Do Immigrants Promote Exports to Countries Other than Their Country of Origin? On the Role of Geographic and Linguistic Proximity

Oleg Firsin

November 2015

Abstract

Immigrants have been found to promote exports to their countries of origin. In this study, we show that immigrants can also increase exports to other countries–specifically, countries that we term "proximate" to the country of origin. The proximity measure for a country pair is based on a geographical factor of sharing a common border and a cultural/linguistic factor of the probability of sharing a common native language; these factors reflect business networks, foreign market information and communication facilitation as channels of trade facilitation. The expected number of non-proximate immigrants is termed "distant" and is found to have a negative effect on exports to a given country. The trade effects of different immigrant groups are more pronounced among state-country pairs with a larger number of immigrants from the export destination country. We obtain state-country level estimates of immigration-export elasticity in the United States that are preferable to previously available estimates due to the use of proper trading-pair fixed effects and more recent data, covering 2003-2013. Analysis is done both at the aggregate state-country level and at the industry level, with the within-industry immigrant effect on industry exports being estimated for the first time for the case of the United States as the host country.

1 Introduction

Since the seminal work of Gould (1994), there has been an explosion of interest in studying the connection between immigration and trade. Previously, it had been analyzed through the Heckscher-Ohlin Model, which treated production factor trade and commodity trade as substitute processes (Mundell (1957)), implying movement in the opposite directions for migration and trade. This suggested reduction in trade could be a negative implication of easier immigration. Gould's work helped shift the analysis framework to that of immigrants as trade facilitators. Despite a large amount of additional evidence supporting the finding that immigration promotes trade, export promotion role of immigrants has not yet become a significant part of immigration or trade policy discussions. Our work highlights a heretofore unexplored direction of immigration-trade link–to countries geographically and linguistically linked to country of origin,–thereby accentuating potential importance of immigration in export promotion discussions and importance of trade-related effects in immigration policy considerations.

The two channels that are generally used to explain the trade facilitation effect are information and networks. The former refers to information about legal, institutional and cultural aspects of export markets; the latter refers to business relationships and contacts that immigrants may have in home countries. Additionally, for trade among countries with different predominant languages, immigrant language skills can improve communication and logistics, and potentially increase trust. Trade facilitation effect applies to both exports (the focus of this paper) and imports, while many researchers also examine home preference effect, which only applies to imports.

In virtually all empirical work to date, the impact of immigrants was assessed with regards to exports to or imports from their countries of origin. The rational behind the trade facilitation effect, however, means that anyone with a network in or relevant legal, institutional, cultural or other knowledge of a given country c, or in possession or rare language skills that facilitate communication, can have a pro-export effect to that country. Thus, if Ecuadorians can promote exports to Ecuador through facilitation in negotiations and logistics/communication due to being native speakers of Spanish, potentially so can Colombians. Moreover, a Colombian can have a business network in Ecuador and know important institutional peculiarities of operating there, given that the two countries share a border and are part of the Andean Community. If we view networks as being more likely to connect businesses that are close by than far apart, the fact that border areas of two countries are much closer to each other than to many parts of the respective own-country interiors suggests that immigrants may promote exports to countries bordering their country of origin. Furthermore, in countries with different official languages there may still be many native speakers of the same language (for example, Russian in the former USSR, Spanish in Portugal, etc.). We operationalize this using a combination of Melitz and Toubal's (2014) work on common native and spoken language probability and an indicator of common border (one could potentially also look at common economic zones) to construct the expected number immigrants, geographically and linguistically proximate to a given country c and analyze their impact on exports to country c. We find evidence that these proximate immigrants do stimulate exports and that it is driven through geographic, rather than linguistic connection; we fail to find the same evidence for non-proximate, or "distant," immigrants, who, conversely, exhibit trade-diversion effect.

Another contribution of the paper to immigration-trade literature concerns estimates of the standard immigration-export elasticity for the United States. The generally positive impact of immigration on both exports and imports has been reaffirmed in many studies with empirical analysis based on data from many immigrant host countries, including the United States. Methodology, however, has not usually been very demanding. In particular, we are not aware of any other paper that analyzes U.S. state-level aggregate exports to multiple foreign countries and uses trading-pair fixed effects. The only papers (that we are aware of) that do use trading pair fixed effects are Millimet and Osang (2007), studying U.S.-Canada state-state level trade, Coughlin and Wall (2011) and Bandhyopadhay et al. (2008).[1] None of these studies use an immigration-trade model comparable to no-fixed-effects models in other studies[2]. Meanwhile, trading-pair fixed effects are important to control for, since there may be idiosyncratic economic ties between particular states and foreign countries, and ignoring this possibility may attribute to immigrants an impact of an omitted variable or variables. We find that whereas state and country fixed effects produce results in line with the central tendency in the literature, trading-pair fixed effects specifications produce a positive and statistically significant effect of immigration on exports only for subsamples of state-country pairs with a large number of immigrants from country c (above a minimum threshold number of immigrants of around 1000, close to the median); even then, the estimate is smaller than what is usually found in the literature, suggesting that the findings of many of the previous papers were likely influenced by lack of proper fixed effects. Part of the difference in results may also be due to us using more recent data, 2003-2013, than any other study of the subject, as most studies focus on the period between 1990 and 2000, which might have differed from periods before and after. However, this would not explain the differences in results from different specifications based on our data. We conduct part of the analysis at the aggregate state level and part at the industry level. The level of analysis here refers to both the export measure and all the explanatory variables. This is the first study to conduct U.S.-based immigration-trade link analysis with explanatory variables also at the state-industry level. The combined analysis of aggregate-level and industry-level results suggests that export-promotion effect of immigrants is not confined to industries they are employed in.

The next section reviews some of the relevant literature. Section 3 describes data and presents descriptive statistics, while Section 4 outlines empirical methodology. Section 5 presents results, and Section 6 concludes.

2 Literature Review

Immigration-trade literature has been largely focused on uncovering immigrant characteristics, country characteristics and export good characteristics that affect the estimated elasticity of trade with respect to immigration. It is not unusual for different studies to conflict in terms of significance of certain factors or to encounter opposite estimated signs of the effects of certain characteristics examined, so we created Table A1 (Appendix) to summarize some of the findings (best judgment was used where results were contradictory). Factors specified in the first column are found to have a positive, negative or no effect on exports or immigration-trade elasticity. A different type of literature overview can be found in Genc et al. (2012) and Lin et al. (2011) meta-analyses.

One of the broad categories of factors considered is country differences. Country characteristics usually enter as separate explanatory variables and interaction terms with the number of immigrants or separate regressions are run for different groups of countries. In the former case, country characteristics have both a direct effect on exports (imports) and an indirect one through interaction. Same language (Rauch and Trindade (2002), Aleksinska and Peri (2014), Egger et al. (2012),

[1] Fixed effects used in Good (2012), studying U.S.-Mexico state-state trade, are not clear, but they cannot be trading-pair, given inclusion of time-invariant trading-pair level variables like distance

[2] Reasons explains in the literature review

Blanes (2008), same colonizer (Melitz and Toubal (2014)) or a former colonial relationship (Meitz and Toubal (2014), Aleksinska and Peri (2014), Blanes (2008)) generally increase trade between a country pair. Same language and religion, theoretically, are expected to decrease immigrants' trade promotion effect, while cultural distance is expected to increase it, since the former reduces relevance of immigrants' language skills and religious fluency and the latter increases relevance of their cultural knowledge; empirically, results are mixed (Melitz and Toubal (2014), Aleksynska and Peri (2014), Kandogan (2005), Tadesse and White (2010)). Weaker rule of law, higher levels of corruption and lower level of economic development should lead to higher immigration-trade elasticity, especially from a developed host country perspective, since immigrants bring knowledge of less formal ways of doing business in their home countries and networks that help overcome hurdles posed by weaker institutions. These predictions are generally confirmed empirically (White (2009*a*), Bratti et al. (2014), Hatzigeorgiou and Ladefalk (2014), Briant et al. (2014), Dunlevy (2006)). Similarly, high institutional similarity, previous colonial ties and common legal system reduce immigration-trade elasticity, while results for the same economic zone vary (Aleksynska and Peri (2014), Herander and Saveedra (2005), Blanes (2008), Hong and Santhaparraj (2006)).

The second major source of heterogeneity in immigration-trade elasticity is immigrant characteristics. Certain types of immigrants are more likely to have and take advantage of professional networks. In particular, several studies show that more educated/skilled immigrants and those in professional and managerial occupations have a greater pro-trade effect (Mundra (2014), Aleksynska and Peri (2014), Blanes (2008), Martin-Montaner et al. (2014)). Male immigrants are found to have a greater trade promotion effect according to Hatzigeorgiou and Ladefalk (2014), but the opposite is true in Bowen and Wu (2014). With regards to immigrant tenure length, Jansen and Piermartini (2009) show that temporary migrants have a greater effect on trade; relatedly and, to some extent, in contrast, Bratti et al. (2014) find that earlier waves of immigrants increase trade more. Location-wise, immigrants within the state/province of export origin matter more than those outside of it (Bratti et al. (2014), Herander and Saveedra (2005), Peri and Requena-Silvente (2010), Good (2012)). Finally, using Danish data, Hiller (2013) asks a question similar to that of this paper–whether there is a trade-diversion effect of immigrants from a given country from exports elsewhere. They look at all countries and countries with only the same or only different official language, but do not find any effect.

A number of studies also look at the type of export, and find immigrant pro-trade effect being greater for differentiated goods (Rauch and Trindade (2002), White (2009*b*), Melitz and Toubal (2014), Peri and Requena (2010)). The likely reason is that the relevance of immigrant networks and knowledge is more pronounced when there is more need for communication and negotiation. Other works have looked at cultural vs. non-cultural, producer vs. consumer goods and goods vs. services (Tadesse and White (2010), Bowen and Wu (2013), Blanes-Cristobal (2008)).

Two additional points of interest are the direction of trade (exports vs. imports) and the margins of trade (extensive vs. intensive). Theory would suggest that immigrants should affect the extensive margin more, since they are thought to affect fixed costs of learning about and established relationships with business in foreign markets, and they should matter more for imports, because there is an additional home-country product preference effect. Empirically, the effect on immigration-trade elasticity is ambiguous on both accounts, and it can be rationalized in both cases. Exports feature higher elasticity in Blanes and Cristobal (2008) and Briant et al. (2014), among others, but imports feature larger effect in Head and Ries (1998) and Wagner et al. (2002); overall, Genc et al. (2012) meta-analysis estimates export and import elasticities to be roughly equal. With respect to the margins of trade, Pennerstrofer (2014) uses data from Central European firms to try to distinguish between extensive and intensive margins, and finds the effect on the intensive margin to be much smaller and not always significant. Coughlin and Wall (2011) reach a similar conclusion studying U.S. state-level exports, whereas findings of Hiller (2013) and Peri and Requena-Silvente (2010) are the opposite. It is important to stress that what is meant by extensive and intensive margins differs between studies. Pennerstrofer (2014) and Hiller (2013) treat extensive margin as more exporting firms to a given destination, whereas Coughlin and Wall (2011) interpret extensive margin as new exports to a country of destination for an entire industry in a state and Peri and Requena-Silvente (2010) operationalize it as new province-country trade relations.

Methodologically, most studies employ various specifications of the gravity equation. In their meta-analysis, Genc et al. (2012) distinguish between studies that use OLS, the Heckman selection model, the Tobit model, the pseudo-Poisson model, IV/3SLS/GMM and other methods (such as FGLS). With regards to fixed-effects, there are some studies at a region-country level that account for trading-pair fixed effects, notably Bratti et al. (2014), Peri and Requena-Silvente (2010), and Requena et al. (2012).

With regards to estimated elasticity, Genc et al. (2012) meta-analysis finds that for the United States as the host country, they range between -0.14 to 0.57, with an unweighted mean of 0.16. The analysis also mentions two trends that stand out in terms of dynamics of the effect size estimated. The first is that the size of the marginal effect of immigrants on trade may be decreasing beyond a certain threshold number (Gould (1994), Egger et al. (2012)), although some authors that address this argument do not find support for it (Serrano-Domingo and Requena-Silvente (2013), Peri and Requena-Silvente (2010)). The second—as found in both Lin's (2011) and Genc et al. (2012) meta-analyses—is that studies using more recent data tend to find lower elasticities, although this also is not a uniform result (Peri and Requena-Silvente (2010)).

Since this paper treats U.S. as the host country, and the impact of immigration on trade may differ across countries (Genc et al. (2012) analysis finds that immigrant-export elasticity ranges from -0.03 for the two U.S.-Canada trade studies to 0.43 for Australia as the host country (15 estimates), with an average 0.17), we briefly summarize studies with the U.S. as the host country in Table 1. Studies in Table 1 range from those with exports from the U.S. as a whole to several foreign countries in one year, to panel studies from U.S. states to foreign countries in a given industry (Coughlin and Wall (2011)). Authors have long acknowledged the possibility of spurious correlation or reverse causality between immigration and trade, whereby observed and unobserved characteristics of trading-pairs could lead to more trade and immigration (e.g., lower observed and unobserved trade and migration costs) or immigrants come to countries with strong economic relations with the home country. It is, therefore, important to at least control for time-invariant trading-pair unobservables through trading pair fixed effects. Among studies with state-country trading pairs, however, only two do so–Bandyopadhyay et al. (2008) and Coughlin and Wall (2011)– and both deviate from the literature in other ways, not producing an estimate directly comparable to the other studies of the type. Bandyopadhyay et al. (2008) do include trading pair fixed effects but depart from the literature by including stock of all 29 export destination countries as separate variables and then impose equality of all coefficients (Bandyopadhyay et al. (2008); Table 3), a procedure that produces different results than what one obtains when estimating just one coefficient on the stock of immigrants, while Coughlin and Wall (2011) use state-country fixed effects, but all of their variables are at the state-country-industry level. The type of fixed effects used can be crucial for results, as a number of studies find that immigration-export elasticity reduces significantly or becomes insignificant with inclusion of fixed effects (Millimet and Osang (2007)).

In addition to time-invariant unobservables, time-variant ones may also bias the estimates, but

Table 1: Studies with the U.S. as the Host Country

Author	Data: # of Countries	Years	Export Elasticity	Method
Bandyopadyay et al. (2008)	29 (From U.S. states)	1990-2000	0.14	OLS with FE (country)
Bardhan and Guhathakurta (2004)	51 (from 17 U.S. states)	1994-1996	0-0.26	OLS with FE (region)
Co et al. (2004)	28 (from U.S. states)	1993	0.29	OLS
Coughlin and Wall (2011)	29 (48 states; 19 industries)	1990-2000	0.19	OLS with FE (state-county)
Dunlevy (2006)	87 (from U.S. states)	1990-92	(0.24-0.47) 0.39	OLS with FE (state, country)
Dunlevy and Hutchinson (1999)	17	1870-1910	0.08	OLS
Good (2013)	Mexican states from U.S. states	2008-2010	0.08	OLS with FE (state)
Gould (1994)	46	1970-1986	0.02*	NLIN LS
Herander and Saavedra (2005)	36	1993-1996	0.18	Tobit (region FE)
Jansen and Piermartini (2009)	>100	1996-2005	0.01-0.25	Tobit, OLS with FE (time)
Millimet and Osang (2007)	Canadian states from/to U.S. states	1993,1997	0.0	OLS with FE (state-state)
Mundra (2005)	47	1973-1980	Not estimated	Semi-par with FE and IV
Mundra (2014)	63	1991-2000	0.25	2SLS
Tadesse and White (2010)	75 (from U.S. states)	2000	0.05	Tobit
White (2009a)	28 (from 48 U.S. states)	1993	0-0.57	OLS
White (2009b)	70	1980-1997	0.0	OLS
This study	131 (50 States + DC; 29 industries)	2003-2013	0.06**	2SLS with FE (state-country)

*Wagner et al. (2002) calculation. **Preferred estimate.

only 2 studies employ instrumental variable strategy to address this possibility. We, therefore, find it important to both include fixed-effects and IV in a model that is otherwise comparable to previous studies. Finally, most of the studies use data on relatively few foreign countries and focus on a time period before 2000, which may not be representative of the subsequent years. Thus, the estimates of immigration-trade elasticity from this study are in certain ways preferable to those previously available for the case of the U.S. as the host country.

3 Data and Descriptive Statistics

The data for the study is taken from several sources, summarized in Table 2. Information on exports from 50 U.S. states and Washington, D.C., to about 200 foreign countries comes from USA Trade Online. This source only presents positive values of exports; rather than imputing exports based on missing values, we restrict analysis only to state-country-industry-year observations with positive export values. We restrict most of the analysis to 131 foreign countries for which there was enough data to estimate the basic model. State aggregate- and industry-level GDP[3] is taken from County Business Patterns of U.S. Census Bureau, while country GDP and population are from World Development Indicators of the World Bank[4]. All GDP and export data are expressed in 2014 U.S. dollars, using GDP deflator from St. Louis Federal Reserve Bank.

Table 2: Data Sources

Variable	Description	Source
Exports	Aggregate and industry (NAICS, 3 digit) exports from states to countries	USA Trade Online
State GDP	State aggregate and industry-level GDP	Country Business Patterns
State Population	State GDP by year	ACS PUMS (calculated)
Country GDP	Country GDP by year	WDI (World Bank)
Country Population	Country population by year	WDI (World Bank)
Immigrants (2003-2013)	Immigrant stock and shares by country of origin and state of residence	ACS PUMS (calculated)
Immigrants (1980)	Immigrant stock by country of origin and state of residence	IPUMS U of Minnesota
Language	Official language, share native and share spoken	Melitz and Toubal (2014)
Borders	Bordering countries	CIA World Factbook
Distance	Distance between largest state cities and foreign county capitals	Calculated [5]
GDP Deflator	Implicit Price Deflator	St. Louis Fed.

The main source of individual-level data is American Community Survey Public Use Microdata Sample (ACS PUMS). For years 2003-2013 we use ACS PUMS data from U.S. Census Bureau and for 1980, IPUMS by University of Minnesota Population Center (also based on ACS PUMS). As is common in the immigration-trade literature, we designate as immigrants anyone who was born abroad. It is worth noting that country of origin is identified as country of birth, rather than place of residence previously to arriving to the United States. We only consider adult population with positive income. The reason is that we are interested in immigrants who are actually employed; if we were estimating impact on imports, even immigrants who are not employed would be relevant.

[3]Some of the industries are grouped together in CBP GDP data, so we aggregate other data across the same industries; the industries that were aggregated are presented as a range in table 4 (ex., Agriculture and Livestock, NAICS 111-112).

[4]Argentina GDP and population estimates come from UNdata of the UN Statistics Division

[5]$D_{12} = acos(sin\phi_1 * sin\phi_2 + cos\phi_1 * cos\phi_2 * cos(\alpha_2 - \alpha_1)) * 6371$, where ϕ=radians(90-latitude) and α=longitude.

3.1 Immigrant Proximity

We estimate the *expected* number of proximate immigrants based on geography and language. For geographic proximity we use a simple rule and assign the value of 1 to a pair of countries that share a border and 0 to those that do not, and denote the value as B. It is certainly the case that there may be some country X that does not border country Y, but is closer to its economic centre than a country Z that does border Y, especially in the case of large countries, like Russia or China. Most of the time, however, a common border is a good measure of proximity to economically important areas, and furthermore, immigrants from bordering countries may have networks close the border area, so this simple rule has intuitive appeal.

Language proximity measure is more nuanced. It is an intuitive argument that common spoken language is vital for communication associated with international trade transactions, be it marketing, logistics or more informal communication. Yet language usually enters the estimating gravity equation as a dummy for common official language between trading countries and interacted with immigrant stock variable. It is not accurate to assume, however, that people from countries with the same official language will be able to communicate or that people from countries with different official languages will not be able to communicate. Realizing this, Melitz and Toubal (2014) further network specificity literature by showing that not only common official language, but also all common native and spoken languages, spoken by a substantial portion of the population in each country, matter for trade. Their results indicate that all relevant languages together have double the effect of just the official language and that native language is especially important, since in addition to basic ability to communicate it allows more nuanced communication and potentially ensures more trust. Hence, common native language is what we use to operationalize language proximity in this study.

To construct a measure of proximity via language we use the data made available through Melitz and Toubal (2014) study. For a large number of countries, they compiled data on the official language, the share of people who report each language as native (with a 5% threshold) and the same measure for a spoken language. To create a value for linguistic proximity, we follow their methodology to create a value for common native language (CNL) for each country, a measure that should be highly correlated with the true (unknown) probability that any two randomly taken people from two countries would share the same native language.[6] Common native language score for a pair of countries is calculated as

$$\text{CNL}_{cj} = \max_k(L_{kc}L_{kj}) + (\alpha - \max_k(L_{kc}L_{kj}))(1 - \max_k(L_{kc}L_{kj})), \tag{1}$$

where L_{kc} is the share of people in country c that speak language k, $\alpha = \sum_{k=1}^{n} L_{kc}L_{kj}$ is the sumproduct of shares of people who report each common language between two countries as native, and $max(L_{kc}L_{kj})$ is the maximum product of the two shares. For example, if in country 1, 90 percent of people report French as a native tongue, 50 percent report German, and 0 percent name Spanish, while in country 2, 80 percent report French, 90 percent report German, and 10 percent report Spanish, then $\alpha = 0.72 + 0.45 + 0 = 1.17$ and $CNL = 0.72 + (1.17 - 0.72)(1 - 0.72) = 0.09 + 0.05 * 0.91 = 0.846$; in practice, this ensures CNL is always between 0 and 1. *Importantly, we do not include English as one of the potential common languages, since it is not expected to be an*

[6]We also experimented with common spoken language, CSL, and common official language, but the results where not substantially different.

advantage exporting from the United States, where almost every native worker is expected to speak English.

Since we want the proximity index to reflect both geographic and linguistic proximity, we need a rule to combine CNL and B. For a given level of CNL, bordering countries should lead to more extensive business networks and greater shared knowledge of and familiarity with the other country's legal and cultural idiosyncrasies; for analogous reasons, for a given bordering country, a higher level of linguistic proximity should lead to a greater proximity index. There are many potential rules that satisfy this criterion. We use a simple average of both measures:

$$PI_{cj} = (CNL_{cj} + B_{cj})/2, \tag{2}$$

meaning proximity index (PI) is equal to the unweighted mean of the linguistic proximity index (CNL) and the binary value of common border (B). If instead of the unweighted mean, the maximum or the minimum of the two values were used, bordering countries with different CNL would be treated the same. While we do not do much experimentation with the relative weights, we do ultimately test the importance of each factor separately. The estimated number of proximate immigrants for any country c in state s in year t is

$$I_{sp^ct} = \sum_{j \in C; j \neq c} I_{sjt} * PI_{cj}, \tag{3}$$

where the summation is over all countries j other than c, the country for which the measure is constructed. If only CNL or B was used, the interpretation would be the expected number of people who speak the same native tongue or are from a bordering country; with the current form of equation (3), the interpretation is the expected number of proximate immigrants based on the proximity index. Additionally, we construct a measure of distant immigrants, equal to all immigrants not from country c less the proximate ones.

3.2 Descriptive Statistics

Table 3: Summary Statisics (Monetary values in millions of 2014 U.S. $)

	Obs	Mean	SD	Min	Max
Exports	46,021	225	1,161	.00258	67,676
State GDP	46,021	392,444	418,907	26,551	2237921
State Population	46,021	7426170	7369286	487,341	3.83e+07
Country GDP	46,021	746,093	1547025	202	1.64e+07
Country Population	46,021	6.33e+07	1.92e+08	49,139	1.36e+09
Distance	45,525	8,386	3,600	216	19,243
Own Immigrant Stock	46,021	6,936	61,957	2	3403282
Proximate Immigrant Stock	46,021	32,529	148,558	0	3145843
Distance Immigrant Stock	46,021	789,411	1428542	7,359	8417627

Table 3 presents main descriptive statistics at the aggregate (across all included industries) level. Mean exports from a state to a country totaled 225 million 2014 U.S. dollars. There is quite a bit of variation, with the range of $2.6 thousand to $68 billion (Texas to Mexico). The average state GDP of the sample is $400,000 billion, with California recording the maximum of 2.3 trillion. Average state population in the sample is 7.4 million about 9 times smaller than the average population of a country in the sample. On the other hand, average country GDP, at 750 billion, is only two times larger than average state GDP. The mean number of immigrants from the country of export destination was 7 thousand, ranging from 2 to 3.4 million. Proximate immigrant stock averaged 32.5 thousand and that of distant immigrants averaged 789 thousand.

Table 4 breaks down descriptive statistics by industry (sorted by immigrant employment share). It only includes exports between trading pairs analyzed, which account for most, but not all of U.S. manufacturing exports. The largest export value is attributable to Computer Electronic Products, followed by Transportation Equipment, Chemicals and Machinery. These industries are also among the most export intensive, with up to 0.38 export to GDP ratio, and employ the most people; Agriculture and Food Beverages also employ large numbers of people, but do not account for as much export. Apparel and Leather is by far the most immigrant-intensive sector, followed by Agriculture, Computer Electronic Products and Food Beverages. Overall, it appears that industries that are more export focused do not necessarily employ more immigrants.

We next look at unconditional relationships between export share going to a given country c and 1) the share of all immigrants in the state from that country or 2), the ratio of proximate to distant immigrants to those of country c. We present relationships for the two states with the highest share of foreign-born, California and New York, for the most recent year, 2013 (Figure 1). Panels A and B show the relationship in levels[7]. The slope of the regression line is 0.42 for country c immigrant share, but it is 0 for proximate/distant ratio. Since almost all countries are close to 0 on both axis, we take the log of both variables to see if this clarifies the relationship.

Panel C shows that now the relationship between export share and country share is 0.92, or almost 1-to-1. Panel D presents a much clearer picture than the one we get from B. The regression slope including all the observations is 0.14. However, there is a number of countries on the left of the graph–most of them do not have any proximate immigrants in any of the years and any of the states. These are island countries that do not border others, such as Japan, Philippines, South Korea (not quite an island, but it only borders North Korea, which we do not include in the analysis), Iceland, Trinidad & Tobago, and Australia, which mostly speak English (which, again, we excluded from the list of relevant native languages) or languages that do not have native speakers elsewhere, such as Japanese, Icelandic, and Tagalog. Thus, they cannot possibly have any proximate immigrants in the export origin states and the log of proximate/distant ratio is zero, since

$$log(proximate/distant) = log(proximate + 1(if proximate = 0)) - log(distant + 1(if distant = 0)) \tag{4}$$

and one is added to zero values to take the log. If we do not consider the countries that cannot have proximate immigrants, we obtain the middle regression fit line, with the slope of 0.31. If we, furthermore, only consider countries with positive proximate immigrant value for the graph, we obtain the steepest line, with the slope of 0.37. Panels E and F present the same relationships for New York. Now the slope for country c share is 0.53, whereas that for proximate/distant ratio without countries that do not have any proximate immigrants is 0.23. Naturally, if we look look at the correlation with distant/immigrant ratio, the sign will be the opposite. This shows that at least correlation between exports and proximate immigrant share is in the expected direction.

[7]Mexico is not shown in panel 1, as it is further to the right, at 0.3

Table 4: Descriptive Statistics by Industry

Industry Name	NAICS	Obs	Exports ('13)	Ind. GDP	Exp/GDP	Ind. Pop.	% Imm.
Apparel and Leather	315-316	240	2,976.97	10,339.83	0.29	309,812.81	0.48
Agriculture and Livestock	111-112	322	21,831.10	195,120.77	0.11	1,615,061.21	0.31
Computer & Electronic Products	334	742	90,042.76	257,958.34	0.35	1,493,321.00	0.27
Food & Beverages	311-312	662	29,259.85	238,765.08	0.12	2,049,169.19	0.26
Textiles and Textile Products	313-314	168	3,430.57	16,235.68	0.21	199,727.57	0.22
Misc. Manufacturing Comm.	339	680	31,769.66	83,520.28	0.38	1,486,805.00	0.21
Furniture & Fixtures	337	216	1,382.75	24,406.35	0.06	495,233.00	0.19
Chemicals	325	668	63,027.59	350,023.18	0.18	1,504,323.00	0.18
Oil & Gas	211	22	1,998.92	204,415.14	0.01	61,724.00	0.18
Petroleum & Coal Products	324	103	14,524.21	145,503.39	0.10	197,149.00	0.18
Nonmetallic Mineral Products	327	210	2,938.37	38,940.23	0.08	486,451.00	0.17
Electrical Equipment	335	330	14,065.42	50,709.58	0.28	519,901.00	0.16
Printing and Related	323	280	1,844.92	35,881.06	0.05	667,374.00	0.15
Wood Products	321	164	2,500.56	24,690.83	0.10	470,987.00	0.15
Fabricated Metal Products	332	499	18,003.60	141,398.75	0.13	1,501,888.00	0.14
Plastics & Rubber Products	326	278	10,756.53	70,985.89	0.15	609,530.00	0.13
Newspapers, Books and Other	511	246	288.85	201,009.50	0.00	770,968.00	0.13
Transportation Equipment	336	711	73,901.57	264,831.58	0.28	2,661,129.00	0.13
Machinery, Except Electrical	333	615	50,406.94	145,748.25	0.35	1,547,993.00	0.13
Paper	322	194	6,344.05	50,485.04	0.13	452,262.00	0.12
Forestry, Fishing and Other	113-115	64	366.29	29,201.87	0.01	147,812.58	0.12
Primary Metal Mfg	331	206	13,894.73	62,889.38	0.22	626,996.00	0.11
Minerals & Ores	212	59	1,129.21	62,152.78	0.02	222,563.00	0.08
Total/Unweighted Average		7679	456,685.41	2,705,212.78	0.17	20,098,180.36	0.18

All calculations are for the largest industry-level sample analyzed.

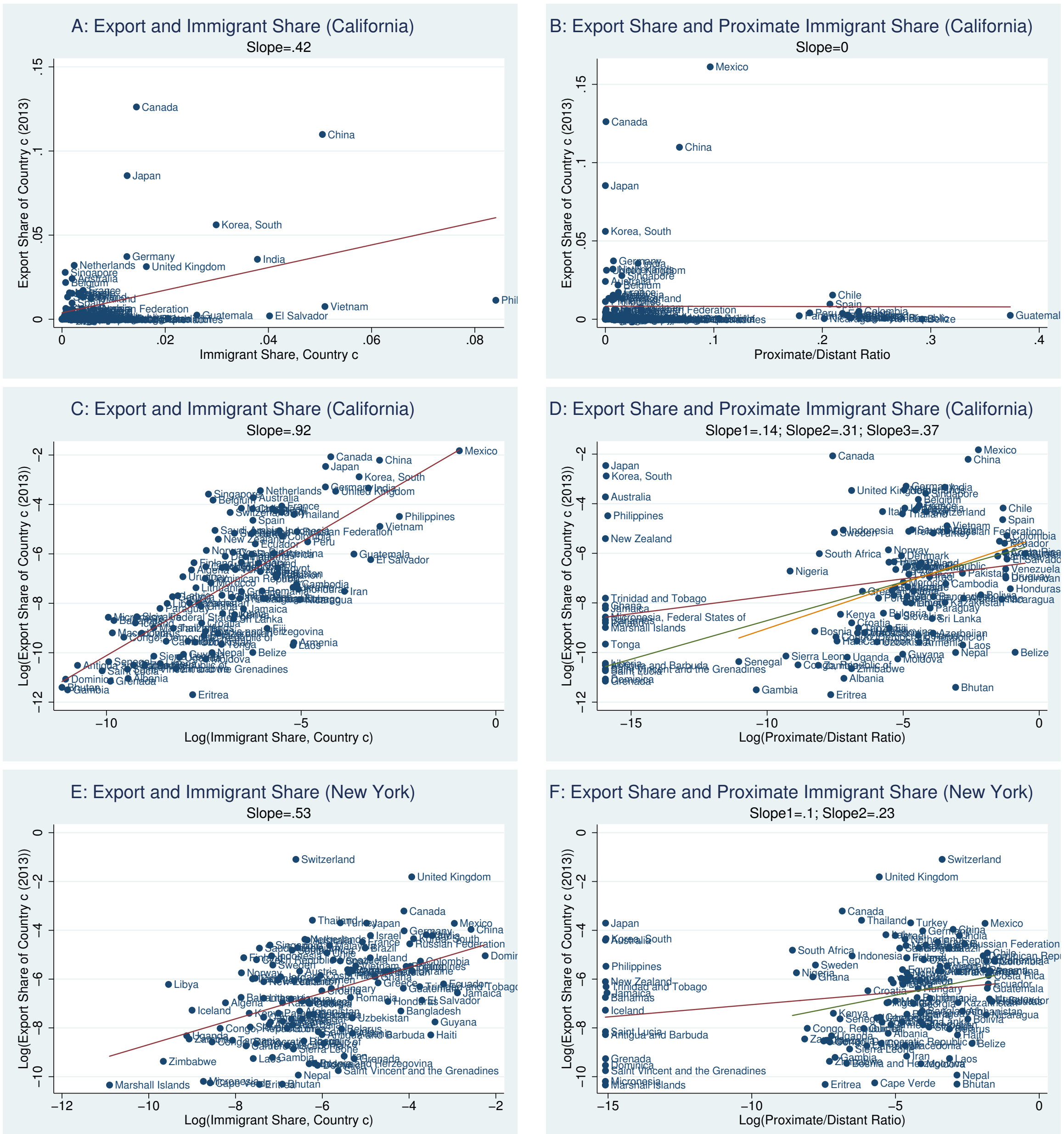

Figure 1: Exports and Immigrant Share.

(Red line–all observations; Green Line–without countries that do not have proximate immigrants overall; Orange line-without countries that do not have proximate immigrants for the state.)

4 Empirical Strategy

4.1 Aggregate Exports

The foundation of the empirical strategy for estimating determinants of trade (exports) is the gravity equation, which has both strong empirical support and theoretical foundation (Anderson, 1979; Bergstrand, 1985). It has the following general form:

$$X_{sct} = B_0 \frac{(Y_{st}Y_{ct})^{\beta_Y}(N_{st}N_{ct})^{\beta_N}}{(D_{sc})^{\beta_D}}, \ (5)$$

where X_{sct} is exports from state s to country c at time t, Y_{st} and Y_{ct} are, respectively, state and country GDP at time t, $N_{st}N_{ct}$ (which is not always featured in the gravity equation) is the product of respective populations, and D_{sc} is the distance between a state and a country. It is important to reiterate at this point that US Trade Online does not record 0 values of trade, so the sample only includes observations with positive export values. Some authors refer to this as the "intensive" margin of trade. Of course, it is not merely intensive at the firm or product level, since there may be plenty of new firms exporting and new products exported over time for the same trading pair. Taking the log of equation (5) we can obtain the following estimating equation:

$$lnX_{sct} = \beta_0 + \beta_Y ln(Y_{st}Y_{ct}) + \beta_N ln(N_{st}N_{ct}) + \beta_D lnD_{sc} + \delta_{sct}, \quad (6)$$

where $\beta_0 = log(B_0)$. If we think of the denominator of (5) as not just distance but also other trade inhibitors and facilitators, which can vary over time, we can add log of immigrant stock from country c in state s at time t (and appropriate fixed effects), and obtain

$$lnX_{sct} = \beta_0 + \beta_Y ln(Y_{st}Y_{ct}) + \beta_N ln(N_{st}N_{ct}) + \beta_D lnD_{sc} + \beta_I lnI_{sct} + \psi_t + FE + \delta_{sct}, \quad (7)$$

where ψ_t are year dummies, $FE = [\psi_s + \psi_c; \psi_{sc}]$,$\psi_s$ and ψ_c are state and country dummies and ψ_{sc} are state-country dummies. We do not include country-year effects, since with state-country fixed effects they are too demanding on the data with almost no "within" explanatory power left ("within" r^2 is less than 0.003), and with state and country effect, results with country-year dummies are virtually the same as with year dummies. We do include country-year effects in industry-level regressions, however.

Since we hypothesize that immigrants proximate to those from country c may matter for exports to country c, we include measures based on borders, language or both, lnI_{spt}, where p stands for proximate. Additionally, we look at the expected number of distant immigrants, lnI_{sdt}, and, separately, all immigrants from countries other than c, $lnI_{sc't}$. To take logs of the immigrant stock variables, we add 1 to observations with zero values. The specification for the equation with different immigrant origin is

$$\begin{aligned} lnX_{sct} = \beta_0 + \beta_Y ln(Y_{st}Y_{ct}) + \beta_N ln(N_{st}N_{ct}) + \beta_D lnD_{sc} + \beta_I lnI_{sct} + [\beta_p lnI_{spt} + \beta_d lnI_{sdt}; \\ \beta_{c'} lnI_{sc't}] + FE + \delta_{sct} \end{aligned} \quad (8)$$

4.2 Industry Level

We make use of the availability of industry-specific export data from USA Trade Online and industry of work in ACS PUMS and investigate the relationship between immigration and trade at the state-country-industry level. Unfortunately, some of the 3 digit NAICS are combined together in the County Business Patterns state data, so we aggregate over these industries, as described in the data section (we also estimated the same model with 1-to-1 matched industries and saw no significant differences in results). Equation (9) is transformed into

$$lnX_{scti} = \beta_0 + \beta_Y ln(Y_{sti}Y_{ct}) + \beta_N ln(N_{sti}N_{ct}) + \beta_D lnD_{sc} + \beta_I lnI_{scti} + \beta_r lnI_{spti} + \beta_d lnI_{sdti} + \\ + FE + \delta_{scti}, \quad (9)$$

where $FE = [\psi_t + \psi_s + \psi_c + \psi_i; \psi_t + \psi_{sc} + \psi_i; \psi_{sc} + \psi_{it}; \psi_t + \psi_{sci}; \psi_{sci} + \psi_{it}]$, lnX_{scti} refers to exports from state s to country c at time t in industry i, Y_{sti} is the GDP of state s in year t in industry i and N_{sti} is the state population that works in industry i in the same year; Y_{ct} and N_{ct} are as before, lnI_{scti} is the number of immigrants in state s from country c in year t working in industry i, and the other variables are defined accordingly.

4.3 2SLS

Although we already control for more potential time-invariant unobservables than any other study of the kind focused on the U.S. as the host country, we may still be concerned about time-variant unobservables that affect both immigrant flows between a country and a state and trade between the two. At the aggregate exports level, we use the instrument previously used by Peri and Requena-Silvente (2010) and Bratti et al. (2014) for immigration-trade link studies and many others used for predicting immigrant stock in other contexts. The instrument uses the assumed orthogonality between immigrant stock at a sufficiently distant period in the past–in our case, 1980–and shocks related to exports in the time period of interest. The instrument for the stock of immigrants from country c in year t in state s is the sum of the number of immigrants from that country in the state in 1980 plus the product of the national share of immigrants from country c in 1980 in the state and national change in the number of immigrants from c between 1980 and year t. The following expression captures the procedure:

$$\hat{I_{sct}} = I_{sc80} + (I_{sct}/I_{ct}) * (I_{ct} - I_{c80}) \quad (10)$$

It is less likely that there may be unobservables that are correlated with exports and *proximate* immigrants, but we do use an instrument constructed the same way as above to account for such possibility. At the industry level, we use the same general process as above, but modify it based on national industry composition of immigrants from country c in year t, and obtain

$$\hat{I_{scti}} = \hat{I_{sct}} * (I_{cti}/I_{ct}), \quad (11)$$

where I_{cti} is the total national number of immigrants from c in industry i in year t. The calculation is analogous for proximate immigrants.

4.4 Geography versus Language

Since our measure of proximity is based on both linguistic and geographic measures, it is useful to see whether both, neither, or only one of them matters. A measure is constructed to estimate the effect of CNL alone, by only summing over the non-bordering countries,

$$I_{sp^ct}(language) = \sum_{j \in C_{NB}; j \neq c;} I_{sjt} * PI_{cj} \quad (12)$$

We term this an "exclusive" language-based measure (since it excludes those from bordering countries). For exclusive border-based measure, we construct the proximity index by subtracting the expected number of common native language speakers from the border value (for all countries)

$$PI_{cj} = max(B_{cj} - CNL_{cj}, 0) \tag{13}$$

We also looked at results with inclusive measures, where $PI = CNL$ or $PI = B$, but the results where not substantially different from exclusive measures.

5 Results

5.1 Aggregate Exports Results

We begin by estimating equations (7) and (8). Table 5 presents results of these aggregate state exports regressions. There are 5606 state-country pairs. Nine columns of the table are divided into 3 blocks: one with just immigrants from the export destination countries (own-country), one with own-country immigrants and those from all other countries, c', and one with own-country, proximate and distant immigrants. The 3 specifications in each block vary only by the type of fixed effects. As indicated in the table, the first equation in each block has only year dummies; the second adds state and country dummies, which is what studies of the immigration-trade link include most often when they have a time dimension; the third column has state-country and year fixed effects.[8]

The first column indicates export elasticity with respect to GDP product of 0.97, which is within the range normally found in the literature, an elasticity with respect to population product of -0.38 and distance elasticity of -0.29. Immigration-export elasticity of 0.3 is above the 0.17 average from Genc et al. (2012) meta-analysis, but within the reported range for both the U.S. and across all host countries. Adding state and country fixed effects in column (2), GDP product elasticity stays about the same, population product elasticity increases in absolute value somewhat, and the coefficent on distance becomes -1.24. Immigration coefficient falls to 0.13, close to the average elasticity reported for the U.S. as the host country. In fact, in the two studies closest to this one, Bandhyopadhay et al. (2008) report the preferred specification estimate of 0.14 and Coughlin and Wall (2011) estimate intensive margin elasticity at 0.14. The r-squared of 0.84 suggests that the model has a high explanatory power. When state-country fixed effects are included, however, the coefficient on immigrants drops to 0 and is not significant. Thus, all that is needed for the commonly reported estimated value to drop to 0 is trading pair fixed effects. Trading pair fixed effects and year dummies is our preferred specification, since there are many reasons to believe trading pair effects are important, and they are empirically found to make a large difference in estimates. Hence, for a larger than usual sample of countries and more recent than usual years, the preferred specification produces an estimate of immigration-trade elasticity of 0. This is not a finding that should necessarily be generalized to immigration-trade link overall, but it is important for the literature with the U.S. as the host country and immigration-export link at the state-country level.

The second block in Table 5 includes immigrants from all countries other than country c. The estimates on all variables included in the first block stay virtually the same, suggesting the coefficients on these variables were not biased due to the omission of $lnI_{sc't}$. Between column 4 and

[8]We are not aware of other estimates of the traditional gravity equation at the state-country level with trading-pair fixed effects. See earlier notes on closest studies: Coughlin and Wall (2011), Bandhyopadhay et al. (2008), Millimet and Osang (2007) and Good (2012)

Table 5: Aggregate Trade Regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GDP Product	0.965**	1.011**	1.060**	0.969**	1.013**	1.063**	0.970**	1.009**	1.060**
	(0.0173)	(0.0597)	(0.0588)	(0.0171)	(0.0598)	(0.0589)	(0.0169)	(0.0599)	(0.0590)
Population Product	-0.372**	-0.427**	-0.365**	-0.378**	-0.415**	-0.343*	-0.355**	-0.434**	-0.349*
	(0.0178)	(0.143)	(0.139)	(0.0176)	(0.146)	(0.140)	(0.0179)	(0.146)	(0.141)
Distance	-0.288**	-1.237**		-0.291**	-1.235**		-0.307**	-1.234**	
	(0.0227)	(0.0932)		(0.0229)	(0.0931)		(0.0239)	(0.0942)	
Same Country:									
Imm_c	0.294**	0.125**	0.00763	0.271**	0.125**	0.00758	0.273**	0.124**	0.00752
	(0.0117)	(0.00874)	(0.00484)	(0.0126)	(0.00865)	(0.00484)	(0.0126)	(0.00858)	(0.00484)
All Other:									
$Imm_{c'}$				0.0602**	-0.0466	-0.0887			
				(0.0144)	(0.111)	(0.0694)			
Proximate:									
Imm_p							-0.0380**	0.00532	0.00919+
							(0.00498)	(0.00864)	(0.00497)
Distant:									
Imm_d							0.0704**	0.00517	-0.113
							(0.0142)	(0.110)	(0.0691)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State and Country	No	Yes	No	No	Yes	No	No	Yes	No
State-Country	No	No	Yes	No	No	Yes	No	No	Yes
Observations	46021	46021	46021	46021	46021	46021	46021	46021	46021
R-squared	0.621	0.838	0.935	0.622	0.838	0.935	0.625	0.838	0.935
R-sq, Within	0.620	0.0827	0.0280	0.621	0.0827	0.0281	0.624	0.0827	0.0282

Standard errors in parentheses. All errors clustered at the state-country level, with 5950 clusters. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

column 5, the coefficient on all immigrants from c' changes from positive and highly significant to negative and insignificant, suggesting it is important to at least control for state and country fixed effects. Further controlling for trading pair fixed effects, the coefficient on $lnI_{sc't}$ becomes more negative, but still insignificant.

The third block includes proximate and distant immigrants instead of all other immigrants lumped together. Going from only year dummies in column 7 to trading-pair fixed effects in column 9, the coefficient on proximate immigrants changes from negative 0.04 to positive 0.01, and is statistically significant at the 90% confidence level. In contrast, the coefficient on distant immigrants for the preferred specification is -0.113 and not statistically different from zero (but statistically different from 0.01). We thus observe a positive and significant effect of proximate immigrants even when the estimate on own-country immigrants is insignificant. The reason for this may be that it is more difficult for the change in the number of own-country immigrants to achieve the magnitudes where an effect is significant enough, because of the generally fewer own-country immigrants. As shown in Table 3, average own-country immigrant stock is 7,000, whereas for proximate immigrants it is 5 times higher, even with a large number of countries that always have 0 associated proximate immigrants. We come back to this point later.

5.2 Industry-Level Results

The first set of industry-level results is presented in Table 6. We have only two blocks in the table: with own-country immigrants only and with proximate and distant immigrants. We omit specifications with c', since the coefficients on c' are very similar to d, as over 90% of immigrants from c' are classified as distant. The 5 columns in each block of Table 8 differ by fixed effects.

Table 6: Industry-Level Regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
GDP Product	0.114**	0.121**	0.0973**	0.201**	0.0527*	0.113**	0.120**	0.0966**	0.200**	0.0525*
	(0.0167)	(0.0163)	(0.0166)	(0.0302)	(0.0262)	(0.0168)	(0.0164)	(0.0166)	(0.0301)	(0.0261)
Population Product	0.457**	0.490**	0.517**	0.537**	0.560**	0.453**	0.497**	0.520**	0.546**	0.563**
	(0.0246)	(0.0237)	(0.0239)	(0.0290)	(0.0292)	(0.0280)	(0.0262)	(0.0265)	(0.0296)	(0.0298)
Distance	-0.914**					-0.909**				
	(0.141)					(0.141)				
Same Country:										
Imm_{ci}	0.134**	0.0705**	0.0715**	-0.00000776	0.000259	0.132**	0.0683**	0.0696**	-0.000336	0.0000899
	(0.0155)	(0.0100)	(0.0101)	(0.00508)	(0.00487)	(0.0153)	(0.00985)	(0.00991)	(0.00507)	(0.00487)
Proximate:										
Imm_{pi}						0.0103+	0.00891*	0.00855+	0.00470*	0.00302
						(0.00568)	(0.00434)	(0.00439)	(0.00223)	(0.00213)
Distant:										
Imm_{di}						-0.00435	-0.0118	-0.00850	-0.00940+	-0.00422
						(0.00903)	(0.00832)	(0.00834)	(0.00537)	(0.00495)
Year	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes	No
State and Country	Yes	No	No	No	No	Yes	No	No	No	No
State-Country	No	Yes	Yes	No	No	No	Yes	Yes	No	No
State-Country-Ind.	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Industry	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Country-Year	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Observations	77192	77192	77192	77192	77192	77192	77192	77192	77192	77192
R-squared	0.687	0.743	0.752	0.919	0.926	0.687	0.743	0.752	0.919	0.926
R-sq within	0.130	0.115	0.118	0.0370	0.0352	0.131	0.116	0.118	0.0371	0.0352

Standard errors in parentheses. All errors clustered at the state-country level, with 3782 clusters. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

The coefficient on the product of industry GDP and foreign country GDP is about ten times lower than in the aggregate-level regression, ranging from 0.05 to 0.2. Unlike population product in the aggregate-level regression, product of industry population and foreign country population is positive, and equal to about 0.5. Distance coefficient is -0.9, close to the estimate at the aggregate level with state and country fixed effects. Immigration-trade elasticity is 0.13 with country, state and industry fixed effects, but it falls to 0.07 with state-country and industry fixed effects; there is little difference in the results when year dummies are replaced with country-year dummies. When we include state-country-industry fixed effects, the coefficient on own-country immigrants drops to zero.

The second block adds proximate and distant immigrant variables. The results for the variables included in the first block stay roughly the same. The coefficient on proximate immigrants in the first 4 specifications ranges from 0.01 to 0.005 and is statistically significant. In contrast, the coefficient on distant immigrants is negative, and significant in the fourth specification only, at -0.01. Again, the effect of proximate immigrants is positive and statistically significant even when the coefficient on own-country immigrants is not. In the last specification, none of the immigrant variables are statistically significant, but this is not the preferred specification, and we include it for completeness only.[9]

[9] Beyond reasons already mentioned, we feel country-year fixed effects are not necessary since our main variable of interest is proximate immigrants, not own-country immigrants, and it is unlikely to be affected by country-specific time dummies. Part of the reason for lack of statistical significance in column 10 is likely lower variation left for the model to exploit after fixed effects.

Table 7: Aggregate Exports Regression with Immigrant Characteristics: IV

	First Stage				2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)
	Dep.=Imm$_c$	Dep.=Imm$_c$	Dep.=Imm$_p$	Dep.=Imm$_p$		
$IV(Imm)_c$	0.254**	0.809**	-0.155**	-0.0385		
	(0.0111)	(0.0344)	(0.0106)	(0.0287)		
$IV(Imm)_p$	-0.134**	-0.0475*	0.211**	0.630**		
	(0.00833)	(0.0226)	(0.00958)	(0.0614)		
GDP Product	0.251**	0.168**	0.341**	0.225**	0.965**	0.916**
	(0.0420)	(0.0377)	(0.0432)	(0.0381)	(0.0628)	(0.0617)
Population Product	0.713**	0.218*	1.473**	0.760**	-0.589**	-0.703**
	(0.112)	(0.101)	(0.122)	(0.111)	(0.168)	(0.158)
Distance	-0.793**		-0.464**		-1.101**	
	(0.0769)		(0.0706)		(0.106)	
Imm_c			0.158**	0.00139	0.222**	0.369**
			(0.00928)	(0.00498)	(0.0368)	(0.0495)
Imm_p	0.162**	0.00177			0.0340	0.132**
	(0.0102)	(0.00633)			(0.0337)	(0.0393)
Imm_d	-1.400**	-0.00312	-1.744**	0.474**	0.267+	-0.115
	(0.0978)	(0.0638)	(0.113)	(0.0698)	(0.155)	(0.0746)
Year	Yes	Yes	Yes	Yes	Yes	Yes
State and Country	Yes	No	Yes	No	Yes	No
State-Country	No	Yes	No	Yes	No	Yes
Observations	46021	46021	46021	46021	46021	46021
R-squared	0.759	0.893	0.942	0.979	0.836	0.925
R-sq, Within	0.173	0.0232	0.157	0.0602	0.0715	-0.127

Standard errors in parentheses. All errors clustered at the state-country level, with 5950 clusters. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

5.3 2SLS Results

We next use the instrumental variable strategy described in Section 4. The first 4 columns of Table 7 show first stage results for the two variables that we instrument for and two fixed effects specifications. We cannot include a specification with country-year fixed effects, since the instrument is constructed in such a way that it uses variation across trading pairs in 1980 and country-year variation, and fixed effects that control for variation in these dimensions would not be left with much variation to exploit. A cursory view of the coefficients and standard errors of the instruments in Table 7 makes it clear that the relevant F-statistics from the first stage are safely above 10.

The last two columns show second stage results. The coefficient on own-country immigrants is 0.22 with state and country fixed effects, up from 0.12 without IV, and 0.37 with state-country fixed effects, up from 0 without the IV. The coefficient on proximate immigrants is 0.03 and 0.13 for the two specifications, being significant only for the second. Distant immigrants coefficient is positive with state and country fixed effects, but negative with trading pair dummies.

Industry-level results are presented in Table 8. We now include specification with country-year effects, since there is an additional dimension of variation in the instrument. As in the aggregate-level regression, we observe an increased and significant coefficient on own-country immigrants; it is close to 0.3 for state-country and industry effects specification and 0.06 for both specifications with state-country-industry fixed effects. The coefficient on proximate immigrants is positive for all three specifications and higher than without instrumenting, ranging from 0.04 to 0.05. Distant immigrants

Table 8: Industry-Level Regression: 2SLS

	First Stage						2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$IV(Imm)_{ci}$	0.500**	0.554**	0.547**	-0.222**	-0.266**	-0.268**			
	(0.0121)	(0.0108)	(0.0114)	(0.0152)	(0.0179)	(0.0188)			
$IV(Imm)_{pi}$		-0.137**	-0.134**	0.787**	0.677**	0.676**			
		(0.00718)	(0.00734)	(0.0115)	(0.0225)	(0.0226)			
GDP Product	0.0733**	0.0251+	0.0246	0.148**	0.0474	0.00192	0.120**	0.194**	0.0502+
	(0.00967)	(0.0149)	(0.0159)	(0.0161)	(0.0305)	(0.0305)	(0.0165)	(0.0223)	(0.0262)
Population Product	0.267**	0.264**	0.270**	0.381**	0.267**	0.288**	0.363**	0.490**	0.511**
	(0.0179)	(0.0225)	(0.0229)	(0.0291)	(0.0400)	(0.0406)	(0.0268)	(0.0303)	(0.0342)
Imm_{ci}				0.173**	0.0726**	0.0623**	0.303**	0.0612**	0.0613**
				(0.0116)	(0.00889)	(0.00881)	(0.0250)	(0.0166)	(0.0180)
Imm_{pi}	0.0284**	0.0204**	0.0175**				0.0386**	0.0495**	0.0408**
	(0.00343)	(0.00242)	(0.00243)				(0.0104)	(0.0104)	(0.0105)
Imm_{di}	-0.0301**	-0.0346**	-0.0393**	0.180**	0.221**	0.212**	-0.00356	-0.0177**	-0.0103+
	(0.00625)	(0.00597)	(0.00605)	(0.0144)	(0.0175)	(0.0170)	(0.00778)	(0.00627)	(0.00551)
Year	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
State-Country	Yes	No	No	Yes	No	No	Yes	No	No
State-Country-Ind.	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Industry	Yes	No	No	Yes	No	No	Yes	No	No
Country-Year	No	No	Yes	No	No	Yes	No	No	Yes
N	77192	77192	77192	77192	77192	77192	77192	77192	77192
R-squared	0.594	0.764	0.768	0.827	0.898	0.900	0.734	0.919	0.925
R-sq within	0.220	0.0994	0.0903	0.272	0.0923	0.0875	0.0822	0.0259	0.0257

Standard errors in parentheses. All errors clustered at the state-country level, with 3782 clusters.+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

coefficient is between -0.02 and -0.01 and statistically significant for the last two specifications.

After instrumenting, the main change in results for both levels of analysis is that the coefficient on immigrants is much higher and is statistically significant in all specifications, while the coefficient on proximate immigrants also increases, but not as much, and remains significant. This may mean that earlier coefficient was downward biased or biased toward zero. This is not the usual concern behind the immigration-export elasticity estimation that invokes the use of instrumental variables. The threat of endogeneity normally concerns unobservables that both increase immigration from country and exports to it, resulting in an upward bias. If we did not have state-country fixed effects, the unobservable could be special economic relations between certain countries and certain states, such as immigrants from East Asia in California and California-East Asia trade (Bardhan et al. (2004)), but since we do, this may not be as a big a concern. Genc et al. (2012) meta-analysis reports an average estimate of 0.35 for the immigration-trade elasticity when using "IV/3SLS/GMM/FGLS" and 0.16 when also using some fixed effects. Thus, our high IV estimates of 0.37 at the aggregate level are similar to those usually found in the literature, but higher than our non-IV results and industry-level results. Importantly, positive and significant effect of proximate immigrants at the industry level in the preferred fixed effects specifications is robust to instrumenting with a commonly used instrument, as is the finding of a different effect of proximate and distant immigrants.

5.3.1 Geography versus Language

Since our measure of proximate immigrants includes both geography- and language-based indicators, we would like to see if both, neither or one of them matter. We start by looking at the aggregate level regression (Table 9). The estimate based on the expected number of geographically but not linguistically proximate immigrants, at 0.01, is similar to that based on the combined measure in the preferred OLS specification (column 2). Conversely, the estimate based on the linguistic measure is small, negative and insignificant. The same is true for the IV estimates, wherein geography-based measure is also more positive, at 0.11, whereas that based on language is more negative.

Table 9: Aggregate-Level Regression: Geography vs. Language

	OLS				2SLS			
	Border		Language		Border		Language	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP Product	1.009**	1.059**	1.007**	1.064**	0.969**	0.924**	0.956**	0.963**
	(0.0598)	(0.0589)	(0.0598)	(0.0588)	(0.0618)	(0.0610)	(0.0614)	(0.0602)
Population Product	-0.435**	-0.360*	-0.448**	-0.329*	-0.581**	-0.736**	-0.653**	-0.543**
	(0.146)	(0.140)	(0.145)	(0.141)	(0.165)	(0.159)	(0.160)	(0.151)
Distance	-1.231**		-1.221**		-1.092**		-1.053**	
	(0.0932)		(0.0943)		(0.106)		(0.107)	
Imm_c	0.124**	0.00757	0.123**	0.00760	0.221**	0.357**	0.225**	0.389**
	(0.00856)	(0.00484)	(0.00869)	(0.00484)	(0.0375)	(0.0495)	(0.0394)	(0.0507)
Imm_p	0.00712	0.00952**	0.0291**	-0.00592	0.0252	0.113**	0.127**	-0.0169
	(0.00659)	(0.00355)	(0.0109)	(0.00746)	(0.0266)	(0.0320)	(0.0321)	(0.0194)
Imm_d	-0.0103	-0.0882	0.0192	-0.124+	0.233+	-0.0862	0.348**	-0.0880
	(0.106)	(0.0688)	(0.0907)	(0.0680)	(0.140)	(0.0735)	(0.128)	(0.0723)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State and Country	Yes	No	Yes	No	Yes	No	Yes	No
State-Country	No	Yes	No	Yes	No	Yes	No	Yes
Observatios	46021	46021	46021	46021	46021	46021	46021	46021
R-squared	0.838	0.935	0.838	0.935	0.836	0.925	0.835	0.925
R-sq, Within	0.0828	0.0282	0.0834	0.0282	0.0720	-0.124	0.0645	-0.130

Standard errors in parentheses. All errors clustered at the state-country level, with 5950 clusters. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Industry-level results without IV are similar to the aggregate-level results. In the first two specifications of the upper panel of Table 10, the estimate on proximate immigrants is somewhat smaller in magnitude than in the model with combined proximity measures, but still positive and significant. Language-based estimate is negative, small and insignificant. The coefficient on distant immigrants is negative and insignificant across all specifications.

When instrumenting for own-country and proximate immigrants (Table 10, lower panel), proximate immigrants coefficient based on geographic measure is between 0.015 and 0.05, whereas it is between 0.07 and 0.11 based on language, all being statistically significant. Thus if non-IV model is to be given preference, geographic proximity among countries of origin is what drives the result on the proximity variable in the main results, whereas in IV case, both measures are important for the result. At the aggregate level, both OLS and 2SLS results point to geographical proximity driving the importance of proximate immigrants.

5.3.2 Other Industries

There are reasons to believe that not only immigrants employed by exporting firms, but those employed by other firms in other industries may also matter. Other industries include service industries, which may include firms that specialize in trade facilitation. We construct an estimate of immigrants employed in all industries other than i and specify index i'; we do so for country c

Table 10: Industry-Level Regression: Geography vs. Language

	Border			Language		
	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
GDP Product	0.121**	0.200**	0.0525*	0.121**	0.200**	0.0525*
	(0.0164)	(0.0301)	(0.0261)	(0.0163)	(0.0301)	(0.0261)
Population Product	0.491**	0.546**	0.564**	0.498**	0.549**	0.565**
	(0.0260)	(0.0297)	(0.0299)	(0.0263)	(0.0298)	(0.0301)
Imm_{ci}	0.0688**	-0.000282	0.000116	0.0696**	-0.000173	0.000165
	(0.00985)	(0.00507)	(0.00487)	(0.00996)	(0.00508)	(0.00488)
Imm_{pi}	0.00752+	0.00361*	0.00166	0.00632	-0.00109	-0.000212
	(0.00387)	(0.00183)	(0.00177)	(0.00546)	(0.00343)	(0.00333)
Imm_{di}	-0.00540	-0.00832	-0.00393	-0.0103	-0.00817	-0.00358
	(0.00800)	(0.00518)	(0.00479)	(0.00840)	(0.00538)	(0.00497)
			2SLS			
GDP Product	0.119**	0.196**	0.0502+	0.120**	0.195**	0.0505+
	(0.0165)	(0.0223)	(0.0261)	(0.0165)	(0.0225)	(0.0262)
Population Product	0.354**	0.504**	0.525**	0.344**	0.452**	0.463**
	(0.0266)	(0.0300)	(0.0339)	(0.0278)	(0.0303)	(0.0357)
Imm_{ci}	0.295**	0.0617**	0.0618**	0.313**	0.0606**	0.0611**
	(0.0247)	(0.0165)	(0.0178)	(0.0251)	(0.0166)	(0.0182)
Imm_{pi}	0.0461**	0.0233**	0.0157*	0.0704**	0.106**	0.114**
	(0.00870)	(0.00645)	(0.00633)	(0.0135)	(0.0163)	(0.0189)
Imm_{di}	0.00498	-0.00764	-0.00253	-0.00793	-0.0244**	-0.0200**
	(0.00737)	(0.00569)	(0.00484)	(0.00774)	(0.00635)	(0.00592)
Year	Yes	Yes	No	Yes	Yes	No
State-Country	Yes	No	No	Yes	No	No
State-Country-Ind.	No	Yes	Yes	No	Yes	Yes
Industry	Yes	No	No	Yes	No	No
Country-Year	No	No	Yes	No	No	Yes
R-squared	0.743	0.919	0.926	0.743	0.919	0.926
R-sq within	0.116	0.0371	0.0352	0.115	0.0370	0.0352

Standard errors in parentheses. All errors clustered at the state-country level, with 3782 clusters.
There are 77192 observations. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

immigrants, proximate and distant immigrants. For own-country immigrants,

$$I_{scti'} = I_{sct} - I_{scti}, \tag{14}$$

and the process for the other two groups is analogous. Table 11 presents OLS and 2SLS results.

Table 11: Industry-Level Regression with Other Industries

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
GDP Product	0.119**	0.195**	0.0517*	0.119**	0.186**	0.0498+
	(0.0163)	(0.0295)	(0.0258)	(0.0166)	(0.0285)	(0.0258)
Population Product	0.493**	0.544**	0.563**	0.360**	0.482**	0.510**
	(0.0261)	(0.0295)	(0.0298)	(0.0303)	(0.0333)	(0.0342)
Same Country:						
Imm_{ci}	0.0685**	-0.000879	-0.0000962	0.304**	0.0598**	0.0618**
	(0.00984)	(0.00506)	(0.00487)	(0.0304)	(0.0174)	(0.0180)
$Imm_{ci'}$	0.0287**	0.0679**	0.0272*	0.0327**	0.0606**	0.0253*
	(0.00985)	(0.0135)	(0.0116)	(0.0103)	(0.0127)	(0.0117)
Proximate:						
Imm_{pi}	0.00904*	0.00479*	0.00301	0.0389**	0.0493**	0.0410**
	(0.00434)	(0.00220)	(0.00210)	(0.0130)	(0.0106)	(0.0105)
$Imm_{pi'}$	0.0241+	0.0337*	-0.00994	0.0304*	0.0293*	-0.0105
	(0.0134)	(0.0155)	(0.0172)	(0.0139)	(0.0148)	(0.0172)
Distant:						
Imm_{di}	-0.0123	-0.00862	-0.00428	-0.00346	-0.0166**	-0.0102+
	(0.00834)	(0.00524)	(0.00487)	(0.00860)	(0.00581)	(0.00540)
$Imm_{di'}$	-0.486*	-0.228	0.0285	-0.342*	-0.464*	-0.00135
	(0.194)	(0.195)	(0.176)	(0.171)	(0.230)	(0.176)
Year	Yes	Yes	No	Yes	Yes	No
State-Country	Yes	No	No	Yes	No	No
State-Country-Ind.	No	Yes	Yes	No	Yes	Yes
Industry	Yes	No	No	Yes	No	No
Country-Year	No	No	Yes	No	No	Yes
N	77192	77192	77192	77192	77192	77192
R-squared	0.743	0.920	0.926	0.734	0.919	0.925
R-sq within	0.116	0.0384	0.0354	0.0823	0.0281	0.0257

Standard errors in parentheses. All errors clustered at the state-country level, with 3782 clusters.
All regressions also include overall state GDP and state population. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Within-industry immigrant groups have similar coefficients as before, with own-country coefficient being 0.7 in the first specification and 0 in the second, proximate immigrants coefficient being 0.01 and 0.05 and distant immigrants coefficient being around -0.01 in both and not statistically significant. The coefficient on own-country immigrants in other industries is positive, at 0.03 to 0.07 and statistically significant, the one on proximate immigrants is about 0.03, and on distant immigrants is -0.49 in the first specification and -0.29 in the second.

IV results from the second half of Table 11 show virtually unchanged proximate immigrant effect for industries i', but it is 10 times higher for industry i, at 0.05 instead of 0.005, in the preferred specification. The effect of country c immigrants in other industries in the preferred specification is similar to OLS results, but that on country c immigrants in the same industry is now positive and significant, close to what it was without explanatory variables for other industries. Distant immigrants now also have a negative and statistically significant effect, at -0.017, while the coefficient on distant immigrants in other industries remains negative and large, at 0.34-0.46 for the preferred specifications.

One possible reason for the observed effect of immigrants from other industries is that foreign-born workers in other industries in firms important for exports tend to help export to their country of origin or related (geographically and linguistically) countries. But then the question arises, why is the aggregate-level estimate of the coefficient on own-country immigrants equal to zero? Part of the reason may be that the different subsampes of trading pairs; subsample that is used for industry regression has trading pairs with larger immigrant presence compared to the aggregate-level sample, since we only have observations with positive state-industry-level exports, which is more restrictive than positive state-level exports.

Table 12: Aggregate Exports Regression with Different Cutoffs

	(1) 50	(2) 100	(3) 250	(4) 500	(5) 1000	(6) 2500	(7) Ind. reg. sample
GDP Product	1.087**	1.130**	1.217**	1.306**	1.351**	1.325**	1.114**
	(0.0595)	(0.0622)	(0.0693)	(0.0796)	(0.0948)	(0.121)	(0.107)
Population Product	-0.357*	-0.410**	-0.449**	-0.405*	-0.244	0.132	-0.478^{+}
	(0.142)	(0.148)	(0.169)	(0.192)	(0.239)	(0.334)	(0.269)
Same Country:							
Imm_c	0.00371	0.00414	0.00719	0.0170	0.0428*	0.0657*	0.0159^{+}
	(0.00535)	(0.00622)	(0.00854)	(0.0115)	(0.0172)	(0.0304)	(0.00929)
Proximate:							
Imm_p	0.00751	0.0113*	0.0113^{+}	0.0133^{+}	0.0134	0.0191	0.0134
	(0.00513)	(0.00545)	(0.00662)	(0.00780)	(0.00971)	(0.0138)	(0.0168)
Distant:							
Imm_d	-0.121^{+}	-0.137^{+}	-0.150	-0.154	-0.277*	-0.388^{+}	-0.496**
	(0.0713)	(0.0768)	(0.0921)	(0.109)	(0.140)	(0.206)	(0.137)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State and Country	No	No	No	No	No	No	No
State-Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	44956	41808	34673	27765	20375	12279	14309
R-squared	0.936	0.937	0.938	0.939	0.938	0.936	0.943
R-sq, Within	0.0293	0.0319	0.0360	0.0418	0.0460	0.0487	0.0254

Standard errors in parentheses. All errors clustered at the state-country level, with 3141 clusters.$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

To explore whether the sample of trading pairs makes a difference in estimates, we look at subsamples of trading pairs with the number of immigrants above specified cutoffs. Table 12 shows that with the increase in cutoff for trading pairs from 50 to 2500, the estimate on own country immigrants with trading pair fixed effects steadily grows from 0.003 to 0.07, reaching a statistically significant estimate of 0.04 at 1000, close to the median value. The last column includes the subsample of trading pairs that are used in the industry level regression, and produces a coefficient value of 0.016. The coefficient on proximate immigrants is between 0.01 and 0.02 for 7 out of 8

Table 13: Industry-Level Regression with Different ("$Imm_{ci'}$") Cutoffs

	(1) 100	(2) 250	(3) 500	(4) 1000	(5) 2500	(6) 5000	(7) 10000
GDP Product	0.191**	0.190**	0.188**	0.189**	0.185**	0.179**	0.182**
	(0.0287)	(0.0288)	(0.0291)	(0.0297)	(0.0315)	(0.0347)	(0.0388)
Population Product	0.540**	0.540**	0.543**	0.548**	0.551**	0.561**	0.567**
	(0.0295)	(0.0296)	(0.0298)	(0.0301)	(0.0314)	(0.0345)	(0.0377)
Imm_{ci}	-0.00152	-0.00189	-0.00176	-0.000948	-0.0000699	-0.000461	-0.00189
	(0.00508)	(0.00508)	(0.00512)	(0.00517)	(0.00542)	(0.00580)	(0.00652)
Imm_{pi}	0.00457*	0.00460*	0.00431+	0.00398+	0.00304	0.00320	0.00251
	(0.00221)	(0.00221)	(0.00222)	(0.00223)	(0.00231)	(0.00244)	(0.00263)
Imm_{di}	-0.00886+	-0.00825	-0.00859	-0.00853	-0.00781	-0.00752	-0.00544
	(0.00523)	(0.00525)	(0.00527)	(0.00533)	(0.00543)	(0.00556)	(0.00604)
$Imm_{ci'}$	0.0932**	0.106**	0.117**	0.130**	0.159**	0.176**	0.181**
	(0.0184)	(0.0201)	(0.0223)	(0.0260)	(0.0348)	(0.0459)	(0.0650)
$Imm_{pi'}$	0.0309*	0.0317*	0.0311*	0.0332*	0.0402*	0.0747**	0.100*
	(0.0151)	(0.0152)	(0.0155)	(0.0160)	(0.0182)	(0.0240)	(0.0399)
$Imm_{di'}$	-0.427+	-0.432+	-0.440+	-0.462+	-0.446	-0.630+	-0.663
	(0.232)	(0.235)	(0.239)	(0.250)	(0.279)	(0.337)	(0.419)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Country	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	76406	75472	73766	70560	62968	53993	43484
R-squared	0.919	0.919	0.918	0.917	0.915	0.914	0.913
R-sq, Within	0.0395	0.0398	0.0402	0.0408	0.0410	0.0426	0.0458

Standard errors in parentheses. All errors clustered at the state-country level, with 728 clusters. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$
All regressions also include overall state GDP and state population.

specifications, but is only precisely estimate in 3. The effect of distant immigrants grows steadily more negative with increasing cutoffs and is statistically significant in 6 out of 8 columns.

At the industry level (Table 13), we vary the cutoff for the number of own-country immigrants in other industries. Analogously to the aggregate-level regression, the estimates of the effect of own-country immigrants in other industries grow with increasing cutoff, from 0.09 at 100 to 0.18 at 10000 (close to the median), always being statistically significant. The coefficient on proximate immigrants in other industries also increases, from 0.03 to 0.1, while that on distant immigrants in other industries becomes more negative, -0.66 from -0.43. Thus, the positive effect of own-country and proximate immigrants and the negative effect of distant immigrants is more pronounced in trading pairs with a larger own immigrant stock. Notably, the number of immigrants in other industries for each of the three groups is larger than within the export industry, which may be part of the reason the effect of immigrants in other industries is more pronounced than of those in the exporting industry.

6 Conclusion

In this study, we combined data on geographic and linguistic proximity of immigrant countries of origin and U.S. export destinations to examine whether immigrants stimulate exports to countries geographically or linguistically proximate to their countries of origin. The rational behind the idea that immigrants may promote exports to countries other than their country of origin is that the channels that are normally cited as the ones through which immigrants affect exports, particularly networks, information, and communication, do not have to apply to just the country of origin: business networks may cross borders, valuable information may extend to neighbouring countries with institutional and cultural similarities, and linguistic (and cultural) connections through common native language may facilitate communication and cooperation between people from different countries of origin.

The results, based on industry-level and aggregate exports from U.S. states to foreign countries over the course of 10 years, 2003-2013, confirm this intuition. The findings suggest that foreign-born in the U.S. do indeed stimulate exports to countries geographically and linguistically proximate to the countries of origin. Moreover, there also appears to be a trade diversion effect from geographically and linguistically distant countries. The results are robust to instrumenting for immigrants with immigrant enclaves in 1980. Comparing the relative importance of the two proximity components, we find some evidence for the importance of both, but that the finding of the importance of geographic proximity is more convincing and robust.

One possible way forward to further investigate the effect of immigrants that we term proximate and distant and assess robustness of the results is to conduct the analysis with other countries as the hosts, as has been done with immigrants from export destination countries. Detailed data from Italy and Spain, in particular, has enabled a number of valuable studies of the effect of own-country immigrants, and can be used for the assessment of the importance of immigrant country of origin proximity. Another possibility is cross-country studies with country-country migration and trade. Additionally, we only looked at geographic and linguistic relationship between immigrant countries of origin, but there may be other factors, such as institutional similarity or level of economic development, that also affect the relevance of workers from some countries for exports to others. As with immigrants from the export destination countries, immigrant and export good characteristics may affect the estimated elasticities of trade with respect to proximate immigrants. Theoretical pathways through which immigration is believed to affect trade can suggest multiple ways of analysis to better understand the relationship between immigrants from some countries and exports to others.

References

Aguiar, Angel, Terrie Walmsley, and Jason Abrevaya. 2007. "Effects of bilateral trade on migration flows: the case of the United States." *10th Annual Conference on Global Economic Analysis, Purdue University, USA.*

Aleksynska, Mariya, and Giovanni Peri. 2014. "Isolating the network effect of immigrants on trade." *The World Economy*, 37(3): 434–455.

Anderson, James E. 1979. "A theoretical foundation for the gravity equation." *The American Economic Review*, 106–116.

Bandyopadhyay, Subhayu, Cletus C Coughlin, and Howard J Wall. 2008. "Ethnic Networks and US Exports*." *Review of International Economics*, 16(1): 199–213.

Bardhan, Ashok Deo, and Subhrajit Guhathakurta. 2004. "Global linkages of subnational regions: coastal exports and international networks." *Contemporary Economic Policy*, 22(2): 225–236.

Bergstrand, Jeffrey H. 1985. "The gravity equation in international trade: some microeconomic foundations and empirical evidence." *The review of economics and statistics*, 474–481.

Bowen, Harry P, and Jennifer Pédussel Wu. 2013. "Immigrant specificity and the relationship between trade and immigration: theory and evidence." *Southern Economic Journal*, 80(2): 366–384.

Bratti, Massimiliano, and Chiara Conti. 2014. "The effect of (mostly unskilled) immigration on the innovation of Italian regions." *IZA Discussion Paper.*

Bratti, Massimiliano, Luca De Benedictis, and Gianluca Santoni. 2014. "On the pro-trade effects of immigrants." *Review of World Economics*, 150(3): 557–594.

Briant, Anthony, Pierre-Philippe Combes, and Miren Lafourcade. 2014. "Product complexity, quality of institutions and the protrade effect of immigrants." *The World Economy*, 37(1): 63–85.

Chaney, Thomas. 2008. "Distorted gravity: the intensive and extensive margins of international trade." *The American Economic Review*, 98(4): 1707–1721.

Coughlin, Cletus C, and Howard J Wall. 2011. "Ethnic networks and trade: Intensive versus extensive margins." *Economics Letters*, 113(1): 73–75.

Dunlevy, James A. 2006. "The influence of corruption and language on the protrade effect of immigrants: Evidence from the American States." *Review of Economics and Statistics*, 88(1): 182–186.

Egger, Peter H, Maximilian Von Ehrlich, and Douglas R Nelson. 2012. "Migration and trade." *The World Economy*, 35(2): 216–241.

Genc, Murat, Masood Gheasi, Peter Nijkamp, and Jacques Poot. 2012. "The impact of immigration on international trade: a meta-analysis." *Migration Impact Assessment: New Horizons*, 301.

Girma, Sourafel, and Zhihao Yu. 2002. "The link between immigration and trade: Evidence from the United Kingdom." *Weltwirtschaftliches Archiv*, 138(1): 115–130.

Good, Michael. 2012. "How Localized is the Pro-trade Effect of Immigration? Evidence from Mexico and the United States." *Economics Research Working Paper Series, Florida International University.*

Gould, David M. 1994. "Immigrant links to the home country: empirical implications for US bilateral trade flows." *The Review of Economics and Statistics*, 302–316.

Hanson, Gordon H. 2012. "Immigration, Productivity and Competitiveness in American Industry." *Competing for Talent: The United States and High-Skilled Immigration.*

Hatzigeorgiou, Andreas, and Magnus Lodefalk. 2014. "Trade, Migration and Integration–Evidence and Policy Implications." *The World Economy.*

Head, Keith, and John Ries. 1998. "Immigration and trade creation: econometric evidence from Canada." *Canadian journal of economics*, 47–62.

Herander, Mark G, and Luz A Saavedra. 2005. "Exports and the structure of immigrant-based networks: the role of geographic proximity." *Review of Economics and Statistics*, 87(2): 323–335.

Hiller, Sanne. 2013. "Does immigrant employment matter for export sales? Evidence from Denmark." *Review of World Economics*, 149(2): 369–394.

Hong, Tan Chuie, and A Solucis Santhapparaj. 2006. "Skilled labor immigration and external trade in Malaysia: A pooled data analysis." *Perspectives on Global Development and Technology*, 5(4): 351–366.

Jansen, Marion, and Roberta Piermartini. 2009. "Temporary migration and bilateral trade flows." *The World Economy*, 32(5): 735–753.

Kandogan, Yener. 2005. "Is Immigration Necessary and Sufficient? The Swiss Case on the Role of Immigrants on International Trade." Working Paper.

Law, David, Murat Genç, and John Bryant. 2013. "Trade, diaspora and migration to New Zealand." *The World Economy*, 36(5): 582–606.

Lewer, Joshua J, and Hendrik Van den Berg. 2009. "Does immigration stimulate international trade? Measuring the channels of influence." *The International Trade Journal*, 23(2): 187–230.

Malchow-Møller, Nikolaj, Jakob R Munch, and Jan Rose Skaksen. 2011. "Do foreign experts increase the productivity of domestic firms?" Discussion paper series//Forschungsinstitut zur Zukunft der Arbeit.

Martín-Montaner, Joan, Francisco Requena, Guadalupe Serrano, et al. 2014. "International trade and migrant networks: Is It really about qualifications?" *Estudios de Economía*, 41(2): 252.

Melitz, Jacques, and Farid Toubal. 2014. "Native language, spoken language, translation and trade." *Journal of International Economics*, 93(2): 351–363.

Millimet, Daniel L, and Thomas Osang. 2007. "Do state borders matter for US intranational trade? The role of history and internal migration." *Canadian Journal of Economics/Revue canadienne d'économique*, 40(1): 93–126.

Mundell, Robert A. 1957. "International trade and factor mobility." *the American Economic Review*, 321–335.

Mundra, Kusum. 2005. "Immigration and international trade: A semiparametric empirical investigation." *The Journal of International Trade & Economic Development*, 14(1): 65–91.

Mundra, Kusum. 2014. "Immigration and Trade Creation for the United States: Role of Immigrant Occupation." *The International Trade Journal*, 28(4): 311–343.

Pennerstorfer, Dieter. 2014. "Export, Migration and Costs of Trade: Evidence from Central European Firms." *Regional Studies*, 1–16.

Peri, Giovanni. 2012. "The effect of immigration on productivity: Evidence from US states." *Review of Economics and Statistics*, 94(1): 348–358.

Peri, Giovanni, and Francisco Requena-Silvente. 2010. "The trade creation effect of immigrants: evidence from the remarkable case of Spain." *Canadian Journal of Economics/Revue canadienne d'économique*, 43(4): 1433–1459.

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2013. "STEM workers, H1B Visas and Productivity in US Cities." *London: Norface Migration.*

Rauch, James E. 1999. "Networks versus markets in international trade." *Journal of international Economics*, 48(1): 7–35.

Rauch, James E. 2001. "Business and social networks in international trade." *Journal of economic literature*, 1177–1203.

Rauch, James E, and Vitor Trindade. 2002. "Ethnic Chinese networks in international trade." *Review of Economics and Statistics*, 84(1): 116–130.

Requena, Francisco, Vicente Pallardó, and Andrés Artal. 2012. "Which immigrants stimulate exports in their host country?(en homenaje a José Vicente Blanes)."

Serrano-Domingo, Guadalupe, and Francisco Requena-Silvente. 2013. "Re-examining the migration–trade link using province data: An application of the generalized propensity score." *Economic Modelling*, 32: 247–261.

Tadesse, Bedassa, and Roger White. 2010. "Cultural distance as a determinant of bilateral trade flows: do immigrants counter the effect of cultural differences?" *Applied Economics Letters*, 17(2): 147–152.

Wagner, Don, Keith Head, and John C Ries. 2002. "Immigration and the Trade of Provinces." *Scottish Journal of Political Economy*, 49: 507–525.

White, Roger. 2009*a*. "Immigration, trade and home country development: state-level variation in the US immigrant–export link." *Journal of International Migration and Integration/Revue de l'integration et de la migration internationale*, 10(2): 121–143.

White, Roger. 2009*b*. "Immigration, trade and product differentiation." *Economic Issues*, 14(1): 43.

7 Appendix 1

List of Countries (131): Afghanistan, Albania, Algeria, Antigua and Barbuda, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahamas, Bangladesh, Barbados, Belarus, Belgium, Belize, Bermuda, Bhutan, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Cambodia, Cameroon, Canada, Cape Verde, Chile, China, Colombia, Democratic Republic of the Congo, Republic of the Congo, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Eritrea, Estonia, Fiji, Finland, France, Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guyana, Haiti, Honduras, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea, South, Kuwait, Laos, Latvia, Lebanon, Liberia, Libya, Lithuania, Macedonia, Malaysia, Marshall

Islands, Mexico, Micronesia, Moldova, Morocco, Nepal, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Russian Federation, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Saudi Arabia, Senegal, Sierra Leone, Singapore, Slovakia, South Africa, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Tanzania, Thailand, Togo, Tonga, Trinidad and Tobago, Turkey, Uganda, Ukraine, United Arab Emirates, United Kingdom, Uruguay, Uzbekistan, Venezuela, Vietnam, Yemen, Zambia, Zimbabwe.

Table A1: Immigration-Trade Link Literature

	Positive	Negative	No difference
Country Type			
Same language (on exports)	Rauch and Trindade (2002); Aleksynska and Peri (2014); Egger et al. (2012)	Blanes-Cristobal (2008)	Wagner et al. (2002)
Same language (on elasticity)	Melitz and Toubal (2014)	Aleksynska and Peri (2014)	
English-speaking country (on elasticity)		Dunlevy (2006)	
Spanish-speaking country (on elasticity)		Dunlevy (2006)	
Same colonizer (on exports)	Melitz and Toubal (2014)		
Former colony of (on exports)	Melitz and Toubal (2014); Aleksynska and Peri (2014); Blanes-Cristobal (2008)		
Cultural distance (on elasticity)	Kandogan (2005)	Tadesse and White (2010)	
Developed countries (on elasticity)		White (2009a); White (2009b); Bratti (2014); Hatzigeorgiu and Ladefalk (2014)	Co et al. (2004)
Institutional similarity high (on easticity)		Herander and Saveedra (2005)	Dunlevy (2006)
Common legal system (on elasticity)		Aleksynska and Peri (2014)	
High corruption (on elasticity)	Dunlevy (2006); Hatzigeorgiu and Ladefalk (2014)		
Same economic zone/union (on elasticity)	Blanes-Cristobal (2008)	Hong and Santhaparraj (2006)	
Former colony (on elasticity)		Blanes-Cristobal (2008)	Aleksynska and Peri (2014)
Same religion (on elasticity)	Aleksynska and Peri (2014)		
Rule of law quality (on elasticity)		Briant et al. (2014)	
U.S. West Coast (vs. East; on elasticity)	Bardhan and Guhathakurta (2004)		
Imigrant Type			
More educated/skilled (on elasticity)	Aleksynska and Peri (2014); Blanes-Cristobal (2008)		Jansen and Piermartini (2009)
Professional occupations (on elasticity)	Mundra (2014)		
Managers (on elasticity)	Aleksynska and Peri (2014); Blanes-Cristobal (2008); Martin-Montaner et al. (2014)		
Entrepeneurs (on elasticity)		Head and Ries (1998)	
Men (on elasticity)	Hatzigeorgiu and Ladefalk (2014)	Bowen (2014)	
In-province (on elasticity)	Bratti et al. (2011); Herander and Saveedra (2005); Requena and Peri (2012)		
Temporary migrants (on elasticity)	Jansen and Piermartini (2009)		
Export Type			
Differentiated goods (on elasticity)	Rauch and Trindade (2002); White (2009b); Melitz and Toubal (2014); Peri and Requena (2010)		
Cultural goods (on elasticity)		Tadesse and White (2010)	
Goods (vs. services; on elasticity)	Bowen and Wu (2013)		
Consumer goods (on elasticity)	Blanes-Cristobal (2008)		
Exports (vs. imports; on elasticity)	Blanes-Cristobal (2008); Briant et al. (2014)		
Extensive margin (vs. intensive; on elasticity)	Hiller (2013); Peri and Requena (2010)	Coughlin and Wall (2011)	