The economics of literary translation: Some theory and evidence

Victor Ginsburgh a,*, Shlomo Weber b,c, Sheila Weyers d

a ECARES, Brussels and CORE, Louvain-la-Neuve, Belgium
b Southern Methodist University, Dallas, USA
c New Economic School, Moscow, Russia
d Université catholique de Louvain, Belgium

Abstract

Books are an important factor of cultural transmission, but often need to be translated to achieve this goal. English is sometimes accused of dominating in terms of translations. We develop a theoretical model, which is estimated using UNESCO translation data. We show that if account is taken of factors such as production in the source and the destination languages, as well as distances between cultures, then translations from English are surpassed in relative terms by translations from other idioms, including Scandinavian languages and French; their position as destination of literary translations is however fairly weak.
© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The economic literature on cultural exchanges deals essentially with media industries, and especially with movies and television programs.1 Much less is written on music, which needs no translation, dubbing or subtitles. But as pointed out by Frith (1996, p. 157) well before the explosion of Internet, MP3 and iTunes, “the point is not that a new technology enabled – or determined – a new music international, but, rather, that the music’s own essential mobility enabled the new technology to flourish, and shaped the way it worked.”

Though television and broadcasting have changed considerably the way culture is transmitted, books (and more generally written material, including the Web) remain essential.

* Corresponding author.
E-mail address: vginsbur@ulb.ac.be (V. Ginsburgh).
1 See Hoskins et al. (1997) and the list of references therein.
As Susan Sontag pointed out when she received the Peace Prize at the Frankfurt Book Fair in 2003:

[W]hat saved me as a schoolchild in Arizona, waiting to grow up, waiting to escape into larger reality, was reading books, books in translation as well as those written in English. To have access to literature, world literature, was to escape the prison of national vanity, of philistinism, of compulsory provincialism, of inane schooling, of imperfect destinies and bad luck. Literature was the passport to enter a larger life; that is, the zone of freedom. (Sontag, 2005, p. 222)

Quantities exchanged may be impressive. As of 2009, sixty million copies of Dan Brown’s *Da Vinci Code*, published in early 2003 were in print or sold. The *Code* was translated into 44 languages, and by October 2004, it had generated some sixteen titles supporting or debunking the code.

In some cases, books even become part of national literary heritage in their translated version. The lovely novel, *Ali and Nino: A Love Story* – which became a national emblem in Azerbaijan – was written in German by Baku-born author Lev Nussimbaum under the pseudonym of Kurban Said and was first published in Vienna in 1937. More generally, Casanova (2002, 2008, p. 202) points out that some authors become renown through the translated version of their work. This was the case, among others, for Jorge Luis Borges, Rabindranath Tagore, James Joyce, Franz Kafka in the past, but it is still so today. Milena Agus’s novel *Mal di Pietre* (2006), written in Italian, became known in its French translation before gaining acceptance in Italy, and it is now translated into 13 languages.

Translations, an obvious form of economic exchange, should have interested economists much more than it did, but the literature is surprisingly small. It includes Melitz’s (2007) seminal paper, but otherwise, the word “translation” appears neither in the sociological classic, *Books*, by Coser et al. (1982), nor in the very comprehensive survey of the book industry by Canoy et al. (2006). Caves (2000) discusses the economics of books’ production at great length but is mute on translations. The economic study closest to our model is the one by Hjorth-Andersen (2001), who estimates a three-equation model of translations, where the first equation identifies the total number of titles in a given country, the second determines the aggregate propensity to translate, and the third disaggregates this total into single languages. Melitz (2007) is, to our knowledge, the first economist who notes that “if one language is sufficiently larger than others in the sales of original-language works, it will tend to crowd out the rest in translations. . . [and] those writing in the dominant language are privileged.” A similar opinion was recently expressed in one of the important French literary bi-monthly, *La Quinzaine Littéraire* (2006), claiming that translations from English into French are dominant in France. The title of the article “Fiction is American” is unequivocal.

Meanwhile, the topic of translation is very widely discussed recently by sociologists, who consider cultural exchanges – and, in particular, exchanges of the literature – as being “a relatively autonomous sphere with economic [but also] political and symbolic dimensions” (Heilbron, 1999, p. 432). Heilbron and Sapiro (2002, 2007) underline the role that the ebb and
flows of dominating and dominated languages\(^5\) have on translations and, therefore, on the transmission of political ideas. Translations may, thus, be motivated not only by economic but also by cultural, social and ideological factors. This makes translations – and, more generally, trade in cultural goods such as movies – very different from exchanges of commodities.

Ganne and Minon (1992) were among the first to show that France, Italy, Spain and Germany translate much more (18, 25, 26 and 15%, respectively) than the United Kingdom (3.3%). Heilbron (1999) describes the system as accounting for uneven flows between language groups: On the European continent, 50–70% of the published translations are from English, though “language groups do not always coincide with nation states, [since] languages have a supranational character” (Heilbron, 1999, p. 342). Janssen (2009) takes a different angle by analyzing the coverage of foreign literatures in American,\(^6\) Dutch, French and German “elite” newspapers between 1955 and 2005.\(^7\) Her empirical results lead her to emphasize the following conclusions, which do not always confirm her \textit{a priori} assumptions: (i) the coverage of foreign literature increases over time; (ii) so does the number of countries and regions from which the literature originates; (iii) this coverage depends on the size and the centrality\(^8\) of the countries’ literary production; (iv) more attention is paid to literature coming from neighboring countries or countries of the same language area; and (v) the coverage of literature in English has become more important.

Sapiro (2010) uses Bourdieu’s approach and compares the markets for translations in France and the US. She finds that while large publishers behave as actors who are dominated by economic incentives, this is less the case for small-scale producers—who, especially in some European countries (France in particular), make use of financial support from the state, as well as from some private bodies. She also points out that in both countries, France and the US, there is a difference between the up-market, which is more diversified, and the market for commercial literature and bestsellers, which is subject to globalization.

All this may be for the better or the worse, but is hardly surprising because the population speaking English as a first language is, with the exception of those speaking Mandarin, the largest

---

\(^5\) The current role of English was predicted by Alphonse de Candolle, a prominent Swiss scientist (not a linguist), in an essay written in French, and published in 1873. Thus writes de Candolle (1873): “[O]ne has to think why some languages are preferred to others. During the 17th and 18th centuries, there existed good reasons for which French succeeded to Latin in Europe. At the time, French, a relatively simple and clear language, was spoken by a large proportion of instructed people. It had the advantage of being close to Latin which was well-known. An Englishman or a German would know half of French if he knew Latin. A Spaniard or an Italian would know three quarters of it. A discussion, or a paper written in French would have been understood by everybody. During the 19th century, civilization has extended to the north [of France] and the population has increased more there than in the south. The use of English tripled as a consequence of the United States. Science is more and more cultivated in Germany, England, Scandinavia and Russia. The following conditions have to be satisfied for a language to become dominant: (1) include enough words, or Germanic and Latin forms to be understood by both German and Latin populations; (2) be spoken by a large number of civilized people. It would also be convenient for such a language to be simple and clear. English is the only language that could, within the next 50 or 100 years, offer all these conditions. The future predominance of the Anglo-American language is obvious.” (Our translation. Thanks are due to Dean Simonton for having made us aware of Candolle.)

\(^6\) In her study, books written in English elsewhere than in America are considered “foreign.” And indeed in 2005, for example, some 18% of books discussed in American national media originate from the UK, Canada, Australia, South Africa, and Ireland. Given that 76% are American, this leaves little space (6%) for others.

\(^7\) It is interesting to point out that a new monthly literary journal, \textit{Books}, is published in France since 2008. It covers international literature, and it includes articles on books that have not yet been translated into French. The website of the journal aims at “sheding light on topical issues and affairs through books published in the whole world. [The journal translates] the best articles that deal with those books in the world press.”

\(^8\) Heilbron (1999, p. 433) defines centrality as follows: “A language is more central, when it has a larger share in the total number of translated books worldwide.”
in the world (see below). Moreover, English is spoken in countries with very distinct cultures. It can therefore be expected that more fiction be produced in English than in many other languages. We nevertheless show that – if account is taken of factors such as production in the source language, reception in the destination language and proximity between cultures – translations from English, in fact, fall behind translations from other languages, including Scandinavian idioms and French. One can also expect that fiction written in English is more diversified, as it originates from writers located in culturally diverse countries: Great Britain, Ireland, Canada, the United States, and former British colonies in Africa, South Africa, Australia, New Zealand and Tasmania, India, or the West Indies.\textsuperscript{9} This may explain why publishers in these countries do not have the same interest to translate from other languages, as the diversity that readers look for is available without translation. And indeed, our results show that in this sense, the exchange between English and other languages is unequal as pointed out by sociologists, but should not be overestimated.

Indeed, let us assume that data exist on the number of literary works produced, as well as on the number of readers in each language. It is reasonable to think that more books will be written in languages that are native to many speakers and these languages are likely to translate less to their own than are small ones (since more books are available in the former's own language). In addition, and as mentioned above, widely used languages (such as English, Spanish or Arabic) are often spoken in several countries, and they yield a larger degree of diversity in cultures and literature than languages that are spoken in smaller countries such as, say, Lithuania.\textsuperscript{10} If this is true, then the English domination in literature could partly be due to the fact that English, the mother tongue of some 400 million people (Crystal, 2001), produces more fiction than any other European language – including Spanish and Portuguese, spoken by 270 and 175 million as a first language (Crystal, 2001)\textsuperscript{11} – and that more books are translated from English.\textsuperscript{12} Table 1, which tabulates (per head in each native language) the number of literary works translated from some European languages illustrates that this is roughly the case. Indeed, even though a large number of books is translated from English (0.95 per 1000 speakers of English), and very few are translated into English (0.043 per thousand speakers of English as well), there are almost as many that are translated from French (0.88 per thousand speakers of French) and many more from Danish, Norwegian and Swedish – with respectively 1.33, 1.32 and 1.76 books per thousand speakers of each language. Countries with smaller populations – such as Denmark, Estonia, Finland, Iceland, The Netherlands, Norway or Slovenia—translate more into their own language than those whose languages are spoken by large populations. This, of course, explains the staggering number, 22.09, of the ratio “From English/To English” in the third column of Table 1—which means that (per head) the number of translations

\textsuperscript{9} To illustrate this diversity, we sampled some authors who were shortlisted by, or winners of, the Booker (now Man Booker) prize since its inception in 1969. Here is a list of such English writing authors born elsewhere than in Great Britain, although some may have later moved there, and whose native culture often impregnates their work and provides diversity: Chinua Achebe (Nigeria), Aravind Adiga (India), Margaret Atwood (Canada), André Brink (South Africa), John M. Coetzee (South Africa), Ashmat Dangor (South Africa), Anita Desai (India), Kiran Desai (India), Amitav Ghosh (India), Nadine Gordimer (South Africa), Romesh Gunesekera (Sri Lanka), Abdulrazak Gurnah (Zanzibar), Mohsin Hamid (Pakistan), Thomas Keneally (Australia), Doris Lessing (Rhodesia), Yann Martel (Canada), Timothy Mo (Hong Kong), V.S. Naipaul (West Indies), Ben Okri (Nigeria), Michael Ondaatje (Sri Lanka), Arundhati Roy (India), Salman Rushdie (India), Indra Sinha (India), and Ahdaf Soueif (Egypt). This may also be the case for France, Spain, and other important languages.

\textsuperscript{10} We often make the assumption that a language is spoken in a unique country. This is roughly the case, with some exceptions.

\textsuperscript{11} And even more than Arabic and Hindi, which are native to less people than English. The only serious exception would be Mandarin Chinese.

\textsuperscript{12} See Pym (1999) for a very similar argument.
English is 22 times larger than (per head) the number of translations into English. With the exception of Swedish, all other ratios are smaller than one.

These considerations take no account of another important factor in comparing the number of translations: the role of cultural proximities—an assumption that is also made by Janssen (2009), although she does not explicitly uses it in her model. Our theoretical and empirical investigations are in line with the view that—since more fiction is produced in English, which is culturally closer to other Indo-European languages than Mandarin, Arabic or Hindi—more will be translated from English than from other languages, even though these languages have large numbers of native speakers (e.g., Mandarin Chinese is spoken by 1.2 billion people; Crystal, 2001). This argument is reinforced by the fact that translations are relatively expensive. De Swaan (1993, p. 45) estimates translation costs at 30% of the price of a 300-page book of which 2000 copies (“certainly not a too conservative estimate,” according to De Swaan) are circulated. It follows that publishers will take little risk to translate books that have a low probability of being read.13

---

Table 1
Number of translations 1979–2002 (per 1000 native speakers).

<table>
<thead>
<tr>
<th>Language</th>
<th>From (1)</th>
<th>To (2)</th>
<th>Ratio (1):(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulgarian</td>
<td>0.158</td>
<td>1.182</td>
<td>0.13</td>
</tr>
<tr>
<td>Czech</td>
<td>0.333</td>
<td>1.702</td>
<td>0.20</td>
</tr>
<tr>
<td>Danish</td>
<td>1.329</td>
<td>5.730</td>
<td>0.23</td>
</tr>
<tr>
<td>Dutch</td>
<td>0.276</td>
<td>2.159</td>
<td>0.13</td>
</tr>
<tr>
<td>English</td>
<td>0.950</td>
<td>0.043</td>
<td>22.09</td>
</tr>
<tr>
<td>Estonian</td>
<td>–</td>
<td>5.480</td>
<td>–</td>
</tr>
<tr>
<td>Finnish</td>
<td>0.312</td>
<td>3.330</td>
<td>0.09</td>
</tr>
<tr>
<td>French</td>
<td>0.884</td>
<td>1.069</td>
<td>0.83</td>
</tr>
<tr>
<td>German</td>
<td>0.451</td>
<td>1.220</td>
<td>0.37</td>
</tr>
<tr>
<td>Greek</td>
<td>–</td>
<td>0.500</td>
<td>–</td>
</tr>
<tr>
<td>Hungarian</td>
<td>0.201</td>
<td>0.667</td>
<td>0.30</td>
</tr>
<tr>
<td>Icelandic</td>
<td>–</td>
<td>14.300</td>
<td>–</td>
</tr>
<tr>
<td>Italian</td>
<td>0.270</td>
<td>0.204</td>
<td>–</td>
</tr>
<tr>
<td>Latvian</td>
<td>–</td>
<td>1.652</td>
<td>–</td>
</tr>
<tr>
<td>Lithuanian</td>
<td>–</td>
<td>1.083</td>
<td>–</td>
</tr>
<tr>
<td>Norwegian</td>
<td>1.319</td>
<td>5.586</td>
<td>0.24</td>
</tr>
<tr>
<td>Polish</td>
<td>0.113</td>
<td>0.482</td>
<td>0.23</td>
</tr>
<tr>
<td>Portuguese</td>
<td>0.023</td>
<td>0.139</td>
<td>0.17</td>
</tr>
<tr>
<td>Romanian</td>
<td>0.073</td>
<td>0.279</td>
<td>0.26</td>
</tr>
<tr>
<td>Russian</td>
<td>0.133</td>
<td>0.156</td>
<td>0.85</td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>0.038</td>
<td>0.291</td>
<td>0.13</td>
</tr>
<tr>
<td>Slovak</td>
<td>0.237</td>
<td>1.248</td>
<td>0.19</td>
</tr>
<tr>
<td>Slovene</td>
<td>–</td>
<td>2.180</td>
<td>–</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.054</td>
<td>0.277</td>
<td>0.19</td>
</tr>
<tr>
<td>Swedish</td>
<td>1.759</td>
<td>1.500</td>
<td>1.17</td>
</tr>
<tr>
<td>Ukrainian</td>
<td>–</td>
<td>0.041</td>
<td>–</td>
</tr>
</tbody>
</table>

Source: UNESCO, Index Translationum.

---

According to an article by Sonia Kolesnikov-Jessop – published by Newsweek on January 11, 2010 – 275,000 titles were published in China in 2008, and 6.9 billion books printed, but only a few dozen of those will likely ever be translated. This situation is changing also. At the 2009 Frankfurt fair where China was invited as the guest of honor, foreign publishers snapped up the rights to about 1300 books. Still, 1300 is a very tiny fraction of 275,000.

---

13
The paper is organized as follows. The theoretical model is discussed first. It leads to a demand equation for translations by a “representative” reader, which lends itself to econometric estimation. Empirical results shed some light on the determinants of translations of fiction and only mildly support the hypothesis of the dominance of English. This may still look worrisome but, at least, it can be explained by demographic, economic and cultural proximity arguments. It leads us to conclude that, although translations from English are indeed dominating in number, they are not disproportionately large.\textsuperscript{14} And translations into English are not that small.

2. The model and data

The main ingredients of the model are the number of titles translated, the number of titles produced in the source language, the number of people who read in the destination language, and the “cultural distance” between the source and the destination language. We expect that the larger the number of titles written in the source language – and the larger the number of people who read in the destination language – then the larger the number of translations. The distance between the two languages should have a negative effect, as more distant cultures will lead to fewer translations. This is very similar to models that are used in international trade, where exchanges between countries are explained (with a high degree of success) by wealth and production in each pair of countries, as well as the distance (geographical, cultural, linguistic, other) between the countries.

Clearly, such a simple framework is not sufficient to account for all the details that generate the number of translations, but it will hopefully help in explaining part of the large inequalities that are observed in Table 1. The part that remains unexplained is due to idiosyncrasies in the partner countries or languages.

Due to very strong data limitations on the production of books, and the number of readers by language, we were forced to rely on other variables than those that are described above. We briefly review these reasons and the data that we were led to use, and also make the reader aware that we restrict our attention to European languages only—thus eliminating from consideration some important languages, such as Mandarin Chinese, Hindi, Arabic, and Japanese.

2.1. Number of titles translated

Data are obtained from UNESCO’s website.\textsuperscript{15} They cover the period 1979–2002. This is how UNESCO describes the data: “The \textit{Index Translationum} is a list of books translated in the world, i.e. an international bibliography of translations. The database contains cumulative bibliographical information on books translated and published in about one hundred of the UNESCO Member States since 1979 and totaling more than 1.7 million entries in all disciplines [including literature].” We concentrated on literature. UNESCO receives the data from bibliography centers or national libraries in the participating countries.

\textsuperscript{14} In fact, some authors even suggest that the share of English in printed and electronic media is declining. Melitz (2007, Table 1, and p. 212) shows data suggesting that the role of English in literary works is decreasing over the last 30 years, and Pfanner (2007) indicates that only 36% of all blog postings on the Internet are in English, while 37% are written in Japanese and the share of Chinese and Spanish blogs is rapidly increasing. Because printed books may (very unfortunately) soon be superseded by the Internet, this may lead to a more balanced world.

\textsuperscript{15} As well as from \url{http://databases.unesco.org/xtrans/stat/xTransList.a}. 
The UNESCO database is often strongly criticized as not very reliable because (a) the definition of a book varies across countries, and (b) the numbers exhibit important fluctuations (Heilbron, 1999). We partly avoid both criticisms because (a) we consider only literature, the definition of which is probably the subject of larger agreement, and (b) we deal with longer periods than single years, so that fluctuations are smoothed out. Note that Heilbron, who offers this critique, claims nevertheless that the UNESCO database is the only readily available international source.

We focus on the main European languages considered as official (thus excluding for instance Catalan), and we chose to discard some languages for which the number of titles translated is too small (Albanian, Moldavian). Appendix A provides an overview of the languages included in our data. As will be seen, some languages (Estonian, Greek, Icelandic, Latvian, Lithuanian, Slovene and Ukrainian) are included as destination languages only, for the number of titles translated from these languages is very small. Our sample includes thus 19 source countries, and 26 destination countries, which leads to 475 (=19 * 26–19) translation flows. We ignored 4 observations (translations from Finnish to Hungarian and Estonian, and from Hungarian to Finnish and Estonian), as there are no data on distances between these three languages. We are thus left with 471 observations on bilateral translations.

2.2. Number of titles produced in the source language

The UNESCO statistics dealing with book publishing seem extremely incomplete and limited – not only by country (which could give us an indication of language, though many countries, including the US, Belgium, Canada, publish in more than one language) – but also when one tries to regroup country statistics into language statistics. The reasons are as follows: (a) The volume edited by Altbach and Hoshino (1995), International Book Publishing, An Encyclopedia – which seems to be the reference on book publishing – contains an article by Gretchen Whitney, who collected all the available UNESCO data on book production by country and for five languages (English, French, German, Spanish and Russian). The data are extremely incomplete, even for these very important languages. For several years there are no data for Canada and the US. Here is what Whitney (1995, p. 168) writes: “North America encompasses 25 countries, including those of Central America, and four large producers: the US, Canada, Cuba and Mexico. The data for the first two are unusable: The US shows figures in the 88,000–99,000 title range for 1970–1976, 1978 and 1981 only; for Canada, there are no data after 1980, when nearly 20,000 titles were produced. Mexico is missing but for six years, etc.” (b) Moreover, even if these data can be retrieved, they are by country, and not by language; and here, we deal with languages and not countries. So we would have had to check how many books in French come from Belgium or from Canada (two languages in each country); how many titles in English were written in Canada, India, East or West Africa; how many titles in Russian or Ukrainian in Ukraine; add for Spanish, the books written in every Latin American country; and so on. (c) It is of course possible to do approximations by taking, for instance, the maximal or the average number of books published for those years for which numbers exist for a country. However, this does not seem a good alternative, for not all countries publish such numbers for the same years.16

---

16 This could also make suspicious the data on translated books. Our understanding is, however, that the translation data that UNESCO produces in its Index Translationum come from another source (national libraries) than the production statistics, so they may hopefully be better.
2.3. Number of readers in the destination country

Except for some surveys organized from time to time in specific countries, no data on reading habits are available. Using the number of books produced would be an alternative, but this does not work either given what is described above.

Given the very incomplete data on the number of titles produced and on the number of readers in the destination country, we rely on the populations of native speakers (Crystal, 2001) in both the source and the destination language. Though this is unfortunately a second best choice, it can be rationalized since the larger the population, the more variety there may be in types of books and situations described.

2.4. Distances between cultures

The data that we use to represent distances between cultures require some explanation. The measure is based on so-called cognate data, established as follows. For each meaning from a list of 200 basic meanings (such as father, mother, digits, etc.) selected by Swadesh (1952)—Dyen et al. (1992) collected the words used in 95 Indo-European speech varieties (i.e., languages and dialects) and classified these into cognate classes. For a given meaning, such a class contains all the words from different speech varieties that have an unbroken history of descent from a common ancestral word. The distance between two languages \(i\) and \(j\) is then equal to the percentage of words in the two languages that do not descend from a common word. This distance will thus be close to 1 if the two languages have completely different roots (say English and Finnish, a non Indo-European language) and close to 0 otherwise (Slovak and Czech). In our context, this distance is not meant to measure the difficulty of translating from \(i\) to \(j\), which is more complex than just the relative proximities of vocabularies. We rather follow Cavalli-Sforza (2000) and assume that linguistic distances are a proxy of cultural distances.

The reader is referred to Appendix B, wherein we develop a microeconomic model that leads to an equation in which the number of titles translated from a source language to a destination language is determined by the following variables: (a) the sizes of the populations that speak the source and the destination languages as first language; and (b) the distances between the two languages.

---

17 Though cultural distances are available to some extent (see Hofstede, 1980, 1991, as well as Hofstede’s websites http://spiswww.uvt.nl/web/iric/hofstede/page3.htm and http://geert-hofstede.international-business-center.com/index.shtml), we will use the linguistic distances computed by Dyen et al. (1992). The reason is twofold: (a) cultural distances are available for countries, not for languages, while UNESCO data are for languages; (b) cultural distances exist only for a small number of countries, and certainly not all those that are in our sample of translations between languages.

18 Words borrowed from an other language are thus excluded.

19 Distances between Finnish, Hungarian and Estonian and Indo-European languages were set to 1.

20 See, for example, Catford (1967) and Nida and Taber (1969). Nida (1975, p. 98) is very explicit about the two questions that confront translation: “The first concerns translation as an art rather than a science and the second raises the issue as to whether translation is even possible.” The German poet Heine claimed that his poems translated into French, were just “moonlight stuffed with straw,” and Nabokov who used to write indifferently in English and Russian notes (in ‘On translating Eugen Onegin’, one of his poems) that translation is “ ‘On a platter a poet’s pale and glaring head, a parrot’s screech, a monkey’s chatter, and profanation of the dead.’ In ancient times, translation (of God’s words) was blasphemy. The Roll of Fasting (first century A.D.) “records the belief that three days of utter darkness fell on the world when the Law was translated into Greek.” See Steiner (1992, pp. 251–252). But this is a matter for hermeneutics.
In a variant, we add as explanatory variables the literacy rate and the average income of the population speaking the destination language. The data are described in Appendix C.

In this model, a unique parameter is estimated for all source languages and another one for all destination languages. This may (and does indeed) hide the disparities between languages that are singled out by sociologists and that are illustrated in a straightforward way in Table 1. To mitigate this, we estimate two other forms of the model in which either source or destination languages are disaggregated. This is further discussed in Section 3.

3. Empirical results

The model discussed in Section 2 and in Appendix B leads us to estimate the following equation, using ordinary least squares:

$$\ln t_{ij} = \alpha_0 + \alpha_1 \ln P_i + \alpha_2 \ln P_j + \alpha_3 \ln D_{ij} + \alpha_4 \ln L_j + \alpha_5 \ln W_j + v_{ij},$$

where all the variables are expressed in logarithms (ln)—$t_{ij}$ is the number of translations from source language $i$ to destination language $j$; $P_i$ and $P_j$ are the sizes of the populations that speak $i$ and $j$ as first language; $D_{ij}$ is the distance between $i$ and $j$; $L_j$ and $W_j$ represent the literacy rate and the average income of the population speaking destination language $j$; and $v_{ij}$ is an error term. Note that we also estimate the model without these two additional variables. The $\alpha$ are parameters to be estimated; they can easily be obtained from combinations of the parameters of Eq. (A4) in Appendix B. The results discussed above (and summarized in the Proposition stated in Appendix B) imply that $\alpha_1$ and $\alpha_2$ should be positive (the larger the source or destination languages, the larger the number of translations), while $\alpha_3$ should be negative (the larger the distance between two languages, the smaller the number of translations). Literacy and income of the population speaking the destination language are expected to have a positive influence, for more books will be read (and thus translated from other languages) in more literate and richer regions; $\alpha_4$ and $\alpha_5$ should therefore also be positive; $\alpha_0$ is an intercept term.

Estimation results for literary translations between 1979 and 2002 are shown in Table 2. They include 19 source and 26 destination languages. The first equation gives the results, ignoring literacy rate and income in the destination population. All parameters are significantly different from zero at the one percent probability level (which means that they can hardly be considered not to influence the number of translations), and they all carry the expected signs discussed earlier. They show that a ten percent difference in the populations of two source languages makes for a 7.6% difference in the number of translations, in favor of the larger population. To make this more explicit, consider the English speaking population (400 million), which is 5.5 times larger than the one that speaks French (72 million). If the number of books translated from French to English is 100, the number of translations from English to French will be $5.5 \times 0.76 = 4$ times larger—that is, of the order of 400, all other things being equal. The coefficient picked up by the destination language is much smaller (0.35). This results from the fact that more books are written in languages that are spoken by many and populations that speak these languages are more self-sufficient. On the other hand, fewer titles are translated into languages spoken by smaller and more homogeneous populations, where word of mouth may work, or where the number of newspapers that carry literary critiques is smaller and less diversified. While the English speaking population is 5.5 times larger than the French speaking.

---

21 The equation is estimated over the whole period, and not year by year.
population, translations to English should only be $5.5 \times 0.35 = 2$ times larger than translations to French.

The coefficient picked up by distance is not significantly different from $-1$. This is a large effect, as it means that a 10% increase in the linguistic distance between two languages decreases the number of translations by the same percentage. Again, a simple example may be useful. The distance between German and Dutch is 0.162; it is equal to 0.812 between German and Greek, thus 5 times larger. For 100 books translated from German to Dutch (and vice versa), there would be merely 20 translations from German to Greek (and vice versa).

In the second equation, we add literacy rates and income per head in the country of destination.22 As expected, both are significantly positive, and the coefficient picked up by the literacy rate is quite important (3.65, which means that a 10% increase in literacy increases translation by 36.5%), and so is the one relative to income (0.45), while the other parameters remain similar to those of the previous equation.

Both equations explain over 40% of the total variance of (the logarithm of) the number of translations. The residual unexplained variance could be reduced in a significant way if one distinguishes the effect of the populations speaking the various source or destination languages.

This can be examined by rewriting Eq. (1) as:23

$$
\ln t_{ij} = \alpha_0 + \sum_{k=2}^{19} \delta_k \beta_k \ln P_k + \alpha_2 \ln P_j + \alpha_3 \ln D_{ij} + v_{ij},
$$

(2)

and

$$
\ln t_{ij} = \alpha_0 + \alpha_1 \ln P_i + \sum_{k=2}^{29} \delta_k \gamma_k \ln P_k + \alpha_3 \ln D_{ij} + v_{ij}.
$$

(3)

22 Literacy rates may not matter in Europe since they are almost equal to 100% in all countries. However, some of the languages (in particular English and Spanish) are spoken in Africa, Asia, Central and South America also, where literacy rates are not as large. The same argument holds for the income variable.

23 To simplify, we ignore the literacy and income variables.
In Eq. (2), the common source language variable is replaced by several source variables, one for each language (except one, here $k = 1$, normalized to 0); the $\delta_k$ are dummy variables that take the value 1 for $i = k$, and 0 otherwise. In Eq. (3) an identical procedure is adopted for destination languages (with dummy variables $\delta_k$ that take the value 1 for $j = k$, and 0 otherwise). Each dummy is multiplied by (the log of) the population that speaks the language, so that each source or destination language will pick up its own parameter, $\beta_k$ or $\gamma_k$ and give an indication of the “power” or other idiosyncrasies of the language that is not explained by the other variables present in the equation. The interesting result here is that it allows ranking source and destination languages.

Estimation results are shown in Table 3, for both equations. Note, first, that the effects of common parameters (destination language in (2) and source language in (3), as well as distance between languages) are qualitatively similar to those derived in Eq. (1).

The idiosyncratic language parameters can be ranked by decreasing order. This is done in the second part of the table, in the two columns headed “Index (German = 100).” As mentioned above, one of the parameters cannot be estimated, and all these index numbers are calculated with respect to German, which is arbitrarily given the index value of 100. But one can, of course, compute the index between any two languages. For instance, translations from Norwegian are $149/115 = 30\%$ larger than those from French, when account is taken for source and destination languages and for distances.

The top tree source languages are Scandinavian (Norwegian, Danish and Swedish), and another Scandinavian language, Icelandic, is the top destination language. The number of translations between Scandinavian languages is obviously important, although clearly, these are translations from small languages into small languages and the numbers cannot reach the numbers reached by large languages. It should also be noted that French is ahead of English as a source language, and in fact, French, English and German are in a close race. On the destination side, and among the languages with large number of speakers, Polish comes first—just behind Icelandic, with a very small number of speakers, but a large numbers of translations per head; German, Finnish, Hungarian and French are close “competitors”; and indeed, English is quite far down in the list—but ahead of Spanish, Italian and Portuguese, three other important languages.

This clarifies the result concerning the disproportionate number of translations from English. It hardly disproves the fact that little is translated into English, but many languages do worse than English. Though Russian, Spanish and Portuguese are spoken by large populations, their role as source languages is often dwarfed by languages spoken in smaller countries, such as Finnish, Czech and Hungarian. The model, however, offers no explanation for this effect, which may be due to the fact that some writers in smaller countries become very famous and are translated very extensively during certain periods (including the one under review) and then disappear after some time. This is unlikely to happen in larger countries with many authors, where the mix of more and less prolific authors is somewhat stable. Unfortunately, the data are aggregate (number of titles translated) and do not allow verifying whether the number of translations is due to many authors or whether it is the consequence of a small number of prolific authors translated because of their fame. It leaves open the question whether the differences are the consequence of popular or upscale authors. This ranking also includes public (or private, through publishers) efforts made by some countries (or linguistic communities) to make their literature known in the rest of the world.

---

24 For technical econometric reasons (normalization), one origin or one destination variable has to be discarded, and all the other estimated parameters are measured with respect to the omitted origin or destination.
as well as those made by some communities to offer to their citizens translated works. This is for example the case of France, where the cost of translating a novel is often subsidized.

Finally, the estimated parameters of the model and the various interpretations to which they lead are based on the world that we observe, and not on a possible ideal world in which there could or should be more equality. But they point to some reasons for which inequalities, that can hardly be changed, arise.

4. Conclusions

We construct and estimate a model that offers some insight into the determinants of literary translations. The formulation of the estimated equations can be derived on intuitive grounds or from the “demand for translated” books equation described in Appendix B. We show that the model fits the data quite well. Translations from English are not disproportionately large, and

<table>
<thead>
<tr>
<th>Source language</th>
<th>Coefficient in Eq. (2)</th>
<th>Index (German = 100)</th>
<th>Source language</th>
<th>Coefficient in Eq. (3)</th>
<th>Index (German = 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norwegian</td>
<td>1.3864</td>
<td>149</td>
<td>Icelandic</td>
<td>1.1684</td>
<td>181</td>
</tr>
<tr>
<td>Danish</td>
<td>1.3705</td>
<td>146</td>
<td>Polish</td>
<td>0.9975</td>
<td>153</td>
</tr>
<tr>
<td>Swedish</td>
<td>1.2828</td>
<td>134</td>
<td>German</td>
<td>0.5751</td>
<td>100</td>
</tr>
<tr>
<td>French</td>
<td>1.1253</td>
<td>115</td>
<td>Finnish</td>
<td>0.5106</td>
<td>94</td>
</tr>
<tr>
<td>English</td>
<td>1.0903</td>
<td>111</td>
<td>Hungarian</td>
<td>0.4560</td>
<td>89</td>
</tr>
<tr>
<td>German</td>
<td>0.9897</td>
<td>100</td>
<td>French</td>
<td>0.4436</td>
<td>88</td>
</tr>
<tr>
<td>Finnish</td>
<td>0.8510</td>
<td>87</td>
<td>Czech</td>
<td>0.2949</td>
<td>76</td>
</tr>
<tr>
<td>Italian</td>
<td>0.7892</td>
<td>82</td>
<td>English</td>
<td>0.2829</td>
<td>75</td>
</tr>
<tr>
<td>Czech</td>
<td>0.7723</td>
<td>80</td>
<td>Russian</td>
<td>0.2817</td>
<td>75</td>
</tr>
<tr>
<td>Hungarian</td>
<td>0.7566</td>
<td>79</td>
<td>Spanish</td>
<td>0.2749</td>
<td>74</td>
</tr>
<tr>
<td>Russian</td>
<td>0.7453</td>
<td>78</td>
<td>Dutch</td>
<td>0.2575</td>
<td>73</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.5774</td>
<td>66</td>
<td>Danish</td>
<td>0.2208</td>
<td>70</td>
</tr>
<tr>
<td>Polish</td>
<td>0.5477</td>
<td>64</td>
<td>Serbo-Croatian</td>
<td>0.1530</td>
<td>66</td>
</tr>
<tr>
<td>Dutch</td>
<td>0.5383</td>
<td>64</td>
<td>Bulgarian</td>
<td>0.1190</td>
<td>63</td>
</tr>
<tr>
<td>Portuguese</td>
<td>0.3238</td>
<td>51</td>
<td>Norwegian</td>
<td>0.0671</td>
<td>60</td>
</tr>
<tr>
<td>Romanian</td>
<td>0.2781</td>
<td>49</td>
<td>Romanian</td>
<td>0.0652</td>
<td>60</td>
</tr>
<tr>
<td>Bulgarian</td>
<td>0.2418</td>
<td>47</td>
<td>Slovak</td>
<td>0.0290</td>
<td>58</td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td>0.1661</td>
<td>44</td>
<td>Italian</td>
<td>0.0196</td>
<td>57</td>
</tr>
<tr>
<td>Slovak</td>
<td>0.000</td>
<td>37</td>
<td>Estonian</td>
<td>0.0000</td>
<td>56</td>
</tr>
<tr>
<td>Swedish</td>
<td></td>
<td></td>
<td>Portuguese</td>
<td>−0.0488</td>
<td>54</td>
</tr>
<tr>
<td>Portuguese</td>
<td></td>
<td></td>
<td>Lithuanian</td>
<td>−0.1454</td>
<td>49</td>
</tr>
<tr>
<td>Greek</td>
<td></td>
<td></td>
<td>Ukrainian</td>
<td>−0.3020</td>
<td>42</td>
</tr>
<tr>
<td>Slovene</td>
<td></td>
<td></td>
<td>Latvian</td>
<td>−1.7646</td>
<td>10</td>
</tr>
</tbody>
</table>

No. of observations 471
Adj. R-squared 0.734
they are not larger than those from French. It is obvious that the more titles are written in a language (and this number is roughly proportional to the population, at least in countries which benefit from a similar degree of literacy), the more will be translated into other languages, as long as cultural traits are sufficiently similar. If these are not, cultural distances will also play a role: the larger the distance, the smaller the number of translations.

We consider the period 1979–2002 as a whole. Heilbron (2008) underlines the importance that changing transnational relations have on translations, and he looks at the patterns of exchanges between France and the Netherlands between 1946 and 1995. He shows that the “trades” between the two countries underwent substantial changes. We did not pursue this issue here, since the time period is only half as long as the one considered by Heilbron and is unlikely to show very large changes. Moreover, as Heilbron (1999, 2008) points out, translations should be looked at in terms of shares in the total market for books in a country or language—which, for reasons given earlier, are not available, and not as a share of various languages in the total market for translations.

This being said, it is true that English has a privileged (rather than a dominating or hegemonic) position on the market for novels. A large number of books are translated from English, and little is translated into English, as the English literature is more heterogeneous and demand for translations is smaller. The sizes of the populations and linguistic distances, which are exogenous to the translation process, explain these disparities, and they cannot be removed. Our results are consistent with what Pym (1999, p. 9) had already in mind more than ten years ago, writing that “since literatures in English constitute a vast and diverse phenomenon, many of the non-translated [that is originally written in] English texts are in fact from non-central cultures, potentially fulfilling much of the diversifying role that translations play in less extensive languages.”

Public or private policies can of course marginally change this state of affairs by subsidizing translations. This is, to some extent, the case in continental Europe. One may wonder why Americans and the British, who are so keen in privately subsidizing music, theaters, or museums, do not subsidize translations. In her survey of American publishers, Sapiro (2010) notes that some publishers are embarrassed by this situation. One of those active in the up-market complains that “in our whole culture, we have been unable to assimilate and learn about other cultures” (Sapiro, 2010, p. 435). A similar view is taken by Esther Allen, a literature professor in a New York college and former director of the PEN Translation Fund: “There is still a very entrenched attitude on the part of mainstream commercial houses that the US consumer of books does not want to read translations.”

For the time being, authors who write in English thus benefit from three advantages. They have a large market in their own language, they also have access to other large markets since their books get translated, and finally, they do not “need” spending on translation costs. This is also reflected in the number of British (and probably American) high school students who learn foreign languages. A recent European survey shows that while some 98–99% of children in the European Union study at least one foreign language, the number is 81% in Ireland and only 48% in the United Kingdom. Some do indeed benefit from free lunches.

Acknowledgments

Susanne Janssen and Kees van Rees, the editors of Poetics, as well as several referees provided extremely useful comments that introduced us to the sociological approach of translations and

---

clarified a large number of issues that were merely sketched in previous versions of the paper. We are also grateful to Georg Kirchsteiger and Abdul Noury for comments on previous versions, as well as to Alain Brion and Mauro Rosi from UNESCO, who helped us to retrieve the data.

Appendix A. Languages included

<table>
<thead>
<tr>
<th>Language</th>
<th>Source</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulgarian</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Czech</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Danish</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Dutch</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>English</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Estonian</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Finnish</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>French</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>German</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Greek</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Icelandic</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Latvian</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Lithuanian</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Norwegian</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Polish</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Portuguese</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Romanian</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Russian</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Serbo-Croatian(^a)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Slovak</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Slovene</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Swedish</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Ukrainian</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) UNESCO deals with Serbian, Croatian and Serbo-Croatian separately. The distance matrix between languages includes Serbo-Croatian only. We therefore aggregated the three languages.

Appendix B. The theoretical model

One of the standard tools in economics is the demand function (for commodities) derived from the behavior of a “representative” consumer who makes his choices at prices and income that are given, by maximizing his utility under a budget constraint that limits his expenditure to given income. Here, the representative reader in a destination country chooses the number of translated titles from each language that she reads.\(^{27}\) This number results from the maximization of her utility under the constraint that she cannot exceed the time or effort devoted to reading, which is exogenously given. This behavior leads to a demand equation for titles in each destination country, endowed with the following properties. If the population that speaks the source language \(A\) is larger than that of another source language \(B\), then on average, and \textit{ceteris paribus}, the number of titles that will be translated from \(A\) will be larger than the one translated from \(B\), as

\(^{27}\) For simplicity, we assume that authors write in their native language and that readers read foreign book in translations.
there will be more titles written in A. Likewise, if the population in the destination language C is larger than in D, individuals in C will look for more variety, and they will demand more translations than those in D. Cultural distances play the role of prices in the usual model, for we assume that the larger the distance between the origin and the destination language, then the more it takes effort to read the novel. The impact of linguistic distances goes in the opposite direction: the larger the distance between the source and the destination language, the fewer titles will be translated. Literacy and income of the population of the destination language are expected to have a positive influence, for more books will be read (individuals spend more time reading) – and thus translated from other languages – in more literate and richer regions.

Consider a society where citizens speak languages from the set \( Q = \{1, 2, \ldots, q\} \). Suppose that every citizen knows one and only one native language in \( Q \). Denote by \( P_j \) the population of citizens whose native language is \( j \), and by \( L_j \) and \( W_j \) their literacy rate and average income, respectively.

 Citizens who know language \( j \) have access to books translated from other (foreign) languages \( i = 1, 2, j - 1, j + 1, \ldots, q \). Subscript \( i \) will denote a source language from which a book is translated into \( j \), the so-called destination language. We assume that for every language \( j \) there is a representative (average) reader \( A_j \) who speaks \( j \) and spends \( R_j \) hours reading translated books. Let \( t_{ij} \) be the number of titles (translated from language \( i \) to language \( j \)) that \( A_j \) reads. We assume that:

**Assumption 1.** The exogenously given time (effort) that this average reader can devote to reading is increasing with her literacy rate and income. \( R_j \) is an increasing function of the average literacy rate \( L_j \) and average income \( W_j \) of population \( P_j \).

It is often more difficult to adjust to novels entrenched in different cultures. Therefore, we assume that it takes longer or more effort to read books translated from languages that are culturally more distant. Denote the cultural distance between languages \( i \) and \( j \) by \( D_{ij} \). Then:

**Assumption 2.** It takes more time (or effort) for an individual to read titles that are translated from more distant source languages. There is a positive constant \( r \) such that \( A_j \) needs \( r(1 + D_{ij}) \) hours or units of effort to read a book translated from \( i \) to \( j \).

\( A_j \)'s reading time constraint can be formulated as:

\[
 r(1 + D_{1j})t_{1j} + \cdots + r(1 + D_{j-1,j})t_{j-1,j} + r(1 + D_{j+1,j})t_{j+1,j} \cdots + r(1 + D_{qj})t_{qj} = R_j.
\]

(A1)

**Assumption 3.** The utility that an individual derives from reading is increasing with the number of titles he reads, but titles translated from different languages may generate different units of utility.

This assumption can be interpreted in two ways: either readers know the kinds of translations they want to read, or publishers impose their mix of translations from source languages. In reality, there is probably a mutual influence as publishers “know” or “guess” the tastes of their readers and choose the right combination of translations from the various source languages. \( A_j \)'s preferences for foreign books translated from languages different from \( j \) are represented by the following Cobb-Douglas utility function with \( q - 1 \) variables:

\[
 u_j(t_{1j}, t_{2j}, \ldots, t_{j-1,j}, t_{j+1,j}, \ldots, t_{qj}) = t_{1j}^{\gamma_1} \times t_{2j}^{\gamma_2} \times \cdots \times t_{j-1,j}^{\gamma_{j-1}} \times t_{j+1,j}^{\gamma_{j+1}} \times \cdots \times t_{qj}^{\gamma_q}
\]

(A2)

**Assumption 4.** These units of utility increase with the population that speaks the destination language, as readers look for more diversity (this is captured by the next assumption). For every language \( i \), \( \gamma_i = \gamma(P_i) \) is an increasing function of the population \( P_i \).
Assumption 3 offers a simple functional form (more complicated forms would lead to very similar results) for the utility derived from translated books. The representative reader has access to translations from all languages, including those from distant cultures giving her the possibility to learn about all possible cultures. Assumption 4 is meant to represent the influence of the source language $i$. If the number of writers per head in population $P_i$ is equal across languages, then the number of books written in a language is an increasing function of the number of its speakers.

For every pair of languages, $i$ and $j$, the average reader $A_j$ chooses the optimal number $t_{ij}'$ of foreign titles in $i$ that she will read by maximizing her utility $u_j(.)$ under her reading time constraint (A1). This leads to the following demand functions:

$$t_{ij}' = \frac{R_j}{r(1 + D_{ij})} \Gamma_i$$

where $\Gamma_i = \gamma_i / \sum_{k \neq j} \gamma_k$. Note that in this formulation $t_{ij}'$ is decreasing in the distance between languages. One could argue that when languages are very close, there is no need for translation as both populations can read each other’s books in the native language. The number of translated titles would then increase with the linguistic distance up to a point, and then it would eventually decline when the linguistic gap between two languages becomes large. It is easy to develop a model that would reproduce such an inverted-U shape for the relation between number of titles translated and distance. However, we could find no evidence for this in the empirical results and did not pursue the idea. The reason is probably due to the fact that the distance for which the inverted-U curve peaks is quite low. The closest languages in our database are Slovak and Czech, and there are nevertheless books translated from one to the other language.

The link between the total number of titles translated from $i$ to $j$ and the preferences of the average reader is determined by geographical and cultural diversity within population $P_j$. Indeed, if the population is perfectly homogeneous, all readers will read the same titles, and the total number of titles translated from $i$ to $j$ will be equal to $t_{ij}'$. This will happen if population $P_j$ is small with a small number of literary critics who recommend the same books. Word of mouth between the few readers who do not live far away from each other will go in the same direction. In the other extreme case of a completely heterogeneous (and large) population $P_j$, in which each reader lives on an “island” and reads different titles, the total number of titles translated from $i$ to $j$ will be equal to $P_j t_{ij}'$. In other words, when $P_j$ is small, the number of titles read there will be small. When $P_j$ is large and diverse (as in the case of English or Spanish whose speakers are scattered across countries and continents), there will be less information flowing between sub-regions, local populations will be more isolated from each other, the number of newspapers carrying literary criticisms will be larger, and the number of translated titles will be relatively large. Therefore, it is reasonable to represent the size and diversity of population $P_j$ by a heterogeneity index $H_j$, an increasing function of the population size $P_j$. Therefore, we assume:

**Assumption 5.** The link between the total number of titles translated from a source language and the preferences of the average reader is determined by the diversity within the population of the destination language (country).
If this population is small, it will be more homogeneous; it will have access to critics who are more likely to share the same tastes; and the number of titles read will be relatively small.\textsuperscript{29} A larger population in a large country (or even in different countries that share the same native language) will be less homogeneous, and the number of titles translated into its native language is likely to be larger. The total number of titles translated from $i$ to $j$ is therefore given by $t_{ij} = H_{ij}^{*}$.

Demand functions for all languages $i \neq j$ can now be fully specified as

$$t_{ij} = H_j(P_j) \frac{R_j(L_j, W_j)}{r(1 + D_{ij})} \Gamma_i$$

where $\Gamma_i = \gamma_i(P_i) / \sum_k \delta_{k \neq j} \gamma_k(P_k)$. It is easy to show that these demands satisfy the following properties.

\textit{Proposition}. Under Assumptions 1–5, the number of titles translated from $i$ to $j$ is

(a) increasing in $P_j$, the population whose native language is $j$,
(b) increasing in $P_i$, the population whose native language is $i$,
(c) decreasing in $D_{ij}$, the linguistic (or cultural) distance between languages $i$ and $j$,
(d) increasing in $L_j$, the literacy level of the population $P_j$,
(e) increasing in $W_j$, the income level of the population $P_j$.

\textbf{Appendix C. Additional data}

\textit{Literacy rates in destination languages.} See UNESCO, Institute for Statistics (2002). We chose rates given for 1990 and computed population weighted rates for Portuguese (Brazil and Portugal) and Spanish (Argentina, Bolivia, Chile, Colombia, Dominican Republic, Ecuador, Guatemala, Honduras, Mexico, Peru, Spain, and Venezuela). For other languages, we took the rate of the country in which the language is native. UNESCO does not provide literacy rates for all Western European or North American countries. We assumed that these were equal to 100%. Though we deal with European languages, some are also spoken in other continents. Therefore literacy rates are not always as high as one would think. The world average rate for Spanish and Portuguese are 89% and 73%.

\textit{GNI per capita 2004 in destination languages.} See World Bank (2005). We used Purchasing Power Parity per capita GNI per head (international dollars) in 2004. No such data were available for 1990. We assumed that relative ranking did not drift too much apart between 1990 and 2005. Population weighted GNIs are computed for Portuguese (Brazil and Portugal), Spanish (Argentina, Bolivia, Chile, Colombia, Dominican Republic, Ecuador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Salvador, Spain, Uruguay, and Venezuela), German (Germany and Austria), and English (Australia, Canada, New Zealand, United Kingdom, and United States). For other languages, we merely use the GNI of the country in which the language is native.

\textsuperscript{29} In France, for example, some 6000 books are published every year. Ginsburgh and Weyers (2008) analyze literary reviews in two important non-specialized but nationally distributed newspapers (\textit{Le Figaro} and \textit{Le Monde}) and two important specialized magazines (\textit{Le Magazine Littéraire} and \textit{La Quinzaine Littéraire}). They find that only 800 books are reviewed, that 294 out of these 800 titles are reviewed in at least two outlets and 157 are reviewed in all four outlets. They conclude that there is too little diversity in reviews. It would of course be useful to have comparable numbers for other countries to substantiate this claim.
References


Victor Ginsburgh is honorary professor of economics. He carries on with his research at ECARES, Brussels and CORE, Louvain-la-Neuve. He wrote and edited a dozen of books and is the author of over 180 papers in applied and theoretical economics, including industrial organization and general equilibrium analysis. His more recent work includes economics of the arts and economics of languages. He has published over 60 papers on these topics. He is coeditor (with D. Throsby) of the first volume of a Handbook of the Economics of Art and Culture (Elsevier, Amsterdam, 2006), and is preparing the second volume. His recent book How Many Languages Do We Need? The Economics of Linguistic Diversity (with S. Weber) has been recently published by the Princeton University Press.

Shlomo Weber is the Dedman Trustee Professor of Economics at Southern Methodist University, Dallas, and Professor of Economics at the New Economic School, Moscow. He holds the M.Sc. in mathematics from the Moscow State University and the Ph.D. from the Hebrew University of Jerusalem. Professor Weber held positions in various universities in North America, Europe and Asia. He was the holder of the Alexander von Humboldt Research Award for Outstanding Foreign Scientists. His main research is in political economics and game theory and has published more than 100 articles in leading economics and political science journals. His recent book How Many Languages Do We Need? The Economics of Linguistic Diversity (with V. Ginsburgh) has been recently published by the Princeton University Press.

Sheila Weyers has a degree in philosophy, and is interested in aesthetics and its relations with art history. She has published on movies, art history, canons of the Flemish and Italian Renaissance painters, and awards and prizes. Her papers appeared in Artibus et Historiae, Journal of Cultural Economics, Empirical Studies of the Arts and Poetics.