2015

The Effects of Foreign Language Acquisition on Wages for US College Graduates

Christopher Nguyen
Oberlin College

The Effects of Foreign Language Acquisition on Wages for US College Graduates

Christopher Nguyen

May 11, 2015

Abstract: This paper seeks to determine whether fluency in a foreign language has a significant impact on future earnings. The sample in question is made up of US college graduates who attained their bachelor's degrees during the 2007-2008 school year. While previous literature suggests that there is a positive impact of 2-3% higher wages on top of a bachelor's degree for those who are fluent in a foreign language, I am unable to find this general premium for fluency in a foreign language. I look at interactions of foreign language and their relative difficulty and, once again, do not find a significant relationship between the difficulty of a language and wages. However, I find that foreign language fluency provides a significant impact on wages only in certain sectors. Specifically, a positive increase in hourly wage is found in personal service and sales occupations. A negative relationship with hourly wage is found in science and technology related as well as "other" (which includes agricultural) fields.

Introduction

According to the 2013 American Community Survey, in 2010 nearly 20% of the US population, or 60 million people, spoke a language other than English at home.[1] With such a significant portion of the United States self-reporting these levels of language proficiency, one may wonder whether or not those who acquire foreign language[2] skills via formal schooling would stand to gain economic benefits during their lifetime.

In 2014 the Economist suggests that not only are economic benefits present for foreign language learners, but also that there is evidence for the prevention of dementia, as well as for better decision making in general. Taking a closer look at economic benefits, Saiz and

[1] Data on the use of foreign languages at work would be useful; however, US Census data and other sources I look at only look at language use at home.

[2] I note that 'foreign language' for the context of this paper is defined as the language not dominant to the country in question. Specifically, since I am taking a look at US college graduates, 'foreign language' refers to any non-English language.

Zoido (2005) find a wage premium of up to 3.8% for US college graduates who acquired German language proficiency, which was the highest language premium of the languages analyzed. Also, while not necessarily a causal relationship, outside of the US, the top 10 richest countries include "countries where trilingualism is typical, like Luxembourg, Switzerland and Singapore, and small countries like the Scandinavian ones, where [a second language, typically English, is excellent.]" Furthermore, a study by James Foreman-Peck of Cardiff Business School has estimated the "lack of foreign-language proficiency in Britain costs the economy £48 billion ($80 billion), or 3.5% of GDP, each year." These findings suggest that economic benefits to bilingualism are a worldwide phenomenon that there are, and that underinvestment in language study may hinder an economy.

This paper aims to explore the wage premiums of native US college graduates for acquiring fluency in a second language in the United States. Intuitively, these economic benefits arise from a demand for the language among job postings and employers, as well as the relative scarcity of proficient speakers of that language in the country. The labor market also favors skills that are outside of actually using the foreign language skills that are signaled by foreign language proficiency. Beyond increases in wages, there are of course other benefits to studying a foreign language. Being able to communicate with natives in a country other than your own, whether you are ordering at a restaurant, asking for directions, or eventually working at a foreign-owned firm, has its own personal and practical significance. While this paper is unable to identify those benefits, it does focus on the labor market returns to studying a foreign language. To avoid overselling the results, Saiz and Zoido note that one extra year of general schooling has proven economically more beneficial than language study, which may explain the pattern of foreign-language

investment in the United States today (Saiz 524). While Saiz and Zoido's results were based upon a particular cohort who graduated in 1992-1993, this paper will directly address if the economic benefits are similar in a more recent cohort of US college graduates.

Literature Review

While there is an increasing literature on the economic benefits of acquiring English as a foreign language, the literature on non-English languages appears to still be limited. However, I find a substantial contribution to the topic with a paper written by Albert Saiz and Elena Zoido titled "Listening to What the World Says: Bilingualism and Earnings in the United States" (2005), which is the basis for answering my research question. Their main data source, the "Baccalaureate and Beyond Longitudinal Study," keeps track of a nationally representative sample of US college graduates who received their bachelor's degrees during the 1992-1993 school year. Saiz and Zoido control for a variety of factors that might affect post graduation earnings, including race, gender, marital status, and work experience, while focusing on the dummy variable for fluency in a second language. Furthermore, they control for cognitive ability through SAT scores as well as GPA. Ultimately, Saiz and Zoido find that there is a 2-3% increase in the wage premium four years after graduation for those who studied a foreign language. These results are confirmed through OLS estimation, panel data methods, and propensity score methods. However, they are also careful to highlight the potential problems with their analysis. First of all, the wage premium may be due to unobserved differences that language learners are likely to have from the general population. Returns for language learners are higher in specific occupations, primarily business and management services. Finally, their sample of college graduates may not allow for generalization to the entire US population (Saiz 535).

Barry R. Chiswick has done research that is relevant to the connection between immigrants, language, and earnings. "Hebrew Language Usage: Determinants and Effects on Earnings among Immigrants in Israel" (1998) has two points of focus: the determinants of Hebrew language use among male immigrants in Israel, and the effects of Hebrew-speaking skills on the earnings of adult male immigrants to Israel. Chiswick uses the 1983 Census of Israel for analyzing the economic standing of immigrants. In analyzing the effect on wages, it is found that those with Hebrew and English language ability have the highest earnings, regardless of country of origin. English in this case is considered to be a foreign language, which parallels with my own analysis of foreign languages within the United States. However, Chiswick also brings up the importance of noting that wage effects also influence the acquisition of a language, in that higher wage effects provide incentive for learning the dominant language. My own data does not allow me to analyze wage effects on language acquisition, so I only take a look at the relationship in one direction. Also, when refugees come to a country, their incentives for learning language can differ drastically from those who come more willingly since they may be more concerned with survival as opposed to the pure economic benefits of studying the dominant language (Chiswick 269).

Taking a broader look at the economic effects of education, Hanushek and Woessman (2007) provide a paper titled "The Role of Education Quality in Economic Growth." While the findings are broad, Hanushek and Woessman essentially conclude that the cognitive skills of a population, rather than the quantity of years in school, have a significant impact on the economic growth. Generally, if an educational institution is already determined to provide a high quality education, then years of schooling has a positive correlation with economic growth. However, if education quality is low, this does not hold true. Thus, if

language skills are a proxy for unobserved cognitive skills, a wage premium for foreign language could be overestimated.

Empirical Strategy

A classic empirical approach that links wage to education starts with Mincer's human capital function (1974). This is well described in Carl (1999) and was also used in Saiz and Zoido (2005).

I pose this Mincerian regression as an example:

$\log w_i = a + bFL_i + cX_i + e_{i.,}$

where log w is the log of earnings, FL is an indicator of whether the individual speaks a second language, and X is a vector of personal characteristics.

Saiz and Zoido have gone into detail on the methodological challenges of measuring foreign language knowledge on earnings, which they take into account with their own OLS Estimates. They include variables for the following characteristics: Speaks FL, Experience, Male, Married, Black, Public college, Normalized College GPA, Spoke FL in 1983 only (as an indicator of whether they continued to study the language), Parents' Education, Major, Graduate Degree, SAT-ACT Quartile, College Quality, and State High School Requirements (for foreign language classes). In addition, they have characteristics of the individual's location using measures of log state income.

Similar to Saiz and Zoido, in addition to using the entire sample, I construct subsamples of native English speakers to isolate the effect of foreign languages on earnings. I extend the work of Saiz and Zoido to specify the acquired foreign language spoken in order to explore the possibly different effects of a language that is considered relatively more difficult. I hypothesize that the more difficult a language is relative to English, the smaller the number

of US college graduates who will have attempted to become fluent. This relative scarcity of proficient speakers may be met by a high demand for employees fluent in these more difficult languages. For example, sales firms attempting to break into foreign markets may experience a need for these language skills. I also take a look at the interaction of foreign language fluency and occupational field. Though this is an exercise carried out by Saiz and Zoido, my analysis uses a more recent cohort and is divided into subsamples to control for those who grew up with English at home.

It is important to keep in mind that the cohort I have looked at is a completely different one from which Saiz and Zoido analyzed. This means that the results cannot be interpreted as a change among the previous cohort's wages. However, since both cohorts graduated with BA's from US institutions, we may assume that these results still reflect differences in economic benefits of foreign language acquisition. While the current data set only provides language knowledge in 2009, about a year after graduation, I look at earnings in 2012, the most recent information on earnings that is available through my data. This allows about 3 to 4 years for respondents to either start working or enroll in a graduate program after graduation.

One may ask why this study needs to be replicated now when similar research was already carried out in a 2005 paper. Conditions surrounding the demand for foreign language acquisition may have changed since Saiz and Zoido's research was carried out. An easy way to look at this is to take a brief look at data from monster.com and careerbuilder.com. Saiz and Zoido use data from these websites as an indicator for foreign language requirements after graduation. Saiz and Zoido search for listings within a 24-hour period that require a bachelor's degree. From these results, they narrow down the search

to postings that required a specific language (Spanish, French, German, Chinese, Arabic, Korean, Japanese, and foreign language in general). They list in a chart how many relevant findings were found for each language, though they do not talk about how many postings originally showed up. I attempted this process at the beginning of my research. However, where Saiz and Zoido find 33 listings for jobs requiring Spanish and 2 requiring Japanese, a recent search in the last few months found 752 and 106, respectively. To explain these differences, English has become an international language, and as the rest of the world continues to acquire English, there may be less need for foreign language acquisition in a country where English is the dominant language. On other hand, as evidenced from the online world, the demand within the US job market for foreign language acquisition has significantly increased. It would be reasonable to assume that the demand for foreign language has risen over the last nine years. However, it is also true that companies have begun to use career websites and the Internet as a whole more often since 2005, and thus all of these effects must be taken into account. A replication and extension of previous research is necessary in order to verify if the demand for foreign language abilities holds today.

Data

The main data source for this research is the "Baccalaureate and Beyond Longitudinal Study" (B&B) from the National Center for Education Statistics. This data keeps track of the experiences of a nationally representative cohort of college graduates who received their bachelor's degrees during the 2007-2008 academic year. It contains information on the

characteristics of the universities attended,[3] demographics, parents' background, earnings and job characteristics, academic major, and courses taken in college and after graduation. This information is based on students' undergraduate transcripts as well as answers to questions posed to the individuals in surveys carried out in 2009 as well as in 2012. The survey include information on the sample that responded to interviews in 2009, which consisted of 17,172 individuals, while the number of respondents in 2012 decreased to 14,397. Of the initial sample, 14,293 of the respondents spoke English at home while growing up.

Respondents were asked in 2009 whether they spoke a foreign language. This question is useful as it is intended to capture knowledge of a second language among native English speakers, which makes this data set distinct from others. Specifically, they were asked if they were at least as fluent in a foreign language in reading, speaking, and writing as their English. Following an affirmative response, the next question was "What second language do you know best?" However, one of the shortfalls of the dataset is that B&B underestimates the actual effect of language fluency in its wording. The specificity of the question is one of the key reasons for this underestimation: the stringent requirement does not acknowledge that a working knowledge of the language may suffice in the workplace, and also discourages a positive response based on how strict it may sound to a respondent.[4] The effect that is being estimated is thus an average over all possible fluency levels. Furthermore, the data fails to address the level of fluency directly through the questions on survey. This causes measurement error as individuals are self-reporting their

[3] The specific institution a respondent attended is not reported.
[4] The B&B survey itself acknowledges this difficulty with the question, noting, "Because of the wording of the items used to construct this variable, interpretation may not be comparable for native English and non-native English speakers" (BB:2012)

language ability, thus providing their own standards for fluency. Ultimately, we are left without a method of measuring different aspects of language fluency.

Regardless, this data proves to be useful because of a specific question about second-language geared towards native English speakers. This question allows isolation of the effect of foreign language acquisition specifically. Additional variables are also helpful. Several variables can be considered cognitive ability proxies, such as parental education levels, SAT or ACT scores, and college grade point average (GPA), which can allay some concerns about language as the lone signal of cognitive ability. Information is also provided on whether the respondent spoke English at home while growing up,[5] if the respondent is a US citizen, and whether the respondent has a parent born in the US. These variables will be used to ideally control for individuals who are learning a foreign language while their native language is English. B&B contains other information on academics, job careers, and earnings as well. Since B&B only captures data for college graduates, we must be careful in generalizing these results to the general population. However, at least this group is homogenous in important ways, so the premium is that earned above and beyond the well-documented premium for a BA.[6] Also a concern of Saiz and Zoido, it is difficult to know how these effects may differ for those who studied foreign language outside of college, for example, or those who did not attend postsecondary education at all.

I include complementary information to the B&B study. Per capita income in the respondent's state of residence in 2013 can be found on the website for the Bureau of

[5] Data is also available for those who speak more than one language; however, this did not include information on whether this language was spoken at home, or which languages they were, and thus I did not find this data relevant for the discussion.

[6] The New York Times reported in 2012 that the mean annual salary for college graduates is $45,500, which is $17,750 higher than those working only with a high school diploma.

Economic Analysis. Their state of residence is assumed to be their state of residence in 2009, about a year after graduation, for lack of the most current information.

I begin my analysis with three tables of descriptive statistics: one for my full sample, one for the subsample of respondents who spoke English growing up at home, and one for the subsample of respondents who spoke English growing up at home, who are US citizens, and have at least one parent born in the US. In explaining the provided variables, I note that log wages in 2012 are derived from a reported annual salary.[7] "Experience" is calculated by subtracting 19 from the respondent's 2012 age, representing the maximum possible years of working experience a respondent may have.[8] "GPA" measures the grade point average on a 4-point scale during the respondent's third year at the institution. This is used over the fourth year simply to incorporate more observations, but admittedly is not recorded as close to graduation. "Parent's education" refers to if one or both parents received at least a high school education.[9] I use SAT composite scores as opposed to both ACT and SAT scores in order to avoid collinearity. Even if a respondent only took the ACT, their scores were converted to the appropriate SAT score employing the College Board's "Correspondence Between ACT and SAT I Scores" (1999) so that every respondent in the sample is measured on the same scale. "Post BA Degree" is a dummy variable that takes on the value 1 if the respondent has obtained a degree higher than a bachelor's since graduation by 2012. "Enrolled" is a dummy variable that takes on the value 1 if a respondent is enrolled in a graduate degree program to any extent in 2012, whether it is full-time or part-time. Finally,

[7] Respondents were actually allowed to report income in any form they liked, which includes hourly wage, weekly wage, monthly wage, and annual salary. This number was then converted into reported annual salary. I divide reported salary by total number of hours worked in the year, multiplying reported weekly hours by 50 weeks. I choose hourly wage for the sake of consistency with past research.
[8] I make the assumption that the minimum experience is 3 years for all respondents who are 21 or younger.
[9] This is a dummy variable that takes on 1 if a respondent has at least one parent with at least a high school education and 0 if the respondent has no parent with as high of an educational background.

"Unemployed" is a dummy variable that takes on the value 1 if the respondent is unemployed or out of the labor force in 2012, and 0 if the respondent is working part-time, full time, or multiple jobs.

Table 1a. Descriptive Statistics, Baccalaureate & Beyond

2012 Survey, Complete Cohort

Variable	N	Mean	Std. Dev.	Min.	Max.
Log hourly wage	14566	2.723	0.93	0	6.4
Fluent in a Foreign Language	14669	0.064	0.244	0	1
Age	16117	24.779	6.545	18	73
Experience	16117	6.094	6.374	3	54
Black	16117	0.11	0.313	0	1
Public	16117	0.517	0.5	0	1
Male	16117	0.414	0.493	0	1
Post BA degree	17172	0.246	0.43	0	1
Married	14566	0.412	0.492	0	1
Enrolled	14566	0.196	0.397	0	1
Employed	14566	0.769	0.422	0	1
SAT	13304	1084.802	184.363	400	1600
GPA	13231	3.275	0.569	0	4

1. All negative log wages (17 respondents) were converted to 0
2. Evident from the sample statistics, there are some outlying respondent ages (2 respondents are in their 70s)
3. All observations with nonresponses are dropped

Table 1b. Descriptive Statistics, Baccalaureate & Beyond

2012 Survey, Subsample of Respondents Who Spoke English Growing Up

Variable	N	Mean	Std. Dev.	Min.	Max.
Log hourly wage	12979	2.758	0.883	0	6.4
Fluent in a Foreign Language	12981	0.021	0.145	0	1
Age	14293	24.798	6.613	18	73
Experience	14293	6.108	6.447	3	54
Black	14293	0.112	0.315	0	1
Public	14293	0.515	0.499	0	1
Male	14293	0.411	0.492	0	1
Post BA degree	14293	0.26	0.438	0	1
Married	12979	0.419	0.493	0	1
Enrolled	12979	0.19	0.392	0	1
Employed	12979	0.786	0.41	0	1
SAT	11937	1087.918	181.377	400	1600
GPA	13231	3.283	0.571	0	4

Table 1c. Descriptive Statistics, Baccalaureate & Beyond

2012 Survey, Subsample of Respondents Who Spoke English Growing Up, Are US Citizens, and Had at Least One Parent Born in the US

Variable	N	Mean	Std. Dev.	Min.	Max.
Log hourly wage	10422	2.773	0.857	0	6.4
Fluent in a Foreign Language	10413	0.018	0.135	0	1
Age	11467	24.819	6.598	18	71
Experience	11467	6.118	6.437	3	52
Black	11467	0.104	0.305	0	1
Public	11467	0.518	0.5	0	1
Male	11467	0.409	0.492	0	1
Post BA degree	11467	0.258	0.437	0	1
Married	10422	0.427	0.495	0	1
Enrolled	10422	0.188	0.391	0	1
Employed	10422	0.794	0.404	0	1
SAT	9593	1089.019	179.698	400	1600
GPA	8811	3.288	0.57	0	4

As I limit the sample further into subsamples, I have a better understanding of which variables reduce the number of observations for regressions, with the lowest number of observations reported for SAT, GPA, and wage. We notice slightly higher means for all three of these variables as we limit the sample, but at a glance, the descriptive statistics do not seem to change dramatically from one table to the next. It may be due to this lack of difference in descriptive statistics between limitations of the sample that we will also find the analysis on the relationship between foreign language fluency and wages to not show drastic differences between subsamples.

Part of the goal of this paper is to extend research that was previously conducted on a cohort and their 1997 wages. In order to fulfill that goal, I run regressions using an equation that is as similar as possible to that used by Saiz and Zoido in 2005 for my more recent cohort. This cohort graduated with bachelor's degrees between 2007-2008, and was interviewed in 2009 and 2012. One of the main differences is that in the 2005 paper, the Baccalaureate and Beyond study lists both language data and wages for 1997, the year the dependent variable is being analyzed. However, for my cohort, I am only able to gather if the respondent reported they were fluent in 2009, while my wages are from 2012. Respondents who were fluent in 2009 may not have maintained their language fluency for the next three years; on the other hand, other respondents may have improved their second language skills in order to become fluent. My paper assumes that fluency or non-fluency in 2009 also means fluency or non-fluency in 2012.

The Returns to Speaking a Second Language

I begin my analysis by presenting my results from an equation that is a modified from the version used by Saiz and Zoido. All regressions only include respondents who have non-missing data for all listed variables. For the sake of consistency, I attempted to replicate the regression used by Saiz and Zoido as much as possible, but I do not include these regression results or discussion, as the results do not change significantly from my own regressions.[10]

In addition to the aforementioned variables, I add a variable that describe characteristics of where the respondent lived. "Log state income 2013" refers to the per-capita income in the state of residence of the respondent. Due to the lack of data, I assume that the state of residence for the respondent in 2012 is the same as 2009, and that the per-capita income has not changed drastically between 2012 and 2013. I include dummy variables for a number of majors. I also include a continuous variable representing the number of years of foreign language was taken in high school by the respondent.[11] Finally, I include dummy variables for unemployment and enrollment in a graduate program.

According to the model, our baseline is a respondent who does not speak a foreign language in 2009, is not male identifying, is not currently married, did not graduate from a public institution, is not black, does not hold any degree beyond a BA, is a science and

[10] A discussion of this regression equation and results, as well as summary statistics for the cohort that graduated in 1992-1993, can be found in the appendix.

[11] Though the variable is continuous, the variables takes 1 if the respondent has studied for 1 to 1.5 years of foreign language, 2 for 2 to 2.5 years, 3 for to 3.5 years, and 4 for 4 or more years.

technology major,[12] is employed part-time, full-time, or with multiple jobs, and is not currently enrolled in a graduate school program.

I use this equation with three samples: (1) for all respondents, (2) for all respondents who spoke English growing up at home, and (3) for all respondents who spoke English growing up at home, are US citizens, and have at least one parent born in the US.

Table 2. New Regression Results

Dependent Variable: Log Wages in 2012

Variable	All Respondents	Spoke English at Home	Spoke English at Home, US Citizen, Parent(s) Born in US
Fluent in a Foreign Language in 2009	**-0.08***	0.003	0.02
	(0.047)	(0.073)	(0.081)
Experience	0.008	0.006	-0.004
	(0.039)	(0.039)	(0.041)
Experience^2	0.0001	0.001	0.002
	(0.003)	(0.004)	(0.004)
Male	0.002	0.005	-0.002
	(0.02)	(0.02)	(0.022)
Married	**0.058*** **	**0.07*** **	**0.066*** **
	(0.02)	(0.02)	(0.021)
Black	**-0.1*** **	**-0.081** **	**-0.095*** **
	(0.033)	(0.033)	(0.021)
Log state income 2013	**0.342*** **	**0.41*** **	**0.409*** **
	(0.07)	(0.071)	(0.075)
Public	0.021	**0.033***	0.03
	(0.019)	(0.02)	(0.021)
GPA	0.017	0.025	0.027
	(0.018)	(0.019)	(0.019)
Parents' education	0.002	0.008	0.019
	(0.046)	(0.047)	(0.05)
Major: Computer Science	**0.314*** **	**0.312*** **	**0.293*** **
	(0.054)	(0.056)	(0.06)

[12] Four respondents claim to be "Undeclared" at graduation and are also part of the baseline if they have non-missing data.

Major: Engineering	**0.458*** **	**0.438*** **	**0.43*** **
	(0.037)	(0.038)	(0.04)
Major: General Studies	0.089	0.079	0.019
	(0.067)	(0.067)	(0.069)
Major: Social Science	**0.076** **	**0.06***	0.032
	(0.033)	(0.033)	(0.035)
Major: Humanities	-0.012	-0.031	**-0.073** **
	(0.035)	(0.035)	(0.037)
Major: Health	**0.334*** **	**0.319*** **	**0.278*** **
	(0.044)	(0.044)	(0.046)
Major: Business	**0.208*** **	**0.204*** **	**0.172*** **
	(0.034)	(0.035)	(0.037)
Major: Education	-0.019	-0.031	-0.049
	(0.037)	(0.037)	(0.039)
Major: Other Applied	0.022	0.012	-0.007
	(0.033)	(0.033)	(0.035)
Post BA degree	0.008	0.012	0.03
	(0.021)	(0.021)	(0.022)
SAT	**0.0004*** **	**0.0004*** **	**0.0004*** **
	(0.00006)	(0.00006)	(0.00006)
Years of foreign language in high school	-0.01	-0.01	-0.013
	(0.009)	(0.008)	(0.009)
Enrolled	**-0.078*** **	**-0.081*** **	**-0.074*** **
	(0.024)	(0.025)	(0.026)
Unemployed	**-0.866*** **	**-0.803*** **	**-0.756*** **
	(0.023)	(0.024)	(0.026)
N	7085	6411	5502
R-Squared	0.2434	0.2295	0.2212

1. It is important to note that there is not much variation in experience due to the relatively young sample
2. ***Statistically significant at the 1% level, ** at the 5% level, * at the 10% level
3. Standard errors are reported in parentheses under the t-statistics

I find a significant relationship between enrollment, unemployment,[13] SAT scores, certain majors, the per capita income of the state, and certain demographics each with wage. However, aside from a significant negative relationship within the full sample, I find

[13] I also run these regressions on selected samples for those who are employed (in place of using a dummy variable for unemployment). I find no change in the results for statistical significance on my variable for foreign language fluency, and thus I do not report those results in this paper.

there is no relationship between foreign language and wages for my subsamples, which are limited to closely represent the respondents I am most interested in: those who spoke English growing up and acquired a foreign language through study.

While a general effect could not be found for foreign language acquisition, I decided to examine if there were any economic benefits to studying a relatively difficult language. I divide the reported languages according to categories created by the Foreign Service Institute (FSI), with languages in each category taking a certain amount of time to master if the speaker's native language is English. The higher the category, the longer it takes for an English speaker to attain reading and speaking proficiency in that language. It must be noted that this categorization is subjective; while Japanese is labeled as the most difficult language for English speakers, one may disagree and make arguments for Mandarin Chinese or Arabic as the most difficult language due to a focus on characters or a script highly differentiated from English letters. Similarly, the relative difficulty of lower category languages is up to debate; German is considered similar to English, but is farther away than Spanish or French and sits in its own category.

According to the FSI, Category I languages take 575-600 hours of study for proficiency; Category II languages take 750 hours, Category III languages take 900 hours, Category IV languages take 1100 hours, and Category V languages take 2200 hours.

Table 3. Language Categories

Category Name	Language	English Native Speakers with Second Best Language as of 2009	Spoke English Growing Up, US Citizen, At Least One Parent Born in US with Second Best Language as of 2009
Category I	French or Canadian French	1349	1093
	Italian	192	150
	Portuguese	71	46
	Spanish	6370	5052
Category II	German	563	466
Category III	Malay	2	0
Category IV	Bengali	8	2
	Greek (ancient/modern)	40	25
	Hebrew (Biblical/modern)	60	47
	Hindi	25	2
	Russian	74	51
	Urdu	26	1
	Vietnamese	45	1
Category V	Arabic	46	18
	Chinese	160	58
	Japanese	122	96
	Korean	69	26

1. Original categories with a wider selection of languages can be found at http://www.effectivelanguagelearning.com/language-guide/language-difficulty
2. Column 3 lists "Respondents whose native language is English, who took classes in or knew another language as of 2009, and had a second best language as of 2009."
3. Column 4 limits Column 3 further by imposing a requirement of at least one parent born in the US and being a US citizen
4. There is a difference in wording between Column 3 and Column 4: in Column 3, we have native English speakers, but in Column 4 we have those speaking English growing up; I leave this difference simply because Column 3 provides the raw data from the codebook, and Column 4 represents the second subsample included in our regressions. I do not believe the wording drastically changes the type of respondent I am looking at.

The FSI lists Category IV to carry the highest number of languages; however, most likely because these are languages not commonly taught in school, the B&B survey does not provide many of those languages as a possible answer to the question. Also, while some

respondents indicated that they spoke English while growing up, but also reported a language in Category IV as their second best language, I assume the majority of these languages are not acquired in school; rather, I assume that most of these languages are in fact learned at home from immigrant families and spoken alongside English while being raised. Thus, any significance on the relationship between Category IV and wages may be imprecise and may not represent acquired foreign language proficiency.

For the following regressions, I take a look at economic benefits for a respondent who speaks a foreign language, with that language falling into a certain category. I combine Categories III and IV as there are not enough observations for Category III by itself. It should be noted that "Other" was an answer reported by respondents, but since I could not know which category those languages would fall under, I choose to leave them out.

Table 4. Interactions with Language Categories

Dependent Variable: Log Hourly Wage in 2012

Language Category X Foreign Language	All Respondents	Spoke English at Home	Spoke English at Home, US Citizen, Parent(s) Born in US
Category I	0.07	0.169	0.213
	(0.099)	(0.158)	(0.19)
Category II	0.438	0.354	0.305
	(0.448)	(0.435)	(0.425)
Category III or IV	0.143	-0.16	0.283
	(0.451)	(0.532)	(0.729)
Category V	0.554 (!)	0.182	0.151
	(0.351)	(0.531)	(0.52)

1. (!) represents significance at the 20%
2. Standard errors are reported in parentheses

For all of the language categories, the coefficient for foreign language fluency (unreported) is statistically insignificant, and so I only focus on the interpretation of the

coefficients on the interaction terms between foreign language fluency and having that foreign language fall under a certain category.

I find no statistical significance between language difficulty and wages. This may be a reflection of the lower demand for foreign language abilities in general by employers in the United States. However, one warning about these regressions is that these coefficients do not capture whether or not the respondent uses the language at work, but simply that the second language acquired falls under that category. Thus, due to a lack of data, we are unable to flesh out this relationship further.

As a contrast to previous literature, in 2005 Saiz and Zoido observe lower wage premiums for Spanish speakers than German or French speakers; since there is a much higher supply of native speakers in the United States, those who acquire Spanish simply face more competition for the same jobs (Saiz 531). While this would have led me to believe that Category I would have a significantly negative relationship or a positive relationship that is lower than other categories with wages, I find that this is not true. Also, I find that while Category V almost provides evidence for a wage premium, there is no consistent relationship that appears as languages become more difficult (despite the majority of the coefficients being positive for all categories). This may be due to grouping languages into categories rather than looking at them individually. In this case, a stronger relationship may be found if enough observations were available to do an analysis for individual languages. Statistical insignificance may also reflect a general change in the demand for foreign language ability since the previous analysis was done in 2005.

The next step in my regression analysis is to understand if there are benefits for acquiring a second language while working in specific sectors. This is an exercise also

carried out by Saiz and Zoido, but I take their research a step further by using my modified model and taking a look at subsamples. I divide reported careers in 2012 into 7 categories and interact them with foreign language fluency.

Table 5. Interactions With Occupation Variables

Dependent Variable: Log Hourly Wage in 2012

Occupation X Foreign Language	All Respondents	Spoke English at Home	Spoke English at Home, US Citizen, Parent(s) Born in US
Personal Services	0.173	**0.502****	**0.518****
	(0.124)	(0.202)	(0.212)
Science & Technology	**-0.344****	**-0.571*****	**-0.682*****
	(0.139)	(0.191)	(0.218)
Sales	**0.621*****	**0.461***	0.364
	(0.217)	(0.265)	(0.291)
Managerial	-0.09	-0.137	0.186
	(0.164)	(0.278)	(0.31)
Educators	0.084	0.073	0.076
	(0.138)	(0.213)	(0.224)
Business Services	0.081	0.185	0.055
	(0.144)	(0.241)	(0.309)
Other	**-0.281***	**-0.542***	**-0.504***
	(0.042)	(0.316)	(0.336)

1. ***Statistically significant at the 1% level, ** at the 5% level, * at the 10% level
2. Standard errors are reported in parentheses

As with the language category analysis, the unreported coefficient for foreign language fluency is statistically insignificant throughout the regressions including occupation. Thus, I only focus on the interpretation of the coefficients on the interaction terms between foreign language fluency and working in a certain occupational sector.

The strongest effect that can be found is within science and technology and "other"[14] fields; there appears to be a negative relationship between foreign language fluency on

[14] These are primarily jobs related to agricultural, according to the descriptions found in the B&B Codebook.

wages in these jobs. This may possibly be due to the opportunity costs of learning a language in such fields, where acquiring other skills related to the occupation would provide a higher premium in wages. For science and technology, these may mean economic benefits to taking more math and statistics classes; for working in agriculture, there may not be so much of a benefit for learning foreign languages over simply mastering the skills relevant to your occupation. However, respondents in personal service and sales positions who are also fluent in a second language experience a positive increase in wages. Respondents working in these fields would find that being able to communicate with different people is an important skill, which may include interactions with international customers or clients with non-English first languages. In an increasingly globalized world, these occupations especially may be experiencing a significant increase in demand for foreign language abilities.

Why might there be such a strong contrast from 15 years ago? The use of English as a global language has increased dramatically over the last decade, perhaps lowering the need by employers for native speakers of English to learn foreign languages. Partly because of this, there may be fundamental differences in how this cohort approached learning foreign languages compared to the original cohort. The US has also been experiencing an increase in the number of students studying from abroad, which may be increasing the supply of potential employees with bilingual and multilingual abilities. Since 2000, the overall number of international students to the US has increased by 72% according to the 2014 Open Doors Report on International Exchange (Haynie 2014). This increase may be increasing the competition between native English speakers learning a foreign language and fluent speakers of non-English languages who would be required to have a highly

English ability in order to study in the United States. All of these factors may be impacting the relationship we see between foreign language proficiency in native English speakers and their earnings. However, respondents in this cohort also graduated into an economic recession, and the wages recorded are most likely for entry-level positions that do not completely take advantage of their specialized skills, which means over time a more statistically significant relationship between foreign language ability and wages may be found.

Conclusions

My findings indicate that for US college graduates who spoke English growing up at home, fluency in a foreign language does not have a significant impact on wages. This differs from Saiz and Zoido's research in 2005, where they find that speaking a foreign language does have a positive impact in the US labor market when it comes to the college graduates in the sample. When regressing the relative difficulty of the language on wages, I still do not find a significant relationship with a wage premium. Finally, I find a positive relationship between foreign language while also working in personal services and sales positions. I also find a negative relationship between foreign language and wages while working in science and technology and agricultural occupations.

These results must be taken with caution. There is a concern for validity in that there is not more language information over time, so assumptions have to be made about whether or not fluency in a second language was maintained since graduation. Even the information that is collected on foreign language is subjective. If respondents do not feel comfortable stating their language abilities as "fluent", the data underestimates the effect foreign language fluency may have on wages. A similar overestimation exists when thinking about

those who may be overselling their language ability. The sample was homogeneous in that all respondents were college graduates with a bachelor's degree, but this in itself means that economic benefits for those without a bachelor's degree are simply not recorded. This effect may be strong in jobs that do not require a college degree, such as a contractor who does not necessarily have a degree but uses a foreign language to speak with his employees. Also, it should be emphasized that when taking a look at wages for recent college graduates, not only are the respondents in my cohort graduating into an economic recession, but also many are enrolled in a graduate program in 2012, and even those who are employed are likely to only be working in entry-level positions. Thus, the wages may not accurately reflect the average earnings the respondents will earn while utilizing their foreign language abilities.

In looking forward, I would like the opportunity to look at more precise data over time in order to verify the economic impact of foreign language acquisition. Further research may continue to investigate specific occupations or specific language categories, perhaps with categories that differ from the ones used here. It would be especially helpful to be able to look at wages that are farther out from a respondent's graduation date, as well as at more frequent intervals. I primarily look at 2012 wages, which are only a few years out from the graduation of the respondents. More information on wages over time would provide a more accurate look at the fluctuation of wages and whether or not any wage premium foreign language acquisition may provide remains consistent. If possible, it would be even more helpful to take a look at continuous and objective measures of language fluency that are tested rather than self-reported. Finally, it must be noted that benefits that cannot be captured through wages absolutely exist for studying foreign language. As the

world continues to change and globalization continues to spread, it is important that better data is captured in order to verify the true effects of foreign language acquisition on earnings.

References

Baccalaureate and Beyond (BB:2012). Codebook. U.S. Department of Education, National Center for Education Statistics. Washington, DC. Web. 1 April 2015.

Card, David. "The Causal Effect of Education on Earnings." *Handbook of labor economics* 3 (1999): 1801-63. Print.

Chiswick, Barry R. "Hebrew Language Usage: Determinants and Effects on Earnings among Immigrants in Israel." *Journal of Population Economics* 11.2 (1998): 253-71. Print.

Chiswick, Barry R. "The Earnings of White and Coloured Male Immigrants in Britain." *Economica* 47.185 (1980): 81-7. Print.

Dewan, Shaila. "Wage Premium from College is Said to be Up." *The New York Times*Feb 11, 2014 2014. Print.

Dorans, Neil J. *Correspondences between ACT and SAT I Scores*. New York: College Entrance Examination Board, 1999. 1999. Web. 5 Feb. 2015.

Dustmann, Christian, and Francesca Fabbri. "Language Proficiency and Labour Market Performance of Immigrants in the UK*." *The Economic Journal* 113.489 (2003): 695-717. Print.

Foreign Service Institute. "Effective Language Learning." *Effective Language Learning RSS*. Foreign Service Institute, n.d. Web. 09 Dec. 2014.

Hanushek, Eric A., and Ludger Wößmann. "The Role of Education Quality for Economic Growth." (2007)Print.

Haynie, Devon. "Number of International College Students Continues to Climb." US News. U.S.News & World Report, 17 Nov. 2014. Web. 09 May 2015.

Hoxby, Caroline M., and Bridget Terry. *Explaining rising income and wage inequality among the college educated* (1999)Print.

"Regional Economic Accounts." *BEA*. N.p., n.d. Web. 19 Dec. 2014.

Mincer, Jacob. "Schooling, Experience, and Earnings. Human Behavior & Social Institutions no. 2." (1974)Print.

Pascarella, Ernest T., Patrick T. Terenzini, and Kenneth A. Feldman. *How College Affects Students*. 2 Vol. Jossey-Bass San Francisco, CA, 2005. Print.

R.L.G. "Johnson: What Is a Foreign Language Worth?" *The Economist*. The Economist Newspaper, 11 Mar. 2014. Web. 09 Dec. 2014.

Saiz, Albert, and Elena Zoido. "Listening to what the World Says: Bilingualism and Earnings in the United States." *Review of Economics and Statistics* 87.3 (2005): 523-38. Print.

"State Requirements for High School Graduation, in Carnegie Units: 2004." *National Center for Education Statistics*. Institute of Education Sciences, n.d. Web. 19 Dec. 2014.

Wooldridge, Jeffrey M. *Introductory Econometrics: A Modern Approach*. Australia: South-Western College Pub., 2003. Print.

Appendix

Comparison to 1992-1993 Cohort and Replicating the 2005 Regression

In order to take a look at the cohort used in previous research by Saiz and Zoido, I include summary statistics describing the cohort that graduated in 1992-1993 that is used in their 2005 paper. While this is not the same cohort that I analyze, I believe the requirements of being a respondent for the survey allows me to generalize (with caution) my own results in a similar manner for US college graduates.

Table A1. Descriptive Statistics, Baccalaureate & Beyond, 1997 Survey

Variable	Mean	Std. Dev.	Min.	Max.
Log hourly wage	2.55	0.42	1.03	3.73
Speaks FL	0.34	0.474	0	1
Age	30.102	6.543	0	73
Experience	8.102	6.543	0	51
Married	0.305	0.46	0	1
Black	0.04	0.196	0	1
Male	0.448	0.497	0	1
Normalized GPA	3.063	0.498	1.17	1
Public College	0.677	0.467	0	1
MBA	0.028	0.167	0	1
PhD	0.029	0.168	0	1
Other master's	0.256	0.436	0	1
Observations: 7940				

Source: Saiz, Albert, and Elena Zoido, "Listening to What the World Says: Bilingualism and Earnings in the United States" (2005)
Note: According to Saiz and Zoido, "This table refers to the subsample of individuals who have hourly wages above \$2.8 (1st percentile) and below \$42.30 (99th percentile), who answer the question on whether they speak a foreign language and with complete data on age, experience, gender, marital status, race, state of residence, college GPA, and type of college attended. Also, experience = max(0, age-22) (528). Wage and postgraduate degrees are measured differently than in my own regressions.

I now include the results of a regression in which I replicate as much as possible the regression equation used by Saiz and Zoido. Variables that include "competitive" in the names are dummy variables that correspond to Saiz and Zoido's measures of college

quality based on a respondent's GPA, SAT scores, and ACT scores, indicators originally compiled by Hoxby and Long (1999). Since the specific institution that a respondent attended is not available, consistent with Saiz's and Zoido's definition, it is assumed that the higher a student's GPA, SAT scores, and ACT scores, the higher the quality of the institution. The table used to create the dummy variables associated with college quality can be found in an appendix of their 2005 paper, and I also replicate the table here. I do not believe these dummy variables to be enough to stand in as a measure of the institution's quality as they are made up of the characteristics of the respondents themselves, and so I do not use them in my own model.

Other changes between my regression and the one used by Saiz and Zoido are that I replace "State Requirements" with "Years of foreign language in high school." The former may be imprecise, as I did not have data on specifically where respondents went to high school. Also, high school foreign language requirements are not necessarily regulated by the state. However, Saiz and Zoido used the former, and so I include it below corresponding to my respondents. These are state high school minimum graduation requirements appropriate to the current cohort from the Digest of Education Statistics, and are matched to the respondent's state of residence upon college graduation in 2007-2008, which is assumed to be the state in which they attended high school.

I find, similar to my own regression for all respondents, there is a negative relationship between foreign language fluency and log wages. However, I only include the regression for all respondents, as the results for subsamples become insignificant, also corresponding to my own results.

Table A2. Quality Grouping Criteria

Category	GPA	SAT	ACT
Most competitive	3.75-4.50	1280-1600	31-36
Highly competitive	3.40-3.74	1120-1279	26-30
Very competitive	3.00-3.39	1000-1119	21-25
Competitive	2.50-2.99	850-999	16-20
Less competitive	2.00-2.49	650-849	12-15
Noncompetitive	1.00-1.99	400-469	4-11

Source: Saiz, Albert, and Elena Zoido, "Listening to What the World Says: Bilingualism and Earnings in the United States" (2005)

Table A3. Replication of Regression from Saiz and Zoido (2005)

Dependent Variable: Log Hourly Wage in 2012

Variable	All Respondents
Fluent in a Foreign Language in 2009	**-0.22*** **
	(0.048)
Experience	0.008
	(0.04)
Experience2	0.0004
	(0.004)
Male	-0.031
	(0.02)
Married	**0.097*** **
	(0.02)
Black	**-0.226*** **
	(0.034)
Log state income 2013	**0.345*** **
	(0.074)
Public	**0.048** **
	(0.02)
GPA	-0.01
	(0.019)
Parents' education	-0.065
	(0.041)
Post BA degree	0.032
	(0.021)
SAT	**0.0004**
	(0.00006)
Major: Computer Science	**0.459*** **
	(0.054)
Major: Engineering	**0.585*** **

	(0.037)
Major: General Studies	0.08
	(0.068)
Major: Social Science	**0.075****
	(0.032)
Major: Humanities	-0.009
	(0.034)
Major: Business	**0.308*****
	(0.034)
Major: Education	0.039
	(0.037)
Major: Other Applied	**0.077****
	(0.032)
Years of Required Foreign Language in High School	0.008
	(0.011)
Most Competitive	-0.037
	(0.056)
Highly Competitive	0.027
	(0.046)
Very Competitive	0.07
	(0.045)
Less Competitive	0.173
	(0.188)
Noncompetitive	**-2.441*****
	(0.853)
N	8218
R-Squared	0.0676

1. ***Statistically significant at the 1% level, ** at the 5% level, * at the 10% level
2. Standard errors are reported in parentheses under the t-statistics
3. The baseline is the same as in my own regression, but now the respondent has a major that is undeclared, science and technology, or health; the respondent is also assumed to be attending a "competitive" institution

Acknowledgements

I thank Professor Barbara Craig for doing so much to assist me in obtaining the data for this research. I thank Professor Ron Cheung for sitting down with me to give detailed feedback on my drafts numerous times and encouraging me to take up this project in the first place. I thank the above as well as Professors Tobias Pfutze, Ellis Tallman, Edward McKelvey, Hirschel Kasper, Viplav Saini, and Martin Saveedra for all of their feedback on this paper. I thank Terri Pleska for everything she has done for me as an economics major. I thank my fellow economics honors students during the 2014-2015 school year for their support. I thank Deborah Serianni, Oberlin College Class of '83 who was a former economics honors student as well, for her mentorship this entire year. I thank Loan Lu for her endless encouragement throughout my time at Oberlin College and especially through this research. Finally, I thank all of the friends, peers, and mentors that have given me feedback on my ideas for this research, supported me to keep moving forward, as well as listened to my presentations. Without you all this would not be possible.