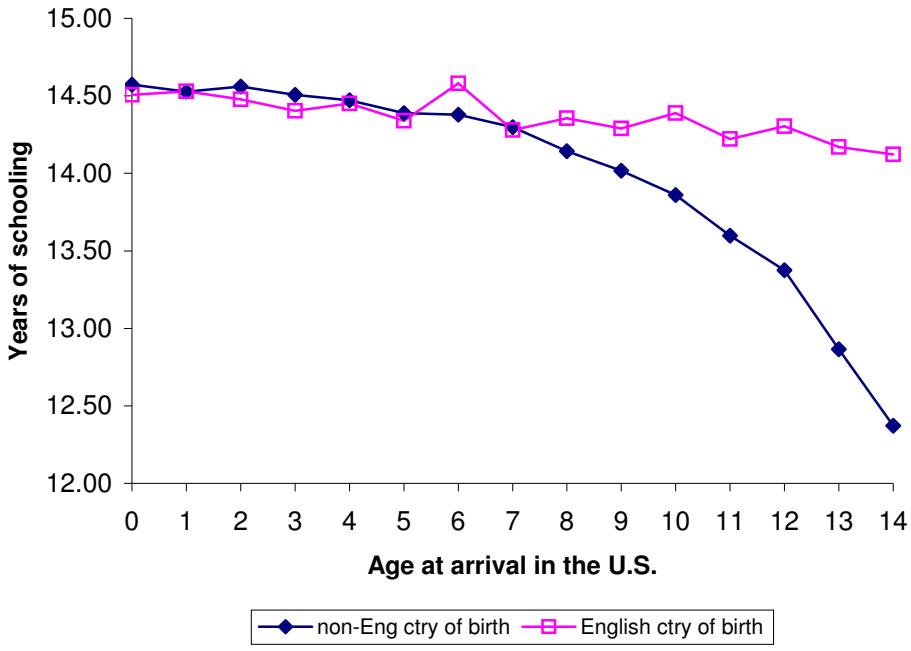
Notes: Data are from the 2000 IPUMS. Sample size is 191,534 (composed of people who immigrated to the U.S. before age 15 and are currently aged 25-55, and with nonmissing English variable). Displayed for each age at arrival is the mean English-speaking ability. Means are weighted by IPUMS weights, and 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



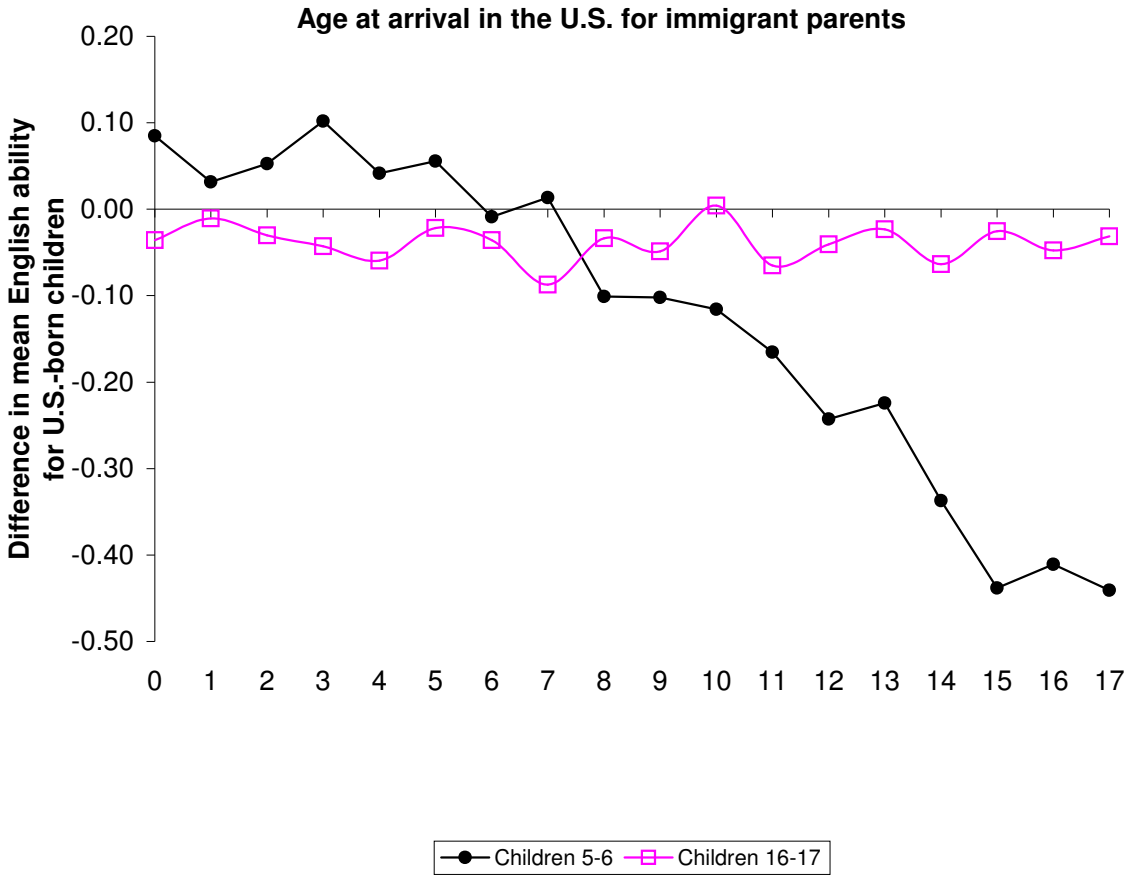
Notes: Data are from the 2000 IPUMS. Sample size is 127,616 (composed of people who immigrated to the U.S. before age 15 and are currently aged 25-55, and with nonmissing English and wage variables). Displayed for each age at arrival is the mean log wage. Means are weighted by IPUMS weights, and regression-adjusted 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



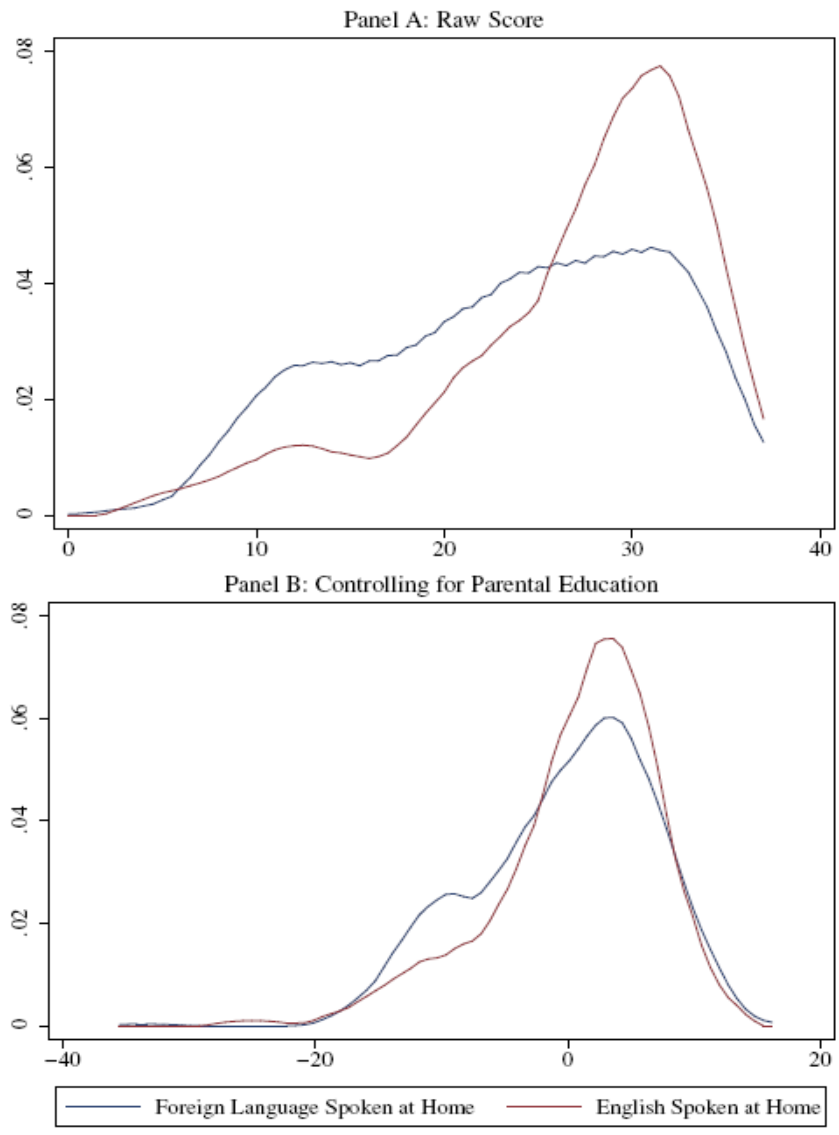
Notes: Data are from the 2000 IPUMS. Sample size is 126,293 (composed of people who immigrated to the U.S. before age 15 and are currently aged 25-55, and with nonmissing English, wage and schooling variables). Displayed for each age at arrival is the mean log wage. Means are weighted by IPUMS weights, and regression-adjusted for age, race, Hispanic and sex dummies. The race categories used were White, Black, Asian & Pacific Islander, Multiracial and Other.

Figure 4. Child's English-Speaking Ability by Immigrant Parent's Age at Arrival and Child's Age



Notes: Data are from the 2000 IPUMS. Sample consists of children who were born in the U.S., currently aged 5-17, living with at least one biological parent who immigrated to the U.S. before age 18 and is currently aged 25-55, and with nonmissing language variable for both child and parent. Displayed for each parental age at arrival is the difference in mean English-speaking ability between between children with parents from non-English-speaking countries and children with parents from English-speaking countries. Means are weighted by child-level IPUMS weights, and regression-adjusted for age, race, Hispanic and sex dummies for parent and age and sex dummies for child. 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 5. Kernel Density Estimates of ASVAB Scores by Home Language



Blue line: Foreign Language Spoken at Home Red line: English Spoken at Home

Notes: Data are from the NLSY-1979. Sample size is 751 (composed of children who were born in the U.S. with at least one foreign-born parent and with nonmissing ASVAB scores). Displayed in each panel is the kernel density estimate for the indicated ASVAB subtest, decomposed by whether a foreign language was spoken at the respondent's childhood home blue line for yes, red line for no).

Appendix Table 1. Individuals by Country of Birth

Panel A. English-speaking countries (=Control Group)

Rank by N	country	N	% of group
1	Canada	8,962	34.6%
2	England	6,121	23.6%
3	Jamaica	3,180	12.3%
4	United Kingdom, ns	1,242	4.8%
5	Trinidad & Tobago	1,014	3.9%
6	Guyana/British Guiana	991	3.8%
7	Scotland	803	3.1%
8	Ireland	565	2.2%
9	Australia	543	2.1%
10	South Africa (Union of)	308	1.2%
11	Barbados	297	1.1%
12	Bermuda	283	1.1%
13	Bahamas	258	1.0%
14	U.S. Virgin Islands	256	1.0%
15	Belize/British Honduras	251	1.0%
16	New Zealand	131	0.5%
17	Antigua-Barbuda	112	0.4%
18	St. Vincent	90	0.3%
19	Liberia	84	0.3%
20	Grenada	82	0.3%
21	St. Kitts-Nevis	73	0.3%
22	Wales	71	0.3%
23	Northern Ireland	69	0.3%
24	Zimbabwe	63	0.2%
25	St. Lucia	59	0.2%
26	British Virgin Islands	1	0.0%
27	Anguilla	1	0.0%
	Total English-spking obs	*25,910*	100.0%

Panel B. Non-English-speaking countries (=Treatment Group)

Rank by N	country	N	% of group
1	Mexico	47611	28.7%
2	Germany	19445	11.7%
3	Puerto Rico	13203	8.0%
4	Cuba	9389	5.7%
5	Vietnam	6334	3.8%
6	Italy	5642	3.4%
7	Japan	5475	3.3%
8	Korea	3926	2.4%
9	El Salvador	3233	2.0%
10	Dominican Republic	3103	1.9%
11	France	2466	1.5%
12	Portugal	2390	1.4%
13	Colombia	2266	1.4%
14	Taiwan	1987	1.2%
15	China	1854	1.1%
16	Laos	1668	1.0%
17	Poland	1499	0.9%
18	Haiti	1468	0.9%
19	Guatemala	1452	0.9%
20	Greece	1427	0.9%
21	Panama	1415	0.9%
22	South Korea	1344	0.8%
23	Ecuador	1316	0.8%
24	Iran	1314	0.8%
25	Spain	1207	0.7%
26	Netherlands	1188	0.7%
27	Nicaragua	1186	0.7%
28	Cambodia (Kampuchea)	1081	0.7%
29	Israel/Palestine	983	0.6%
30	Peru	933	0.6%
31	Argentina	926	0.6%
32	Thailand	818	0.5%
33	Honduras	801	0.5%
34	Austria	781	0.5%
35	Brazil	751	0.5%
36	Africa, ns/nec	680	0.4%
37	Venezuela	644	0.4%
38	Lebanon	637	0.4%
39	Hungary	550	0.3%
40	Turkey	544	0.3%
41	Azores	492	0.3%
42	Yugoslavia	487	0.3%
43	Costa Rica	466	0.3%
44	Chile	465	0.3%
45	Egypt/United Arab Rep.	454	0.3%
46	Iraq	443	0.3%
47	Other USSR/Russia	416	0.3%
48	Belgium	392	0.2%
49	Romania	358	0.2%
50	Indonesia	358	0.2%
	subtotal, top 50 countries	159,268	96.1%
	subtotal, other (91) countries	6,397	3.9%
	Total non-Eng-spking obs	*165,665*	100.0%

Notes: Information on each country's official languages is from the World Almanac. Recent adult immigrants from the 1980 IPUMS were used to divide English-official countries into English-speaking (at least 50% of recent adult immigrants did not speak a language other than English at home) or Other. The countries in the "Other" category are the Philippines, India, Hong Kong, Guam, Pakistan, Nigeria, American Samoa, Fiji, Tonga, Ghana, Kenya, Singapore, Dominica, Tanzania, Uganda, Sierra Leone, Senegal, Malta, Zambia, Micronesia, Marshall Islands, Papua New Guinea, Kiribati, Palau, Gambia, Malawi, Mauritius and Swaziland; people from these countries have been dropped from the empirical analysis. Above tabulations by country of birth use following sample: individuals from the 2000 1% and 5% PUMS files who are currently aged 25-55, immigrated to the U.S. before age 15 and has nonmissing age, year of immigration, country of birth and English variables. Country refers to IPUMS detailed birthplace code.