Working Paper

CEPII

One Nation, One Language? Domestic Language Diversity, Trade and Welfare

Tamara Gurevitch, Peter R. Herman, Farid Toubal & Yoto Yotov

Highlights

- We examine how language diversity affects trade and economic welfare.

- We use the new version of the Ethnologue.com dataset and build linguistic ties between populations in and across nations.

- In our quantitative general equilirium model, a change in domestic linguistic similarity simultaneously affects both internal as well as international trade frictions. It has direct effects on trade between populations domestically and globally and affects economic welfare.

- We illustrate the effect of changing domestic linguistic diversity simulating the repeal of Quebec's Bill 101, which made French an official language in Canada.

- We show large impacts of domestic linguistic policies on domestic welfare, but not only. Linguistic policies have significant economic consequences that stretch far beyond the country's own borders because of trade linkages.

RESEARCH AND EXPERTISE
ON THE WORLD ECONOMY

▊ Abstract

Using new data on linguistic diversity across and within countries, we examine novel channels though which language affects trade patterns and economic welfare. We find that linguistic similarity within a country accounts for about 10 percent of estimated `home bias', demonstrating the importance of shared languages for domestic integration. To highlight the general equilibrium implications of domestic language proximity, we simulate the repeal of Quebec's Bill 101, which made French an official language in Canada and established fundamental language rights for French-speakers. The analysis demonstrates that domestic language diversity has significant implications for Canada's welfare but also sizable economic consequences that stretch far beyond its borders.

▊ Keywords

Common Language, Ethno-linguistic Diversity, Identity, International Trade, Domestic Trade, Domestic Trade Costs, Welfare.

▊ JEL

D60, F14, F19, C54, Z13.

Working Paper ▪

CEPII

RESEARCH AND EXPERTISE
ON THE WORLD ECONOMY

One Nation, One Language? Language Diversity, Trade and Welfare[1]

Tamara Gurevitch [*] Peter R. Herman [†] Farid Toubal[‡] and Yoto Yotov [§]

1. Introduction

Since the classic work of Marschak (1965), there has been a growing consensus within economics that language is an important determinant of economic activity.[2] Much of this influence is due to the fact that language is integral to much of our lives. Not only is it a critical means of communication, it is also a medium through which values, beliefs, and customs are transmitted to others. In many countries, language acts as an important marker of identity within diverse populations (Alesina and Reich, 2015).[3] The influences of language are so deep that linguistic homogeneity is crucial for promoting integration between groups and improving economic welfare.

How does language diversity affect trade patterns and economic welfare? Within the economics literature, there are many competing views.[4] Some argue that standardization or the widespread

[*]U.S. International Trade Commission, Office of Economics: Tamara.Gurevich@usitc.gov.
[†]U.S. International Trade Commission, Office of Economics: Peter.Herman@usitc.gov.
[‡]University of Paris-Dauphine – PSL, LEDa, CEPII, CESifo and CEPR: farid.toubal@dauphine.psl.eu.
[§]Drexel University and CESIfo: yotov@drexel.edu

[2]See the insightful review of the literature by Ginsburgh and Weber (2020) and the compendium of papers compiled by Gazzola and Wickström (2016) and Ginsburgh et al. (2016), which describe many studies on the impacts of language on varied economic outcomes and public policies.

[3]To illustrate, a recent cross-national survey conducted in 14 countries by the Pew Research Center shows that language matters more to national identity than birthplace (Stokes, 2017). Similarly, the vast related academic literature stresses the importance of languages on the construction of nations (Anderson, 1983; Gellner and Gellner, 1983; Hobsbawm et al., 1992).

[4]The costs and benefits of linguistic diversity and the need for sharing a common language is thoroughly described

knowledge of a common language increases efficiency as it allows for the discovery of preferences and lowers transaction costs associated with communication. It therefore helps establish trade between populations and improves welfare (Melitz and Toubal, 2014; Egger and Toubal, 2016). Linguistic diversity might also lower redistribution (Desmet et al., 2009), hinder growth (Alesina et al., 2003), limit the provision of public goods (Desmet et al., 2012), or hamper a country's level of development (Desmet et al., 2016). Others claim that standardizing languages may result in the disenfranchising of linguistic groups within a country, restricting their linguistic rights and potentially causing social and economic damages (Ginsburgh and Weber, 2005, 2011). Domestic policies governing language throughout the world ought to consider these competing issues because they can have significant economic consequences domestically as well as far beyond their own borders. However, it is difficult to empirically quantify these effects as doing so requires detailed data on linguistic ties between populations and a quantitative framework that takes into account linguistic diversity between populations throughout the world.

Our contribution to this literature is to quantify the partial and general equilibrium (GE) effects of language diversity on trade and welfare by using a new comprehensive language dataset and a contemporary trade framework. A key novelty of our analysis is the focus on linguistic similarity *within* countries. A change in domestic linguistic similarity simultaneously affects both internal as well as international trade frictions, with global implications for trade and welfare. Admittedly, our specific focus of the impact of language through trade does not capture a number of potentially important channels through which language similarity may impact welfare. However, the benefits of our analysis are that (i) it contributes to an extensive literature that studies the links between language and trade, and (ii) it enables us to quantify both the partial and the GE effects of *domestic* and international language proximity on trade and welfare within the same theoretical framework.[5]

_____

by Ginsburgh and Weber (2011).

[5]Consistent with the voluminous trade literature we use the term 'welfare' to denote 'real expenditure'. We refer the reader to Melitz (2012), who defines welfare more broadly and proposes a general theoretical framework to analyze

While our model is sufficiently general to quantify the effects from many potential policy scenarios, we chose to simulate the impact of changing domestic linguistic diversity on trade and welfare using one of the most important linguistic shifts in Canada's history. Language has long been a sensitive and contentious issue in Canada. So much so that by the end of the 1960s, language had become the main issue dominating Quebec's economic, social, and political life. The linguistic shift began with what has come to be known as the *Révolution Tranquille* (Quiet Revolution), which led to the implementation in 1977 of the *Charter of the French Language*, commonly known as Bill 101. French became the only official language of Quebec and the main language of the workplace, education, and several other areas of public life in the province.[6] Motivated by the interesting history and important cultural and economic implications of Quebec's Révolution Tranquille, we quantify the effects of language policies on trade and welfare by examining it as a case in point. In particular, we estimate the economic impacts of a counterfactual experiment in which Quebec were to become fully bilingual in English—the other prominent language in Canada—as well as French.

Our analysis uses novel data on shared languages between populations within and across countries from the widely used Ethnologue database of languages maintained by the Summer Institute of Linguistics. These data allow us to construct a new common language index (CL) that reflects multiple different aspects of languages such as translation and interpretation, communication, and linguistic proximity between ethno-linguistic groups. We follow the procedure described in Melitz and Toubal (2014) to construct an aggregate index of common languages that measures linguistic similarities between the populations of different countries. Additionally, our new index also measures linguistic similarities between populations *within* each country and thus reflects the extent of each nation's ethno-linguistic fractionalization. Our CL index is therefore composed of an international

---

the many different influences that language may have on welfare.

[6]François Vaillancourt, who worked on the drafting of Bill 101 as a former economic adviser to Parti Québécois cultural development minister Camille Laurin, proposes an interesting special issue on the supply and demand of English and French in Quebec over the 40 years since Bill 101 (See, Vaillancourt, 2019).

(ICL) component that describes the common languages between pairs of countries and a novel domestic (DCL) component that characterizes the degree of linguistic similarity within countries. In constructing the indices, we make use of the entire set of information on ethno-linguistic groups, representing 6,534 languages around the world (Simons and Fennig, 2018). We combine the CL index with detailed information on international trade between countries as well as domestic (intra-national) trade within countries.

We propose a quantitative trade model to examine the GE effects of language on trade and welfare. The analysis is performed using a representative 'new' quantitative trade model (c.f., Arkolakis et al. (2012)). Anderson et al. (2018) identify internal distance and economic development as key determinants of domestic trade costs. We view the degree of language similarity as an additional key factor that shapes internal trade frictions. Further, a change in language diversity within a country also affects its linguistic connections to each and every foreign trade partner. We embed these insights in a framework similar to that of Ramondo et al. (2016), which assumes that internal trade frictions are not uniform across countries.[7] Changes in domestic linguistic diversity affect both internal and international trade costs and may divert trade from domestic to international sales and across trading partners. The GE impacts are a redistribution of welfare across countries.

Following Melitz and Toubal (2014), which is the predecessor of our ICL index, we begin the empirical analysis by showing that the new ICL measure outperforms the widely used common official language indicator throughout the literature that investigates the impact of common language on international trade. Our analysis confirms that the effects of international language proximity on bilateral trade are large, positive, and statistically significant, implying that greater linguistic similarities between countries increase international trade.

More importantly, we extend the evaluation of language and international trade to domestic trade

---

[7]Agnosteva et al. (2019) show large and widely heterogeneous domestic trade costs across Canada's provinces.

and provide novel estimates on the impacts of diversity between different ethno-linguistic groups within the same country. We find that the effects of DCL are large, positive, and statistically significant. In the context of Canada, our estimates imply that a fully bilingual Quebec would double Canada's domestic trade. Thus, our quantitative results uncover the potential for very large gains from domestic integration.

In investigating the impact of language on domestic trade, we contribute to the aforementioned literature that has explored the effect of linguistic diversity on other important domestic outcomes.[8] We also contribute to a small but growing literature that underscores the importance of proper treatment of domestic trade costs.[9] Specifically, we find that once we control explicitly for shared languages within a country, our estimate of the 'home bias' impact on trade decreases by about 10 percent.[10] In the context of the literature, we extend the typical treatment of home bias using dummy variables by isolating and identifying the impact of language as a proxy for shared identity within a nation.

The quantitative analysis considers a counterfactual experiment in which Quebec becomes fully bilingual, speaking both French and English. In order to highlight the relative importance of the competing forces that act in response to the changes in Canada's domestic language composition, we quantify the GE impact of the changes in the CL index in two steps. First, we modify only the DCL component of Canada's CL index, reflecting a case in which Canada's domestic trade costs are lowered but its international trade costs remain unchanged. Canada's total exports to foreign

---

[8]See Alesina and Ferrara (2005) and Ginsburgh and Weber (2020) for surveys of the literature.

[9]See for example: Ramondo et al. (2016), who demonstrate that the standard findings (i) that larger countries should be richer than smaller countries and (ii) that real income per capita increases too steeply with country size, disappear when domestic trade costs are taken into account; Donaldson (2018), who studies the implications of domestic trade costs in the form of railroad network in India for productivity and welfare; Coşar and Demir (2016) and Coşar and Fajgelbaum (2016), who consider the impact improvements in transportation infrastructure and internal geography when trade must pass through gateway locations

[10]These estimates relate our work to the extensive literature on the home bias in international trade. See, for example, Head and Mayer (2013), Wolf (2000), and Millimet and Osang (2007).

countries would decrease by 2.61 percent in response to the change in its DCL component. This is due to diversion away from foreign trade and toward domestic sales. This step provides insight into the domestic effects of the change without any international spillovers. Second, we modify the ICL component to reflect changes in language ties between Canada and the rest of the world. This step estimates the full global impact of the change such as reduced trade costs with English-speaking partners. Canada's international trade costs with English-speaking countries would fall and Canada would trade more with them, mitigating the negative effects of the DCL on Canada's exports to English-speaking countries. However, the trade creation effect of the decrease in Canada's international trade costs would not be strong enough to overcome the trade diversion effect from the change in Canada's domestic trade costs in non-English speaking countries. The combined DCL and ICL impact of making Quebec bilingual would be a 1.52 percent reduction in exports, which is still negative but significantly smaller than if only DCL was affected.

Turning to the welfare implications of our experiment, we find that the reduction in Canada's domestic trade costs would lead to a 0.94 percent increase in Canada's real GDP. We find heterogeneous effects across countries but the impact on the rest of the world would be negative for most. The United States would suffer the largest losses of about a 0.02 percent decrease in real GDP. While this percentage change may seem small, it represents a more than 30 billion dollar reduction in real GDP. When international trade costs are considered as well, the gain in GDP for Canada increases to 1.22 percent, the losses for English-speaking countries shrink, and the losses for most other countries grow.

Our work has some potential caveats. First, our CL measures are imperfect representations of language and identity. Because the measures are constructed to represent multiple aspects of language, the estimated impacts reflect many different influences.[11] For this reason, our experiment

---

[11] The old aphorism *"qui perd sa langue, perd sa foi"* expresses the many influences of languages beyond its impacts on communication and trade costs.

may overestimate the impacts of bilingualism. Second, the analysis may not properly quantify many important aspects of language and identity that do not manifest themselves in trade statistics. For example, the recent discussion in Canada around the discriminatory effects of required bilingualism in federal positions is one such example of an important impact that is not meaningfully reflected in our analysis.[12]

The remainder of the paper is organized as follows. Section 2 presents a brief overview of the impact of language diversity on domestic outcomes. Section 3 describes the data sources and the methods that we employed to construct the novel common language indices. Section 4 reviews the structural model with a focus on the partial and the GE effects of language. Section 5 obtains and analyzes the direct, partial equilibrium (PE) effects of language on domestic and foreign trade. Section 6 translates the PE estimates into GE effects on trade and welfare and presents and some history regarding the Quiet Revolution in Quebec. Finally, Section 7 summarizes our findings and points to further extensions. Supplementary discussions, results, and analysis are included in the Appendix.

2. Linguistic diversity and economic outcomes

In recent years, a vast literature has argued that linguistic diversity within countries affects a wide range of economic outcomes. Most studies show that linguistic diversity leads to communication barriers and divisions in national unity. In their famous paper, Alesina and Reich (2015) provide an example of how the 19th century European elites took steps to homogenize their population in order to create a national identity. Linguistic standardization and rationalization can lower the cost of communication, improve education, and increase the potential for economic growth. However, as shown by Ginsburgh and Weber (2011), standardization may meet considerable resistance as it inevitably restricts the usage of some languages and alienates some individuals, often minority

---

[12]For recent discussions, see Keung (2020) and McCullough (2020). However, the argument has circulated for many decades (Carr, 1985).

groups whose cultural, societal, and historical values are not represented by the official languages.[13] Consequently, it creates linguistic disenfranchisement and might exacerbate civil unrest.

There is a rich literature tying linguistic diversity to economic outcomes such as growth, development, and trade. Comparing growth rates in Africa and East Asia in 1969–1990, Easterly and Levine (1997) find that about 40 percent of the 3.5 percent annual growth differential between the two regions can be attributed to the effects of linguisitic diversity. Alesina et al. (2003) find that linguistic diversity undermines growth by generating political instability or reducing the quality of government. Alesina and Ferrara (2005) find nuanced effects of linguistic diversity on growth in a democratic society. Desmet et al. (2012) and Desmet et al. (2016) find that fine and coarse linguistic differences between populations have different impacts on economic outcomes. Desmet et al. (2012) find that fine linguistic differences hinder growth and the provision of public goods while coarse linguistic differences affect civil conflict and redistribution. In addition, Desmet et al. (2016) show that fine linguistic differences are sufficient to generate adverse effects on the level of development measured by per capita GDP.

In their paper on the impact of language and trade, Melitz and Toubal (2014) show that linguistic diversity leads to important communication barriers across countries. Aside from direct and indirect communication (through interpreters and translation), they show that languages have other influences on international trade such as ethnic ties and trust. Egger and Lassmann (2015) study a multilingual sample of native speakers in German, French, and Italian in Switzerland and are therefore able to distinguish between trade among partners possessing the same native language and trade among partners possessing a common language that differs from their native one. They find that native language has a semi-elasticity of influence on bilateral trade of around 0.3. Given that language can be inherited or learned, Egger and Toubal (2016) examine the impact of common

---

[13]The identification of groups in a nation is often based on languages. Fearon (2003), Laitin (2000), and Alesina and Ferrara (2005) suggest that there are other important dimensions of a group identity such as ethnicity or religion.

spoken native and acquired languages on various margins of international trade. They also quantify the GE impact of a five-percentage-point increase in common spoken native and acquired language. They do not explicitly model domestic trade and show that the increase in common spoken native and acquired language has a positive impact on welfare. The meta-analysis conducted by Egger and Lassmann (2012) further confirms the positive impact of language similarity on international trade. Based on 701 language effects on international trade collected from 81 academic articles, they show that that an increase of 10 percent in linguistic similarity raises international trade by almost 5 percent.

3. Language connections within and across countries

This section describes the sources and methods that we use to construct our extended index for international common language and the novel index for domestic common language. The common language indexes are built using a single and unique source of data, which is the 21st edition of Ethnologue: Languages of the World (Simons and Fennig, 2018).[14] This is the most comprehensive linguistic database available and describes in detail nearly 7,100 languages in 242 countries and territories. We make use of 6,534 languages for which Ethnologue reports the number of native speakers.[15] As seen from Table 1, there is a large variation in the number of languages spoken in each country. Some countries only have one language reported in the dataset, while others have several hundred languages listed. On average, there are 47 languages spoken in a country, although the median number is about third that (17). Papua New Guinea has the largest number of native and total spoken languages. Our data covers a large number of speakers most of them being native speakers. Mandarin Chinese has the largest number of native speakers, while English is most widely

---

[14] The 23rd edition of the data became available in February, 2020. For more information about the data see https://www.ethnologue.com/

[15] As mentioned in Lewis (2005), the distinction between languages and dialects is very difficult to draw. To mitigate the effect of possible ambiguities, our common language indexes incorporate information on cladistic distances between languages.

spoken by all speakers.

As in Melitz and Toubal (2014), our index of common language (CL_{ij}) between countries i and j reflects multiple different aspects of languages such as translation and interpretation, communication, and linguistic proximity between ethno-linguistic groups. We gather three types of information: the official language(s) of each country; the native language(s) spoken in each country; and linguistic trees, which are used to construct cladistic distances between languages and measure linguistic proximity. From this data, we construct our CL_{ij} index as an average of three subindices, each based on one of these three categories of information. We extend the approach of Melitz and Toubal (2014) in two directions: (i) by vastly increasing the set of languages used to construct our indices, and (ii) by measuring domestic common language. Thus, our index of common language has two components: a traditional international component labeled *international common language* (ICL_{ij}), which informs on common languages between pairs of countries, and a novel domestic component labeled *domestic common language* (DCL_{ii}), which characterizes the degree of linguistic similarity within a country.

For the official languages subindex, we follow the Ethnologue definition of official language, which specifies the function for which each language is recognized in a country. We define common official language such that if a language is jointly recognized for any official function—whether national or provincial, *statutory* or *de facto*—in a pair of countries, the two countries share a common official language. We build a common official language (COL_{ij}) indicator variable for each pair of countries i and j that takes value of one if both countries share at least one official language. Within a country, this indicator (COL_{ii}) takes a value of one if the country has at least one official language.

For the primary *native* languages subindex, we base our measure on the products of the percentages of speakers in each country pair. The product represents the probability that two people chosen at random from each pair of countries share a common native language k. The common native

language (CNL_{ij}) subindex is computed as $CNL_{ij} = \sum_{k \in K}(l_{ki} \times l_{kj}) \; \forall i, j \in \Omega$. Ω denotes the set of countries in our dataset, l_k is the percentage of speakers of a specific language k in country i or j, and K is the set of spoken native languages. We make the full use of the 6,534 languages available in the dataset to compute the subindex. The common native language between populations of the same country is given by $CNL_{ii} = \sum_{k \in K} l_{ki}^2$, which is the opposite of the commonly employed ethno-linguistic fractionalization (ELF) index (See, Alesina et al., 2003).

Finally, for the linguistic proximity (LP_{ij}) subindex, we turn to linguistic trees. This subindex was inspired by ideas from Laitin (2000) and Fearon (2003) who thought to base calculations of linguistic proximities on the Ethnologue classification of languages into trees, branches, and sub-branches. Within the classification, languages are broadly grouped into language families with common characteristics based on root proto-languages.[16] However, as individual languages evolved overtime and diverged from there ancestors, they formed new branches in the tree with their closest relatives and split from their more distant relatives. These linguistic trees provide a means for evaluating the similarities between two languages based on the proximity of their respective branches.

We develop an algorithm that makes use of all 6,534 languages in our dataset. It measures linguistic distance between languages by how far two languages are from their closest common proto-language. Using the linguistic family classification provided by Ethnologue, we compare the proportion of the linguistic tree that each pair of languages shares to the proportion in which they diverge. We construct a subindex that takes a value of zero if two languages do not originate from a common proto-language and share no part of a language tree. For languages that stem from the same proto-language, a value between zero and one is assigned. Languages that diverge early in their family tree receive lower proximity score than languages that share many common branches prior to splitting.[17]

---

[16] For example, Proto-Indo-European or Proto-Turkic.

[17] Linguistic proximity between languages k and k' is given by the number of common branches the two languages share starting at the proto-language divided by the average length of the branches that terminate in each language. For example, if two languages split after the seventh branch and one of those languages is 10 steps removed from

To account for the fact that most countries have multiple languages, we use the shares of native speakers of languages in a country to aggregate the linguistic proximity between languages in a country or a pair of countries. The full index CL_{ij}, which we use for our empirical analysis, is the simple average of the subindices COL_{ij} CNL_{ij} and LP_{ij} $\forall i, j \in \Omega$.

The ICL and DCL indices depict a complex web of linguistic connections between populations around the world. To illustrate, figure 1 plots a network of the ICL indices for a sample of 67 major trading countries.[18] The links connecting each country reflect the value of their ICL such that thicker and darker links reflect stronger linguistic ties. Similarly, countries are positioned in the figure based on linguistic ties, resulting in clusters based on certain languages and geographic regions. Notably, there are clusters of Spanish, English, and Arabic speaking countries with large indices and strong linguistic ties within each group. Similarly, many countries in central and eastern Europe as well as east Asia are grouped together with moderately high indices. By comparison, countries with less commonly spoken languages, such as Japan, Thailand, and Indonesia, have few strong ties and reside on the periphery of the network.

The DCL indices depict a similarly complex picture of the world. Figure 2 presents a map of the DCL indices for each the same 67 countries and highlights the significant variation in domestic language diversity. Some countries, such as Kenya, Cameroon, and India, have high language diversity, resulting in relatively low DCL indices. Meanwhile, countries like Iceland, South Korea, and Colombia have relatively low language diversity and high DCL indices. Canada, which has a fairly low index value (0.62), is an informative example of domestic linguistic diversity. To illustrate part of its diversity, Figure 3 depicts the share of French speakers in each of Canada's 13 provinces

---

its proto-language while the other language is 11 steps removed, then the proximity for these two languages is $Prox_{kk'} = \frac{pl_{kk'}}{0.5(pl_k + pl'_k)} = \frac{7}{10.5}$.

[18]While our indices cover 242 countries, we present those for only 67 here for simplicity and visual clarity. These countries, which are listed in table 2, are those included in our empirical analysis.

and territories in 2016.[19] The figure demonstrates the significant linguistic differences between Quebec (77 percent French speaking), New Brunswick (31 percent), and the rest of Canada (less than 4 percent).[20] If differences in language act as an impediment to trade, as has been found in the extensive existing literature, then efforts to reduce diversity within countries like Canada ought to be a strong means of promoting greater economic activity and welfare.

4. Quantifying the Role of Language

To perform the empirical analysis, we rely on the simplest and most widely accepted version of the new generation of quantitative trade models.[21] We use the following representative structure as a tool to describe and quantify novel channels that link domestic language diversity and—more broadly—changes in domestic policy and trade costs to international trade and welfare:

$$X_{ij,t} = \frac{Y_{i,t}E_{j,t}}{Y_t}\left(\frac{\tau_{ij,t}}{P_{j,t}\Pi_{i,t}}\right)^{-\theta}, \tag{1}$$

$$\Pi_{i,t}^{-\theta} = \sum_j \left(\frac{\tau_{ij,t}}{P_{j,t}}\right)^{-\theta}\frac{E_{j,t}}{Y_t}, \tag{2}$$

$$P_{j,t}^{-\theta} = \sum_i \left(\frac{\tau_{ij,t}}{\Pi_{i,t}}\right)^{-\theta}\frac{Y_{i,t}}{Y_t}, \tag{3}$$

$$p_{j,t} = \frac{(Y_{j,t}/Y_t)^{\frac{1}{-\theta}}}{\gamma_j\Pi_{j,t}}. \tag{4}$$

Arkolakis et al. (2012) famously demonstrate that the system (1)–(4) can be derived from a wide class of theoretical micro-foundations. Here, $X_{ij,t}$ denotes trade flows from exporter i to destination j at time t; $E_{j,t}$ is the total expenditure in importer j; $Y_{i,t}$ is the value of total production in exporter

---

[19]The underlying data, which reflects French as a "mother tongue", is from Statistics Canada/Statistique Canada, 2016 Data Tables (accessed October 14, 2020).
[20]Notably, this illustration under-represents the true extent of linguistic diversity in Canada, which also has extensive populations of indigenous and immigrant language speakers in addition to the large English and French speaking populations.
[21]Costinot and Rodríguez-Clare (2014), Head and Mayer (2014), and Yotov et al. (2016) offer recent surveys of the related literature.

i; Y_t is the value of world output; $\tau_{ij,t}$ denotes bilateral trade frictions between partners i and j; and θ is the trade elasticity. Equation (1) intuitively links bilateral exports to market size (the first term on the right-hand side, $\frac{Y_{i,t} E_{j,t}}{Y_t}$) and trade frictions (the second term on the right-hand side, $\left(\frac{\tau_{ij,t}}{P_{j,t} \Pi_{i,t}}\right)^{-\theta}$). The numerator of the trade cost term includes direct bilateral trade frictions ($\tau_{ij,t}$). Importantly, unlike many other studies, we will not treat internal trade frictions as uniform and equal to zero. Our approach to allow for heterogeneous domestic trade costs is consistent with theory, (c.f., Yotov et al. (2016)), and is crucial to our analysis for two reasons. First, Equation (1) will enable us to simultaneously identify the heterogeneous effects of international and domestic common language on trade. We elaborate on this point in Section 4.2. Second, changes in domestic common language (i.e., domestic trade costs) will also, by definition, affect international common language (i.e., international trade costs). Both of these changes in direct trade costs will lead to further GE effects, which are captured in (1)–(4). We discuss these GE effects in detail in Section 4.1.

The two remaining variables in Equation (1), $P_{j,t}$ and $\Pi_{i,t}$, are the inward and outward multilateral resistance (MR) terms, respectively, as coined by Anderson and van Wincoop (2003). As defined in equations (2)–(3), the MRs are consistent aggregates of bilateral trade costs for each country and can be interpreted as buyers' and sellers' incidence of trade costs, respectively (Anderson and Yotov, 2010).[22] As will be demonstrated in the next section, the MRs are the key vehicle that enables us to translate the direct changes in domestic and international common language into GE effects on trade and welfare throughout the world.

Finally, equation (4), which is a restatement of the market-clearing condition ($Y_{i,t} = \sum_j X_{ij,t}$), captures the link between trade and national income (via the outward MRs).[23] Specifically, (4)

---

[22]To add intuition, we note that the inward multilateral resistance, $P_{j,t}$, has a dual theoretical interpretation as an ideal consumer price index, while, equation (4) reveals an inverse mapping between the outward multilateral resistance, $\Pi_{i,t}$, and the producer factory gate prices.

[23]Depending on the underlying structural foundation of system (1)–(4), the parameter γ_j could be the CES share parameter, such as that employed by Anderson and van Wincoop (2003), or the technology parameter, such as that employed by Eaton and Kortum (2002).

reflects an inverse relationship between the outward MR, $\Pi_{i,t}$, which captures the incidence of trade costs on the producers in j, and the factory-gate price received by the suppliers in j. The intuition for this result in the context of language is that when consumers and producers communicate more easily (both internationally and internally), the incidence of trade costs on the producers will be lower and, through (4), they will internalize these gains by enjoying higher factory-gate prices. In turn, the higher factory-gate prices will translate into higher nominal income, which will result in more trade via equation (1). We capitalize on these relationships in the next section where we discuss the structural links between language, trade, and welfare.

4.1.   On the structural links between language, trade, and welfare

Our new common language indices, and possible changes associated with them, offer an opportunity to generate novel and broader insights about the GE impact of changes in domestic trade costs within the standard quantitative trade framework described by equations (1)–(4). The focus on language is particularly interesting because, by definition, an improvement in the DCL index for a given country is also (almost) always associated with a favorable change in this country's ICL index.[24] Thus, our setting provides an example of how a policy targeted at reducing domestic trade costs can simultaneously also lead to a reduction in international trade costs. The discussion in this section aims to explain and disentangle these effects as well as trace their separate and combined GE impact. For the sake of clarity and simplicity, and to maintain consistency with our counterfactual experiment, we focus on Canada as an example throughout this discussion. However, the concepts expressed are readily generalizable.[25]

---

[24]The unrealistic exception is when the language(s) that become more common domestically are not spoken in any other country in the world. In this case, the rest of the countries in the world will still be affected, but exclusively through GE forces.

[25]See, for example, Ginsburgh and Weber (2011) and Ginsburgh and Weber (2020), who present several cases of linguistic policies across many countries. Given the richness of our dataset, our methodology could be applied to a wide set of public policies that concern linguistic diversity in different countries in particular those related to the treatment of minority languages.

The hypothetical scenario of making Quebec bilingual would directly impact both the DCL and ICL components of the CL index for Canada. Improved domestic communication would decrease internal trade costs in Canada. Additionally, because more Canadians would speak English, the trade costs between Canada and other English-speaking countries would fall. While making Quebec bilingual would have no direct impact on Canada's trade costs with countries that do not speak English, trade between Canada and those countries would be affected through GE (trade-diversion) forces. Even though we cannot provide an analytical decomposition of the impact of these changes in Canada's language indices on trade and welfare, we can rely on system (1)–(4) to characterize the relative strength and the direction of these effects.[26]

The strongest impact of making Quebec bilingual would be on domestic trade in Canada. Under the assumption that improved communication promotes trade, domestic trade in Canada would increase because a larger fraction of people in Quebec speaking English would cause that province to trade more with the other primarily English-speaking provinces and territories in Canada. In effect, it would be akin to reducing Quebec's trade costs with the rest of Canada. This effect is captured by Equation (1) and we argue that it would be the strongest effect for two reasons. First, because it is a direct, first-order effect. Second, because it would further reduce the already lowest trade costs that Canada faces—its domestic trade costs.

Barring other changes, increased domestic trade in Canada would be associated with a decrease in Canada's exports to foreign countries. This is a second-order, GE trade diversion effect, which is captured by the multilateral resistances (2)–(3) in system (1)–(4).[27] The intuition is that if internal trade costs within Canada decrease, producers and consumers in Canada would enjoy lower

---

[26]We have structured the counterfactual analysis in Section 6 to complement the discussion in this section in order to provide further insight into the alternative channels through which changes in domestic common language may impact trade and welfare in the world.

[27]Anderson et al. (2020) label these effects 'Conditional General Equilibrium' effects. They are 'general equilibrium' effects because they affect all countries, but they are dubbed 'conditional' because the size variables remain unchanged. Alternatively, Head and Mayer (2014) label these effects 'Modular Trade Impact'.

multilateral resistances to shipping and buying goods, respectively. However, according to (2)–(3), producers and consumers throughout the rest of the world would face higher multilateral resistances because of the inverse relationship between the MRs in Canada and those of other countries. Thus, all other countries—regardless of whether or not they speak English—would be negatively affected. Further, the more integrated a nation is with Canada, the stronger the negative trade-diversion impact on it would be. For this reason, we expect the United States would be one of the most negatively affected countries.[28]

To gauge the impact of lower domestic trade costs on Canada's welfare, we utilize Equation (4), which conveniently links factory gate prices $(p_{j,t})$ to the corresponding outward MRs $(\Pi_{i,t})$ on the producer side. As noted earlier, the outward MR faced by Canada's producers would fall. Through (4), this change would translate into higher factory-gate prices and higher nominal income for the Canadian producers. In combination with the lower prices for the Canadian consumers, reflected in lower inward MRs, the increase in the factory-gate prices implies that Canada's welfare/real GDP would rise. The opposite would happen throughout the rest of the world where, all else equal, producers would suffer lower prices and consumers would suffer higher prices. However, the negative impact on the rest of the world would be mitigated by a third-order size effect, which is captured by Equation (1). Specifically, *ceteris paribus*, if Canada were richer and larger, it would trade more with all other countries.[29]

As noted earlier, making Quebec bilingual would also directly impact Canada's international trade

---

[28]We also recognize that the decrease of domestic trade costs in Canada may exhibit third-order GE effects. For example, the stronger resistance that the outside countries would face for trading with Canada would induce them to trade more with each other. However, we note that these effects would be very small empirically for two reasons: by construction, based on system (2)–(3), and because most of the diverted trade will actually be absorbed domestically in the outside countries due to the large home bias in trade.

[29]In principle, it is even possible that this third-order trade-creation size effect can dominate the second-order trade-diversion effect. For example, Anderson et al. (2020) find that, through its size effects on the EU and the US, TTIP would have lead to increased exports from Hong Kong, Vietnam, and Malaysia—even though these countries were not part of the planned trade agreement.

through changes in international trade costs. The initial impact would be direct trade creation between Canada and other English-speaking countries, channeled through lower bilateral trade costs in Equation (1). There would also be GE trade diversion effects, which are captured by the MRs (2)–(3). The impact on the MRs for Canada's consumers and producers, as well as for the consumers and the producers in all English-speaking countries, would be positive. For countries that do not speak English, however, the effects on their consumers and producers would be negative. In principle, it is possible that the trade creation effects experienced by English-speaking countries could outweigh the negative impact of increased domestic trade within Canada, both for Canada and for its English-speaking partners. The effects on trade for non-English-speaking countries are negative in both cases and compound. Finally, the combined changes in the MRs triggered by the improvement in Canada's ICL component would translate into positive welfare effects for Canada and all other English-speaking countries. Non-English-speaking countries would suffer welfare losses.

The joint effect of the changes in Canada's DCL and ICL components would be positive and reinforce each other for Canada. For other English-speaking countries, the joint effect is ambiguous. As will be demonstrated in the empirical analysis in Section 5, it is possible that the positive welfare effects experienced by Canada's English-speaking trading partners could outweigh the negative impact of increased domestic trade within Canada, resulting in a net benefit. However, this is not assured in general, and these countries could experience net losses instead if the domestic trade effects dominate. Finally, countries that do not speak English will suffer unambiguous welfare losses because the changes in both the DCL and ICL components imply negative impacts on these countries.

The preceding discussion offers some intriguing insights and implications about the economic impacts of domestic policies. Many of these go beyond the scope of this paper so we include only a short discussion of two illustrative examples. First, the trade-diversion effects triggered by making Quebec bilingual or, more broadly, by any domestic policy that reduces internal trade costs and pro-

motes internal integration, have political economy implications and point to potential opportunities for rent seeking. These policies can make a country effectively more closed to international trade, which—naturally—will hurt its trading partners. Importantly, however, this is not achieved through any trade protection policy such as tariffs that could be subject to international condemnation and multilateral penalties. Instead, it is a unilateral, welfare-promoting domestic policy with implications for international trade, which—as demonstrated empirically in Section 5—can be economically sizable.

A second general implication is that one can apply the same logic, discussion, end even our empirical (estimation and simulation) methods to study the impact of any domestic policy on international trade. For example, consider the recent surge in the implementation of various standards (e.g., SPS and TBT measures), which play a central role in the modern trade policies of many countries. Continuing our focus on Canada, consider a simple hypothetical scenario where, in order to export to Quebec, a firm from British Columbia were required to satisfy a packaging requirement in French.[30] Once the fixed cost of meeting this domestic requirement is sunk for Canadian firms, it may also decrease international trade costs for trade with other French-speaking partners. Thus, in effect, the domestic standard also favorably alters international trade costs too, thus promoting international trade.[31]

---

[30] According to Subsection 6(2) of the *Consumer Packaging and Labelling Regulations* in Canada, "all" mandatory label information must be shown in English and French, except the dealer's name and address, which can appear in either language. Any label information in addition to the mandatory requirements discussed above (i.e., directions for use, promotional statements, etc.) does not have to appear in a bilingual manner. However, the province of Quebec has additional requirements concerning the use of the French language on all products marketed within its jurisdiction. The Province of Quebec has a special unit, handled by the Quebec French Language Bureau and the Ministere de l'Agriculture, des Pecheries et de l'Alimentation du Quebec (MAPAQ), which monitors and enforces the regulations governing labelling in Quebec. Labels that are not in compliance with Quebec's language regulations are subject to large fines, which increase for subsequent offences.

[31] It also possible that the stricter standards for domestic trade may raise international trade costs, in which case trade and welfare would decrease. This is ultimately an empirical question. Importantly, our methods can be extended to both estimate the initial impact of the domestic policy in question, as well as to translate it into a total, GE effect on trade and welfare.

4.2. Language and trade: From theory to empirics

Guided by theory, we follow the latest recommendations from the related empirical literature to translate equation (1) into an econometric model, which will enable us to obtain our direct/PE estimates of the impact of language on trade.[32] First, we estimate a panel version of the model with exporter-time and importer-time fixed effects, which control for the unobservable multilateral resistances as well as any other country-specific determinants of trade. Second, we follow the recommendations of Santos Silva and Tenreyro (2006, 2011) to use the Poisson Pseudo Maximum Likelihood (PPML) estimator.[33] Third, and most important for our purposes, we follow the recommendation of Yotov et al. (2016) that the estimations should be performed with domestic trade flows in addition to international trade flows to be consistent with the underlying theory. The inclusion of domestic trade flows is crucial for our analysis because it enables us to estimate the impact of common *domestic* language. To get access to consistently constructed international and domestic trade flows data, we utilize the dataset of Baier et al. (2019), which provides domestic and international trade in manufacturing for 68 countries during the period 1988–2006.[34] After taking into account the aforementioned considerations, our estimating model becomes:

$$X_{ij,t} \;\; = \;\; \exp[\alpha_1 ICL_{ij} + \beta_1 DCL_{ij} + \textbf{GRAVITY}'_{ij,t}\alpha + \textbf{SMCNTRY}'_{ij,t}\beta + \pi_{i,t} + \chi_{j,t}] + \epsilon_{ij,t}. \quad (5)$$

Here, $X_{ij,t}$ denotes nominal trade flows from exporter i to importer j at time t, including domestic trade flows. The two key covariates of interest to us are our measures for international common language, ICL_{ij},

---

[32]Head and Mayer (2014) and Yotov et al. (2016) offer recent surveys of the related empirical literature.

[33]The benefits of using PPML are: (i) the estimator successfully handles the heteroskedasticity in trade data that would otherwise lead to inconsistent OLS estimates; and (ii) due to its multiplicative form, the PPML estimator enables us to take advantage of the information contained in the zero trade flows. In a sensitivity analysis, we also experiment with the OLS estimator.

[34]Domestic trade flows are constructed as the difference between the gross value of total production and total exports. The original international trade data come from the United Nations (UN) COMTRADE database, accessed via WITS. The data on total gross production come from the CEPII TradeProd database and the UNIDO IndStat database. Of the 68 countries covered in the dataset of Baier et al. (2019), Macau is the only region that lacks some of the language indices that we use for the analysis. Therefore, our estimating sample includes 67 countries in 1988–2006. The list of countries appears in Table 2, and we refer the reader to Baier et al. (2019) for further details on the construction of the manufacturing data.

and for domestic language proximity, DCL_{ij}. The vector $GRAVITY_{ij,t}$ includes a number standard covariates like bilateral distance that are commonly used in the related empirical 'gravity' literature, which we describe in detail below. $SMCNTRY_{ij,t}$ is a vector of covariates designed to control for domestic trade costs and we introduce its components below. Finally, $\pi_{i,t}$ denotes the set of time-varying exporting-country dummies, which control for all observable and unobservable exporter-specific factors that may influence bilateral trade, and $\chi_{i,t}$ encompasses the set of time-varying destination country dummy variables, which serve the same purpose on the importer side.

5.   The impact of language on foreign and domestic trade

The objective of this section is twofold. First, in subsection 5.1, we use our new common language index to quantify the impact of language on international trade and compare its performance to the most successful and widely used existing language indices from the related literature. Second, in subsection 5.2, we employ our novel index for domestic common language and study its impact on internal domestic trade. In both cases, we find that common language as measured by our index has a large, positive, and significant impact on trade. Finally, a series of sensitivity experiments confirm the robustness of our main findings.

5.1.   Common language and international trade

We begin with an investigation of the impact of language on international trade. For this analysis, we employ the most widely used version of the empirical gravity equation, which does not include domestic trade flows. Thus, by definition, we can only identify the impact of covariates that proxy for international trade costs (the variables in vector $GRAVITY_{ij,t}$ and our ICL index) and do not include the proxies for domestic trade costs (the variables in vector $SMCNTRY_{ij,t}$ and the DCL index). Guided by the voluminous empirical gravity literature, we replace $GRAVITY_{ij,t}$ with a series of trade costs proxies. Our estimating equation becomes:

$$
\begin{aligned}
X_{ij,t} = \ & \exp[\alpha_1 ICL_{ij} + \alpha_2 DIST_{ij} + \alpha_2 CNTG_{ij,t} + \alpha_3 CLNY_{ij,t}] \times \\
& \exp[\alpha_4 EIA_{ij,t} + \alpha_5 EU_{ij,t} + \alpha_6 WTO_{ij,t} + \pi_{i,t} + \chi_{j,t}] + \epsilon_{ij,t}, \ \forall \ i \neq j.
\end{aligned}
\tag{6}
$$

Here, in addition to the key language variable ICL_{ij},[35] we have introduced three standard time-invariant gravity variables including the logarithm of bilateral distance ($DIST_{ij}$), an indicator variable for contiguous borders ($CNTG_{ij}$), and an indicator variable for colonial relationships ($CLNY_{ij,t}$). We also control for a series of time-varying policy variables such as economic integration agreements ($EIA_{ij,t}$), European Union membership ($EU_{ij,t}$), and GATT and WTO membership ($WTO_{ij,t}$). All of the these variables come from the *Dynamic Gravity Dataset* (DGD) of Gurevich and Herman (2018).

Our estimates for the impact of language on international trade are presented in Table 3. The estimates in column (1) are obtained with the ICL indicator of Mayer and Zignago (2011), which is one of the most widely used proxies for common international language throughout the trade and gravity literature. The estimates from column (1) are as expected. Specifically, we find that distance is a significant impediment to trade, while common borders, colonial ties, economic integration agreements, and EU membership promote trade. Similar to Rose (2004), we do not obtain a significant estimate of the impact of GATT/WTO. Overall, the estimates of the effects of the standard gravity variables in column (1) are readily comparable to corresponding estimates from the related literature such as those reported in the gravity estimates meta analysis of Head and Mayer (2014). Turning to the effects of language, the positive and significant estimate on $ICL_{ij,t}$ implies that, all else equal, sharing a common official language increases bilateral trade by 15.72 percent.[36]

Our main results, which are obtained with our newly constructed ICL index, are presented in column (2) of Table 3. Once again, the estimate on ICL_{ij} is positive, large, and statistically significant. To get a sense of the quantitative importance of this result, we calculate two trade volume effects. First, we find that an improvement in the common language index by one standard deviation (0.182) would lead to a statistically significant increase in bilateral trade by 7.28 percent (std.err. 1.931).[37] Second, to set the stage for our counterfactual experiment in which Quebec were to become bilingual, we calculate the direct effect of the corresponding increase in the ICL index for trade between Canada and the United States. All else equal,

---

[35]We also experiment with the indicator variable for commonly spoken official language from the Centre d'Études Prospectives et d'Informations Internationales' (CEPII) database (Mayer and Zignago, 2011).
[36]Calculated as $[\exp(0.146) - 1] * 100 = 15.72$.
[37]Calculated as $[\exp(0.386 * 0.182) - 1] * 100 = 7.28$ with standard errors constructed using the delta method.

our estimates imply that this increase would lead to a 2.31 percent (std.err. 0.600) increase in Canada's exports to the United States.

We conclude this section with a series of sensitivity experiments using our new ICL index. First, in column (3) of Table 3, we include our index together with the indicator for common official language of Mayer and Zignago (2011). The estimates reveal that once our new language index is introduced, the estimate of the indicator for common official language of Mayer and Zignago (2011) is no longer statistically significant, implying that the new language index dominates. This result is consistent with the findings from Melitz and Toubal (2014), who were the first to demonstrate that the continuous measure of international language proximity dominates the standardly used dummy variables for common language. The natural explanation for this result is the wider variability in the ICL index due to the greater coverage of languages.

Table 4 offers a number of additional robustness checks. For comparison purposes, column (1) reproduces our main estimates from column (2) of Table 3. The results in column (2) are obtained with positive trade flows only. They are virtually identical to the results in column (1), implying that the zero trade flows do not affect the standard gravity estimates. This result is common in the literature and is explained by the fact that PPML assigns less weight to the small countries in gravity estimations. Columns (3), (4), and (5) employ data with 2-, 3-, and 5-year intervals, respectively. As can be seen from Table 4, these estimates are very similar to the results from column (1). This supports our decision to use all years in our sample. Finally, the results in column (6) are obtained with the OLS estimator. We find that the estimate of the language variable remains positive and statistically significant, however, its magnitude is significantly larger. Other notable differences in the OLS results are that $CNTG_{ij,t}$ looses statistical significance while the estimate on $EU_{ij,t}$ becomes negative and statistically significant. In combination with the arguments from Santos Silva and Tenreyro (2006) that the OLS estimates are inconsistent, these changes in the results cast doubt on the OLS estimator and favor our choice of PPML to obtain our main findings.

5.2. Common language and domestic trade

This section offers novel estimates of the impact of common language on domestic trade. To obtain

the results, we extend specification (6) from the previous section to a theory-consistent empirical model

that includes domestic trade flows. First, we introduce observations for domestic trade flows ($X_{ii,t}$) to

the dependent variable ($X_{ij,t}$). Then, we add a number of proxies for domestic trade costs. Our new

estimating equation becomes:

$$X_{ij,t} = \exp[\alpha_1 ICL_{ij} + \alpha_2 DIST_{ij} + \alpha_3 CNTG_{ij} + \alpha_4 CLNY_{ij} + \alpha_5 EIA_{ij,t} + \alpha_6 EU_{ij,t} + \alpha_7 WTO_{ij,t}] \times$$

$$\exp[\alpha_8 DCL_{ij} + \alpha_9 HOME_{ij} + \alpha_{10} DIST\_INTRA_{ij} + \alpha_{11} GDPPC \times HOME_{ij,t}] \times$$

$$\exp[\alpha_{12} INST \times HOME_{ij,t} + \pi_{i,t} + \chi_{j,t}] + \epsilon_{ij,t}, \quad \forall \, i,j. \tag{7}$$

The new covariates in (7) are motivated by the related literature. Consistent with the extensive

literature on the home bias effect in international trade, $HOME_{ij}$ is an indicator variable that is

equal to one for domestic trade and equal to zero for international trade.[38] Following Anderson et

al. (2018), we include $DIST\_INTRA_{ij}$ as a proxy for domestic geography/transportation costs.

The internal distance measure is constructed as a population-weighted index, following Mayer and

Zignago (2011), who build on the methods of Head and Mayer (2000). Importantly, by construc-

tion and similar to our language indices, the measure of internal distance is consistent with the

corresponding measure of international distance.[39]

---

[38]Without an intent to be exhaustive, we refer the reader to Hillberry and Hummels (2003), Head and Mayer (2010), and Anderson et al. (2018) for some examples from the voluminous literature concerned with the 'home bias' in international trade.

[39]Specifically, distance—both between countries and (internally) within countries—is calculated as $DIST_{ij} = \sum_{k \in i} Pop_k/Pop_i \sum_{l \in j} Pop_l/Pop_j D_{kl}$, where Pop_k is the population of city k in exporter i, Pop_l is the popu-lation of city l in importer j, and D_{kl} is the bilateral distance in kilometers between city k and city l (using the Great Circle Distance formula). All data on latitudes, longitudes, and population are from the World Gazetteer web page. The same procedure is used to construct consistent international and internal distances. Cities with population of at least 300,000 are used in the calculation. For more details see Gurevich and Herman (2018).

Following Beverelli et al. (2017) and Anderson et al. (2018), we include $GDPPC \times HOME_{ij,t}$ and $INST \times HOME_{ij,t}$ as additional proxies in the vector of domestic trade costs. $GDPPC \times HOME_{ij,t}$, which is constructed as the product of GDP per capita and the home bias indicator $HOME_{ij}$, is designed to proxy for economic development.[40] The motivation for its inclusion is that, all else equal, more developed countries are expected to be more open to international trade and less 'home biased'. $INST \times HOME_{ij,t}$, which is constructed as the product between a measure of country-specific institutional quality and the $HOME_{ij}$ dummy variable, controls for the impact of institutional quality.[41] The motivation for the inclusion of institutional quality as a proxy for domestic trade costs is that, for example, stronger institutions may signal less uncertainty and, therefore, lead to more international trade and lower home bias. Our contribution to the vector of proxies for domestic trade costs is the introduction of domestic common language $DCL_{ij,t}$, as a category designed to capture domestic "cultural proximity".

Our main findings are presented in Table 5. For comparison purposes, column (1) reproduces the main estimates from column (3) of Table 3 discussed in the previous section. Column (2) introduces domestic trade flows and the corresponding proxies for domestic trade costs from specification (7). Without going into details, we note that the results are consistent with previous findings and are as expected. Specifically, the very large, positive, and significant estimate on $HOME_{ij}$ in column (2) reveals the presence of significant home bias in international trade. The estimate on $DIST\_INTRA_{ij}$ is negative, sizable, and significant, suggesting that internal distance is a significant obstacle to domestic sales, just like international distance.[42] The negative estimate on $GDPPC \times HOME_{ij,t}$ implies that, all else equal, more developed countries are more open to

---

[40]Data on GDP and population are from the Penn World Tables.

[41]To ensure maximum coverage of our proxy for institutional quality, we use the Civil Liberties indicators of the Freedom House initiative, https://freedomhouse.org/report/freedom-world-2016/methodology.

[42]The estimate on internal distance is slightly smaller but readily comparable to the corresponding estimate of the impact of international distance. This is encouraging for the specification of trade costs in our model because large differences would have suggested possible omission of variables in the vector of domestic trade costs.

international trade relative to domestic trade. This result is consistent with the findings of Anderson et al. (2018). Finally, similar to Beverelli et al. (2017), we find that stronger institutions promote international relative to domestic trade from the estimate on $INST \times HOME_{ij,t}$.

The estimates on all but one of the international trade cost covariates are very similar between columns (1) and (2) in Table 5. The only notable difference is that the estimate on $GATT\_WTO_{ij,t}$, which is small and not statistically significant in column (1), becomes positive, large, and statistically significant in column (2). This result is consistent with the main finding of Larch et al. (2019) who argue that GATT and WTO have been effective in promoting international trade among member countries. The explanation is that the introduction of domestic trade flows in the estimating gravity equation allows identification of the impact of GATT/WTO on bilateral trade creation at the expense of diversion from domestic trade flows. Overall, the findings from column (2) of Table 5 are consistent with those from the related literature, which is reassuring for the representativeness of our estimating sample.

Two important findings stand out from the estimates in column (3) of Table 5, which introduces our novel domestic language index. First, we obtain a large, positive, and statistically significant estimate of the effects of DCL_{ij}. We postpone the analysis and interpretation of this estimate for later when we discuss the results from the next column, which we view as our main specification. The second key finding from column (3) is that the estimates on all but one of the other gravity covariates have not changed. Importantly, the single estimate that changed significantly is the one on $HOME_{ij}$, which decreased from 9.657, in column (2), to 8.849, in column (3). The fall in the proxy for home bias after the introduction of DCL_{ij} was expected. The implication is that domestic language plays an important role in the home bias in trade. Thus, once we control explicitly for the impact of domestic language/culture, the estimate of the home bias effect becomes significantly smaller. That the estimates of the other domestic variables are largely unchanged implies that

language represents a new factor to home bias that is separate from the other factors.

The estimates in column (4) of Table 5 are obtained after allowing for the home bias in trade to vary over time. Instead of a single home bias dummy, we introduce a home bias dummy variable for each year in the sample. The estimates are virtually identical to those from column (3). Most importantly, once again, we obtain a large, positive and statistically significant estimate of the impact of DCL_{ij}, implying that, all else equal, a one standard deviation improvement (an increase of 0.149) in domestic language similarity will lead to a 16.02 percent (std.err. 6.604) increase in domestic trade. Extending our Canadian example from the previous section, our estimate on DCL_{ij} implies that if all people in Quebec spoke English in addition to French, then domestic trade in Canada would have been twice as large, representing a 106.68 percent increase (std.err. 57.475).

We acknowledge that our estimate of a 107 percent increase in Canada's domestic trade due to increased communication is probably an overestimate of the true effect of language on trade for two reasons. First, it is only a PE effect, which—by construction—overestimates the total impact of common domestic language on trade (we demonstrate this in the next section.) Second, this is an overestimate of the true impact of language because our DCL index is broadly a cultural proxy and reflects more than just linguistic similarity between populations in the same country. Nevertheless, we do believe that our estimate is an indicator for the significant benefits that can be gained from internal integration in Canada and elsewhere. Furthermore, the same criticisms apply to the estimated impact of language on international trade, which are now widely accepted in the literature.

We conclude this section by testing the robustness of our main findings on domestic common language. Table 6 presents the results from a series of sensitivity experiments. For comparison purposes, column (1) of Table 6 reproduces our main estimates from column (4) of Table 5. The results in column (2) are obtained with positive trade flows only and, as was the case in the previous

section, they are virtually identical to the results in column (1). The implication is that the zero trade flows do not affect the standard gravity estimates. Columns (3), (4), and (5) employ data with 2-, 3-, and 5-year intervals, respectively. As can be seen from Table 6, these estimates are very similar to the results from column (1). This confirms our findings that were based on the international sample and supports our decision to use all years in our sample. Finally, the results in column (6) are obtained with the OLS estimator. Similar to the results from the previous section, we find that the estimate of the domestic language variable remains positive and statistically significant. However, its magnitude becomes significantly larger. Other notable differences in the OLS estimates are that, as before, $CNTG_{ij}$ looses statistical significance, while the estimate on $EU_{ij,t}$ becomes negative and statistically significant. In addition, we see two other changes; $GATT\_WTO_{ij,t}$ and $INST \times HOME_{ij,t}$ also loose statistical significance. Once again, our conclusion, in combination with the arguments from Santos Silva and Tenreyro (2006), is that the OLS estimates are inconsistent and these results favor our choice of the PPML estimator.

6. The trade and welfare effects of a bilingual Quebec

Our previous findings suggest that linguistic diversity leads to communication and trade barriers across countries and is an impediment to domestic integration. Ginsburgh and Weber (2011) and Ginsburgh and Weber (2020) provide many examples of linguistic policies aimed at reducing the economic costs of linguistic diversity and show how language policies have been used to shape preferences and homogenize societies. While our model may accomodate many different scenarios, our focus is on Canada where language has been a contentious issue since the British Conquest of 1760. Many historians and linguists describe the attempts to impose uniformity and create a national Canadian identity by restricting French language education (Dickinson and Young, 2014).[43]

---

[43]Many regulations in different provinces were introduced in order to regulate or to eliminate French language instruction. The New Brunswick Common Schools Act of 1871 prescribed English language textbooks in all schools within the province. In Prince Edward Island, the Public Schools Act of 1877 established obligatory, non-denominational

Over the years, these attempts created a separatist sentiment in Quebec and inspired the Révolution Tranquille that unfolded in the province at the beginning of the 1960s. The policies created as a part of this movement offer an opportunity to understand how a shift in language policy can affect patterns of domestic and international trade as well as the distribution of welfare across countries.

The Révolution Tranquille was a time of intense social, political, and economic changes in Quebec.[44] It was characterized by the affirmation of the *Québecois* identity, which was and remains closely related to the expansion of the French language. Plans to replace English terminology were supported by the *Office Québécois de la Langue Française* established in 1961. In 1969, Quebec introduced its first language policy—Bill 63—which guaranteed parents the right to choose the language of instruction for their children. The Ministry of Education ensured that children taught in English acquire "a working knowledge of French". In 1974, the *Official Language Act* (Bill 22) was enacted making French the language of Quebec's government administration and required that business be conducted in French, including most services, signage, and labor relations. However, the law was not sufficiently binding and its application remained vague. In 1977, the *Parti Québécois* proposed and passed into law Bill 101 (also known as the Charter of the French Language), which made French the sole official language of the province and established the fundamental language rights of French-speakers.[45] Since then, French has been recognized as the language of the legislature and the courts, although judgments and proceedings may be in another language, if the parties so agree. It has been the language used in business and the workplace, and educational instruction for kindergarten through secondary school. Most firms with 50 or more employees must maximize the

---

instruction; textbooks were standardized; and French language textbooks were withdrawn because of their religious content. In Ontario, a regulation in 1889 under the Public School Act established exclusive use of English as the language of instruction.

[44]A literature on language policy and the relationship between language and economic outcomes appeared in Canada during this period (c.f. Vaillancourt, 1983). In particular, Breton (1978)—from a perspective based on nationalism—initiated the trend of applying economic analysis to language phenomena.

[45]Quebec is bilingual on the constitutional and federal levels but officially allows only French in its provincial institutions.

use of French given the environment they operate in (Vaillancourt, 2019). All children must attend a school taught primarily in French. The only exceptions are cases in which at least one of the child's parents received instruction from an English school in Canada. In particular, all immigrant children must attend French school, even if they are from an English-speaking country.

Motivated by the interesting history and important cultural and economic implications of Quebec's Révolution Tranquille, we quantify the effects of language policies on trade and welfare by examining it as a case in point. In particular, we estimate the economic impacts of a counterfactual experiment in which Quebec were to become fully bilingual in English—the other prominent language in Canada—as well as French. This analysis also provides novel insight into potential consequences of the more recent discussion surrounding the requirement that many federal jobs in Canada be bilingual in both English and French. To date, much of that discussion has been based on the effects that the requirement has had on the diversity of government employees, but our analysis indicates it may have considerable economic effects as well.[46] We recognize that the proposed analysis may not be fully realistic but, nevertheless, we find the results informative. To broaden the context of the analysis, we propose two additional or alternative interpretations and motivations for this experiment. First, there has been a natural trend of increasing bilingualism in Quebec, which has nearly doubled from about 25.5 percent in 1961 to 44.5 percent in 2016 (Turcotte, 2019). Thus, one interpretation of our experiment is as an *ex-ante* evaluation of the extreme scenario in which the bilingualism trend is complete and 100 percent of Quebec's population speaks both French and English.[47] Second, our experiment can be thought of as an *ex-post* quantification of the economic

---

[46]For example, see McCullough (2020) and Keung (2020).

[47]A related interpretation of our findings is that they are the result of an education reform. Such an interpretation highlights two caveats of our analysis. First, for various historical, political, and practical reasons, it is unrealistic to believe that everyone in Quebec can and will learn English. Second, even if Quebec were to become fully bilingual, it may not lead to the results that we obtain and describe in this section. The reason is that language is also a cultural proxy in our analysis. Therefore, by changing language, we are effectively simulating a broader cultural/identity change because the empirical estimates do not disentangle theses different aspects of language. While people can learn a language, cultural change is slower and more difficult.

impacts of the significant decline in the percentage of English speakers in Quebec, which started in the 1960s because of worsening economic conditions and was reinforced further by the Révolution Tranquille.

The counterfactual experiment is performed in three steps.

1. **Solve the model in the baseline.** To obtain the baseline indices of interest, we solve system (1)–(4) using estimates for bilateral trade costs ($\tau_{ij,t}$) based on the estimates from our preferred specification (7). Data from the last year from our sample is used for the trade, output, and expenditure values, making 2006 our baseline year. Consistent with most of the related literature, we focus on total exports ($X_i = \sum_{j \neq i} X_{ij}$) and real GDP ($W_i = Y_i / P_i$) as a proxy for welfare. To construct our welfare indices, we need a value for the trade elasticity θ. We select a conventional value, $\hat{\theta} = 7$ (c.f. Head and Mayer (2014)). In general, a higher value of θ will lead to lower welfare effects because costumers do not value variety as much.

2. **Define the counterfactual scenario.** We combine the (baseline) estimates from specification (7) with a new, hypothetical language index for Canada, which assumes that all people in Quebec speak both English and French. Modelling a bilingual Quebec requires changes to both the domestic and international language indices for Canada. Specifically, the DCL index for Canada will increase as will the values of Canada's ICL index with English-speaking countries. All other language index values, for Canada and other countries, remain unchanged between the baseline and the counterfactual analysis. In order to highlight the relative importance of the competing forces that act in response to the changes in the domestic versus international trade costs, we construct two counterfactual trade cost vectors. The first takes into account only the change in Canada's DCL index, representing a scenario in which only domestic trade costs in Canada were lower but Canada's international trade costs remained unchanged. The second takes into account the changes in both Canada's DCL and ICL indices, thereby representing the full impact

of the change on trade costs between Canada and the rest of the world.

3. **Solve the model in the counterfactual scenario.** We solve system (1)–(4) again using the counterfactual trade cost vectors while keeping all other parameters and exogenous variables at their baseline level. Utilizing the two counterfactual trade cost vectors from the previous step, we obtain two sets of counterfactual results: one that is in response to changing only Canada's DCL index and one in response to simultaneous changes in Canada's DCL and ICL index with its English-speaking partners. Finally, we construct the percentage changes in total exports $\left(\%\Delta X_i = \%\Delta \sum_{j\neq i} X_{ij}\right)$ and welfare $(\%\Delta W_i = \%\Delta(Y_i/P_i))$ between the baseline and the counterfactual scenarios for all countries in our sample.

Our findings from the two experiments are presented in Table 7. Panel A reports the impact of Quebec being bilingual on total exports and panel B reports the impact on welfare. We rank the countries in Table 7 by the welfare effect of a change in the DCL component of the language index and have isolated Canada as the country of central interest. We also visualize our findings in Figures A.1–A.4 in the Appendix. Column (1) of Table 7 lists the percentage changes in total exports for each country in response to a change in only in the DCL index for Canada. Several interesting and intuitive findings stand out. First, as expected, we find that Canada's total exports decrease significantly by 2.61 percent. The explanation for this result is trade diversion away from Canada's international trade and toward domestic sales within Canada. The impact on all other countries is also negative and expected, with intuitive variation. For example, the United States is the most severely affected country and experiences a 0.67 percent decrease in total exports. The geographical proximity and tight economic integration between the United States and Canada are natural explanations for this result.

The next group of countries whose exports decrease the most include Senegal, Panama, Kenya, Ecuador, and Tanzania. The negative impacts on those countries are an order of magnitude smaller

than the effect on the United States, varying between 0.06 percent for Ecuador to 0.08 percent for Senegal. Although small in relative magnitude, we find this result interesting, especially from a development perspective. The United Kingdom follows these countries with a relatively small decrease in exports of about 0.06 percent, which we largely attribute to the colonial relationships between Canada and the United Kingdom. The countries that suffer the least trade diversion are a group of European economies including Norway, Austria, Germany, and the Netherlands as well as Singapore. A combination of relative remoteness from Canada and relatively large economic size are the natural explanations for these findings.

The results in column (2) of Table 7 present the percentage changes in total exports in response to the full change in both of Canada's indices. Four main results stand out. First, while still negative, the impact on Canada's exports is significantly smaller (-1.52 percent compared to -2.61 percent). The difference is due to trade creation with the other English-speaking countries in the sample. Second, the impact on the United States remains negative but is more than three times smaller in absolute value. The explanation is a combination of geographical and economic proximity combined with lower trade costs stemming from the increase in the language index. Third, some countries such as Senegal suffer even larger decreases in exports. The explanation is that the first-order GE trade-diversion effect for Senegal from the increased domestic trade in Canada is now magnified by a second-order GE effect due to the fact that few people in Senegal speak English. Thus, the diversion of Canada's trade toward English-speaking countries will further hurt Senegal. Finally, we note that the trade creation effects from the improvement in Canada's ICL component are strong enough to overcome the initial negative trade-diversion effects for a small group of English-speaking countries. These countries include the United Kingdom, Ireland, and Australia, who actually exhibit increased total exports in column (2).

Panel B of Table 7 reports the real GDP/welfare effects of both experiments. In column (3), we

first report the percentage changes that are obtained in response to the change in only the DCL index for Canada. The main result, as expected, is that Canada's welfare increases as a result of the lower domestic trade costs. Specifically, we estimate a 0.94 percent increase in Canada's real GDP. The impact on the rest of the countries in the sample is negative. Again, the United States experiences the largest decline of 0.02 percent. While this percentage change may seem small, the corresponding impact on real GDP in levels is larger than 30 billion dollars! The next group of countries whose real GDP decreases the most (about 0.01 percent) includes Trinidad and Tobago, Qatar, and Kuwait. The impact of the language index change is relatively large for those countries because of the close, long-standing, and uninterrupted trade relationships between Canada and each of them. Finally, the countries that are affected the least, but still with some losses, are exclusively Asian economies including Japan, Korea, India, and China.

Column (4) of Table 7 reports the percentage change in real GDP in response to the full change in both components of Canada's language indices. The main findings from column (4) are as follows. Canada's welfare gains further grow to a total 1.22 percent increase in real GDP, underscoring the importance of Canada's trade relations with its English-speaking partners. As expected, the additional impacts are positive for other English-speaking countries and significantly mitigate their welfare losses. However, in most cases, these positive effects are not strong enough to overcome the negative trade-diversion effect on their real GDP from the change in Canada's DCL. Ireland is the only country that would see a reversal from a negative to a positive welfare impact. As with exports, most non-English speaking countries face exacerbated welfare losses due to the combined trade diversion towards Canada's domestic market and other English-speaking countries.

In addition to providing new insights into the economic impacts of language in Canada, the counterfactual analysis in this section demonstrates more broadly that domestic policies that changes domestic trade costs can have large international effects and far reaching impacts on economic wel-

fare. Understanding these relationships is critical for better understanding the interconnected nature of the global economy.

7.   Conclusion

International economics has long touted the ways in which international diversity influences economic activities. In this paper, we demonstrate that the same can be said of within-country diversity in the form of differing languages and identity. Such diversity can represent significant barriers to domestic and international trade and welfare.

To study these issues, we construct a new index reflecting both international and domestic language diversity. Empirical estimates of the impact of language on trade indicate that linguistic proximity is an influential determinant of bilateral trade—both internationally and domestically. In particular, we find new evidence that domestic common language is a key part of the prevalent home bias present in international trade.

We further explore the role of domestic language and identity by examining the impacts of Canada's Révolution Tranquille and the rise of the French language in Quebec, Canada's only majority French-speaking province. Using a representative "new" quantitative trade framework, we simulate the effects of Quebec becoming bilingual in English and French to cast light on its likely economic effects. The simulation results show that domestic language diversity has significant trade and welfare effects. Quebec becoming bilingual would promote domestic trade within Canada at the expense of foreign trading partners. For many English-speaking countries, these losses would be at least partially offset by trade creation stemming from lower international trade costs and wealth effects. For non-English-speaking countries, however, the impacts would be entirely negative.

Our analysis highlights several important new aspects of trade and presents a framework with which to analyze them. The role of language diversity and Quebec's language policies is informative of

the likely impacts of other types of domestic policies that contribute to home bias. These policies can have significant, far reaching impacts on both domestic and international welfare that should not be overlooked.

References

Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov, "Intra-national trade costs: Assaying regional frictions," *European Economic Review*, 2019, *112* (C), 32–50.

Alesina, Alberto and Eliana La Ferrara, "Ethnic diversity and economic performance," *Journal of economic literature*, 2005, *43* (3), 762–800.

__ , **Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg**, "Fractionalization," *Journal of Economic Growth*, 2003, *8* (2), 155–194.

Alesina, Alberto Francesco and Bryony Reich, "Nation building," *Unpublished*, 2015.

Anderson, Benedict Richard O'Gorman, *Imagined communities : reflections on the origin and spread of nationalism*, Verso, 1983.

Anderson, James E. and Eric van Wincoop, "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 2003, *93* (1), 170–192.

__ **and Yoto V. Yotov**, "The Changing Incidence of Geography," *American Economic Review*, 2010, *100* (5), 2157–2186.

Anderson, James E., Ingo Borchert, Aaditya Mattoo, and Yoto V. Yotov, "Dark costs, missing data: Shedding some light on services trade," *European Economic Review*, 2018, *105* (C), 193–214.

Anderson, James E. Mario Larch, and Yoto V. Yotov, "Transitional Growth and Trade with Frictions: A Structural Estimation Framework," *Economic Journal*, 2020, *130* (630), 1583–1607.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Baier, Scott L., Yoto V. Yotov, and Thomas Zylkin, "On the Widely Differing Effects of Free Trade Agreements: Lessons from Twenty Years of Trade Integration," *Journal of International Economics*, 2019, *116*, 206–226.

Beverelli, Cosimo, Alexander Keck, Mario Larch, and Yoto V. Yotov, "Institutional Quality, Trade and Development: A Quantitative Analysis," unpublished manuscript, 2017.

Breton, Albert, "Nationalism and language policies," *Canadian Journal of Economics*, 1978, pp. 656–668.

Carr, J, "Bilingualism in Canada. Is the use of the English language a natural monopoly," *Economic et langue*, 1985, pp. 27–37.

Coşar, A. Kerem and Banu Demir, "Domestic Road Infrastructure and International Trade: Evidence from Turkey," *Journal of Development Economics*, 2016, *118*, 232–244.

__ **and Pablo D. Fajgelbaum**, "Internal Geography, International Trade, and Regional Specialization," *American Economic Journal Microeconomics*, 2016, *8* (1), 24–56.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," Chapter 4 in the Handbook of International Economics Vol. 4, eds. Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogoff, Elsevier Ltd., Oxford, 2014,

pp. 197–261.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Romain Wacziarg, "The political economy of linguistic cleavages," *Journal of development Economics*, 2012, *97* (2), 322–338.

— , — , **and** — , "Linguistic cleavages and economic development," in "The Palgrave handbook of economics and language," Springer, 2016, pp. 425–446.

— , **Shlomo Weber, and Ignacio Ortuño-Ortín**, "Linguistic diversity and redistribution," *Journal of the European Economic Association*, 2009, *7* (6), 1291–1318.

Dickinson, John and Brian Young, *A short history of Quebec*, McGill-Queen's University Press, 2014.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, apr 2018, *108* (4-5), 899–934.

Easterly, William and Ross Levine, "Africa's growth tragedy: policies and ethnic divisions," *The quarterly journal of economics*, 1997, *112* (4), 1203–1250.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Egger, Peter H. and Andrea Lassmann, "The language effect in international trade: A meta-analysis," *Economics Letters*, 2012, *116* (2), 221 – 224.

Egger, Peter H and Andrea Lassmann, "The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design," *The Economic Journal*, 2015, *125* (584), 699–745.

— **and Farid Toubal**, "Common spoken languages and international trade," in "The Palgrave handbook of economics and language," Springer, 2016, pp. 263–289.

Fearon, James D, "Ethnic and cultural diversity by country," *Journal of economic growth*, 2003, *8* (2), 195–222.

Gazzola, Michele and Bengt-Arne Wickström, *The economics of language policy*, MIT Press, 2016.

Gellner, E. and D.C.S.N.E. Gellner, *Nations and Nationalism* Cornell paperbacks, Cornell University Press, 1983.

Ginsburgh, Victor and Shlomo Weber, "Language disenfranchisement in the European Union," *JCMS: Journal of Common Market Studies*, 2005, *43* (2), 273–286.

— **and** — , *How many languages do we need?: The economics of linguistic diversity*, Princeton University Press, 2011.

— **and** — , "The economics of language," *Journal of Economic Literature*, 2020, *58* (2), 348–404.

— , **Jacques Melitz, and Farid Toubal**, "Foreign Language Learning: An Econometric Analysis," CEPR Discussion Papers 10101, C.E.P.R. Discussion Papers August 2014.

— , **Shlomo Weber, and Palgrave Macmillan**, *The Palgrave handbook of economics and language*, Springer, 2016.

Gurevich, Tamara and Peter Herman, "The Dynamic Gravity Dataset: 1948-2016," 2018. USITC Working Paper 2018-02-A.

Head, K. and T. Mayer, "Illusory Border Effects: Distance Mismeasurement Inflates Estimates of Home Bias in Trade," *in Brakman, S. and P. van Bergeijk eds. The Gravity Model in International Trade: Advances and Applications, Cambridge*, 2010.

Head, Keith and Thierry Mayer, "Non-Europe: The Magnitude and Causes of Market Fragmentation in the EU," *Review of World Economics (Weltwirtschaftliches Archiv)*, June 2000, *136* (2), 284–314.

___ **and** ___ , "What Separates Us? Sources of Resistance to Globalization," *Canadian Journal of Economics*, nov 2013, *46* (4), 1196–1231.

Head, Keith and Thierry Mayer, "Gravity Equations: Workhorse, Toolkit, and Cookbook," Chapter 3 in the Handbook of International Economics Vol. 4, eds. Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogoff, Elsevier Ltd., Oxford, 2014, pp. 131–195.

Hillberry, R. and D. Hummels, "Intra-national Home Bias: Some Explanations," *Review of Economics and Statistics*, 2003, *85* (4), 1089–1092.

Hobsbawm, E.J., E.J.E. Hobsbawm, Cambridge University Press, History E-Book Project, E.J.H. Hobsbawm, Queen's University of Belfast, American Council of Learned Societies, H.E. J, and P.E.J. Hobsbawm, *Nations and Nationalism Since 1780: Programme, Myth, Reality* Cambridge paperbacks : history, Cambridge University Press, 1992.

Keung, Nicholas, "Fill the six vacancies on the Federal Court with Black, Indigenous and People of Colour judges, legal groups urge Ottawa," *The Toronto Star*, September 14 2020.

Laitin, David D, "What is a language community?," *American Journal of political science*, 2000, pp. 142–155.

Larch, Mario, José-Antonio Monteiro, Roberta Piermartini, and Yoto V. Yotov, "On the effects of GATT/WTO membership on trade: They are positive and large after all," WTO Staff Working Papers ERSD-2019-09, World Trade Organization (WTO), Economic Research and Statistics Division 2019.

Lewis, Bernard, *From Babel to dragomans: interpreting the Middle East*, Oxford University Press, 2005.

Marschak, Jacob, "Economics of language," *Behavioral Science*, 1965, *10* (2), 135–140.

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's Distances Measures: The GeoDist Database," *CEPII Working Paper 2011 - 25*, 2011.

McCullough, J.J., "Trudeau says he wants to tackle systemic racism. He should start with this policy.," *The Washington Post*, September 30 2020.

Melitz, Jacques, "A framework for analyzing language and welfare," CEPR Discussion Papers 9091, C.E.P.R. Discussion Papers August 2012.

___ **and Farid Toubal**, "Native Language, Spoken Language, Translation and Trade," *Journal of International Economics*, 2014, *93* (2), 351–363.

Millimet, Daniel L and Thomas Osang, "Do state borders matter for U.S. intranational trade? The role of history and internal migration," *Canadian Journal of Economics*, 2007, *40* (1), 93–126.

Ramondo, Natalia, Andres Rodriguez-Clare, and Milagro Saborio-Rodriguez, "Trade, Domestic Frictions, and Scale Effects," *American Economic Review*, 2016, *106* (10), 3159–3184.

Rose, Andrew K., "Do We Really Know That the WTO Increases Trade?," *American Economic Review*, 2004, *94* (1), 98–114.

Santos Silva, João M.C. and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641–658.

_ **and** _ , "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator," *Economics Letters*, 2011, *112* (2), 220–222.

Simons, Gary F and Charles D Fennig, "Ethnologue: Languages of the world, twenty," *Dallas, Texas: SIL International. Online version: http://www. ethnologue. com Accessed*, 2018, *26* (12), 2018.

Stokes, Bruce, "What It Takes to Truly Be 'One of Us'," *Pew Research Center*, 2017, *1*.

Turcotte, Martin, *Results from the 2016 Census: English-French Bilingualism Among Canadian Children and Youth*, Statistics Canada= Statistique Canada, 2019.

Vaillancourt, François, "The economics of language and language planning," *Language Problems and Language Planning*, 1983, *7* (2), 162–178.

_ , "Assessing Bill 101 after 40 years: An introduction," *Language Problems and Language Planning*, 2019, *43* (2), 103–112.

Wolf, H. C., "Intra-national Home Bias in Trade," *Review of Economics and Statistics*, 2000, *82* (4), 555–563.

Yotov, Yoto V., Roberta Piermartini, José-Antonio Monteiro, and Mario Larch, *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, Geneva: UNCTAD and WTO, 2016.

Figure 1 – International Common Language in the World
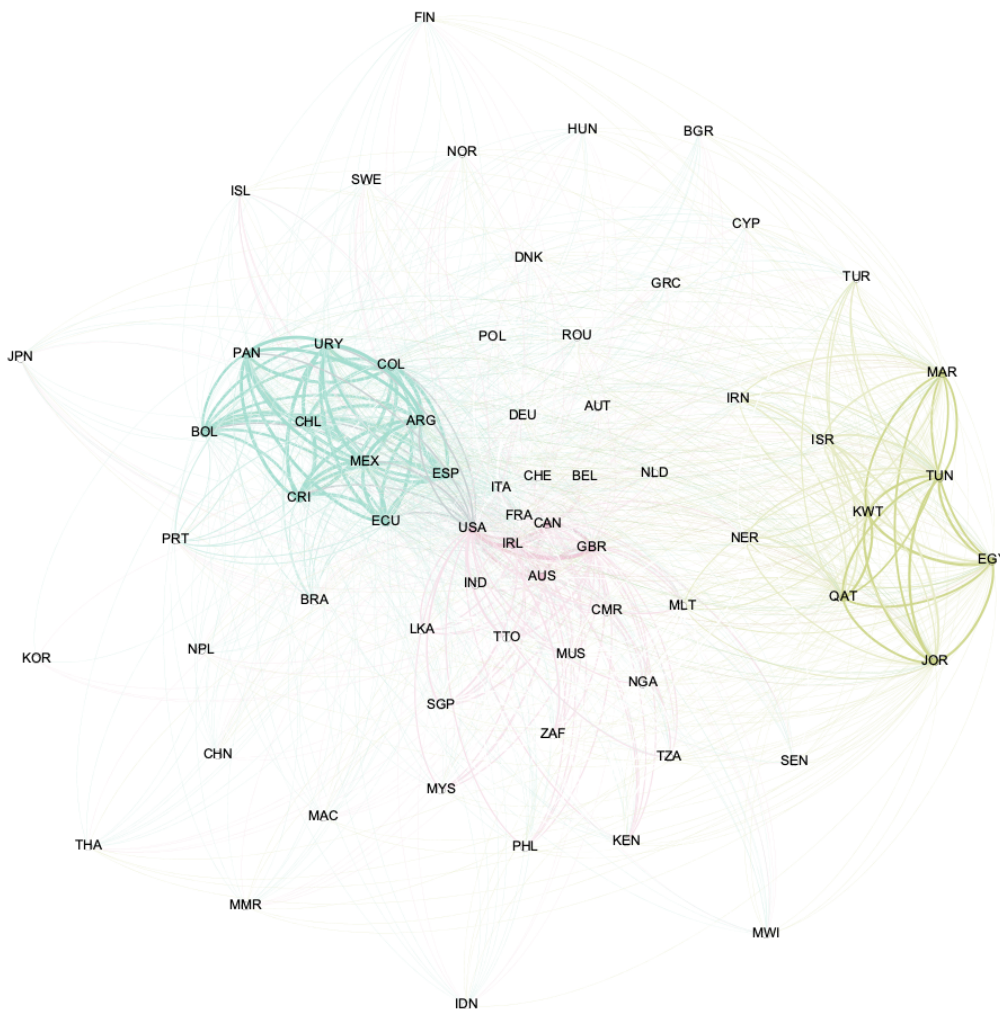
Notes: This figure depicts the ICL index values between 67 major trading countries (see table 2). The thickness of links between countries reflects their ICL index such that thicker links represented larger index values and greater linguistic ties. Similarly, the countries are positioned based on the magnitude of their ICL indices. The different colors represent various primary languages as define by Ginsburgh et al. (2014).

Figure 2 – Domestic Common Language in the World

0.49   0.62   0.74   0.87   0.99
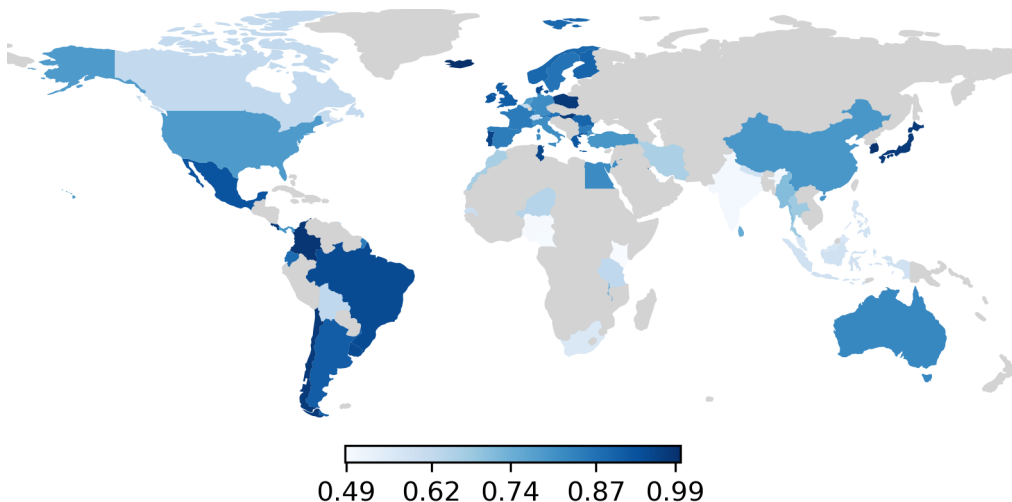
Notes: This figure depicts the DCL index values for 67 major trading countries (see table 2). Countries are shaded based on their DCL such that darker colors reflect larger DCL values and less domestic language diversity.

Figure 3 – Domestic Common Language within Canada

0.00   0.25   0.50   0.75   1.00
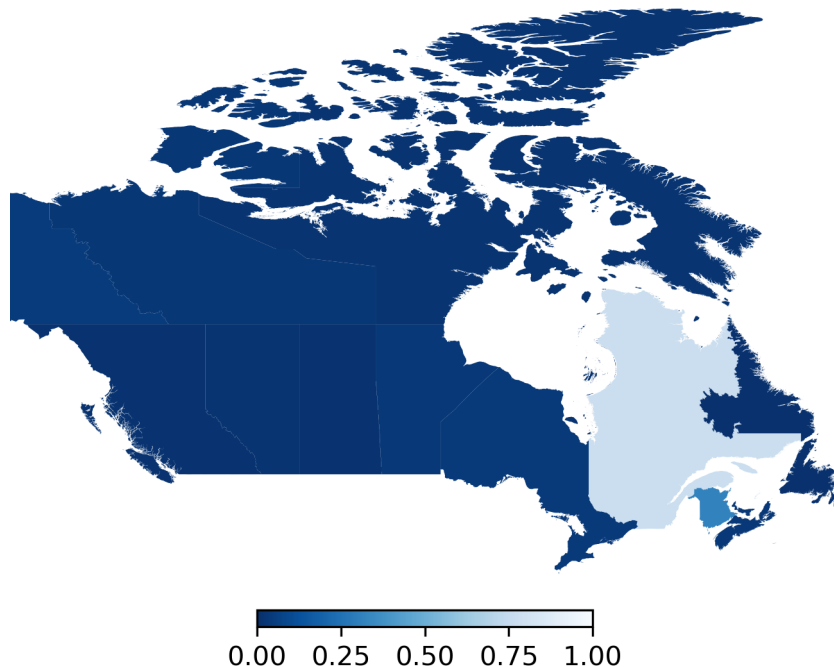
Notes: This figure depicts the proportion of the population in each of Canada's provinces and territories for which French is considered their "mother tongue".

Table 1 – Summary Statistics for Language Data

	Languages per country		Speakers per language	
	Native	Total	Native	Total
Mean	41	47	647,601	958,228
Median	14	17	7,320	10,000
Standard Deviation	90	99	10,589,815	13,885,341
Minimum	1	1	0	0
Maximum	837	853	896,000,000	1,074,000,000

Notes: The source of the data is Simons and Fennig (2018).

Table 2 – Included Countries

Argentina, Australia, Austria, Bulgaria, Belgium-Luxembourg, Bolivia, Brazil, Cameroon, Canada, Chile, China, Colombia, Costa Rica, Cyprus, Denmark, Ecuador, Egypt, Finland, France, Germany, Greece, Hong Kong, Hungary, India, Indonesia, Iceland, Iran, Ireland, Israel, Italy, Japan, Jordan, Kenya, Kuwait, Malawi, Malta, Mauritius, Mexico, Morocco, Myanmar, Malaysia, Nepal, Netherlands, Niger, Nigeria, Norway, Panama, Philippines, Poland, Portugal, Qatar, Romania, Senegal, Singapore, South Africa, South Korea, Spain, Sri Lanka, Sweden, Switzerland, Tanzania, Thailand, Trinidad & Tobago, Tunisia, Turkey, United Kingdom, United States, Uruguay.

Notes: This table lists the countries that are included in our estimating sample. The source of the data is Baier et al. (2019).

Table 3 – Common Language indices and International Trade

	Label	(1)	(2)	(3)
Distance (log)	$DIST_{ij}$	-0.744	-0.721	-0.721
		(0.037)**	(0.036)**	(0.036)**
Contiguity	$CNTG_{ij}$	0.389	0.360	0.360
		(0.065)**	(0.063)**	(0.063)**
Ex colonizer/colony	$CLNY_{ij}$	0.689	0.669	0.669
		(0.089)**	(0.096)**	(0.095)**
Economic integration agreements	$EIA_{ij,t}$	0.710	0.714	0.714
		(0.061)**	(0.061)**	(0.061)**
GATT/WTO membership	$WTO_{ij,t}$	-0.009	-0.009	-0.009
		(0.209)	(0.212)	(0.212)
European Union membership	$EU_{ij,t}$	0.518	0.550	0.549
		(0.089)**	(0.086)**	(0.087)**
ICL index (Mayer & Zignago, 2011)	ICL_{ij}	0.146		-0.002
		(0.059)*		(0.081)
ICL Index (Gurevich *et al.*, 2020)	ICL_{ij}		0.386	0.388
			(0.099)**	(0.137)**
Exporter × Year FE		Yes	Yes	Yes
Importer × Year FE		Yes	Yes	Yes
Estimator		PPML	PPML	PPML
Sample		International trade flows only		
Obs.		89508	89508	89508

Notes: This table reports the results from a series of specifications that focus on the impact of language on international trade. All estimates are obtained with the PPML estimator and the dependent variable is always nominal bilateral trade (in levels). All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. The differences between the columns are as follows: Column (1) uses the standard indicator variable for common official language of Mayer and Zignago (2011); Column (3) uses our newly constructed common language index; Column (4) includes our new language index together with the indicator for common official language. Standard errors are clustered by country-pair and are reported in parentheses. $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 4 – Common Language and International Trade, Robustness

	Label	(1)	(2)	(3)	(4)	(5)	(6)
Distance (log)	$DIST_{ij}$	-0.721	-0.722	-0.725	-0.731	-0.733	-1.291
		(0.036)**	(0.036)**	(0.036)**	(0.036)**	(0.036)**	(0.036)**
Contiguity	$CNTG_{ij}$	0.360	0.359	0.362	0.356	0.353	0.069
		(0.063)**	(0.063)**	(0.063)**	(0.063)**	(0.062)**	(0.138)
Ex colonizer/colony	$CLNY_{ij}$	0.669	0.668	0.670	0.664	0.667	0.895
		(0.096)**	(0.095)**	(0.096)**	(0.095)**	(0.101)**	(0.137)**
Economic integration agreements	$EIA_{ij,t}$	0.714	0.713	0.665	0.607	0.906	0.140
		(0.061)**	(0.061)**	(0.063)**	(0.064)**	(0.066)**	(0.062)*
GATT/WTO membership	$WTO_{ij,t}$	-0.009	-0.010	-0.022	0.083	0.178	0.425
		(0.212)	(0.219)	(0.208)	(0.205)	(0.228)	(0.210)*
European Union membership	$EU_{ij,t}$	0.550	0.551	0.542	0.519	0.506	-0.580
		(0.086)**	(0.086)**	(0.086)**	(0.087)**	(0.086)**	(0.081)**
ICL index	ICL_{ij}	0.386	0.388	0.379	0.361	0.417	1.232
		(0.099)**	(0.099)**	(0.100)**	(0.102)**	(0.095)**	(0.116)**
Exporter × Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Importer × Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Estimator		PPML	PPML	PPML	PPML	PPML	OLS
Sample		International trade flows only					
		Full	Positive trade flows	2-year interval	3-year interval	5-year interval	Positive trade flows
Obs.		89508	82184	47092	18816	18816	82184

Notes: This table reports the results from a series of robustness experiments designed to evaluate the impact of language on international trade. The estimates in columns (1) through (5) are obtained with the PPML estimator and the dependent variable is always nominal bilateral trade (in levels). All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. Column (1) reproduces our main estimates from column (3) of Table 3. Column (2) uses only positive trade flows. Columns (3), (4), and (5) employ data with 2-, 3-, and 5-year intervals, respectively. Finally, column (6) estimates the model with the OLS estimator and the logarithm of bilateral trade as the dependent variable. Standard errors are clustered by country-pair and are reported in parentheses. $^+$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 5 – Common Language, International and Domestic Trade

	Label	(1)	(2)	(3)	(4)
Distance (log)	$DIST_{ij}$	-0.721	-0.538	-0.547	-0.547
		(0.036)**	(0.050)**	(0.052)**	(0.052)**
Contiguity	$CNTG_{ij}$	0.360	0.541	0.500	0.498
		(0.063)**	(0.082)**	(0.086)**	(0.085)**
Ex colonizer/colony	$CLNY_{ij}$	0.669	0.555	0.563	0.567
		(0.096)**	(0.163)**	(0.158)**	(0.158)**
Economic integration agreements	$EIA_{ij,t}$	0.714	0.625	0.634	0.640
		(0.061)**	(0.096)**	(0.103)**	(0.104)**
GATT/WTO membership	$WTO_{ij,t}$	-0.009	0.440	0.451	0.449
		(0.212)	(0.135)**	(0.142)**	(0.151)**
European Union membership	$EU_{ij,t}$	0.550	0.460	0.487	0.493
		(0.086)**	(0.093)**	(0.098)**	(0.097)**
ICL index	ICL_{ij}	0.386	0.490	0.413	0.423
		(0.099)**	(0.143)**	(0.145)**	(0.145)**
Home bias	$HOME_{ij,t}$		9.659	8.849	
			(0.533)**	(0.659)**	
Home bias ×					
– Distance (log)	$DIST_{ij}$		-0.459	-0.478	-0.481
			(0.059)**	(0.056)**	(0.056)**
– Per capita GDP	$GDPPC$		-1.657	-1.658	-1.686
			(0.099)**	(0.097)**	(0.106)**
– Institution	$INST$		-0.514	-0.505	-0.524
			(0.055)**	(0.054)**	(0.054)**
– DCL index	DCL_{ij}			1.020	0.996
				(0.383)**	(0.382)**
Exporter × Year FE		Yes	Yes	Yes	Yes
Importer × Year FE		Yes	Yes	Yes	Yes
Estimator		PPML	PPML	PPML	PPML
Sample		International and domestic trade flows			
Obs.		90821	90821	90821	90821

Notes: This table reports the results from a series of specifications designed to evaluate the impact of language on domestic trade. All estimates are obtained with the PPML estimator and the dependent variable is always nominal bilateral trade (in levels). All results are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. Column (1) reproduces our main estimates from column (3) of Table 3. Column (2) introduces a series of proxies for domestic trade costs. Columns (3) adds the key variable for domestic language. Finally, column (4) allows the home bias effects to vary over time. Standard errors are clustered by country-pair and are reported in parentheses. $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 6 – Common Language, International and Domestic Trade, Robustness

	Label	(1)	(2)	(3)	(4)	(5)	(6)
Distance (log)	$DIST_{ij}$	-0.547	-0.547	-0.554	-0.559	-0.536	-1.255
		(0.052)**	(0.051)**	(0.050)**	(0.051)**	(0.058)**	(0.038)**
Contiguity	$CNTG_{ij}$	0.498	0.499	0.493	0.483	0.537	0.125
		(0.085)**	(0.085)**	(0.084)**	(0.084)**	(0.090)**	(0.138)
Ex colonizer/colony	$CLNY_{ij}$	0.567	0.564	0.583	0.586	0.574	0.910
		(0.158)**	(0.158)**	(0.155)**	(0.147)**	(0.164)**	(0.137)**
Economic integration agreements	$EIA_{ij,t}$	0.640	0.639	0.638	0.610	0.612	0.154
		(0.104)**	(0.104)**	(0.100)**	(0.095)**	(0.120)**	(0.063)*
GATT/WTO membership	$WTO_{ij,t}$	0.449	0.441	0.399	0.345	0.259	0.295
		(0.151)**	(0.149)**	(0.157)*	(0.167)*	(0.140)$^{+}$	(0.214)
European Union membership	$EU_{ij,t}$	0.493	0.493	0.507	0.490	0.443	-0.508
		(0.097)**	(0.097)**	(0.094)**	(0.096)**	(0.107)**	(0.081)**
ICL index	ICL_{ij}	0.423	0.428	0.409	0.394	0.405	1.271
		(0.145)**	(0.145)**	(0.143)**	(0.149)**	(0.158)*	(0.117)**
Home bias ×							
– Distance (log)	$DIST_{ij}$	-0.481	-0.479	-0.498	-0.521	-0.450	-0.871
		(0.056)**	(0.055)**	(0.054)**	(0.055)**	(0.064)**	(0.211)**
– Per capita GDP	$GDPPC$	-1.686	-1.687	-1.679	-1.569	-1.544	-1.882
		(0.106)**	(0.106)**	(0.106)**	(0.100)**	(0.107)**	(0.229)**
– Institution	$INST$	-0.524	-0.529	-0.521	-0.458	-0.409	0.111
		(0.054)**	(0.054)**	(0.054)**	(0.049)**	(0.057)**	(0.153)
– DCL Index	DCL_{ij}	0.996	0.996	0.864	0.767	1.203	4.281
		(0.382)**	(0.381)**	(0.374)*	(0.371)*	(0.421)**	(1.809)*
Exporter × Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Importer × Year FE		Yes	Yes	Yes	Yes	Yes	Yes
Estimator		PPML	PPML	PPML	PPML	PPML	OLS
Sample			International and domestic trade flows				
		Full	Positive trade flows	2-year interval	3-year interval	5-year interval	Positive trade flows
Obs.		90821	83497	47789	19047	19047	83497

Notes: This table reports the results from a series of robustness experiments designed to evaluate the impact of language on domestic trade. The estimates in columns (1) through (5) are obtained with the PPML estimator and the dependent variable is always nominal bilateral trade (in levels). All results are obtained with exporter-time importer-time fixed effects as well as time-varying home bias effects, whose estimates are omitted for brevity. Column (1) reproduces our main estimates from column (4) of Table 5 Column (2) uses only positive trade flows. Columns (3), (4), and (5) employ data with 2-, 3-, and 5-year intervals, respectively. Finally, column (6) estimates the model with the OLS estimator and the logarithm of bilateral trade as the dependent variable. Standard errors are clustered by country-pair and are reported in parentheses. $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table 7 – Language, Trade, and Welfare

Country	A. Total Exports		B. Welfare: Real GDP	
	ΔDCL	ΔCL	ΔDCL	ΔCL
	(1)	(2)	(3)	(4)
Canada	**-2.609**	**-1.524**	**0.939**	**1.220**
United States	-0.671	-0.187	-0.017	-0.005
Trinidad and Tobago	-0.045	-0.029	-0.011	-0.008
Qatar	-0.010	-0.015	-0.010	-0.014
Cameroon	-0.025	-0.038	-0.009	-0.014
Kuwait	-0.011	-0.018	-0.009	-0.014
Iceland	-0.014	-0.008	-0.009	-0.006
Panama	-0.074	-0.097	-0.008	-0.010
Mexico	-0.036	-0.054	-0.006	-0.009
Tunisia	-0.018	-0.027	-0.006	-0.009
Bolivia	-0.046	-0.062	-0.006	-0.008
Ecuador	-0.061	-0.078	-0.006	-0.008
Iran	-0.024	-0.029	-0.006	-0.007
Norway	-0.006	-0.004	-0.006	-0.005
Ireland	-0.011	0.006	-0.005	0.001
Mauritius	-0.037	-0.054	-0.005	-0.007
Niger	-0.059	-0.085	-0.005	-0.007
Chile	-0.058	-0.067	-0.005	-0.006
Cyprus	-0.013	-0.013	-0.005	-0.005
United Kingdom	-0.060	0.004	-0.005	-0.000
Egypt	-0.026	-0.040	-0.004	-0.006
Morocco	-0.036	-0.054	-0.004	-0.006
Senegal	-0.078	-0.113	-0.004	-0.006
Argentina	-0.033	-0.041	-0.004	-0.005
Colombia	-0.058	-0.073	-0.004	-0.005
Costa Rica	-0.053	-0.067	-0.004	-0.005
Finland	-0.014	-0.021	-0.004	-0.005
Portugal	-0.019	-0.024	-0.004	-0.005
Uruguay	-0.029	-0.036	-0.004	-0.005
Greece	-0.016	-0.015	-0.004	-0.004
Switzerland	-0.016	-0.011	-0.004	-0.003
Australia	-0.050	0.007	-0.004	-0.000
Malta	-0.021	-0.031	-0.003	-0.005
Tanzania	-0.061	-0.090	-0.003	-0.005
Bulgaria	-0.011	-0.012	-0.003	-0.004
Israel	-0.022	-0.030	-0.003	-0.004
Jordan	-0.025	-0.039	-0.003	-0.004
Malaysia	-0.013	-0.021	-0.003	-0.004
Sri Lanka	-0.032	-0.038	-0.003	-0.004
Denmark	-0.009	-0.007	-0.003	-0.003
Nepal	-0.027	-0.031	-0.003	-0.003
Sweden	-0.012	-0.008	-0.003	-0.003
Austria	-0.007	-0.002	-0.003	-0.002
Turkey	-0.029	-0.042	-0.002	-0.004
France	-0.017	-0.021	-0.002	-0.003
Hungary	-0.011	-0.018	-0.002	-0.003
Kenya	-0.063	-0.094	-0.002	-0.003
Malawi	-0.044	-0.066	-0.002	-0.003
Nigeria	-0.055	-0.082	-0.002	-0.003
Spain	-0.018	-0.023	-0.002	-0.003
Thailand	-0.014	-0.021	-0.002	-0.003
Italy	-0.016	-0.017	-0.002	-0.002
Romania	-0.016	-0.016	-0.002	-0.002
Singapore	-0.006	-0.005	-0.002	-0.002
Belgium	-0.011	-0.005	-0.002	-0.001
Netherlands	-0.008	-0.000	-0.002	-0.001
Brazil	-0.038	-0.045	-0.001	-0.002
Burma (Myanmar)	-0.014	-0.021	-0.001	-0.002
Japan	-0.022	-0.033	-0.001	-0.002
Philippines	-0.009	-0.019	-0.001	-0.002
Poland	-0.014	-0.015	-0.001	-0.002
South Korea	-0.017	-0.026	-0.001	-0.002
China	-0.014	-0.022	-0.001	-0.001
Germany	-0.008	-0.001	-0.001	-0.001
India	-0.032	-0.037	-0.001	-0.001
Indonesia	-0.016	-0.025	-0.001	-0.001
South Africa	-0.044	-0.05	-0.00	-0.003

Notes: This table reports the results from a counterfactual experiment that simulates a hypothetical scenario where all people in Quebec speak both English and French. The first two columns report effects on trade, while the last two columns report effects on welfare/real GDP. The results in columns (1) and (3) are obtained in response to only changing domestic component of the common language index, while the results in columns (2) and (4) are obtained in response to a simultaneous change in the domestic and the international language components for Canada. See text for specific details on the design of the experiment and a discussion of our findings.

Appendix

Figure A.1 – Change in total export resulting from the change in the DCL component of the CL index
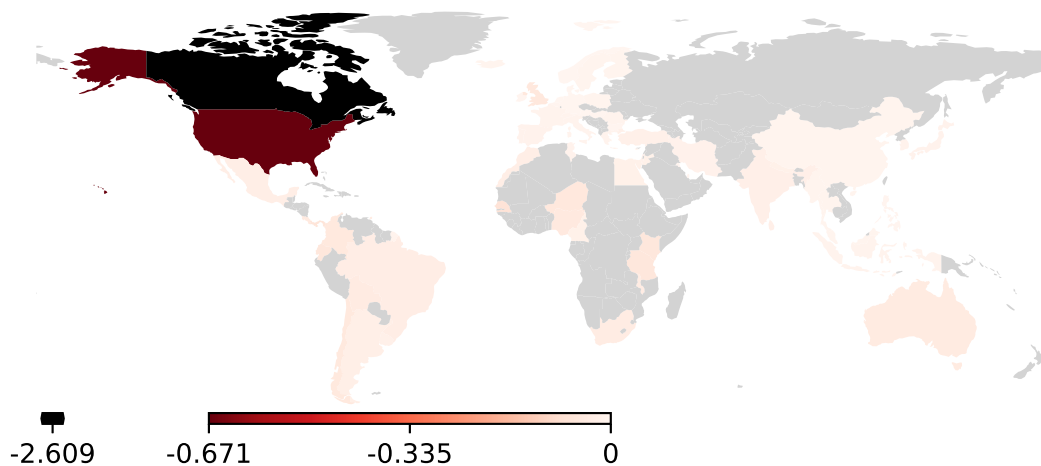

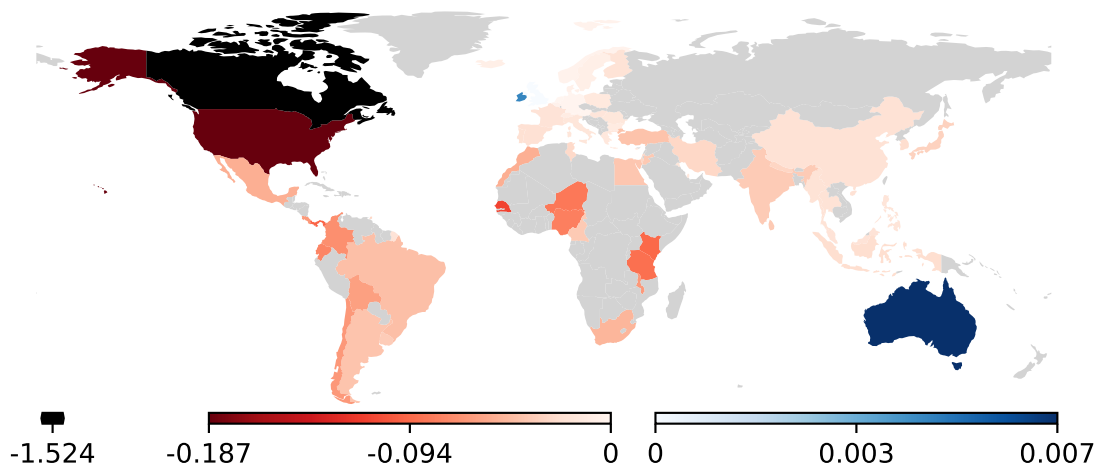**Figure A.2 – Change in total export resulting from the change in the CL index**

Figure A.3 – Change in real GDP resulting from the change in the DCL component of the CL index
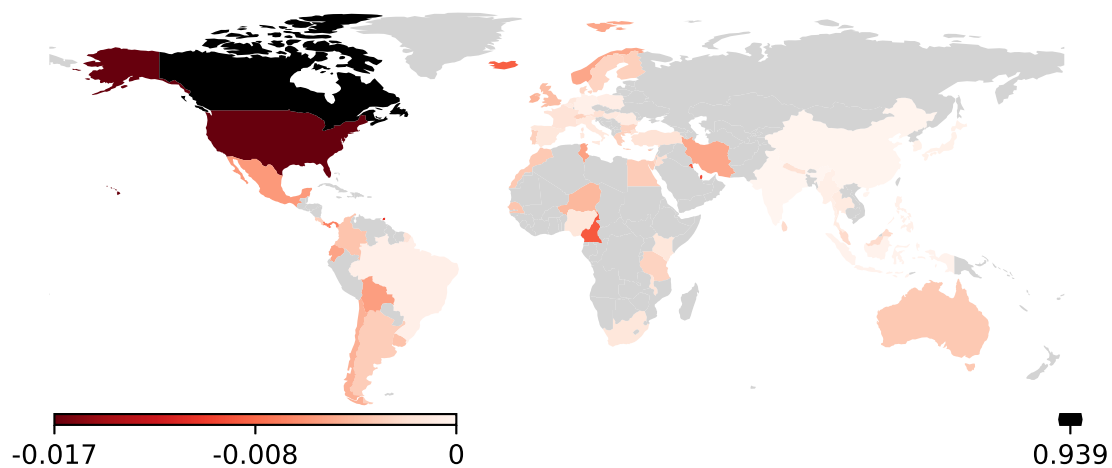

Figure A.4 – Change in real GDP resulting from the change in the CL index.