Age at Arrival, English Proficiency and Social Assimilation among US Immigrants: A Replication

Microeconometrics

Group 8:

Markus Edvard Relbo (59537)

Thiago Miguel Gomes (43217)

Giani Ranzetti (58575)

Referee Report:

This is a replication of "Age at Arrival, English Proficiency and Social Assimilation among US Immigrants" by Bleakley and Chin (2010). The paper focuses on the role of English language skills in immigrant integration and social outcomes: quantifying the impact of English proficiency on marriage, fertility, and residential location choices. The data sample is divided into individuals from non-English speaking countries and countries that have English as an official language and predominantly speak it. The first category is the 'treatment group' and the second is the 'control group'. Since the English proficiency is an endogenous variable, the paper instrumentalizes the immigrant's age at arrival by assuming that there is no difference in English level between the treatment and control before the age-at-arrival of 9. Any differences between groups regarding socioeconomic outcomes for young vs old arrivers, is attributed to the English proficiency.

An interesting feature about the paper is that it seems to be the first that takes the endogeneity of the English proficiency on marriage, fertility, and residential outcomes into consideration. Another contribution is that it estimates all these socioeconomic outcome variables by using the same dataset and the same methods. This provides a homogenous overview on how English proficiency influences these socioeconomic factors, compared to other studies, which used different datasets and methods for different outcome variables. Additionally, the paper assesses the robustness of its findings by changing the model specifications and the samples. For example, it omits Canadian and Mexican immigrants and still finds similar English proficiency effects.

A weakness of the paper is the ability to measure whether people live in ethnic enclaves or not. The narrowest geographic aggregation is public-use microdata area (PUMA) where each PUMA contains at least 100,000 people. It is therefore hard to assess if people live in ethnic enclaves, since these areas are so big that they will contain people from multiple different backgrounds. To adjust for this weakness, the paper transforms the data into the percentage of people with the same country of birth and ancestry within the PUMA. However, the paper only finds weak evidence of English proficiency negatively affecting the likelihood of living in ethnic enclaves. The effects are only statistically significant when Mexican immigrants are omitted from the sample. Another weakness is that the spousal characteristics of interest can only be measured for those who live together. Therefore, it is not possible to measure the characteristics of former spouses of divorced individuals. This can lead to sample selection bias if being divorced is related to any of the regressors. The paper argues that the possible bias is not big enough in magnitude to overturn the estimated effects of English proficiency on socioeconomic outcomes. Observing the omitted variables would probably only strengthen the trend that is already discovered. That is because immigrants with high English proficiency are more likely to marry someone from other countries and especially likely to marry Americans. Since Americans on average are more educated and earn more than immigrants, observing those former spouses would probably only increase the hypotheses about how immigrants' level of English influences the characteristics of their spouses.

Also, one of the foundational arguments of the paper is that younger children learn English better during the "critical learning period" before the age of 9. However academic literature has found evidence against this theory like (Newport, 2002 & Birdsong, 2006). Does the ability to acquire new languages decline sharply or gradually? What is the driving mechanism behind the negative relationship between age and language proficiency?

The data is primarily derived from the 2000 US Census of Population and Housing. Specifically, the authors utilize the 1 percent and 5 percent samples from the Integrated Public Use Microsample Series (IPUMS). These datasets are extensive and provide a robust framework for analyzing the English proficiency and various social integration measures among immigrants in the United States. This survey approach provides a weighting to each observation to recover the true population demographics. The weight directly indicates how many persons in the U.S. population are represented by a given person in an IPUMS sample.

Given the endogeneity of English proficiency, OLS regressions of social outcomes on English proficiency will most likely not estimate the causal effect. Hence, a 2SLS (IV) regression design is implemented. To ensure the validity of causal inference, the following assumptions were made: Relevance of the instrumental variable (The IV must be strongly correlated with the endogenous regressor, which in this case is English proficiency); Exogeneity of the IV (The IV should not be correlated with the error term in the outcome equation) and the Exclusion Restriction (The IV should not directly affect the dependent variable except through its effect on the endogenous regressor - English proficiency).

The researchers propose a few extensions that could be implemented in order to explain English proficiency's effect on the reduction in fertility. Namely if increasing the pool of suitors, adopting US cultural norms and raising female bargaining power has a causal effect. Instead, as an extension we will attempt to implement an IV- probit model on certain binary social outcomes.

Main Replication Report:

The original paper focused on 4 different facets of immigrant integration in the US. However, given that the same econometric methodology is applied across each aspect and there are constraints with regards to report length, only marriage and residential outcomes will be investigated. The table in *Appendix 1* reports the means of each variable for the treatment (Non-English speaking country of origin) and control group (English speaking country of origin) with decompositions by age categories.

To accurately characterize 'ethnic enclaves', the researchers created four measures. The first captures the proportion of residents from one's PUMA that shares the same country of birth. Recognizing that an enclave is where there is a high concentration of people from one's own group. The second variable represents whether for someone from your country, you tend to live in a neighborhood with an above-average number same country of birth individuals. A second set of measures based on ancestry accounts for US-born people with the same ethnic background.

The ordinal English speaking ability variable is divided into 4 arbitrary English mastery levels, with 3 being the best level and 0 meaning no English is spoken at all. In *Figure 1 Panel A*, we plot the average English level for each year of arrival for treatment and control groups. The means for each age at arrival is weighted by survey weight, and controlled for age, race and sex. In accordance with the original paper, individuals who arrived before the age of 9 from non-English speaking countries speak as well as their native counter parts if not better, on average. We further investigate the relationship by plotting the difference in means in Figure 1's Panel B.

Panel A: Regression adjusted means *Panel B: Difference in means*

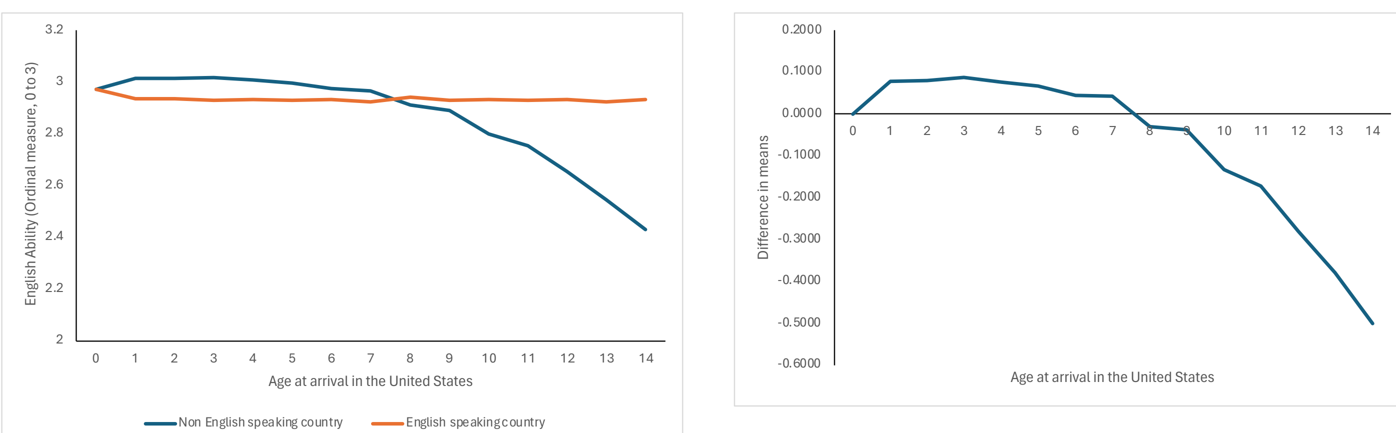


Figure 1: English speaking ability by age at arrival and differences in means

The shape of the curve of the difference in means is the same as the treatment group curve. This indicates that immigrants from English speaking countries have a consistent English ability regardless of their age at arrival. However, for immigrants from non-English countries, there is a clear difference in English proficiency depending on age of arrival. These findings are match with the critical period of language acquisition assumption. Nevertheless, from the plots, it could be argued that the critical period already ends at age 8 since the difference in means already becomes negative at that age. However, using this assumption negligibly influences the results of the paper and being more generous with the cutoff may favour a more realistic gradual decline in language absorption.

Considering this finding, we impose a restriction that there is no difference in English proficiency between childhood immigrants from non-English speaking countries and English-speaking countries that come before the age of 9. Although the relationship for immigrants arriving after the age of 9 looks like a decreasing exponential function, such an assumption would imply that past a certain age it would be impossible to learn English. Hence, we will follow the linear restriction set by the original paper. This should theoretically capture the relationship in Figure 1. It can be parameterized as follows:

(1) $$k_{ija} = \max(0, a - 9) \times 1[j\ is\ a\ non\ English\ speaking\ country],$$

where a is the age of arrival, $1(\cdot)$ is an indicator function and j is country of birth.
The relationship between English and age at arrival is estimated using the following equation:

(2) $$ENG_{ija} = \alpha_1 + \pi_1 k_{ija} + \delta_{1a} + \gamma_{1j} + w'_{ija}\rho_1 + \varepsilon_{1ija},$$

for individual i born in country j arriving in the United States at age a, ENG_{ija} is a measure of English language skills, δ_{1a} is a set of age-at-arrival dummies, γ_{1j} is a set of country of birth dummies, and w'_{ija} is a vector of exogenous explanatory variables (e.g. sex, age, race).

The estimated coefficients π_1 from equation (2) can be found in Table 2. Due to there being only exogenous variables in the model, parameters can be consistently estimated using OLS. We find that for each additional years after the age of 9, the probability of speaking any English at all decreases by 0.6%. The probability of speaking English "well" decreases by 3% and the probability of speaking English "very well" decreases by 7%. All these estimates are statistically significant to some conventional significance level. Finally, the ordinal measure of English ability decreases by 10% of a unit for each year past the age of arrival of 9. This effect represents about one sixth of a standard deviation in the ordinal measure for immigrants from non-English speaking countries with an age at arrival greater than 9. Intuitively, someone arriving at age 14 would see their English ordinal measure lowered by half an ordinal unit compared to pre-9-year-old immigrants.

	Coefficient on indentifying instrument	t-statistic
Panel A: English Profficiency		
Speaks English not well or better	-0.00610*	-1.98
Speaks English well or better	-0.0293*	-2.37
Speaks English very well	-0.0689***	-5.12
English-speaking ability ordinal measure	-0.104***	-3.62
Panel B: Marital Status		
Is currently married with spouse present	0.0112**	2.8
Is currently divorced	-0.00545**	-3
Has ever married	0.00745**	2.92
Number of children living in the same household	0.0460**	3.23
Has children living in the same household	0.00762	1.97
Is a single parent	-0.00219	-0.81
Is a never-married single parent	0.000251	0.14
Panel C: Residencial Location		
Fraction of PUMA population from same country of birth	0.000747	1.02
Fraction from same country of birth is above the mean for the country of birth	0.00159	0.23
Fraction of PUMA population with same primary ancestry	0.00185	1.4
Fraction with same country ancestry is above the mean for the national mean for the primary ancestry	0.00228	0.39

* p<0.05, ** p<0.01, *** p<0.001"

Table 2: Reduced-Form Effects on the coefficient of identifying instrument

As Table 2 suggests, age-at-arrival effects are significant drivers of social outcomes for immigrants of non-English speaking countries. In *Appendix 2 panels A-D*, we visualize some of these effects for: currently being married, the number of children living in the same household, currently being divorced and living in ethnic enclaves. The plots in Appendix 2 and 3 in this replication only visualizes the sample means for each age arrived which give slightly different plot than the original paper. In the actual report, a regression was run to retrieve the expected value of the social outcome for each age arrived and group. This allows proper inference for which age arrived causally affect socioeconomic outcomes. Given the extensive data set, it was difficult to implement such a large number of regressions.

For each of these pre-selected social outcome measures, the outcomes diverge between the treatment and control groups for later arriving immigrants. The curve for the control group is often much flatter than the treatment group (except *Panel A*). Therefore, the difference in means between the two groups always approximates to the curve from *Figure 1, Panel B*. The differences for immigrants with age-at-arrival less than 9 are not statistically different from 0 and after the threshold, age effects become approximately linear. This can be seen in more detail in *Appendix 3*.

These empirical findings give argument for the following causal mechanisms: childhood immigrants with exposure to English after the critical period attain less English proficiency as adults, which in turn influences their social outcomes. As outlined by the referee report, English proficiency is a highly endogenous variable. In fact, it may suffer from reverse causality. Given the endogeneity of English proficiency, an instrumental variable design is in order, considering the following regression:

$$y_{ija} = \alpha_1 + \beta ENG_{ija} + \delta_a + \gamma_j + w'_{ija}\rho + \varepsilon_{ija}, \tag{3}$$

for individual i born in country j arriving in the United States at age a. y_{ija} is the social outcome, ENG_{ija} is a measure of English language skills (endogenous variable), δ_a is a set of age-at-arrival dummies, γ_j is a set of country-of-birth dummies, and w'_{ija} is a vector of exogenous variables.

Since English skills are endogenous to the model, unbiased estimates cannot be recovered using OLS. Instead, k_{ija}, the excess age at arrival effect for the treatment group, is used as the instrumental variable to identify the effect of English proficiency (β). The first stage regression for all IV models is equation (2). Since there are as many endogenous variables as instruments, the model is just-identified. Ergo, dividing the reduced form coefficient by the first stage coefficient should provide the two stage least squares (2SLS) estimate. This means that, π_1 from equation (2) is rescaled to have a direct interpretation as an effect of English proficiency. From Table 2, Panel A, Regression 4, we find that the first stage coefficient is -0.104. This indicates that the 2SLS estimates will be opposite in sign and almost ten times larger in magnitude compared to reduced form effects from *Table 2, Panel B & C*.

	All childhood immigrants				Women				Men			
	OLS	t-statistic	IV	t-statistic	OLS	t-statistic	IV	t-statistic	OLS	t-statistic	IV	t-statistic
Panel A: Marital Status												
Is currently married with spouse present	0.00832	0.88	-0.108**	-2.72	-0.00409	-0.47	-0.0760*	-2.38	0.0192	1.73	-0.141**	-2.63
Is currently divorced	0.0100***	3.54	0.0524**	3.28	0.0146***	5.76	0.0643*	2.37	0.00582	1.65	0.0379	1.72
Has ever married	-0.00165	-0.25	-0.0717*	-2.56	-0.000575	-0.11	-0.0152	-0.55	-0.00363	-0.42	-0.134**	-2.69
* p<0.05, ** p<0.01, *** p<0.001"												

Table 3: Effect of English Language Skills on Marriage Outcomes

In *Table 3* we display the OLS and 2SLS estimates of the effect of English level on marital status divided between men and women. The 2SLS estimates suggest that English proficiency significantly decreases the probability of being currently married with a spouse present. This is attributable to English proficiency significantly increasing the probability of divorce for women and decreasing the probability of every marrying for men. As the original paper correctly points out, we only observe participants' marital status at a single point in time. So, it is possible that the large negative effect on having ever married estimated for men may be due to postponing marriage rather than foregoing it. English proficiency most likely raises the likelihood of divorcing because it increases opportunities to such a large extent that better English speaking immigrants divorce at lower levels of marital discord. Perhaps that greater English proficiency causes individuals to have higher expectations of their spouse's outcomes and leads to divorce. One final channel that might affect this relationship is that better English-speaking immigrants may be more accustomed or accepting to American's liberal attitude towards divorce.

	All childhood immigrants				Women				Men			
	OLS	t-statistic	IV	t-statistic	OLS	t-statistic	IV	t-statistic	OLS	t-statistic	IV	t-statistic
Panel B: Residencial Location												
Fraction of PUMA population from same country of birth	-0.00997***	-7.17	-0.00710	-0.84	-0.0117***	-6.33	-0.0136*	-2.2	-0.00822***	-8.39	-0.000223	-0.02
Fraction from same country of birth is above the mean for the country of birth	-0.0497***	-7.03	-0.0152	-0.22	-0.0581***	-10.52	-0.0376	-0.69	-0.0410***	-4.65	0.00654	0.07
Fraction of PUMA population with same primary ancestry	-0.0131***	-7.77	-0.0177	-1.04	-0.0168***	-6.91	-0.0327*	-2.52	-0.00938***	-8.82	-0.00172	-0.08
Fraction with same country ancestry is above the mean for the national mean for the primary ancestry	-0.0413***	-6.09	-0.0218	-0.35	-0.0496***	-7.58	-0.0880	-1.43	-0.0333***	-4.37	0.0557	0.78
* p<0.05, ** p<0.01, *** p<0.001"												

Table 4: Effect of English Language Skills on Residencial Location

Table 4 examines the effects of English language on residential outcomes, in particular the extent to which a participant lives in an ethnic enclave. We only find statistical significance for the effect of English proficiency on the share of PUMA population from the same country of birth and that share ancestry for women. In other words, English proficiency significantly decreases the odds of living in an enclave for women. All the OLS estimates that were highly significant due to endogeneity are now mostly all insignificant.

The interpretation of the regression coefficient relies on several assumptions. If age-at arrival effects, not related to language, differ between the treatment and control group, the identification of effects of English proficiency is invalid. An example of this could be different integration processes for English speaking countries and non-speaking countries. As a robustness check, we should consider different additional samples and specifications.

We first control for the possibility that immigrants from poorer countries face more barriers to adaptation, increasing with age at arrival. This is a plausible hypothesis given that on average non-English speaking countries are poorer than English speaking ones. It might affect outcomes through channels like education. In that case, the 2SLS estimate would express the marginal effect of English skills but also the differential returns to origin-country schooling. To test this, an interaction between country GDP per capita in 1980 and age at arrival is added to the model. In Table 5, we find similar results to the base case with this modified specification.

The second model specification adjustment was to control for the fertility rates of the origin counties. Assimilation to US norms means that immigrants need to adjust their fertility, generally, developing countries have higher fertility rates than industrialized countries. English-speaking countries are more likely to be industrialized. An interaction between age at arrival and total fertility rate of the country of birth was implemented to control for this. The estimated coefficients are like those of the base case. The effect of English on having ever married becomes insignificant.

	Base Results	t-statistic	Control for origin GDP and country of arrival	t-statistic	Control for origin fertility and country of origin	t-statistic	Drop Canada	t-statistic	Drop Mexico	t-statistic
Panel A: Marital Status										
Is currently married with spouse present	-0.108**	-2.72	-0.0822*	-2.55	-0.0972*	-2.03	-0.0950*	-2.09	-0.141	-1.66
Is currently divorced	0.0524**	3.28	0.0513***	3.84	0.0536**	2.72	0.0515**	2.62	0.0789*	2.24
Has ever married	-0.0717*	-2.56	-0.0432*	-1.98	-0.0568	-1.76	-0.0542	-1.97	-0.0940	-1.68
Panel B: Residencial Location										
Fraction of PUMA population from same country of birth	-0.00710	-0.84	0.000316	0.11	-0.0137	1.11	-0.000123	-0.03	-0.0140**	-2.71
Fraction from same country of birth is above the mean for the country of birth	-0.0152	-0.22	0.0223	0.48	-0.0351	-0.52	0.0316	0.64	-0.193*	-2.21
Fraction of PUMA population with same primary ancestry	-0.0177	-1.04	-0.00796	-0.65	-0.0319	-1.6	-0.00951	-0.7	-0.0576***	-4.23
Fraction with same country ancestry is above the mean for the national mean for the primary ancestry	-0.0218	-0.35	0.0305	0.85	-0.0318	-0.51	0.0227	0.51	-0.156*	-2.41
* p<0.05, ** p<0.01, *** p<0.001"										

Table 5: 2SLS Effect of English Language Using Alternative Samples and Specifications

The next hypothesis is that immigrants from English-speaking countries might be more culturally and institutionally similar to the US. This would make adjustment easier for the control group, regardless of age at arrival. Namely, Canada accounts for one third of the English-speaking country sample, yet it might be a poor control. Because of its geographical proximity and status as a British colony, Canadian immigrants might not experience the same barriers to integration. Nevertheless, when dropping Canadian observations, the results resemble the base case.

The final sample adjustment is dropping Mexican immigrants. They account for 29 percent of observations in the treatment group. This sample change helps determine if the effect of English proficiency is driven by Mexican immigrants alone or if it is common to other groups. The results show a few differences to the base case. First, the residential outcomes are now all significant and negative, living in an ethnic enclave is more likely for lower English levels. This finding could be explained by Mexicans being by far the largest group of immigrants and possess deep ethnic community ties. Due to cultural and marriage preferences, Mexican enclaves may even be attractive for English-proficient Mexicans to live in. The original paper also argues that Mexicans have a stronger preference to marry other Mexicans and live with each other regardless of English proficiency. Estimated coefficients on marital status are similar to the base case estimates. Only English effects on being divorced stays significant and it even increases in magnitude.

Table 6 suggests that the findings of the paper are robust to changes in sample and specifications. It is relatively unaffected by nonlanguage at-arrival effects that may differ between immigrants from English speaking and non-English speaking countries. It can be argued that the sample is not being representative of population demographics when dropping many observations.

This concludes the replication of the original paper by Bleakley and Chin (2010). However, given that many of the independent variables were binary, a linear index model may be superior to the linear methodologies used so far. Firstly, linear models like OLS or IV are not restricted to the [0,1] range of the binary variable. Secondly, they do not impose diminishing effects of the dependent variables on the probability of the independent variable. The difference in outcome between an individual that speaks English 'well' and 'very well' may be different to the difference between speaking 'not at all' and 'not well'.

For example, a linear index model can be used for the dependent variable "is a single parent". Given the binary nature of the variable, it must follow some properties of Bernoulli random variables, namely:

(4) $$E[y|x] = P(y = 1|x) = P(x)$$

(5) $$Var(y|x) = p(x)[1 - p(x)]$$

Therefore, the model is estimating conditional probabilities which in this case is correspondent of modelling conditional means, much like OLS. In fact, P(x) is often called the response probability. A binary variable will possess a natural heteroskedasticity because the conditional variance of y depends on x, except in the special case where probability of the outcome does not depend on x.

In the context of this research, only the probit linear index model will be used, it assumes that there is a latent variable model, not observed by the researcher, that generates the observed value of the binary variable. The probit model is based on a non-linear transformation of a linear probability model (LPM). Mathematically, it can be represented as:

(6) $$P(y = 1 \mid X) = \Phi(X\beta), \text{ for } \Phi: \mathbb{R} \rightarrow (0,1),$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $X\beta$ is the LPM specification. Since these models are non-linear, they must be solved using a maximum likelihood estimator.

A probit model would assume that the LPM residual u is distributed normally. The latent variable is not observable, and as a result, the distribution of u cannot be tested. In order to retrieve interpretable coefficients of the linear index models, we must analyze the average marginal effects. The average partial effect (APE) for continuous variables is represented by:

$$\widehat{APE}_K = \hat{\beta}_j[N^{-1}\textstyle\sum_{i=1}^{N}\phi(X_i\hat{\beta})] \tag{7}$$

The marginal effects represent the change in predicted probability of the outcome for a 1-unit change in English proficiency/age arrived, holding all other variables constant. Since we have already established that English proficiency is an endogenous regressor, we will use an instrumentalized version of probit using the 2SLS approach. The full model is:

$$ENG = \delta_1 \cdot agearrived + \delta_2 Z + \varepsilon \tag{8}$$

$$y^* = \beta_1 ENG + \beta_2 Z + u \tag{9}$$

$$y = y^*\ if\ y^* > 0\ and\ y = 0\ if\ y^* \le 0 \tag{10}$$

Where the errors (ε, u) follow bivariate distribution:

$$\binom{\varepsilon}{u} \sim N\left(\binom{0}{0}, \begin{pmatrix} 1 & \rho\sigma_2 \\ \rho\sigma_2 & \sigma_2^2 \end{pmatrix}\right) \tag{11}$$

In this model, *agearrived* is the instrument, Z is a set of exogenous variables, y* is the latent dependent variable and y is the observed dependent variable. ENG is endogenous because:

$$if\quad corr(\varepsilon, u) \neq 0\ \ and\ \ corr(ENG, \varepsilon) \neq 0\ \longrightarrow\ corr(ENG, u) \neq 0$$

The first equation of the model can be estimated directly using standard OLS. In fact, it is the same regression as the first stage in the IV 2SLS procedure. This model is essentially an augmented version of the IV 2SLS methodology. For the second equation of the model, we can decompose the error term u in to the endogenous and exogenous components. Given the bivariate distribution outlined in equation (11), then, conditional on ε, u will have the following distribution:

$$u \sim N\left(\rho\frac{\varepsilon}{\sigma_2}, \sqrt{1-\rho^2}\right), hence\colon u = \rho\frac{\varepsilon}{\sigma_2} + v \tag{12}$$

Such that v represents the exogenous component uncorrelated with y.

We plug in equation (12) into (9) to retrieve:

(13) $$y^* = \beta_1 ENG + \beta_2 Z + \rho \frac{\varepsilon}{\sigma_2} + v$$

Given that ε is observable, since we know the true English proficiency of each participant, the equation can finally be estimated directly. The final step is to rescale the equation such that the error term v is standardized and has a variance of 1. This is achieved by dividing every term by $\sqrt{1-\rho^2}$, and estimating the following model using the standard probit model.

(14) $$y^* = \frac{\beta_1}{\sqrt{1-\rho^2}} ENG + \frac{\beta_2}{\sqrt{1-\rho^2}} Z + \frac{\rho}{\sqrt{1-\rho^2}} \frac{\varepsilon}{\sigma_2} + \frac{v}{\sqrt{1-\rho^2}}$$

The average partial effects under iv linear indices are a modified version of equation (13) with respect to the predicted values of ENG:

(15) $$\widehat{APE}_{ENG} = \left(\frac{\beta_1}{\sqrt{1-\rho^2}}\right) \cdot g\left(\frac{\beta_1}{\sqrt{1-\rho^2}} ENG + \frac{\beta_2}{\sqrt{1-\rho^2}} Z + \frac{\rho}{\sqrt{1-\rho^2}} \frac{\varepsilon}{\sigma_2}\right)$$

We run the IV probit model on the marital status outcome variables, we were unable to standardize the exogenous component of the error term like in equation 14 or 15. This is due to the *ivprobit* command being too computationally intensive with all the country of origin and age dummies. Instead, the reported results from Table 6 display the unstandardized marginal effects which will exhibit some bias. We note that for all estimated partial effects (except for probability of having ever married for women) values are close to standard IV estimates but have a much higher significance level. In fact, in some cases (like the probability of being divorced for men) the IV probit model uncovers significance when the standard IV model did not at all. The fact that marginal effects are relatively similar provides evidence that the instrument “age at arrival” is likely valid and captures the variation in English proficiency between participants effectively.

	All childhood immigrants				**Women**				**Men**			
	IV	t-statistic	IV probit	t-statistic	IV	t-statistic	IV probit	t-statistic	IV	t-statistic	IV probit	t-statistic
Panel A: Marital Status												
Is currently married with spouse present	-0.1080**	-2.72	-0.1086***	-4.72	-0.0760*	-2.38	-0.0766**	-2.61	-0.1410**	-2.63	-0.1418***	-3.95
Is currently divorced	0.0524**	3.28	0.0593***	4.27	0.0643*	2.37	0.0683***	3.53	0.0379	1.72	0.0483**	2.44
Has ever married	-0.0717*	-2.56	-0.0797***	-4.34	-0.0152	-0.55	-0.0299	-1.36	-0.1340**	-2.69	-.1354***	-4.48

* p<0.05, ** p<0.01, *** p<0.001"

Table 6: IV probit marginal effect of English Language

The greater significance levels in the linear index model estimated average partial effects may reflect stronger distributional assumptions of the residuals compared to the IV model. Furthermore, linear index models are better suited for binary variables, because they model the probability of an event rather than the level of an outcome directly. If the binary nature and the diminishing effects of the dependent variables imposed captured by the linear index models fits the data better, it will lead to a much greater significance. However, given the computational intensity of these models, especially the calculation of marginal effects, they should mainly be used as robustness checks of the main paper's methodology. The IV regressions provide strong enough estimates for inference. A final extension to this could be attempt a logit model for events that are more extreme like being divorced, due to the fatter tails of the logistic distribution. However, the difference in distributional assumptions makes this sort of model much more complicated to implement than a probit model.

Bibliography:

Avila, F.R. (Retrieved in 2024). "The IV-Probit Model". *Github*. https://friosavila.github.io/playingwithstata/main_ivprobit.html

Bleakley, H. & Chin, A. (2010). "Age at Arrival, English Proficiency and Social Assimilation Among US Immigrants". *American Economic Journal: Applied Economics, Vol. 2, No. 1 (pp. 165-192).* Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants - American Economic Association (aeaweb.org)

Birdsong, David. (2006). "Age and Second Language Acquisition and Processing: A Selective Overview." Language Learning, 56(S1): 9–49.

Li, C., Poskitt, D.S, Windmeijer, F. & Zhao, X. "Binary Outcomes, OLS, 2SLS, and IV Probit". (2019). *Monash Business School*. https://www.monash.edu/business/ebs/research/publications/ebs/wp05-2019.pdf

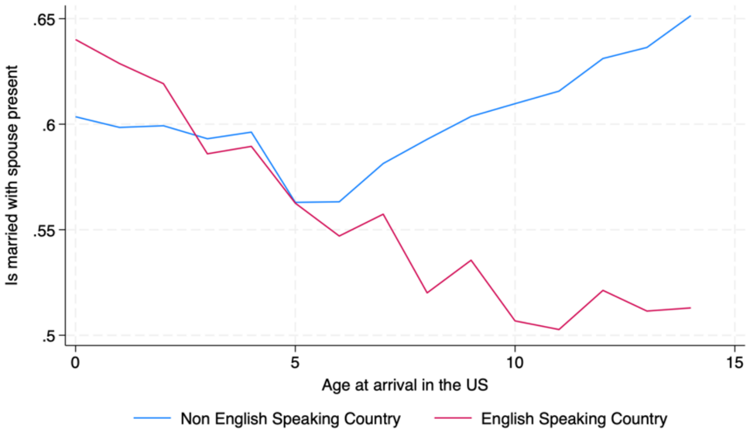
Newport, Elissa L. (2002). "Critical Periods in Language Development." In *Encyclopedia of Cognitive Science*, ed. Lynn Nadel. London: Macmillan Publishers and Nature Publishing Group.

Appendix 1: Descriptive statistics for treatment and control groups, decomposed by age at arrival

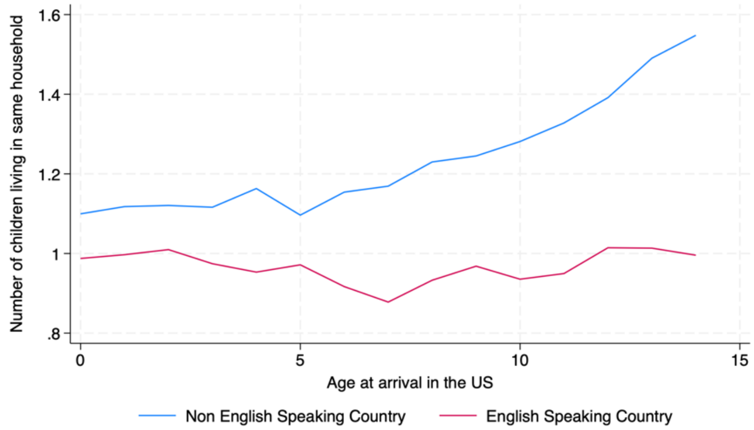
		Born in non-english speaking country			Born in english speaking country				
	Full sample	Total	Arrived 0-9	Arrived 10-14	Total	Arrived 0-9	Arrived 10-14	Min	Max
Panel A: Regressors									
English-speaking ability ordinal measure	2.754223	2.718557	2.87225	2.440772	2.980476	2.980831	2.979495	0	3
Age	36.80163	36.54926	36.83917	36.02527	38.40262	38.90563	37.01371	25	55
Female	.5036406	.4998331	.5118721	.4780736	.5277947	.5129425	.5688053	0	1
White	.569191	.5539557	.6091734	.4541547	.6658409	.7656241	.3903162	0	1
Black	.0598547	.0302889	.0306379	.029658	.247414	.1640271	.4776649	0	1
Asian/ Pacific-Islander	.1097994	.1227306	.1056105	.1536737	.0277662	.0239012	.0384384	0	1
Other single race	.2099049	.2400298	.2014068	.3098374	.0187988	.0155428	.0277892	0	1
Multi-racial	.05125	.052995	.0531714	.0526762	.0401801	.0309048	.0657914	0	1
Hispanic	.4511446	.5203673	.4519836	.6439643	.0120103	.0122179	.011437	0	1
Years of Schooling	13.32843	13.13825	13.75288	12.01352	14.52655	14.59304	14.34192	0	20
Panel B: Marriage Outcomes									
Is currently married with spouse present	.598028	.6038349	.588234	.6321225	.5612445	.5843406	.4972327	0	1
Is currently divorced	.0998466	.0966646	.1072601	.0774526	.1200033	.1230343	.1116029	0	1
Has ever married	.7631284	.767479	.7567636	.7869083	.7355694	.7554436	.6804871	0	1
Number of children living in the same household	1.20923	1.246333	1.141432	1.43593	.9738589	.9694233	.9861067	0	9
Has children living in the same household	.5703437	.580438	.5540421	.628146	.5063078	.5022431	.5175313	0	1
Is a single parent	.1015798	.1007803	.1006974	.1009308	.1066438	.0931879	.1439375	0	1
Is a never-married single parent	.0304784	.0293582	.0275753	.032591	.0375743	.0281293	.0637516	0	1
Panel C: Residencial location									
Fraction of PUMA population from same country of birth	.0546737	.0620286	.0514001	.0813047	.0082219	.0061931	.0138302	0	.4880442
Fraction from same country of birth is above the mean for the country of birth	.336983	.3348322	.3116	.3769663	.3505669	.3474912	.359069	0	1
Fraction of PUMA population with same primary ancestry	.104515	.1140583	.1059805	.1285033	.0438068	.0466625	.0359595	0	0.735508
Fraction with same country ancestry is above the mean for the national mean for the primary ancestry	.3635943	.3637854	.35221	.3844848	.3623786	.3639262	.3581257	0	1

Appendix 2: Plot of means between treatment and control group for 4 social outcomes

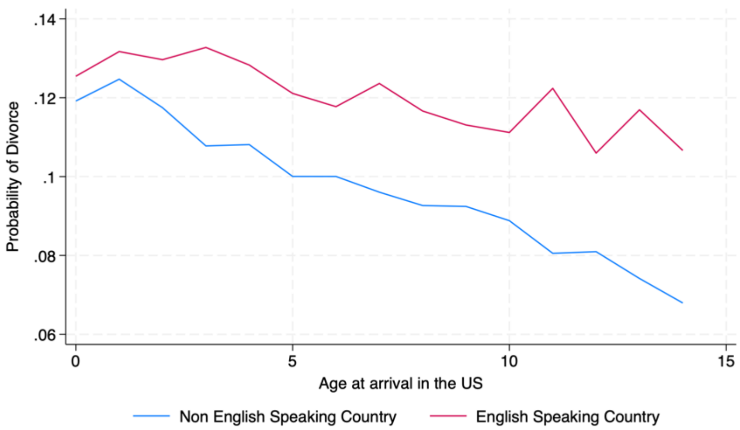
Panel A: Currently married with spouse present



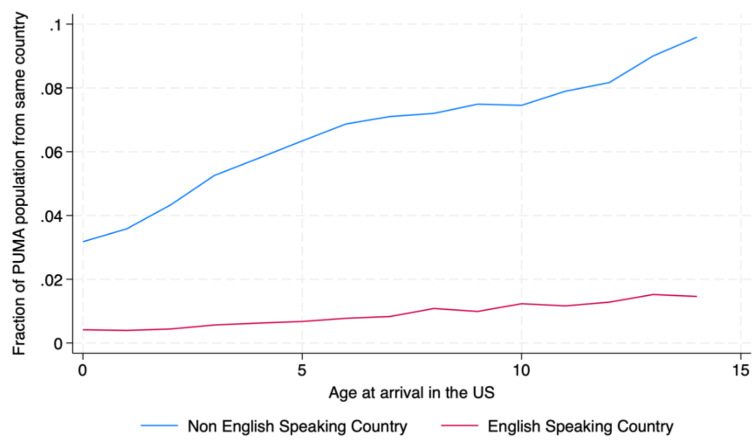
Panel B: Children living in same household



Panel C: Is currently divorced

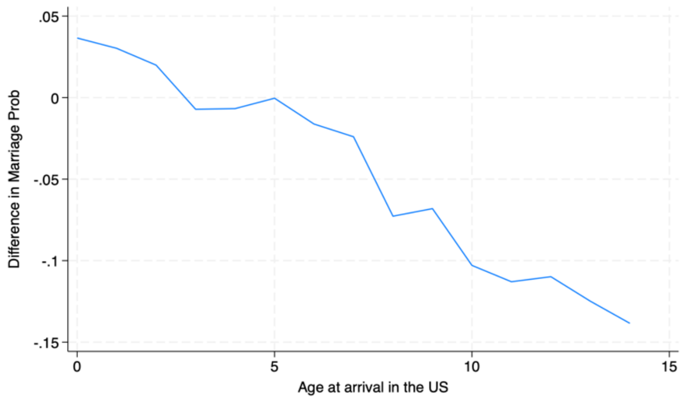


Panel D: Fraction of PUMA from same country

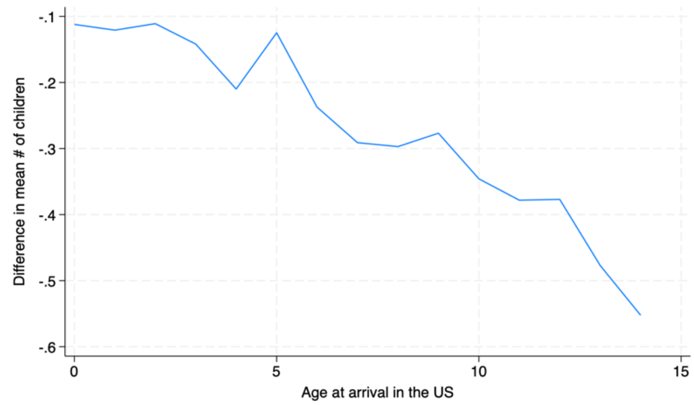


Appendix 3: Differences in means between treatment and control group for 4 social outcomes

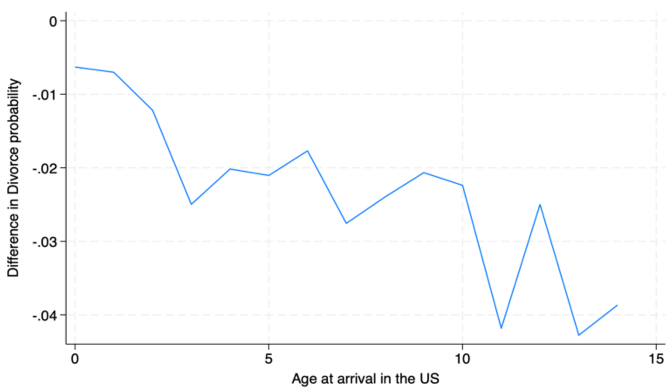
Panel A: Differences in marriage outcomes



Panel B: Differences in number of children



Panel C: Differences in divorce outcomes



Panel D: Differences in ethnic enclaves