

English Proficiency and Labor Market Performance: Evidence from the Economics Profession

William W. Olney*

April 22, 2016

Abstract

This paper investigates whether the global spread of the English language provides an inherent advantage to native English speakers. This question is studied within the context of the economics profession, where the impact of being a native English speaker on future publishing success is examined. English speakers may have an advantage since they are writing in their native language, the quality of writing is a crucial determinant of publishing success, and all the top economics journals are published in English. Using a ranking of the world's top 2.5% of economists, this paper confirms that native English speakers are ranked 100 spots higher (better) than similar non-native English speakers. A variety of extensions examine and dispel many other potential explanations.

Keywords: English Proficiency; Labor Market Performance; Economics Profession

JEL Codes: A11; J44; F66

*Department of Economics, Williams College, Williamstown, MA 01267 (email: william.w.olney@williams.edu). I am grateful to Jon Bakija, Dan Friedman, Dan Hamermesh, Paul Oyer, Romain Wacziarg, Christian Zimmermann, and numerous seminar participants for helpful comments and suggestions.

1 Introduction

The spread of the English language has been profound. It is used throughout the business world, intergovernmental agencies, and academia and there is really no viable alternative.¹ While the diffusion of the English language has been remarkable, the implications of these changes are not well understood.² This paper investigates whether the spread of English provides an inherent advantage to native English speakers. This question is studied within the context of the economics profession, where specifically the impact of being a native English speaker on future publishing success is examined.

With the rapid pace of globalization, it is increasingly useful to have a common language of discourse. For better or worse, English serves this role. Multinationals throughout the world now use English as their official language, including numerous firms from non-native English speaking countries. For instance, Lufthansa a German airline company, Lenovo a Chinese computer company, Audi a German car maker, Rakuten a Japanese technology company, and Aventis a French/German pharmaceutical firm all use English. *The Economist* writes that "native English speakers often assume that the spread of their language in global corporate life confers an automatic advantage on them."³ Whether this is in fact true remains an open question.

English proficiency is also crucial in academic disciplines where most top journals are published in English. For instance, in 1995 87.2% of all publications in physical sciences and 82.5% of all publications in social sciences were in English and these shares are growing (Ammon 2003). More specifically, all of the top economics journals are published in English (see Engemann and Wall 2009 and IDEAS/RePEc rankings).⁴ Since tenure and compensation often hinge on research output, the need to publish articles in English speaking journals is a crucial determinant of future success in academic disciplines.

Despite the importance of English, it is not an easy language to master. Even native English

¹By some measures Mandarin is more common, due to the sheer size of China. However, it is one of the most difficult to master, least computer friendly languages, and is not even universal within China ("The English Empire" *The Economist*. February 15th, 2014).

²For instance, David Crystal in "English as a Global Language" (2003) documents the impressive spread of English around the globe since 1950. For instance, he says "there has never been a language so widely spread or spoken by so many people as English" but he only speculates at some of the implications of this growth.

³"The English Empire" *The Economist*. February 15th, 2014.

⁴Some journals will accept an article in a language other than English (for instance the Canadian Journal of Economics accepts articles in French) but these journals ultimately publish the article in English.

speakers can have difficulty articulating complex ideas in their first language. For non-native speakers, it is that much more difficult.⁵ The most fluent non-native English speakers can still struggle to write technical papers in a second language and presumably this is a more time consuming process. Thus, English proficiency can affect both the quality and quantity of research.

Abundant anecdotal evidence indicates that the quality of writing is crucially important in the publishing process and that native English speakers have an advantage in this endeavour. For instance, Robert Moffitt (former editor of the *American Economic Review*) said “I should also note that non-native-English speakers should work hard to get the English right and, if necessary, hire native English speakers to edit their papers. It is no doubt unfair, but editors and referees often take poor English as a signal of low quality.”⁶ In a similar vein, Lawrence Katz (editor of *The Quarterly Journal of Economics*) advises that “Papers need to be well-written and self-contained. A paper will get desk rejected for sure if it is sloppy regardless of the quality of the ideas.”⁷ Patricia Anderson (former editor of *Journal of Human Resources*) admits that “As an editor, when I get a paper that is riddled with typos, I can’t help but have that color my view of the overall enterprise.”⁸ Daniel Hamermesh (former editor of *Journal of Population Economics*) recommends that if you want to publish your paper in a top economics journal then “get a native English speaker to read it *carefully* for you.”⁹

To summarize, three important characteristics of the economics profession are especially relevant for this analysis. First, all the top economics journals are published in English. Second, publishing success within economics depends crucially on the quality of the writing. No matter how groundbreaking the idea or how sophisticated the analysis, a poorly written paper will be rejected. Third, as these editors allude to, English speakers tend to have an advantage at writing papers in their native language. The goal of this paper is to examine whether there is empirical support for this basic intuition that native English speakers have an inherent advantage in the publishing process.¹⁰

⁵ “The English Empire” *The Economist*. February 15th, 2014.

⁶ “How to Get Published in an Economics Journal.” CSWEP Newsletter, Spring 2011.

⁷ “Q&A with Larry Katz, editor of QJE.” Berk Ozler and David McKenzie Blog. World Bank. January 4, 2012.

⁸ “How to Get Published in an Economics Journal.” CSWEP Newsletter, Spring 2011.

⁹ “How to Publish in a Top Journal (I wish that I knew!).” Daniel S. Hamermesh. University of Texas at Austin presentation.

¹⁰ An alternate and more nuanced story is that the characteristics of the English language itself might be more conducive to economic thought or research. To the best of my knowledge there is no evidence that supports this intuition and in fact existing research indicates the opposite. For instance, Chen (2013) shows that the structure and grammar of the English language adversely affects future economic and health outcomes.

Despite these editor's comments, perhaps, being born and raised in a non-native English speaking country provides benefits that more than compensate for the difficulty of publishing in a second language. For instance, being exposed to different countries and cultures may offer insights and expertise that is useful in economic research. Or maybe, the process of learning a second language stimulates parts of the brain that are useful in future academic pursuits. Alternatively, perhaps the perceived advantages associated with writing in their native language leads native English speakers to be complacent and rest on their laurels.¹¹ For all of these reasons, the overall impact of being a native English speaker is ambiguous and is ultimately an empirical question that this paper hopes to answer.

This analysis utilizes data provided by IDEAS/RePEc on the ranking of economists. Specifically, the top 2.5% of economists worldwide (or 1,082 economists) are identified using data on their quality adjusted number of publications and on their quality adjusted number of citations. Then the biographical information for each of these economists was individually gathered, which was the most challenging and time consuming aspect of this project. Specifically, a variety of sources are used to identify information about the economist including their country of birth, date of birth, gender, date of death (when applicable), undergraduate institution and year of degree, and graduate institution and year of degree. A rigorous data collection process was followed in order to maintain the integrity of the data set.

Focusing on economics is appealing for a couple of reasons. First, there is an objective and publicly available rankings of economists based on research output. It would be much more difficult, for instance, to examine the impact of English proficiency on success in other professions where analogous rankings are unavailable.¹² Second, there are a variety of factors that should, if anything, mitigate the impact of English in the economics profession. For instance, economics is a math and statistical based discipline where the benefits of English proficiency should be less pronounced compared to other more language intensive academic fields. In addition, success in economics is based on written output which is easier to improve relative to spoken English which is often needed in the business world. For example, non-native English speaking economists can hire proof-readers

¹¹"The English Empire" The Economist. February 15th, 2014.

¹²For instance, as Fourcade et al. say (2015): "in no other social science can one find the extraordinary volume of data and research about rankings (of journals, departments, and individuals) that economists produce."

to address writing and grammatical issues.¹³ Finally, coauthoring is common in economics, so to the extent that native and non-native English speakers collaborate, the disadvantage of being a non-native English speaker will be mitigated.¹⁴ For all of these reasons, English proficiency should have less of an effect in economics relative to other fields and professions. Thus, the results of this paper should be considered a lower bound on the broader effects of English proficiency.

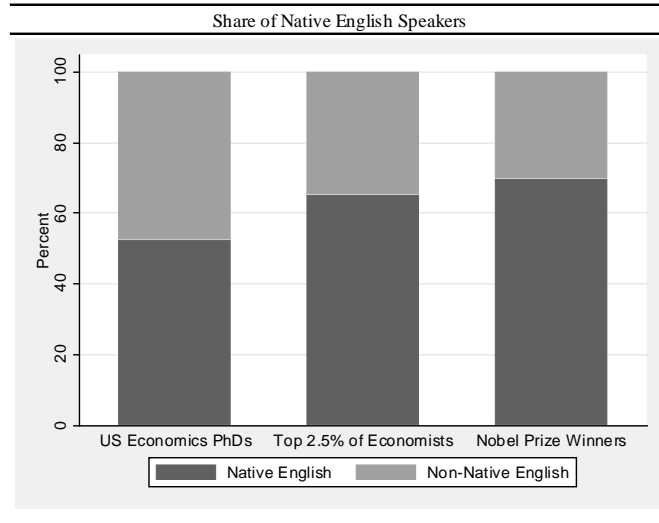
A preliminary look at the data suggests that native English speakers do in fact have an advantage. Figure 1 shows the share of native English speakers in the economics profession, in the sample of the top 2.5% of economists, and among Nobel Prize winning economists.¹⁵ A comparison between the first two columns indicates that the share of native English speakers in the overall population of economists (52%) is much lower than the share among the top 2.5% of economists (65%). Column 3 then shows that the share of native English speakers among the 76 Nobel Prize winning economists is even higher (70%). Overall, Figure 1 illustrates that the share of native English speakers increases as the sample of economists becomes more talented. While this basic comparison in Figure 1 suggests that native English speakers may in fact have an advantage, a more rigorous analysis is needed that controls for other factors that likely contribute to success in the economics profession.

¹³However, proof readers likely have a harder time helping convey complex economic ideas.

¹⁴Consistent with this story, Freeman and Huang (2015) find that papers coauthored by scientists of different ethnicities tend to be published in better journals and cited more often. Although surprisingly cross-ethnicity coauthorship happens less frequently than expected.

¹⁵Estimating the share of native English speakers in the overall population of economists is difficult, due to a lack of data. Column 1 provides a rough estimate by calculating the share of U.S. citizens who graduated from U.S. economics PhD programs using data from the NSF. This is probably an overestimate since the native English speaking share is likely even lower at programs outside the U.S.

FIGURE 1



The first bar uses data from the NSF on economics PhD graduates from U.S. institutions over the years 1966-2006. Native English speakers are defined as graduates who are U.S. citizens. The second bar uses RePEc/IDEAS rankings from February 2015 and biographical information obtained by the author. The final bar uses data provided by the Nobel Prize website for the years 1969-2015. In these latter two bars, native English speakers are defined as those born in the 6 native English speaking countries.

Focusing more carefully on the top 2.5% of economists, the empirical analysis examines whether an economists' ranking is increasing with English proficiency. Testing this hypothesis faces a couple of important empirical challenges. First, English proficiency is hard to measure. Second, it likely depends on factors such as innate ability to learn languages, education, and other characteristics which could be correlated with research output. To address these concerns, the analysis uses information about whether an economist was born in a native English speaking country (i.e. U.K., Ireland, Australia, New Zealand, Canada, and the U.S.) as a proxy for English proficiency. Country of birth is unaffected by future publishing success which mitigates endogeneity concerns and allows the causal effect of being a native English speaker on research output to be identified. Furthermore, using country of birth to define native English speakers, rather than self-reported language ability, is appealing because it reduces misclassification errors and has better data coverage.

A second potential econometric challenge is selection into the profession. Given the importance of English proficiency in economics, perhaps only those mathematically-inclined people with sufficiently strong English skills choose to become an economist. Although information about non-economists is unavailable, this selection effect should attenuate the empirical results of this paper. A related issue is that, due to data constraints, the sample focuses on an extremely talented and

successful group of economists. An individual that is able to make it into this upper echelon of the profession will presumably have relatively strong English skills (which is consistent with Figure 1). Thus, within this group of well educated and talented economists, language differences should play a relatively minor role which will further attenuate the results.

The empirical results show that, after controlling for other characteristics of the economist, native English speakers have a significant advantage in the economics profession. Specifically, being born in an English speaking country increases the rank of an economist by about 100 spots. Furthermore, native English speakers have an advantage in both components of the ranking: they are more highly ranked according to quality adjusted publications and according to quality adjusted citations.

A series of extensions examine and address many other alternate explanations for this finding. First, the sensitivity of the results to using different definitions of native English speakers is examined. Similar results are found if English proficiency is defined using information on a country's official language, the percentage of the population that speaks English, the linguistic distance between English and other native languages, and whether the undergraduate institution was in an English speaking country. Second, the findings are robust to using alternate country-of-birth*year controls such as real GDP per capita and average educational attainment. This indicates that being a native English speaker is not inadvertently capturing some other advantage of being born in this group of countries. Third, these results are not simply driven by Americans, who comprise a large share of the top 2.5% of economists. Fourth, similar results are obtained using a totally separate ranking of economists provided by Coupe (2003), which dispels concerns that these results are specific to the IDEAS/RePEc data. Fifth, the results are significant across different age cohorts, which indicates the findings are not being driven by compositional effects (i.e. maybe older economists are more likely to be native English speakers and have a larger publication portfolio). Sixth, results including field of study, Ph.D. institution, and employer controls indicate the results are not being driven by differences in subdisciplines, differences in network effects, or differences in research expectations across countries. Finally, the results are robust to using different subsamples of economists, including for instance the top 800 and the bottom 800 economists. Overall, these extensions dispel numerous alternate explanations and thus leave English proficiency as the most plausible driver of these results.

It is important to emphasize that these results do not imply that non-native English speakers are doomed to failure or that native English speakers are guaranteed success. To the contrary, many of the most brilliant economists ever are non-native English speakers and presumably some of the worst economists of all time are native English speakers. Rather the findings of this paper suggest that a native English speaker has, on average, a slight advantage over a similar non-native English speaker. Given the importance of a well-written paper emphasized by many editors of prominent economics journals, this finding should not be surprising.

However, the implications of these findings are important. Specifically, they indicate that some economists may have an advantage in the publishing process simply due to their country of birth and their native language. This is troubling since many important decisions are based on publishing success. Employers should be cognizant of these discrepancies when making hiring, tenure, promotion, and compensation decisions. Furthermore, the fact that significant results are found in a math based discipline like economics, suggests that the advantages of being a native English speaker are likely even more profound in other professions. For instance, native English speakers in more language intensive disciplines (such as other humanities) or in professions that require more spoken English (such as services or business) are likely to have an even greater advantage. Given the rapid pace of globalization and the growing prevalence of English, these findings are especially noteworthy.

This paper contributes to a large literature that focuses on the publishing process within the economics profession. For instance, existing studies show that an economist's research output can be affected by a variety of factors, including their age, their name, their gender, their initial job placement, and whether they coauthor (Hudson 1996; Maske, Durden, and Gaynor 2003; Einav and Yariv 2006; McDowell, Singell, and Stater 2006; Oyer 2006; Oyer 2007; van Praag and van Praag 2008; Hamermesh 2013; Hamermesh 2015). Other papers examine trends in the publication process at top economics journals (Ellison 2002; Card and DellaVigna 2012; Card and DellaVigna 2013). This large body of work provides numerous interesting insights. However, to the best of my knowledge, this is the first paper to specifically examine the impact of English proficiency on research output.

There is also a large related literature that focuses on the English proficiency of immigrants. This research shows that the language skills of immigrants has a strong positive impact on their

labor market outcomes such as employment, wages, and educational attainment in a variety of English speaking countries (Rivera-Batiz 1990; Chiswick 1991; Chiswick and Miller 1995, 2010, and 2014; Borjas 2000; Carnevale et al. 2001; Dustmann and Fabbri 2003; Bleakley and Chin 2004; Bleakley and Chin 2010; Ferrer, Green, and Riddell 2006; Adsera and Pytlikova 2015). Peri and Sparber (2009) highlight the importance of English proficiency by showing that immigrants in the U.S. specialize in manual labor intensive occupations while similarly skilled natives tend to specialize in language intensive jobs. My paper does not focus on immigrants per se, but it is related to this literature in that it highlights the importance of English proficiency. Consistent with this existing evidence, I find that English proficiency has a strong positive impact on labor market outcomes within the economics profession.

The results of this paper also offer an intriguing potential explanation for cross-country differences in income inequality. Specifically, research shows that top income shares have increased substantially in English speaking countries but not in continental Europe or in Japan (Atkinson, Piketty, and Saez 2011; Alvaredo et al. 2013; Bakija, Cole, and Heim 2012). These cross-country differences have been hard to reconcile with common explanations of income inequality, such as globalization and skill biased technical change, which should be prevalent in all of these developed countries. However, the findings of this paper point to a new and interesting potential explanation. First, the business world is increasingly using English as the common language of discourse. Second, this paper shows that English proficiency has a significant positive impact on labor market performance. Thus, managers and executives from English speaking countries, who represent a large fraction of top income earners (Bakija, Cole, and Heim 2012), are in a unique position to disproportionately benefit from the "spread of their language in global corporate life."¹⁶ Certainly additional research is needed, but this seems like one possible explanation for the increase in top income shares in English speaking countries.

The remainder of the paper is organized as follows. Section 2 discusses the data used in this analysis and presents a number of descriptive statistics. Section 3 focuses on the empirical specification and section 4 presents the baseline results. A variety of potential alternate explanations are examined in the extensions discussed in Section 5. Finally, Section 6 concludes.

¹⁶ "The English Empire" The Economist. February 15th, 2014.

2 Data

2.1 Rankings Data

IDEAS/RePEc provides one of the most comprehensive and commonly used rankings of authors, journals, and institutions in the economics profession (Zimmermann 2013). Over 1,700 archives provide bibliographic data to RePEc on over 1,600,000 items of research (including books, articles, working papers, book chapters and software components) spanning over 2,000 journals (including many non-English journals), over 13,000 institutions, and over 42,000 individuals.¹⁷ While omissions are unavoidable, the scale and scope of this dataset are impressive. The IDEAS/RePEc data is preferable for these reasons, however section 5.3 examines whether the results are robust to using an older ranking of economists produced by Tom Coupe (2003).

For the purposes of this analysis, the publicly available February 2015 rankings of the top 2.5% of economists (1,082 economists) across a variety of categories was obtained from IDEAS/RePEc. These categories include various methods of measuring the number of papers, the number of citations, the number of journal pages published, and the number of abstract views and downloads. IDEAS/RePEc constructs a total rank for each economist as the mean of these, often times repetitive, individual measures. One unfortunate implication of this total ranking is that some authors have a very high overall ranks simply because of high rankings in a couple of these criteria. For instance, economists who often publish in statistical software journals are ranked very highly due to abstract views and downloads but are ranked lower according to more traditional measures such as publications and citations.

Instead of using IDEAS/RePEc's composite rank, this paper focuses on two of the best specific measures of publishing success. The first measure is the weighted number of distinct papers divided by the number of authors on each work. The second measure is the weighted number of citations divided by the number of authors on each work. The weights in both measures reflect quality by using a recursive impact factor weighting method (Zimmermann 2013). These two measures represent the best and most common method for judging research output.¹⁸ An average of these two individual measures is then calculated and used as the total ranking of economists in this paper.

¹⁷See IDEAS/RePEc website: <https://ideas.repec.org>.

¹⁸See IDEAS/RePEc website: <https://ideas.repec.org>.

These three variables successfully capture prominence within the profession and are consistent with other metrics. For instance, among the top 10 in each of these measures there are seven Nobel prize winners and six John Bates Clark winners, with no doubt more to come.

2.2 Biographical Information

Collecting the biographical information of the top 2.5% of economists was the most challenging and time consuming aspect of this project. IDEAS/RePEc only provides the name and rank of each individual, so additional information about each of the economists had to be individually gathered. Specifically, data on the country of birth, date of birth, gender, date of death (when applicable), undergraduate institution, undergraduate graduation year, Ph.D. institution, Ph.D. graduation year, among other variables was collected from a variety of different sources. To maintain the integrity of the data set a rigorous data collection process was utilized, which often required each piece of information to be confirmed by multiple sources. If there was any doubt about the accuracy of the data being provided or if different sources provided conflicting information, then that particular observation was left missing. See Appendix A.1 for additional details about how this biographical information was collected.

This analysis defines a native English speaker as a person born in the U.S., Canada, U.K., Australia, Ireland, or New Zealand.¹⁹ Using country of birth to define native English speakers is appealing for a few reasons. First, data on country of birth is more readily available than self-reported information on language proficiency. However, when both are available they are almost always consistent. Second, using country of birth eliminates the potential for misclassification errors associated with self-reported language ability (Dustmann and van Soest 2001). Finally, self-reported language proficiency is potentially endogenous. For instance, both research output and the capacity to learn a language are likely correlated with unobserved ability. Using country of birth to determine language proficiency eliminates these endogeneity concerns.

With that being said, it is certainly possible that an economist was born in one of these countries but is not a native English speaker or conversely was born in a different foreign country but is a native English speaker. However, the data collection process suggested that this is rare and it should, if anything, attenuate the empirical results that follow.

¹⁹Section 5.1 shows that the results are robust to alternate definitions of native English speaking countries.

For some economists, information on their country of birth or date of birth was not available or was unreliable. To maintain as complete a dataset as possible, information on undergraduate institution and graduation year, which are almost always available, are used to make inferences about these missing values. Specifically, it is assumed that the individual was born in the same country as their undergraduate institution and that they were 22 years old when they graduated. Appendix A.2 confirms that these assumption are reasonable and also demonstrates that the results are actually stronger if these country of birth and date of birth observations are left missing.

In addition to English proficiency, the biographical information is also used to construct a number of additional individual level controls. Specifically, the empirical analysis that follows includes: age, age squared, gender, and deceased. A control variable indicating the first letter of economists' last name is also included, where the variable takes on values 1 through 26 that reflect the letters of the alphabet. Given the practice in economics to list coauthors alphabetically, economists with last names beginning with a letter earlier in the alphabet may have an advantage (Einav and Yariv 2006; van Praag and van Praag 2008).

Finally, the empirical analysis will also control for the level of development of the country in which the economists was born. Specifically, using 1990 World Bank data, the country of birth is classified as either low income, low-middle income, upper-middle income, or high income.²⁰ It is important to control for the level of development to ensure that the empirical analysis identifies the benefits of being a native English speaker and doesn't inadvertently capture the benefits of being born in these developed countries. An extension reported in Section 5.2 shows that the results are robust to using other country of birth controls, such as GDP per capita and educational attainment, despite the fact that this leads to many more missing observations.

2.3 Descriptive Statistics

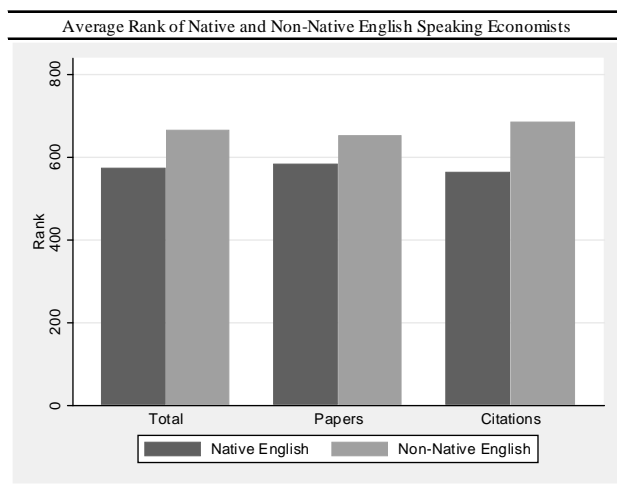
Table 1 reports summary statistics of the key dependent and independent variables. Specifically, 65% of top economists are native English speakers, they are on average 59 years old, they are 94% male, on average their last name begins with the letter K, and 3% of them are deceased. Table A2 in the appendix reports the country of birth of these top economists. Clearly, the economists in

²⁰If data was not available for a particular country in 1990, then information from a subsequent year was used instead. This was relevant for only 9 countries and entailed using data predominantly from 1991 or 1992.

this sample were born in a diverse set of countries spanning different stages of development.

To gain insight into the relationship of interest, Figure 2 shows the average ranking of native English speakers and non-native English speakers. Specifically, on the left side of Figure 2, the average total rank of native English speakers is 573 while the average total rank of non-native English speakers is 667 (more successful economists have a lower rank). Thus, even within this group of exceptional economists, whose unobserved ability is likely similar, there appears to be advantages to being a native English speaker. Similarly, on the right hand side of Figure 2, native English speakers tend to outperform non-native English speakers according to both the number of papers they write and the number of citations they receive. This basic bar chart suggests that native English speaking economists do in fact have an advantage in the publishing process.

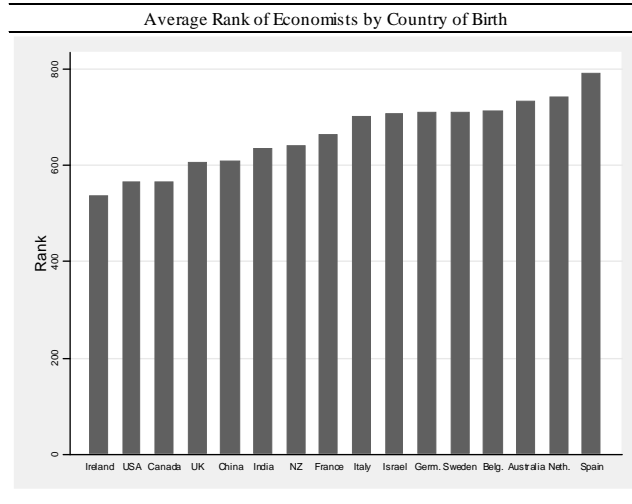
FIGURE 2



The average rank of native and non-native English speaking economists is reported (with a lower number indicating a better ranking). Rankings provided by RePEc/IDEAS and biographical information obtained by the author.

Figure 3 then shows the average rank of economists by country of birth. Specifically, the average total rank is reported for each country with more than a dozen economists in the rankings. Despite their small size, Ireland and New Zealand are also included in Figure 3 since they are defined as native English speaking countries and are thus of particular interest in this analysis. In Figure 3, the four countries with the lowest (i.e. best) average ranking are all native English speaking countries (i.e. Ireland, USA, Canada, and UK) which is consistent with the hypothesis of this paper. Interestingly, economists from China and India rank relatively highly too.

FIGURE 3



The average rank of top 2.5% of economists by country of birth is reported (with a lower number indicating a better ranking). Countries with more than a dozen economists in the top 2.5% plus Ireland and New Zealand are reported. Rankings provided by RePEc/IDEAS and biographical information obtained by the author.

One drawback is that the basic bar charts in Figures 2 and 3, do not control for demographic characteristics that may influence research output. However, what is surprising and interesting is that such strong results emerge from these basic cuts of the data. Both figures suggest that native English speakers do in fact have an advantage in the economics profession. The remainder of the paper examines to what extent these findings survive a more rigorous econometric analysis.

3 Empirical Specification

Figure 1 indicates that native English speakers are disproportionately represented in the sample of the top 2.5% of economists. The empirical analysis focuses more carefully on this sample and asks if being a native English speaker increases your ranking within this group of highly talented economists. Specifically, the empirical analysis estimates the following equation:

$$(1) \quad rank_i = \beta_0 + \beta_1 English_i + \beta_2' X_i + WB_c + \varepsilon_i.$$

where $rank$ is the opposite sign of the rank of individual economist i . Ranks are switched from positive to negative values in order to ease the interpretation of the coefficients. Thus, a positive

coefficient in equation (1) indicates that that variable improves the ranking of the economist.

English is a binary variable indicating whether economist i was born in a native English speaking country. X is a vector of individual level controls including age, age squared, the first letter of the economist’s last name, a binary variable indicating male, and a binary variable indicating whether the economist is deceased.²¹ Finally, WB are the World Bank fixed effects which capture the level of development of the country of birth. Specifically, binary variables are included which indicate whether the country of birth is a low income, low-middle income, upper-middle income, or high income country. Finally, the standard errors in all the subsequent regressions are clustered at the country of birth level.

If native English speakers have an advantage in publishing in the economics profession then $\beta_1 > 0$. Conversely, if growing up speaking another language and being exposed to other cultures offers an unique perspective that is useful in economic research then $\beta_1 < 0$.

One especially appealing aspect of this empirical specification is that all of the independent variables included in equation (1) are exogenous. Specifically, since date of birth, country of birth, name, gender, and date of death do not respond to the rank or unobserved ability of the economist, endogeneity is of minimal concern within this context. Other potential controls such as field of study, Ph.D. institution, and current employer are likely a function of unobserved ability or English proficiency itself, and are thus not included in the baseline analysis. However, these factors will be controlled for in an extension reported in section 5.5.

4 Results

4.1 Baseline

The results from estimating equation 1 are reported in Table 2. All regressions have standard errors clustered at the country level in brackets. Column 1 does not include the control variables and finds that native English speakers are on average ranked about 94 spots higher (better) than non-native English speakers. This is consistent with the results from Figure 2. Column 2 includes the individual level control variables and still finds that English proficiency has a significant positive impact on the ranking of an economist. Finally, column 3 includes the individual controls and

²¹For economists who have died, the age at death is used rather than what the current age would be.

the level of development fixed effects and shows that native English speakers are on average more highly ranked. Specifically, holding other factors constant, being a native English speaker improves an economist's ranking by 100 spots.

The control variables are of the expected sign in Table 2 and most are significant. Research output is increasing with age but at a diminishing rate. Males are on average ranked higher (although as section 5.4 shows, this effect has dissipated over time) and being deceased has an insignificant impact. In addition, economists with last names beginning with a letter towards the end of the alphabet are on average ranked worse. This finding is consistent with Einav and Yariv (2006) and van Praag and van Praag (2008) and is likely due to the norm within economics to list coauthors alphabetically. Thus, over time authors with a name starting with a letter towards the beginning of the alphabet, will receive more notoriety and thus be ranked higher.

Overall, the results in Table 2 provide evidence that native English speakers do in fact have an advantage in the economics profession. The strength and significance of this finding is surprising given the host of reasons one would expect a relatively small impact in economics. For example, this is a sample of extremely well educated, talented economists where we would expect the differences in English proficiency to be minimal. Second, economics is a math and statistical based discipline that relies less on eloquent language relative to other fields. Third, since economists often collaborate, one would expect the impact of native language to be minimal as native English and non-native English speakers coauthor together. Fourth, the economists on this list have large resources at their disposal which can be used to hire English proofreaders if need be. Thus, for a variety of reasons one would expect a minimal impact, and yet Table 2 shows that native English speakers have a significant advantage in the publication process over similar non-native English speaking colleagues.

4.2 Components

Table 3 examines the impact of being a native English speaker on each of the components of the total rank variable. Specifically, column 1 presents the baseline results while columns 2 and 3 use the quality adjusted number of papers and citations as the dependent variable respectively. The results indicate that being a native English speaker has a positive and significant impact on both of these components. In column 2 of Table 3, native English speakers are ranked 58 spots

higher according to their quality adjusted number of papers. Consistent with the editor’s comments about the importance of well-written papers, the results in column 2 indicate that economists with a better command of the English language are more successful at publishing their work.

In addition, column 3 shows that native English speakers are ranked 142 spots higher according to the (quality adjusted) number of times their research is cited. If native English speakers are on average writing clearer and more persuasive papers, as the editors quotes indicate, then it is not surprisingly that they are relatively more successful at publishing their work (column 2) and that their research is cited more often (column 3). Overall, Table 3 confirms that native English speakers have an advantage in terms of both the quantity and quality of their papers and citations.

The results in Table 3 use the rank of an economist as the dependent variable. However, Appendix A.4 investigates whether the findings are robust to using the underlying (quality adjusted) number of paper and citation ‘score’ variables instead. Consistent with the baseline results, this extension indicates that native English speakers have approximately 9% more quality adjusted papers and 22% more quality adjusted citations. Overall, Table A3 shows that the results are not sensitive to whether the rank or the score of the economist is used as the dependent variable.

5 Extensions

The baseline results provide compelling evidence that native English speaking economists do in fact have an advantage in the publishing process. However, perhaps there are other potential explanations that could be contributing to these strong and significant results. Thus, the goal of this section is to examine and hopefully refute these alternate hypotheses.

5.1 English Definition

This subsection investigates whether the results are robust to alternate definitions of English proficiency. This includes using broader definitions of native English speaking countries, an analysis that takes into account the linguistic distance between foreign languages and English, and using undergraduate institution rather than the country of birth as a proxy for English proficiency.

First, this analysis has defined native English speakers as those economists born in the U.S., Canada, the UK, Australia, Ireland, and New Zealand. This is a logical but conservative definition

of English speaking countries. There are other countries that either use English as their official language or have a significant English speaking population. Thus, two alternate definitions of native English speaking countries are constructed. Using data from CEPII, countries were defined as English speaking if English is the official language or if at least 20% of the population speaks English.

The results from this robustness analysis are reported in Table 4. Column 1 defines English speaking countries as those where English is the official language. According to CEPII, this includes the six baseline countries as well as Hong Kong, India, Jamaica, and Zimbabwe.²² Column 2 defines a country as English speaking if at least 20% of the population speaks English. According to CEPII, this includes the six baseline countries as well as Hong Kong, India, Israel, Jamaica, Lebanon, Pakistan, Rwanda, South Korea, and Zimbabwe. The results in columns 1 and 2 indicate that English proficiency has a positive and significant impact on the ranking of economists regardless of how native English speaking countries are defined.

Second, the linguistic distance between English and an economist's native language may be important. For instance, economists that grew up speaking a foreign language that is more similar to English may have an easier time learning English and subsequently publishing in economic journals. To investigate this possibility, I use data on linguistic distance between countries constructed by Fearon (2003) and Spolaore and Wacziarg (2009). Specifically, Fearon (2003) identifies how many common nodes in the linguistic tree two languages share. Spolaore and Wacziarg (2009) then create a weighted measure of linguistic similarity between countries using the number of nodes two randomly chosen people (one from each country) would share.²³ Following Fearon (2003) and Spolaore and Wacziarg (2009), I then transform this variable so that it is bounded by 0 and 1 and increasing in linguistic distance.²⁴

Column 3 in Table 4 shows that this measure of linguistic distance has a significant negative effect on subsequent publishing success within the economics procession. This means that an economist's ranking is decreasing as the linguistic distance between their country of birth and the

²²Obviously, there are many other countries where English is the official language but these are the ones where at least one economists in the top 2.5% was born.

²³My results are similar if I instead use an unweighted measure that simply measures the number of common nodes shared by the most commonly spoken language in the two countries.

²⁴There are a maximum of 15 nodes so the transformation takes the following form:

$$Linguistic_Distance = \sqrt{\frac{(15 - \#common_nodes)}{15}}$$

U.S. increases. This is consistent with the idea that it is more difficult for some economists to learn English and thus publish in English speaking journals than others. It is important to emphasize that this continuous measure of linguistic distance is totally distinct from the measure of English proficiency used in the baseline analysis. However, the results are very similar.

Third, perhaps English proficiency is driven more by the undergraduate institution rather than the country of birth (Hamermesh and Pfann 2012). To test this hypothesis, it is possible to define an English speaker as someone who went to an undergraduate institution in a native English speaking country rather than someone who was born in a native English speaking country. Column 4 of Table 4 shows that the coefficient on this alternate definition of native English speakers is positive and significant.²⁵ Finally, column 5 defines economists as native English speakers if they were born *and* went to an undergraduate institution in a native English speaking country. Again the results are positive and significant. Overall, Table 4 shows that the results are robust to a wide variety of alternate definitions of native English speakers.

5.2 Country Controls and Results

Perhaps the English proficiency variable is inadvertently capturing some other benefit of being born in these six countries or maybe the results are being driven by one particular country and not reflecting the benefits of being a native English speaker more generally. This subsection investigates both of these issues.

While the World Bank fixed effects, which control for the level of development of the country of birth, should mitigate this former concern, maybe this is too coarse a measure. Although few potential time varying country-of-birth controls span the countries and years in this data set, fortunately both real GDP per capita and average years of schooling have reasonably good coverage. Data on real GDP per capita was obtained for numerous countries from 1950-2010 from the Penn World Tables. Data on average years of schooling was obtained from the Barro and Lee (2013) Educational Attainment Dataset.²⁶

The results in Table 5 account for the possibility that the level of income or the educational

²⁵ Additional results show that the benefits of going to an undergraduate institution in a native English speaking country are, not surprisingly, even larger for economists born in non-English speaking countries.

²⁶ The average years of schooling data of those over 25 years old is reported every five years from 1950-2010 and the intervening years are calculated using linear interpolation.

system within a country affects the economist's upbringing which in turn is correlated with their future success. Specifically, column 1 controls for the real GDP per capita and the average years of schooling in the country and year of the economist's birth. Neither real GDP per capita nor average years of schooling are significant and importantly the coefficient on English remains positive and significant. This result is similar to the baseline findings, despite the fact that a third of the sample is lost because of missing GDP and schooling values.²⁷

In order to reduce the number of missing observations, Column 2 uses real GDP per capita and the average years of schooling in the country of birth when the economist was 18 years old instead. Using information from when the economist was 18 is appealing since this is when they are in the midst of their education and thus the impact of their country of birth may be most profound. However, the results in column 2 remain similar. Once again GDP and schooling are insignificant and the coefficient on English remains positive and significant.

Another concern is that maybe the results are being driven by Americans who represent a large fraction of the top economists and who may have an advantage because a disproportionate number of the best economics departments and journals happen to be based in the U.S. Although Figure 3 refutes this concern, columns 3 and 4 decompose native English speakers into those born in the U.S. and those born in the other five native English speaking countries. The coefficients on both variables are positive and significant, and importantly not statistically different from one another. This indicates that the results are not being driven solely by U.S. born economists.²⁸

Table 5 shows that including time-varying country specific controls does not offer any benefit over the World Bank fixed effects. The results remain similar but these controls significantly reduce the number of observations which is why the World Bank fixed effects are used in the baseline analysis. Furthermore, the results in Table 5 support the assertion that it is these country's shared language that is driving the results and not some idiosyncratic factor specific to Americans. Overall, these findings provide no evidence that the English proficiency variable is inadvertently capturing other characteristics of the home country or that one particular English speaking country is driving the results.

²⁷Due to a lack of data, any economist that was born prior to 1950 is automatically dropped from the sample and in many countries economists born much later were also dropped.

²⁸Additional results show that all six native English speaking countries individually have significant positive effects on total rank, although these results are a bit noisier given the small sample sizes.

5.3 Alternate Rankings

This subsection investigates whether the results are robust to using other rankings of economists. The IDEAS/RePEc dataset provides the most comprehensive, the most rigorous, and the most commonly used ranking of economists. However, perhaps there is some unobserved bias against non-English speaking economists in the IDEAS/RePEc rankings that are affecting the results in this paper. While there is nothing in the data that suggests this is in fact the case, it is worth exploring whether the results are robust to the use of other rankings.

The only other comprehensive rankings of economists, that I am aware of, was constructed by Tom Coupe (2003). Specifically, he ranked economists according to publications from 1969-2000 and ranked economists according to citations from 1975-2000.²⁹ The analysis focuses on quality adjusted articles and the citation count which are broadly consistent with the measures used in the IDEAS/RePEc data.³⁰ However, there are some differences between these two datasets that are worth noting: the Coupe data covers an earlier time period, uses a different database of articles, includes fewer economists, and measures publications and citations in a slightly less sophisticated way.

Nonetheless, the results using the Coupe rankings are reported in Table 6. This analysis focuses on the top 611 economists according to total rank, which is an average of the publication rank and the citation rank.³¹ The results indicate that native English speaking economists are ranked higher overall (column 1), ranked higher according to their publications (column 2), and ranked higher according to their citations (column 3). Not only are these results significant and identical in sign to the baseline results, but amazingly the magnitudes are also very similar too. Thus, the results are virtually the same when using a completely different data set, that covers a different sample of years, includes a different group of economists, and uses slightly different rank variables. Overall this provides support for the conclusions of the paper and indicates that the baseline results are robust to the use of a totally distinct rankings of economists.

²⁹A more limited version of this data spanning the years 1990-2000 was used in Tom Coupe's 2003 paper. The more comprehensive data used in this analysis is available on his archived website.

³⁰The results are robust to using other rankings of publications and citations.

³¹The sample is limited due to the fact that some economists are not ranked according to either the citation or the publication metric.

5.4 Cohorts

It is possible that the composition of native English speaking economists could be changing over time. For instance, perhaps older economists are more likely to be native English speakers and given their longer working careers they may also be more highly ranked. This story would be problematic for this analysis because it would generate a spurious positive coefficient on English simply do to compositional changes. The inclusion of age and the other demographic controls should mitigate this concern. However, this subsection examines more carefully whether these types of compositional changes could be affecting the baseline results.

Table 7 splits the sample roughly in half based on birth year. The contrasting findings between column 2 which reports the results for those born before 1955 and column 3 which reports the results for those born in 1955 or after are interesting. While both coefficients are significant, being a native English speaker has a stronger impact on the rank of younger economists in column 3. This refutes the compositional concerns associated with older economists and instead indicates that the importance of English proficiency is becoming even more important over time. This finding is interesting in light of the recent spread of English and the emphasis placed on learning the language in many foreign countries. One potential explanation is that due to the massive increase in submissions and declining acceptance rates at top journals (Card and DellaVigna 2013), editors are increasingly relying on writing ability as a signal of paper quality (see quotes in the introduction).

Another interesting result is that being male has a significant positive impact on the research success of the older cohort of economists in column 2 but an insignificant impact on the younger cohort in column 3. This is encouraging and indicates that the advantage of being male in the economics profession has diminished over time.

5.5 Individual Controls

The baseline analysis only includes independent variables that are plausibly exogenous. Country of birth, date of birth, name, gender, and date of death do not depend on the rank or unobserved ability of the economist, which is an appealing feature of the econometric specification. With that said, there are other factors, such as field of study, Ph.D. institution, and country of employment

that may influence the research output of an economist. While potentially important, these factors were not included in the baseline analysis since they are likely endogenous to unobserved ability and English proficiency. However, this section examines whether the results survive the inclusion of these additional variables.

First, field of study may influence an economist's ranking since it could be easier for some fields to publish in top journals relative to others. Furthermore, economist's in larger subdisciplines may have more citations to their work simply due to the sheer size of their field. Second, whether an economist received their Ph.D. from an American institution may also have an important impact on future research output. Many of the top programs are located in the U.S., attending a Ph.D. program in the U.S. may improve English proficiency, and attending an American institution may provide professional networks that prove useful in the future. Third, research and tenure expectations as well as professional networks may vary according to the country of employment. For instance, U.S. employers may have higher research expectations, better professional networks, and attract more native English speakers.

Table 8 includes these various control variables. Column 1 reports the baseline results while column 2 includes binary controls for the field of study of the economist.³² Most of these large subdisciplines have a significant positive effect on an economist's ranking (relative to the excluded smaller fields), but importantly the English proficiency coefficient remains positive, significant, and is actually larger in size. The results in columns 3 and 4 indicate that economists that obtained their Ph.D. from a U.S. university and that currently work for a U.S. institution are ranked higher.³³ Given the endogeneity of both variables it is hard to distinguish between whether this implies that attending and working for American institutions is beneficial for research output or whether a selection effect means that talented economists are more likely to attend and work at U.S. institutions. Nonetheless, the more important point is that the coefficient on English proficiency remains positive and significant in both specifications. Not surprisingly, the magnitude of this coefficient is slightly smaller in column 4, since where an economist works is a direct function of his or her research. Thus, controlling for U.S. employment absorbs a lot of potentially useful variation in research output

³²The subdiscipline of the economist was determined using the field designation in IDEAS/RePEc. Note, it is possible for the same economist to be listed in different subfields.

³³Current employer is not available for deceased economists which limits the sample in column 4.

Finally, column 6 includes all of these additional controls together. The results show that English proficiency has a significant positive effect on research output after controlling for the economist’s field of study, whether they obtained a U.S. Ph.D., and whether they work in the U.S. It is reassuring that the results remain significant despite the inclusion of all of these potentially endogenous controls. Overall, Table 8 shows that subdisciplines, network effects, and research expectations are not the primary drivers of the results in this paper. Even after controlling for all of these factors, native English speaking economists still have an advantage.

5.6 Sample

The baseline specification examines the impact of English proficiency on the ranking of the top 2.5% of economists. This subsection examines to what extent the results are sensitive to using different samples of economists. Specifically, Table 9 restricts the analysis to the top 1000, the top 900, the top 800, the top 700, the bottom 1000, the bottom 900, the bottom 800, and the bottom 700 economists within the top 2.5% sample. This provides insight into whether an unusual composition of economists at the top or bottom of the rankings could be driving the results.

Table 9 shows that the results are not sensitive to the sample of economists included in the analysis. The coefficient on English is positive and significant in all eight specifications. This is reassuring and indicates that the baseline results are robust to the use of a variety of alternate samples.

6 Conclusion

This paper investigates whether being a native English speaker has a significant impact on labor market performance. This question is studied within the context of the economics profession, where the effect of English proficiency on publishing success is examined. A comprehensive data set on the rankings of top economists and their biographical information is constructed to examine this issue.

The results provide compelling evidence that native English speakers do in fact have an advantage in the economics profession. Controlling for a host of other factors, the results indicate that native English speakers are ranked, on average, about 100 spots higher than non-native English

speakers. Similar results are found when separately examining the impact on paper and citation ranks. Additional results using the underlying paper and citation scores rather than the ranks (see appendix A.4), confirm these baseline results and indicate that the findings are meaningful. English proficiency leads to approximately a 9% increase in the quality adjusted number of papers and a 22% increase in the quality adjusted number of citations.

A host of extensions and robustness checks dispel other potential explanations for these findings and suggests that English proficiency is leading to a significant advantage in research output in the economics profession. This confirms anecdotal evidence that indicates that native English speakers have an advantage in writing economic papers and that well written papers are more likely to be accepted by economics journals.

Due to data constraints, this analysis focuses on an elite group of economists. However, these are the leaders in the field who have enormous influence within the profession, in policy circles, and in public debates more generally. Thus, the finding that this elite group of economists disproportionately consists of native English speakers may have important implications within and outside the profession. Furthermore, since this sample of economists are incredibly talented and have enormous resources at their disposal, the impact of English proficiency should if anything be attenuated.

More generally, these results have important implications. First, they indicate that some individuals have an advantage at publishing within economics simply due to their country of birth and native language. Since research output is a crucial factor in tenure, promotions, and compensation this finding is potentially troubling. Second, these results have broader implications for other disciplines and professions. The fact that such strong results are found in a math and written based discipline like economics suggests that the impact of being a native English speaker is likely even larger in other fields and professions. Given the rapid pace of globalization and the spread of the English language, this is an important finding that warrants further research.

References

- Adsera, Alicia and Mariola Pytlikova. 2015. "The Role of Language in Shaping International Migration." *The Economic Journal*, 125(586): F49-F81.
- Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez. 2013. "The Top 1 Percent in International and Historical Perspective" *The Journal of Economic Perspectives*, 27(3): 3-20.
- Ammon, Ulrich. 2003. "The International Standing of the German Language." in J. Maurals and M. Morris (Eds.), *Languages in a Globalising World*, Cambridge University Press.
- Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature*, 49(1): 3-71.
- Bakija, Jon, Adam Cole, and Bradley T. Heim. 2012. "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data." Williams College Working Paper #2010-22.
- Barro, Robert and Jong-Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics*, 104: 184-198.
- Blaug, Mark and Howard R. Vane. 2003. "Who's Who In Economics, Fourth Edition." Edward Elgar Publishing.
- Bleakley, Hoyt and Aimee Chin. 2004. "Language Skills and Earnings: Evidence from Childhood Immigrants." *The Review of Economics and Statistics*, 86(2): 481-496.
- Bleakley, Hoyt and Aimee Chin. 2010. "Age at Arrival, English Proficiency, and Social Assimilation Among US Immigrants." *American Economic Journal: Applied Economics*, 2(1): 165-192.
- Borjas, George J. 2000. "Foreign-Born Teaching Assistants and the Academic Performance of Undergraduates." *The American Economic Review*, 90(2): 355-359.
- Card, David and Stefano DellaVigna. 2012. "Revealed Preferences for Journals: Evidence from Page Limits." NBER Working Paper No. 18663.

- Card, David and Stefano DellaVigna. 2013. "Nine Facts about Top Journals in Economics." *Journal of Economic Literature*, 51(1): 144-161.
- Carnevale, Anthony P, Richard A. Fry, and B. Lindsay Lowell. 2001. "Understanding, Speaking, Reading, Writing, and Earnings in the Immigrant Labor Market." *The American Economic Review*, 91(2): 159-163.
- Chen, M. Keith. 2013. "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets." *The American Economic Review*, 103(2): 690-731.
- Chiswick, Barry R. 1991. "Speaking, Reading, and Earnings among Low-Skilled Immigrants." *Journal of Labor Economics*, 9(2): 149-170.
- Chiswick, Barry R. and Paul W. Miller. 1995. "The Endogeneity between Language and Earnings: International Analysis." *Journal of Labor Economics*, 13(2): 246-288.
- Chiswick, Barry R. and Paul W. Miller. 2010. "Occupational Language Requirements and the Value of English in the US Labor Market." *Journal of Population Economics*, 23(1): 353-372.
- Chiswick, Barry R. and Paul W. Miller. 2014. "International Migration and the Economics of Language." IZA Discussion Paper No. 7880.
- Crystal, David. 2003. "English as a Global Language." Cambridge University Press, Cambridge, UK.
- Dustmann, Christian. 1994. "Speaking Fluency, Writing Fluency, and Earnings of Migrants." *Journal of Population Economics*, 7(2): 133-156.
- Dustmann, Christian and Arthur van Soest. 2001. "Language Fluency and Earnings: Estimation with Misclassified Language Indicators." *The Review of Economics and Statistics*, 83(4): 663-674.
- Dustmann, Christian and Francesca Fabbri. 2003. "Language Proficiency and Labor Market Performance of Immigrants in the UK." *The Economic Journal*, 113(489): 695-717.
- Einav, Liran and Leeat Yariv. 2006. "What's in a Surname? The Effects of Surname Initials on Academic Success." *Journal of Economic Perspectives*, 20(1): 175-188.

- Ellison, Glenn. 2002. "The Slowdown of the Economics Publishing Process." *Journal of Political Economy*, 110(5): 947-993.
- Engemann, Kristie M. and Howard J. Wall. 2009. "A Journal Ranking for the Ambitious Economist." *Federal Reserve Bank of St. Louis Review*, May/June: 127-140.
- Fearon, James. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth*, 8(2): 195-222.
- Ferrer, Ana, David A. Green, and W. Craig Riddell. 2006. "The Effect of Literacy on Immigrant Earnings." *Journal of Human Resources*, 41(2): 380-410.
- Fourcade, Marion, Etienne Ollion, and Yann Algan. 2015 "The Superiority of Economists." *Journal of Economic Perspectives*, 29(1): 89-114.
- Freeman, Richard B., and Wei Huang. 2014. "Collaborating with People Like Me: Ethnic Co-Authorship within the U.S." NBER Working Paper No. 19905.
- Hamermesh, Daniel S. and Gerard A. Pfann. 2013. "Reputation and Earnings: The Roles of Quality and Quantity in Academe." *Economic Inquiry*, 50(1): 1-16.
- Hamermesh, Daniel S. 2013. "Six Decades of Top Economics Publishing: Who and How?" *Journal of Economic Literature*, 51(1): 162-172.
- Hamermesh, Daniel S. 2015. "Age, Cohort and Co-Authorship" NBER Working Paper No. 20938.
- Hudson, John. 1996. "Trends in Multi-Authored Papers in Economics." *Journal of Economic Perspectives*, 10(3): 153-158.
- Maske, Kellie L., Garey C. Durden, and Patricia E. Gaynor. 2003. "Determinants of Scholarly Productivity Among Male and Female Economists." *Economic Inquiry*, 41(4): 555-64.
- McDowell, John M., Larry D. Singell Jr., and Mark Stater. 2006. "Two to Tango? Gender Differences in the Decisions to Publish and Coauthor." *Economic Inquiry*, 44(1): 153-168.
- Oyer, Paul. 2006. "Initial Labor Market Conditions and Long-Term Outcomes for Economists." *Journal of Economic Perspectives*, 20(3): 143-160.

- Oyer, Paul. 2007. "Is There an Insider Advantage in Getting Tenure?" *The American Economic Review*, 97(2): 501-505.
- Peri, Giovanni, and Chad Sparber. 2009. "Task Specialization, Immigration, and Wages." *American Economic Journal: Applied Economics*, 1(3): 135-69.
- Rivera-Batiz, Francisco L. 1990. "English Language Proficiency and the Economic Progress of Immigrants." *Economics Letters*, 34(3), 295-300.
- Spolaore, Enrico and Romain Wacziarg. 2009. "The Diffusion of Development." *The Quarterly Journal of Economics*, 124(2), 469-529..
- Thursby, Jerry G. 2000. "What Do We Say about Ourselves and What Does It Mean? Yet Another Look at Economics Department Research." *Journal of Economic Literature*, 38(2): 383-404.
- van Praag, C. Mirjam and Bernard M.S. van Praag. 2008. "The Benefits of Being Economics Professor A (rather than Z)." *Economica*, 75(300): 782-796.
- Zimmermann, Christian. 2013. "Academic Rankings with RePEc." *Econometrics*, 1(3): 249-280.

TABLE 1
Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Total Rank	1,059	606	337	5	1,186
English	1,059	0.65	0.48	0	1
Age	1,059	59	10	35	97
Age Squared	1,059	3,588	1,286	1,225	9,409
Male	1,059	0.94	0.24	0	1
1st Letter Last Name	1,059	11	7	1	26
Deceased	1,059	0.03	0.18	0	1

Summary statistics of the top 2.5% of economists according to papers and citations using data from RePEc/IDEAS.

TABLE 2
Impact of English Proficiency on Total Rank

	Total Rank		
	(1)	(2)	(3)
English	94.17*** [17.24]	68.81*** [16.53]	99.74*** [15.90]
Age		43.95*** [7.82]	43.20*** [7.54]
Age Squared		-0.32*** [0.06]	-0.31*** [0.06]
Male		48.26* [24.79]	49.41** [23.95]
1st Letter Last Name		-1.66** [0.82]	-1.46* [0.81]
Deceased		25.98 [53.35]	26.72 [53.79]
COB WB FE	No	No	Yes
Observations	1,059	1,059	1,059
R-squared	0.018	0.064	0.074

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the total rank of each economist, which is the average of their rank according to papers and citations.

TABLE 3
Impact of English Proficiency on Components of Total Rank

	Total Rank	Paper Rank	Citation Rank
	(1)	(2)	(3)
English	99.74*** [15.90]	57.71*** [21.40]	141.78*** [21.17]
Age	43.20*** [7.54]	33.97*** [8.03]	52.43*** [8.74]
Age Squared	-0.31*** [0.06]	-0.21*** [0.07]	-0.42*** [0.08]
Male	49.41** [23.95]	78.96*** [28.88]	19.86 [38.85]
1st Letter Last Name	-1.46* [0.81]	-2.53 [1.55]	-0.40 [1.51]
Deceased	26.72 [53.79]	38.88 [46.28]	14.56 [89.66]
COB WB FE	Yes	Yes	Yes
Observations	1,059	1,059	1,059
R-squared	0.074	0.075	0.064

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. "Paper Rank" is the rank of an economist according to their quality adjusted number of distinct papers. "Citation Rank" is the rank of an economist according to their quality adjusted number of citations.

TABLE 4
Impact of English Proficiency on Rank, by Alternate English Definitions

	Total Rank				
	(1)	(2)	(3)	(4)	(5)
English Official	94.68*** [16.09]				
English 20%		70.83*** [17.85]			
Linguistic Distance			-102.57*** [16.49]		
English Undergrad				113.01*** [15.82]	
English COB & Undergrad					112.42*** [16.51]
Controls	Yes	Yes	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes	Yes	Yes
Observations	1,059	1,059	1,059	1,046	1,046
R-squared	0.073	0.067	0.073	0.084	0.084

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. Column 1 defines countries as English speaking if their official language is English (using data from CEPII) which includes the 6 baseline countries as well as Hong Kong, India, Jamaica, and Zimbabwe. Column 2 defines countries as English speaking if more than 20% of the population speaks English (according to CEPII) which includes the 6 baseline countries as well as Hong Kong, India, Israel, Jamaica, Lebanon, Pakistan, Rwanda, South Korea, and Zimbabwe. Column 3 instead uses a measure of linguistic distance from the U.S. (Spolaore and Wacziarg 2009 and Fearon 2003) to examine whether it is more difficult for some economists to learn English than others. Column 4 defines a native English speaker as an economist that went to an undergraduate institution in one of the 6 native English speaking countries. Finally, column 5 defines a native English speaker as an economist that was born and went to an undergraduate institution in one of the 6 native English speaking countries.

TABLE 5
Impact of English Proficiency on Rank, Country Specific Controls and Results

	Total Rank			
	(1)	(2)	(3)	(4)
English	138.87*** [21.50]	94.28*** [24.55]		
English U.S.			153.98*** [23.05]	119.32*** [28.89]
English Other			129.31*** [19.71]	80.00*** [23.05]
ln (RGDPPC)	-31.03 [39.89]		-40.36 [42.09]	
ln (Schooling)	-23.83 [40.78]		-34.13 [41.48]	
ln (RGDPPC) at 18		20.36 [43.67]		-1.56 [47.06]
ln (Schooling) at 18		11.15 [46.17]		-11.05 [50.26]
Controls	Yes	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes	Yes
Observations	717	1,009	717	1,009
R-squared	0.085	0.076	0.085	0.077

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. Column 1 and 3 include real GDP per capita and average years of schooling at the time of the economist's birth. Columns 2 and 4 use instead real GDP per capita and average years of schooling when the economist was 18 years old. Columns 3 and 4 separate the U.S. from other native English speaking countries.

TABLE 6
Impact of English Proficiency on Alternate Coupe Ranking of Economists

	Total Rank	Paper Rank	Citation Rank
	(1)	(2)	(3)
English	117.57*** [38.90]	117.28** [45.38]	110.02*** [39.64]
Controls	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes
Observations	611	611	611
R-squared	0.023	0.082	0.043

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. This table uses rankings of economists produced by Tom Coupe rather than from IDEAS/RePEc.

TABLE 7
Impact of English Proficiency on Rank, by Cohort

	Baseline	<1955	1955+
	(1)	(2)	(3)
English	99.74*** [15.90]	74.28** [30.53]	112.79*** [19.72]
Male	49.41** [23.95]	150.98*** [45.23]	16.72 [32.64]
Controls	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes
Observations	1,059	479	580
R-squared	0.074	0.036	0.097

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is total rank. Column 1 reports the baseline results. Column 2 includes only economists that were born prior to 1955 while column 3 includes economists born in 1955 and after.

TABLE 8
Impact of English Proficiency on Rank, with Additional Individual Controls

	Baseline	Fields	US PhD	US Empl	All
	(1)	(2)	(3)	(4)	(5)
English	99.74*** [15.90]	123.69*** [17.61]	72.02*** [21.34]	40.64** [17.20]	54.55*** [16.53]
Development		161.67*** [14.93]			166.70*** [18.27]
Labor		128.97*** [15.54]			131.40*** [14.06]
Macro		152.88*** [28.08]			155.98*** [31.04]
Econometrics		145.04*** [33.43]			158.18*** [29.26]
Micro		155.76*** [23.14]			166.09*** [29.14]
Monetary		103.03*** [23.21]			86.10*** [22.86]
Public		108.15*** [14.07]			118.80*** [13.08]
Urban		36.54 [31.79]			42.83 [37.92]
US PhD			101.22*** [24.65]		75.25*** [24.86]
US Employer				133.45*** [26.09]	118.92*** [26.41]
Controls	Yes	Yes	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes	Yes	Yes
Observations	1,059	1,059	1,046	1,021	1,009
R-squared	0.074	0.217	0.089	0.101	0.256

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. Column 1 reports the baseline results. Column 2 includes binary control variables indicating the field of study according to IDEAS RePEc. Column 3 includes a binary control indicating whether an economist received their PhD from an institution within the US. Column 4 includes a binary control indicating whether the economist currently works at an institution within the US. Finally, column 5 includes all of these controls.

TABLE 9
Impact of English Proficiency on Rank, by Sample Size

	Top 1000	Top 900	Top 800	Top 700	Bottom 1000	Bottom 900	Bottom 800	Bottom 700
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
English	86.34*** [17.34]	81.23*** [20.11]	87.56*** [20.60]	78.75*** [20.36]	83.08*** [13.07]	60.70*** [13.90]	38.50** [17.63]	29.61** [13.56]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,000	900	800	700	999	900	800	700
R-squared	0.076	0.081	0.087	0.084	0.054	0.037	0.033	0.024

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is total rank. Columns 1-4 restrict the sample to the top X number of economists, and columns 5-8 restrict the sample to the bottom X number of economists.

A Data Appendix

A.1 Data Gathering Process

Biographical information for each of the top economists in the sample was individually gathered from a variety of sources. This proved to be the most daunting aspect of this project but also the most important. Great care was taken to ensure that each piece of information was correct, which often entailed confirming it across multiple sources. The specific data collection process for each individual took the following form.

First, the curriculum vitae (CV) of the individual was used which proved to be the most important source of biographical information. Almost all economists publicly post their CVs and the information is deemed highly accurate, since the economist creates the document themselves. The CV almost always includes current employment, undergraduate institution, undergraduate graduation year, Ph.D. institution, and Ph.D. graduation year. Often the economist's CV also includes information on the country of birth, year of birth, citizenship, and languages (including "mother tongue"). Since citizenship can change, it was not used to identify country of birth unless it could be independently confirmed with other information from different sources. Furthermore, the "mother tongue" and country of birth were almost always consistent. However, country of birth had far better coverage. For economists that provided all of this biographical information, the data gathering process was complete. However, even for these individuals, the data provided in the CV was cross-checked against information provided by other sources. Unfortunately, many economists do not provide for instance country and year of birth in their CVs and so I turned to other sources.

Economists with missing or incomplete CVs, were looked up in *Who's Who in Economics* (Blaug and Vane 2003). This book provides biographical information, including name, education, country of birth, date of birth, and current and past employment, for the most noteworthy economists of the twentieth century. Numerous older and more prominent economists are listed in the book, and thus *Who's Who in Economics* proved to be a valuable source of information. When applicable, rarely did an entry in *Who's Who in Economics* conflict with the information provided on the economist's CV.

The short autobiographies written by Nobel Prize winning economists also proved to be a useful source of information. This was a first hand account of the economists past, although obviously it is

only available for a select group of economists. For economists that unfortunately passed away, there are often obituaries that include extensive biographical information. In addition, economists often grant interviews which are available online and contain information about their past. Numerous economists also provide a short biography on their website. This often has information such as country of birth, date of birth, or the place that they grew up which was not readily available on their CV.

Many economists have entries in Wikipedia that contain biographical information. However, given the open source nature of Wikipedia, I used this information sparingly and would cross-check it with data from other sources before using it. Finally, background check websites (such as Radaris) contain information on the age of an individual which can be used to infer the year of birth.³⁴ This information was only utilized if the full name (including middle initial) matched the economist, if the place of residence was consistent with the economist's current employer, and if past locations were consistent with past employers of the economist. Finally, before using this information, the date of birth was cross checked against the year of undergraduate degree to ensure that the implied age at graduation was plausible. For economists with common names it was not possible to conclusively identify them on these websites and thus the observation was left missing.

A.2 Missing Values

Despite my best efforts to construct a comprehensive dataset using this data collection process, some biographical information for certain economists remains unavailable. Furthermore, if there was any ambiguity or inconsistencies across sources, I erred on the side of caution and left the observation missing. Thus, a portion of economists within my sample have missing biographical information. For this paper, the country of birth and date of birth missing values are potentially problematic since they are used to construct the English, age, and age squared variables utilized in the empirical analysis. Information on gender, the first letter of the economists last name, and the year of death (when applicable) were never missing.

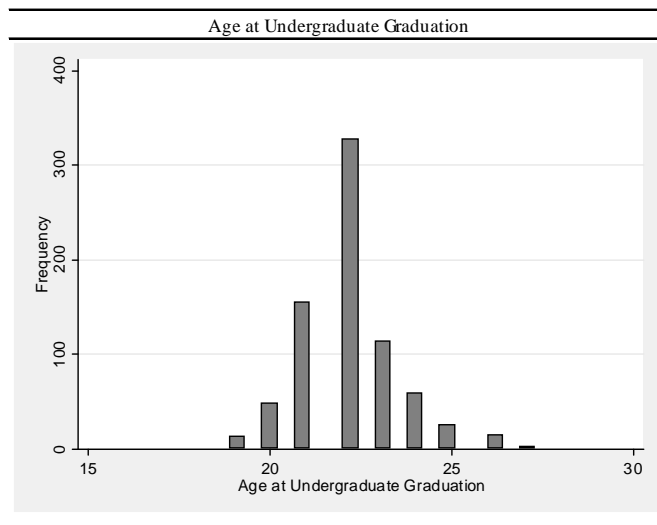
In the sample of the top 2.5% of economists, 9% of the year of birth observations are missing and 20% of the country of birth observations are missing. It is unclear whether to leave these

³⁴The amount of publicly available information on these websites is surprisingly comprehensive. For instance, they have information on any person that resided within the U.S. for any period of time (which includes most economists in my sample).

observations missing or rely on other biographical information to make inferences about these missing values. In the baseline specification, I chose the latter option to ensure, to the extent possible, that the full sample of top economists was maintained. Thus, specifically it is assumed that individuals graduated from their undergraduate institution at the age of 22.³⁵ In addition, it is assumed that the economist obtained an undergraduate degree from an institution in their country of birth. Since information on the undergraduate institution is seldom missing, this proved to be a useful way of identifying both missing values.

How reasonable is it to assume that the economist graduated from their undergraduate institution at the age of 22? Figure A1 provides information on the age at graduation, for those economists that have non-missing data on year of birth and year of graduation. 43% of these individuals graduated at the age of 22 and 79% graduated between the ages of 21-23. Thus, approximating the date of birth by subtracting 22 from the economist’s graduation year appears to be remarkably accurate.

FIGURE A1



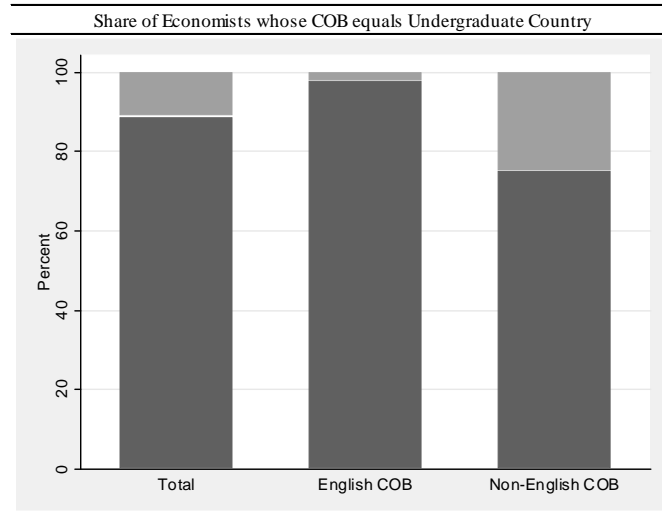
This histogram includes the top 2.5% of economists with non-missing date of birth and undergraduate year of graduation.

Next it is worth asking how reasonable it is to assume that the economist graduated from an undergraduate institution in their country of birth. Among those economists with non-missing values, Figure A2 shows the share of individuals whose country of birth matches the country of

³⁵For five economists whose undergraduate information was also missing, I inferred their date of birth by using their Ph.D. graduation year and assuming that they were 28 years old when they graduated.

their undergraduate degrees. The first column shows that 89% of economists graduated college in the country in which they were born. Column 2 shows that the share is 98% for those born in the six English speaking countries and 75% for non-English speaking countries. Overall, Figure A2 demonstrates that it is surprisingly common for economists to attend an undergraduate institution in their country of birth.

FIGURE A2



This bar chart reports the share of economists, with non-missing values, whose country of birth equals the country of their undergraduate studies. This share is reported for all economists in column 1, for those born in the 6 native English speaking countries in column 2, and for those born in all non-native English speaking countries in column 3.

Despite the fact that the assumptions used to approximate the age and country of birth of the economist seem remarkably accurate, it is worth investigating what would happen if these observations were left missing. Table A1 replicates Table 5 but leaves these values missing. Despite the drop in the number of observations, the results remain very similar. Specifically, English remains positive and significant in all three specifications. Furthermore, the magnitude of the coefficients, if anything, becomes larger in Table A1. Thus, the treatment of these missing values has no significant bearing on the results.

TABLE A1
Impact of English Proficiency on Components of Rank, Missing Values

	Total Rank	Papers	Citations
	(1)	(2)	(3)
English	125.56*** [20.84]	87.65*** [27.93]	163.47*** [25.82]
Controls	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes
Observations	794	794	794
R-squared	0.087	0.075	0.092

Robust standard errors clustered at the country of birth level in brackets. ***
p<0.01, ** p<0.05, * p<0.1. Missing country of birth and age observations are not filled.

A.3 Country of Birth of Top Economists

TABLE A2
Country of Birth of Top Economists

Country	WB Code	Country	WB Code
Algeria	LM	Lithuania	UM
Argentina	LM	Malaysia	LM
Australia	H	Mexico	UM
Austria	H	Monaco	H
Belarus	UM	Myanmar	L
Belgium	H	Netherlands	H
Brazil	UM	New Zealand	H
Bulgaria	LM	Norway	H
Canada	H	Pakistan	L
Chile	LM	Peru	LM
China	L	Poland	LM
Costa Rica	LM	Portugal	UM
Cuba	LM	Romania	LM
Cyprus	H	Russia	UM
D. R. of Congo	L	Rwanda	L
Denmark	H	Serbia	LM
Finland	H	Slovakia	LM
France	H	South Korea	UM
Germany	H	Spain	H
Greece	UM	Sweden	H
Hong Kong	H	Switzerland	H
Hungary	UM	Taiwan	H
India	L	Tunisia	LM
Iran	LM	Turkey	LM
Iraq	UM	UK	H
Ireland	H	USA	H
Israel	H	Uruguay	UM
Italy	H	Venezuela	UM
Jamaica	LM	Vietnam	L
Japan	H	Zimbabwe	LM
Lebanon	LM		

World Bank 1990 country classification data is used where countries are defined as Low Income (L), Lower-Middle Income (LM), Upper-Middle Income (UM), and High Income (H).

A.4 Score Instead of Rank

The baseline analysis uses the rank of an economist as the dependent variable of interest. However, perhaps the mapping of papers and citations data into rankings is affecting the distribution of economists and thus the results. Fortunately, in addition to the rank data, IDEAS/RePEc also provides the underlying "score" data which in this case refers to the weighted (i.e. quality adjusted) number of papers and citations. This section examines whether the results are sensitive to whether the rank or score variables are used.

Column 1-3 of Table A3 reports the results using total average score, papers score, and citations score as the dependent variables. The coefficients on English remain significant and positive in all of these specifications. However, unlike the rank data, the magnitude of the coefficients now vary due to the different underlying distributions of the dependent variables. This makes comparisons across these specifications difficult.³⁶

For comparison purposes, columns 5-6 standardize the paper and citation scores to have a mean zero and a standard deviation of one (the total score in column 4 is then the average of these two standardized variables). The coefficients on English in columns 4-6 are all still positive and significant but now the magnitude of the coefficients are also similar. We see that being a native English speaker increases the economists total score, papers score, and citations score by about 0.2 standard deviations.

Finally, columns 7-9 take the natural log of the total average score, the paper score, and the citation score. The results indicate that native English speakers have approximately 9% more quality adjusted papers and approximately 22% more quality adjusted citations. Overall, Table A3 indicates that the results are similar across a variety of different score measures and all of these findings are consistent with the baseline rank results.

³⁶For instance, the Papers Score variable has a mean of 94 while the Citations Score variable has a mean of 866.

TABLE A3
Impact of English Proficiency on Components of Score

	Total Score	Papers Score	Citations Score	Total (std)	Papers (std)	Citations (std)	ln Total	ln Papers	ln Citations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
English	107.56*** [23.78]	11.20*** [3.09]	203.92*** [46.01]	0.21*** [0.04]	0.18*** [0.05]	0.25*** [0.06]	0.20*** [0.04]	0.09*** [0.03]	0.22*** [0.04]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
COB WB FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,059	1,059	1,059	1,059	1,059	1,059	1,059	1,059	1,059
R-squared	0.061	0.071	0.058	0.074	0.071	0.058	0.074	0.084	0.071

Robust standard errors clustered at the country of birth level in brackets. *** p<0.01, ** p<0.05, * p<0.1. Columns 1-3 uses the unadjusted score data rather than the rank data as the dependent variables. Columns 4-6 standardizes these score variables to have a mean of zero and a standard error of one for ease of comparison. Columns 7-9 uses instead the ln of these scores as the dependent variable.