# The Effects of Language Skills on the Economic Assimilation of Female Immigrants in the United States

Mary A. Silles
Department of Economics, Hull University Business School, Hull HU6 7RX, United Kingdom. Email: M.Silles@hull.ac.uk; Tel.: +441482463828; Fax: +441482 463484

**Abstract**

This paper uses recent data from the American Community Survey between 2010 and 2015 to investigate the effect of language skills on women's economic assimilation who immigrated to the United States as children. The problem of endogenous language acquisition and measurement error in the language variable is addressed utilizing the phenomenon that younger children learn languages more easily than older children to construct an identifying instrument. Two-stage-least-squares estimates suggest that greater English proficiency has a positive effect on a number of indicators of economic assimilation of adult women including several measures of labor supply and earnings. A range of sensitivity tests are undertaken to check the validity of these results.

Key words: Immigration; language skills; labor market performance
JEL Codes: J24, J15

* This paper was accepted in May 2018 for publication in *The Manchester School.*

## 1. Introduction

New interest in the language capital of immigrants has been fostered by the dramatic increase in immigration in recent years to the United States. Since 1990, the size of the foreign-born population has increased from 19.8 million in 1990 to 42.4 million in 2014 (Batalova and Zong 2016)[1]. Although the gender composition of immigration has fluctuated slightly over this period, female immigration represented approximately 51% of the foreign-born population in 2014. Despite the high proportion of female immigrants, little is known about the assimilation process of women. The purpose of this study is to examine the extent to which language ability influences the economic integration of women who immigrated to the United States when children. Language skills are among the most important forms of human capital and may affect central aspects of immigrant women's lives both in and outside the labor market.

In terms of labor market performance, the vast majority of previous studies have predominantly analyzed the importance of language skills on earnings for male immigrants (Carliner, 1981; Chiswick, 1991; Dustmann, 1994; Grenier, 1984; Kossoudji, 1988; McManus et al., 1983; Rivera-Batiz, 1990). These studies using simple cross sectional analysis generally conclude that migrants who are fluent in the dominant language of the host country earn higher wages. There is also a growing literature examining the effect of language skills across a variety of domains outside the labour market. These include studies that estimate the correlation between language fluency and marriage (Stevens and Swicegood 1987; Davila and Mora 2001; Meng and Gregory 200; Duncand and Trejo 2007), fertility (Sorenson 1988; Swicegood et al. 1988), residential location (Funkhouser and Ramos 1993; Toussaint-Comeau and Rhine 2004) and educational attainment (Ruiz-de-Velasco et al. 2000).

Research in this area presents two challenges. First, language proficiency may be positively correlated with a wide array of other factors such that immigrants who speak English poorly may well be the same individuals whose economic status would have been lowest in any case. Second, language measures usually reported in survey data suffer from serious measurement error. This problem usually leads to downward bias in the estimates of language effects on economic outcomes. Previous attempts to address both sources of bias have relied

[1] The numbers reported by Batalova and Zong (2016) are based on the 1990 decennial census and the 2014 ACS survey.

on instrument variables (Angrist and Lavy 1997; Bleakley and Chin 2004, 2010; Budría and Swedberg 2015; Chiswick and Miller 1995; Di Paolo and Raymond 2012; Dustmann and Soest 2002; Dustmann and Fabbri 2003; Guven and Islam 2015; Miranda and Zhu 2013). The common finding across these labour market studies is that the effect of language on earnings is underestimated in OLS regressions. Most notable and closest to the present study is Bleakley and Chin (2004) for the United States who utilize the subsample of foreign-born childhood immigrants from the 1990 census. Their instrumental strategy, which is the one adopted in this study, is based on the well-known finding that differences in language proficiency open up early with younger children having a far greater capacity for learning a new language than do older children. Following the same instrumental strategy, a handful of other studies have examined the causal effect of language skills on social integration (Bleakley and Chin 2010), health (Guven and Islam 2015) and the educational success of US-born children of immigrants (Bleakley and Chin 2008).

This paper contributes to the large and growing literature on the effect of language fluency on immigrants' assimilation by focusing on female immigrants who came to the United States as children. Much of the literature to date has concentrated on estimating the effects of language fluency on earnings functions with most previous studies considering only male migrants. There are few papers that investigate the economic assimilation of female immigrants[2]. This paper focuses attention on women and the causal effect of English-speaking ability on a broader range of outcome variables than used in existing studies including labour force participation, measures of the intensity of work, earned income, poverty status, health insurance coverage, assimilation into home ownership and completed years of schooling. The analysis is based data from the American Community Survey (ACS) between 2010 and 2015. This data set when pooled is large enough to produce results that can be generalized to the overall population of childhood immigrants in the United States. Given the large presence of women as part of the immigrant flow, understanding the process of economic assimilation for this group is important both for predicting future changes in the size of the labour force and for analyzing the sources of the differences in economic outcomes. Separate estimates of

[2] Bleakley and Chin (2004) include female immigrants in their sample but they do not analyze language effects separately for males and females. Dustmann and van Soest (2002) study wage effects of language skills for women but they have difficulties in finding suitable instrumental variables. Miranda and Zhu (2013) study the immigrant-native wage gap for female employees in the UK. Yao and van Ours (2015) for the Netherlands consider the effect of language skills on labor market performance separately for men and women.

women's language effects should also be undertaken as women are unlikely to experience the same process of assimilation as men and may differ in their incentives to acquire a new language.

My main findings are that improvements in language ability positively impact all measures of labour supply and substantially raise women's earned income. Furthermore, language proficiency generates important benefits beyond those linked to the labour market including raising the numbers covered by health insurance and lowering the proportion living in poverty. While there also appears to be a large positive effect on years of schooling, there is a sizeable direct impact of English-speaking ability on economic assimilation once educational attainment is controlled for in estimation.

This paper continues as follows. Section two provides a description of the data. Section three describes the estimation method. Section four discusses the main results. Section five provides an extensive series of sensitivity tests. Section six presents some concluding comments.

## 2. Data

I use the Integrated Public Use Micro Sample (IPUMS) data from the ACS between 2010 and 2015. These surveys contain information on each individual's year of arrival in the United States, English language proficiency, labor market performance and other relevant variables. An attractive feature of the ACS is that it contains information on the exact year of arrival in the United States, making the IV strategy used here more precise than it would be if multiyear intervals were used instead[3]. For this analysis, attention is restricted to foreign-born individuals who arrived in the United States before age 15 (that is age 0 – 14). I use age 14 as the maximum age of arrival so as to mitigate concerns surrounding adolescents who immigrated of their own volition particularly from Mexico and Central America to look for work. For immigrants younger than age 15, age at arrival is not a choice variable but a decision taken by their parents. Given the primary focus is on the impact of language fluency on labor market performance, the sample is further restricted to individuals from 25 to 55 years of age. Thus, the sample consists of individuals who moved to the United States between 1955 and 2003.

[3] Previous work by Bleakley and Chin (2004) utilized the 1990 census which contains year of arrival in multiyear intervals.

The language questions in the ACS are from self-reported responses to two questions beginning with: “Does this person speak a language other than English at home?” For individuals who respond affirmatively, a follow-up question is asked: “How well does this person speak English?” Individuals can choose from four possible responses: very well, well, not well, or not at all. These two questions are coded as one ordinal measure of English-speaking ability that is equal to 3 for those who speak only English or speak it very well, equal to 2 for those who speak it well, equal to 1 for those who speak it not well, and 0 for those who do not speak English at all.

The ACS contains information on a variety of dimensions of labour supply behavior and it is helpful to consider each of them separately. Here I consider five: (1) whether the respondent participated in the labor force at the time of the survey; (2) whether the individual worked for pay during the calendar year preceding the survey; (3) the number of hours usually worked in the calendar year preceding the survey; (4) the number of weeks worked during the calendar year preceding the survey; and (5) annual earned income in the calendar year preceding the survey[4]. I also focus on four other indicators of assimilation: (5) home ownership; (6) health insurance coverage; (7) poverty status (below the poverty line)[5]; and (8) years of completed schooling.

The ACS survey contains information on country of origin. I classify these countries into three language groups: non-English-speaking countries of birth; countries of birth where English is spoken by the majority of the population; and other countries of birth with English as an official language but spoken by less than the majority of the population[6]. The immigrants from non-English-speaking countries of birth constitute the “treatment” group, and individuals from predominantly English-speaking countries make up the “control” group. Individuals from countries of birth with English as an official but not predominant language

[4]I set earnings to zero for those who did not work during the year prior to the survey.
[5] The original variable expresses each family's total income for the previous year as a percentage of the poverty threshold established by the Social Security Administration in 1980, adjusted for inflation. I construct a binary variable which takes a value equal to 1 if an individual is living in a family whose income is on or below this threshold, and zero otherwise.
[6] The categorization of countries is undertaken using the CIA World Factbook, except for Puerto Rico which I classify as non-English speaking even though English is an official language.

are not included in the main analysis as I cannot be sure how much exposure to the English language these individuals would have had prior to immigrating. Table A1 in the Appendix presents the division of the sample by country of birth, and also provides the classification of countries by English-speaking status.

The number of original observations drawn from the six ACS surveys was 690,386. This sample was comprised of all women who were born outside the United States and 25 to 55 years of age at the time of the survey. This study focuses on the subset of these women who arrived in the United States between age 0 and 14 (inclusive). This requirement reduced the number of observations to 159,090. For the reason mentioned above, I also dropped those who were born in countries where English is an official language but not a predominant language. This restriction further reduced the sample to 144,030 observations. Finally, 1,528 observations with missing values on variables used in the analysis were deleted[7]. This yielded a final sample of 142,502 women. It is quite possible that my sample of childhood immigrants have different characteristics from other samples of immigrants. So the results presented here may be showing evidence of a marginal effect of English-speaking ability over a sample that is not characteristic of most immigrants in the United States. In addition, the present instrumental variable design does not allow me to investigate the importance of English-speaking language ability for assimilation among those who arrived as adults.

Table 1 provides summary statistics for immigrants from non-English-speaking and English-speaking countries of birth used in this study, disaggregated by age at arrival. Almost 90% of female immigrants in the sample were born in a country where English is not spoken by the majority of the population. Approximately half of immigrants who arrived before age 14 came to the United States when they were less than six years old. Reported English-speaking ability is somewhat higher for younger arrivers in the non-English-speaking subsample. The mean score on English-speaking ability is 2.9 for the sample of immigrants that arrived before the age of six, and 2.5 for the sample that arrived between age 7 and 14. The fairly high level of self-reported language proficiency among immigrants from non-English-speaking countries suggests that there might be considerable measurement error in this subjective measure of language skill. The tabulations also show that young immigrants from non-English speaking countries are generally better off than those who arrived after age 6. In

[7] Specially, home ownership status was missing for 1,397 individuals and labor market information was inaccurate for a further 131 individuals.

terms of labor market performance, early arrivers have a higher probability of being employed, work for more weeks per year and hours per week and also have higher annual earnings as compared with those who arrive past age six. In terms of other indicators of assimilation, schooling, home ownership and health insurance rates are higher while poverty rates are lower for younger immigrants from non-English speaking countries.

## 3. Method

The objective of this paper is to identify the causal effect of language proficiency on a number of economic indicators of assimilation. I begin by assuming that language fluency is exogenous to each outcome and measurement error in reporting is absent. Hence, I estimate the following model by OLS:

$$Y_{ija} = \beta_0 + \beta_1 l_{ija} + \beta_2 X_{ija} + \delta_a + \gamma_j + \varepsilon_{ija} \qquad (1)$$

for individual $i$ born in country $j$ arriving in the United States at age $a$. Here $Y_{ija}$ denotes the outcome variable (e.g. employment participation, earned income), $l_{ija}$ is a (ordinal) measure of English fluency, $\delta_a$ is a full set of age of arrival fixed effects, and $\gamma_j$ is a series of dummy variables for country of birth fixed effects[8]. The vector $X_{ija}$ includes exogenous explanatory variables (such as age, race, Hispanic, and survey year). The error term $\varepsilon_{ija}$ captures the effects on the dependent variable of any omitted or unobserved variables.

The assumption of exogeneity of language problems and absence of measurement errors may lead to the OLS estimates being biased. Language endogeneity is likely to lead to an upward bias in the OLS estimate of the effect of language skills. This might arise if more able immigrants find it easier to learn English and at the same time obtain other useful skills that may make them more productive in the labor market[9]. By contrast, measurement error in self-reported language information is likely to lead to underestimation of the language effects (Dustmann and van Soest 2002).

To correct for these two potential sources of bias, I adopt an instrumental variable approach similar to that introduced by Bleakley and Chin (2004). This strategy is based on the "early

[8] I introduce age at arrival as a right-hand side control variable in outcome equations to allow for the fact that age at arrival can affect labor market performance and other outcomes directly through non-language channels.

[9] Additionally, employed immigrants have more opportunities to practice speaking English leading to reverse contributions from labor market success to language proficiency.

sensitivity" theory of language acquisition from psychology whereby younger children have a language-learning advantage over older children and adults (Lenneberg 1967)[10]. The mechanism is neurological with the cerebral organization for language learning much more suited to acquiring a second language in early childhood. The consequence of these biological constraints is that younger immigrants should reach higher levels of final proficiency in English than otherwise similar children who arrived at older ages.

The "early sensitive period" hypothesis appears to be borne out by the ACS data. Figure 1A plots the mean English-speaking ability among childhood immigrants from both English and non-English speaking countries by age of arrival in the United States. The diamond-marked line shows that for immigrants from English-speaking countries there is no relationship between age of arrival and eventual language fluency as indicated by the nearly flat age of arrival profile for this group. This is not surprising as children from English-speaking countries tend to arrive fully fluent in English. For immigrants from non-English speaking countries there is a strong negative association between age of arrival and language skills as predicted. Although there is no sharp break at a particular age, the square-marked line shows that immigrants who arrive before age six from non-English speaking countries have similar levels of speaking ability to their counterparts from English-speaking countries. This result is in line with previous research which finds that the younger the child, the greater their ability to learn a new language (Bleakley and Chin 2004, 2008).

The empirical strategy used here compares younger and older arrivers from non-English speaking countries after removing the age-of-arrival effects for immigrants from English-speaking countries. Thus Figure 1B graphically displays by age of arrival the mean English-speaking ability of immigrants from non-English speaking countries with the mean for English-speaking countries subtracted out. Individuals who arrived at age six or earlier from non-English speaking countries speak at least as well as their counterparts from English-speaking countries. Past age of arrival six, immigrants from non-English speaking countries have significantly lower English-speaking ability, and the severity increases almost linearly with age of arrival thereafter.

[10] Lenneberg (1967) proposed 14 years of age as the critical turning point for native-like acquisition of a new language. See for a recent review of the psychological literature in this field Newport (2002).

Because of this, in the main part of the analysis, the instrumental variable is $z_{ija}$ is defined as the interaction between a dummy variable for having arrived by age six, labeled as $A_a$, and the dummy variable $N_j$ indicating whether or not an immigrant originated from a non-English speaking country of origin. The key assumption underlying this instrumental variable is that past age of arrival six significant differences in language skill between immigrants from English-speaking countries and non-English speaking countries emerge. The following first-stage equation is then estimated by OLS:

$$l_{ija} = \alpha_0 + \alpha_1 z_{ija} + \alpha_2 X_{ija} + \delta_{1a} + \gamma_{1j} + \nu_{ija} \qquad (2)$$

where $z_{ija} = A_a N_j$. As before the $\delta_{1a}$ represents age of arrival fixed effects, and $\gamma_{1j}$ denotes country of birth fixed effects. In the second-stage equation, the predicted values of $l_{ija}$ are entered into equation (1) to yield a consistent estimate of the effect of language capital. Standard errors in the second stage are adjusted to account for the use of a predicted probability. The 2SLS coefficient can be interpreted as the effect of language proficiency under the assumption that non-language age-of-arrival effects are the same for immigrants from non-English-speaking countries as those from English-speaking countries.

It is worth emphasizing that the instrument for language proficiency is not age of arrival. This is because age at arrival probably affects immigrants' economic outcomes through channels other than language acquisition. For example, younger immigrants not only receive greater exposure to more English, they also may find it easier to acquire knowledge of the culture, values and institutions of the United States. Therefore, age of arrival may reflect social acculturation as well as language ability. For this reason, the identifying instrument is an interaction term between age at arrival and non-English-speaking country of birth which removes non-language effects correlated with age of arrival. The English proficiency of immigrants from English-speaking countries is not sensitive to the age at which they arrive in the United States but these children experience everything else that children from non-English speaking countries experience. Thus, the key assumption is that any differences in indicators of assimilation between early and late arrivers from non-English speaking countries that arise over and above the differences experienced by those from English-speaking can be attributed to language ability. This instrumental variable strategy is equivalent to estimating a difference-in-difference model.

## 4. Results

Before exploring the socioeconomic effects of language proficiency, I consider the first-stage estimates from OLS models which show the effect of the instrument on English-speaking ability. These results are displayed in table 2. Just as graphically illustrated in figure 1, age of arrival is a strong predictor of language fluency. The estimates in column 1 of table 2 show that immigrants who arrived before the age of six from non-English speaking countries have statistically significant higher English-speaking ability as adults than those who arrived between the ages of six and 14. For each year past the age six that an immigrant from a non-English-speaking country arrives in the United States their English speaking ability (on a scale of 0-3 units) decreases by 0.058 (SE = 0.001) of a unit. The associated *F*-statistic is 4,100, which indicates that there is ample explanatory power in the first-stage regression (Bound et al. 1995; Staiger and Stock 1997)[11]. The partial R-squared on the instrument is 0.007. This estimate is a difference-in-difference effect that captures the difference in English-speaking ability between immigrants who arrived before and after age six from non-English-speaking countries that is over and above the difference between immigrants who arrived before and after age six from English-speaking countries.

The remaining three columns in the table replicate the first-stage equation using the non-English-speaking country of birth dummy interacted with other age-of-arrival cut-offs (before age 7, 8 and 9) to capture difference-in-difference effects between young and older arrivers at different age discontinuities. This additional analysis is undertaken because there is no clear age discontinuity after which the ability to learn a language markedly diminishes. The coefficients on these instruments are always statistically significant and grow somewhat more pronounced as the cut-off age for native-like language acquisition is increased, which is to be expected. The standard errors around the estimated coefficients also increase with the age-of-arrival cut-off, presumably reflecting the growing divergence of responses as children age. For example, the results in the last column imply that on average for each year past the age of nine that immigrants from non-English-speaking countries move to the United States there is a 0.097 (SE = 0.002) unit disadvantage in English-speaking ability in comparison with those who arrived at younger ages[12]. These results confirm the hypothesis that learning English becomes more difficult with age for children from non-English-speaking countries of origin,

[11] Staiger and Stock (1997) suggest that if the first-stage *F*-statistic is less than ten it would raise concerns that the instrument were weak.

[12] It is noteworthy that the coefficient of -0.097 on this instrument is virtually identical to the -0.104 coefficient reported by Bleakley and Chin (2010) in their first-stage results using the same instrument and the 2000 census.

leading to slower acquisition of language skills among those who immigrate later in childhood.

The OLS and 2SLS results for the effect of language skills on five labor market outcomes and four other indicators of assimilation are displayed in the first two columns of table 3, respectively[13]. The identifying instrument is max(0, age of arrival – 6) × non-English-speaking country. The results reveal that immigrant women with higher English-speaking ability have substantially better labor market outcomes across all indicators of performance. The results in the first row for the 2SLS estimates imply that an additional unit increase in English-speaking ability leads to a 12.2 percentage point increase in labor force participation. To put the magnitude of this effect in perspective, since 76 percent of immigrants from non-English-speaking countries participated in the labor force, this implies that a 1 unit exogenous improvement in language skill caused the labor force participation rate to rise by 16.1 percent. Similar effects are reported in row 2 when working for pay is the dependent variable.

Rows 3 and 4 describe the effects of language proficiency on the intensity of work measured by weeks per year and hours per week. The 2SLS estimated coefficients imply that one unit improvement in English-speaking ability results in a statistically significant increase of 7.5 weeks per year while time at work per week rises by 6.8 hours.

Row 5 presents the effect of language skill on women's earned income (wages, salaries, and self-employment income). As one would expect, the increases in hours and weeks of work associated with improvements in language ability results in higher incomes for women. The 2SLS results reveal that a 1 unit increase in English speaking ability raises annual earnings by $15,980[14]. This translates to 50.7% of the overall sample mean among immigrants from non-English-speaking countries. These estimates of the effects of language ability for women are somewhat larger than those found in the 1990 census by Bleakley and Chin (2004) for men

[13] In analysis not reported, I also estimated the language effects using IV-probit models for the dichotomous dependent variables (LFP, employment, home ownership, poverty, health insurance,) and because of the large number of women who reported zeroes for the continuous outcomes (weeks of work, hours of work and earned income), IV-Tobit models. In these models, I replaced the first-stage equations with an ordered probit model because the measure of English fluency is constructed as a categorical variable. These results are qualitatively similar to those reported here.

[14] These results and all other dollar amounts discussed in the text are expressed in 1999 dollars.

and women combined. They found that on average improving English speaking ability by 1 unit leads to an earnings advantage of 33%. As an additional point of comparison, my findings for the United States closely resemble the wage disadvantage for immigrant women with poor levels of proficiency in the dominant language of the Netherlands produced by Yao and van Ours (2015). They found that poor Dutch language skills significantly reduced the wages of female immigrants by approximately 48%.

The effects of language fluency on non-labor market outcomes are given next. Despite the fact that better English ability considerably raises earned income on average, the 2SLS estimates do not reveal any statistically significant effect on the probability of home ownership. However, the income effects associated with improved language skill are mirrored in the results on poverty rates. These results imply that a unit increase in English-speaking ability reduces the fraction of women living in poverty by 13.5 percentage points, which amounts to a reduction of 88.8 percent in the sample average poverty rate.

Row 8 of table 3 reports the coefficient estimates for the effects of language skill on health insurance coverage, which is of concern to some policymakers. The estimated coefficient reveals that women with better English fluency are more likely to have health insurance coverage. The 2SLS estimates suggest that each increment to English-speaking ability increases the probability of having health coverage by 16.0 percentage points, about 20.2% of the mean health insurance rate for individuals from non-English-speaking countries.

An important issue in the study of language skills is its effect on school completion. Investments in education are likely to be more productive if one can communicate in the dominant language in school. The final row of the table displays these results. As expected women with better English-language skills face a lower cost of education in the United States. According to the 2SLS results increasing English-speaking ability by 1 unit raises average completed schooling by 2.6 years. This result can be compared with that reported by Bleakley and Chin (2004) using the 1990 census for men and women combined who found that a one-unit increase in English speaking increases years of schooling by 3.9 years.

The last three columns of table 3 examine the sensitivity of these conclusions to alternative identifying instruments. The identifying instruments are max(0, age of arrival - 7) × non-English-speaking country in column (3), max(0, age of arrival - 8) × non-English-speaking

country in column (4), and max(0, age of arrival - 9) × non-English-speaking country in column (5). The 2SLS estimated coefficients using these alternative specifications of the instrument are virtually identical across the table and the conclusions are unchanged. In the remainder of the paper, I use the instrument, max(0, age of arrival - 6) × non-English-speaking country, as individuals who came to the United States by age six would have received all their schooling in the country which may greatly have enhanced their acquisition of the English language.

A comparison of OLS and 2SLS results show that the 2SLS estimates are always larger in magnitude[15]. Thus the OLS estimates appear to be downward biased, which is contrary to what endogenous choice of language acquisition would predict. It was not known *a priori* whether OLS estimates would be higher than 2SLS estimates due to ability bias or lower than 2SLS estimates due to classical measurement error in the language variable. The language measure used in the present paper is based on respondents' self-reported assessments of their own English-speaking ability, which is therefore likely to suffer substantially from measurement error. It is well known that 2SLS can correct for both attenuation bias arising from measurement error and upward bias arising from endogeneity. My results are in line with those of earlier studies which also find that downward bias induced by measurement error overcompensates the upward bias induced by unobserved heterogeneity (Dustmann and van Soest 2002; Dustmann and Fabbri 2003; Bleakley and Chin 2004).

## 5. Additional robustness checks

Table 4 restricts the analysis to certain groups of countries. The first two columns display the base model for reference. As English-speaking countries might share greater cultural and institutional similarities with the United States, immigrants from these countries perhaps find it easier to adjust irrespective of age of arrival. In particular Canadian immigrants may be a poor comparison group for the assimilation process of the average immigrant due to Canada's geographical proximity to the United States and its similar cultural features. To allay

[15] Hausman tests indicate that the 2SLS estimate and the OLS estimate differ at the 5% level of significance for the following outcomes: number of weeks worked last year (*p-value* is 0.034), usual hours worked per week last year (*p-value* is 0.008), income earned last year (*p-value* is 0.001), health insurance (*p-value* is 0.002), poverty (*p-value* is 0.006) and years of schooling (*p-value* is 0.000). For the other outcomes the *p-value* for the Hausman tests are 0.212 for labour force participation, 0.133 for employed last year, and 0.188 for home ownership.

concerns in this regard, the next two columns exclude immigrants from Canada. These immigrants accounts for 27% (4,623 observations) of individuals from English-speaking countries. I find that the results using this restricted sample are remarkably similar to those using the entire sample of countries. In columns 5 and 6, I also drop immigrants from other countries that seem have to have similar economic and cultural environments to the United States. In addition to Canada, these are England, Scotland, Wales, Ireland, Australia and New Zealand. The OLS and 2SLS results are essentially the same as those using the total sample. Thus my main findings are very robust to alternative samples that might make the immigrants from English-speaking countries better controls for the non-language age-of-arrival effects experienced by immigrants from non-English-speaking countries.

Finally, in the last two columns of the table, I drop immigrants from Mexico to explore whether the estimated effect of English proficiency is generated by Mexicans alone, or whether the effect is common to immigrants from other countries as well. Immigrants from Mexico account for 31% (38,995 observations) of immigrants from non-English-speaking countries. Although almost all my results hold up for the sample without Mexicans, one interesting difference arises with respect to earnings. Each one unit increment in English proficiency generates a rise in earnings of $25,866 (SE = 4,501) compared to the base result of $15,980 (SE = 2,386). Although the estimates are imprecise, immigrants from non-English-speaking countries other than Mexico appear to fare better in terms of the English-speaking wage premium. This may be due in part to the fact that Mexican are the most numerous group of immigrants in the United States and tend to geographically sort into large ethnic enclaves where they can live and work in a Spanish language environment.

*How comparable are treatment and control countries?* If there are differential rewards associated with the schooling obtained in a non-English speaking country as opposed to an English-speaking one, the exclusion restriction in the first-stage equation would be invalid. In this case the 2SLS estimate may reflect not only differential English-language skills but also foreign education effects. One way to approach this potential problem is to control explicitly for attributes of the country of birth that influence the quality of education. The country of birth data that I employ to proxy for school quality include: the 1980 levels of per capital GDP, per-pupil school expenditures at the primary level, and the teacher-pupil ratio at the

primary level provided by Lee and Barro (1997)[16]. Table 5 shows the estimation results after controlling for these school quality interactions one by one. As a reference point, the first two columns display the OLS and 2SLS coefficients for the English-language measure in the base specification. In the remaining columns of the table, the indicators of origin-country school quality are included as interactions with age at arrival to allow age-of-arrival effects to differ by school quality. It turns out that these differences in schooling characteristics do not play an important role in explaining the impact of English proficiency in my analysis; OLS and 2SLS estimates of the impact of English proficiency remain similar to those presented in the base specification.

## 6. Conclusion

In this paper I focus on the importance of language skills for economic assimilation among immigrant women in the United States. Estimating the effect of language proficiency on economic outcomes is complicated by the endogeneity of language acquisition. That is, factors that affect the decision to learn a new language could also affect baseline indicators of labor market performance as well as other indicators of economic integration. This paper addresses this problem by focusing on a sample of women from the ACS between 2010 and 2015 who came to the United States during childhood. In order to find exogenous variation in English fluency, following Bleakley and Chin (2004), I utilized an instrumental strategy based on the fact that age of arrival in the United States is a determinant of English-speaking ability among immigrants from non-English-speaking countries but not for those from English-speaking countries.

Using this IV strategy, the results of this study show that language proficiency is critical to a broad range of outcomes for immigrant women. Stronger language skills significantly raise levels of labor supply as measured by labor force participation, employment, usual working hours per week and the number of weeks worked per year. Additionally, women who acquire higher speaking ability receive much higher wages than their counterparts with lesser proficiency. Better English skills also significantly reduce the probability of falling below the poverty line and increase the probability of being covered by health insurance.

[16] These variables are from the data sets constructed and described by Lee and Barro (1997). I use the values for 1980 because immigrants in the 2010-2015 ACS would be exposed not to contemporaneous conditions in their country of birth but conditions prevailing when they were children.

The results of this study suggest that enabling immigrant women to better their language proficiency will substantially improve their economic assimilation into American society. At the same time, I caution the reader not to generalize from these findings to women who arrive in the United States as adults as different selection mechanisms may apply to children and adult immigrants. I have considered only a limited segment of the immigrant population who arrived in the United States as children. Further research is needed before anything can be said about the economic importance of language acquisition for adult immigrants. Nevertheless, the evidence presented here lends credence to the importance of language acquisition for the process of economic assimilation.

**References**

Angrist, J. and Lavy, V. 1997. The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco. Journal of Labor Economics 15(1), pp. 48-76.

Batalova, J. and Zong, J. 2016. Frequently Requested Statistics on Immigrants and Immigration in the United States. Migration Policy Institute Online Journal, April 2016.

Bleakley, H. And Chin, A. 2004. Language Skills and Earnings: Evidence from Childhood Immigrants. Review of Economics and Statistics 86(2), pp. 481-496.

Bleakley, H. and Chin, A. 2008. What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital among Immigrants. Journal of Human Resources 43(2), pp. 267-298.

Bleakley, H. and Chin, A. 2010. Age at Arrival, English Proficiency, and Social Assimilation Among U.S. Immigrants. American Economic Review: Applied Economics 2 (1), pp. 165-192.

Bound, J., Jaeger, D. and Baker, R.1995. Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variables is Weak. *Journal of the American Statistical Association* 90(430), pp. 443–50.

Budría, S. and Swedberg, P. 2015. The Impact of Language Proficiency on Immigrants' Earnings in Spain. Revista de Economia Aplicada 23(67), pp. 63-91.

Carliner, G. 1981. Wage Differences by Language Group and the Market for Language Skills in Canada. Journal of Human Resources 16(3), pp. 384–399.

Chiswick, B. 1991. Speaking, Reading, and Earnings among Low- Skilled Immigrants. Journal of Labor Economics 9(2), pp.149–170.

Chiswick, B. and Miller, P. 1995. The Endogeneity between Language and Earnings: International Analyses. Journal of Labor Economics 13(2), pp. 246–288.

Dávila, A. and Mora, M. 2001. The Marital Status of Recent Immigrants in the United States in 1980 and 1990. International Migration Review 35(2), pp. 506–524.

Di Paolo, A. and Raymond, J. 2012. Language knowledge and earnings in Catalonia. Journal of Applied Economics 15 (1), pp. 89-118.

Duncan, B. and Trejo, S. 2007. Ethnic Identification, Intermarriage and Unmeasured Progress by Mexican Americans. In: Borjas, George, editor. Mexican Immigration to the United States. National Bureau of Economic Research and the University of Chicago Press; Chicago, pp. 229-267.

Dustmann, C. 1994. Speaking Fluency, Writing Fluency and Earnings of Migrants. Journal of Population Economics 7(2), pp. 133-156.

Dustmann, C. and Fabbri, F. 2003. Language Proficiency and Labour Market Performance of Immigrants in the UK. Economic Journal 113 (489), pp. 695-717.

Dustmann, C. and van Soest, A. 2002. Language and the Earnings of Immigrants. Industrial and Labor Relations Review 55(3), pp. 473-492.

Friedberg, R. 2000. You Can't Take it with You? Immigrant Assimilation and the Portability of Human Capital. Journal of Labor Economics 18(2), pp. 221–251.

Funkhouser, E., and Ramos, F. 1993. The Choice of Migration Destination: Dominican and Cuban Immigrants to the Mainland United States and Puerto Rico. International Migration Review 27(3), pp. 537-556.

Grenier, G. 1984. The Effect of Language Characteristics on the Wages of Hispanic American Males. Journal of Human Resources 19(1), pp. 35-52.

Guven, C. and Islam, A. 2015. Age at Migration, Language Proficiency and Socio-economic Outcomes: Evidence from Australia. Demography 52(2), pp. 513-542.

Kossoudji, S. 1988. English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men. Journal of Labor Economics 6(2), pp. 205-228.

Lee, J. and Barro, R. 1997. Schooling Quality in a Cross Section of Countries. NBER Working Paper No. 6198. Cambridge, Mass.: National Bureau of Economic Research.

Lenneberg, E. 1967. Biological Foundation of Language. Wiley & Sons; New York.

McManus, W., Gould, W. and Welch, F. 1983. Earnings of Hispanic Men: The Role of English Language Proficiency. Journal of Labor Economics 1(2), pp. 101-130.

Meng, X. and Gregory, R. 2005. Intermarriage and the Economic Assimilation of Immigrants. Journal of Labor Economics 23(1), pp. 135-175.

Miranda, A. and Zhu, Y. 2013. English deficiency and the Native-Immigrant Wage Gap. Economics Letters 118(1), pp. 38-41.

Newport, E. 2002. Critical Periods in Language Development. In Encyclopedia of Cognitive Science, ed. Lynn Nadel. London: Macmillan Publishers and Nature Publishing Group.

Rivera-Batiz, F. 1990. English Language Proficiency and the Economic Progress of Immigrants. Economics Letters 34(3), pp. 295-300.

Ruiz-de-Velasco, J., Fix, M. and Clewell, B. 2000. Overlooked and Underserved: Immigrant Students in U.S. Secondary Schools. Washington, DC: Urban Institute.

Sorenson, A. 1988. The Fertility and Language Characteristics of Mexican-American and Non- Hispanic Husbands and Wives. Sociological Quarterly 29(1), pp.111-130.

Stevens, G. and Swicegood, G. 1987. The Linguistic Context of Ethnic Endogamy. American Sociological Review 52(1), pp. 73-82.

Staiger, D., and Stock, J. 1997. Instrumental Variables Regression with Weak Instruments. Econometrica 65(3), pp. 557–86.

Swicegood, G., Bean, F., Stephen, E., Opitz, W. 1988. Language Usage and Fertility in the Mexican-Origin Population of the United States. Demography 25(1), pp. 17–33.

Toussaint-Comeau, M., and Rhine, S. 2004. Tenure Choice with Location Selection: The Case of Hispanic Neighborhoods in Chicago. Contemporary Economic Policy 22(1), pp. 95-110.

Yao, Y., and van Ours, J. 2015. Language Skills and Labor Market Performance of Immigrants in the Netherlands. Labour Economics 34, pp. 76-85.

Table 1 Summary statistics

| | From Non-English Speaking Countries | | | From English Speaking Countries | | |
|---|---|---|---|---|---|---|
| | Overall | Arrived Aged 0-6 | Arrived Aged 7-14 | Overall | Arrived Aged 0-6 | Arrived Aged 7-14 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| English-speaking ability (scale of 0 to 3, 3 = best) | 2.692 | 2.857 | 2.525 | 2.983 | 2.984 | 2.983 |
| | (0.671) | (0.478) | (0.785) | (0.156) | (0.158) | (0.154) |
| Control variables: | | | | | | |
| Age of arrival in the U.S. | 6.674 | 2.594 | 10.790 | 6.305 | 2.567 | 10.601 |
| | (4.635) | (1.993) | (2.325) | (4.540) | (1.970) | (2.305) |
| Age | 38.485 | 38.852 | 38.115 | 40.607 | 41.229 | 39.892 |
| | (8.862) | (8.975) | (8.732) | (9.012) | (9.117) | (8.836) |
| White | 0.579 | 0.622 | 0.536 | 0.531 | 0.694 | 0.343 |
| | (0.494) | (0.485) | (0.499) | (0.499) | (0.461) | (0.475) |
| Black | 0.042 | 0.040 | 0.043 | 0.346 | 0.203 | 0.510 |
| | (0.200) | (0.196) | (0.203) | (0.476) | (0.402) | (0.500) |
| Asian and Pacific Islander | 0.172 | 0.151 | 0.193 | 0.057 | 0.048 | 0.067 |
| | (0.377) | (0.358) | (0.395) | (0.232) | (0.214) | (0.250) |
| Other non-white race | 0.170 | 0.142 | 0.200 | 0.026 | 0.019 | 0.034 |
| | (0.376) | (0.349) | (0.400) | (0.160) | (0.137) | (0.182) |
| Multiracial | 0.037 | 0.046 | 0.028 | 0.041 | 0.036 | 0.046 |
| | (0.188) | (0.209) | (0.165) | (0.197) | (0.186) | (0.210) |
| Hispanic | 0.534 | 0.459 | 0.609 | 0.020 | 0.022 | 0.017 |
| | (0.499) | (0.498) | (0.488) | (0.139) | (0.146) | (0.130) |
| Survey year | 2013 | 2013 | 2013 | 2012 | 2012 | 2013 |
| | (1.698) | (1.699) | (1.698) | (1.700) | (1.699) | (1.701) |

Table 1. Summary statistics (concluded)

| | From Non-English Speaking Countries | | | From English Speaking Countries | | |
|---|---|---|---|---|---|---|
| | Overall | Arrived Aged 0-6 | Arrived Aged 7-14 | Overall | Arrived Aged 0-6 | Arrived Aged 7-14 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent variables: | | | | | | |
| Labour force participation (current year) | 0.764 | 0.779 | 0.749 | 0.817 | 0.802 | 0.834 |
| | (0.425) | (0.415) | (0.433) | (0.387) | (0.398) | (0.372) |
| Employed last year | 0.761 | 0.778 | 0.744 | 0.813 | 0.806 | 0.821 |
| | (0.427) | (0.416) | (0.437) | (0.390) | (0.395) | (0.383) |
| Number of weeks worked last year | 35.038 | 35.910 | 34.159 | 37.557 | 37.074 | 38.112 |
| | (22.081) | (21.684) | (22.440) | (20.748) | (20.942) | (20.510) |
| Usual hours per week last year | 28.671 | 29.477 | 27.857 | 31.067 | 30.634 | 31.565 |
| | (18.770) | (18.611) | (18.894) | (18.273) | (18.531) | (17.961) |
| Income earned last year (1999 dollars) | 31,545 | 34,021 | 29,047 | 39,467 | 39,442 | 39,495 |
| | (43296) | (45481) | (40823) | (48088) | (50184) | (45563) |
| Home ownership | 0.627 | 0.646 | 0.609 | 0.676 | 0.703 | 0.645 |
| | (0.484) | (0.478) | (0.488) | (0.468) | (0.457) | (0.479) |
| Health insurance | 0.792 | 0.833 | 0.751 | 0.879 | 0.888 | 0.869 |
| | (0.406) | (0.373) | (0.432) | (0.326) | (0.315) | (0.337) |
| Poverty | 0.152 | 0.129 | 0.175 | 0.099 | 0.095 | 0.103 |
| | (0.359) | (0.335) | (0.380) | (0.299) | (0.294) | (0.304) |
| Years of schooling | 13.553 | 14.081 | 13.020 | 14.853 | 14.936 | 14.758 |
| | (3.332) | (2.895) | (3.643) | (2.487) | (2.462) | (2.513) |
| Number of observations | 125,386 | 62,975 | 62,411 | 17,116 | 9,152 | 7,964 |

Table 2. First-stage results for the total sample

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Identifying instrument: | | | | |
| max(0, age at arrival - 6) x (non- | -0.058*** | | | |
| English-speaking country of birth) | (0.001) | | | |
| max(0, age at arrival - 7) x (non- | | -0.067*** | | |
| English-speaking country of birth) | | (0.001) | | |
| max(0, age at arrival - 8) x (non- | | | -0.080*** | |
| English-speaking country of birth) | | | (0.001) | |
| max(0, age at arrival - 9) x (non- | | | | -0.097*** |
| English-speaking country of birth) | | | | (0.002) |
| Observations | 142,502 | 142,502 | 142,502 | 142,502 |

Notes: The dependent variable is English-speaking ability defined as: 0 = no English, 1 = not well, 2 = well, and 3 = very well. All regressions contain dummies for country of birth, age at arrival, age, race (white (default), black, Asian and Pacific Islanders, other, multiracial), Hispanic origin and survey year. Robust standard errors are shown in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence, double 95%, and triple 99%.

Table 3. The effect of English-speaking ability on labor market and other socioeconomic outcomes – base results

| | OLS | 2SLS | 2SLS | 2SLS | 2SLS |
|---|---|---|---|---|---|
| Labour force participation | 0.097*** | 0.122*** | 0.120*** | 0.120*** | 0.122*** |
| | (0.002) | (0.020) | (0.020) | (0.020) | (0.021) |
| Employed last year | 0.098*** | 0.130*** | 0.130*** | 0.129*** | 0.134*** |
| | (0.002) | (0.021) | (0.021) | (0.021) | (0.021) |
| No. of weeks worked last year | 5.119*** | 7.456*** | 7.484*** | 7.571*** | 7.893*** |
| | (0.106) | (1.107) | (1.100) | (1.103) | (1.121) |
| Usual hours per week last year | 4.192*** | 6.752*** | 6.684*** | 6.681*** | 6.837*** |
| | (0.088) | (0.975) | (0.968) | (0.971) | (0.987) |
| Income earned last year | 8194.248*** | 15980.352*** | 15116.715*** | 14545.812*** | 14601.822*** |
| | (119.140) | (2385.737) | (2348.159) | (2322.966) | (2317.394) |
| Home ownership | 0.066*** | 0.033 | 0.029 | 0.027 | 0.030 |
| | (0.002) | (0.025) | (0.025) | (0.025) | (0.025) |
| Health insurance | 0.096*** | 0.153*** | 0.146*** | 0.146*** | 0.150*** |
| | (0.002) | (0.018) | (0.018) | (0.018) | (0.019) |
| Poverty | -0.084*** | -0.129*** | -0.123*** | -0.124*** | -0.126*** |
| | (0.002) | (0.017) | (0.016) | (0.017) | (0.017) |
| Years of schooling | 1.746*** | 2.606*** | 2.569*** | 2.538*** | 2.524*** |
| | (0.017) | (0.137) | (0.137) | (0.138) | (0.141) |
| Observations | 142,502 | 142,502 | 142,502 | 142,502 | 142,502 |

Notes: Each cell reports the coefficient on English-language ability from a separate regression that contains dummies for country of birth, age at arrival, age, race, Hispanic origin and survey year. The "2SLS" columns are estimated using 2SLS with max(0, age at arrival - 6)×non-English-speaking country as the identifying instrument in column 2, with max(0, age at arrival - 7)×non-English-speaking country in column 3, max(0, age at arrival - 8)×non-English-speaking country in column 4, and max(0, age at arrival - 9)×non-English-speaking country in column 5. Robust standard errors are shown in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence, double 95%, and triple 99%.

Table 4. The effect of English-speaking ability on labor market and other socioeconomic outcomes – alternative countries in sample

| | Total Sample Base results | | Excluding Immigrants from Canada | | Excluding Immigrants from Countries Similar to the US | | Excluding Immigrants from Mexico | |
|---|---|---|---|---|---|---|---|---|
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Labour force participation | 0.097*** | 0.122*** | 0.097*** | 0.106*** | 0.097*** | 0.100*** | 0.085*** | 0.113** |
| | (0.002) | (0.020) | (0.002) | (0.022) | (0.002) | (0.024) | (0.003) | (0.038) |
| Employed last year | 0.098*** | 0.130*** | 0.098*** | 0.113*** | 0.099*** | 0.117*** | 0.091*** | 0.137*** |
| | (0.002) | (0.021) | (0.002) | (0.023) | (0.002) | (0.026) | (0.003) | (0.038) |
| No. of weeks worked last year | 5.119*** | 7.456*** | 5.115*** | 6.815*** | 5.119*** | 6.582*** | 4.559*** | 7.670*** |
| | (0.106) | (1.107) | (0.106) | (1.208) | (0.106) | (1.385) | (0.165) | (2.044) |
| Usual hours per week last year | 4.192*** | 6.752*** | 4.186*** | 5.942*** | 4.191*** | 5.788*** | 4.110*** | 8.135*** |
| | (0.088) | (0.975) | (0.088) | (1.062) | (0.088) | (1.203) | (0.138) | (1.803) |
| Income earned last year | 8194.248*** | 15980.352*** | 8193.417*** | 15215.711*** | 8166.949*** | 10381.742*** | 10875.780*** | 25865.725*** |
| | (119.140) | (2385.737) | (118.453) | (2524.562) | (118.506) | (2762.786) | (235.417) | (4500.625) |
| Home ownership | 0.066*** | 0.033 | 0.066*** | 0.016 | 0.065*** | -0.042 | 0.072*** | -0.007 |
| | (0.002) | (0.025) | (0.002) | (0.027) | (0.002) | (0.033) | (0.003) | (0.045) |
| Health insurance | 0.096*** | 0.153*** | 0.096*** | 0.125*** | 0.096*** | 0.115*** | 0.078*** | 0.115*** |
| | (0.002) | (0.018) | (0.002) | (0.020) | (0.002) | (0.024) | (0.003) | (0.034) |
| Poverty | -0.084*** | -0.129*** | -0.084*** | -0.108*** | -0.084*** | -0.078*** | -0.077*** | -0.139*** |
| | (0.002) | (0.017) | (0.002) | (0.019) | (0.002) | (0.022) | (0.003) | (0.030) |
| Years of schooling | 1.746*** | 2.606*** | 1.744*** | 2.425*** | 1.743*** | 2.059*** | 1.634*** | 2.548*** |
| | (0.017) | (0.137) | (0.017) | (0.148) | (0.017) | (0.173) | (0.027) | (0.252) |
| Observations | 142,502 | 142,502 | 137,879 | 137,879 | 132,781 | 132,781 | 103,507 | 103,507 |

See notes for table 3. The 2SLS columns are estimated using 2SLS with max(0, age at arrival - 6)×non-English-speaking country as the identifying instrument. In columns (5) and (6) the excluded countries are Canada, England, Scotland, Wales, Northern Ireland, Republic of Ireland, Australia, and New Zealand.

Table 5. The effect of English-speaking ability on labor market and other socioeconomic outcomes – school quality controls

| | Total Sample Base results | | max(0, age at arrival - 6) x ln(per capita PPP GDP) | | max(0, age at arrival - 6) x ln(teacher-pupil ratio) | | max(0, age at arrival - 6) x ln(school exp. per child) | |
|---|---|---|---|---|---|---|---|---|
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Labour force participation | 0.097*** | 0.122*** | 0.097*** | 0.130*** | 0.097*** | 0.135*** | 0.097*** | 0.140*** |
| | (0.002) | (0.020) | (0.002) | (0.018) | (0.002) | (0.030) | (0.002) | (0.029) |
| Employed last year | 0.098*** | 0.130*** | 0.098*** | 0.130*** | 0.098*** | 0.141*** | 0.098*** | 0.152*** |
| | (0.002) | (0.021) | (0.002) | (0.019) | (0.002) | (0.031) | (0.002) | (0.030) |
| No. of weeks worked last year | 5.119*** | 7.456*** | 5.107*** | 7.935*** | 5.098*** | 8.412*** | 4.157*** | 7.060*** |
| | (0.106) | (1.107) | (0.115) | (0.997) | (0.116) | (1.639) | (0.095) | (1.382) |
| Usual hours per week last year | 4.192*** | 6.752*** | 4.164*** | 6.473*** | 4.149*** | 7.001*** | 4.157*** | 7.060*** |
| | (0.088) | (0.975) | (0.095) | (0.874) | (0.095) | (1.440) | (0.095) | (1.382) |
| Income earned last year | 8194.248*** | 15980.352*** | 7583.023*** | 13245.142*** | 7529.744*** | 16597.659*** | 7526.405*** | 14534.261*** |
| | (119.140) | (2385.737) | (120.019) | (2100.603) | (119.784) | (3539.398) | (119.188) | (3406.094) |
| Home ownership | 0.066*** | 0.033 | 0.064*** | 0.065** | 0.064*** | 0.042 | 0.064*** | 0.062 |
| | (0.002) | (0.025) | (0.002) | (0.022) | (0.002) | (0.036) | (0.002) | (0.035) |
| Health insurance | 0.096*** | 0.153*** | 0.104*** | 0.171*** | 0.103*** | 0.145*** | 0.104*** | 0.171*** |
| | (0.002) | (0.018) | (0.002) | (0.017) | (0.002) | (0.027) | (0.002) | (0.026) |
| Poverty | -0.084*** | -0.129*** | -0.086*** | -0.128*** | -0.085*** | -0.123*** | -0.085*** | -0.111*** |
| | (0.002) | (0.017) | (0.002) | (0.015) | (0.002) | (0.024) | (0.002) | (0.023) |
| Years of schooling | 1.746*** | 2.606*** | 1.757*** | 2.628*** | 1.751*** | 2.790*** | 1.751*** | 2.457*** |
| | (0.017) | (0.137) | (0.018) | (0.125) | (0.018) | (0.207) | (0.018) | (0.200) |
| Observations | 142,502 | 142,502 | 112,226 | 112,226 | 112,226 | 112,226 | 112,226 | 112,226 |

See notes for table 3. The 2SLS columns are estimated using 2SLS with max(0, age at arrival - 6)×non-English-speaking country as the identifying instrument. GDP per capital, teacher-pupil ratio, and school expenditure per child are those for the country of birth and are from 1980 data provided by Lee and Barro (1997).

Figure 1A. English speaking ability at age of arrival – regression adjusted means

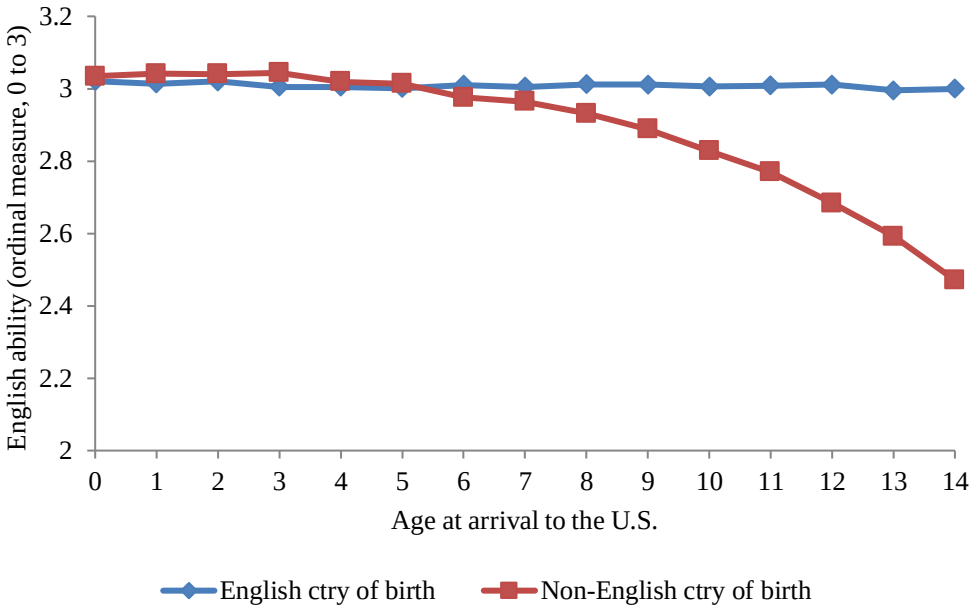
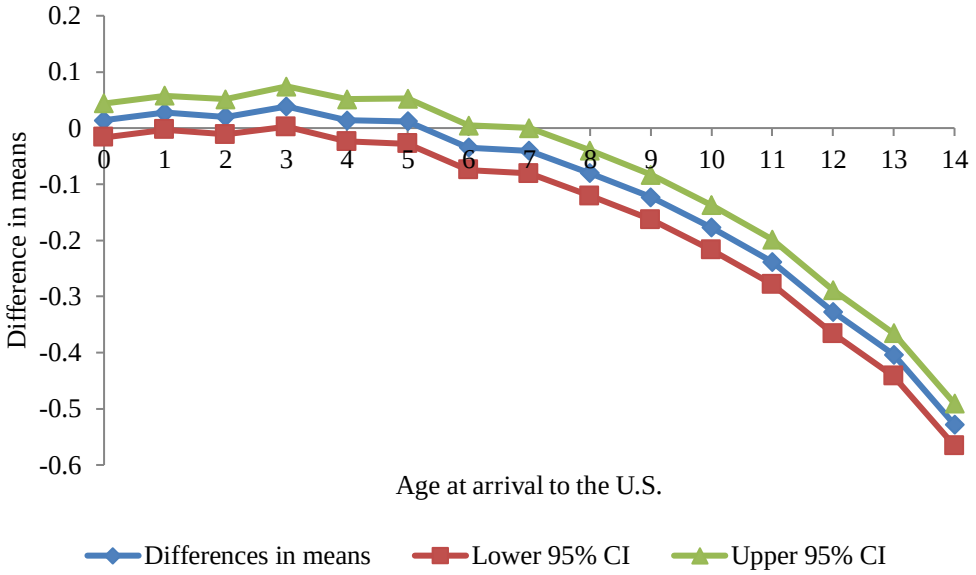



Figure 1B. English speaking ability at age of arrival – difference in means




Notes: Data from 2010-2015 IPUMS. Sample size is 142,502 (composed of all female individuals who arrived in the United States by age 14 between 1955 and 2003 and are currently age 25 to 55). Means have been regression adjusted for age, race, and Hispanic dummies.

# Appendix

Table A1. Immigrants by country of birth: English speaking countries

| Rank by N | Country of birth | *N* | Share of total *N* (%) |
|---|---|---|---|
| 1 | Canada | 4,623 | 27.01 |
| 2 | Jamaica | 3,030 | 17.7 |
| 3 | England | 2,829 | 16.53 |
| 4 | United Kingdom, ns | 1182 | 6.91 |
| 5 | Guyana/British Guiana | 1136 | 6.64 |
| 6 | Trinidad and Tobago | 1054 | 6.16 |
| 7 | Scotland | 356 | 2.08 |
| 8 | South Africa (Union of) | 347 | 2.03 |
| 9 | Australia | 342 | 2 |
| 10 | U.S. Virgin Islands | 314 | 1.83 |
| 11 | Barbados | 276 | 1.61 |
| 12 | Bahamas | 246 | 1.44 |
| 13 | Ireland | 245 | 1.43 |
| 14 | Belize/British Honduras | 211 | 1.23 |
| 15 | Liberia | 175 | 1.02 |
| 16 | Bermuda | 158 | 0.92 |
| 17 | New Zealand | 121 | 0.71 |
| 18 | Grenada | 103 | 0.6 |
| 19 | St. Vincent | 94 | 0.55 |
| 20 | Antigua-Barbuda | 92 | 0.54 |
| 21 | St. Lucia | 84 | 0.49 |
| 22 | Zimbabwe | 53 | 0.31 |
| 23 | Northern Ireland | 23 | 0.13 |
| 24 | St. Kitts-Nevis | 22 | 0.13 |
| | Total | 17,116 | 100 |

Table A1. Immigrants by country of birth: Non-English speaking countries (concluded)

| Rank | Country of birth | *N* | % | Rank | Country of birth | *N* | % |
|---|---|---|---|---|---|---|---|
| 1 | Mexico | 38,995 | 31.1 | 33 | Israel/Palestine | 647 | 0.5 |
| 2 | Germany | 11,001 | 8.8 | 34 | Greece | 585 | 0.5 |
| 3 | Puerto Rico | 6,615 | 5.3 | 35 | Argentina | 581 | 0.5 |
| 4 | Vietnam | 6,222 | 5.0 | 36 | Romania | 464 | 0.4 |
| 5 | Korea | 5,959 | 4.8 | 37 | Lebanon | 433 | 0.4 |
| 6 | Cuba | 3,983 | 3.2 | 38 | Turkey | 397 | 0.3 |
| 7 | El Salvador | 3,409 | 2.7 | 39 | Armenia | 389 | 0.3 |
| 8 | Japan | 3,115 | 2.5 | 40 | Netherlands | 375 | 0.3 |
| 9 | Dominican Republic | 3,062 | 2.4 | 41 | Chile | 365 | 0.3 |
| 10 | China | 2,513 | 2.0 | 42 | Costa Rica | 346 | 0.3 |
| 11 | Colombia | 2,024 | 1.6 | 43 | Iraq | 322 | 0.3 |
| 12 | Taiwan | 1,960 | 1.6 | 44 | Africa, ns/nec | 327 | 0.3 |
| 13 | Italy | 1,850 | 1.5 | 45 | Egypt/United Arab Rep. | 310 | 0.3 |
| 14 | Guatemala | 1,845 | 1.5 | 46 | Afghanistan | 277 | 0.2 |
| 15 | Haiti | 1,652 | 1.3 | 47 | Bangladesh | 275 | 0.2 |
| 16 | Thailand | 1,606 | 1.3 | 48 | Bosnia | 260 | 0.2 |
| 17 | Laos | 1,407 | 1.1 | 49 | Ethiopia | 224 | 0.2 |
| 18 | Poland | 1,319 | 1.1 | 50 | Bolivia | 221 | 0.2 |
| 19 | Nicaragua | 1,277 | 1.0 | 51 | Azores | 221 | 0.2 |
| 20 | Portugal | 1,232 | 1.0 | 52 | Asia, nec/ns | 209 | 0.2 |
| 21 | France | 1,170 | 0.9 | 53 | Belgium | 214 | 0.2 |
| 22 | Ecuador | 1,077 | 0.9 | 54 | Indonesia | 214 | 0.2 |
| 23 | Honduras | 1,084 | 0.9 | 55 | Switzerland | 182 | 0.2 |
| 24 | Peru | 1,072 | 0.9 | 56 | Yugoslavia | 176 | 0.1 |
| 25 | Iran | 1,047 | 0.8 | 57 | Saudi Arabia | 177 | 0.1 |
| 26 | Cambodia (Kampuchea) | 1,009 | 0.8 | 58 | Burma (Myanmar) | 161 | 0.1 |
| 27 | Panama | 929 | 0.7 | 59 | Morocco | 163 | 0.1 |
| 28 | Other USSR/Russia | 903 | 0.7 | 60 | Cape Verde | 165 | 0.1 |
| 29 | Ukraine | 900 | 0.7 | | Subtotal, top 60 countries | 121,222 | 96.7 |
| 30 | Spain | 853 | 0.7 | | Subtotal, other (58) ctries | 4,164 | 3.3 |
| 31 | Brazil | 801 | 0.6 | | Total non-Eng. spking obs. | 125,386 | 100.0 |
| 32 | Venezuela | 651 | 0.5 | | As % of total observations | | 89.3 |

Notes: Information on each country's official languages was taken from the CIA World Factbook. Recent adult immigrants from the 1980 Census were used to separate English-official countries into English-speaking countries (i.e., where more than half the recent adult immigrants did not speak a language at home other than English) and other countries. These countries, which are excluded from the main analysis, are American Samoa, Samoa, Guam, Dominica, Hong Kong, Philippines, Singapore, India, Pakistan, Gambia, Ghana, Nigeria, Senegal, Sierra Leone, Kenya, Tanzania, Uganda, Zambia, Fiji, Tonga, Marshall Islands, and Micronesia. The above tabulations by country of birth use data from the 2010 to 2015 ACS. Sample size is 142,502 composed of all female individuals who arrived in the US by age 14 between 1955 and 2003 and are between 25 and 55 at the time of the survey. Country refers to IPUMS detailed birthplace code.