Scottish Journal of Political Economy, Vol. 49, No. 5, November 2002

IMMIGRATION AND THE TRADE OF PROVINCES

Don Wagner, Keith Head** and John Ries***

Abstract

A link between immigration, imports, and exports has been found by a number of papers that have used the gravity equation to analyze bilateral trade patterns. We discuss what this research implies about the mechanisms through which immigrants expand trade and identify strengths and weakness of the various approaches. This paper also contributes to this literature by estimating immigrant effects for Canada using cross-province variation in international trade and immigration patterns. We derive an alternative functional form capturing the relationship between immigration and trade based on the proposition that immigrants use their connections and superior 'market intelligence' to exploit trade opportunities that non-immigrants do not access. We find that the average new immigrant expands exports to his/her native country by $312 and expands imports by $944.

I Introduction

The gravity model of international trade has consistently revealed a strong association between immigration and trade. Not only have different studies revealed a robust relationship for different samples and specifications, the strength of the immigration effect varies in sensible ways for different trading partners, products, and types of immigrants. The estimated magnitude of the immigration effect, however, differs greatly across studies. The analysis of Head and Ries (1998), Dunlevy and Hutchinson (1999, 2001), Rauch and Trindade (2002), Girma and Yu (2002) and Combes *et al.* (2002) based on cross-sectional information find large effects whereas Gould's (1994) estimation based on times series variation indicates smaller effects. The discrepancy from an econometric standpoint is easy to explain—cross-sectional estimates may be upwardly biased due to unobserved characteristics of trading relationships whereas fixed effect estimates may have the opposite bias due to the magnification of measurement error caused by this technique. Alternative explanations for the discrepancy are differences in specifications and samples.

* University of Prince Edward Island
** University of British Columbia

This paper reviews studies of the immigration-trade link to establish what we know so far about the mechanisms by which immigrants increase trade and the magnitude of the effects. We identify the consistent findings in the literature as well as conflicting ones. We also evaluate the different approaches employed in existing studies and discuss their strengths and weaknesses. Finally, we introduce new estimates based on the trade of Canadian provinces. These results are derived from specifications explicitly linked to theories underlying the immigration-trade link. The estimates exploit cross-sectional information on trade and immigration across Canadian provinces and control for fixed effects between Canada and its trading partners. We also report how controls for language commonality between a Canadian province and a foreign trading partner influence the immigration effect on trade.

In our review of the literature, we discuss a number of methodological issues. The issues addressed in our commentary extend beyond immigration-trade studies to the large number of recent papers using the gravity model to evaluate the factors influencing trade such as a common currency, common language, and foreign aid on trade levels.[1] One important issue is specification. Gravity models posit a log-linear relationship between trade volumes, source and recipient country GDPs, and trading distances. Generally, researchers simply insert measures of other factors into a gravity model without regard to theoretical considerations. A second issue is regression technique. Estimates may alternatively be based on cross-sectional (country) variation or time-series variation. A third issue concerns endogeneity. Is the association between trade and immigration causal or driven by unmeasured common factors?

The paper has the following structure. Section II reviews the literature and identifies issues warranting further study. We argue that our study based on Canadian provincial trade addresses some of these issues. Our first specification is a standard one that employs country-fixed effects and introduces a detailed language variable measuring the overlap between languages spoken in Canada and its trading partners. Our ensuing specifications examine how the results change when we employ specifications based on theories predicting how immigrants expand trade. These results are reported and discussed in Section III. We conclude by summarizing our current knowledge of the trade and immigration link and suggesting areas for future research.

II The Literature

The papers we review are Gould, Head and Ries, Dunlevy and Hutchinson, Rauch and Trindade, Girma and Yu, and Combes *et al.* Gould studies US trade with 47 trading partners over the 1970–1986 period, whereas Head and Ries consider Canadian trade with 136 countries from 1980 to 1992. Dunlevy and Hutchinson evaluate US imports from 1870–1910 and they analyze US exports ova the same period in their 2002 manuscript. Rauch and Trindade do not

[1] Wagner (forthcoming) considers the effects of overseas development assistance; Frankel and Rose (2002) a common currency; and Helliwell (1999) common language.

consider immigration directly. Instead, they investigate whether the presence of large numbers of ethnic Chinese residents in partner countries is associated with more trade. They compute the product of the ethnic Chinese population shares for each trading pair and add this variable to a gravity equation. Girma and Yu examine UK trade and immigration. Combes *et al.* examine trade and people flows between French departments (95 sub-national units roughly equivalent to counties in the US or UK). Thus, unlike the other studies, they examine intra-national trade and migration.

In the following sections we discuss these papers in three regards. First, we identify the extent to which a relationship between trade and immigration is revealed in the data. Second, we discuss mechanisms through which immigrants affect trade. Third, we discuss specification and data issues to evaluate the accuracy of the estimates and understand the differences in the magnitude of the immigration effect observed across studies.

A robust association between trade and immigration

The five studies examining immigration—Gould, Head and Ries, Dunlevy and Hutchinson, Girma and Yu, and Combes *et al.*—all find a statistically significant relationship between trade flows and immigration. Rauch and Trindade find that country pairs with higher concentrations of ethnic Chinese residents trade more with each other. Table 1 presents a comparison of the elasticities of trade with respect to immigration found in these papers. Each of the papers in the table contains multiple estimates. We use the estimates reported for the full sample when available. For Girma and Yu, we chose their estimates for non-commonwealth immigration since they found no statistically significant effects for immigrants from commonwealth members. For Rauch and Trindade

TABLE 1
Results in the literature

Authors	Sample countries and period	Export elasticity	Import elasticity
Gould (1994)	US and 47 Trade Partners: 1970–86	0·02*	0·01*
Head and Ries (1988)	Canada and 136 Partners: 1980–92	0·10	0·31
Dunlevy and Hutchinson (1999, 2002)	US and 17 Partners: 1870–1910	0·08	0·29
Girma and Yu (2002)	UK and 48 Partners: 1981–1993	0·16†	0·10†
Combes *et al.* (2002)	95 French *Departments*: 1993	0·25	0·14
Rauch and Trindade (2002)	63 Nations: 1980, 1990	0·47‡	0·47‡

Notes:
* Calculated for immigration levels in 1986 (see text for explanation).
† Trade with non-commonwealth countries.
‡ Calculated for differentiated products in 1990 (see text for explanation).

we focused on the sample of countries with at least 1% of the population being ethnically Chinese.

With the exception of Gould (1994) and Rauch and Trindade (2002), the authors estimated constant elasticity relationships between trade and immigration. That is, they regressed the log of bilateral exports or imports on the log of the stock of immigrants. This facilitates comparisons across studies.

In the functional form adopted by Gould (1994), the elasticity of immigrants decreases with the volume of immigrants. Thus, his estimates are not directly comparable, but we use his results to calculate approximate elasticities based on 1986 US immigration levels.[2] Gould's export elasticity is about one-fifth of the elasticities of the other researchers, and his imports less than one twentieth. Later on we attribute these differences to his estimation method.

Trindade and Rauch also opted not to estimate a constant elasticity (linear-in-logs) specification relating exports to immigration. Instead their dependent variable is the log of the sum of bilateral exports and imports. Their key explanatory variable (other than the gravity controls) is the product of the two trade partners' ethnic Chinese share of their respective populations. We calculate a single elasticity of trade with respect to ethnic presence by computing the effect that a 1% increase in a country's Chinese population exerts on trade. This equals $100^*(e^{\beta(0{\cdot}01)z} - 1)$, where z is the mean of the Chinese share variable and β is this variable's coefficient estimate. We apply this calculation to Rauch and Trindade's results involving 1990 trade between countries with at least a 1% share of Chinese people, using their conservative aggregation method to classify goods. For countries with significant (at least 1%) Chinese populations, we compute this elasticity to be 0·47 for trade in differentiated goods and 0·21 for trade in homogenous goods traded on organized exchanges.

The results we show in Table 1 represent results from baseline specifications in the literature. The papers estimate supplementary regressions that consider different subsets of the data in order to explore how immigration effects vary across different trading partners, commodities, and types of immigrant. We discuss some of these additional results in the next section.

The mechanisms underlying the relationship

To understand whether immigrants generate higher trade, it is necessary to identify the mechanisms through which trade creation would occur. Theory suggests the types of immigrants and trading relationships for which the immigration effect should be most pronounced and serves as a basis for

[2] These results are computed as follows. Gould computes the marginal trade increase from adding one new immigrant to the 1986 stock of immigrants in the US and reports these results for 47 countries in an appendix. We compute a weighted average of these marginal trade increases. We then adjust the export and import figures by $1/(1 - \lambda)$, where λ is the coefficient for his lagged dependent variable. (We make this adjustment to capture the effect of immigrants on the lagged dependent variable.) We then compute elasticities using this calculated figure, the total stock of immigrants from those 47 countries, and the total US exports to (and imports from) those countries.

statistical hypothesis testing. If the magnitude of the immigration effect is systematically related to factors that theory indicates it should be related to, then there is evidence that the measured effects are not simply spurious but instead represent a causal relationship.

One mechanism applies only to imports, not exports. Immigrants may prefer certain goods from their country of origin, based on tastes developed before migration. These preferences would generate more imports from the country of origin. A second reason pertains to both imports and exports. Immigrants may face lower barriers to trade because of their knowledge of their countries of origin. One common barrier to trade is simply ignorance of trading opportunities. In this area, immigrants often hold an advantage because they are more likely to be connected to business networks that enable them to find customers or suppliers in the country from which they emigrated. Another reason they are better positioned to seize opportunities is that they are more familiar with the market needs in their country of origin. In addition, immigrants normally face lower communication barriers. That is, they usually know the language of their home country, and they know the culture, enabling them to understand the values of their former compatriots and their ways of thinking. Immigrants are also better connected to the business networks of their native countries and are better informed on whom to trust and whom not to trust. Finally, immigrants are more likely to know the local business laws and practices. For all the reasons listed above, immigrants may be in a better position than other people are to conduct trade with their countries of origin because of the information they possess. This 'information effect' implies that immigrants should raise imports and exports.

Rauch and Trindade posit that immigrants facilitate the enforcement of contracts. An immigrant may be plugged into a business network in her country of origin. If either party to a business transaction acts opportunistically, that party's reputation would suffer within that network. Thus, the business network will provide a means of deterring opportunism and serves to enforce contracts. Rauch and Trindade propose that this enforcement effect will benefit both imports and exports.

The arguments outlined above have two implications. First, both the information and enforcement effect of immigration should be associated with greater trade. Second, the preference effect will cause immigrants to exert a larger effect on imports than exports. One might argue that the information and enforcement effects are reflected in the export elasticity whereas the import elasticity includes a preference effect in addition to the information and enforcement effects. Table 1 shows conflicting evidence on this proposition. Head and Ries' and Dunlevy and Hutchinson's results imply that the preference effect is approximately double the information/enforcement effect (since the immigration effect on imports is approximately three times the effect on exports). However, Girma and Yu obtain smaller immigration effects on imports than on exports. As we show in Table 1, Gould's results imply a lower elasticity for imports than exports. This elasticity represents the average effect across all immigrants. His estimates reveal, however, that when the stock of immigrants exceeds 5000, the import elasticity associated with further immigration exceeds the export elasticity. One explanation for the persistence of the immigration effect on imports found by Gould is the presence of a preference effect.

Another implication ensues from the information hypothesis; immigrants should expand trade for goods in which information is most valuable. Potential trading partners may need information about product specifications for differentiated goods but not for homogeneous goods. Thus, if immigrants serve to bridge information gaps, their effects should be strongest on differentiated goods. Gould finds that the immigration-trade link is stronger in the consumer manufactured goods sector than in the producer goods sector. Dunlevy and Hutchinson find that migration strongly affects imports of finished and semi-finished goods, and does not strongly affect imports of crude goods. Both sets of results support the information hypothesis that immigrants have a stronger effect for trade in goods that can be considered differentiated products. Rauch and Trindade classify goods into three categories: organized exchange, reference priced and differentiated. They posit that goods listed on organized exchanges (primarily commodities) require the least and differentiated goods the most knowledge. Their results indicate that the Chinese network affects trade for all three classes of goods, but that the network effect gets progressively stronger as products become more differentiated: we calculate the elasticity for homogeneous goods traded on organized exchanges as 0·21, less than half the value for differentiated products.

Immigrants should also expand trade for *trading relationships* for which information is most valuable. Girma and Yu distinguish 'individual-specific' advantages and 'non-individual-specific' advantages. They argue that individual-specific mechanisms involve the unique advantages that a particular immigrant would have—such as personal contacts and connections to networks. These mechanisms should apply to immigrants from either commonwealth or non-commonwealth countries. In contrast, non-individual-specific mechanisms involve familiarity with legal systems, business systems and communication systems of a foreign country. Girma and Yu hypothesize that these types of mechanisms would not apply very strongly to immigrants from commonwealth countries, since many of the institutions in those countries have British roots. They find a strong immigrant effect for non-commonwealth countries, but no trade benefit from immigrants who come from commonwealth countries, suggesting that the advantage immigrants possess are not individual-specific.

Combes *et al.* study trade within a single country France. Thus all migrants are coming from and going to places with the same legal systems, language, and a variety of other institutional frameworks common to all locations within France. Thus, if Girma and Yu's conclusions were correct, we would expect to see no link between trade and migrant stocks. In fact, Combes *et al.* find that departments import more from departments that have been the origin and destination of migration. This supports the view that migrants carry with them information about specific opportunities rather than mere institutional information.

Head and Ries use Canadian immigration policy as a means to assess the mechanisms through which immigration affects trade. Canada admits three types of immigrants—independent, family, and refugee. The independent gains entry according to a point system based on personal characteristics. Within the independent class are entrepreneurs and investors that can gain enough points for admission by creating jobs or making investments in Canada. Head and Ries examine immigration effects for the family, refugee, entrepreneur, investor, and other independent categories. They find that the 'other' independent class of immigrants, but not the investor or entrepreneur classes, has a significant effect on trade that is larger than the family sponsored immigrant effect. The refugee class effect is significantly lower. These results provide partial support for the proposition that immigrants increase trade due to information advantages. On one hand, we would expect independent immigrants, a highly skilled group likely to have knowledge that might benefit trade, to have larger effects than refugees or family-sponsored immigrants. It is somewhat surprising, however, that relatively large effects are not observed for entrepreneurs. The explanation may be that entrepreneur-class immigrants could enter with lower language skills and education due to the points they were awarded based on their intention to start businesses.

Overall, we observe that the differences in the impact of immigrants on trade across types of goods, classes of immigrants, and trading partner pairs supports a causal relationship running from immigrants to trade. The evidence seems consistent with the information hypothesis of trade. There is conflicting evidence over whether immigrants' information advantages relate to specific trading opportunities or to more general country-level institutions. Surprisingly to us, the existence of a preference effect stimulating trade does not show up in all studies: several authors find export effects that are stronger than import effects.

Specification issues

Reliable statistical tests of immigrants' effects on trade require that regression specifications utilize the proper functional form and control for variables likely to be correlated with the explanatory variables of interest. An incorrect functional form can bias the estimates, and omitted variables that capture forces that promote both trade and immigration levels will lead to an overestimate of the immigrant effect on trade. This section discusses the specifications used in the literature.

Head and Ries, Dunlevy and Hutchinson, Rauch and Trindade, and Girma and Yu employ log-linear specifications that assume a constant elasticity for the effect of immigration on trade. This specification is natural as the basic gravity equation is log linear. However, it may not correspond to theoretical explanations of immigrant effects on trade. Specifically, a constant elasticity implies that a 10% increase in immigrants has the same percentage increase on trade regardless of the level of trade or the number of immigrants in the country. A reasonable depiction of the information hypothesis is that immigrants should have a decreasing marginal effect on trade—the first immigrants may utilize information to complete a large volume of trade whereas later immigrants would have decreasing opportunities to generate new trade.

Gould's specification allows for decreasing marginal effects on trade as the number of immigrants increases. He uses the term in his specification:

$$\rho[\mathrm{IMM}_i/(\theta + \mathrm{IMM}_i)]$$

where ρ and θ are parameters and IMM_i is the number of immigrants from country *i*. Gould argues that ρ measures the size of the immigrant information effects on transaction costs, and θ measures the degree to which there are decreasing returns to immigration. Gould obtains a significant estimate of θ for imports (though not for exports), thereby providing limited evidence for decreasing marginal effects. His estimates reveal that relative to imports, the marginal effect of an additional immigrant on exports is initially high but then diminishes rapidly. Gould's specification, while allowing for decreasing marginal effects, is not derived from theory. In the second half of the paper, we introduce a decreasing marginal effects specification that we derive from a random encounter model.

Numerous factors affect both immigration levels and trade levels. Examples of such factors include distance, historical ties, cultural ties, and overlapping political systems between trading partners as well as individual countries' openness to trade and investment and level of economic development. Since these factors will promote both trade and immigration, failure to control for them will lead to upward bias in estimates of the immigration effect. The broad empirical trade literature serves as a guide to what control variables should be included and all of the papers in the literature include a number of them. There are differences across studies, partly due to data availability.

Even with the inclusion of observable country characteristics, there remains a significant risk that unobserved variables are biasing the results. Thus, omitted variable bias may underlie the large elasticities estimated in the cross-sectional studies. One good way to minimize the distorting effect of unobserved variables is to use country fixed effects. The use of fixed effects, however, comes at the cost of losing the informational content of cross-sectional variation. Gould uses country fixed effects and thus his estimates are based on temporal variation in trade and immigration levels. His tests still produce statistically significant results in favour of the proposition that immigrants expand trade but the magnitude of the effect is smaller than what is estimated in other studies. As shown by Griliches (1986), fixed effect estimation increases the 'noise to information' ratio in the data and exacerbates measurement error, biasing estimates downward. This could explain the relatively small immigration effects reported by Gould. Another drawback of his specification is that, since his specification does not include year dummy variables or a time trend, there is the concern that his results partially capture the simultaneous growth in immigration populations and trade.

In Section III of this paper, we present results of tests that use fixed effects at the *national* level, and use observations at the *provincial* level. We believe that this approach enables us to capture most of the advantages of fixed effects, since the special relationships that affect both trade and immigration likely occur politically at the national level. Yet, by using province-level data, we can still make use of cross-sectional variation and need not rely solely on temporal variation. The tests will still enable us to see whether provinces that draw relatively large numbers of immigrants from a particular country enjoy greater trade with those countries than other provinces enjoy (controlling for economic size, distance, etc.). We also use year dummies to ensure that the simultaneous growth in trade and immigration does not distort our results.

Controls for language commonality have interesting implications for the immigration effect on trade. Girma and Yu, Dunlevy and Hutchinson, and Rauch and Trindade all use a language variable in their statistical tests. How we interpret the language effects depends on how the language variable is constructed, because language may be part of the reason immigrants enjoy trading advantages. Girma and Yu and Dunlevy and Hutchinson use a dummy variable that equals one when both trading partners are English-speaking countries. Girma and Yu consider UK trade and Dunlevy and Hutchinson US trade. Since the ability to speak English does not give these immigrants any advantage over other people in their new country, the English dummy variable in these studies is not measuring an immigrant effect.

Rauch and Trindade construct a more complex language variable by computing the probability that a randomly chosen individual from the exporting nation shares a common language with a randomly chosen individual from the importing nation. This variable may reflect a particular skill that an immigrant or a member of an ethnic group possesses. Consequently, the variable effectively strips out the language effect from the ethnic Chinese effect. The fact that the ethnic Chinese effect exists despite controlling for language shows that there are factors beyond language knowledge that cause countries with large ethnic Chinese populations to trade with one another.

Our results reported in Section III include a variable similar to the one used by Rauch and Trindade reflecting the probability that randomly chosen individuals from two locations would speak the same language. Unlike Rauch and Trindade, we estimate the regression equations with and without this variable to determine the degree to which the immigration effect is attributable to language.

III Results from the International Trade of Candian Provinces

In this section we present new estimates of the effects of immigration on trade to contribute to the existing literature in three ways. First, we use country fixed effects to control for unmeasured characteristics raising immigration and trade between Canada and each of its trading partners. Rather than relying on time-series information, we use the substantial variation in trade and immigration across Canadian provinces to estimate immigrant effects on trade. Second, we employ a common-language variable that measures the degree to which trading partners have languages in common, including minority languages. Third, we experiment with a specification rooted in a random encounter model, to investigate whether such a model provides some insight into the significance of networks.

To carry out these tests, we use 1992–1995 data on each province's trade with each foreign country and on immigration to each province from each country. We use the same data as Wagner (2000), who reports the sources of this data in detail.

Province-level data are supposed to be recorded based on country of consumption but it appears that some of the import trade data is allocated to the wrong province. The imports of some products seem to be allocated to the province of entry. For example, almost all automobiles imported from Europe are reported as Nova Scotia's imports, almost all automobiles imported from the United States show up as Ontario's imports; and almost all automobiles from Japan are included in British Columbia's imports. Most other products appear to be reported properly based on the province of consumption. In the case of exports, trade data are allocated to the province where the goods were produced. We have not noted any industries in which the export trade patterns look unreasonable. We have no practical way of countering the problem with the allocation of imports. All we can do is remain cognizant of these errors when interpreting the results of our statistical tests.

Another problem is that some of the provinces are quite small. To avoid attaching too much weight to the effects of small provinces, we group them into regions. Our groupings follow Statistics Canada's normal groupings of the provinces. The 'provinces' we use are (1) British Columbia; (2) the prairie provinces (Alberta, Saskatchewan and Manitoba); (3) Ontario; (4) Quebec; and (5) the Atlantic provinces (New Brunswick, Nova Scotia, Prince Edward Island and Newfoundland). Throughout this paper, the term 'provinces' refers to these three provinces and two regions. For each foreign country, there are five observations per year. Consequently, each of the 160 foreign countries has 20 observations (five provinces times four years).

In addition to enabling us to use country fixed effects, the use of provincial level data is appealing for another reason. We can consider language variation, since Quebec is predominantly French speaking while the other provinces are predominantly English speaking. Since it is testable, we hypothesize that part of immigrants' advantage in carrying out trade with their countries of origin is knowledge of the language. To carry out this test, we compute a common language variable. For a particular country-province pair, this variable equals the probability that a randomly chosen person from the foreign country and a randomly chosen person from the province can speak the same language. For example, suppose the foreign country is Morocco and the province is Quebec. In Morocco, 20% of the population speaks French and 74% speaks Arabic. In Quebec 92·4% speaks French and 1·1% speaks Arabic. As a result, the common language variable for Morocco-Quebec observations would equal 0·193 ($20\% \times 92{\cdot}4\% + 74\% \times 1{\cdot}1\%$).[3] We consider 30 languages in total.[4]

[3] This formula double counts individuals who speak both languages. We to not adjust for this double counting because we to not have the data necessary to do so. This measurement error is likely immaterial.

[4] Wagner (2000) describes the sources of our language data.

We test the links between immigration and trade using two different models. The first model is a standard linear-in-logs model like those of Head and Ries, Dunlevy and Hutchinson, and Girma and Yu. We then formulate a second, random-opportunity model based on the theoretical reasoning described in previous sections.

For each specification, we run six tests—three involving exports and three on imports. Out of the three tests, the first test considers neither country fixed effects nor language effects. The next test considers fixed country effects, but still does not consider language effects. The last test constitutes the full model—with fixed country effects and language effects.

Constant elasticity (linear-in-logs) specification

Our specification is as follows:

No fixed country effects: Reported in columns 1 and 4 of Table 2

$$\begin{aligned}\ln T_{pf} = {} & \beta_1 \ln \mathrm{IMM}_{pf} + \beta_2 N\mathrm{ID}_{pf} + \beta_3 \ln(Y_p Y_f / Y_w) + \beta_4 \ln D_{pf} \\ & + \beta_5 \ln(R_f) + \beta_6 \mathrm{MILLS}_{pf} + \beta \Lambda_{pf} + \varepsilon_{pf}\end{aligned}$$

With fixed country effects: Reported in columns 2 and 5 of Table 2

$$\begin{aligned}\ln T_{pf} = {} & \beta_1 \ln \mathrm{IMM}_{pf} + \beta_2 \mathrm{NID}_{pf} + \beta_3 \ln(Y_p Y_f / Y_w) + \beta_4 \ln D_{pf} \\ & + \beta_5 \ln(R_f) + \beta_6 \mathrm{MILLS}_{pf} + \beta \Lambda_{pf} + \mathrm{FE}_f + \varepsilon_{pf}\end{aligned}$$

With fixed country effects and language effects: Reported in columns 3 and 6 of Table 2

$$\begin{aligned}\ln T_{pf} = {} & \beta_1 \ln \mathrm{IMM}_{pf} + \beta_2 \mathrm{NID}_{pf} + \beta_3 \ln(Y_p Y_f / Y_w) + \beta_4 \ln D_{pf} \\ & + \beta_5 \ln(R_f) + \beta_6 \mathrm{MILLS}_{pf} + \beta \Lambda_{pf} + \mathrm{FE}_f \\ & + \beta_7 \mathrm{COMLANG}_{pf} + \varepsilon_{pf}\end{aligned}$$

where:

T_{pf} = exports from province p to country f (for test involving exports), or province p's imports from country f (for tests involving imports),

IMM_{fp} = the number of immigrants from foreign country f living in province p,

$\ln \mathrm{IMM}_{fp}$ = is overridden to equal 0 if the number of immigrants is less than or equal to 5,

NID_{fp} = 'no immigrant dummy' and equals 1 if the number of immigrants in province p from country f is no more than 5, but equals 0 if the number of immigrants exceeds 5,

Y_p, Y_f, Y_w = the GDPs of province p, country f and the world,

D_{ps} = the distance between the province and the foreign country,

R_f = the remoteness of country f,

$\beta =$ the row vector of coefficients for the respective elements in Λ_{pf},
$\Lambda_{pf} =$ the column vector of dummy variables for provinces, years and the constant,
$\text{MILLS}_{fp} =$ Mills' ratio,
$\text{FE}_f =$ fixed effect for country f, and
$\text{COMLANG}_{fp} =$ the common language variable described above.

The 'no immigrant dummy' is needed to handle cases where there are no immigrants. If $\text{IMM}_{fp} = 0$, then ln IMM_{fp} becomes undefined. In that case, NID_{fp} takes a value of one and ln IMM_{fp} is set equal to zero. A province is considered to have no immigrants if there are five or less, because Citizenship and Immigration Canada altered some figures by plus or minus five to protect the anonymity of the immigrants in the data set. The remoteness variable measures a country's distance from world markets. The more remote a country is, the more one would expect that country to trade with a partner of a given distance, because the remote country tends to have less trade diverted away to closer countries. In this paper, the remoteness variable for country f (i.e. R_f) equals $1/\Sigma_i[(Y_i/Y_w)/D_{if}]$. When $i = f$, a measure for a county's distance from itself is needed. We follow Leamer (1997) and Nitsch (2000) and assume internal distance $D_{ff} = (\text{AREA}_f/\Pi)^{1/2}$. The statistical tests in this paper all use a Heckman procedure. Mills' Ratio (MILLS_{fp}) is computed based on the probability that the observation has a positive trade amount. That probability comes from a probit equation with the same independent variables (except MILLS_{fp}) as the second stage regression.

Table 2 reports the estimated coefficients along with their robust standard errors.[5] The estimates for ln IMMD_{fp} are similar to those of Head and Ries and Dunlevy and Hutchinson. For exports, we obtain a coefficient of 0·156 when fixed effects are not considered (column 1), which is somewhat above the aforementioned studies' results of 0·099 and 0·08. When we introduce fixed country effects, the coefficient estimate decreases to the 0·081 to 0·089 range (columns 2 and 3). These immigration effects are statistically significant at a 99% significance level without country fixed effects, and at the 90% to 95% level with country fixed effects. The coefficient for ln IMM_{fp} can be interpreted as an elasticity. A coefficient of 0·08 implies that a 10% increase in immigrants from a country will be associated with a 0·8% increase in exports to that country.

We now turn to imports. Without fixed country effects we estimate the coefficient for ln IMM_{fp} at 0·413 (column 4). This result is again slightly above Head and Ries' result of 0·309 and Dunlevy and Hutchinson's result of 0·29. The introduction of fixed country effects reduces our coefficient estimate to about 0·25 (columns 5 and 6). These coefficient estimates are statistically significant at the 99% level even with fixed effects and even using robust errors. A coefficient of 0·25 implies that a 10% increase in immigrants from a country is associated with a 2·5% increase in imports from that county.

[5] We use robust/cluster feature of STATA to obtain robust errors.

TABLE 2
Linear in lcgs specification

	Exports			Imports		
	No fixed country effects and no language effects	With fixed country effects but no language effects	With fixed country effects and language effects	No fixed country effects and no language effects	With fixed country effects but no language effects	With fixed country effects and language effects
ln IMM_{fp}	0·156	0·089	0·081	0·413	0·272	0·251
	(0·028)***	(0·040)**	(0·045)*	(0·055)***	(0·081)***	(0·086)***
NID_{fd}	−0·16	0·35	0·34	1·43	0·72	0·71
	(0·30)	(0·30)	(0·30)	(0·58)**	(0·55)	(0·54)
COMLANG			0·16			0·43
			(0·28)			(0·69)
ln $(Y_p Y_f / Y_w)$	0·96	0·50	0·50	1·07	−0·05	−0·05
	(0·04)***	(0·24)**	(0·24)**	(0·05)***	(0·39)	(0·39)
$\ln(D_{ei})$	−1·18	−1·85	−1·85	−1·05	−1·43	−1·44
	(0·13)***	(0·18)***	(0·18)***	(0·20)***	(0·31)***	(0·31)***
$\ln(R_f)$	0·09	−1·72	−1·73	0·12	−0·27	−0·27
	(0·12)	(0·83)**	(0·83)**	(0·14)	(0·83)	(0·83)
MILLS	0·31	−0·22	−0·22	−0·21	0·89	0·89
	(0·32)	(0·26)	(0·26)	(0·82)	(0·65)	(0·65)
Adj. R^2	0·748	0·842	0·842	0·647	0·785	0·785
Root MSE	1·51	1·23	1·23	2·56	2·06	2·06
No. of obs.	2556	2556	2556	2545	2545	2545

Notes:
Figures in parentheses are robust estimated errors (using the robust/cluster feature in STATA.) Estimates for the dummy variables and the constant are not shown.
*, ** and *** denote 90%, 95% and 99% confidence levels.

Columns 3 and 6 introduce language effects. For both exports and imports, language effects are not statistically significant. This finding surprised us, since our prior expectation was that language differences pose a major barrier to trade. This result differs from the findings of other researchers using the gravity model, who find that common languages significantly increase trade levels. We carried out some other tests and found that our insignificant language results are restricted to the case where the regression includes both fixed effects and immigration. If we remove fixed effects from the regression, the common language variable is statistically significant with a coefficient of 0·88 and a standard error of 0·17 when predicting for exports, and with a coefficient of 0·95 and a standard error of 0·28 when predicting for imports. (These results are not reported in a table.) If instead we retain fixed effects but remove immigration variables, then the common language variable bears some prediction value, albeit weakly. In export predictions, the coefficient is 0·37 with a standard error of 0·23, making the variable almost statistically significant at the 90% level; in import predictions, the coefficient is 1·17 with a standard error of 0·66, making the variable statistically significant at the 90% level. The lower coefficient for exports may be attributable to Canada's high export levels of homogenous goods, where common languages may be less important to trade. These results *may* be suggesting that the immigration effect dominates the language effect. Thus the prime benefit of immigrants might be their access to networks rather than their knowledge of foreign languages.

Within province-country pairs of trading partners, residuals from year to year are correlated with each other. Regressing residuals on one-year-lagged residuals yields a coefficient of 0·78. Switching the independent variable to a two-year lag yields a coefficient of 0·71 and a three-year lag yields 0·68. All these residual regressions are highly significant. Since these coefficients decline somewhat as the years grow farther apart, there appears to be some serial correlation and possibly some other correlation amongst observations involving the same trading partners. These observations underscore the importance of reporting robust standard errors that accounts for correlated errors across trading partner pairs. The robust errors are nearly double the estimated errors produced by a normal OLS regression.

Many researchers include the lagged dependent variable as an independent variable. This variable is recommended by Eichengreen and Irwin (1996), who contend that the gravity model is often used without adequate care for considering omitted variables, which can lead to erroneous conclusions. Omitted variables can involve persistent errors for a pair of trading partners or can involve historical relationships. Eichengreen and Irwin conclude that the gravity model should always have the lagged dependent variable as a regressor. However, missing variables can often be handled in other ways that permit a more natural interpretation of the coefficient estimates. In this paper we use fixed country effects to capture omitted variables and adjust our standard errors to reflect correlation across cross sectional units (province-county pairs).

For comparison, we introduced the lagged dependent variable to check the robustness of our results reported in Table 1. In the case of exports without fixed effects, the coefficient on ln IMM_{fp} remains statistically significant at the 99% level, but decreases from 0·156 (reported in Table 1) to 0·058 (not reported in a table). As noted earlier, using the lagged dependent variable allows immigration to affect trade in two ways directly and indirectly. The direct effect comes from the immigration variable itself while the indirect effect comes through the lagged dependent variable, because immigration affected prior years' trade, which affects the current year's trade. The total effect of immigration is $\beta/(1-\lambda)$, where β is the coefficient estimate for the immigration variable and λ is the coefficient estimate of the lagged dependent variable. In this case $\beta = 0{\cdot}058$ and $\lambda = 0{\cdot}696$, producing a total effect of 0·191, which is somewhat higher than the coefficient estimate reported in Table 2. With imports, the coefficient estimate for immigration again remains statistically significant and the computed total elasticity of immigration amounts to 0·492, which is also somewhat higher than the 0·413 reported in Table 2. The use of the lagged dependent variable continues to yield similar results that are statistically significant.

Random-opportunities model

We now introduce a new specification based on the proposition that immigrants are endowed with trading advantages, but that these opportunities become exhausted as the immigrant population grows. We assume there are a multitude of trading opportunities between a country and a province. The size of each trading opportunity is assumed to be proportional to the GDP of the exporting province (or country) and proportional to the GDP of the importing country (or province). These trading opportunities can be divided into two classes—the easy opportunities and the hard opportunities. There are a times as many hard opportunities as easy opportunities. The easy opportunities do not require the facilitation of an immigrant. We assume that all of the easy opportunities get exhausted.

The hard opportunities require the facilitation of an immigrant. The immigrant needs to have skills, knowledge or connections for a particular industry both in the country of origin and in the Canadian province. For a specific hard trade opportunity, one immigrant has a probability p of being able to facilitate the trade. So the probability that there exists at least one immigrant who can facilitate a potential hard trade opportunity is $1-(1-p)^{\mathrm{IMM}pf}$. This specification implies that there are diminishing returns on immigration.

Therefore, the full model is:

$$T_{pf} = [(Y_p Y_f / Y_w)^{\beta 1} D_{pf}^{\beta 2} R_f^{\beta 3} e^{\beta 4 \mathrm{COMLANG} + \mathrm{FE}_f + \beta \Lambda + C}] \times [1 + \alpha(1-(1-p)^{\mathrm{IMM}pf})] e^{\varepsilon_{pf}}.$$

Since this specification cannot be transformed into a linear relationship, it is solved iteratively using a maximum likelihood estimation.

Table 3 reports the results of the statistical tests employing this specification. The results provide evidence that many immigrants do possess trading advantages that they exploit. Both variables (α and p) for this trade-expanding mechanism produce statistically significant coefficients at the 90% level or better in all six columns. Had we been able to compute robust errors, some of the results of the tests with fixed effects (reported in columns 2, 3, 5 and 6) would likely lose their significance. The results of the tests without fixed effects (reported in columns 1 and 4) would retain their significance.

For exports, we estimate α to be about 0·77 (in the fixed effects model). This implies that there are fewer hard trading opportunities than easy opportunities. The estimate for p is 0·00048, which implies that a province needs 1,444 immigrants from a country to achieve half of its potential hard exports and 4,796 immigrants to achieve 90% of its potential hard exports.[6]

In the import regressions, we estimate a larger α and a smaller probability, p. With fixed country effects, estimates for α range between 3·00 and 3·55. Thus α is about four times larger for imports than exports. The estimate for the imports p is about 1/3 of the size of that for exports. The estimate of p of 0·0015 implies that a province needs 4,621 immigrants from a country to realize half of its potential hard exports, and 15,350 immigrants to achieve 90% of its potential. Like Gould, we find that the benefit of new immigrants tapers off more quickly on the export side than on the import side. The reason for this outcome may be

[6] $1-(1-0{\cdot}00048)^{1444} = 0{\cdot}5$; $1-(1-0{\cdot}00048)^{4796} = 0{\cdot}9$.

TABLE 3
MLE of 'Random opportunities' specification

	Exports			Imports		
	No fixed country effects and no language effects	With fixed country effects, but no language effects	With fixed country effects and language effects	No fixed country effects and no language effects	With fixed country effects, but no language effects	With fixed country effects and language effects
α	1·85	0·77	0·73	13·27	3·55	3·00
	(0·31)***	(0·25)***	(0·24)***	(3·12)***	(1·34)***	(1·18)**
P	0·00179	0·00048	0·00049	0·00045	0·00015	0·00016
	(0·00051)***	(0·00025)**	(0·00026)*	(0·00011)***	(0·00007)**	(0·00007)**
COMLANG			0·19			0·61
			(0·22)			(0·38)
$\ln(Y_p Y_f / Y_W)$	0·96	0·49	0·49	1·08	0·07	0·07
	(0·02)***	(0·24)**	(0·24)**	(0·03)***	(0·40)	(0·40)
$\ln(D_{pf})$	−1·19	−1·85	−1·85	−1·09	−1·46	−1·47
	(0·07)***	(0·12)***	(0·12)***	(0·12)***	(0·20)***	(0·20)***
$\ln(R_f)$	0·07	−1·78	−1·78	0·08	−0·13	−0·12
	(0·06)	(0·93)*	(0·93)*	(0·09)	(1·55)	(1·54)
σ^2	2·28	1·42	1·42	6·53	3·99	3·99
	(0·06)***	(0·04)***	(0·04)***	(1·18)***	(0·11)***	(1·11)***
Log Likelihood	−2330·9	−1726·0	−1725·6	−3659·3	−3033·3	−3032·0
No. of obs.	2556	2556	2556	2545	2545	2545

Notes:
Figures in parentheses are estimated errors. (Unlike the log linear specification reported on in Table 1, the robust/cluster cannot be used in this maximum likelihood estimation). Estimates for the dummy variables and the constant are not shown.
*, **and ***denote 90%, 95% and 99% confidence levels.

attributable to a persistent preference effect operating to increase imports even after opportunities for new trades based on market knowledge have been exhausted.

As in the previous tests, language effects are not statistically significant, and do not materially affect the immigration coefficients. This outcome implies that the strength of immigrants' trading advantages do not come from language.

We use the estimates in Table 3 to calculate the marginal effect on trade of additional immigrants in 1995 and their corresponding trade elasticity.[7] The computed elasticities are 0·013 for exports and 0·092 for imports. The export figure is below the 0·021 elasticity that we derived for Gould and is well below our earlier result of 0·081 (Table 2, column 3). The import elasticity is well above the Gould result of 0·013, but well below our previous estimate of 0·251 (Table 2, column 6). We estimate an additional immigrant in Canada beyond the 1995 level would raise exports by $312 and imports by $944.

IV Conclusion

Our survey indicates there are a number of aspects of the trade-immigration relationship that are supported in the empirical literature. First, there is a positive association between the two variables that seems robust across very different samples and econometric methods. Second, the magnitude of the immigration effect varies in sensible ways across samples, groups of immigrants, and products.

The studies also shed light on the mechanisms that underlie immigrants' effects on trade. Head and Ries, Dunlevy and Hutchinson, and to some extent Gould all produce evidence that supports the contention that part of the growth in imports can be attributed to immigrants' preferences for goods from their native countries. However, Girma and Yu, and Combes *et al.* do not detect such a preference effect. There is strong evidence that immigrants possess information that they use to facilitate trade. Information would have its greatest value for differentiated goods, and Gould, Rauch and Trindade, and Dunlevy and Hutchinson all find this to be true in their studies. Gould and Rauch and Trindade find that immigrants also promote trade in homogenous goods, a result consistent with alternative hypotheses such as that immigrant networks facilitate contract enforcement. Girma and Yu's result that immigration does not expand trade for commonwealth countries suggest to them that the immigrant advantage does not reflect knowledge that is idiosyncratic to an individual. Our tests find that the explanatory effect of language disappears in a test with immigrants and country-level fixed effects. Combes *et al.* find that migrants influence trade even within a nation.

[7] We compute the increase in trade associated with a 1% increase in immigrants from each source country. The percentage change in trade that this number serves as an estimate of the trade elasticity. Dividing the increase in trade for a 1% increase in immigration by 1% of the immigrant stock yields an estimate of the marginal effect of an additional immigrant.

Of more uncertainty in the literature is the magnitude of the effect of immigration on trade. The diversity of data sets may explain a portion of the discrepancies in results, but most of the variation is likely attributable to differences in specification. The most important issue concerns the handling of unmeasured variables. Our use of province level data and national-level fixed effects is a reasonable way of addressing this problem. Most of the unmeasured factors likely occur at the national level, and are therefore captured by the country fixed effects.

We use country fixed effects and exploit the considerable variation in trade and immigration levels across provinces to estimate the effect of immigration on trade. In our constant-elasticity specification, we obtain results similar to those of studies that omit country-fixed effects. A second specification issue is whether to allow immigration to have a decreasing marginal effect on trade. We employ a decreasing marginal effect specification that is similar to the one used by Gould and, like him, we obtain smaller immigration effects. These results indicate to us that future research should continue to explore specification issues in order to increase our confidence in estimates of magnitude of the effect immigrants exert on trade.

References

Combes, P.-P., Lafourcade, M. and Mayer, T. (May 16, 2002 draft). Can business and social networks explain the border effect puzzle?

Dunlevy, J. A. and Hutchinson, W. K. (1999). The impact of immigration on American import trade in the late nineteenth and early twentieth centuries. *Journal of Economic History*, pp. 1043–62.

Dunlevy, J. A. and Hutchinson, W. K. (2001). The pro-trade effects of immigration on American exports during the period 1870 to 1910. Vandabilt University Working Paper 01-W25.

Eichengreen, B. and Irwin, D. A. (1997). The role of history in bilateral trade flows. In J. A. Frankel (ed.), *The Regionalization of the World Economy*. Chicago: University of Chicago Press.

Frankel, J. and Rose, A. K. (2002). An estimate of the effect of currency unions on trade and growth. *Quarterly Journal of Economics*.

Girma, S. and Yu, Z. (2002). The link between immigration and trade: evidence from the UK *Weltwirtschaftliches Archiv*.

Gould, D. M. (1994). Immigration links to the home country: empirical implications for U.S. bilateral trade flows. *The Review of Economics and Statistics* 76, 2, pp. 302–316.

Griliches, Z. (1986). Economic data issues. In Z. Griliches and M. D. Intriligator (ed.), *Handbook of Econometrics, Volume III*. Elsevier Science Publishers BV, Amsterdam.

Head, K. and Ries, J. (1998). Immigration and trade creation: econometric evidence from Canada. *Canadian Journal of Economics*, pp. 47–62.

Helliwell, J. F. (1999). Language and trade. In A. Breton (ed.), *Exploring the Economics of Language*. Canada: Canadian Heritage.

Leamer, E. (1997). Access to western markets and eastern effort levels. In S. Zecchini (ed.), *Lessons from the Economic Transition: Central and Easter Europe in the 1990s*. Kluwer Academic Publishers, Boston.

Nitsch, V. (2000). National borders and international trade: evidence from the European Union. *Canadian Journal of Economics*, 22, 4, pp. 1091–105.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, 48, pp. 7–35.

Rauch, J. E. and Trindade, V. (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics*.

WAGNER, D. (2000). Essays on the mobility of goods and people. Dissertation at the University of British Columbia.
WAGNER, D. (forthcoming). Aid and trade—an empirical study. *Journal of the Japanese and International Economies*.

Date of receipt of final manuscript: 3 July 2002.