A Threshold Model for the Migration-Trade Link

Hisham Foad*

June 2009

Abstract

How are immigration and trade related? The existing literature has found a positive effect of migration on trade, with migration working through both network and preference channels to increase international trade. I extend this analysis by examining this relationship across different levels of both migration and economic development. Migration from poor countries has a larger effect on trade than that from rich countries, reflecting the weaker trading institutions in poor countries that migrant networks may replace. For poor country migration, the marginal impact of migration is largest at high levels of migration. For rich country migration, this marginal effect declines as migration increases, suggesting that local production may be replacing native country exports. Sample splits are endogenized using Hansen's (2000) technique and greater support is found for a regime change across income levels. These results are relevant from both an academic and policy perspective as they suggest that the existing literature has underestimated the impact of poor country migration on trade, a key result when discussing migration policy.

1 Introduction

How are immigration and trade related? A growing literature in economics has examined this question finding compelling evidence for a positive and causal relationship between cross-border migration and international trade. These studies have focused on two explanations for the migration-trade link: network effects and transplanted home bias. When people immigrate, they bring with them not only their own labor but also knowledge about and connections with their native country. As limited information about foreign markets can raise the cost of trade, the formation of cross-border networks through immigration should lower these trade costs. Immigrant networks may also help enforce contracts that would otherwise be difficult to enforce across borders. Grief (1989.

*Department of Economics, San Diego State University, San Diego, CA, USA 92182. Please direct all correspondence to the author at hfoad@mail.sdsu.edu.

1993), for example, cites the Maghribi traders coalition, a distinct social group within the Jewish diaspora living in the 11th century. Trade costs within the coalition were low because the traders shared information and were willing to engage in collective punishment of any members who breached their contracts. Rauch and Trindade (2002) find similar evidence of the importance of ethnic networks on trade, with countries having large ethnic Chinese populations engaging in brisk trade with one another.

At the same time, there is a well documented home bias in trade, with more trade within a country than across borders. For example, McCallum (1995) and Helliwell (1996) find that Canadian provinces trade approximately ten times more with other provinces than they do with equidistant and similarly sized US states. Wolf (2000) finds that home bias even exists within US states, suggesting that this phenomenon is not only being driven by formal and informal trade barriers, but perhaps a local preference effect. If indeed home bias is driven by a preference for local (familiar) goods, then immigration may increase international trade by increasing the population within a country who have a preference for foreign goods (familiar to the immigrants).

The existing literature has generally treated the link between migration and trade as log-linear. For example, White (2007) follows the convention in earlier studies and uses a gravity equation in which each country in the world consumes its own output as well as the output of every other country in the world. Define bilateral imports as:

$$Imports_{i,j} = s_{i,j} y_j \tag{1}$$

Imports into country i from country j are equal to the share of country j's output consumed by country i, where y_j is country j GDP. If there is balanced trade between countries i and j and no trade barriers, then $s_{ij} = y_i / y_w$, where y_w is world GDP. For example, if the US represents 20% of world GDP ($y_{US}/y_w =$ 20%) then we would expect the US to consume 20% of every country's output. Obviously we do not live in a fully integrated world with no trade barriers, so the assumption that $s_{i,j} = y_i / y_w$ is too restrictive. Rather, define the bilateral trade share as:

$$s_{i,j} = \frac{y_i}{y_w} * \tau_{i,j} \tag{2}$$

The bilateral trade share is equal to country i's share of world demand multiplied by $\tau_{ij} = exp\{Z_{i,j}\}$, a measure of trade inhibiting and trade promoting factors. For example, if $\tau_{i,j} > 1$, then country i actually consumes relatively more of j's output than would be predicted in a perfectly integrated world. If $\tau_{i,j} < 1$, then country i is underweight in country j's output. A common estimation technique has been to plug equation 2 into 1 and solve for:

$$ln(Imports_{i,j}) = ln(\frac{y_i y_j}{y_w}) + Z_{i,j} \tag{3}$$

Trade between countries i and j will be a function of the relative market sizes of countries i and j as well as a vector of trade inhibiting and trade promoting factors. Examples of trade inhibiting factors include distance (a proxy for transportation costs), whether or not countries i and j have a history of military conflict, or if countries i and j have any sanctions levied against them. Trade promoting factors include whether or not countries i and j share a common language, colonial history, or border; are members of a free trade area; and how much migration has occurred between these nations.

According to the two channels discussed above, immigration should cause trade flows to increase by strengthening cross-border networks and through transplanted home bias. The existing literature has estimated significantly positive coefficients when regressing log bilateral trade on log bilateral migration, confirming this theory. However, the relationship between migration and trade may not be linear or even monotonic. My paper allows the effect of migration on trade to differ across both economic development and the level of migration. For example, migration may not affect trade until a certain threshold level of migration has been reached. Alternatively, if an immigrant community becomes large enough, any transplanted home biases may be served through local production and FDI rather than exports from the immigrants' native countries. The failure to account for these potential non-linearities in the existing literature has led to biased estimates of the true migration-trade elasticity.

2 Literature Review on the Migration-Trade Link

One of the first empirical studies to document complementarity between migration and trade was Gould (1994), who inserts the lagged immigrant stock from a range of countries into a gravity equation estimating exports and imports separately for the United States. He finds that a 10% increase in immigration to the US raises exports to the immigrants' native country by 4.7% and imports from the native country by 8.3%. That imports are more sensitive to immigration reflects the fact that immigrants retain some preference for native country goods (the transplanted home bias channel). Head and Ries (1998) follow up on Gould's work using Canadian data and estimate the relationship with a Tobit model to account for the large number of zeros in bilateral trade data. Though they also find a positive relationship between immigration and trade, their elasticity estimates are lower. Head and Ries argue that the Canada's trade is primarily in commodities or in US bound automotive goods which do not significantly benefit from the migration-induced reduction in transaction costs. Girma and Yu (2002) find additional evidence for the transaction costs explanation by examining immigration and trade between the UK and 48 trading partners. They find that the migration-trade link is significantly positive only for countries that are not former British colonies. They argue that the trade promoting effects of immigration (contract enforcement, more symmetric information, etc.) do not matter so much for former colonies, as these nations already share many similarities with the UK (such as in the legal system) and information about foreign markets is plentiful.

Several studies have looked at the composition of trade, finding that immigration most strongly promotes trade in differentiated goods. Rauch and Trindade (2002) show that the presence of an ethnic Chinese network in a country significantly increases differentiated goods trade with China. Further evidence of the relatively stronger effect on differentiated goods is given by Dunleavy and Hutchinson (1999) looking at immigration and trade in the late 19th century as well as Blanes (2005) and White (2007) who look at the effects of immigration on intra-industry trade for Spain and the US respectively. These studies present evidence that labor migration induces international trade, most notably for differentiated goods.

My paper adds to the literature in several ways. First, my paper is the first I am aware of to allow the migration-trade link to differ at different levels of migration. I also allow the threshold at which the migration-trade link changes to be endogenous, using Hansen's (2000) sample splitting and threshold estimation technique. This methodology is less restrictive than the existing literature which imposes a log-linear monotonicity constraint on the migration-trade elasticity. My data-set brings together a much wider range of countries than has been used in previous studies, allowing for greater variation across income levels. Finally, I address the issue of potential endogeneity in the migration variable, something that has not consistently been done in the existing literature.

3 Non-Linearities in the Migration-Trade Link

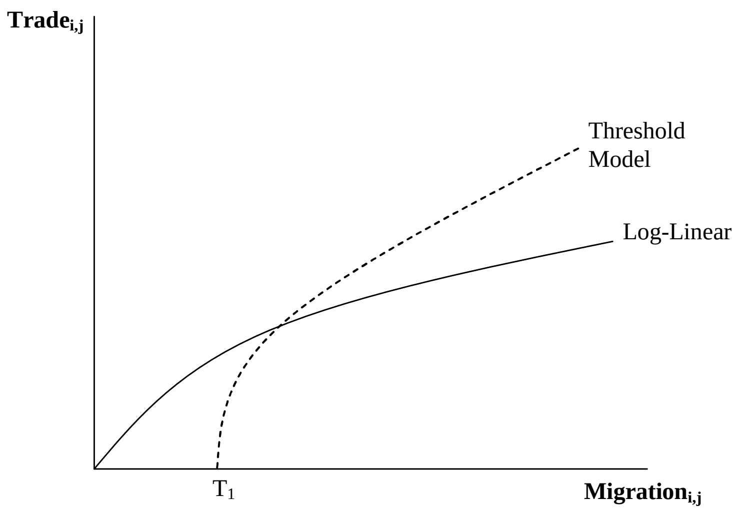
3.1 A minimum threshold for migration and trade

Theoretically, migration is predicted to affect trade through two channels. Migration strengthens cross-border network effects, reducing both the cost and risk of international trade. Second, migration may lead to a transplanted home bias in which the migrants retain a preference for their native country's goods. In both of these cases, the effect of migration on trade may actually vary across the level of migration.

For example, suppose that a firm located in country j will only export to country i if doing so yields positive profits: $\Pi_{i,j} = R_{i,j} - C_{i,j} > 0$. In this case, $R_{i,j}$ and $C_{i,j}$ are the firm's revenues and costs (both fixed and variable) from exporting and are both functions of variables specific to country j as well as migration between countries i and j. We expect that $R_{i,j}$ will increase with migration as people moving from country i to j will create a built in market for the exporter in j to sell to. At the same time, $C_{i,j}$ should be a decreasing function of migration, as greater migration should increase information flow and lead to greater contract enforcement between countries i and j.

Holding all other variables constant, exporting is more likely to be profitable as migration increases. At low levels of migration, we may simply see zero trade as exporters choose not to take on the fixed costs of market entry. In fact, the data supports this hypothesis as there are a large number of zeros in bilateral trade data. Holding all else constant, there is some threshold level of migration at which trade becomes profitable. Below this threshold, there is no trade

Figure 1: Threshold Migration Model



and the true migration-trade elasticity is zero. Above this threshold, migration should positively affect trade through both network and preference effects.

Figure 1 displays two hypothetical relationships between migration and trade. The solid line is the simple log-linear specification used in the existing literature, while the dashed line represents the threshold model discussed above. In the hypothetical relationship depicted in Figure 1, trade will not occur between countries i and j until a threshold level of migration T_1 has been reached. At levels of migration less than this threshold, there will be no trade. This suggests that the existing literature (log-linear specification) has overestimated the migration-trade elasticity at low levels of migration and underestimated its effect at high levels of migration. The larger the migration-trade threshold (T_1), the larger the bias in the log-linear model.

The discussion above assumes that all other variables are held constant. Obviously there are many other variables that will affect trade, with migration being of secondary importance when compared to income and proximity. As migration becomes less of a factor in determining trade flows, the threshold level of migration at which trade will occur becomes smaller and smaller. If migration, however, is a more economically significant determinant of trade flows, then addressing this potential non-linearity becomes more important.

3.2 A negative relationship between migration and trade?

Nearly all studies looking at migration and trade have found a positive relationship. Some studies that have broken trade down by commodity have found no relationship for more homogeneous goods, but no study that I am aware of has found a negative effect of migration on trade. I argue that this is not necessarily immutable proof that migration positively affects trade, but rather that migration below a certain threshold positively affects trade. Above this threshold, we may actually see a reduction in trade brought on by migration.

The basic intuition for this result is that when an immigrant community becomes large enough within a country, the demand in this community for native country goods and services will be met by local production (whether by domestic firms or FDI) rather than imports from abroad. For example, consider the sizable Turkish community in Germany. As this community first started to grow in the 1960's, demand by the migrants (primarily guest workers) for Turkish goods was met by imports from Turkey. Over time the community expanded, gained greater permanence, and we observed the formation of German firms that produced goods and services specifically targeted to the Turkish community (often founded and run by the Turkish immigrants themselves.) Local production became a substitute for imports, suggesting that at high enough levels of migration, there may actually be a decline in trade with the native country.

This result fits with both channels through which migration is expected to affect trade. Immigrant network effects should increase FDI as the information flow and contract enforcement generated by these networks increase the expected profitability of direct investments.[1] If FDI and immigration are substitutes, then we may see a reduction in trade if migration benefits FDI more than trade at high levels. This may indeed be the case when we consider the preference channel. If immigrants retain a preference for their native country's goods, then more immigrants will lead to more demand for native country goods. Considering that starting a new business or shifting operations to meet the demands of the immigrant community incurs large fixed costs, we should not observe local production until the immigrant community is large enough to generate net revenues in excess of these fixed costs.

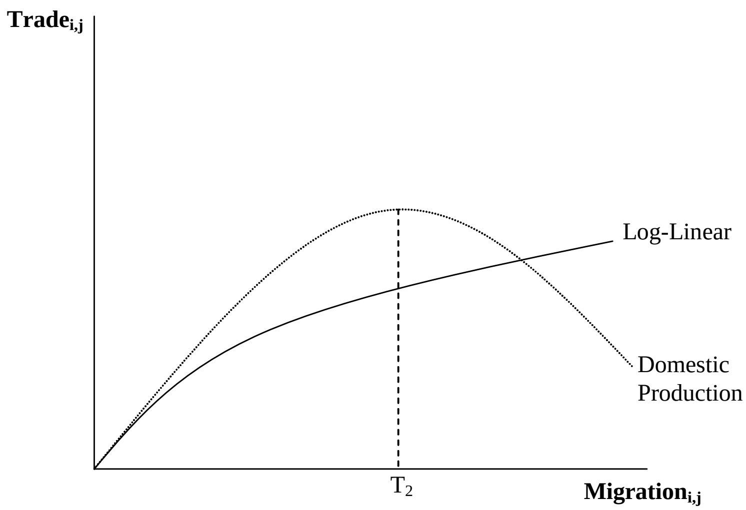
Figure 2 depicts the relationship described above alongside the log-linear model. Below a certain threshold given by T_2, the log-linear and domestic production models look very similar. However, at this threshold level of migration, we start to see domestic firms producing specifically for the immigrant market. This presumably eats into the market share of exporters and we may observe a decrease in trade (though a potential increase in FDI). If this is indeed the case, then the existing literature may in fact be overestimating the true migration-trade elasticity at high levels of migration.

3.3 Differences across development levels

Another potential source of variation in the migration-trade elasticity is the development level of both the sending and receiving countries. In other words, migration from a rich country to another rich country may have a different impact on trade than migration from a poor country to a rich country. This

[1] See Foad (2008), Javorcik et. al. (2006), and Kugler and Rapoport (2005) for examples of the positive relation between migration and FDI

Figure 2: Local Production Threshold



difference may arise from the migrants themselves or from the kind of trade occurring between nations. On one hand, migrants coming from poorer nations tend to have lower incomes, suggesting a lower per capita effect through the preference channel. On the other hand, migrants from these poorer nations may be less assimilated into their new homes, in which case retained home bias may be even stronger despite lower migrant income.

Rauch (1999) finds that impact of migration on trade is strongest for differentiated goods and weakest for commodities. This result fits with both the network and preference channels through which immigration may affect trade. The quality of differentiated goods may be difficult to ascertain across borders, so the greater information flow and contract enforcement provided by migrant networks will facilitate trade in these goods. Furthermore, differentiated goods are clearly the kind of goods that immigrants may retain a preference for. Homogeneous goods, by definition, should be of more uniform quality regardless of where they were produced, and thus less likely to benefit from migrant networks or retained home bias. This result matters if trade patterns differ by development level. As numerous studies have documented, trade between developed countries (north-north) better fits the new trade theory models based on returns to scale and product differentiation, whereas trade between developed and developing countries (north-south) is more likely to involve trade in homogeneous goods. If this is indeed the case, then we should observe a larger migration-trade elasticity for developed countries.

4 Empirical Methodology and Results

4.1 Data Sources

The two key variables in this analysis are trade and migration. Bilateral trade data is from the NBER-United Nations trade database provided by Feenstra and Lipsey (2005), while bilateral migration is from Docquier and Marfouk (2007). Bilateral migration data is only available for two years, 1990 and 2000, so all estimates will be made with cross-sectional data. To address potential endogeneity, migration in 2000 is instrumented with 1990 migration and a host of other variables, to be explained in further detail below. The migration database covers the birthplace of the immigrant stock in 29 OECD countries from nearly all source countries in the world. After eliminating countries with missing observations or minimal migration (generally tiny Oceanic nations), I compile a data-set with 29 destination countries (OECD nations) and 163 source countries. This yields 4,698 observations after eliminating common sending and receiving countries. A list of countries included in the sample is given in Table 1.

Proximity variables such as distance, common border, common language, and colonial ties are from the CEPII Geodesic Distance database. Bilateral distance is defined as the distance in kilometers between city centers. Two countries are considered to share a common language if at least 10% of the population in each country speaks the same language.[2] Data on armed conflict is from the Upssala Conflict Data Program. The conflict variable is equal to 1 if two countries were in armed conflict against one another (either alone or as part of a coalition) at any time between 1980 and 2000.[3] Finally, membership in a free trade area is taken from the World Trade Organization.

4.2 Benchmark Model

The existing literature has generally estimated the migration-trade elasticity with a gravity equation. This is essentially an estimate of equation 3 above, with various trade inhibiting and promoting variables (including migration) augmenting the relationship between trade and country size. They regress log trade between countries i and j (where trade may be total trade, imports, or exports) on log GDP in countries i and j and a vector of trade inhibiting and promoting variables. Past studies have also found log per capita GDP to be a significant determinant of trade flows, so in keeping with convention, I include those as well. The benchmark model to be estimated is given by:

$$\begin{aligned} lnTrade_{i,j} &= \alpha + \beta * lnMig_{i,j} + \gamma_{1*} ln(GDP_i * GDP_j) + \\ &\quad \gamma_2 * ln(PGDP_i * PGDP_j) + Z_{i,j}\Phi + u_{i,j} \end{aligned} \tag{4}$$

[2] A more conservative definition using only the official language was also used, but the results did not change

[3] These are relatively arbitrary cutoff dates, but moving the window forward or back does not appreciably change the results.

The vector of trade inhibitors/promoters is given by $Z_{i.j}$ and encompasses:

$$Z_{i,j} = \{lnDist_{i,j}, Border_{i,j}, Language_{i,j}, Colony_{i,j}, Conflict_{i,j}, FTA_{i,j}\} \quad (5)$$

Trade is expected to fall as distance rises, reflecting transportation costs and information asymmetries. Adjacent countries should trade more with each other, as should countries sharing a common language and colonial history. Members of a free trade area should trade more, while countries that have been engaged in armed conflict recently are less likely to trade with one another.

One problem with estimating equation 4 is that migration may be endogenous. While we would expect migration to affect trade through the network and preference channels discussed above, there may also be reverse causality. Greater trade between two countries may lead to increased familiarity and warmer political relations between nations, lowering the cost of migration. Furthermore, there may be some unobserved or poorly measured variable (such as information flow) that affects both trade and migration, leading us to falsely attribute the effect of this variable on trade to migration. To address this issue, I instrument the migrant stock following Javorcik et al (2006) and McKenzie (2005). The migrant stock in 2000 is instrumented with the migrant stock in 1990, the cost of obtaining a passport in the sending country as a percentage of gross national income (from McKenzie, 2005), bilateral telephone traffic between the sending and receiving countries (from the Telegeography's international telephone traffic database), population density in the sending country, two dummy variables equal to 1 if the sending country has cultural restrictions on women traveling alone and legal restrictions that require residents to get government approval before traveling abroad (from McKenzie, 2005), and all explanatory variables in 4.

4.2.1 IV Estimates by Migration Level

Table 2 list benchmark IV estimates of equation 4 with the 2000 migrant stock instrumented with the variables discussed above. Estimates are given with total trade, imports, and exports as the dependent variable. For each of the three dependent variables, estimates are reported for the entire sample and two subsamples defined as low-migration pairs and high-migration pairs. The cutoffs for low and high migration are chosen somewhat arbitrarily as the 25th and 75th percentiles for the bilateral migration variable, roughly 2,000 and 12,000 migrants from country j residing in country i respectively. A more precise measure of the migration threshold will be estimated in the next section, but these cutoffs do suggest that the migration-trade elasticity differs across migration levels.

Looking at the first column, we see an estimated migration-trade elasticity of 0.2, suggesting that a 10% increase in the migrant stock from country j living in country i will increase total trade between these nations by 2%. These estimates are slightly smaller than those found in the existing literature, with the difference most likely due to the larger sample size as well as my correction for potential endogeneity. Looking at the other explanatory variables, we see that GDP in both countries has a positive and significant effect on trade, though per-capita GDP in the receiving country does not appear to matter. Trade falls with bilateral distance and conflict, rises with colonial ties, and is unaffected by common language or membership in a free trade area.

When we restrict the sample to bilateral observations with less than 2,000 migrants, we see a slight increase in the migration-trade elasticity (up to 0.25). The income variables have the same effects as before, but now the effects of common language and membership in a free trade area increase in size and significance. Looking at the high immigration sample (more than 12,000 migrants), the migration-trade elasticity nearly doubles to 0.45. That the marginal effect of migration is larger at higher levels of migration lends support to the minimum threshold model discussed in section 3.1. Interestingly, the common language variable is negative and significant in this sample. One possible explanation for this reversal is that when immigration is high, the trade-promoting effects of having a common language are reduced since a sizable immigrant group from a country that speaks a different language makes this different language less of a trade barrier.

The results for imports and exports support those found for total trade. Across the entire sample, a 10% increase in migration is predicted to increase imports and exports by 2.2%. Restricting the sample to low-migration pairs, a 10% increase in migration will should cause imports to rise by 2.8% and exports to rise by 3.0%. For high-migration pairs, a 10% increase in migration will increase imports by 3.8% and exports by 4.9%. The results in Table 2 suggest that migration has a larger effect on trade when the immigrant community in a country is large. This supports the minimum threshold model of trade and suggests that existing literature has overestimated the migration-trade elasticity at low levels of migration and underestimated the elasticity at high levels.

4.2.2 IV Estimates by Income Level

The migration-trade link may also differ by economic development. Table 3 presents IV estimates of equation 4 across different income levels for total trade, imports, and exports. The column Y_S^L restricts the sample to sending countries with per-capita GDP below \$5,000 while the column Y_S^H restricts the sample to sending countries with per-capita GDP above \$14,000 (roughly the 25th and 75th percentiles of sending country per-capita GDP).[4] The estimates in Table 3 suggest that migration has a much stronger effect on trade when the migrants are coming from relatively poor countries. I estimate that a 10% increase in migration from a poor country will increase total trade by 2.9%, while a 10% increase from a rich country will increase total trade by a statistically insignif-

[4]The sample is grouped based on *sending* country GDP rather than receiving country because all receiving countries are OECD members and have relatively high income levels. This group of receiving countries does include such middle-income nations as the Czech Republic, Hungary, Mexico, and Poland, so there is some variation in income that would be interesting to explore further, however.

icant 0.4%. A 10% increase in immigration from a poor country is predicted to increase imports by 3.3% and exports by 2.9%, while a similar increase in migration from a rich country will increase imports by 1.5% and exports by only 0.3%.

There are several explanations for why immigration from relatively poor countries affects trade more than immigration from rich countries. First, immigrants from poor countries may be less assimilated into their new homes than immigrants coming from rich nations. As a result, they may display a stronger retained home bias for native country goods. Second, trade-promoting institutions such as property rights, minimal bureaucracy, and transparency tend to be more established in rich countries. As immigrant networks often serve as a proxy for these institutions, networks should have a stronger effect on trade when these institutions are weak. For example, trade between the United States and Great Britain will benefit less from Americans living in London than trade between the US and Kenya would from Americans living in Nairobi simply because the institutions proxied for by an immigrant network are already in place when trading with the UK, but not necessarily so when trading with Kenya.

4.2.3 IV Estimates by Migration and Income Level

The estimates in Table 2 show that the marginal effect of migration on trade is larger at high levels of migration, suggesting that there may be some minimum threshold level of migration needed to induce trade. As an immigrant community grows, both the network and preference effects on trade increase to a point where the benefits of trade exceed the fixed costs of market entry. The estimates in Table 3 indicate that migration from low-income countries has a larger effect on trade than that from high income countries. This suggests that immigrants from poor countries may display a stronger preference for native country goods and/or that immigrant networks have a stronger effect when institutions in the sending country are weak. Both of these results suggest that the migration-trade link is a complex one that may vary across both the level of migration and economic development.

While the estimates in Tables 2 and 3 examined these sub-samples separately, the estimates in Table 4 allow for joint variation in both migration and development. Looking first at the low-income sub-sample (defined as before by sending country GDP per capita below $5,000), we see that the migration-trade elasticity is much larger at high levels of migration. A 10% increase in migration from a poor country is predicted to increase trade by 2.8% when migration is low (below 2,000 migrants). When migration is high (more than 12,000 migrants), a 10% increase in poor country migration is predicted to increase trade by 10.9%, a nearly four-fold increase. For poor countries, the minimum threshold model of migration appears to fit the best.

Looking at the estimates for high-income migration, the situation is reversed. A 10% increase in migration from a rich country is predicted to increase trade by 0.9% at low levels of migration and by less than 0.1% at high levels of migration. Though neither estimate is statistically different from the zero, the lower magnitude of the coefficient at high levels of migration may suggest that local production is substituting for trade when an immigrant community reaches a certain size.

4.3 Threshold Model

The estimates given in Tables 2-4 arbitrarily split the sample based on income and migration. A more accurate estimation would be to endogenize these breakpoints. In other words, to estimate the immigration level or income level at which a structural change occurs. We can do this using a procedure developed by Hansen (2000). Define $X_{i,j}$as a vector containing all the right hand side variables in equation 4 and $T_{i,j}$as log trade, we can write:

$$T_{i,j} = X_{i,j}\Psi_1 + u_{i,j} \quad if\, q_{i,j} \leq \gamma \tag{6}$$

$$T_{i,j} = X_{i,j}\Psi_2 + u_{i,j} \quad if\, q_{i,j} > \gamma \tag{7}$$

The migration-trade elasticity (a component of Ψ) may differ depending on whether or not some variable $q_{i,j}$ is above or below a threshold γ. For example, if we expect that the migration-trade elasticity is larger for high levels of immigration (the minimum-threshold hypothesis), then $q_{i,j}$ is the migrant stock from country j residing in country i and γ is the threshold (T_1) at which the elasticity changes. Rather than arbitrarily choosing this threshold, we estimate it from the data.

To do so, first define $\Delta = \Psi_1 - \Psi_2$ as the threshold effect. Now define a dummy variable $d_{i,j}(\gamma) = 1$ if $q_{i,j} \leq \gamma$ and $X_{i,j}(\gamma) = d_{i,j}(\gamma) * X_{i,j}$ We can then combine equations 6 and 7 into:

$$T_{i,j} = X_{i,j}\Psi_2 + X_{ij}(\gamma)\Delta + u_{i,j} \tag{8}$$

Note that equation 8 will turn into equations 6 or 7 depending on whether $q_{i,j}$ is above or below the threshold γ. This gives us three parameters to estimate: Ψ_2, Δ, γ. We could estimate these parameters through OLS, yielding a sum of squared errors given by:

$$S(\Psi_2, \Delta, \gamma) = [T_{i,j} - X_{i,j}\Psi - X_{i,j}(\gamma)\Delta]'[T_{i,j} - X_{i,j}\Psi - X_{i,j}(\gamma)\Delta] \tag{9}$$

Since γ is unobserved, we cannot estimate equation 8 nor obtain the sum of squared errors given in 9. However, we can obtain least square estimates through concentration. Conditional on γ, equation 8 is linear in both $\Psi$and Δ. Regressing $T_{i,j}$ on $X_{i,j}$ and $X_{i,j}(\gamma)$ yields the conditional estimators $\hat{\Psi}(\gamma)$ and $\hat{\Delta}(\gamma)$. Plugging these estimates into equation 9 yields the concentrated sum of squares given by:

$$S_n(\gamma) = S[\hat{\Psi}(\gamma), \hat{\Delta}(\gamma), \gamma] \tag{10}$$

We can then estimate $\hat{\gamma} = \underset{\gamma \in \Gamma}{argmin} S_n(\gamma)$. A computationally intensive, though relatively straightforward estimation method is to compute the concentrated sum of squares (10) across wide range of potential values of γ. The range (Γ) of possible values for γ is constrained by the variable that we are testing threshold effects on. For example, if we are testing the migration threshold at which the immigration-trade elasticity changes, then we know that γ is bounded below by 0 (can't have negative migration) and above by the largest migration value in the sample. While it is theoretically possible that the migration threshold exceeds the largest observed value, such a scenario is in practice the same as having no threshold at all. Hansen (2000) provides an asymptotic distribution for $\hat{\gamma}$, which we can then use to construct confidence intervals.

I consider two potential regime changes: differences in the level of migration and in the level of sending country development. Table 5 gives estimates from this procedure for total trade, imports, and exports. Looking first at total trade, the threshold at which there is a regime change is estimated to be 251.4 migrants with a confidence interval between 149.0 and 292.8 migrants. Given that this is a relatively small number of migrants (i.e. the threshold is very low), it is not surprising that the migration-trade elasticity does not differ above or below this threshold. Distance, common language, and colonial ties all have smaller absolute effects above this threshold, suggesting that these variables decline in importance as migration increases. This makes sense as these variables all proxy for the kind of trade costs that migration is expected to reduce.

I estimate a larger threshold of 3,045 migrants for imports, but no appreciable difference in the migration-trade elasticity. However, there is a small difference for exports. I estimate that a 10% increase in migration from the sending to receiving country will increase exports from the receiving to sending country by 3.0% when the immigrant community is less 809. When this community is larger than this threshold, a 10% increase in migration is predicted to increase exports by 4.2%.

Table 6 lists threshold estimates when sorting on sending country income. These results confirm those in Table 3 indicating a much larger immigration-trade elasticity for poor countries. I estimate that the gravity model with total trade as a dependent variable undergoes a regime change when sending country GDP per capita reaches \$2,904. Below this threshold, the migration-trade elasticity is relatively high at 0.44. This same elasticity is still significant above the threshold, but very low at 0.09. A similar pattern emerges when imports and exports are used as dependent variables. The estimated income threshold for imports is \$7,682, while that for exports is \$16,033.[5] Below the income threshold, a 10% increase in migration is predicted to raise imports by 2.8% and exports by 2.7%. Above this income threshold, a 10% increase in migration is predicted to raise imports by 1.3% and exports by only 0.2%.

[5] Why the threshold for exports is so much higher is an interesting question that needs further exploration.

5 Conclusion

The results presented in Tables 2-6 suggest that there are indeed non-linearities in the migration-trade relationship. Migration from poor countries to rich countries has a larger effect on trade than migration from rich to rich countries, perhaps due to weaker assimilation of immigrants from poor countries or weaker institutions in these poor countries. Looking within income groups, the marginal effect of migration is much stronger at high levels of migration from poor countries, while this marginal effect does not change or may even decline at high levels of rich country migration. This result suggests that migration from poor countries may need to reach a certain minimum threshold before any trade effects are seen, while high migration between rich countries could result in local production replacing exports as a way to serve immigrant preferences.

There are several limitations in this study. First, there are a large number of zeros in both the trade and migration databases. As the estimating equation is a double-log specification, there is a question of how to deal with these zeros in estimation. Rather than throw these observations out, I replaced them with an arbitrarily small number (0.001). However, an alternative method such as Tobit may improve accuracy. Second, there are numerous idiosyncratic country effects that may be biasing the results. The preferred method of dealing with these effects is to utilize panel estimation. I am limited, however, by the paucity of comparable migration data across time. I only have this data for two years (1990 and 2000) and must use the 1990 data for the instrumental variables analysis. However, there may be more frequently updated migration data available for a smaller subset of countries.

Given these limitations, it is important to take the following policy implications with a grain of salt. However, the results in this study suggest that the positive effect that migration will vary across both migration and source country income. Past estimates of this effect have, especially for migration from relatively poor countries, overestimated the impact of migration on trade at low levels of migration and underestimated the impact at high levels of migration. The bias in the existing literature is not as bad for rich country migration, although there is some evidence that rich-country migration may actually reduce trade at high enough levels as local production supplants exports.

6 References

Blanes, Jose V. 2005. Does Immigration help to Explain Intra-Industry Trade? Evidence from Spain. *Review of World Economics* 141:2, 244-270

CEPII Geodesic Distances Database, available at: http://www.cepii.fr/anglaisgraph/bdd/distances.htm. Accessed on May 1st, 2009

Docquier, Frederic and Abdeslam Marfouk. 2007. Brain Drain in Developing Countries. *World Bank Economic Review* 21:2, 193-218

Dunleavy, James A. and William K. Hutchinson. 1999. The Impact of Immigration on American Import Trade in the Late Nineteenth and Early Twentieth Centuries. *The Journal of Economic History* 59:4, 1043-1062

Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Henry Mo. 2005. World Trade Flows: 1962-2000. NBER Working Paper No W11040

Foad, Hisham. 2008. FDI and Immigration: A Regional Analysis. Mimeo available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1092286

Girma, Sourafel and Zhihao Yu. 2002. The Link between Immigration and Trade: Evidence from the United Kingdom. *Review of World Economics* 138:1, 115-130

Gould, David. 1994. Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows. *Review of Economics and Statistics* 76:2, 302-316

Greif, Avner. 1989. Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders. *The Journal of Economic History* 49:4, 857-882

Greif, Avner. 1994. Cultural Beliefs and the Organization of Society: A Historical Reflection on Collectivist and Individualist Societies. *The Journal of Political Economy* 102:5, 912-950

Hansen, Bruce E. 2000. Sample Splitting and Threshold Estimation. *Econometrica* 68:3, 575-604

Head, Keith and John Ries. 1998. Immigration and Trade Creation: Econometric Evidence from Canada. *The Canadian Journal of Economics* 31:1, 47-62

Helliwell, James. 1996. Do National Borders Matter for Quebec's Trade? *Canadian Journal of Economics* 29:3, 507-522

Javorcik, Beata, Caglar Ozden, Mariana Spatareanu, and Cristina Neagu . 2006. Migrant Networks and FDI. *World Bank Policy Research Working Paper Series* 4046

Kugler, Maurice and Hillel Rapoport. 2005. Skilled Emigration, Business Networks, and Foreign Direct Investment. *CESifo Working Paper Series* 1455

McCallum, John. 1995. National Borders Matter: Canada-U.S. Trade Patterns. *American Economic Review* 85:3, 615-623

McKenzie, David. 2005. Paper Walls are Easier to Tear Down: Passport Costs and Legal Barriers to Emigration. World Bank Policy Research Paper No 3783

Rauch, James. 1999. Networks vs. Markets in International Trade. *Journal of International Economics* 48:1, 7-35

Rauch, James and Vitor Trindade. 2002. Ethnic Chinese Networks in International Trade. *Review of Economics and Statistics* 84:1, 116-130

Telegeography International Traffic Database, available at http://www.telegeography.com/. Accessed on February 1st, 2009

Upssala Conflict Data Program, available at http://www.pcr.uu.se/research/UCDP/index.htm. Accessed on May 15th, 2009

White, Roger. 2007. Immigrant Trade Links, Transplanted Home Bias and Network Effects 39:7, 839-852

Wolf, Holger. 2000. International Home Bias in Trade. *Review of Economics and Statistics* 82:4, 555-563

Table 1: List of Receiving and Sending Countries

Receiving	Sending
Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States	Afghanistan, Albania, Algeria, Angola, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahamas, Bahrain, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bolivia, Bosnia and Herzegovina, Brazil, Brunei, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Cape Verde, Central African Republic, Chile, China, Colombia, Congo, Rep., Costa Rica, Cote d'Ivoire, Croatia, Cuba, Cyprus, Czech Republic, Denmark, Djibouti, Dominican Republic, DR Congo, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Ethiopia, Fiji, Finland, France, Gabon, Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea, Kuwait, Kyrgyzstan, Laos, Latvia, Lebanon, Liberia, Libya, Lithuania, Macedonia, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Myanmar, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russia, Rwanda, Samoa, Saudi Arabia, Senegal, Serbia and Montenegro, Sierra Leone, Singapore, Slovakia, Slovenia, Somalia, South Africa, Spain, Sri Lanka, Sudan, Suriname, Sweden, Switzerland, Syria, Taiwan, Tajikistan, Tanzania, Thailand, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Venezuela, Vietnam, Yemen, Zambia, Zimbabwe

Table 2: IV Estimates by Migration Level

	Total Trade			Imports			Exports		
	All	M_L	M_H	All	M_L	M_H	All	M_L	M_H
Migration	0.20	0.25	0.45	0.22	0.28	0.38	0.22	0.30	0.49
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
GDP_R	1.38	1.69	0.39	1.6	1.83	0.77	0.89	1.51	-1.14
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
GDP_S	1.36	1.52	0.66	1.64	1.81	0.86	1.32	1.38	1.11
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$PGDP_R$	-0.07	0.18	-1.19	-0.29	-0.19	-1.48	0.05	0.59	-1.41
	[0.72]	[0.41]	[0.01]	[0.20]	[0.50]	[0.01]	[0.81]	[0.02]	[0.08]
$PGDP_S$	0.44	0.37	0.84	0.46	0.38	0.79	0.78	0.74	1.27
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Distance	-1.16	-1.41	-0.21	-1.15	-1.26	-0.24	-1.46	-1.74	-0.37
	[0.00]	[0.00]	[0.16]	[0.00]	[0.00]	[0.18]	[0.00]	[0.00]	[0.14]
Border	-1.33	-0.94	-0.33	-1.3	0.00	-0.37	-2.18	-2.67	-1.91
	[0.00]	[0.40]	[0.47]	[0.01]	[1.00]	[0.51]	[0.00]	[0.04]	[0.02]
Language	0.04	0.48	-0.81	0.28	0.60	-0.63	-0.13	0.42	-0.28
	[0.84]	[0.12]	[0.01]	[0.28]	[0.12]	[0.09]	[0.63]	[0.24]	[0.60]
Colony	0.60	-0.34	0.92	1.05	0.50	1.12	1.51	-0.05	1.96
	[0.05]	[0.58]	[0.01]	[0.01]	[0.51]	[0.01]	[0.00]	[0.94]	[0.00]
Conflict	-6.33	-3.59	-6.02	-5.83	-3.61	-5.93	-6.96	-3.98	-5.70
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
FTA	0.19	0.52	0.30	0.65	1.32	0.71	0.11	0.26	0.26
	[0.35]	[0.08]	[0.38]	[0.01]	[0.00]	[0.08]	[0.67]	[0.45]	[0.66]
Constant	2.59	0.67	9.80	0.12	-1.18	9.28	2.50	-3.66	15.56
	[0.15]	[0.76]	[0.03]	[0.96]	[0.66]	[0.09]	[0.27]	[0.14]	[0.05]
R^2	0.61	0.56	0.63	0.57	0.51	0.57	0.50	0.50	0.55
Obs	4698	3525	580	4698	3525	580	4698	3525	580

IV estimates of equation 4. P-values given in brackets below each coefficient estimate. M_L is a sub-sample defined by less than 2000 migrants from the sending country living in the receiving country. M_H is a sub-sample defined by more than 12000 migrants from the sending country living in the receiving country.

Table 3: IV Estimates by Income Level

	Total Trade			Imports			Exports		
	All	Y_S^L	Y_S^H	All	Y_S^L	Y_S^H	All	Y_S^L	Y_S^H
Migration	0.20	0.29	0.04	0.22	0.33	0.15	0.22	0.29	0.03
	[0.00]	[0.00]	[0.38]	[0.00]	[0.00]	[0.04]	[0.00]	[0.00]	[0.62]
GDP_R	1.38	1.47	1.24	1.60	1.65	1.42	0.89	0.82	1.03
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
GDP_S	1.36	1.37	1.34	1.64	1.64	1.54	1.32	1.31	1.39
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$PGDP_R$	-0.07	-0.23	0.50	-0.29	-0.71	-0.04	0.05	-0.01	0.77
	[0.72]	[0.45]	[0.07]	[0.20]	[0.06]	[0.91]	[0.81]	[0.98]	[0.02]
$PGDP_S$	0.44	0.00	-0.07	0.46	-0.17	0.32	0.78	0.52	0.05
	[0.00]	[1.00]	[0.83]	[0.00]	[0.25]	[0.47]	[0.00]	[0.00]	[0.89]
Distance	-1.16	-1.40	-1.03	-1.15	-0.85	-1.02	-1.46	-2.05	-1.1
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Border	-1.33	-2.48	-0.97	-1.30	-0.72	-1.72	-2.18	-5.32	-0.80
	[0.00]	[0.10]	[0.02]	[0.01]	[0.69]	[0.00]	[0.00]	[0.00]	[0.12]
Language	0.04	-0.44	1.05	0.28	-0.22	1.03	-0.13	-0.48	0.32
	[0.84]	[0.18]	[0.00]	[0.28]	[0.58]	[0.02]	[0.63]	[0.24]	[0.39]
Colony	0.60	0.91	-0.23	1.05	1.47	-0.07	1.51	2.51	0.58
	[0.05]	[0.07]	[0.60]	[0.01]	[0.02]	[0.91]	[0.00]	[0.00]	[0.26]
Conflict	-6.33	-7.05	n.a.	-5.83	-6.21	n.a.	-6.96	-7.26	n.a.
	[0.00]	[0.00]	[n.a.]	[0.00]	[0.00]	n.a.	[0.00]	[0.00]	n.a.
FTA	0.19	0.55	0.41	0.65	1.54	0.93	0.11	1.04	0.42
	[0.35]	[0.37]	[0.06]	[0.01]	[0.04]	[0.00]	[0.67]	[0.18]	[0.10]
Constant	2.59	8.70	2.68	0.12	5.60	-0.22	2.50	10.51	-0.54
	[0.15]	[0.01]	[0.53]	[0.96]	[0.14]	[0.97]	[0.27]	[0.01]	[0.92]
R^2	0.61	0.49	0.64	0.57	0.46	0.58	0.5	0.35	0.55
Obs	4698	2378	1194	4698	2378	1194	4698	2378	1194

IV estimates of equation 4 across income levels. Y_S^L is a sub-sample defined by sending country GDP per capita below \$5,000 while Y_S^H is a sub-sample defined by sending country GDP greater than \$14,000. P-values given in brackets below the coefficient estimates.

Table 4: IV Estimates by Income and Migration

	Low Income		High Income	
	M_L	M_H	M_L	M_H
Migration	0.28	1.09	0.09	0.00
	[0.00]	[0.00]	[0.20]	[0.97]
GDPR	1.80	0.13	1.54	0.79
	[0.00]	[0.65]	[0.00]	[0.00]
GDPS	1.56	0.15	1.54	0.89
	[0.00]	[0.45]	[0.00]	[0.00]
PGDPR	0.34	-6.59	0.60	0.28
	[0.32]	[0.00]	[0.11]	[0.24]
PGDPS	0.03	1.22	-0.15	0.82
	[0.79]	[0.01]	[0.72]	[0.00]
Distance	-1.84	0.98	-1.23	-0.65
	[0.00]	[0.02]	[0.00]	[0.00]
Border	-3.48	-6.83	0.19	0.17
	[0.11]	[0.02]	[0.91]	[0.36]
Language	-0.01	-2.04	2.55	0.39
	[0.98]	[0.00]	[0.00]	[0.01]
Colony	0.06	0.85	-2.28	0.18
	[0.95]	[0.26]	[0.05]	[0.28]
Conflict	-4.13	-6.01	n.a.	n.a.
	[0.00]	[0.00]	n.a.	n.a.
FTA	0.88	0.79	0.58	0.28
	[0.31]	[0.49]	[0.06]	[0.08]
Constant	4.65	50.4	1.68	-1.22
	[0.17]	[0.00]	[0.77]	[0.73]
R2	0.47	0.57	0.57	0.88
Obs	1939	203	808	213

Table 5: Migration Threshold Estimates

	Total		Imports		Exports	
	$\gamma = 251.4$		$\gamma = 3045.1$		$\gamma = 808.9$	
	(149.0, 292.8)		(2440.6, 8989.6)		(665.1, 1096.6)	
	Below	Above	Below	Above	Below	Above
Migration	0.31	0.30	0.26	0.28	0.30	0.42
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
GDP_R	1.87	0.68	1.82	0.94	1.67	-0.47
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
GDP_S	1.61	0.88	1.79	1.05	1.42	1.05
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$PGDP_R$	0.16	-0.69	-0.14	-1.25	0.79	-2.16
	[0.57]	[0.00]	[0.61]	[0.00]	[0.00]	[0.00]
$PGDP_S$	0.36	0.64	0.41	0.70	0.75	0.83
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Distance	-1.6	-0.59	-1.28	-0.52	-1.82	-0.78
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Border	-1.34	-0.40	-0.28	-0.30	-3.38	-1.89
	[0.47]	[0.21]	[0.83]	[0.51]	[0.04]	[0.00]
Language	0.50	-0.55	0.58	-0.36	0.73	-0.66
	[0.25]	[0.00]	[0.11]	[0.21]	[0.08]	[0.04]
Colony	-1.73	0.86	0.86	0.98	-0.55	2.01
	[0.08]	[0.00]	[0.21]	[0.00]	[0.50]	[0.00]
Conflict	-4.37	-7.22	-2.77	-7.77	-4.38	-7.46
	[0.00]	[0.00]	[0.02]	[0.00]	[0.00]	[0.00]
FTA	1.13	0.31	1.16	0.42	0.58	0.21
	[0.02]	[0.08]	[0.00]	[0.16]	[0.15]	[0.51]
Constant	1.35	8.06	-1.6	8.9	-5.85	26.88
	[0.63]	[0.00]	[0.54]	[0.02]	[0.03]	[0.00]
R^2	0.53	0.61	0.52	0.52	0.50	0.50
Obs	2607	2091	3666	1031	3143	1555

Concentrated IV estimates of equation 4 using Hansen's sample splitting technique. γ is the estimated threshold using migration as the threshold variable. 95 percent confidence intervals are given below the γ estimates. "Below" refers to the sub-sample defined by observations with migration below this threshold, while "Above" refers to the sub-sample defined by migration above this threshold.

Table 6: Income Threshold Estimates

	Total		Imports		Exports	
	$\gamma = 2904.3$		$\gamma = 7682.0$		$\gamma = 16033.1$	
	(2477.1, 2988.9)		(7653.9, 8232.3)		(15637.1, 18191.8)	
	Below	Above	Below	Above	Below	Above
Migration	0.44	0.09	0.28	0.13	0.27	0.02
	[0.00]	[0.01]	[0.00]	[0.02]	[0.00]	[0.76]
GDP_R	1.41	1.36	1.67	1.43	0.85	1.05
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
GDP_S	1.01	1.50	1.70	1.51	1.34	1.40
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$PGDP_R$	-0.71	0.28	-0.44	-0.07	-0.07	0.74
	[0.09]	[0.15]	[0.16]	[0.8]	[0.79]	[0.05]
$PGDP_S$	0.33	0.76	-0.07	-0.20	0.62	-0.52
	[0.19]	[0.00]	[0.57]	[0.34]	[0.00]	[0.33]
Distance	-1.31	-1.22	-1.06	-1.11	-1.64	-1.10
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Border	-3.97	-1.00	0.34	-1.25	-2.41	-0.82
	[0.11]	[0.01]	[0.77]	[0.01]	[0.00]	[0.14]
Language	-1.29	0.82	-0.09	0.92	-0.29	0.47
	[0.00]	[0.00]	[0.80]	[0.01]	[0.37]	[0.25]
Colony	0.89	0.23	1.46	0.34	1.82	0.51
	[0.19]	[0.48]	[0.01]	[0.49]	[0.00]	[0.36]
Conflict	-7.45	-4.17	-6.18	-1.40	-7.25	n.a.
	[0.00]	[0.01]	[0.00]	[0.69]	[0.00]	n.a.
FTA	1.62	-0.05	1.15	0.86	0.90	0.40
	[0.09]	[0.81]	[0.05]	[0.00]	[0.06]	[0.15]
Constant	11.40	-3.55	3.89	6.36	6.49	5.49
	[0.01]	[0.08]	[0.21]	[0.09]	[0.02]	[0.39]
R^2	0.43	0.65	0.51	0.55	0.43	0.55
Obs	1479	3219	2956	1741	3675	1023

Concentrated IV estimates of equation 4 using Hansen's sample splitting technique. γ is the estimated threshold using sending country income per capita as the threshold variable. 95 percent confidence intervals are given below the γ estimates. "Below" refers to the sub-sample defined by observations with income below this threshold, while "Above" refers to the sub-sample defined by income above this threshold.