

# **Has the Return to English Proficiency Decreased for Immigrants?**

Jack Katz

State University of New York at Geneseo

February 2024 draft, not for citation

Advisor: Léonie Stone

## **Abstract**

New immigrants need to be successfully economically and socially integrated if they are to bring positive economic effects. One important piece of integration into a new society is language proficiency. As the U.S. economy increasingly relies on immigrant labor, it will become even more important to understand the returns to speaking English as an immigrant as well as the determinants of English proficiency among immigrants. I focus specifically on the return to learning English as a Spanish speaker. As the U.S. becomes a more Spanish-speaker accessible country, as more government and other agencies are offering forms in Spanish, as well as broader use of Spanish and Spanish media in American culture, the incentive to learn English may be decreasing for Spanish immigrants.

This paper hypothesizes that the return to learning English has decreased over the years, both because the United States has become a more Spanish-friendly country and because the number of Spanish speakers has increased in the country over time. I analyze data from the American Community Survey (ACS) from 1980-2019. I first consider the determinants of English proficiency. I then use an instrumental variables approach to consider the effect of English proficiency on income, the return to proficiency.

While I find no significant change in returns to English proficiency from decade to decade, there is a significant decrease between the 1980 sample and the 2019 sample. The results also show a significant reduction in the income gap between female and male Hispanic immigrants.

## **Introduction and Literature Review**

At a time of slowing domestic population growth, immigration has the potential to mitigate some of the reduction in both domestic population and labor force participation. New immigrants need to be successfully economically and socially integrated if they are to bring positive economic benefits, and one important part of this is language proficiency.

Adsera and Pytlkova (2012), Bleakley and Chin (2004, 2010), and Chiswick and Miller (1995) have identified language skills as an important part of transferring human capital from one country to another. Researchers have largely used instrumental variables approaches to control for the endogeneity of English proficiency, finding a positive relationship between English proficiency among immigrants to the U.S. and earnings.

It is also important to recognize that immigrants may have different incentives to learn English upon arrival to the United States. One of these is the enclave effect. Chiswick and Miller (1995, 1996), Baur et al. (2003), and Danzer et al. (2022) show that immigrants from a new country who are deciding where to settle are generally influenced by three factors: proximity to ports; where other people with the same ethnic or cultural background live; and third places with economic activity. If new immigrants largely settle in areas in which there are already concentrations of people who speak their native language, the incentive to learn English can be reduced.

Immigrants who have poor language skills may find barriers to work outside of their linguistic enclave in the forms of proficiency tests or discrimination. Then even if they can find work, productivity may suffer from difficulty in communicating, as in Chiswick and Miller (1996). For this reason, it is easier for immigrants to stay within their cultural linguistic enclave, which would take away the need to become fully proficient in the dominant language of the new

country. Bauer et. al (2003) research variables affecting the probability that Mexican immigrants to the United States would move to a cultural enclave. They found that the probability of a Mexican immigrant to the United States moving to a high percentage Hispanic population becomes higher as English proficiency decreases. Danzer et al. (2022) use panel data of children of immigrants to Germany and find a negative relationship between German proficiency and population percentage of the child's ethnicity.

Another mechanism that influences the incentive to learn English is family structure. The child-parent relationship is significant in influencing the English proficiency of the parent. In some ways, having a child can negatively impact the parent's English proficiency. If the children act as interpreters for the parents this takes away the incentive for the parents to learn English. Or if the parents find it important to preserve their mother tongue by speaking it in the home this can also make it more difficult for the parents to improve their English ability. Having children can also have positive effects on English language learning. Children of immigrants normally have an easier time learning the new language and culture because of their exposure at school. The children then might speak English around the home making it easier for the parents to acquire English skills, as demonstrated in Chiswick and Miller (2004).

This study examines immigrants from Spanish-speaking countries. While immigration in general has leveled off in recent years, immigrants from South and Central American countries are immigrating to the United States in larger numbers. As more Spanish-speaking people enter the labor force, it is important to know how to create policy to properly socially and economically integrate Spanish-speaking immigrants. Additionally, Spanish-speaking immigrants are in a unique linguistic position in the U.S. Although the U.S. is a monolingual country, Spanish has a special place within American culture, media, and education. In 2000,

60% of all high school students took Spanish for at least some credit. This number had increased to 70% by 2019 (National Center for Education Statistics, 2022). Spanish is a language choice for many types of forms and services, and entertainment and social media for Spanish speakers has proliferated. With more accessible media and other resources, the returns to speaking English may have declined. These characteristics also make Spanish-speaking immigrants a good group to test the enclave effect variable.

### **Data and Methodology**

This study uses data from the Integrated Public Use Microdata Series (IPUMS). The data set includes respondents from 5 sample years of the American Community Survey (ACS), 1980, 1990, 2000, 2010, and 2019 (2020 was not used due to unreliable data from the pandemic). The ACS is administered every month and collects data on education, income, and family situations, as well as information about immigration and country of origin.

The data was filtered as in Bleakley and Chin (2004). I filtered the data for people who arrived in the United States 14-30 years before the sample year and who were ages 0-18 years of age when they immigrated. I used this same filter for two reasons. Being in the country for 14-30 years gives immigrants enough time to learn English to a plateau point where they will probably have reached their peak ability, and being between 0-18 years of age at immigration makes this variable not a choice variable. In that age range, these immigrants would have followed their parents to the United States and would not have taken English proficiency into account when moving. This takes away the endogeneity factor in the age at arrival variable. The data was further cleaned to remove respondents who attended graduate school, because they are not represented in the 1980 sample. Finally, the data includes only immigrants from Spanish-

speaking countries. The data was further filtered to include only observations from U.S. counties in which there was a significant Hispanic population, using only responses from the top ten Hispanic counties in any sample year. While this limits the size of the sample, within the top ten counties, population percentages vary from around 10% to 80%, giving significant variability, and moving to more counties reduces survey responses to very small numbers. Sample sizes then range from 910 (1980) to 4,783 (2000).

I first estimate English proficiency as a function of age at arrival, age of oldest child in household, age of youngest child in household, years of education, and the percent of county population that is Hispanic.

The English proficiency variable is qualitative and, following the conventions of the ACS, is coded as:

- 1 – Speaks only English
- 2 – Speaks English very well
- 3 – Speaks English well
- 4 – Speaks English but not well
- 5 – Does not speak English

Thus, it is important to recognize that a *lower* number represents *greater* proficiency.

Age at arrival plays an important role in predicting dominant language proficiency. Human development finds that the human brain has a much easier time acquiring new language skills from birth through early childhood, whereas adolescents and adults cannot learn new languages as easily. This has a predicted positive sign since the older a respondent was when they immigrated the higher their ranking in the English proficiency variable will be, or in other words, English skills will be worse when age of arrival is later.

Parents who have older *eldest* children may be able to rely on them more for translation help and thus have less of a need to learn English, so the predicted sign for this variable is positive. However, as the age of the youngest child increases there is a chance that the English that is being learned in school will permeate the home and support better English proficiency among parents.

The education variable measures the amount of years the respondent was in school, coded as 5 for elementary school, 8 for middle school, 12 for high school, and 16 for college. As years of education increase the value of English proficiency variable will decrease, and so the expected sign is negative.

Finally, in areas with greater Hispanic populations, the enclave effect is likely to be more important. Thus the expected sign is positive, because a greater enclave effect leads to lower English proficiency.

In the second regression, I take an instrumental variables approach due to the endogeneity of English proficiency. The dependent variable is annual wage and salary income, again following the literature. In addition to the variables used as instruments from the first regression, I add gender and experience as independent variables. Gender is a dummy variable, coded as 1 if female. Experience, as in Chiswick and Miller (1995) and other wage studies, is

$$\frac{(Age - years\ of\ schooling - 5)^2}{100}.$$

## **Results**

I estimate two equations for each of the five sample periods. The first regressions estimate coefficients for the variables predicting English proficiency (Table 1). The second equation uses the significant variables from the first equation as instruments for the English proficiency variable (Table 2). In the first equation, independent variables generally matched

most of the *a priori* assumptions for partial effects and significance, with slight variability across sample years. In particular, I note the significant and positive effect of the Hispanic population variable, supporting the existence of an enclave effect.

In the second estimation, again, variables are generally highly significant and have the expected signs. The primary research objective of this study is to evaluate whether the return to English proficiency has changed over the decades covered by this data. Thus I then tested for the difference in the English proficiency coefficients between the five samples. From year to year, there is no significant difference between coefficients, likely because this variable changes slowly. However, there is a significant difference over the entire period. In 1980, a one-unit decrease in the English proficiency variable implies a 64% higher income. In 2019, the last sample year, a one-unit decrease in the English proficiency variable means a 37% higher income.

The gender variable produced a similar result. Immigrant women generally make less money compared to their male counterparts, but that gap is decreasing. Long (1980) found that often a wife will enter the labor force upon arrival in the new country and work until the husband can join the rest of the family. After the husband arrives, the wife will shift to non-market activities. Thus, women tend to have less labor market experience, if they work at all. Over the sample period, the coefficient on gender decreases from -0.91 in 1980 to -0.6 in 2019, indicating a significant reduction in the gender wage gap.

Although the adjusted r-squares are small in instrumental variables regression (Table 2) these are cross-sectional models. There are many factors that determine income, so relatively small goodness-of-fit measures are normal for studies of this nature.

# Table 1: OLS Regression

Dependent variable:					
	English Proficiency				
	1980 (1)	1990 (2)	2000 (3)	2010 (4)	2019 (5)
AAA	0.030*** (0.005)	0.037*** (0.003)	0.056*** (0.003)	0.047*** (0.003)	0.052*** (0.004)
ELDCH	0.022*** (0.006)	0.007** (0.003)	0.003 (0.003)	0.013*** (0.003)	0.016*** (0.004)
YNGCH	-0.022*** (0.006)	-0.014*** (0.003)	-0.007** (0.003)	-0.008** (0.003)	-0.017*** (0.004)
Hispanic_Population	0.006*** (0.002)	-0.0005 (0.001)	0.003** (0.001)	0.002* (0.001)	0.006*** (0.002)
Years_Education	-0.127*** (0.007)	-0.108*** (0.004)	-0.102*** (0.004)	-0.111*** (0.005)	-0.095*** (0.006)
Constant	3.573*** (0.133)	3.573*** (0.082)	3.147*** (0.081)	3.414*** (0.089)	2.928*** (0.128)
Observations	910	3,077	4,783	3,646	2,655
R2	0.353	0.274	0.262	0.246	0.225
Adjusted R2	0.350	0.273	0.261	0.245	0.223
Residual Std. Error	0.713 (df = 904)	0.789 (df = 3071)	0.832 (df = 4777)	0.858 (df = 3640)	0.861 (df = 2649)
F Statistic	98.723*** (df = 5; 904)	231.440*** (df = 5; 3071)	338.339*** (df = 5; 4777)	237.771*** (df = 5; 3640)	153.598*** (df = 5; 2649)

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

# Table 2: Instrumental Variables

Dependent variable:					
	Log of Total Income				
	1980 (1)	1990 (2)	2000 (3)	2010 (4)	2019 (5)
SPEAKENG_new	-0.649*** (0.095)	-0.555*** (0.045)	-0.406*** (0.039)	-0.534*** (0.045)	-0.379*** (0.053)
Experience_Squared	0.053*** (0.015)	0.036*** (0.008)	0.039*** (0.007)	0.054*** (0.007)	0.030*** (0.008)
Gender	-0.920*** (0.067)	-0.733*** (0.035)	-0.658*** (0.028)	-0.586*** (0.031)	-0.595*** (0.034)
Constant	11.825*** (0.210)	11.570*** (0.106)	11.049*** (0.093)	11.236*** (0.113)	10.957*** (0.128)
Observations	910	3,077	4,783	3,646	2,655
R2	0.146	0.088	0.068	0.020	0.056
Adjusted R2	0.143	0.087	0.068	0.019	0.055
Residual Std. Error	0.978 (df = 906)	0.960 (df = 3073)	0.940 (df = 4779)	0.927 (df = 3642)	0.868 (df = 2651)

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01



## **Conclusion and Future Directions of Research**

In this study, I examined determinants of English proficiency among Hispanic immigrants to the United States. I found evidence that family structure and the enclave effect were significant in determining English proficiency. From a policy perspective, the enclave effect suggests that growing Hispanic populations may actually deter greater English proficiency, at least for new immigrants.

Secondly, I found that, decade to decade, the returns to English proficiency do not change significantly, which is not surprising. But over the fifty year period spanned by the sample, there has been a significant decrease in the returns to English proficiency.

At third result, which merits further study, is the significant decrease in the wage gap between male and female Hispanic immigrants.

This study overall confirms the work of Chiswick and Miller (1995), which found a positive relationship between dominant language proficiency and income and the enclave effect variable, with updated data. The significance of the age-at-arrival variable also confirms the later works of Bleakley and Chin (2004, 2010).

The findings here have important implications from a policy perspective. If the returns to English proficiency are decreasing, immigrants may be slower to adapt. Thus programs that increase incentives to acquire language proficiency would be desirable. Also, native English speakers may have more incentive to learn Spanish, which could be supported through further emphasis on Spanish language education from an earlier age to children born to native English-speaking parents.

Further directions of research take at least two paths. First, it would be desirable to adjust the Hispanic population variable to include smaller geographic locations and test a wider variety

of places around the United States. I also plan a panel study of this data, and given the low explanatory power of the IV regressions, additional variables both improve the fit and make a better argument for robustness. Second, both the significant reduction of the gender wage gap and the policy implications of the implied trend towards lower levels of proficiency merit study.

## Appendix 1: Summary Statistics

### Summary Statistics: 1980

Statistic	N	Mean	St. Dev.	Min	Max
AAA	910	11.215	4.996	0	18
ELDCH	910	11.244	6.487	0	31
YNGCH	910	6.915	5.682	0	29
Hispanic_Population	910	28.826	12.926	13.900	81.300
Years_Education	910	11.510	3.424	5	16
SPEAKENG_new	910	2.702	0.884	1	5
Experience_Squared	910	4.427	3.351	0.000	14.440

### Summary Statistics: 1990

Statistic	N	Mean	St. Dev.	Min	Max
AAA	3,077	11.366	5.188	0	18
Citizenship_Status	3,077	0.390	0.488	0	1
ELDCH	3,077	12.096	6.568	0	69
YNGCH	3,077	6.880	5.884	0	69
Hispanic_Population	3,077	33.298	11.748	13.600	49.200
Years_Education	3,077	11.742	3.492	5	16
SPEAKENG_new	3,077	2.703	0.925	1	5
Experience_Squared	3,077	4.116	2.793	0.000	14.440

### Summary Statistics: 2000

Statistic	N	Mean	St. Dev.	Min	Max
AAA	4,783	12.105	5.305	0	18
Citizenship_Status	4,783	0.457	0.498	0	1
ELDCH	4,783	11.575	6.142	0	36
YNGCH	4,783	6.023	5.017	0	36
Hispanic_Population	4,783	35.878	9.218	19.900	48.400
Years_Education	4,783	11.128	3.107	5	16
SPEAKENG_new	4,783	2.784	0.968	1	5
Experience_Squared	4,783	4.012	2.827	0.000	14.440

Summary Statistics: 2010

Statistic	N	Mean	St. Dev.	Min	Max
AAA	3,646	13.197	4.869	0	18
Citizenship_Status	3,646	0.383	0.486	0	1
ELDCH	3,646	13.209	6.362	0	39
YNGCH	3,646	7.308	5.683	0	34
Hispanic_Population	3,646	36.183	13.133	11.100	47.700
Years_Education	3,646	11.453	3.096	5	16
SPEAKENG_new	3,646	2.961	0.988	1	5
Experience_Squared	3,646	4.683	2.995	0.000	14.440

## References

- Alexander M. Danzer, Carsteen Feuerbaum, Marc Piopiunik and Ludger Woessmann, “Growing up in Ethnic Enclaves: Language Proficiency and Educational Attainment of Immigrant Children,” *Journal of Population Economics* 35, (April 2022): 1297 – 1344. Springer Link
- Alicia Adsera and Mariola Pytlikova, explore the impact of language on international migration in their article “The Roles of Language in Shaping International Migration” published in *The Economic Journal* (2012).
- Barry R. Chiswick, Paul W. Miller, “The Endogeneity between Language and Earnings: International Analyses,” *Journal of Labor Economics* 13, No.2 (April 1995): 246-288. JSTOR
- Barry R. Chiswick, Yew Liang Lee and Paul W. Miller, “Parents and Children Talk: The Family Dynamics of English Language Proficiency,” *Review of Economics of the Household* 3, (September 2005), 243 – 268
- Borjas, George J. "Immigration Economics and Immigration Policy." *Harvard University* 24 (2018).
- Gordon Hanson, Pia Orrenius, and Madeline Zavodny, “US Immigration from Latin America in Historical Perspective” *Journal of Economic Perspectives* 37, No. 1, (Winter 2023): 199-222, <https://doi.org/10.1257/jep.37.1.199>
- Hoyt Bleakley and Aimee Chin, “Language Skills and Earnings: Evidence From Childhood Immigrants,” *Review of Economics and Statistics* 86, No. 2 (May 2004): 481 – 496. JSTOR
- Long, James E. “The Effect of Americanization on Earnings: Some Evidence for Women”, *Journal of Political Economy* 88, No. 3, 620-629 (June 1980).
- Mevlude Akbulut-Yuksel, Hoyt Bleakley, and Aimee Chin, “The Effects of English Proficiency Among Childhood Immigrants: Are Hispanics Different?”, *Latinos and the Economy*, (November 2010): 255-283, [https://doi.org/10.1007/978-1-4419-6682-7\\_13](https://doi.org/10.1007/978-1-4419-6682-7_13)
- Pew Research Center, *Hispanic Population Growth and Dispersion Across U.S. Counties, 1980-2020*

Thomas Bauer, Gil S. Epstein, and Ira N. Gang, "Enclaves, Language, and the Location Choice of Migrants," *Journal of Population Economics* 18, (October 2005), 649-662, 10.1007/s00148-005-0009-z

U.S. Department of Education, *Number and percentage of public and private high school graduates who took foreign language courses in high school and average number of credits earned, by language and number of credits: 2000, 2009, and 2019*, National Center for Education Statistics, (2022)

