TV or not TV?

The impact of subtitling on English skills*†

Augusto Rupérez Micola
Luxembourg School of Finance

Ainhoa Aparicio Fenoll
Collegio Carlo Alberto

Albert Banal-Estañol
Universitat Pompeu Fabra, Barcelona GSE,
City University of London

Arturo Bris
IMD, Yale ICF, ECGI

December 2018

*We thank Antonio Cabrales, Jan Fidrmuc, Sven-Olof Fridolfsson, Karsten Jonsen, Javier Ortega, Thomas Tangerås, an associate editor, two reviewers, and seminar participants at City University, Universitat Pompeu Fabra, Institute for Research in Industrial Economics, Universitat de Barcelona, the European Economic Association, and the Work Pensions and Labour Economics Conferences for helpful feedback. Eduard Bartoll, Natàlia Izard, Agnieszka Szarkowska and Patrick Zabalbeascoa helped with the historical and linguistic accounts. We also thank Stefania Basiglio and Luís Rodrigo Úbeda for research assistance. Banal-Estanol gratefully acknowledges financial support from the Ministerio de Economía and Feder (ECO2016-76998-P) and Severo Ochoa Programme (SEV-2015-0563).

†Corresponding author: Albert Banal-Estañol, Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain, email: albert.banalestanol@upf.edu. Tel.: +34 93 542 2871, Fax: +34 93 542 1746.

Abstract

We study the influence of television translation techniques on the worldwide distribution of English-speaking skills. We identify a large positive effect for subtitled original version broadcasts, as opposed to dubbed television, on English proficiency scores. We analyze the historical circumstances under which countries opted for one of the translation modes and use it to account for the possible endogeneity of the subtitling indicator. We disaggregate the results by type of skills and find that television works especially well for listening comprehension. Our paper suggests that governments could promote subtitling as a means to improve foreign language proficiency.

JEL codes: I21, Z11

Keywords: Television, subtitling, foreign language skills

I Introduction

English is the language of the globalized world, and the lingua franca for the international communities in, among others, science, business, finance, advertising, tourism, and technology. Sixty-eight percent of citizens in the EU rate English as the most useful foreign language – far above the second position of French with 25 percent (European Commission, 2006).

Not surprisingly, English is the most widely learned foreign language, and this trend is expected to continue growing fast in the coming decades (Graddol, 2006). Graddol (1997) estimates that about one billion people are currently learning English worldwide, with 200 million in China alone.[1] More than 80 percent of the EU's school students learn English. The duration of foreign language as a compulsory subject ranges between six and 13 years in the non-English-speaking EU (Eurydice, 2005).[2] In comparison, students in England and Wales have foreign languages for five and three years, respectively, and there are no requirements in Ireland and Scotland.[3]

Despite the huge amounts of time and money spent, disparities in English proficiency across non-English speaking countries are large. In places such as the Netherlands, Denmark, and Sweden, more than 80 percent of citizens state that they are able to hold a conversation in English, but the proportion is below 60 percent in some of their neighboring countries like Austria, Germany, and France (European Commission, 2006). Portuguese-takers of the Test of English as a Foreign Language (TOEFL) score 95 on average (placing them 10th in a ranking of 135 countries), compared to 89 by their Spanish counterparts (rank 28). The reasons for these disparities between seemingly similar countries do not seem straightforward.

In this paper, we argue that the method used to translate foreign films and programs on television is an important driver of English skills in non-English-speaking countries. Subtitled original version programs provide continuous exposure to foreign languages as spoken by natives, which, we argue, is bound to improve the listeners' foreign-language skills. The US produces most of the successful films (and series) worldwide,[4] so that when someone watches films or series on

[1] Japan has created one hundred "super English high schools" where classes are taught exclusively in that language (Newsweek, 2007).

[2] The minimum is in the region of Flanders in Belgium and the maximum in the Netherlands, Norway, and Luxembourg.

[3] In 2004 a British survey discussed by the BBC showed that only one in 10 UK workers could speak a foreign language and less than 5 percent could count to 20 in a second language (http://news.bbc.co.uk/2/hi/uk_news/3930963.stm).

[4] The films produced by the Hollywood studios in Los Angeles represent 80% of world cinema's box office receipts (European Commission, 2006b).

television, the source language is very likely to be English.[5] Thus, the citizens of countries where television is broadcast in the original version would have better English vocabulary, grammar and, in particular, listening comprehension, than those of countries where programs are dubbed. Surprisingly, only 12 percent of Europeans think that television is useful for learning foreign languages (European Commission, 2006).[6]

We show that the average English proficiency of a country is positively associated with the country's expenditures in the education system and with the linguistic proximity of the local language to English. But, one of the most important significant explanatory factors appears to be the television translation mode. Our results suggest that, ceteris paribus, English skills are better in countries where television films and programs are subtitled. The magnitude of our effect is large, equivalent to 16.9 percent (one and a half standard deviations) of the average level of English skills. We disaggregate the results by types of skills – listening comprehension, speaking, reading, and writing – and find that television is an especially beneficial tool for listening comprehension.[7]

We use the insights of the history of cinema literature to account for the possible endogeneity of the translation mode, instrumenting it with language size at the time of the choice of translation mode. Indeed, we identify and analyze the historical circumstances under which countries opted for one of the alternatives in the years around World War II. According to the standard historical account, the use of subtitles was not due to a higher ability to understand the English language, nor to the idea that it would be beneficial for people to hear actors speak foreign languages (Crystal, 1997). Rather, limited box office receipts and a significant number of imported films induced small countries or, more precisely, countries with "small languages," to favor the

[5] In 1995, the US television exports into the EU amounted to US $6.8 billion, whereas the total US television imports amounted to $532 million (Ávila, 1997). On commercial television channels, the percentage of US fiction programs in the EU ranges from 60.7 percent of the total in France to 79.5 percent in the Belgian region of Flanders (De Bens and de Smaele, 2001). On public television channels, US fiction productions range from 19.6 percent in Germany to 52.9 percent in France. To these numbers one would have to add a significant number of series and films produced in other English-speaking countries (the UK, Canada, etc.).

[6] Europeans think that the best way to learn English is either at school (57% of the interviewed) or through lessons with a teacher, either one-to-one or in groups (40 and 42%, respectively). Other ways in which they think they can learn the language is by visiting the country, either as a tourist or while taking a language course (50 and 44%), or through conversation with native speakers, both through language exchanges and informally (36 and 33%).

[7] Our paper thus forms part of an emerging literature on the effects of television on educational and social phenomena. Gentzkow and Shapiro (2008), for example, find a positive effect of television on verbal skills in the US, which is particularly strong for those children whose mother tongue is not English. Television also influences violent crime (Dahl and DellaVigna, 2006), voting turnout (Gentzkow, 2006), democratic/republican patterns (DellaVigna and Kaplan, 2007) and international policy (Eisensee and Stromberg, 2007). There is further research on television and social capital in rural communities (Olken, 2006), anti-Americanism (Shapiro and Gentzkow, 2004) and even on the effect of soap operas on women's fertility (Chong et al., 2008).

low-cost subtitling option. Second, authoritarian regimes would have promoted dubbing in the local language to strengthen national identity. In any case, national media markets coordinated around one of the translation technologies at that time (Gottlieb, 1997), and have not deviated since. Using historical data, we provide evidence that, indeed, subtitling tended to be adopted in countries whose national languages were less widely used internationally. But, in our estimations, dictatorial regimes did not adopt dubbing significantly more often than more democratic countries.

Our paper suggests that governments could promote subtitling as a means to improve English language proficiency. This can come in addition to recent policy efforts to promote foreign language education at school.[8] The widespread knowledge of foreign languages, particularly English, has been linked to improvements in trade (Fidrmuc and Fidrmuc, 2009; Ku and Zussman, 2010; Melitz and Toubal, 2014), migration flows (Aparicio and Kuehn, 2016) and, more generally, income per capita (Ufier, 2015).[9,10] Of course, proficiency in foreign languages also has a direct impact on business. A survey conducted by the European Commission among nearly 2,000 small and medium European enterprises (European Commission, 2007) reports that a significant amount of business is being lost as a result of a lack of language skills.

As illustrations of our mechanism, consider again the cases of Austria and the Netherlands and Spain and Portugal. Austria and the Netherlands are two relatively small countries (less than 20 million inhabitants) that have similar levels of public education expenditure per student (3.08 and 3.07 percentage points of GDP per capita, respectively). But Austria shares a common language with Germany whereas Dutch is only spoken in the Netherlands and part of Belgium. Probably because of this, Austria broadcasts television dubbed in German while the Netherlands uses subtitles. This may contribute to explaining why 87 percent of the Dutch are able to hold a conversation in English while only 53 percent of Austrians can do so (European Commission, 2006). Similarly, Spain and Portugal share many geographical and cultural traits. But the number of Spanish speakers is double that of Portuguese speakers. Again, maybe in part because

[8] Over the last 50 years, most European countries have implemented reforms to introduce foreign languages in their compulsory education (Aparicio-Fenoll and Kuehn, 2015).

[9] More generally, previous literature has shown that countries that share a common language have higher bilateral trade flows (Anderson and van Wincoop, 2003; Frankel and Rose, 2002; Egger and Lassmann, 2015), cross-border activity (Coeurdacier et al., 2008), and cross-listings (Pagano et al., 2002).

[10] At the micro level, the literature has shown that better English skills allow immigrant populations in the US to earn more (Bleakley and Chin, 2004 and 2008; McManus, 1985; McManus et al., 1983). The ability to speak foreign languages has also been found to generate positive returns for non-immigrants in the EU (Ginsburgh and Prieto-Rodriguez, 2006), the US (Saiz and Zoido, 2005), and South Africa (Levinsohn, 2004).

of this, Portugal uses subtitling while in Spain television is dubbed. And, as a result, Portugal's results in the TOEFL exams are much better than Spain's. Better English skills may serve to increase the trade flows of the Netherlands and Portugal.

This paper also suggests that the translation mode could be used as an additional instrument for English proficiency. Linguistic proximity has been traditionally used in the literature as a determinant of English proficiency (e.g., Ku and Zussman, 2010; Ufier, 2015). Of course, the validity of each of these two variables as an instrument depends on the variable of interest. But, if we want to study the impact of English proficiency on macroeconomic variables such as trade or migration, linguistic distance may not satisfy the exclusion restriction because it may be capturing cultural similarities (Chen, 2013; Santacreu-Vasut et al., 2013) that may have a direct influence on trade and migration. In contrast, the choice of television translation mode does not depend on any kind of similarity between English and non-English speaking countries and in that sense it may be more likely to fulfill the exclusion restriction.

Nevertheless, the use of television translation mode as an instrument is not without drawbacks, either. First, it can only be used for trade with or migration to English-speaking countries. This may not be especially problematic, as most academic papers and databases focus on the US. Second, and more importantly, the main television translation mode has very little variation, none over time and very little within countries that share the same language. So, it cannot be used in regressions that include country or language fixed effects. More generally, our study inevitably needs to rely on a relatively small sample that mainly uses cross-section, between-language variation in translation mode to make inference.

The rest of the paper proceeds as follows. In the next section we provide an overview of the translation modes and a brief history of the choice between dubbing and subtitling. The data is introduced in section III. Section IV provides a description of the empirical strategy. In section V we present our main results on the influence of the translation mode on English skills, as well as those on why there are subtitles in some countries and dubbing in others. In section VI we conclude and discuss the limitations of our data and approach.

II Television translation modes: Background and history

There are three main foreign language translation traditions: subtitling, dubbing, and voice-over. Subtitling consists of supplying a translation of the spoken source language dialogue into the

target language in the form of synchronized captions, usually at the bottom of the screen, while the sound is in the original version. Hence, we use the terms "subtitled" and "original version" interchangeably. Dubbing is the method by which the foreign dialogue is translated, adjusting to the mouth movements of the actors so that the audience feels as if they are listening to actors speaking the target language. Finally, in voice-over, the translation is provided by a single person who does not imitate the action. For the purpose of this paper, we consider voice-over to have the same effects as dubbing because the viewer mainly hears his own language.

A Subtitling vs. dubbing

The film history literature provides a detailed account of the introduction of dubbing and subtitling in the cinema. In the times of silent cinema, inter-titles interrupted the course of a film to provide additional explanations to the audience. It was then easy to replace the original language titles with local-language text. But, with the introduction of sound, language became a serious problem for the cinema.[11] The Hollywood studios rapidly understood that one could not force audiences to watch films in a language they did not understand.[12] They therefore quickly started to promote dubbing around the world. In the 1930s, Paramount Pictures, for example, dubbed films into 14 European languages, including not only French and Spanish, but also Dutch and Swedish. A few years later, some countries moved on to subtitling while others continued with dubbing. The film history literature discusses two reasons for this shift.

First, there are economies of scale arguments. Countries with small languages, like the Netherlands, Sweden or Greece, moved to subtitling as the major translation mode. "The [dubbing] process was difficult, cumbersome, and far too expensive to be worthwhile in a small country" (Gottlieb, 1997). Limited box office receipts, combined with the relative low cost of subtitling and a significant number of imported films, meant that "the production of movies started to require much higher budgets than most of these countries could afford" (Danan, 1991). Note that some small countries who share large languages with others (e.g., Austria, Switzerland or the French-speaking Wallonia region in Belgium) also adopted dubbing.[13]

[11] In those times, those few in Europe with access to education overwhelmingly chose to learn either French or German. Widespread English language learning did not start taking place until the 1960s (Crystal, 2007).

[12] Hollywood was concerned with losing its leading position in the world market. D.W. Griffith, one of the founders of the Academy of Motion Picture Arts and Sciences said in 1923: "Only 5% of the world's population speak English. Why should I lose 95% of my audience?" (cited by Gottlieb, 1997). Film had developed into a universal language which all of a sudden would be divided into many languages when sound was added.

[13] This is consistent with Bridgman's (2013) findings that movie exporters use more intensive modes, i.e., those that require them to pay a higher share of distribution costs, in larger markets.

Second, there seem to be political motives. During the 1930s, countries like Germany, Italy, Japan, and Spain were taken over by authoritarian regimes that sought to strengthen national identity.[14] Dictators may have promoted the local language to strengthen national pride and may thus have favored dubbing (Mera, 1998). For example, the Spanish dictator Franco ruled against any non-dubbed version and published a number of ministerial guidelines (órdenes) to make showing films in a foreign language difficult (Szarkowska, 2005). In Italy, Mussolini introduced a law which ruled that all imported films had to be dubbed into standard Italian, with the idea of using cinema as a means of creating a common language (Szarkowska, 2005).

In sum, according to the standard account provided by film historians, the combination of these two factors would have resulted in the development of either dubbing or subtitling industries in the 1930–1940 period. The introduction of sound was parallel to the expansion of US cinema around the world. Television generally followed the country cinema translation choice upon its introduction in the 1950s (Ávila, 1999).[15] In particular, US "telefilms" and series became very popular and created the demand necessary for the growth of national translation industries.

B The persistence of the translation technology

The choice of the (main) television translation mode in each country, either dubbing or subtitling, has persisted to the present day (Szarkowska, 2005).[16] None of the countries of the OECD have moved from one to the other since World War II. This even applies to countries that later endured dictatorships, such as Greece. This persistence in the translation technology, which will be at the core of our identification strategy, can be explained by sunk costs and coordination on the supply side, and habit formation on the demand side (Blinn, 2008).[17]

Indeed, on the demand side, viewers now have strong preferences for the translation method used in their country. According to a European Commission (2006) survey, more than 90 percent

[14] Abramitzky and Sin (2014) show that authoritarian regimes can shape knowledge diffusion through language policies.

[15] Still, there are some countries that ended up using different translation modes in television and cinema. According to a report prepared for the European Commission (2007), in Europe, this is the case for Bulgaria, the Czech Republic, Hungary, Latvia, Lithuania, Poland, and Slovakia. Our analysis is based on the mode used in television. But we include a discussion of the translation mode used in the cinema as an additional source of variation in the conclusion.

[16] This classification is, of course, a simplification. Children's programs, for example, are dubbed in most countries and some late-night, less commercial films are broadcast in the original version in dubbing countries such as France or Spain.

[17] Digital technology has started to produce a slow convergence process and it is now possible to watch original version films in traditionally dubbing countries and dubbed versions in countries where subtitling is prevalent.

of the respondents in Sweden, Finland, Norway, Denmark, and the Netherlands agree with the following statement: "I prefer to watch foreign films and programs with subtitles, rather than dubbed." Around 30 percent of French, Spanish and Italians and less than 20 percent of Germans agree with this statement. A change from voice-over to the original version in one of the public television channels in Poland in 2008 was met with strong opposition.[18]

On the supply side, the existence of a consolidated industry also makes the change difficult. Countries have created and organized their local translation industries. The subtitling costs are double the European average in France, Germany, Spain, and Italy, where dubbing is prevalent. In contrast, dubbing costs are 66 percent more expensive than the European average in subtitling Scandinavia and the Netherlands (MCG, 2007).

III Data

We use data combining measures of English skills, translation mode, and demographic and educational variables for the period 2008–2015, as well as historical data of the time of sound cinema diffusion. Our data set includes all the 135 countries worldwide for which: (i) there is information on our measure of English proficiency, the internet TOEFL score, plus the television translation mode, and (ii) English is not the official language. Table A.1 in the Appendix shows the country list used in our regressions, separated by the main television translation mode, together with the official language and average measurements of English skills.[19]

A Translation mode

Our main explanatory variable is dichotomous, taking the value of one if foreign television programs are mainly subtitled, and a value of zero if they are dubbed or voice-overed. We collected information from multiple sources to create a database of the main translation mode used in

[18] See http://www.wirtualnemedia.pl/artykul/dwojka-z-pasmem-z-napisami-zamiast-lektora.

[19] From the set of countries with Internet TOEFL score data, we exclude, because English is official, Australia, Bahamas, Botswana, Cameroon, Canada, Gambia, Ghana, Gibraltar, Ireland, Jamaica, Kenya, Liberia, Malawi, Mauritius, Namibia, New Zealand, Nigeria, Sierra Leone, South Sudan, Swaziland, Uganda, the UK, US and Zambia. The official language of each country is obtained from Ethnologue (Simons and Fennig, 2017). In case of a country with more than one official language, we assigned the most widely-spoken official language in that country (e.g., we assigned German to Switzerland). We also exclude, because we could not find reliable information on the main television translation mode, Bahamas, Bhutan, Cabo Verde, Myanmar, Palestine, Papua New Guinea, South Sudan, Suriname, and United Arab Emirates. As we explain below, we also exclude Belgium from the sample because dubbing is used in one region and subtitling in another. The information source for the translation mode of each country is included in the Supplementary Appendix.

each country worldwide. As shown in Table A.1 in the Appendix, 67 of the 135 countries use subtitling, and 68 dubbing or voice-over, as the preferential translation method. Belgium is an interesting case as dubbing is used in the French-speaking Wallonia region but subtitling in the Dutch-speaking Flanders. We excluded it from the sample but we checked that all results are robust to the inclusion of Belgium as a subtitling country. We assigned subtitling because the population of the subtitling region (Flanders) is larger than that of the dubbing region (Wallonia).

A priori, the overall list is suggestive of some patterns. French- and German-speaking countries in Europe (Austria, France, Germany, Switzerland) all use dubbing. Small language (e.g., Finnish, Dutch, Greek) and Northern European countries mainly subtitle. Many Arabic-speaking countries in Africa use dubbing (e.g., Algeria, Morocco, Tunisia).

Interestingly, countries with the same language tend to use the same translation technology (which is consistent with the economies of scale argument discussed in section II.A).[20] We can say that there are "subtitling" and "dubbing" languages: the fraction of countries with the same language that subtitle is almost always between 0 and 0.2 (i.e., countries with a "dubbing language") or between 0.8 and one (i.e., countries with a "subtitling language"). The only languages in which the translation mode is relatively evenly split (fractions between 0.2 and 0.8) are Korean (North Korea uses dubbing and South Korea subtitles) and Mandarin (China and Taiwan use dubbing and Singapore subtitles).[21]

B English skills

We measure English skills, our main dependent variable, using national score averages of the TOEFL (Test of English as a Foreign Language) exams, designed and administered by the Educational Testing Service (ETS), a private non-profit organization. The TOEFL is an English-as-a-foreign-language exam accepted by most colleges and universities around the world. Its standardization means that it is relatively fair and accurate. The fact that everybody takes a similar test eliminates the inconsistency of interviews and other softer methods.

There are two versions of the test: paper-based and internet-based. The paper-based test

[20] In our sample of 135 countries, there are 72 languages. A bit less than half (61 out of 135) have a language that is unique in our sample. The mean number of countries per language is 1.8 but there are some languages shared by many countries (e.g., Arabic (17), French (16), and Spanish (21)).

[21] The other languages that use more than one translation mode are (i) Arabic: majority subtitling (14 countries), exceptions are Algeria, Morocco, Tunisia, (ii) French: majority dubbing (15 countries), exception is Haiti, (iii) Portuguese: majority subtitling (4 countries), exception is Mozambique, (iv) Spanish: majority subtitling (18 countries), exceptions are Chile, Equatorial Guinea, and Spain.

is the traditional version of the test, used since 1995, which aggregates three scores (reading, understanding, and writing). The internet-based test is, according to ETS itself, an "improved" version of the paper-based test, which is more reflective of communicative competence models, and it also includes a speaking category. Because of this, we display the results of the internet-based version in the main text and relegate those of the paper-based version to the Appendix.

Although the TOEFL scores are available yearly, our main independent variable (the translation mode) is time-invariant. Thus, our main regressions use time-averaged data over the sample period (2008−2015). As we explain in the next section, though, we replicate the main analysis with yearly data and report the resulting regression results in the Appendix.

As we can observe in Table 1, there are striking differences in English proficiency between subtitling and dubbing countries. Subtitling countries score 3.4 points higher in the overall internet-based TOEFL and obtain one point more in the paper-based TOEFL. At the disaggregated level, the differences in internet TOEFL scores are most pronounced for the listening comprehension tasks (1.4 points for internet-based and 2 for paper-based). Differences are statistically different from zero (p-value < 0.1) for the overall, listening, and speaking scores.

<<TABLE 1: DEPENDENT VARIABLES>>

A potential concern of the TOEFL score measures is that they may suffer from self-selection issues. TOEFL-takers may be those who are more interested in pursuing studies abroad. Hence, our measures may not reflect the English skills of the population as a whole but of a subsample of those with sufficient educational attainment or income to study overseas.[22] So, we have checked that our measures of English proficiency are consistent with other possible measures of English proficiency. Although they are highly correlated, a number of reasons deterred us from displaying the regression results obtained using these other measures, as we explain hereafter.

We first tried the percentage of people in each country who declare themselves able to hold a conversation in English in the three Eurobarometer surveys (e.g., European Commission, 2006). On average, 58 percent of people state they are able to use English in subtitling countries compared to 32 percent in dubbing countries. The correlation between this "Eurobarometer measure" and our overall TOEFL measures is significant, 0.44 for the paper-based and 0.56 for

[22] Unfortunately, information on the number of TOEFL test-takers across countries is not available for the years 2008-2015. Still, an earlier working paper version of this paper (Ruperez-Micola et al., 2009), which used an earlier (and smaller) sample, presented robustness checks for the scores corrected by the proportion of the national population that took the test each year. Results were qualitatively the same.

the internet-based versions of the test (statistically different from zero with a p-value < 0.01). Consistent with the nature of the question in the Eurobarometer, the highest correlation is with the score of the speaking part (0.65), followed by those with the writing, listening, and reading tests (0.58, 0.54, and 26, respectively). Unfortunately, the Eurobarometer measure is available for a limited number of (European) countries and the regression results were not significant.

We have also tried to make use of worldwide data on the fraction of English speakers per country, which is available in Appendix 1 of Melitz and Toubal (2014). These data, drawn from the "list of countries by English-speaking population" from the web encyclopedia Wikipedia, complement the information of the 2006 Eurobarometer survey from multiple sources worldwide. The fraction of English speakers is 15.98 percent, on average, for dubbing countries and 18.54 percent for subtitling countries (or 2.56 percentage points higher). The correlation between this measure and our overall TOEFL measures is also significant, 0.49 for the paper-based and 0.61 for the internet-based versions of the test (statistically different from zero with a p-value < 0.01). Unfortunately, maybe because the information sources are more heterogeneous, the regression results were not significant either.

C Other explanatory variables

The remaining explanatory variables, and their descriptive statistics, are shown in Table 2. As main control variables, we include demographic indicators (language size and country population), linguistic proximity with English, and a proxy for the quality of the education system (public expenditure in education per student as a percentage of GDP per capita).

<<TABLE 2: OTHER EXPLANATORY VARIABLES>>

As shown at the bottom of the table, dubbing countries do not differ significantly from subtitling countries in terms of language size or expenditures in education. But, as compared to dubbing countries, subtitling countries are smaller and have languages that are more similar to English. We also include a set of other "education controls" (staff/student ratios in primary and secondary school) and a set of "colonial past controls" (dummies for having been a UK or a US colony), which may also improve English proficiency.

In terms of data sources, population data are obtained from the World Bank Economic Indicators. Language size is measured as the sum of the populations of countries worldwide that use the same official language. The variable on education expenditures, as well as the

education controls, are provided by the IMD world competitiveness yearbook data set. Colonial past controls are obtained from CEPII (Mayer and Zignago, 2011). Linguistic proximity is based on an idea by Laitin (2000) and Fearon (2003), which has been taken up in several studies (including Melitz and Toubal, 2014). The idea is to use the Ethnologue classification of language trees into trees, branches, and sub-branches. Using English as a reference point, a country gets assigned a value of 0 if the local language belongs to a separate family tree (e.g., Mandarin), 0.25 if it belongs to a different branch of the same family tree (e.g., French, Spanish, Portuguese) and 0.50 if it belongs to the same branch (e.g., German, Dutch). Following Melitz and Toubal (2014), countries with more than one official language are assigned a weighted average of these values (where the weights are based on the size of the languages in the country).[23]

D Historical variables

In our main analysis, we instrument the television translation mode with the language size at the time of sound cinema diffusion, based on the arguments provided by the film history literature. As explained earlier, the film history literature points at language size and political situation at the time of sound cinema diffusion as the most important factors behind the choice of translation mode. Thus, we use the 1930–1940 average of language size and the 1930–1940 average of the Polity IV index (a measure of democracy that ranges from -10 to +10).[24]

In the main analysis, we use language size at the time of sound cinema diffusion as an instrument together with current language size as a control. To have a strong instrument, we need them to differ from each other. The correlation between historical and current language size is high (0.95). But, there was a lot of variation in language size growth. French-, German- and Russian-speaking countries experienced a low growth rate (the number of speakers increased by 9%, 10%, and 25%, respectively) while Spanish, Portuguese, Arabic, and Mandarin experienced

[23] This makes linguistic proximity a continuous variable, although most countries have values equal to 0, 0.25, or 0.5. However, our regression results are qualitatively the same if we use, rather than a continuous variable, two dummy variables: one that takes a value of one for linguistic proximities higher than zero and smaller or equal to 0.25 and another dummy variable that takes a value of one if linguistic proximity is higher than 0.25 and smaller or equal to 0.5 (leaving linguistic proximity equal zero as the reference category).

[24] For our measure of historial language size, we collected and merged 1930–1950 country population data from Maddison (2003), the United Nations and the Institute for Demographic Studies. We combined these data with 1950–1960 data from the Penn World Tables and with 1960–2016 data from the World Bank (the detailed data sources are available in the Supplementary Appendix). We then ran an unconditional population growth model (Snijders, 2011) to impute population size for the country-years for which we had missing population data at the time of cinema diffusion (1930–1940). Polity IV index can be found in the Polity IV project website at: http://www.systemicpeace.org/polity/polity4.htm.

relatively high growth rates (134%, 91%, 87%, and 78%, respectively). The higher fertility rates of former colonies with respect to developed countries may explain these differences. Figure 1 represents current versus historical language size to illustrate the kind of variation that our instrument is providing. We exploit departures of language size at the time of sound cinema diffusion from the fitted line.

<<FIGURE 1: EVOLUTION OF LANGUAGE SIZE>>

IV The empirical strategy

In our main analysis, we estimate the effect of subtitling on English proficiency using the following linear specification based on time-average data:

$$\log(TOEFL_i) = \beta_0 + \beta_1 S_i + \beta_2 Lansize_i + \beta_3 Pop_i + \beta_4 \log(Edex_i) + \beta_5 Linsim_i + C_i + \varepsilon_i \quad (1)$$

where $TOEFL_i$ represents the average English proficiency in country i over the 2008–2015 period, as measured by one of the TOEFL scores (paper or internet-based, overall or disaggregated by skill), S_i is a dummy variable equal to one if country i uses subtitles, $Lansize_i$ represents the time-average size of its language, Pop_i its time-average population, $Edex_i$ its time-average education expenditures, $Linsim_i$ its linguistic similarity index, C_i a vector of time-average education and colonial past controls, and ε_i the residual. In an alternative specification, reported in the Appendix, we also use yearly data for each country (over the same 2008–2015 period), include year fixed effects and cluster standard errors at the country level.[25]

We estimate this model using both ordinary least squares (OLS) and two-stage least squares (2SLS). The OLS estimation may be biased if countries decide to use subtitling depending on their level of English proficiency (reverse causality) or if, for instance, countries with open cultures are more likely to have citizens that know English and prefer subtitles (omitted variables). We instrument the variable S_i with language size at the time of sound cinema diffusion.[26] The

[25] We use the logarithmic transformation of the dependent variable to ease the interpretation of the effects. As shown in Table A.2 in the Appendix, results do not change when we do not log transform the dependent variable or when we do log transform language size or population.

[26] We include the instrument in absolute values rather than its logarithmic transformation because (i) this is consistent with the argument that it is the scale of the worldwide film market at the time of cinema diffusion (in millions of potential viewers) that determined the choice of translation mode, and (ii) the instrument becomes weak if we use the logarithmic transformation. The choice of functional form of the instrument should not affect the consistency of the estimates.

validity of this variable as an instrument for subtitling relies on the assumption that it affects English proficiency only through the subtitling decision (conditional on the controls). For this condition to hold it is crucial that we also control for contemporaneous language size, which may be related to the government's incentives to invest in English classes in the public education system, or to the population's incentives to study English. In the absence of contemporaneous language size as a control, language size at the time of sound cinema diffusion may be artificially capturing some of these aspects due to its natural correlation with current language size.

To explore the use of the historical variables as instruments of translation mode, we run a simple linear probability model:

$$\begin{aligned} S_i &= \beta_0 + \beta_1 HistLansize_i + \beta_2 HistPolit_i + \beta_3 Lansize_i + \beta_4 Polit_i + \\ &+ \beta_5 Pop_i + \beta_6 \log(Edex_i) + \beta_7 Linsim_i + C_i + \nu_i \end{aligned} \tag{2}$$

where $HistLansize_i$ and $HistPolit_i$ are the language size and the Polity IV index at the time of sound cinema diffusion of country i, $Polit_i$ represents its average Polity IV index over the 2008–2015 period and ν_i the residual.

As a robustness check, we also estimate the reduced form version of the main model, which consists of regressing English proficiency on language size at the time of sound cinema diffusion directly (in addition to all the controls). Figure 2 depicts the relationship between English proficiency and language size at the time of sound cinema diffusion. We observe a strong and negative correlation between those variables. But, of course, in order to interpret this negative relationship as evidence in favor of the impact of subtitling, we need to control for contemporaneous language size (and the other explanatory variables), as we do in the regressions.

<<FIGURE 2: HISTORICAL LANGUAGE SIZE AND TOEFL SCORES>>

V Results

A The determinants of English proficiency (OLS)

Table 3 reports the results of OLS regressions on several factors that could plausibly influence the level of English proficiency in a country, as specified in Equation (1). The dependent variable is the overall internet-based TOEFL score in each country. We control for language size, population

size, education expenditures and linguistic proximity. We depart from a specification with neither the education and colonial past controls nor language fixed effects. We then sequentially add the education controls (staff/student ratios in primary and secondary education) and the colonial past controls (dummies for having been a UK or US colony). We then restrict the sample to countries with languages spoken in more than one country and compare the results with and without language fixed effects.

<<TABLE 3: SUBTITLING AND ENGLISH PROFICIENCY (OLS)>>

The coefficient for subtitles is positive and highly significantly different from zero in the first four regressions. The magnitude is stable across the three specifications, indicating that subtitling is associated with an increase of about 4% in TOEFL scores. This magnitude is equivalent to an increase in 2.35% in education expenditures (a 1% increase in education expenditures is associated with a 1.7% higher TOEFL score). The correlation is even higher in the subsample of languages spoken in more than one country in regression four. Unfortunately, as shown in regression five, there is not enough within-language variation in translation mode to identify the coefficient of subtitling when introducing language fixed effects. As shown above, countries with the same language tend to use the same translation technology, thus forming "subtitling" and "dubbing languages."

B The determinants of the translation mode

The first column of Table 4 contains the results of the empirical examination of the historical account of the dubbing/subtitling decision, as specified in Equation (2). We jointly test whether language size and political regime, both measured at the time of sound cinema diffusion, can explain the adoption of a certain translation mode. We use the same controls as before, including contemporaneous language and population size. Positive parameter estimates indicate that the variable is more conducive to subtitling, while negative estimates suggest a propensity for dubbing.

<<TABLE 4: SUBTITLING AND ENGLISH PROFICIENCY (IV)>>

The coefficient of the political regime at the time of sound cinema diffusion is, albeit positive, not significant. Thus, we do not find that more democratic countries adopt subtitling significantly more often than more dictatorial regimes. In contrast, language size at the time of sound cinema

diffusion has a very significant and negative correlation with the probability of adopting subtitles. An increase of one million in the number of speakers of a particular language at the time of sound cinema diffusion is associated with a reduction of 0.007 in the probability of using subtitles in the countries where the language is official. These results explain why we focus on language size at the time of sound cinema diffusion as the main shifter of the translation mode and use it as an instrument in the regressions that explain English proficiency.

C The determinants of English proficiency (IV)

The rest of the columns of Table 4 show our results for the estimation of the causal effect of subtitling on English proficiency. The second column refers to the reduced form (RF) regressions of English proficiency on language size at the time of sound cinema diffusion. The third column contains the first stage (FS) regression of subtitling on language size at the time of sound cinema diffusion. The last column shows the instrumental variables (IV) estimates for the impact of subtitles on TOEFL scores.

In the reduced form regression, the coefficient of language size at the time of sound cinema diffusion is negative and significant. It indicates that larger languages at the time of sound cinema diffusion (associated with a higher probability of dubbing) imply lower English proficiency. In particular, an increase of one million in language size at the time of sound cinema diffusion is associated with a reduction of 0.1 percent in TOEFL scores.

The coefficients associated with language size at the time of sound cinema diffusion in the first stage regression is also negative and highly significant (and similar to that reported in the previous subsection). This specification indicates that an increase of one million speakers at the time of sound cinema diffusion decreases the probability of using subtitles by 0.8 percent. The F-statistic of the excluded instrument is 39.83. This value is well above the rule-of-thumb "critical value" of the Stock and Yogo (2005) test so we conclude that our instrument is not weak in the context of our specification. Moreover, the coefficient for contemporaneous language size is positive and significant, which reassures us that our instrument is capturing the desired variation, i.e., at the time of sound cinema diffusion in larger markets it was found to be more profitable to introduce dubbing, but the contemporaneous size of the markets is capturing other factors.

The instrumental variable coefficient resulting from the ratio of the reduced form and first stage estimates is positive and significant. The coefficient indicates that a change from dubbing

to subtitling translation mode in a country improves test scores by 16.9 percent. Education expenditures and linguistic proximity present a positive and significant correlation with English proficiency, as we had expected.

D Types of language skills

Table 5 reports IV regressions of the four parts of the internet TOEFL exam: reading, writing, speaking, and listening. All coefficients are positive and significant. The highest effect is found for listening (25.2%), followed by reading (18.3%), writing (12.6%), and speaking (11.9%). The coefficient for listening is significantly higher than the one of the average effect (16.9%).

<<TABLE 5: SUBTITLING AND ENGLISH PROFICIENCY BY SKILL>>

E Robustness

Table A.3 in the Appendix replicates the main regressions of tables 3, 4, and 5 using the different measures of the paper-based version of the test as dependent variables. Results are in line with those obtained using the internet-based version. The coefficient for subtitles in the instrumental variable estimation indicates that a change from dubbing to subtitling translation mode in a country improves paper-based test scores by 4.6 percent. By type of skill, the strongest and most significant effect is found again for listening (7.5%).

As our main-regressor-of-interest is time-invariant, all our previous regressions use time-averaged data over the 2008–2015 period. But, while disaggregating by year may overstate the sample size, averaging may hide the underlying noise in the yearly data that is important for understanding the statistical relationship and precisely estimating the other coefficients. Table A.4 in the Appendix replicates the main regressions of tables 3, 4, and 5 using yearly data, all the controls, year dummies, and clustering standard errors by country. The estimated coefficients are similar (and statistically indistinguishable) from those in the regressions that use time-averaged data.

Table A.5 in the Appendix again replicates the annual analysis using the paper-based variables as dependent variables. Results are in line with those obtained using the internet-based version of the test. The effect on the overall TOEFL scores is significant and, as before, the strongest effect by type of skill is found for listening.

VI Conclusion and discussion

The general message in this paper is simple. Continuous exposure to English-language media contents help people learn English and, thus, the citizens of countries where foreign films and programs are shown in their original version in television will likely speak, on average, better English than those that live in countries where television is dubbed. This is relevant because previous studies have shown that better English language skills improve economic performance.

Dubbing countries in our sample invest the same in education as the subtitling countries. Yet subtitling countries score 3.4 points higher in the TOEFL exams. We show that the television translation methods can explain part of the skills gap. We identify a subtitling effect equivalent to 16.9 percent of the overall TOEFL score. We also analyze the differential impact of subtitling by type of English skill (listening, speaking, reading, and writing). We find that the strongest effect is for listening (19.4%). Our results are robust to the inclusion of other determinants of language skill including language proximity, demographic indicators and proxies for the quality of the education system. Interestingly, the choice of translation technology at the time of sound cinema diffusion did (could) not take into account the benefits of improved English skills. In fact, subtitling may have appeared undesirable at first because it forced audiences to read, but it turned out to be beneficial ex-post in terms of English proficiency (and audiences got used to subtitling). This paper thus shows that how countries adopt foreign "cultural" products matters in the long term, as it may create externalities.

Our results can therefore help in both raising awareness and overcoming resistance in a context of the increasing importance of English proficiency.[27] As an example, take the government of Poland's decision in 2008 to introduce subtitling in the public television channels. It was met with strong opposition. A poll had found that only 19 percent of Poles would welcome subtitled television. Still, this percentage reached 32 percent among young, educated individuals. Probably because of this, subtitling was finally introduced in the channel TVP2 which targets young audiences, who may be less reluctant to change from voice-over to subtitling.

Our paper is a first attempt to measure the impact of television translation mode on English proficiency. Statistical analysis, though, may not provide a definite answer. There is no variation in television translation mode over time and very little within countries that share the same language. Sunk costs and coordination in the translation industry on the supply side, and,

[27] In Asia this phenomenon is particularly accurate to the extent that experts have coined the term "English fever" (Park, 2009).

importantly, habit formation on the demand side explain the persistence in translation technology (Blinn, 2008). Economies of scale in translation mode costs explain why countries with the same language tend to use the same translation technology. Although these two features help us in our identification strategy, they also mean that our main source of variation is cross-sectional and between-language. This implies, even after making every possible effort to include as many countries as possible, that our results hinge on a relatively small, historical, and non-experimental sample.

There may be different ways in which one could obtain more variation within and across countries. A change in the main translation mode, such as the one that had been originally proposed in Poland, would have been a good opportunity to estimate the impact of translation mode on English proficiency. Another source of variation could be to find out which countries use (and to which extent they use) a "mixed" system, in which the translation mode is different in cinema than in television. More generally, one could also exploit cross-country (and time-varying) variation in terms of media content exposure. Television penetration, internet usage, cinema attendance, and radio receivers per capita, for instance, could be playing a mediating role in the relationship between translation mode and English skills across subtitling countries.

Ideally, one would like to work with individual rather than country-level data. The linguistics and education literatures have already analyzed, in surveys or small samples, the role of television translation mode on foreign language learning. Webb (2010), for instance, analyzing movie scripts, concludes that low-frequency words can be learned by watching movies regularly. This is also in line with teachers' perceptions, as declared in interviews and surveys conducted by education researchers (Seferoğlu, 2008; Demet, 2009). Individual-level data may also be able to account for personal differences in foreign language skills. Some surveys (e.g., European Commission, 2005) find substantial differences between men and women, the young and the old, city and countryside residents, and across education attainment levels. We believe that the analysis of the causes and the consequences of English proficiency at the micro level could be a fruitful area for future research.

References

[1] Abramitzky, R. and Sin, Isabelle (2014): "Book Translations as Idea Flows: The effects of the collapse of Communism on the diffusion of knowledge." Journal of the European Economic Association, 12(6): 1453-1520.

[2] Aparicio-Fenoll, A. and Zoe Kuehn (2016): "Does Foreign Language Proficiency Foster Migration of Young Individuals within the European Union?" The Economics of Language Policy, edited by Bengt-Arne Wickstroem and Michele Gazzola, MIT Press (ISBN: 978-0-262-03470-8).

[3] Anderson, J.; van Wincoop, E. (2003): "Gravity with Gravitas: A Solution to the Border Puzzle." American Economic Review. 93: 170–192

[4] Ávila, A. (1997): La historia del doblaje cinematográfico. Cims editores. ISBN 84-89643-37-3.

[5] Ávila, A. (1999): Así se crean doblajes para cine y televisión. Cims editores. ISBN 84-8411-518-6.

[6] Bleakley, H., and Chin, A. (2004): "Language skills and earnings: evidence from childhood immigrants." Review of Economics and Statistics 86, 481-496

[7] Bleakley, H., and Chin, A. (2008): "What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital Among Immigrants." Journal of Human Resources, 43:267–298.

[8] Blinn, Miika (2008): "Path Dependence of the Language Transfer Technique for Foreign Films in the German Cinema Market." Free University Berlin Working Paper.

[9] Bridgman, Benjamin (2013): "Export Mode and Market Entry Costs," BEA Working Papers 0089, Bureau of Economic Analysis.

[10] Chen, M Keith (2013): "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets." American Economic Review, 103(2): 690-731.

[11] Chong, A., Suzanne Duryea , Eliana La Ferrara (2008): Soap operas and fertility in Brazil. CEPR working paper. http://www.cepr.org/pubs/new-dps/dplist.asp?dpno=6785

[12] Coeurdacier, N., R.A. De Santis, and A. Aviat (2008): "Cross-border mergers and acquisitions: Financial and institutional forces." Working paper.

[13] Crystal, D. (1997): English as a global language. Cambridge University Press, ISBN 978-0-521-53032-3.

[14] Dahl, Gordon and Stefano DellaVigna (2006): "Does movie violence increase violent crime?" NBER working paper. http://www.nber.org/papers/w13718

[15] Danan, M. (1991): "Dubbing as an expression of nationalism." Meta, vol. 36, no. 4, pp. 606-614.

[16] De Bens and de Smaele (2001): "The inflow of American television fiction on European broadcasting channels revisited." European Journal of Communication, 16 (1), pp. 51-76.

[17] DellaVigna, Stefano and Ethan Kaplan (2007): "The Fox News effect: Media bias and voting." Quarterly Journal of Economics, 122, pp. 1187-1234.

[18] Demet, Y. (2009): "Pre-service EFL Teachers in Practice: Integrating Films into Language Teaching." STEM Journal, 10, 77-98.

[19] Egger, Peter H. and Andrea Lassmann, (2015): "The Causal Impact of Common Native Language on International Trade: Evidence from a Spatial Regression Discontinuity Design" The Economic Journal, 0(584), pages 699-745, 05.

[20] Eisensee, Thomas and David Stromberg (2007): "News Floods, News Droughts, and US Disaster Relief." Quarterly Journal of Economics, vol. 122, issue 2, pp. 693-728.

[21] Eurydice (2005): Key data on teaching languages at school in Europe. ISBN 1830-2076.

[22] European Commission (2005): Europeans and languages. Eurobarometer 237.

[23] European Commission (2006): Europeans and their languages. Special Eurobarometer 243.

[24] European Commission (2006b): The economy of culture in Europe, report prepared for the Directorate-General for Education and Culture.

[25] European Commission (2007): Effects on the European Economy of Shortages of Foreign Language Skills in Enterprise.

[26] European Commission (2012): Europeans and their languages. Special Eurobarometer 387.

[27] Fearon, James D. (2003): "Ethnic and Cultural Diversity by Country." Journal of Economic Growth, June 2003, Volume 8, Issue 2, pp. 195-222.

[28] Fidrmuc, Jan and Jarko Fidrmuc (2009): "Foreign Languages and Trade." CEDI Discussion Paper Series 09-03, Centre for Economic Development and Institutions (CEDI), Brunel University.

[29] Frankel, J., and Rose, A., (2002): "An estimate of the effect of common currencies on trade and income." Quarterly Journal of Economics, 117, pp. 437–466.

[30] Gentzkow, Matthew (2006): "Television and Voter Turnout". Quarterly Journal of Economics, 121(3), pp. 931-972.

[31] Gentzkow, Matthew and Jesse Shapiro (2008): Preschool television viewing and adolescent test scores historical evidence from the Coleman Study. Quarterly Journal of Economics. CXXIII (1). February 2008, 931-972.

[32] Gottlieb, H. (1997): Subtitles, translation & idioms. PhD thesis, University of Copenhagen. ISBN 87-89065-39-5.

[33] Graddol, D. (1997): The future of English? A guide to forecasting the popularity of the English language in the 21st century. Published by the British Council, available at: http://www.britishcouncil.org/learning-elt-future.pdf

[34] Graddol, D. (2006): English next. Published by the British Council, available at: http://www.britishcouncil.org/learning-research-english-next.pdf

[35] Ginsburgh, V. and J. Prieto-Rodriguez (2006): "Returns to Foreign Languages of Native Workers in the EU." Working paper.

[36] Ku, Hyejin and Asaf Zussman (2010): "Lingua franca: The role of English in international trade." Journal of Economic Behavior & Organization, 75(2), pp. 250-260.

[37] Laitin, David D. (2000): "What is a Language Community?" American Journal of Political Science, Vol. 44, No. 1 (Jan. 2000), pp. 142-155.

[38] Levinsohn, J. (2004): "Globalization and the Returns to Speaking English in South Africa." NBER working paper 10985.

[39] Maddison, Angus (2003): "The World Economy". Historical Statistics. Paris, OECD.

[40] Mayer, T. and S. Zignago (2011): "Notes on CEPII's distances measures : the GeoDist Database." CEPII Working Paper 2011-25.

[41] McManus WS. (1985): "Labor market costs of language disparity: an interpretation of Hispanic earnings differences." The American Economic Review 75: pp. 818-827.

[42] McManus W, Gould W, Welch F. (1983): "Earnings of Hispanic men: The role of English language proficiency." Journal of Labor Economics 1(2): 101-130.

[43] Melitz, Jacques, and Farid Toubal (2014): "Native language, spoken language, translation and trade." Journal of International Economics 93(2), pp. 351-363.

[44] Mera, Miguel (1998): "Read my Lips: Re-Evaluating Subtitling and Dubbing in Europe" Links & Letters 6, 1999, pp.73-85.

[45] MCG (2007): "Dubbing and subtitling needs and practices in the European audiovisual industry." Report for the Information Society and Media Directorate General of the EuropeanCommission.

[46] Newsweek (2007): English for everyone. In issue 20-27 August. By Jonathan Adams and Max Hirsch.

[47] Olken, Benjamin (2006): "Do Television and Radio Destroy Social Capital? Evidence from Indonesian Villages." NBER Working Paper 12561. Available from http://www.nber.org/~bolken/televisionandsocialcapital.pdf

[48] Pagano, M., A.A. Röell, and J. Zechner (2002): "The geography of equity listing: Why do companies list abroad?" The Journal of Finance 57, 2651-2694.

[49] Park, J.-K. (2009): "'English fever' in South Korea: its history and symptoms." English Today, 25(1), pp. 50–57.

[50] Rupérez-Micola, A, A. Bris and A. Banal-Estañol (2009): "TV or not TV? Subtitling and English skills," Economics Working Papers 1156, Department of Economics and Business, Universitat Pompeu Fabra.

[51] Saiz, A. and E. Zoido (2005): "Listening to what the world says: Bilingualism and earnings in the United States." Review of Economics and Statistics, 87, 523-538.

[52] Santacreu-Vasut, Estefania, Amir Shoham and Victor Gay (2013): "Do female/male distinctions in language matter? Evidence from gender political quotas," Applied Economics Letters, Taylor & Francis Journals, vol. 20(5), pages 495-498, March.

[53] Seferoğlu, G. (2008): Using feature films in language classes. Educational Studies, 34(1), 1-9.

[54] Shapiro, J.M and Matthew Gentzkow (2004): "Media, Education, and Anti-Americanism in the Muslim World." Journal of Economic Perspectives: vol 18, pp. 117-133.

[55] Simons, Gary F. and Charles D. Fennig (2017): Ethnologue: Languages of the World, Twentieth edition. Dallas, Texas: SIL International.

[56] Snijders, Tom AB (2011): "Multilevel analysis.", International Encyclopedia of Statistical Science. Springer Berlin Heidelberg, 2011. 879-882.

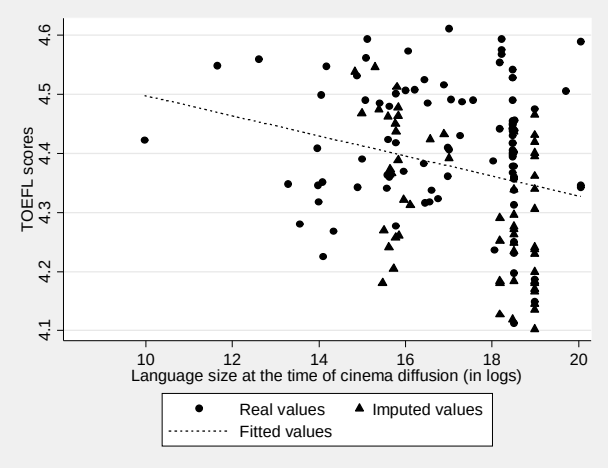
[57] Stock, James H. and Motohiro Yogo (2005): Testing for weak instruments in linear IV regression. In D.W.K. Andrews and J.H. Stock, eds. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg. Cambridge: Cambridge University Press, 2005: 80-108.

[58] Szarkowska, A. (2005): "The power of film translation." Translation Journal, 9(2).

[59] Ufier, A. (2015): "The Impact of English Language Skills on National Income: A Cross-National Comparison." University of Oklahoma. Working Paper.

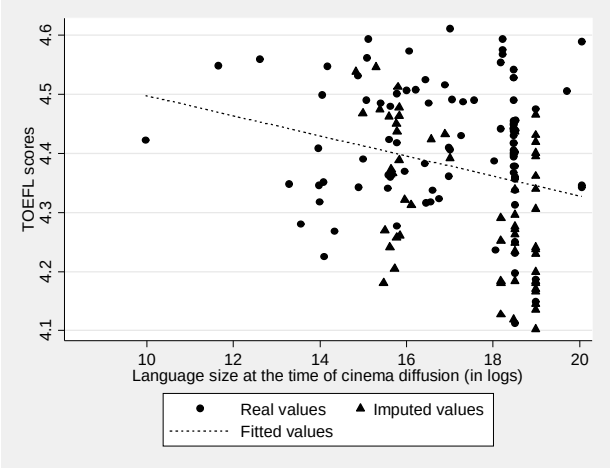
[60] Webb, S. (2010): "A corpus driven study of the potential for vocabulary learning through watching movies." International Journal of Corpus Linguistics, 15(4), 497-519.

Figure 1: Evolution of language size



The information on current language size is obtained by aggregating country population data (World Bank) by official language (Ethnologue). For the measure of historical language size, we merged 1930–1950 country population data from Maddison (2003), the United Nations, and the Institute for Demographic Studies. We combined these data with 1950–1960 data from the Penn World Tables and with 1960–2016 data from the World Bank. We then ran an unconditional population growth model (Snijders, 2011) to impute population size for the country-years for which we had missing population data at the time of sound cinema diffusion (1930–1940). We then aggregated country population data by official language (Ethnologue).

Figure 2: Historical language size and TOEFL scores



The TOEFL scores are overall averages for test-takers resident in each country in the sample period (2008–2015). For the measure of historical language size, we merged 1930–1950 country population data from Maddison (2003), the United Nations and the Institute for Demographic Studies. We combined these data with 1950–1960 data from the Penn World Tables and with 1960–2016 data from the World Bank. We then run an unconditional population growth model (Snijders, 2011) to impute population size for the country-years for which we had missing population data at the time of sound cinema diffusion (1930-1940). We then aggregated country population data by official language (Ethnologue).

Table 1: Dependent variables

	Internet-based TOEFL Score					Paper-based TOEFL score			
	Overall	Reading	Writing	Speaking	Listening	Overall	Reading	Writing	Listening
Dubbing									
Obs.	68	68	68	68	68	68	68	68	68
Mean	78.722	18.358	20.176	20.786	19.404	51.575	51.007	51.259	52.505
Median	78.25	18.375	20	20.75	19.313	51.353	50.732	50.813	52.225
St. Dev.	9.994	3	2.181	2.052	3.169	3.117	3.022	3.325	3.471
Min	60.5	13	16.125	16.625	13.875	41.4	41.5	39.5	43
Max	98.875	24	24.375	25.875	25.75	58.833	58.667	58.833	59
Subtitling									
Obs.	67	67	67	67	67	67	67	67	67
Mean	82.145	19.029	20.682	21.598	20.815	52.662	51.965	51.537	54.559
Median	82.875	19.375	20.875	21.5	21	52.563	52.55	51	53.583
St. Dev.	9.743	3.166	2.102	1.732	3.145	4.188	4.242	4.624	3.993
Min	61.125	12	15.125	18.25	12	44.35	42	42	48.75
Max	100.6	24.4	25.5	26	26.2	61.233	59.667	61	63.333
Diff. Means	-3.424**	-0.671	-0.505	-0.812**	-1.411**	-1.087*	-0.958	-0.278	-2.054***
Std. Errors	(1.699)	(0.531)	(0.369)	(0.327)	(0.543)	(0.646)	(0.645)	(0.705)	(0.653)

TOEFL scores are national score averages over the sample period (2008-2015).

Table 2: Other explanatory variables

	Main control variables				Education controls		Colonial past controls		Historical variables	
	Language size	Population	Linguistic proximity	Education expenditures	Staff/student primary	Staff/student secondary	Former UK colony	Former US colony	Historical language size	Historical Polity IV
Dubbing										
Obs.	68	68	68	68	68	68	68	68	68	68
Mean	118.846	42.716	0.109	3.064	16.664	13.471	0.324	0.044	81.846	0.431
Median	47.735	9.073	0	3.064	16.636	13.457	0	0	23.678	0.417
St. Dev.	177.337	134.739	0.148	0.008	0.263	0.194	0.471	0.207	109.393	3.679
Min	0.367	0.367	0	3.034	16.212	13.126	0	0	0.586	-8.333
Max	907.757	890.634	0.5	3.082	17.825	14.263	1	1	510.586	10
Subtitling										
Obs.	67	67	67	67	67	67	67	67	67	67
Mean	140.972	12.487	0.153	3.064	16.635	13.451	0.373	0.03	72.067	-0.587
Median	191.240	4.751	0.122	3.064	16.636	13.457	0	0	105.76	0.417
St. Dev.	141.839	23.117	0.143	0.01	0.231	0.175	0.487	0.271	73.261	5.363
Min	0.211	0.211	0	3.026	16.182	13.069	0	0	0.773	-10
Max	9.7.757	140.596	0.5	3.09	17.528	14.453	1	1	510.586	10
Diff. Means	-22.126	30.229*	-0.044*	-0.0003	0.029	0.02	-0.05	0.014	9.779	1.018
Std. Errors	0.309	0.213	0.025	0.002	0.043	0.032	0.083	0.033	16.292	0.803

All variables use averages over the sample period (2008-2015). The data on language size is obtained by computing the sum of the populations of the countries in which each language is official (from Ethnologue). Language proximity is based on Melitz and Toubal's (2014) index. Education expenditures are equal to public expenditures per student over GDP per capita. It is expressed in logarithms.

Table 3: Subtitling and English proficiency (OLS)

	(1)	(2)	(3)	(4)	(5)
Subtitling	0.042	0.043	0.04	0.054	0.02
	(0.015)***	(0.016)***	(0.016)**	(0.025)**	(0.046)
Language size	-.00005	-.00005	-.00005	0.0001	0.0002
	(0.00009)	(0.00009)	(0.00008)	(0.00009)	(0.0002)
Country population	0.0002	0.0002	0.0001	-.00005	-.00006
	(0.0001)*	(0.0001)*	(0.0001)	(0.0001)	(0.0001)
Education expenditures	1.677	1.653	1.568	0.853	3.105
	(0.706)**	(0.718)**	(0.764)**	(1.159)	(1.431)**
Linguistic proximity	0.31	0.311	0.324	0.507	0.191
	(0.056)***	(0.057)***	(0.056)***	(0.116)***	(0.243)
Education controls	N	Y	Y	Y	Y
Colonial past controls	N	N	Y	Y	Y
Restricted sample	N	N	N	Y	Y
Language dummies	N	N	N	N	Y
Obs.	135	135	135	75	75

*** p<0.01, ** p<0.05, * p<0.1 All variables are aggregated using time-averages. Education expenditures are included in logs. Estimations are done by OLS. We also performed regressions for each of the 2008-2015 years separately and obtained coefficients of subtitles from 0.029 to 0.048. All those coefficients are significant at the 5% level. Our regression results are robust to the inclusion of a variable that controls for the proportion of imputed observations.

Table 4: Subtitling and English proficiency (IV)

	S	RF	FS	IV
	(1)	(2)	(3)	(4)
Subtitling				0.169
				(0.051)***
Language size at	-.007	-.001	-.008	
sound cinema diffusion	(0.001)***	(0.0003)***	(0.001)***	
Language size	0.005	0.0007	0.005	-.0001
	(0.0008)***	(0.0002)***	(0.0008)***	(0.0001)
Country population	-.001	0.00006	-.001	0.0003
	(0.0003)***	(0.0001)	(0.0003)***	(0.0002)*
Education expenditures	-.019	2.207	1.959	1.877
	(6.107)	(0.797)***	(5.879)	(1.007)*
Linguistic proximity	0.181	0.281	0.321	0.227
	(0.358)	(0.057)***	(0.349)	(0.091)**
Polity	0.006			
	(0.012)			
Obs.	135	135	135	135

*** p<0.01, ** p<0.05, * p<0.1 All variables are aggregated using time-averages. Education expenditures are included in logs. All regressions include education controls (staff/student ratio in primary and secondary education) and colonial past controls (former UK colony and former US colony). The dependent variable is internet-based TOEFL scores in the reduced form and instrumental variables estimations and subtitles in the first stage. Subtitles is instrumented by language size at the time of sound cinema diffusion. Estimations are done by OLS and 2SLS. We also performed regressions for each of the 2008-2015 years separately and obtained coefficients of subtitles from -0.001 to -0.0008 in the RF, from -0.0068 to -0.0061 in the FS and 0.118 to 0.165 in the IV. All those coefficients are significant at the 1% level. Our regression results are robust to the inclusion of a variable that controls for the proportion of imputed observations.

Table 5: Subtitling and English proficiency by skill

	Reading	Writing	Speaking	Listening
Subtitling	0.183	0.126	0.119	0.252
	(0.069)***	(0.042)***	(0.036)***	(0.067)***
Language size	-.0001	-.0001	-.0001	-.0002
	(0.0001)	(0.0001)	(0.00008)	(0.0001)
Country population	0.0004	0.0003	0.0002	0.0004
	(0.0002)**	(0.0001)*	(0.0001)	(0.0002)*
Education expenditures	2.326	1.331	1.586	2.244
	(1.172)**	(0.813)	(0.874)*	(1.398)
Linguistic proximity	0.285	0.175	0.2	0.259
	(0.116)**	(0.077)**	(0.069)***	(0.122)**
Obs.	135	135	135	135

*** p<0.01, ** p<0.05, * p<0.1. All variables are aggregated using time-averages. Education expenditures are included in logs. Subtitles is instrumented by language size at the time of sound cinema diffusion. All regressions include education controls (staff/student ratio in primary and secondary education) and colonial past controls (former UK colony and former US colony). Estimations are done by 2SLS. Our regression results are robust to the inclusion of a variable that controls for the proportion of imputed observations.

Appendix

Table A.1: Translation mode, language and English skills by country

DUBBING				SUBTITLING			
Country	Language	Internet	Paper	Country	Language	Internet	Paper
Algeria	Arabic	72	50.2	Afghanistan	Afghan Persian	70.9	49.1
Armenia	Armenian	78.6	52.2	Albania	Albanian	79	46.6
Austria	German	98.9	49.9	Angola	Portuguese	65.6	50
Azerbaijan	Azerbaijani	76.9	51	Argentina	Spanish	92.5	55.6
Bangladesh	Bangla	83.4	50.8	Bahrain	Arabic	79.6	45.9
Belarus	Russian	86.1	54.5	Bolivia	Spanish	81.9	52.9
Benin	French	64.75	49.05	Bosnia	Bosnian	84.5	54.8
Bulgaria	Bulgarian	88.1	56.6	Brazil	Portuguese	84.9	52.8
Burkina Faso	French	65.4	48.6	Cambodia	Khmer	67	48.3
Burundi	Kirundi	69.5	48.04	Colombia	Spanish	81	51
Chad	French	66.6	48.4	Costa Rica	Spanish	92.5	56.5
Chile	Spanish	83.9	54.5	Croatia	Croatian	91.1	58.2
China	Mandarin	76.9	51.5	Cuba	Spanish	79.6	52.7
Congo	French	63.1	46.6	Cyprus	Greek	82.9	55.8
Cote d'Ivoire	French	64.5	48.5	Denmark	Danish	98.9	60.8
Czech Republic	Czech	90.8	54.6	Dominican Republic	Spanish	81.3	54.3
DR of the Congo	French	69.3	47.4	Ecuador	Spanish	78.8	51.6
Equatorial Guinea	Spanish	61.5	47.5	Egypt	Arabic	81.6	50.3
Eritrea	Tigrinya	78.8	55.6	El Salvador	Spanish	85	49.8
Ethiopia	Oromo	74.6	52.1	Estonia	Estonian	94.3	-
France	French	87.8	53.9	Faeroe Islands	Faroese	87.75	-
Gabon	French	69.5	49	Finland	Finnish	95.8	61
Germany	German	96.3	54.9	Georgia	Georgian	80.6	53.5
Guinea	French	62.5	48.7	Greece	Greek	90.1	52.5
Hong Kong	Cantonese	82.1	51.6	Guatemala	Spanish	81.5	54
Hungary	Hungarian	90.6	58.8	Haiti	French	63.4	49.9
India	Hindi	90.5	53.9	Honduras	Spanish	84.9	50.8
Iran	Persian	80	50.9	Iceland	Icelandic	94.5	61.2
Italy	Italian	89.1	53.8	Indonesia	Bahasa	80.4	51.1
Japan	Japanese	69.1	50.8	Iraq	Arabic	70.1	48
Kazakhstan	Kazakh	78.3	49.5	Israel	Hebrew	93.5	53.8
Kyrgyzstan	Kyrgyz	77.6	51.8	Jordan	Arabic	76.5	50.6
Laos	Lao	65.4	41.4	Kosovo	Albanian	75.3	56.2
Latvia	Latvian	87.1	56	Kuwait	Arabic	69	44.35
Lesotho	Sesotho	77.3	46.5	Lebanon	Arabic	84.5	51.1
Lithuania	Lithuanian	86.7	55.1	Libya	Arabic	70	47.6
Luxembourg	Luxembourgish	95.5	56.5	Macedonia	Macedonian	85.6	54.6
Madagascar	French	78.4	52	Malaysia	Bahasa Malaysia	88.6	53.5
Mali	French	60.5	47.7	Mauritania	Arabic	65.6	49.7
Monaco	French	87	51.9	Mexico	Spanish	85.6	54
Morocco	Arabic	78	51.2	Moldova	Romanian	84.1	50.4
Mozambique	Portuguese	70.25	49	Mongolia	Mongolian	72.3	48.5

DUBBING				SUBTITLING			
Country	Language	Internet	Paper	Country	Language	Internet	Paper
Nepal	Nepali	78.5	50.2	Montenegro	Montenegrin	80.5	59.7
Niger	French	68.8	52.6	Netherlands	Dutch	100.6	60.2
North Korea	Korean	78.4	49.2	Nicaragua	Spanish	84.5	49.1
Pakistan	Urdu	89.3	53.3	Norway	Bokman Norwegian	92.9	57.5
Philippines	Filipino	88.6	53.6	Oman	Arabic	66.5	47.8
Poland	Polish	88.9	53.8	Panama	Spanish	82.9	54.6
Russia	Russian	84.6	53.4	Paraguay	Spanish	84.9	53.8
Rwanda	Kinyarwanda	71.4	50.7	Peru	Spanish	85.5	50.9
Senegal	French	65.6	49.9	Portugal	Portuguese	95	54
Slovakia	Slovak	89.1	55.5	Puerto Rico	Spanish	86	48.2
Somalia	Somali	70.6	50.3	Qatar	Arabic	71.6	45.3
South Africa	IsiZulu	96.8	56.5	Romania	Romanian	91.4	58.1
Spain	Spanish	89.1	54.7	Saudi Arabia	Arabic	61.1	46.5
Sri Lanka	Sinhala	83.4	53.5	Serbia	Serbian	86.8	56.1
Switzerland	German	97	56.2	Singapore	Mandarin	98.4	59.4
Taiwan	Mandarin	77.1	49.9	Slovenia	Slovenian	94.4	59.5
Tajikistan	Tajik	68.4	49.1	South Korea	Korean	82.3	51.1
Tanzania	Swahili	71.5	53.4	Sudan	Arabic	74.6	48.9
Thailand	Thai	74.9	48.2	Sweden	Swedish	92.3	56.9
Togo	French	65.75	49	Syria	Arabic	76.6	49.4
Tunisia	Arabic	78.3	52.1	Timor Leste	Portuguese	62	-
Turkey	Turkish	76.5	49.9	Uruguay	Spanish	93.9	56.6
Turkmenistan	Turkmen	77.1	50.7	Venezuela	Spanish	82.8	53.4
Ukraine	Ukranian	83.9	53.2	Vietnam	Vietnamese	75.4	50.9
Uzbekistan	Uzbek	76.8	50	Yemen	Arabic	68.8	49.7
Zimbabwe	Shona	89.88	56.2				

English skills are measured using the national score averages of the TOEFL scores over the sample period (2008-2015), in both the "internet" and "paper"-based versions of the test.

Table A.2: Robustness of the results to the log transformations

Panel A: Adding language size and country population in logs

	OLS	RF	FS	IV
	(1)	(2)	(3)	(4)
Subtitling	0.042			0.153
	(0.017)**			(0.049)***
Language size at sound cinema diffusion		-.001	-.008	
		(0.0003)***	(0.001)***	
Language size	0.0001	0.0008	0.004	0.0001
	(0.0001)	(0.0002)***	(0.0008)***	(0.0001)
Country population	0.00002	0.00004	-.00005	0.00005
	(0.0002)	(0.0002)	(0.0004)	(0.0001)
Language size in logs	-.019	-.012	0.089	-.026
	(0.01)**	(0.011)	(0.04)**	(0.011)**
Country population in logs	0.006	-.002	-.142	0.02
	(0.01)	(0.01)	(0.043)***	(0.014)
Education expenditures	1.655	2.248	2.301	1.896
	(0.732)**	(0.776)***	(5.850)	(0.924)**
Linguistic proximity	0.321	0.281	0.257	0.241
	(0.056)***	(0.058)***	(0.324)	(0.084)***
Obs.	135	135	135	135

Panel B: Dependent variable in absolute value

	OLS	RF	FS	IV
	(1)	(2)	(3)	(4)
Subtitling	3.155			12.133
	(1.201)***			(3.753)***
Language size at sound cinema diffusion		-.093	-.008	
		(0.023)***	(0.001)***	
Language size	-.003	0.051	0.005	-.009
	(0.007)	(0.014)***	(0.0008)***	(0.008)
Country population	0.008	0.004	-.001	0.021
	(0.01)	(0.01)	(0.0003)***	(0.012)*
Education expenditures	148.823	194.218	1.959	170.448
	(62.687)**	(65.845)***	(5.879)	(75.725)**
Linguistic proximity	25.724	22.789	0.321	18.892
	(4.422)***	(4.558)***	(0.349)	(6.763)***
Obs.	135	135	135	135

*** p<0.01, ** p<0.05, * p<0.1 This table shows the result of replicating our main estimations including logarithmic transformations of language size and country population in Panel A and using the dependent variable in absolute value in Panel B. All regressions include education controls (staff/student ratio in primary and secondary education) and colonial past controls (former UK colony and former US colony). The dependent variable is internet-based TOEFL scores in the OLS, reduced form and instrumental variables estimations and subtitles in the first stage. We instrument subtitling with the language size at the time of sound cinema diffusion in levels, Our regression results are robust to the inclusion of a variable that controls for the proportion of imputed observations.

Table A.3: Paper-based TOEFL scores

	Overall test score				Disaggregated by skill		
	OLS	RF	FS	IV	reading	writing	listening
	(1)	(2)	(3)	(4)	(1)	(2)	(3)
Subtitling	0.021			0.046	0.043	0.018	0.075
	(0.011)*			(0.022)**	(0.024)*	(0.028)	(0.02)***
Language size at sound cinema diffusion		-.0003	-.007				
		(0.0002)**	(0.001)***				
Language size	-.00005	0.0002	0.005	-.00006	-.00004	-.00009	-.00006
	(0.00006)	(0.0001)	(0.0007)***	(0.00006)	(0.00006)	(0.00008)	(0.00006)
Country population	0.00008	0.00005	-.001	0.0001	0.0001	0.0001	0.0001
	(0.00007)	(0.00007)	(0.0003)***	(0.00008)	(0.00007)	(0.0001)	(0.00008)
Education expenditures	1.022	1.138	-.279	1.151	0.82	1.440	1.182
	(0.727)	(0.755)	(5.985)	(0.704)	(0.699)	(0.777)*	(0.716)*
Linguistic proximity	0.152	0.149	0.444	0.129	0.14	0.129	0.122
	(0.043)***	(0.046)***	(0.355)	(0.052)**	(0.055)**	(0.059)**	(0.048)**
Obs.	135	135	135	135	135	135	135

*** p<0.01, ** p<0.05, * p<0.1 This table shows the result of replicating our main estimations including all controls using paper-based TOEFL score as an alternative dependent variable. All regressions include education controls (staff/student ratio in primary and secondary education) and colonial past controls (former UK colony and former US colony). The dependent variable is subtitles in the first stage. Our regression results are robust to the inclusion of a variable that controls for the proportion of imputed observations.

Table A.4: Subtitling and English proficiency with yearly data

	Overall test score				Disaggregated by skill			
	OLS	RF	FS	IV	reading	writing	speaking	listening
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Subtitling	0.037			0.133	0.143	0.1	0.109	0.225
	(0.015)**			(0.045)***	(0.06)**	(0.04)**	(0.033)***	(0.061)***
Language size at sound cinema diffusion	0.038			0.142				
	(0.015)**			(0.046)***				
		-9.12e-10	-6.40e-09					
		(2.39e-10)***	(8.91e-10)***					
Language size	-2.76e-11	2.93e-10	2.64e-09	-8.33e-11	-5.63e-11	-8.12e-11	-7.34e-11	-1.21e-10
	(4.77e-11)	(8.02e-11)***	(3.24e-10)***	(6.50e-11)	(7.50e-11)	(6.61e-11)	(4.97e-11)	(8.06e-11)
Country population	8.76e-11	5.56e-11	-8.51e-10	1.77e-10	2.29e-10	1.68e-10	9.85e-11	2.17e-10
	(6.88e-11)	(5.97e-11)	(2.22e-10)***	(8.82e-11)**	(9.13e-11)**	(8.53e-11)**	(8.15e-11)	(1.15e-10)*
Education expenditures	0.06	0.082	0.034	0.077	0.093	0.054	0.069	0.098
	(0.031)*	(0.033)**	(0.219)	(0.038)**	(0.044)**	(0.031)*	(0.035)**	(0.053)*
Linguistic proximity	0.374	0.353	0.36	0.302	0.391	0.235	0.244	0.354
	(0.051)***	(0.052)***	(0.286)	(0.075)***	(0.096)***	(0.064)***	(0.059)***	(0.102)***
Obs.	1,054	1,054	1,054	1,054	1054	1,054	1,054	1,054

*** p<0.01, ** p<0.05, * p<0.1 This table uses yearly data. Education expenditures are included in logs. All regressions include education controls (staff/student ratio in primary and secondary education) and colonial past controls (former UK colony and former US colony). Standard errors are clustered at the country level. Our regression results are robust to the inclusion of a variable that controls for the proportion of imputed observations.

Table A5: Paper-based TOEFL scores with yearly data

	Overall test score				Disaggregated by skill		
	OLS	RF	FS	IV	reading	writing	listening
	(1)	(2)	(3)	(4)	(1)	(2)	(3)
Subtitling	0.006			0.038	0.035	0.009	0.067
	(0.008)			(0.018)**	(0.019)*	(0.02)	(0.019)***
Language size at		-2.36e-10	-6.20e-09				
sound cinema diffusion		(1.05e-10)**	(1.01e-09)***				
Language size	-1.03e-11	7.07e-11	2.67e-09	-3.07e-11	-1.48e-11	-3.92e-11	-3.72e-11
	(2.63e-11)	(3.84e-11)*	(3.75e-10)***	(3.23e-11)	(2.75e-11)	(3.80e-11)	(3.48e-11)
Country population	3.08e-11	2.63e-11	-8.46e-10	5.84e-11	4.75e-11	4.91e-11	7.51e-11
	(2.86e-11)	(2.68e-11)	(2.39e-10)***	(3.61e-11)	(3.00e-11)	(4.20e-11)	(3.95e-11)*
Education expenditures	0.022	0.028	0.021	0.027	0.014	0.035	0.032
	(0.022)	(0.023)	(0.222)	(0.02)	(0.019)	(0.025)	(0.021)
Linguistic proximity	0.189	0.182	0.434	0.166	0.172	0.169	0.162
	(0.029)***	(0.029)***	(0.293)	(0.034)***	(0.037)***	(0.036)***	(0.034)***
Obs.	717	717	717	717	717	717	717

*** p<0.01, ** p<0.05, * p<0.1 This table replicates the results of the estimations with paper-based TOEFL as dependent variable scores including all controls using yearly data, and controlling for year dummies. All regressions include education controls (staff/student ratio in primary and secondary education) and colonial past controls (former UK colony and former US colony). The dependent variable is subtitles in the first stage. Standard errors are clustered at the country level. Our regression results are robust to the inclusion of a variable that controls for the proportion of imputed observations.