The Impact of English Language Skills on National Income: A Cross-National Comparison

Alex Ufier[a*]

[a] *Department of Economics, University of Oklahoma,.308 Cate Center Drive, Norman OK, 73019*

Abstract:

A common language lowers the transaction costs of international trade, and English is increasingly the language of international business. As a result, proficiency with English is often associated with higher incomes as well as increased employment, trade and other economic opportunities and is promoted as a policy to improve the wellbeing of people in developed and developing countries alike. However, it is not clear whether this effect is causal or just a correlation, as instruction in English or studying it as a new language is costly and may be associated with other bad outcomes that may negate some of its benefits. This paper estimates the impact of English skills as measured by the Test of English as a Foreign Language, TOEFL, and addresses the endogeneity problem using the difficulty of learning English given one's native language as an instrument. After accounting for several covariates and endogeneity, I consistently find a strong effect of English abilities on income and net exports. However, there is no effect of English on FDI or Emigration, suggesting that the impacts of English language emerge may come from the changing nature of domestic industries rather than remittances or foreign investment. This suggests improving English abilities may be a useful tool on the path to development.

*Email Address: alex.ufier@ou.edu, 610-470-3539

I) Introduction

While education and literacy have received a great deal of attention in economics, languages receives far less attention in the role it plays in economic development (Arcand and Grin 2012.) This subject deserves far more study, as a common language between two countries may raise incomes in each by facilitating trade and foreign direct investment via lower transaction costs. In addition, better foreign language abilities often open up more lucrative employment opportunities both within the country and for those emigrating (Lee 2012, Angrist and Lavy 1997, Lein et al 2011, Melitz 2007, Meltiz and Toubal 2014.) Language specific transaction costs are more important for services than for manufacturing, as services generally imply more interpersonal interactions. This suggests that the structural changes inherent in the development process, moving from agrarian to manufacturing to service industries as countries grow richer, may reinforce the importance of improving foreign language skills in the future (Warschauer 2000.)

English has generally become this de facto international language (Nunan 2003.) English is the common language for business, advertising, numerous academic fields, media, and numerous other fields (Crystal 2003.) As a result, many developing countries have begun to teach English as a part of school curriculums, either as standalone subjects or by changing the language of instruction to English in hopes of raising international trade with a better prepared workforce. At a micro level, citizens often work to improve their skills in English in hopes of raising their wages, such as seeking better employment in the service sector like call centers (Nunan 2003, Casale and Posel 2011, Chakboraty and Kapur 2009, Birdsall 2011.) USAid, the US governmental agency responsible for foreign development work, promotes English education in a number of developing countries. In a world with ever increasing international competitive

pressures, English abilities are proposed to be an important component of participating in the global economy (Wedell 2008, and Warschauer 2000.)

However, the benefits from better English skills are currently not well understood. While micro studies find individuals with better language skills earn more, macro level studies are scarce. The benefits of learning English to an individual are readily identified- higher paying employment- the costs may prove to be too high as one is forced to choose between language skills and other forms of human capital (Wedell 2008.) The focus on English language skills may be to the detriment of other useful skills, and this may negatively affect poorer students without opportunities to put English skills to use the most (Bruthiaux 2002.) While having a shared language definitely increases trade volume, which hopefully leads to higher incomes, it is not clear that English itself plays a special role in international trade. Hejazi and Ma (2011)finds that two countries that both speak English have more bilateral trade than two countries who have a similar shared language that is not English, while Melitz (2007) and Melitz and Toubal (2014) find that English plays no special role compared to other common languages.

Empirically identifying the impact of language skills is problematic, as it is not clear if English language skills lead to higher incomes and trade, or higher incomes and trade leads people to study English. From a policy perspective, English education is costly and it should be promoted if it can be shown to improve outcomes, as some micro level studies have shown; otherwise, resources may be better spent elsewhere. To answer whether on a macro level English skills improve incomes and employment opportunities, I employ an index of English language skills, the Test of English as a Foreign Language, or TOEFL, to estimate the impact of English language proficiency on income per capita, net exports, FDI and emigration over the period of

1992-2012 for 139 countries.[1] I use an instrumental variables framework to obtain a consistent estimate of English skill's effects. The instrument, hours required to learn English conditional on a country's official language, is correlated with English skills as it represents how much effort would be required to learn English. However, aside from this channel, official language is not correlated to any of the outcome variables of income, net exports, FDI and emigration.

I find that while the impact of English abilities on income and net exports is positive using OLS, the effect is even larger after using instrumental variables. This effect is robust to several different ways of parameterizing the instrumental variables. There is no such observed effect of language abilities on emigration or FDI. This indicates that English abilities are correlated with higher incomes in part through higher net exports, and this is not driven by either foreign investment or remittances. This would support the story that domestic industries are better able to find profitable export markets for their products and services when their employees have higher English abilities, or countries with higher English abilities are able to go into more profitable export-based industries. Better English abilities may also allow countries to have access to better technology, raising incomes through a channel other than net exports. This shows the promise of English language education in a positive light, as higher English abilities opens countries to more markets and ultimately higher standards of living.

This paper is organized as follows: section I outlines the goals of this paper, section II introduces the relevant literature, section III discusses the estimation strategy, section IV discusses the data, section V discusses results, and section VI concludes.

II) Literature Review

[1] Data availability varies by year: for the 2012 year that gets a large portion of the analysis, 139 countries have data available on TOEFL scores. 120 are available in 1992, the start of the period. 173 countries have TOEFL data available for some years. Data is available from the author upon request.

USAid in particular promotes English education in its development work, in recent years advertising its English education programs as success stories in development. For example, engineers in Kabul taught English are now able to employ electronic control equipment, whose use requires one to know English, instead of manual control equipment to run the local hydroelectric power plant. With their better English abilities, they are also able to seek specialized education abroad. Because of access to better education and technology through higher English skills, there are now fewer accidents and service outages (USAid 2013.) This argument would lend credence to the idea that higher English abilities would allow workers in the developing world to become more productive, as they are able to use better equipment and have better access to information. USAid assisted Rwanda's changing of its official language of instruction from French to English, a change intended to raise incomes and improve international trade (USAid 2014.) USAid claims its English training programs in countries such as Ethiopia and the Philippines are a means of improving the lives of people in these countries by broadening job opportunities, going so far as to draw fire from US lawmakers fearing the development program will cause some service jobs such as call centers to locate in these developing countries instead of the US (Koch 2014, De Lotbiniere 2012, May 2012.) English education is thus considered by some development agencies as a useful and effective tool, and powerful enough to arouse the concern of protectionist lawmakers.

Despite all these promised benefits, the true effects of higher English language abilities and English language education are unclear. There is mounting evidence that countries promoting instruction in English may also be exacerbating income inequalities as well as other problems. Glewwe et al. (2009) found that the use of English-language textbooks in primary school was useless to most students in Kenya. Students lacked the ability to read them with any

real level of comprehension. Only high achievers and wealthier students saw any benefit from receiving these textbooks. Bruthiaux (2002) warns that a focus on English language education may worsen income gaps in the developing world. Only the rich will have the resources to benefit from learning English or being instructed in English, as the poor would never have access to the networks and international opportunities that would make learning English worthwhile. Poorer students may have better earning potential if they received instruction in their native language, with more emphasis on non-language related subjects such as mathematics, science, trades, or writing skills in their native language. This highlights a tradeoff between more general human capital, and English specific human capital, as well as the conflict between promoting a more internationally useful language against one that would preserve local culture and prepare students for domestic-focused employment (Prah 2008.)[2]

There is some literature on the microeconomic returns to language skills, specifically English, on income, that will guide the empirics of this paper. Their findings and methodology are difficult to generalize to the national level due to data availability, data heterogeneity, and choice of control variables. There also exist several papers using gravity models to study the impact of common languages on trade that, while econometrically different from the approach taken here, are instructive with their choice of control variables and parameterization of linguistic similarity.

Several microeconomic studies have found a positive impact of English skills on individual earnings in India and South Africa (Azam et al 2013, Casale and Posel 2011.) Other

[2] Many developing nations face the challenge between teaching their native languages, or teaching a more widely spoken foreign language, often English (Bambgbose 2009, Wolhuter et al 2006.) Wedell find that educators worldwide are dissatisfied with this tradeoff, being forced to choose between subjects without adequate information about their payoffs (2008.) In order to improve literacy rates in sub-Saharan Africa, some countries have found success in switching from teaching a colonial language- French or English- instead instructing in in the native language of students (Bamgbose 2009.) English language skills deserve more empirical study as educators and policymakers alike want to maximize opportunities of students through the choice of curricula. Due to these expensive tradeoffs and even potential downsides to English instruction and study of the English language, the recent moves of some countries, such as the Gambia, away from English are now far more understandable.

studies have considered immigrants in the United States (Kossoudji 1988.) However, these studies and likely any study comparing English skills and incomes are inherently complicated by endogeneity problems, as language abilities affect income and income affects language abilities. Lang and Sinvier show that returns to English language skills in the form of higher incomes, even in countries where English is not an official language (2006.)

This pattern is also not limited to the English language. Languages may affect trade or employment opportunities, as seen in Angrist and Lavy's (1997) study of French in Morocco and Rendon's (2007) study of Catalan in Catalonia. Angrist and Lavy's paper use an instrumental variables regression to explore the switch from French to Arabic instruction as the language of official instruction in Morocco. This change in instruction reduced student's French writing skills, which in turn limited student employment opportunities and as a result lowered income. Silvio created a theoretical model of language learning, which demonstrated that knowledge of more languages led to higher paying job opportunities in Catalonia. These studies are microeconomic in nature, and so a macroeconomic study may find different results even if one uses a similar methodology. For example, microeconomic literature going back at least as far as Mincer (1974) generally find positive returns to education on an individual level, but macro studies often finds no effect of countries increasing average levels of education on incomes (Pritchett 2001.)

Other papers have considered the interaction between language skills and bilateral trade. Egger and Lassman (2013) use a case study in Switzerland to find that areas with more linguistic similarity tend to trade more, ascribing this effect to cultural similarity. Lein et al. (2011), using a gravity model, find that shared languages lead to both an increase in bilateral trade and FDI inflows. They propose choosing to study a language is an optimization decision where

participants maximize benefits of such education conditional on its costs. Molnar (2013) looks at the role translation costs play in international trade, finding a significant effect on bilateral trade. Finally, Egger and Lassman's (2012) survey and meta-analysis of the literature consistently finds that linguistic similarity is a major determinant of trade, even once controlling for factors such as legal origin, colonial history, and exchange rates.

Looking more directly at the role played by English, Hejazi and Ma (2011) employ a gravity model to find if English has a premium as a shared language over other common languages among OECD countries. They find results that agree with previous authors in that shared languages are positive in their effect on trade, but that sharing English has an even greater effect. Melitz (2007) divides linguistic similarity into an extensive, whether the country pair has a shared language, as well as intensive margin, how many people in those countries share that language. While both effects are positive, the intensive margin's effect is stronger. However, while shared languages matter, and shared European languages as a whole have a stronger effect on bilateral trade, English itself has no larger positive marginal effect on bilateral trade than other European languages. Melitz and Toubal (2014) extend this paper, using a more complex measure of linguistic proximity, and find similar results. Thus, while common languages raise bilateral trade, whether English plays a special role is unclear.

So far, such a comprehensive study of English language skills and education worldwide has been hampered by lack of data. While an international database on English language education would be ideal, with information on time and money committed to the subject of English language study by country as well as their languages of instruction, it currently does not exist. A few regional surveys exist, such as the one Nunan (2003) performed in East Asia. Even if the data did exist, such cross-national comparisons are likely limited due to heterogeneity in

education as well as data quality. Studies looking at the impact of English education may thus be limited to considering only one country.

Due to these limitations, this paper considers national English language skills rather than English language education. This allows me to evaluate the value of the end product of English language education, the stock of English language skills.[3] To this end, I use data from the TOEFL, a measure of academic English language skills, to proxy for national level English language abilities. I will discuss estimation of these effects, the challenges posed, and possible solutions in the next section.

III) Estimation Strategy

In order to obtain consistent estimates of the impact of English skills on per capita income, one must solve the problem of endogeneity: Outcome Y_i, most importantly income but also net exports, emigration, and FDI in this paper, is a function of English ability E_i. This higher ability enables easier communication, translation, and access to technology and information. However, English abilities are also a function of income of these other outcome variables. Countries with higher income likely have an easier time improving their language abilities, and inhabitants of countries already experiencing a great deal of trade or foreign investment are likely to seek to improve their language abilities.[4] Alternatively, very poor countries may be the only ones selecting to improve English abilities. Both the outcome variables and language abilities are also likely influenced by some historical elements. The process is likely to be self-reinforcing, with changes in one also affecting the other. The model can be described as follows:

[3] Whether a particular education program or method produces the most human capital per dollar spent is beyond the scope of this paper, as this paper only considers the total stock of capital present in a country.

[4] Because the cost of learning a language is in large part time, the opportunity cost for people to learn a language in poor countries with poor employment opportunities may actually be lower than that of wealthier countries relative to their outside options. People have after all become multi-lingual throughout history, when even wealthier people by that era were quite poor by modern standards.

$$(1) \quad Y_i = A_0 + A_1 E_i + A_2 X_i$$

$$(2) \quad E_i = B_0 + B_1 Y_i + B_2 X_i$$

Where X_i is a vector of common exogenous covariates. Using OLS, the estimate A_1will be upwardly biased and inconsistent due to this endogeneity if there is self-reinforcement with richer countries focusing more on English abilities, or downwardly biased if improving English abilities is a treatment only undertaken by already poorer countries. Trying to estimate A_1 consistently will thus require a different estimation method, such as instrumental variables.

Previous research on the impact of language skills on individual level income has employed instrumental variable estimation. Akbulut-Yuksel et al (2011) as well as Bleakley and Chin (2004) employ an individuals' age of immigration as an instrument for English ability, as younger children are better able to learn a new language than adults, and age of immigration itself should have no direct impact on income. Dustmann and Van Soest (2002) as well as Rooth (2001) employ parental education as an instrument, as it should be correlated with language ability but not directly with income. Chiswick and Miller (1995) employ country of origin and marriage status as instruments, as married immigrants with a spouse who speaks the same language will have less incentive to learn the language of his new country. However, marriage status and country of origin may still be correlated with income. Finally, Chakboraty and Kapur (2009) use a policy change on English instruction in an Indian state as an instrument. As this paper considers cross national variation rather than cross individual variation, different instruments will be needed from those previously considered, as these instruments cannot be generalized to a national level.

A more feasible instrument for this paper, suggested by Shastry (2007) and also implemented by Isphording and Otten (2014), is linguistic distance. Shastry considers the fact

that many provinces in India speak languages very linguistically different from Hindi, and others speak ones that are more similar. Learning Hindi is more common in the latter group, while learning English is more common in the former. Worldwide, some languages are linguistically close to English and this makes it easier for native speakers to learn English compared to native speakers of other languages. Thus, one's native language should raise the time required to learn English and this increase in the time required to learn English should be exogenous. As a result, linguistic distance from English should be negatively correlated with English language skills.[5] Isphording and Otten (2014) find that this increased difficulty of learning English can be considerable depending on their original languages and leads to easily observable differences in incomes of immigrants. However, these studies can be hampered by a lack of concrete measure of distance, with Shastry (2007) relying on estimates of how technically different languages are to learn as classified by linguists and Isphording and Otten (2014) employing a method of linguistic similarity based on comparing grammatical structure and cognates.

In this paper, I employ a source that gives quantitative measures of difficulty for native speakers of one language to learn English. The Foreign Service Institute (FSI) produces information on how many hours of instruction are required for a native English speaker to learn a foreign language, and assuming this time is symmetric, as assumed by Chiswick and Miller (1999), it will also give an unbiased estimate of the time it would take a given native language speaker to learn English. Languages are classified in 7 categories of difficulty, going from requiring 600 hours of study to over 2200 hours of study, with a summary of hours provided on Table 1. Adding in linguistic distance, D_i the estimation of the earlier system of equations (1)

[5] There is no reason to expect linguistic distance from English will impact income or trade other than through the channel of ease of learning English, as there is nothing inherently different between languages' ability to communicate ideas or structure that should affect income. Crystal (2003) argues that the supremacy of English in international trade, business and media is entirely accidental due to the course of history. It is plausible that an alphabet-based language, like English, is superior to a pictograph based one, such as Chinese, due to the ease of which it can change and add new ideas. But Romanization of Chinese through pinyin, for example, illustrates that a pictographic language can be converted to an alphabet based one. Additionally, as children learn to speak their first language at roughly the same age, from an equal starting point no language is more difficult to learn or communicate with than any other (Crystal 2003.)

and (2) thus becomes:

$$(3) \quad Y_i = A_0 + A_1 E_i + A_2 X_i$$

$$(4) \quad E_i = B_0 + B_1 Y_i + B_2 X_i + B_3 D_i$$

While one cannot identify the English language equation (4), one can identify the outcome equation (3), using linguistic distance as an instrument for English language skills to obtain a consistent estimate of their impact on per capita income, trade as a share of GDP, or emigration.

This proposed functional form of this instrumental variable strategy requires two assumptions that may be considered restrictive. First, while the costs of learning a language may not be linear, the instrument would require this relationship to be linear, or at least require a fixed functional form. For example, in a linear form, languages that take 800 hours to learn require half the resources of a language that takes 1600 hours to learn. This is not necessarily true. Second, the assumption of symmetry may be inaccurate due to the complexity of English. While an English speaker may have an easy time learning Spanish in 600 hours, a Spanish speaker may take more hours to learn English. However, the difficulty of learning English would thus come from two sources- its overall difficulty as a language, which should be a constant regardless of one's native language, and its closeness to another language, which varies based on one's native language. Spanish speakers should thus find it easier to learn English than Mandarin speakers, which should preserve the ordinal rank of the time required to learn English for other language speakers, even if it does not preserve the cardinality of difficulty.

Both of these concerns can be addressed by using a series of dummy variables for the different categories of language difficulty instead of the hours required to learn them, as this does not restrict the linguistic distance to be linear in hours required to learn English, or require it to

take the same number of hours to go from English to Spanish as it does to go from Spanish to English. This instrumental variables strategy is parameterized in two ways. In one specification, each of the seven categories on Table 1 receives their own dummy variable, with native English speaking countries being the eighth category omitted for perfect co-linearity. This set of dummies is then used as the instrument. However, many of these categories have very few languages or include very few countries, so I also employ specification that reduces the number of categories dummy variables to three (four including omitted native English for perfect co-linearity.) The first dummy category is equal to 1 if a native language speaker of a country takes 600, 750, or 900 hours (first three categories of Table 1) to learn English and 0 otherwise, the second 1 if it takes 1100 or 1100+ (fourth and fifth categories) to learn English and 0 otherwise, and the third category 1 if it takes 2200 or 2200+ (sixth and seventh categories) to learn English. As a robustness check, I use a specification created from a separate measure of linguistic similarity based on technical proximity of words, the Linguistic Proximity-2, or LP2 variable, used by Melitz and Toubal (2014.) Higher values of this variable represent more linguistic proximity, and lower values less proximity.

Table 2 presents basic correlations that are consistent with these exclusion restrictions for the 20 years of data available, 1992-2012 excluding 2011. TOEFL is positively correlated with natural log per capita real GDP at .3968, and TOEFL scores are negatively correlated with hours to learn a language at -.2837. Hours to learn English are negatively correlated with per capita income at .0376. This evidence would support omitting linguistic distance from the income equation, and suggesting it is an appropriate instrument for English abilities in that same equation. LP2 has a stronger correlation with score than hours to learn, .4972, but also is highly correlated with the outcome of income at .5382.

IV) Data

The TOEFL is an assessment of a student's ability to use English in the classroom, considering four skills: reading, listening, speaking and writing, and I use this assessment as a measure of average English abilities across countries being studied in this paper. 27 million students having taken the test to date in over 4500 testing sites and across 165 countries since it was organized in 1964, giving a broad base of support for the test scores. TOEFL exams were given as a paper test for much of its history, switching to local computer based examinations in mid-1998 and then an internet based examination in mid-2005.

The most recent version of the TOEFL has a number of advantages. A major strength of the TOEFL is that it allows the scores to be comparable internationally. For internet based tests, students should be facing roughly the same testing environment; although paper or local computer based testing students may face a different environment. It has a broad history, with data available from 1992-2012 for most countries, allowing one to construct a relatively balanced panel or observe variation over time. Finally, the examination is generally accepted to be the best assessment of English language abilities that is widely available, as measured by the number of institutions who accept it. TOEFL scores are accepted by over 9000 colleges. This is more than its largest alternative, the International English Language Test IELTS at 8000, which also has a shorter history, fewer test takers, fewer testing centers, and presence in fewer countries, although the difference is relatively small.

The TOEFL index has several shortcomings, none of which are critical. First, there is the incentive to cheat or otherwise compromise the integrity of the test. TOEFL scores are usually a major requirement for admission to an English speaking university and so high scores in a

country may either represent higher ability or better cheating. This would introduce an attenuation bias, biasing towards the effects of English language abilities on income or other outcomes towards zero. If one observes a positive effect, it is possible the true effect is even higher as the current on is observed with attenuation from cheating. However, cheating should be more difficult on internet based tests than on paper exams, where the central testing body has more oversight on how the test is executed and how its results are transported for evaluation.

Second, while test takers may self-select, it is not clear in which direction this selection would bias results. Only more competent English speakers will be motivated to take an examination. But since these are national level aggregates, one would require nation-specific differences in the personal selection into the programs, and selection to take the exam is likely determined on an individual rather than national level. Thus, even if only the best English speakers in each nation select to take the exam, one would still be comparing the scores of the best individuals across nations, which would also be the individuals most relevant to conducting international business interaction. A variety of skill levels are present in the index, and most countries are represented, so it appears countries with poor language abilities are not avoiding the exam altogether.

Finally, the index shows the average English proficiency of a test taker, a quality measure, not the share of inhabitants that speak English, a quantity measure. However, if one can determine that average English proficiency is correlated with share of inhabitants that speak English, one could claim that the TOEFL is a good summary of both quantity and quality of English speakers in a country. Data on English speakers are available from a number of sources such as national censuses as well as Crystal (2003). Using this data, I can compare English speakers as a share of population to English proficiency for about half the countries with TOEFL

data available. Table 1 shows correlations that indicate it is indeed the case that the share of population that speaks English is correlated with the TOEFL at .3321.

Table 3 presents summary statistics for the variables used in the regression. Note that there is a reasonably large range of abilities for the TOEFL. Countries that have multiple official languages with different linguistic distances from English according to the FSI use the value of the language most distant from English among all their official languages. Additionally, following Meltiz and Toubal (2014), I include a control dummy, equal to 1 for a country having English as one of their official languages, in order to capture the effect of official endorsement or presence of English aside from the actual abilities of a country.

Left hand side variables include net exports(exports minus imports as a share of GDP). FDI(net inflows from foreign countries as a share of GDP), natural log of real GDP per capita, and emigration rate data from the OECD from 2000-2008. This emigration data includes 200 countries of origin and 38 OECD and affiliated countries as destinations, but is not available as a panel.

For covariates, a natural and important inclusion would be an overall measure of education. There is a great deal of interaction between language ability and overall education level, with well-educated individuals best being able to leverage the benefits of English in other studies (Azam et al. 2013.) Barro and Lee (2013) produce a measure of overall education, updated in five year windows, that I employ for this study. I use a weighted average of this education measure for the years where measures of education are not present, so if education values are available 1995 and 2000, 1996 uses 20% of the value from 2000 and 80% from 1995, 1997 uses 40% of the value from 2000 and 60% from 1995, and so on.

Persistently high inflation may negatively affect real incomes, so I include inflation as an independent variable. Government size as a share of GDP may also impact long term growth, so this is also added as a right hand variable. As I am analyzing linguistic factors, it is sensible to include a measure of ethnolinguistic fractionalization, as countries with more languages may have lower incomes as well as face different incentives to learn English (Arcand and Grin 2012.)

Geographic factors such as latitude, landlocked status, or island status may also affect incomes or other outcomes being considered. Easterly (2001) and Bloom et al (1998), among others, consider these geographic features to be potential determinants of income. The size of the country may also matter, as larger countries have lower incentives to learn non-native languages as they have many more commercial opportunities within their borders, so I include natural log of population as well as land area as a covariate. I also include continent level dummies with Europe being the omitted continent due to perfect co-linearity.

Ethnolinguistic fractionalization comes from Roeder (2001), latitude data is from the CIA World Factbook, and continents are specified by the ETS in their TOEFL publications. All other covariates come from the World Development Indicators.

The linguistic distance instrument, hours to learn English as listed by the FSI, is parameterized in two different ways as discussed earlier. I also employ Meltiz and Toubal's (2014) LP2 measure as a robustness check.[6]

V) Results

Table 4 shows the impact of English language abilities and various covariates on GDP per capita. The results here are given with standardized betas, where a 1 standard deviation

[6] Several countries lack official languages that are featured on the FSI list of language learning difficulty. The number of hours to learn English by nearest linguistic neighbor is used instead. Additionally, for countries with multiple official languages of varying linguistic distance from English, the language with the largest distance from English is used.

change in the X variable leads to a B (listed in table) standard deviation change in the Y variable. The OLS specification (1) and (2) shows a positive and significant effect of English abilities on incomes, but the magnitude drops substantially once covariates are included. The only covariate value that is particularly surprising is the negative effect of official English status in a country. Once controlling for abilities, officially recognizing English is negative in its effect on income, showing that official endorsement of English is not enough to raise incomes and abilities are what matter.

All three IV specifications, specification 1 breaking linguistic distance into three categories, specification 2 breaking it more finely into seven categories, and the LP2 measure using a continuous measure of linguistic distance, show a positive and significant effect of English language abilities on incomes after accounting for covariates. This effect is also much larger than the effect observed using OLS. Depending on the specification a one standard deviation increase in TOEFL scores raised log income by between .358 and .768 standard deviations. Presented using non-standardized coefficients, raising the average TOEFL score of a country by 1 point on a 120 point scale raises its real per capita incomes by between 6 and 13 percent.

These findings could indicate either that the IVs corrected for an attenuation effect, as TOEFL scores may not accurately English abilities, or that lower income countries were the ones seeking to improve their English abilities to better compete internationally. Both stories are consistent with the increased size of the coefficient on the TOEFL score. The argument that higher income countries choose to improve, or would find it easier to improve, their English abilities is either inconsistent with this evidence or dominated by the other two effects. Otherwise, IV results should have a smaller coefficient than OLS results. The Hausman test

strongly rejects the null of no difference between OLS and IV in all cases, indicating a statistically significant systematic difference between IV and OLS estimations in all IV specifications.

First stage results for IVs in Table 4a are unsurprising given these findings. The coefficients for dummies in specifications 3-4 are steadily decreasing as the hours needed to learn English increases, indicating more linguistic distance from English leads to lower TOEFL scores. They are decreasing less steadily for specifications 5-6, possibly because several of the categories contain a small number of countries. Similarly, for 7-8, where higher values of LP2 mean higher levels of linguistic proximity, the correlation is positive.

Having established that English is correlated with higher incomes, I now explore channels by which this effect may work. As higher English abilities is supposed to open more export markets due to ease of communication, one may expect higher net exports to be a natural result of higher English language abilities. Table 5 shows exactly this, following a similar pattern as seen in Table 4 with income. There is a positive effect of TOEFL scores on net exports both with and without covariates in the OLS specifications (1-2). All IV specifications with covariates (4, 6, 8) show the same positive, statistically significant, and statistically different from OLS (as by the Hausman test) effect of TOEFL scores on net exports as a share of GDP. This effect is comparatively large, with a 1 standard deviation change in TOEFL scores raising net exports as a share of GDP by between .6 and 1.3 standard deviations. In unstandardized terms, a 1 point increase in average TOEFL scores on a 120 point scale raises net exports between .8 to 1.9 percentage points as a share of GDP. First stage results in Table 5a are similar to Table 4a. This would agree with Hejazi and Ma (2011), finding that English is significant in generating trade specifically, as well as numerous authors finding shared languages raise trade.

While the evidence presented so far would promote the idea that higher English abilities are associated with higher incomes, possibly through higher net exports from domestic industries or possibly through access to better technology directly raising productivity or incomes, other channels could be at work. International corporations may take advantage of higher English skills in these countries, leading to a boost in foreign direct investment (Lein et al 2011) or increased migration to opportunities abroad that leads to higher remittances albeit with loss of human capital from the original country (Gupta et al 2008.)

Table 6 looks at the impact of TOEFL scores on FDI inflows as a share of GDP. For all three instrumental variable specifications as well as OLS, there is generally no effect of TOEFL scores on FDI after accounting for covariates. For specification (4), there is a negative impact of TOEFL scores on FDI, but the Hausman test fails to reject the null of no systematic difference between the estimations. There is no evidence that higher TOEFL scores leading to increased FDI. It is possible that higher incomes and net exports may come from domestic industries taking advantage of more international export opportunities as well as having easier access to more information and better technologies, rather than from foreign investment. Foreign corporations may have criteria other than English ability in mind for locating a production facility. Table 6a shows similar results to previous first stage estimation results.

Finally, Tables 7 and 8 consider the impact of English ability on emigration. Unlike previous variables, panel information for emigration is not available, I use cross sectional aggregate migration from 2000-2008 from the World Bank. I consider two cross sections using this migration data, one with covariates at the start of the dataset, 1992, and the other using covariates at the end, 2012. Table 7 shows the 2012 specification results which show, once accounting for covariates, there is generally no effect of TOEFL scores on emigration save

specification (4), where it is negative. However, there is no systematic difference between IV and OLS estimates according to the Hausman test even in this case. Table 7 shows similar results using a 1992 cross section of determinants instead of 2012 cross section, across the board finding no effect of TOEFL scores on the emigration stock. Tables 7a and 8a show first stage results which, while less significant than previous tables due to the smaller sample size, are still roughly similar in direction and magnitude as table 4a seen earlier.

The net result of these regressions is that English abilities are indeed statistically significantly correlated with several important outcomes. Countries with higher TOEFL scores tend to have higher incomes, with a 1 point increase in national average TOEFL scores on a 120 point scale being associated with a 6 to 13 percent increase in real per capita income. This could be due to access to better technology that raised worker productivity, lower transaction costs, or more export opportunities. There is evidence of the latter, with net exports as a share of GDP increasing by .8 to 1.9 percentage points for a 1 point national average TOEFL score increase. Using these parameters as baseline estimates, policy makers can decide whether the particular costs of improving English abilities are worth these stated benefits. However, when looking at other potential channels, such as FDI or emigration, there are no such effects. Higher TOEFL scores do not lead to more foreign direct investment, and they do not increase emigration. The higher income and exports likely come from domestic industries being able to better compete internationally, possibly due to lower communication costs and easier access to information and technology.

VI) Conclusion

In an increasingly globalized world, a common language greatly simplifies international trade and business, lowering transaction costs and making it easier to invest in foreign countries, with the end result of raising incomes. Due to historical accident, English has arisen as this common international language. Many countries have begun to teach English to students in hopes of giving them a competitive edge, after seeing that English language skills are correlated with higher incomes and more employment opportunities. However, it is not clear whether this relationship is causal; higher English skills may allow an individual to earn higher income, or higher income may lead to one seeking more English skills. Any association of English language skills with income may thus be spurious. Even if there is a causal relationship between English skills and income or trade, the effects of a dollar spent on acquiring English language skills may be no different from acquiring other forms of human capital through education.

To answer the question as to whether English language skills actually lead to higher national incomes, or more exports, FDI, or emigration, this paper employs data on English test scores from an international education testing companies' records. To account for the fact that income and English proficiency are endogenous, I use the difficulty of learning English as an instrument for English language skills, as it should be correlated with average English skills, but uncorrelated with income. There is a positive effect of English abilities on income and net exports, but not on FDI or emigration. This suggests that English abilities raise incomes and exports but not necessarily through higher foreign investment or remittances, perhaps either due to better access to technology or fundamental changes in domestic industries. However, while this paper is able to estimate some of the effects of national level English abilities, it is unable to identify the exact costs and benefits from investing in English language education, and I leave the analysis of marginal tradeoffs to future papers.

Works Cited

Adams Jr, Richard H., and John Page. "Do international migration and remittances reduce poverty in developing countries?" *World development* 33.10 (2005): 1645-1669.

Akbulut-Yuksel, Mevlude, Hoyt Bleakley, and Aimee Chin. "The Effects of English Proficiency among Childhood Immigrants: Are Hispanics Different?" *Latinos and the Economy* (2011): 255-283.

Angrist, Joshua D., and Victor Lavy. "The Effect of a Change in Language of Instruction on the Returns to Schooling in Morocco." *Journal of Labor Economics* (1997): 48-76.

Arcand, J. L., and F. Grin. "Language in economic development: Is English special and is linguistic fragmentation bad?" *English and international development.* Bristol: Multilingual Matters (2012).

Azam, Mehtabul, Aimee Chin, and Nishith Prakash. "The returns to English-language skills in India." *Economic Development and Cultural Change* 61.2 (2013): 335-367.

Bambgbose, Ayo. "Language of instruction policy and practice in Africa." Dakar, Senegal: Regional Office for Education in Africa, UNESCO. Retrieved June 5 (2004): 2009.

Barro, Robert J., and Jong Wha Lee. "A new data set of educational attainment in the world, 1950–2010." *Journal of development economics* 104 (2013): 184-198.

Birdsall, Nancy. “The New Education Strategy at the World Bank: Time for a Millennium Learning Goal?” Center for Global Development Blog post, May 3rd 2011.

Bloom, David E., et al. "Geography, demography, and economic growth in Africa." Brookings Papers on Economic Activity 1998.2 (1998): 207-295.

Bleakley, Hoyt, and Aimee Chin. "Language Skills and Earnings: Evidence from Childhood Immigrants." *Review of Economics and Statistics* 86.2 (2004): 481-496.

Bruthiaux, Paul. "Hold your courses: Language education, language choice, and economic development." *TESOL quarterly* 36.3 (2002): 275-296.

Casale, Daniela, and Dorrit Posel. "English language proficiency and earnings in a developing country: The case of South Africa." *The Journal of Socio-Economics* 40.4 (2011): 385-393.

Chakraborty, Tanika, and Shilpi Kapur. "English language premium: Evidence from a policy experiment in India." (2009). Washington University in St. Louis unpublished paper (2008).

Chiswick, Barry R., and Paul W. Miller. "The endogeneity between language and earnings: International analyses." *Journal of Labor Economics* (1995): 246-288.

Chiswick, Barry R., and Paul W. Miller. "Language skills and earnings among legalized aliens." *Journal of Population Economics* 12.1 (1999): 63-89.

Crystal, David. English as a Global Language. Cambridge University Press, 2003.

De Lotbiniere, Max. "Threat to US jobs shuts USAid training." Guardian Weekly, May 15 2012.

Dustmann, Christian, and Arthur Van Soest. "Language and the Earnings of Immigrants." *Industrial and Labor Relations Review* (2002): 473-492

Easterly, William. The Elusive Quest for Growth: Economists Adventures and Misadventure in the Tropics. The MIT press, 2001.

Egger, Peter, and Andrea Lassmann. *The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design*. Centre for Economic Policy Research, 2013.

Egger, Peter, and Andrea Lassmann. The language effect in international trade: A meta-analysis. Economic letters 116 (2012) 221-224.

Fink, Carsten, Aaditya Mattoo, and Ileana Cristina Neagu. "Assessing the impact of communication costs on international trade." *Journal of International Economics* 67.2 (2005): 428-445.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. "Many children left behind? Textbooks and test scores in Kenya." *American Economic Journal: Applied Economics* 1.1 (2009): 112-135.

Gupta, Sanjeev, Catherine A. Pattillo, and Smita Wagh. "Effect of remittances on poverty and financial development in Sub-Saharan Africa." *World Development* 37.1 (2009): 104-115.

Hejazi, Walid, and Juan Ma. "Gravity, the English language and international business." *Multinational Business Review* 19.2 (2011): 152-167.

Isphording, Ingo E., and Sebastian Otten. "Linguistic barriers in the destination language acquisition of immigrants." *Journal of Economic Behavior & Organization* 105 (2014): 30-50.

Koch, J. "Bringing Improved English teaching to scale: The USAID TELL Program " Paper presented at the annual meeting of the 55th Annual Conference of the Comparative and International Education Society, Fairmont Le Reine, January 10 2014.

Kossoudji, Sherrie A. "English language ability and the labor market opportunities of Hispanic and East Asian immigrant men." *Journal of Labor Economics* (1988): 205-228.

Lang, Kevin, and Erez Siniver. "The return to English in a non-English speaking country: Russian immigrants and native Israelis in Israel." *The BE Journal of Economic Analysis & Policy* 9.1 (2009).

Lee, Chew Ging. "English language and economic growth: Cross-country empirical evidence." *Journal of Economic and Social Studies* 2.1 (2012): 5-20.

May, Caroline. "USAID training foreign workers for English speaking jobs." Daily Caller April 20, 2012.

Melitz, Jacques. "Language and foreign trade." *European Economic Review* 52.4 (2008): 667-699.

Melitz, Jacques, and Farid Toubal. "Native language, spoken language, translation and trade." *Journal of International Economics* (2014).

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

Molnar, Alejandro. "Language barriers to foreign trade: evidence from translation costs." (2013).

Nettle, Daniel. "Linguistic Fragmentation and the Wealth of Nations: The Fishman-Pool Hypothesis Reexamined." *Economic development and cultural change* 48.2 (2000): 335-348.

Nunan, David. "The Impact of English as a Global Language on Educational Policies and Practices in the Asia-Pacific Region." *Tesol Quarterly* 37.4 (2003): 589-613.

Oh, Chang Hoon, W. Travis Selmier, and Donald Lien. "International trade, foreign direct investment, and transaction costs in languages." *The Journal of Socio-Economics* 40.6 (2011): 732-735.

Prah, Kwesi Kwaa. "The language of instruction conundrum in Africa." Meeting on the Implications of Language for Peace and Development (IMPLAN). Retrieved April. Vol. 4. 2008.

Pritchett, Lant. "Where has all the education gone?" *The World Bank Economic Review* 15.3 (2001): 367-391.

Rendon, Sílvio. "The Catalan premium: language and employment in Catalonia." *Journal of Population Economics* 20.3 (2007): 669-686.

Rooth, Dan-Olof. "The effect of language proficiency on employment for immigrants–the case of Sweden." Submitted for publication (2001.)

Shastry, Gauri Kartini. "Speaking English in a Globalizing World: Information Technology and Education in India." IZA/World Bank Conference Employment and Development, Bonn. 2007.

Song, Jae Jung. "English as an official language in South Korea: Global English or social malady?" *Language Problems & Language Planning* 35.1 (2011): 35-55.

USAID, “English for Engineers in Afghanistan” USAID Leadership Success stories, December 06, 2013

USAID, “Rwandan Teachers Switch to English Instruction” USAID Leadership Success stories, March 06, 2014

Wedell, Martin. "Developing a capacity to make “English for Everyone” worthwhile: Reconsidering outcomes and how to start achieving them." *International Journal of Educational Development* 28.6 (2008): 628-639.

Warschauer, Mark. "The changing global economy and the future of English teaching." *Tesol Quarterly* 34.3 (2000): 511-535.

Wolhuter, C. C., and S. C. Steyn. "Learning from south-south comparison: the education systems of South Africa and Madagascar." *South African Journal of Education* 23.1 (2006): 29-35.

-, CIA World Factbook. New York: Skyhorse Publishing, 2012.

-, Eurostat Europe in Figures Yearbook 2008 on Education .

-, OECD Database on Immigrants in OECD and non-OECD Countries: DIOC, Emigration Rates, http://www.oecd.org/els/mig/dioc.htm, Accessed 2/18/2014

-, World Development Indicators, http://data.worldbank.org/data-catalog/world-development-indicators, The World Bank, Accessed 2/28/2013.

Thanks to my committee members Kevin Grier, Daniel Hicks, Gregory Burge, Scott Linn, and chair Robin Grier for their helpful revisions. Thanks to the entire University of Oklahoma Department of Economics as well as my family and friends for their support. This paper was entirely written while I was working as an instructor at the University of Oklahoma, but no specific funding was tied to its findings.

Table 1: Hours Required for an English Language Speaker to Learn another Language

Hours	600	750	900		1100			1100+	2200	2200+
Languages	Afrikaans	German	Indonesian	Albanian	Hebrew	Maninka	Slovenian	Estonian	Arabic	Japanese
	Catalan	Luxembourgish	Kirundi	Amharic	Herero	Marshallese	Somali	Finnish	Berber	
	Danish		Malaysian	Armenian	Hindi	Montenegrin	Soninke	Georgian	Cantonese	
	Dutch		Swahili	Azerbaijani	Hiru motu	Moore	Swait	Hungarian	Dzongkha	
	Faroese			Bambara	Icelandic	Ndebele	Tagalog	Mongolian	Mandarin	
	French			Belarussian	Igbo	Nepali	Tajik	Thai	Korean	
	Italian			Bemba	Jola	Nyanja	Tamil	Vietnamese		
	Norwegian			Bengali	Kaonde	Oshiwambo	Tetum			
	Papiamento			Bislama	Kazakh	Palauan	Tigrinya			
	Portuguese			Bosnian	Khmer	Pashto	Tok Pisin			
	Romanian			Bulgarian	Kikongo	Persian	Tonga			
	Spanish			Burmese	Kinyarwanda	Polish	Tongan			
	Swedish			Carolinian	Kongo	Rukwangali	Tshiluba			
				Chamorro	Kurdish	Russian	Turkish			
				Chewa	Kyrgyz	Samoan	Turkmen			
				Chichewa	Lao	Sango	Tuvaluan			
				Comorian	Latvian	Sena	Ukrainian			
				Croatian	Lingala	Serbian	Urdu			
				Czech	Lithuanian	Serer	Uzbek			
				Damara	Lozi	Sesotho	Voro			
				Fiji hindi	Lunda	Setswana	Wolof			
				Fijian	Luvale	Setu	Xhosa			
				Fula	Macedonian	Shona	Yoruba			
				Gilbertese	Makhuwa	Silozi	Zulu			
				Greek	Malagasy	Sinhala				
				Hausa	Mandinka	Slovak				

For specification one, The 600, 750, and 900 categories constitute the first dummy variable, 1100 and 1100+ the second dummy variable, and 2200 and 2200+ constitute the third dummy variable. For specification two, each has its own dummy variable. The omitted dummy in each is 1 for a country where the sole official language is English and 0 if not. Countries use the hours to learn of the official language with the highest hours to learn.

Table 2: Correlations, 1992-2012

	Score	Hours to Learn	LP2	GDPPC, Ln
Score	1			
Hours to Learn	-0.2837	1		
LP2	0.4972	-0.1427	1	
GDPPC, Ln	0.3968	0.0376	0.5382	1

Correlation between percent speaking English and TOEFL score is .3321

Table 3: Summary Statistics, TOEFL Scores and Covariates, 1992-2012

Variable	Obs	Mean	Std. Dev.	Min	Max	Time Invariant
Emigration, % of Population	3440	7.14	10.11	0.10	43	Y
Net Exports, % of GDP	3612	87.70	51.51	0.31	531.74	N
Real GDP Per Capita, Ln	3813	8.07	1.65	3.91	11.98	N
FDI, % of GDP	3579	4.84	12.91	-82.89	366.36	Y
TOEFL Score	3050	83.43	9.67	51.00	105.74	N
Hours to Learn English	4297	967.35	637.30	0	2200	Y
Percent English Speaking	2280	45.59	31.95	0.15	100	Y
Linguistic Proximity-2	3634	0.74	0.73	0	3.77	Y
Official English	4297	0.34	0.47	0	1.00	Y
Schooling, Average Years of	2880	7.51	2.86	0.93	13.18	N
Inflation, GDP Deflator	3840	43.48	549.47	-32.81	26762.02	N
Investment, % of GDP	3469	23.13	11.25	-2.42	227.48	N
Island	4297	0.26	0.44	0	1.00	N
Land Area, Ln	4232	10.82	3.03	0.69	16.61	N
Landlocked	4297	0.20	0.40	0.00	1	Y
Population, Ln	4265	15.06	2.35	9.12	21.02	N
Government Consumption, % of GDP	3740	12.20	8.92	0.93	68.19	N
Ethnolinguistic Fractionalization	3400	0.46	0.27	0	0.98	Y
Latitude	4237	25.47	17.01	0	72	Y
Europe	4297	0.24	0.43	0	1	Y
Pacific	4297	0.09	0.28	0	1	Y
Middle East	4297	0.09	0.28	0	1	Y
Asia	4297	0.15	0.36	0	1	Y
Africa	4297	0.23	0.42	0	1	Y
Americas	4297	0.21	0.40	0	1	Y

Table 4: Determinants of Ln Real GDP per Capita, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.394***	0.078***	-0.235***	0.768***	0.267***	0.358***	1.082***	0.575***
	(23.03)	(4.99)	(-3.71)	(7.84)	(6.41)	(7.06)	(24.44)	(7.89)
Official English		-0.051***		-0.247***		-0.130***		-0.223***
		(-3.51)		(-7.37)		(-6.35)		(-8.43)
Schooling, Average Years of		0.470***		0.283***		0.394***		0.324***
		(25.77)		(7.92)		(16.91)		(11.21)
Inflation, GDP Deflator		-0.025**		-0.017		-0.022*		-0.019
		(-2.27)		(-1.08)		(-1.82)		(-1.41)
Investment, % of GDP		0.045***		0.105***		0.069***		0.095***
		(3.71)		(5.71)		(5.12)		(5.84)
Island		0.094***		0.093***		0.094***		0.102***
		(6.63)		(4.84)		(6.21)		(5.91)
Land Area, Ln		-0.018		0.137***		0.045*		0.088***
		(-0.95)		(4.04)		(1.93)		(2.97)
Landlocked		-0.184***		-0.195***		-0.188***		-0.184***
		(-14.40)		(-11.20)		(-13.84)		(-12.14)
Population, Ln		-0.046**		-0.161***		-0.092***		-0.121***
		(-2.40)		(-5.29)		(-4.23)		(-4.39)
Government Consumption, % of GDP		-0.133***		-0.110***		-0.124***		-0.114***
		(-11.09)		(-6.61)		(-9.60)		(-7.71)
Ethno-linguistic Fractionalization		-0.031**		-0.046**		-0.037**		-0.037**
		(-2.19)		(-2.38)		(-2.45)		(-2.04)
Latitude		0.123***		0.029		0.085***		0.051*
		(5.32)		(0.84)		(3.33)		(1.71)
Pacific		0.006		0.074***		0.033**		0.068***
		(0.40)		(3.36)		(2.06)		(3.53)
Middle East		0.069***		0.320***		0.170***		0.257***
		(4.06)		(7.67)		(6.81)		(7.63)
Asia		-0.189***		0.008		-0.109***		-0.042
		(-9.81)		(0.20)		(-4.44)		(-1.35)
Africa		-0.104***		0.107**		-0.018		0.056
		(-3.90)		(2.30)		(-0.58)		(1.41)
Americas		-0.018		0.016		-0.004		0.017
		(-0.90)		(0.57)		(-0.20)		(0.67)
N	2885	2324	2885	2324	2885	2324	2707	2241
R^2	0.155	0.713	-0.240	0.467	0.139	0.672	-0.312	0.590
F	530.6	336.2	13.73	184.3	41.05	296.5	597.0	230.8
Underid F Stat P-Value			0.000	0.000	0.000	0.000	0.000	0.000
Hausman P value			0.000	0.000	0.000	0.006	0.000	0.000

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 4a: Determinants of Ln Real GDP per Capita, First Stage Results on TOEFL Scores, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.062	0.387***				
			(1.43)	(9.17)				
1100-2200 Hour Dummy			-0.030	0.242***				
			(-0.68)	(5.74)				
2200 Hour and Up Dummy			-0.306***	0.165***				
			(-8.53)	(3.83)				
600 Hour Dummy					0.012	0.340***		
					(0.31)	(8.45)		
750 Hour Dummy					0.254***	0.229***		
					(11.38)	(10.66)		
900 Hour Dummy					-0.039*	0.187***		
					(-1.69)	(8.06)		
1100 Hour Dummy					-0.030	0.248***		
					(-0.72)	(6.24)		
1100+ Hour Dummy					-0.010	0.122***		
					(-0.43)	(5.09)		
2200 Hour Dummy					-0.292***	0.229***		
					(-8.55)	(5.52)		
2200+ Hour Dummy					-0.115***	-0.123***		
					(-6.35)	(-6.83)		
Linguistic Proximity-2							0.497***	0.264***
							(29.81)	(12.68)
Covariates			No	Yes	No	Yes	No	Yes
N			2885	2324	2885	2324	2707	2241
R^2			0.107	0.509	0.172	0.541	0.247	0.523
F			115.3	125.9	85.46	117.6	888.5	143.3

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 5: Determinants of Net Exports as a share of GDP, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.106***	0.214***	-0.272***	1.289***	-0.002	0.903***	0.278***	0.602***
	(5.73)	(8.77)	(-4.52)	(8.82)	(-0.04)	(10.61)	(7.02)	(6.04)
Official English		-0.207***		-0.507***		-0.399***		-0.324***
		(-9.17)		(-10.12)		(-11.61)		(-8.96)
Schooling, Average Years of		0.372***		0.081		0.185***		0.248***
		(12.96)		(1.47)		(4.68)		(6.17)
Inflation, GDP Deflator		-0.013		0.002		-0.003		-0.005
		(-0.71)		(0.08)		(-0.16)		(-0.26)
Investment, % of GDP		-0.212***		-0.116***		-0.151***		-0.109***
		(-11.11)		(-4.05)		(-6.51)		(-4.76)
Island		0.024		0.033		0.030		0.035
		(1.06)		(1.10)		(1.15)		(1.46)
Land Area, Ln		0.280***		0.517***		0.432***		0.335***
		(9.20)		(10.00)		(10.97)		(8.15)
Landlocked		-0.155***		-0.171***		-0.165***		-0.143***
		(-7.72)		(-6.27)		(-7.10)		(-6.72)
Population, Ln		-0.152***		-0.329***		-0.265***		-0.193***
		(-5.05)		(-7.00)		(-7.13)		(-5.02)
Government Consumption, % of GDP		-0.100***		-0.060**		-0.074***		-0.102***
		(-5.30)		(-2.29)		(-3.37)		(-4.91)
Ethno-linguistic Fractionalization		0.060***		0.038		0.046*		0.059**
		(2.67)		(1.24)		(1.76)		(2.34)
Latitude		-0.280***		-0.426***		-0.374***		-0.307***
		(-7.69)		(-8.06)		(-8.60)		(-7.31)
Pacific		-0.005		0.096***		0.059**		0.047*
		(-0.22)		(2.83)		(2.15)		(1.75)
Middle East		0.289***		0.679***		0.539***		0.450***
		(10.88)		(10.80)		(12.74)		(9.72)
Asia		0.117***		0.419***		0.311***		0.219***
		(3.87)		(7.33)		(7.47)		(5.10)
Africa		0.017		0.341***		0.225***		0.164***
		(0.40)		(4.79)		(4.14)		(3.00)
Americas		-0.124***		-0.068		-0.088**		-0.075**
		(-3.94)		(-1.59)		(-2.42)		(-2.17)
N	2863	2332	2863	2332	2863	2332	2683	2250
R^2	0.011	0.285	-0.132	-0.314	-0.000	0.040	-0.018	0.191
F	32.80	54.25	20.39	31.60	0.00190	43.59	49.29	40.63
Underid F Stat P-Value			0.000	0.000	0.000	0.000	0.000	0.000
Hausman P value			0.000	0.000	0.008	0.000	0.000	0.000

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 5a: Determinants of Net Exports as a share of GDP, First Stage Results on TOEFL Scores, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.019	0.408***				
			(0.43)	(9.84)				
1100-2200 Hour Dummy			-0.069	0.261***				
			(-1.56)	(6.28)				
2200 Hour and Up Dummy			-0.340***	0.179***				
			(-9.27)	(4.22)				
600 Hour Dummy					-0.028	0.366***		
					(-0.69)	(9.26)		
750 Hour Dummy					0.240***	0.237***		
					(10.65)	(11.14)		
900 Hour Dummy					-0.059**	0.197***		
					(-2.47)	(8.56)		
1100 Hour Dummy					-0.067	0.270***		
					(-1.60)	(6.90)		
1100+ Hour Dummy					-0.031	0.131***		
					(-1.29)	(5.54)		
2200 Hour Dummy					-0.326***	0.247***		
					(-9.33)	(6.01)		
2200+ Hour Dummy					-0.124***	-0.118***		
					(-6.79)	(-6.56)		
Linguistic Proximity-2							0.491***	0.270***
							(29.20)	(12.99)
Covariates			No	Yes	No	Yes	No	Yes
N			2863	2332	2863	2332	2683	2250
R^2			0.109	0.508	0.176	0.539	0.241	0.521
F			116.7	125.9	86.94	117.4	852.4	142.8

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 6: Determinants of FDI as a share of GDP, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.020	-0.033	-0.085	-0.548***	0.107**	0.029	0.076*	0.003
	(1.07)	(-1.19)	(-1.51)	(-4.28)	(2.30)	(0.36)	(1.87)	(0.03)
Official English		-0.003		0.131***		-0.019		-0.002
		(-0.10)		(3.09)		(-0.58)		(-0.06)
Schooling, Average Years of		0.132***		0.274***		0.115***		0.136***
		(4.06)		(5.60)		(2.99)		(3.15)
Inflation, GDP Deflator		-0.014		-0.022		-0.013		-0.015
		(-0.72)		(-1.01)		(-0.68)		(-0.72)
Investment, % of GDP		0.210***		0.157***		0.216***		0.197***
		(9.80)		(5.99)		(9.54)		(7.86)
Island		-0.034		-0.018		-0.036		-0.038
		(-1.31)		(-0.64)		(-1.39)		(-1.40)
Land Area, Ln		-0.058*		-0.173***		-0.045		-0.037
		(-1.71)		(-3.78)		(-1.18)		(-0.84)
Landlocked		0.018		0.026		0.017		0.025
		(0.80)		(1.07)		(0.76)		(1.12)
Population, Ln		-0.173***		-0.076*		-0.184***		-0.196***
		(-4.99)		(-1.74)		(-4.96)		(-4.53)
Government Consumption, % of GDP		-0.021		-0.026		-0.020		-0.030
		(-0.96)		(-1.10)		(-0.94)		(-1.33)
Ethno-linguistic Fractionalization		0.103***		0.119***		0.101***		0.088***
		(4.05)		(4.33)		(3.97)		(3.24)
Latitude		-0.052		0.027		-0.061		-0.081*
		(-1.25)		(0.57)		(-1.43)		(-1.77)
Pacific		-0.043*		-0.095***		-0.037		-0.045
		(-1.67)		(-3.12)		(-1.38)		(-1.58)
Middle East		-0.094***		-0.279***		-0.072*		-0.081
		(-3.18)		(-5.09)		(-1.82)		(-1.64)
Asia		-0.036		-0.183***		-0.019		-0.030
		(-1.07)		(-3.61)		(-0.47)		(-0.66)
Africa		-0.086*		-0.232***		-0.068		-0.065
		(-1.81)		(-3.75)		(-1.32)		(-1.13)
Americas		-0.060*		-0.074*		-0.058*		-0.065*
		(-1.70)		(-1.96)		(-1.65)		(-1.79)
N	2796	2292	2796	2292	2796	2292	2648	2209
R^2	0.000	0.114	-0.011	-0.022	-0.007	0.112	-0.001	0.108
F	1.153	17.21	2.293	15.92	5.281	17.10	3.489	15.62
Underid F Stat P-Value			0.000	0.000	0.000	0.000	0.000	0.000
Hausman P value			0.046	0.388	0.0415	1.000	0.123	0.005

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 6a: Determinants of FDI as a share of GDP, First Stage Results on TOEFL Scores, 1992-2012

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.030	0.411***				
			(0.71)	(10.02)				
1100-2200 Hour Dummy			-0.044	0.265***				
			(-1.03)	(6.42)				
2200 Hour and Up Dummy			-0.338***	0.172***				
			(-9.42)	(4.05)				
600 Hour Dummy					-0.006	0.373***		
					(-0.14)	(9.60)		
750 Hour Dummy					0.218***	0.237***		
					(10.08)	(11.30)		
900 Hour Dummy					-0.050**	0.197***		
					(-2.14)	(8.73)		
1100 Hour Dummy					-0.045	0.274***		
					(-1.09)	(7.09)		
1100+ Hour Dummy					-0.015	0.143***		
					(-0.61)	(6.06)		
2200 Hour Dummy					-0.324***	0.248***		
					(-9.42)	(6.09)		
2200+ Hour Dummy					-0.119***	-0.135***		
					(-6.42)	(-7.47)		
Linguistic Proximity-2							0.480***	0.267***
							(28.15)	(12.82)
Covariates			No	Yes	No	Yes	No	Yes
N			2796	2292	2796	2292	2648	2209
R^2			0.114	0.514	0.167	0.549	0.231	0.526
F			120.3	126.6	79.78	120.2	792.7	143.2

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 7: Determinants of Emigrant Stock 2000-2008, 2012 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.278***	-0.174	0.613**	-1.717*	0.443**	-0.149	0.348**	-1.989
	(3.27)	(-1.29)	(2.08)	(-1.82)	(2.17)	(-0.57)	(2.12)	(-1.61)
Official English		0.251**		0.590**		0.246**		0.672**
		(2.12)		(2.21)		(2.08)		(2.10)
Schooling, Average Years of		-0.091		0.600		-0.102		0.746
		(-0.65)		(1.30)		(-0.63)		(1.23)
Inflation, GDP Deflator		-0.109		-0.047		-0.110		-0.003
		(-1.12)		(-0.32)		(-1.25)		(-0.02)
Investment, % of GDP		0.046		-0.172		0.049		-0.046
		(0.45)		(-0.86)		(0.50)		(-0.26)
Island		0.219*		0.195		0.219**		0.176
		(1.92)		(1.15)		(2.13)		(0.91)
Land Area, Ln		-0.249*		-0.595**		-0.244*		-0.672*
		(-1.72)		(-2.00)		(-1.73)		(-1.79)
Landlocked		-0.128		-0.015		-0.130		0.029
		(-1.26)		(-0.09)		(-1.40)		(0.16)
Population, Ln		-0.089		0.190		-0.094		0.214
		(-0.59)		(0.68)		(-0.66)		(0.63)
Government Consumption, % of GDP		0.130		0.059		0.131		0.056
		(1.38)		(0.40)		(1.53)		(0.34)
Ethno-linguistic Fractionalization		-0.087		-0.025		-0.088		0.011
		(-0.81)		(-0.16)		(-0.91)		(0.05)
Latitude		0.019		0.383		0.013		0.565
		(0.12)		(1.18)		(0.08)		(1.26)
Pacific		-0.109		-0.159		-0.108		-0.187
		(-0.95)		(-0.93)		(-1.04)		(-0.94)
Middle East		-0.396***		-0.837**		-0.388***		-0.938**
		(-3.07)		(-2.57)		(-2.91)		(-2.14)
Asia		-0.458***		-0.590**		-0.456***		-0.575*
		(-2.77)		(-2.30)		(-3.04)		(-1.92)
Africa		-0.663***		-0.938***		-0.658***		-0.993**
		(-3.03)		(-2.59)		(-3.27)		(-2.26)
Americas		-0.011		0.056		-0.012		0.102
		(-0.07)		(0.23)		(-0.09)		(0.37)
N	129	96	129	96	129	96	119	91
R^2	0.078	0.491	-0.034	-0.363	0.051	0.490	0.076	-0.655
F	10.67	4.420	4.262	1.773	4.626	4.336	4.429	1.372
Underid F Stat P-Value			0.008	0.234	0.002	0.004	0.000	0.076
Hausman P value			0.236	1.000	0.376	1.000	0.969	1.000

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 7a: Determinants of Emigrant Stock in 2012, First Stage Results on TOEFL Scores, 2012 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.027	0.360*				
			(0.14)	(1.86)				
1100-2200 Hour Dummy			-0.072	0.279				
			(-0.40)	(1.53)				
2200 Hour and Up Dummy			-0.303*	0.219				
			(-1.78)	(1.00)				
600 Hour Dummy					-0.016	0.307*		
					(-0.09)	(1.71)		
750 Hour Dummy					0.268**	0.204*		
					(2.53)	(1.86)		
900 Hour Dummy					-0.088	0.153		
					(-0.80)	(1.40)		
1100 Hour Dummy					-0.075	0.236		
					(-0.44)	(1.46)		
1100+ Hour Dummy					-0.012	0.033		
					(-0.11)	(0.31)		
2200 Hour Dummy					-0.292*	0.286		
					(-1.79)	(1.38)		
2200+ Hour Dummy					-0.107	-0.212**		
					(-1.21)	(-2.45)		
Linguistic Proximity-2							0.538***	0.175
							(6.90)	(1.62)
Covariates			No	Yes	No	Yes	No	Yes
N			129	96	129	96	119	91
R^2			0.092	0.658	0.177	0.720	0.290	0.667
F			4.244	7.682	3.705	8.056	47.67	8.595

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 8: Determinants of Emigrant Stock 2000-2008, 1992 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
TOEFL Score	0.037	-0.116	0.295	-0.744	0.217	-0.083	0.283	-0.664
	(0.43)	(-0.98)	(1.41)	(-1.38)	(1.26)	(-0.29)	(1.57)	(-1.57)
Official English		0.002		0.171		-0.007		0.163
		(0.02)		(0.97)		(-0.06)		(1.02)
Schooling, Average Years of		-0.088		0.151		-0.101		0.095
		(-0.68)		(0.63)		(-0.64)		(0.48)
Inflation, GDP Deflator		0.013		0.025		0.012		0.037
		(0.16)		(0.29)		(0.17)		(0.43)
Investment, % of GDP		-0.043		-0.061		-0.042		0.013
		(-0.51)		(-0.68)		(-0.55)		(0.13)
Island		0.201*		0.120		0.205**		0.126
		(1.98)		(0.95)		(2.09)		(1.05)
Land Area, Ln		-0.365**		-0.655**		-0.350**		-0.630***
		(-2.60)		(-2.31)		(-1.96)		(-2.67)
Landlocked		-0.194**		-0.135		-0.197**		-0.142
		(-2.26)		(-1.32)		(-2.41)		(-1.53)
Population, Ln		0.001		0.207		-0.009		0.170
		(0.01)		(0.93)		(-0.06)		(0.92)
Government Consumption, % of GDP		0.122		0.179*		0.119		0.166*
		(1.42)		(1.76)		(1.46)		(1.72)
Ethno-linguistic Fractionalization		-0.050		0.062		-0.056		0.052
		(-0.45)		(0.42)		(-0.50)		(0.38)
Latitude		0.085		0.273		0.075		0.290
		(0.53)		(1.19)		(0.45)		(1.34)
Pacific		-0.030		-0.064		-0.028		-0.054
		(-0.27)		(-0.53)		(-0.28)		(-0.45)
Middle East		-0.371***		-0.582***		-0.360**		-0.576***
		(-3.02)		(-2.67)		(-2.51)		(-2.92)
Asia		-0.456***		-0.599***		-0.448***		-0.594***
		(-3.07)		(-3.05)		(-3.03)		(-3.22)
Africa		-0.531***		-0.645***		-0.525***		-0.582***
		(-2.67)		(-2.81)		(-2.81)		(-2.81)
Americas		-0.060		-0.064		-0.060		-0.041
		(-0.42)		(-0.43)		(-0.46)		(-0.27)
N	137	101	137	101	137	101	127	97
R^2	0.001	0.529	-0.065	0.367	-0.031	0.528	-0.063	0.413
F	0.186	5.477	1.948	4.128	1.572	5.419	2.424	4.193
Underid F Stat P-Value			0.000	0.148	0.000	0.057	0.000	0.004
Hausman P value			0.178	1.000	0.227	1.000	0.097	1.000

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 8a: Determinants of Emigrant Stock in 2012, First Stage Results on TOEFL Scores, 1992 Covariates

	(1) OLS	(2) OLS	(3) IV, Spec 1	(4) IV, Spec 1	(5) IV, Spec 2	(6) IV, Spec 2	(7) IV, Lp2	(8) IV, Lp2
600-900 Hour Dummy			0.037	0.089				
			(0.19)	(0.42)				
1100-2200 Hour Dummy			-0.156	-0.054				
			(-0.82)	(-0.26)				
2200 Hour and Up Dummy			-0.418**	0.176				
			(-2.56)	(0.88)				
600 Hour Dummy					-0.022	0.110		
					(-0.12)	(0.53)		
750 Hour Dummy					0.271***	0.051		
					(2.68)	(0.48)		
900 Hour Dummy					-0.020	0.086		
					(-0.18)	(0.68)		
1100 Hour Dummy					-0.161	-0.025		
					(-0.90)	(-0.13)		
1100+ Hour Dummy					-0.019	-0.003		
					(-0.19)	(-0.03)		
2200 Hour Dummy					-0.406**	0.272		
					(-2.60)	(1.40)		
2200+ Hour Dummy					-0.124	-0.137		
					(-1.51)	(-1.55)		
Linguistic Proximity-2							0.508***	0.279***
							(6.59)	(2.75)
Covariates			No	Yes	No	Yes	No	Yes
N			137	101	137	101	127	97
R^2			0.176	0.612	0.254	0.646	0.258	0.648
F			9.474	6.727	6.269	6.110	43.47	8.558

Standardized beta coefficients; *t* statistics in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Appendix 1: Countries Included

Countries with at least one observation of TOEFL scores for the years studied (173):

Afghanistan; Albania; Algeria; American Samoa; Andorra; Angola; Argentina; Armenia; Aruba; Australia; Austria; Azerbaijan; Bahrain; Bangladesh; Belarus; Belgium; Benin; Bhutan; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Brunei Darussalam; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Cape Verde; Central African Republic; Chad; Chile; China; Colombia; Congo, Dem. Rep.; Congo, Rep.; Costa Rica; Cote d'Ivoire; Croatia; Cuba; Cyprus; Czech Republic; Denmark; Djibouti; Dominican Republic; Ecuador; Egypt, Arab Rep.; El Salvador; Eritrea; Estonia; Ethiopia; Fiji; Finland; France; French Polynesia; Gabon; Gambia, The; Georgia; Germany; Ghana; Greece; Guatemala; Guinea; Guinea-Bissau; Haiti; Honduras; Hong Kong SAR, China; Hungary; Iceland; India; Indonesia; Iran, Islamic Rep.; Iraq; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Korea, Dem. Rep.; Korea, Rep.; Kosovo; Kuwait; Kyrgyz Republic; Lao PDR; Latvia; Lebanon; Lesotho; Liberia; Libya; Lithuania; Luxembourg; Macao SAR, China; Macedonia, FYR; Madagascar; Malawi; Malaysia; Mali; Marshall Islands; Mauritania; Mauritius; Mexico; Micronesia, Fed. Sts.; Moldova; Monaco; Mongolia; Montenegro; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Caledonia; New Zealand; Nicaragua; Niger; Nigeria; Northern Mariana Islands; Norway; Oman; Pakistan; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Puerto Rico; Qatar; Romania; Russian Federation; Rwanda; San Marino; Saudi Arabia; Senegal; Sierra Leone; Singapore; Slovak Republic; Slovenia; Somalia; South Africa; Spain; Sri Lanka; Sudan; Suriname; Swaziland; Sweden; Switzerland; Syrian Arab Republic; Tajikistan; Tanzania; Thailand; Togo; Tonga; Tunisia; Turkey; Turkmenistan; Uganda; Ukraine; United Arab Emirates; United Kingdom; United States; Uruguay; Uzbekistan; Venezuela, RB; Vietnam; Yemen, Rep.; Zambia; Zimbabwe.

Appendix 2: Data Definitions

Africa: 1 if area is in Africa, as defined by ETS.

Americas: 1 if area is in the Americas, as defined by ETS.

Asia: 1 if area is in Asia, as defined by ETS.

Emigration, % of Population: OECD immigration data showing how large the expatriate population of a country is relative to the total population of that origin country, updated 2008.

Ethnolinguistic Fractionalization: Greener sum of squares index constructed by Roeder 2001, comparing whether two randomly selected individuals in a country share the same mother tongue.

Europe: 1 if area is in Europe, as defined by ETS. Omitted in most specifications.

FDI, % of GDP: Net foreign investment as a share of GDP from World Development indicators.

GDP Per Capita, Ln real 2005 US dollars: Natural log of GDP per capita as measured in Real 2000 US dollars, from the World Development Indicators.

Government Spending, % of GDP: Government consumption as a share of GDP from Penn world tables.

Hours to Learn English: How many hours for a native English speaker to learn the official languages of a given country. This is assumed to be symmetric, so it would take the same time for the native official language speaker to learn English. From Foreign Service Institute.

Inflation, GDP Deflator: Inflation as measured by the GDP deflator, from the World Bank.

Institutional Quality: Sum of World Government indicators from the World Bank.

Investment, % of GDP: Gross capital formation as a share of GDP, from the World Bank.

Island: Equal to 1 if a country has no other countries it borders by land, from author's calculations.

Land Area, Ln: Natural log of square kilometers of land area of a country.

Landlocked: Equal to 1 if a country has no land bordering an ocean, from author's calculations.

Latitude: The latitude of a country from CIA World Factbook.

Linguistic Proximity-2: Measure of linguistic proximity with higher values being closer to English, from Melitz and Toubal (2014.)

Middle East: 1 if area is in the Middle East, as defined by ETS.

Official English: Equal to 1 if a countries' official languages include English, CIA world Factbook.

Population: Natural Log of population, from the World Bank

Pacific: 1 if area is in the Pacific, as defined by ETS.

Schooling 1990: Average years of education per person in 1990, from Barro-Lee.

Speaks English, % of Population: % of the population that reports they are Proficient in English. Provided from various national census bureaus and Crystal (2003)

TOEFL Score: Test of English as a Foreign Language, the Average TOEFL score in a country, from ETS 1992-2012 omitting 2011 as data is unavailable. TOEFL scores are rescaled here so that paper, computer and internet based test scores can be compared. Paper scores are rescaled such that the minimum score in the first year available for the paper exam, 223 in 1992, is the same as the lowest score in the first year recorded for the internet based exam, 0 in 2006, and so

that the maximum scores are the same, and 677 and 120 respectively. This yields a (rounded) linear transformation of Y = (X-223)/3.78, where X is the paper score and Y is the comparable transformed score. Using similar methodology with a reference year of 1998, Y = (X–7)/2.44 is used to transform computer test scores into internet based test scores using a base year of 1998 with a minimum score of 7 and maximum score of 300.