

IZA DP No. 9058

**The Causal Impact of Migration on US Trade:
Evidence from a Natural Experiment**

Walter Steingress

May 2015

The Causal Impact of Migration on US Trade: Evidence from a Natural Experiment

Walter Steingress

*Banque de France
and IZA*

Discussion Paper No. 9058
May 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Causal Impact of Migration on US Trade: Evidence from a Natural Experiment^{*}

Immigrants can increase international trade by shifting preferences towards the goods of their country of origin and by reducing bilateral transaction costs. Using geographical variations across US states for the period 1970 to 2005, we quantify the impact of immigrants on intermediate goods imports. We address endogeneity and reverse causality – which arises if migration from a country of origin to a US state is driven by trade opportunities between the two locations – by exploiting the exogenous allocation of refugees within the US refugee resettlement program. Our results are robust to an alternative identification strategy, based on the large influx of Central American immigrants to the United States after hurricane Mitch. We find that a 10 percent increase in recent immigrants to a given US state raises intermediate imports from those immigrants' country of origin by 1.5 percent.

JEL Classification: F14, F22, J61

Keywords: international trade, international migration, political refugees, hurricane Mitch

Corresponding author:

Walter Steingress
Banque de France
International Macroeconomics Division
49-1374 DGEI-DERIE-SEMSI
31 Rue Croix des Petites Champs
75001 Paris
France
E-mail: walter.steingress@banque-france.fr

^{*} I would like to thank Andriana Bellou, Rui Castro, Jean Francois Maystadt, Baris Kaymak, Anna Maria Mayda, Giovanni Peri and Sebastian Stumpner for their useful comments and suggestions. All remaining errors are mine. The views expressed in this paper are those of the author alone and do not reflect those of the Banque de France. All remaining errors are mine. The views expressed in this paper are those of the author alone and do not reflect those of the Banque de France.

1 Introduction

Work on individual attitudes shows that public opinion is not at all favorable to increases in the number of immigrants, see [Mayda \(2006\)](#). Yet increasing evidence in the literature shows a range of beneficial effects for the destination country. An important channel that might increase the welfare of the host country is international trade. If migrants reduce transaction costs between countries, either by holding information about relevant market characteristics or by attenuating frictions due to imperfect contract enforcement, gains from trade are realized. Alternatively, if migrants shift preferences towards the goods of their country of origin, consumers in the host country will benefit from more consumption variety. However, no definitive answer has yet been reached on the causal impact of migration on trade.

This paper studies the impact of immigrants on intermediate goods imports by studying geographical variations across US states for the period 1970 to 2005. To generate our results, we follow a gravity approach and regress the log of intermediate imports on the log of recent immigrants. We address reverse causality, and more broadly endogeneity, and find evidence of a significantly positive effect of immigrants on US intermediate imports. To estimate a causal relationship, we analyze the exogenous allocation of refugees within the US refugee resettlement program. Our results are robust to an alternative identification strategy, based on the large influx of Central American immigrants to the United States after hurricane Mitch.

There are two important differences between our analysis and previous studies. First, we focus on regional variations in immigration and intermediate imports *within* the United States. This approach allows us to estimate the trade-migration relationship leaving aside the potential correlation between trade and migration policy at the national level. Since there are no data on the distribution of intermediate imports across US states for the time period under review, we apply a procedure based on [Anderson and Yotov \(2010\)](#) to infer this distribution. Using data from the County Business Pattern and applying the techniques from [Autor et al. \(2013\)](#) for each US state, we calculate the intermediate imports as the sum of industry-specific imports. The latter are modeled as depending on the relative size of the industry and on bilateral trade costs with respect to each trading partner (geographical distance, border and the value-to-weight ratio). This method allows us to project the observed national import flows by industry onto the state level. To assess the appropriateness of our method, we use data available for recent years on the distribution of intermediate imports across US states and compare the constructed import flows with the observed ones. We find a high correlation between the two flows, i.e. a correlation coefficient of 0.85. In addition, using recent observed import data, we find estimates of the impact of migration on trade to be qualitatively similar to the ones based on constructed import flows for the 1970-2005 period.¹

Second, and most importantly, we address endogeneity by focusing on exogenous shocks to

¹Before 2008, only national US import data are available.

immigration. Endogeneity arises because immigrants' decisions regarding their settlement within the United States are likely to be correlated with several variables, such as income, employment opportunities and/or preferences, which in turn are correlated with trade, see [Borjas \(1999\)](#). An additional source of endogeneity arises in the form of reverse causality, i.e. immigrants from a given country are likely to settle in states that trade a lot with their country of origin. As a first step towards addressing these issues, we focus on recent immigrants and include the number of non-recent immigrants living in the state as a control variable. The non-recent immigrant population of a given origin will capture the unobserved preferences of immigrants of that origin for a specific location. It will also account for the impact of better income and employment opportunities in a state that are specific to immigrants of that origin, see [Clark et al. \(2007\)](#). In addition, and more importantly, we address endogeneity and reverse causality by estimating an IV specification. First, we use the exogenous allocation of political refugees across US states within the refugee resettlement program. Second, we take advantage of an exogenous shock that forces people to migrate. In particular, we look at of the large influx of Central American immigrants to nearby US states following hurricane Mitch in October 1998. The IV approach removes the endogenous component of migration decisions whereby individuals might move to those regions with the best trading opportunities.

Starting with the pioneering work of [Gould \(1994\)](#), there exists ample empirical literature that argues that immigrants increase trade across international borders. [Gould \(1994\)](#) studies the effect of migration on aggregate US exports and imports for the years 1970-1986. He estimates a gravity model of trade on migration and finds evidence of a strong positive relationship. Many authors follow Gould and study immigration into a single country, for example [Head and Ries \(1998\)](#) examine Canada and [Girma and Yu \(2002\)](#) the United Kingdom. The drawback of focusing on the aggregate level is that these studies cannot account for correlation in migration and trade policies. More recent studies exploit the regional distribution of immigrants and look at the bilateral trade relationship between US regions and foreign countries, see [Bardhan and Guhathakurta \(2004\)](#), [Dunlevy \(2006\)](#) and [Parsons and Vézina \(2014\)](#). While these studies focus exclusively on US exports, this paper is the first to present evidence on US imports at the state level.

In a recent literature review, [Felbermayr et al. \(2012\)](#) argue that the main concern for the identification of the causal effect of immigration is reverse causality. To deal with this issue, authors have adopted different approaches. Drawing upon the seminal work of [Card \(2001\)](#), several papers (see for example [Peri and Requena-Silvente \(2010\)](#)) instrument changes in immigrants at the sub-national level by applying the growth in immigrants from a given country of origin at the national level to the distribution of immigrants from a given country of origin across regions at a previous point in time. Still, historical migrant stocks can have direct effects on trade even many years after their arrival, therefore violating the exclusion restriction.

The only other paper that has recourse to a natural experiment for identification is [Parsons and](#)

Vézina (2014). Their strategy is based on the combination of an immigration shock caused by the exogenous allocation of Vietnamese Boat People across US states and a concurrent trade embargo. They use the cross section variation in Vietnamese immigrants and US exports across the 50 US states to estimate the impact of Vietnamese migration on US exports. Rather than focusing on a particular group of migrants at a given point in time, this paper uses data on political refugees to the U.S. from all countries between the years of 1970 and 2005. To establish the causal relationship, our analysis takes advantage of the fact that political refugees to the U.S. are exogenously allocated across locations once we control for the existence of families ties of migrants. As a robustness check, we use the migration flows from several Central American countries following hurricane Mitch. Overall, our findings show that the trade-enhancing effect of immigrants found by Parsons and Vézina (2014) also holds in a general setting where we consider regional intermediate imports from many countries.

The rest of the paper is organized as follows. Section 2 describes the identification strategy as well as the details of the natural experiments, the political refugees and hurricane Mitch, which allow us to establish causality between trade and migration. Section 2 also contains the import demand model used to project the aggregate US data onto the state level. Section 3 covers the data and the summary statistics. Section 4 discusses the obtained OLS and IV results. Section 5 discusses the robustness of our findings and section 6 concludes.

2 Identification strategy

Our identification strategy uses variation in intermediate imports and the number of recent immigrants across US states. Once migrants settle in a US state, international trade between the state of residence and their respective home countries is expected to increase. This effect is likely to take place through the following channels. First, migrants might provide information that reduces transaction costs. Second, trade might increase simply because migrants have a preference for the goods from their country of origin. However, in our analysis we do not explicitly distinguish between the two mechanisms.

The reason why we focus on intermediate - as opposed to final - goods' imports is that, for our sample period, we have no information on regional trade data in the United States. Instead, we infer the distribution of trade flows across US states via an intermediate input demand model. This model allows us to project the observed aggregate trade data onto the state level. A key element in the calculation of regional intermediate imports is production data by region and industry. Since consumption data by region and industry is not available, we cannot assess regional imports for final goods.

2.1 Regression specification

To analyze the impact of immigration on intermediate imports, we follow the literature, [Felbermayr et al. \(2012\)](#), and employ the gravity equation. The gravity equation relates trade flows between state i and country j to the relative sizes of the participating economies (which in our specification are captured by fixed effects). The log of the import flow of state i from country j for the period t , X_{ijt} , is given by

$$\log(X_{ijt}) = \beta_1 \log(Imm_{ijt}) + \beta_2 \log d_{ij} + \beta_3 b_{ij} + \beta_4 \log PImm_{ijt} + f_{jt} + f_{it} + \varepsilon_{ijt}. \quad (1)$$

The regressor Imm_{ijt} indicates the number of foreign-born residents in state i above the age of 21 who immigrated from country j in any of the 5 years prior to time t . The other regressors are the log of the distance, d_{ij} , between the capital of state i and the capital in country j , measured in kilometers. b_{ij} represents a dummy variable that indicates whether state i shares a border with country j . $PImm_{ijt}$ is the number of foreign-born residents in state i above the age of 21 who immigrated from country j more than 5 years prior to time t .² In addition, we also introduce time varying state and country fixed effects, f_{it} and f_{jt} respectively. The coefficient of interest is β_1 . If $\beta_1 > 0$, then the presence of recent immigration increases intermediate imports.

Equation 1 represents the ideal regression in the case of import flows being observed. Given that we approximate import flows within an intermediate input demand model, we introduce a measurement error, u_{ijt} . The estimation equation becomes

$$\log(\hat{X}_{ijt}) = \beta_1 \log(Imm_{ijt}) + \beta_2 \log d_{ij} + \beta_3 b_{ij} + \beta_4 \log PImm_{ijt} + f_{it} + f_{jt} + v_{ijt} \quad (2)$$

and

$$\log(\hat{X}_{ijt}) = \log(X_{ijt}) + u_{ijt}$$

where \hat{X}_{ijt} represents the estimated state import flows and v_{ijt} the error in the OLS regression. β_1 is consistently estimated if $Cov(Imm_{ijt}, u_{ijt}) = Cov(Imm_{ijt}, v_{ijt}) = 0$. Due to the presence of the measurement error, the standard errors are larger and the likelihood of not rejecting the null hypothesis of $\beta_1 = 0$ increases, see [Wooldridge \(2010\)](#). Thus, a significant effect of migration on trade will reinforce our confidence in the results obtained. Equation 2 will be our main regression specification.³

In terms of identification of β_1 , it is important to control for geography, since both migration and trade are correlated with distance and border, see [Head and Mayer \(2013\)](#). States import more from countries that are relatively close by and with whom they share a common border.

²Note that Imm_{ijt} and $PImm_{ijt}$ represent the change in the stock of immigrants in the last five years and the preceding period. Thus, our specification regresses trade flows on migration flows.

³In addition since the dependent variables is estimated, we address heteroskedasticity by estimating robust standard errors. Quote: Trade Reforms and Income Inequality in Colombia, (with O. Attanasio and P. Goldberg), *Journal of Development Economics*, 74, August 2004, pp. 331-366.

At the same time, relatively more people migrate from neighboring countries that are close by. Neglecting these effects will introduce an omitted variable bias.

Note that, by analyzing the trade migration relationship across US states, we directly address the criticism of [Hanson \(2010\)](#) with respect to the earlier literature. He argues that “It is difficult to draw causal inference from results based on international trading and migration patterns, since immigration may be correlated with unobserved factors that also affect trade, such as the trading partners’ cultural similarity or bilateral economic policies (e.g., preferential trade policies or investment treaties that raise the return to both migration and trade)”. Trade policies and investment treaties are negotiated at the federal level and are thus controlled for by fixed effects specific to the country of origin, f_{jt} . These fixed effects also control for any effects that are common to all regions in the United States. For example, if a country experiences a positive productivity shock, trade might increase since all regions will face lower import prices from this country and emigration might decrease because of better employment opportunities.

A further concern for the identification of the parameter β_1 is the presence of time-varying state-specific characteristics that may be correlated with trade flows as well as immigration. One such candidate is, for instance, economies of agglomeration, i.e. more immigrants are likely to settle in larger states and those states have higher demand for intermediate goods. For this reason, we include time-varying state-fixed effects, f_{it} , that control for any state-specific effects, such as local demand and income shocks, which are common to all migrants and vary over time.

By looking at regional variations and including state and country time-varying fixed effects, we follow the recent literature, see [Bardhan and Guhathakurta \(2004\)](#) and [Dunlevy \(2006\)](#). The key difference of this paper with respect to the literature is that we offer a new approach to resolving endogeneity. Endogeneity arises because immigrants’ decisions regarding settlement within the United States is likely to be correlated with several variables, such as income, employment opportunities and/or preferences, which in turn are correlated with trade, see [Borjas \(1999\)](#). An additional source of endogeneity arises in the form of reverse causality, i.e. immigrants from a given country of origin are likely to go to states that trade a lot with that country.

To tackle endogeneity, we proceed in two steps. First, we distinguish between different groups of immigrants according to their arrival date. Specifically, we focus on newly arrived immigrants and include the previous number of immigrants as an additional control variable, $PImm_{ijt}$. This allows us to account for the fact that immigrants are not randomly distributed across geographical locations. Consider for example Austrian immigrants, who have a specific knowledge of the production of skis and decide to emigrate to the United States in order to make use of this knowledge. They may prefer to go to Colorado rather than Texas because of the relatively higher demand for skis, and thus the greater business and employment opportunities for them in the Rocky Mountains. In a similar vein, Italians may prefer to move to California in order to start

producing wine rather than to Montana where climatic conditions prevent wine from growing. In both scenarios, the settlement choice differs across migrants of different origins and is correlated with business and employment opportunities and/or preferences, which in turn are correlated with imports from their countries of origin. We mitigate the resulting bias by looking at the impact of recent immigrants on trade and, at the same time, controlling for the previous number of immigrants from the same country of origin already living in the region. Note that the two immigrant groups, recent and previous, are observed at the same point in time in a given state. Thus, $Pimm_{ijt}$ allows to control for time varying shocks to location preferences of immigrants from a given country of origin. The only difference between the two groups is the arrival date. Hence, we assume that recent immigrants' settlement preferences are not fundamentally different from those of previous immigrants.

Most importantly, as a second step, we focus on exogenous shocks to migration. These exogenous shocks address reverse causality, i.e. immigrants from a given country of origin are likely to go to states that trade a lot with that country, and more generally endogeneity. [Gould \(1994\)](#) argues that immigration occurs before the onset of trade and is therefore predetermined. This is true if the migration decision is based on current or past levels of trade. However, if the migration decision is forward-looking and dependent on expected future trade (for example people emigrate in order to take advantage of information arbitrage, which leads to trade), past immigration is endogenous. As a result, the number of immigrants and the level of trade are jointly determined. Generally, migration is endogenous due to omitted variables such as income, employment opportunities and/or preferences which are correlated with trade. Thus, the OLS estimates are likely to be inconsistent.

To solve for the resulting bias, the literature follows [Card \(2001\)](#) and instruments changes in immigrants at the sub-national level by applying the growth of immigrants from a given country of origin at the national level to the distribution of immigrants from that country of origin across regions at a previous point in time, see for example [Peri and Requena-Silvente \(2010\)](#). Still, historical migrant stocks could have established long-standing trade relationships, with direct effects on the current level of trade. This will violate the exclusion restriction of the instrument. Instead, we take advantage of an exogenous shock that forces people to migrate to a given state independently of trading opportunities. This shock insulates our results from any endogenous migration decisions whereby individuals might move to those regions with better trading opportunities.

The only other paper that provides evidence of a natural experiment on the trade migration relationship is [Parsons and Vézina \(2014\)](#). Their identification strategy is based on the combination of an immigration shock driven by the exogenous allocation of Vietnamese Boat People by the US government across US states and a concurrent trade embargo. They use the cross section variation in the share of Vietnamese immigrants and exports to Vietnam of the 50 US states. They show that after the end of the 1994 trade embargo US states with a higher share of Vietnamese

immigrants exported significantly more to Vietnam. The key identification assumption is that the settlement choice of Vietnamese immigrants before the trade embargo is exogenous with respect to US exports after the embargo. In contrast, our analysis considers 2 different natural experiments. First, we use the exogenous allocation of political refugees across US states within the refugee resettlement program. Second, we make use of a natural disaster, hurricane Mitch. As we argue below in more detail, this hurricane led to the mass immigration of several hundred thousand people from Central America to the US.

The main differences with respect to [Parsons and Vézina \(2014\)](#) are the following. First, we focus on regional differences in US imports, whereas they focus on US exports. Second, and more importantly, our analysis resolves endogeneity by exploiting the *panel* variation of exogenous migration along both, the time and country of origins, dimensions within two natural experiments. As a result, the number of observations increases significantly compared to Parsons and Vezina's cross section of 50 US states. A further benefit from working with the panel structure is that we can exploit the bilateral variation over time and control for any state and country time-varying fixed effects. The third distinction is that we include the pre-existing number of immigrants as an additional regressor. This allows us to control for unobserved settlement preferences, which are specific for a state country pair at a given point in time. Finally, we use the exogenous allocation of the US refugee resettlement program, where migrants' only influence in the settlement decision is to decide whether to be close to family members already living in the United States or not. Given that we control for the number of previous immigrants, the exogenous variation in the migration decision is solely driven by the allocation of "free" application cases, i.e. those where the political refugee has no family ties and friends in the US.

Overall, we see our approach as complementary to that of [Parsons and Vézina \(2014\)](#) and as a test of external validity in a multi-country multi-period setting. [Parsons and Vézina \(2014\)](#) focus on a specific group of migrants at a given point in time. This paper uses data on refugees to the U.S. from all countries in the period of 1970 to 2005, leading to a more comprehensive sample, both in terms of years and countries of origin.

2.2 Refugees

Refugees are people who have fled their home country and cannot return because they have a well-founded fear of persecution based on religion, race, nationality, political opinion or membership of a particular social group (Immigration and Nationality Act, Sect. 101[a][42]). Each fiscal year, the President of the United States sends a proposal to congress for the maximum number of refugees to be admitted. After a congressional debate, the overall refugee admissions limit for the upcoming fiscal year is set. The limit varies from year to year. For example, in 2005, 53,738 refugees were admitted to the United States, primarily from Laos (8,487), Somalia (10,106) and the former

USSR (11,272). For more details on the number of refugees per country of origin , see Table 2.⁴

Note that this study focuses exclusively on refugees, who are distinct from asylum seekers. To become a refugee, people present their individual case before an Immigration and Naturalization Service officers in one of the refugee processing centers around the world. Once the application was successful, the Bureau of Population, Refugees and Migration (PRM), part of the State Department, handles the oversea processing and transportation to the United States. Upon arrival in the United States, the PRM allocates the refugees to voluntary resettlement agencies, who (1) place the refugees in one of their regional offices in the US and provide social services that foster the integration of the refugees.⁵

The allocation of refugees is not random. The voluntary agencies are responsible to settle refugees close to regional offices across the US states. They receive information from the State department about each refugee's characteristics, such as country of origin and demographic information. With this information, the voluntary agencies try to place refugees in the same community as their family members already living in the United States. Otherwise, refugees are placed in a different state to avoid the concentration of ethnic groups, as it was the case for Cuban refugees in Florida in the 1970s, see [Kerwin \(2012\)](#). The number of refugees per regional office also depends on the average employment statistics at the regional office. Offices in regions with lower unemployment rates will thus receive relatively more refugees. The Bureau of Population, Refugees and Migration encourages refugees to find employment quickly as it reduces the economic costs for their social services and it helps refugees to integrate faster into American society, see [Martin \(2004\)](#). Figure 3 plots the distribution of refugees across US states over the whole sample period, i.e. 1985 to 2010. The state that took in the most refugees is California with 277,395 refugees. New York ranked second with 144,273 refugees, followed by Florida with 111,657 refugees.

Since the resettlement of refugees depends on the capacity of the local organizations to take care of the refugees, we include state and country fixed effects. The time varying state fixed effects control for the potential correlation between a state's capacity to host refugees and the level of its income, and in further instance trade. To control for family reunification, we include the previous number of immigrants living in the state as a control variable. Finally, the country of origin country fixed effects control for country specific changes in the quota of refugees common to all states. For example, after the war on terror started the number of refugees from Iraq and Afghanistan increases significantly compared to earlier periods. Overall, controlling for time-varying state and country of origin fixed effects in combination with the number of previous immigrants living in the state will ensure the validity of the instrument.

⁴The United States has a special concern for a designed group of refugees related to religious activists or minorities in certain countries. This group includes Jews and Christians in the former Soviet Union with close family ties in the United States, civil right activists from Cuba, political refugees from Burma, Iranian members of certain religious minorities and Sudanese Darfurians.

⁵Figure 2 in the appendix shows the regional distribution of the voluntary resettlement agencies.

In the first stage, we regress the log of the number of immigrants (Imm_{ijt}) who arrived in the 5 years prior to year t from country j and settle in state i on the log of the number of refugees who arrived in the past 5 years (Ref_{ijt}). The first-stage equation is given by:

$$\log(Imm_{ijt}) = \alpha_0 + \alpha_1 \log Ref_{ijt} + \alpha_2 \log d_{ij} + \alpha_3 \log PImm_{ijt} + f_{it} + f_{jt} + \epsilon_{ijt} \quad (3)$$

where f_{it} and f_{jt} are country and state year pair fixed effects. As already mentioned, we include the number of previous immigrants, $PImm_{ijt}$, in state i from j to account for the fact that refugees may have a family member already living in that state. The state-year effects, f_{it} , control for any state specific change in the allocation of refugees over time that is common to all countries of origin. Country-year fixed effects, f_{jt} , control for country of origin specific effects that are common to all states in the United States, such as the nature of the conflict that forced people to emigrate or any other macroeconomic condition in the country of origin. We exclude the border effect because there are no refugees from Mexico or Canada who migrated to the United States in our sample period.

The main advantages of using refugees are that (1) refugees cannot freely choose where they settle in the United States, (2) the refugee sample includes a wide variety of countries and (3) these countries represent a large share of US imports, see Table 1⁶. However, the correlation between refugees and migrants may potentially be weak because only a limited number of refugees are allowed to enter the United States every year. In addition, the underlying reasons why refugees flee their home country vary greatly across countries and individuals and we have no information on the underlying circumstances that trigger the migration choice.

2.3 Consequences of Hurricane Mitch for Migration

To address the concerns about the generality of migration and trade effects based on political refugees, we also focus on the trade-migration relationship of Central American countries with the United States after hurricane Mitch. We argue that the natural disaster forced people to migrate, while they would not have done so otherwise. We show that (1) the increase in migrants in far away states is higher than in close-by states and (2) the increase in immigrants from hurricane-affected countries reduces the effect of distance on trade.

Hurricane Mitch hit Central America during the last week of October 1998. The countries impacted the strongest were Honduras and Nicaragua, but Guatemala and El Salvador (and to a much lesser extent Belize) were also affected. The hurricane destroyed a large part of these countries' transportation and social infrastructure, including roads, hospitals and schools, see [World-bank \(2001\)](#). According to the World Bank, people from Central America responded to the disaster by migrating to the United States. As a formal response to the migration generated by hurricane

⁶We discuss the summary statistics in detail below after presenting the data sources.

Mitch, the Immigration and Naturalization Service (INS) announced in a news release the designation of Temporary Protected Status (TPS) to countries hit by the hurricane for a period of 18 months, which was later extended, see [Kugler and Yuksel \(2008\)](#). As a consequence, Central Americans, who entered during this period, were not subject to deportation from the United States and, at the same time, were eligible to work in the United States. The INS estimates that by 2003, close to 50,000 Hondurans and Nicaraguans had been granted Temporary Protected Status (TPS) to allow them to stay and work in the United States. This massive inflow of Central Americans was accompanied by a large amount of illegal immigration, such that the number of Central Americans who came from these countries was probably higher than this official number. According to the immigration data from the US census, a total 637 thousand persons from El Salvador, Honduras, Nicaragua, Guatemala and Belize migrated to the United States, which was significantly higher than in previous years.⁷

Based on Figure 1, the migration decision is correlated with distance. Because these migrants have a preference for the goods from their home country and they carry information about the economic conditions in both countries, we consequently expect an increase in trade flows between the two regions. In particular, we expect a lower effect of distance on trade flows. The basic idea is that if the increase in immigrants in far away states is greater than in close-by states and if immigrants have trade enhancing effects, then far away states will experience a greater increase in trade relative to close-by states. As a result, we expect that the effect of distance on trade between the respective state and country of origin decreases due to hurricane Mitch.

The first-stage equation for the IV estimates is:

$$\log(Imm_{ijt}) = \alpha_0 + \alpha_1 \log d_{ij} dPost_{jt} + \alpha_2 \log d_{ij} + \alpha_3 b_{ij} + \alpha_4 \log PImm_{ijt} + f_{it} + f_{jt} + \epsilon_{ijt} \quad (4)$$

where we restrict our sample to Latin American countries only.⁸ We include destination-specific time effects (f_{jt}) as well as origin specific time effects (f_{it}) to check whether the increase in the share of immigrants in close-by states simply is due to an ongoing trend or whether there is indeed a discernible break after Mitch. $dPost_{jt}$ is a dummy that takes the value of 1 if the immigrant from country j arrived after hurricane Mitch and zero otherwise. Our instrument is represented by the interaction between a post-Mitch dummy and distance from Central America to various states in the United States.

An important assumption for the validity of our instrument is that the hurricane did not affect directly the export capabilities of the immigrants' country of origin and that the immigrants

⁷In the appendix we formally estimate the increase in the number of immigrants per US state following a difference in differences approach. The results show that the increase in the number of immigrants was 53 percent higher than the increase in immigrants from other Latin American countries.

⁸Latin American countries affected by hurricane Mitch are Belize, El Salvadore, Guatemala, Honduras and Nicaragua. The remaining Latin American countries, which constitute our control group, are Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, Mexico, Panama, Peru, Uruguay and Venezuela.

were the only factor that reduced the effect of distance on trade. Given our discussion about the hurricane destroying large parts of the local infrastructure and production facilities, exports from hurricane-affected countries decreased significantly after Mitch. For this reason, we include a country of origin-year fixed effect (f_{jt}) in regression 4.

In the previous paragraphs we described our estimation equation together with the IV strategy. We next describe the intermediate input demand model that generates the distribution of intermediate import flows across US states. The OLS and IV estimation results will follow in section 4.

2.4 Inferring state level imports

To project trade flows onto the state level, we use data on the industry production structure by state and apply the intermediate input demand model put forward by [Anderson and Yotov \(2010\)](#). The model distributes the aggregate industry specific trade flows from a given exporting country across states. To begin, consider the unobserved true level of intermediate imports of state i from country j in year t , (X_{ijt}). These state level imports can be written as the sum over all industries within the state:

$$X_{ijt} = \sum_{k=1}^K w_{ijt}^k X_{jt}^k \quad (5)$$

where w_{ijt}^k is the import share of industry k in state i from country j and X_{jt}^k represents the aggregate US import expenditure of industry k from country j . These aggregate US industry level imports from country j are observed. Hence, we estimate the unobserved regional import share. To do so, we follow [Anderson and Yotov \(2010\)](#). Assume that final good producers in a particular state i have the following CES demand for the aggregate quantity of imported intermediate inputs in industry k (Q_{it}^k):

$$Q_{it}^k = \left(\sum_{j=1}^J \beta_{jt}^k (q_{ijt}^k)^{1-1/\sigma_k} \right)^{\sigma_k/(1-\sigma_k)}$$

where q_{ijt}^k is the quantity of intermediate inputs from country j demanded by industry k in state i . β_j^k is a share parameter specific to the country of origin and the industry but common to all regions within the United States. σ_k is the industry specific elasticity of substitution. Industry specific intermediate import expenditure, $X_{it}^k = \sum_{j=1}^J X_{ijt}^k$, is then given by the quantity demanded, Q_{it}^k , times the price index, P_{it}^k . Given the CES demand function, we can express the bilateral import expenditure, X_{ijt}^k , as the value of shipments at destination prices, p_{ijt}^k , from the country of origin j to the state i in industry k .

$$X_{ijt}^k = \left(\frac{\beta_{jt}^k p_{ijt}^k}{P_{it}^k} \right)^{1-\sigma_k} X_{it}^k \quad (6)$$

Destination prices p_{ijt}^k are the product of the factory gate prices p_{jt}^k and trade costs $t_{ijt}^k \geq 1$ of the shipment of goods from j to i in industry k . The key assumption is that producers in country

j charge the same factory gate price to all buyers within the United States. Prices within the United States vary only due to industry-state specific trade costs. The CES price index is given by

$$P_{it}^k = \left[\sum_{j=1}^J \left(\beta_{jt}^k p_{ijt}^k \right)^{1-\sigma_k} \right]^{1/(1-\sigma_k)}.$$

Note that US imports from a given country simply the sum over all states, i.e. $X_{jt}^k = \sum_{i=1}^I X_{ijt}^k$. We write the normalized industry specific state import share as:

$$w_{ijt}^k = \frac{X_{ijt}^k}{\sum_{i=1}^I X_{ijt}^k} = \frac{\left(t_{ijt}^k / P_{it}^k \right)^{1-\sigma_k} X_{it}^k}{\sum_{i=1}^I \left(t_{ijt}^k / P_{it}^k \right)^{1-\sigma_k} X_{it}^k} \quad (7)$$

which depends on import expenditure of state i on the intermediate goods of industry k , the price level and trade costs. States with lower trade costs and higher expenditure import more.

In order to calculate the regional import share, we follow [Stumpner \(2013\)](#) and assume the following trade cost function

$$\log(t_{ijt}^k) = \delta_1 \log(\text{Distance}_{ij}) - \delta_2 \log(\text{Distance}_{ij}) \log(\text{Value-to-weight}_t^k) - \delta_3 \text{Border}_{ij}. \quad (8)$$

Trade costs increase with distance and decrease if the state shares a border with Mexico or Canada. Modeling trade costs using an interaction between distance and the value-to-weight ratio will capture the heterogeneous effect of distance on trade flows across industries. Trade costs in industries with a high value-to-weight ratio, for example computer chips, are less sensitive to distance than in industries with a low value-to-weight ratio, for example car engines. [Stumpner \(2013\)](#) presents evidence that this parameterization of trade costs reflects well US inter-state trade flows at the industry level. Unlike [Stumpner](#), we include a dummy variable if the region borders Canada or Mexico to account for potential border effects.

Equation 8 allows us to construct trade costs. To that goal, one needs estimates of the trade elasticities $\delta_1 - \delta_3$, which can be obtained from a gravity model of bilateral trade flows on the right hand side variables of equation 8. However, bilateral trade flows are not observed at the state level. To circumvent this problem, we offer two alternative approaches. First, we follow [Stumpner \(2013\)](#) and use his elasticities based on US inter-state trade: $\delta_1 = 0.667$, $\delta_2 = 0.081$ and $\delta_3 = 0$. However, these estimates are derived from trade between US states and do not take into account international flows. The resulting trade cost function may not be suitable if there are discontinuities in trade costs when crossing an international border. In an alternative approach, we estimate the trade elasticities using aggregate US industry specific international trade data and impose the estimated elasticities on the state level. Following the estimation procedure described in detail in the appendix, we obtain elasticities similar to [Stumpner](#), i.e. $\delta_1 = 0.691$, $\delta_2 = 0.154$ and $\delta_3 = 1.897$. Note that the results reported below are based on the elasticities using the international trade data and are robust to the values used by [Stumpner \(2013\)](#).

To calculate w_{ijt}^k in equation 7, we require information on X_{it}^k , which following [Autor et al. \(2013\)](#) can be proxied by L_{it}^k , industry k 's employment in state i . This assumption is reasonable if expenditure on intermediate imports is proportional to total production, which again is proportional to employment.⁹ Finally, we assume a common demand elasticity $\sigma_k = \sigma$ and that the law of one price within an industry k holds across the United States, i.e. $P_i^k = P^k$. The resulting weights are then given by

$$\hat{w}_{ijt}^k = \left(\frac{\left(t_{ijt}^k\right)^{1-\sigma} L_{it}^k}{\sum_{i=1}^I \left(t_{ijt}^k\right)^{1-\sigma} L_{it}^k} \right)$$

The intermediate import flow at the state level is simply the sum of the industry specific estimated weights times the industry specific national import flow. The calculated import flows in equation 5 form the basis of our empirical analysis. The following section explains in detail the underlying data needed to apply equation 5. Afterwards, we discuss the estimation results.

3 Data and summary statistics

This section describes the data used and provides basic summary statistics. Particular emphasis is placed on trade flow and industry data. The latter allows us to map aggregate trade flows across states at the industry level. We then provide an overview of the other datasets and discuss the summary statistics

3.1 Import data

The import data comes from the US Census Bureau provided by [Schott \(2008\)](#) for the years 1972 to 2005.¹⁰ The data contains import flows at the 5 digit SITC product level to 183 trading partners by mode of transportation, i.e. either air or sea. The import value of shipments is defined as the net selling value exclusive of freight charges and excise taxes. To calculate the intermediate import levels at the industry level, we use a correspondence table from the World Bank that converts 5 digit SITC codes into the Broad Economic Categories (BEC) developed by the United Nations. The BEC classification enables us to group each SITC code into intermediate and final goods so that we obtain product specific intermediate imports for each trading partner.¹¹ To make the trade flows directly comparable to the industry employment records at the state level, we convert the 5 digit SITC codes into SIC87 industry codes using the procedure described in [Feenstra \(1996\)](#). As a result, we obtain the intermediate import value for each 4 digit SIC 87 industry code (409 industries). These values are used as X_{jt}^k in equation 5.

⁹If we assume a Cobb-Douglas production function with labor (L), intermediate inputs and potentially other input factors, then demand for intermediate inputs is proportional to total labor employed.

¹⁰We consider the trade data of 1970 to be similar to the trade flows in the year 1972. This enables us to use an additional year of immigration data. In the appendix, we present the results when the year 1970 is excluded.

¹¹Intermediate goods are defined as the sum of the categories: Processed food and beverages (12), Industrial supplies (2), Capital goods (4) and Parts of transport equipment (53).

3.2 Trade cost data

The average value-to-weight ratio per SIC 87 industry is calculated as follows. For each 4 digit SIC 87 code, we divide the total import value by the total weight and convert it into tons. To compute the average value-to-weight ratio for each industry, take the average value-to-weight ratio across all trading partners for each mode of transportation, i.e. air or sea, and weight the importance of the transportation mode by its share in the total import value per industry.

To calculate the bilateral distances used in the trade cost function, we adopt the procedure used by [Mayer and Zignago \(2011\)](#). d_{ij} is the distance between the capital of state i and the capital of country j measured in kilometers and calculated by the Great Circle Distance Formula. Like [Mayer and Zignago \(2011\)](#), we use 32.19 kilometers as inner-city distance. All data on latitudes and longitudes are from the Global Administrative Areas (GADM) database, see [Hijmans et al. \(2010\)](#).

3.3 Industry data

To measure the industry structure at the state level, we follow [Autor et al. \(2013\)](#). We take the local employment structure from the County Business Patterns (CBP) for the years 1970, 1980, 1990, 2000 and 2005. The CBP is an annual data series that provides information on employment, firm size distribution and payroll by county and industry and covers all of the United States. Based on the data and the approach of [Autor et al.](#), we calculate total employment per state and SIC 1987 industry code and use it as a proxy for intermediate input demand in the regional import weight.¹²

3.4 Immigration data

Our measures of the immigrant population are based on data from the Census Integrated Public Use Micro Samples compiled by [Ruggles et al. \(2004\)](#) for the years 1970, 1980, 1990 and 2000, and the American Community Survey (ACS) for 2005. The 1980, 1990 and 2000 Census samples include 5 percent of the US population, while the pooled ACS and 1970 Census samples include 3 and 1 percent of the population respectively. The main explanatory variable in our regressions, i.e. recent immigrants, is defined as the number of immigrants who immigrated up to 5 years prior to the census year. In the robustness section, we control for pre-existing immigrants by all those immigrants who live in the respective state and immigrated ten years or more prior to the census year. We focus only on immigrants who are older than 21 at the census year. We then aggregate the number of immigrants at the state level using the census sampling weight.

¹²Given a Cobb-Douglas production function, the demand for intermediate inputs will then be proportional to total employment per industry. As long as this proportionality factor does not vary across US states within an industry, total employment will be a proxy for intermediate input demand within an industry.

As an alternative identification strategy, we use political refugees as an instrument for immigration in the import regression. Data on the number of refugees per US state come from the Office of Refugee Resettlement (ORR). The ORR provides the option to download refugee arrival data sorted by country of origin and state of initial resettlement in the United States for the years 1984 to 2005. Each fiscal year, the US government sets an overall refugee admissions limit based on regional allocations. The limit for refugee admissions varies from year to year depending on the Congress and the geopolitical situation. Figure 3 plots the number of refugees for each state over the period 1985 to 2005.¹³ In order to make the refugee data compatible with the immigration data, we add up the refugee data per country of origin and state for all 5 years prior to the census years 1990, 2000 and 2005 census years. Table 2 contains the total number of refugees per country of origin who arrived in the United States during our sample period.

3.5 Country treatment

Finally, we should mention that several countries changed their names and borders during the 1970-2005 sample period. Since we cannot identify the origin of earlier immigrants, for example the number of Ukrainian immigrants prior to 1990 in the United States, we keep the country borders of the year 1970. We group all imports and migrants from former Soviet republics into one country, i.e. the USSR¹⁴. Immigrants from Bosnia and Herzegovina, Croatia, Macedonia, Serbia and Montenegro, Kosovo, and Slovenia are all included in Yugoslavia. Slovakia and the Czech Republic form Czechoslovakia and Eritrean immigrants count as Ethiopians even though Eritrea became a separate nation in 1993. The exception is Germany, which we assume to be unified as of 1970.

3.6 Summary statistics

Table 1 presents the summary statistics for each of the 5 census years, 1970, 1980, 1990, 2000 and 2005. Columns 1 to 5 show the total number of immigrants, the number of newly arrived immigrants and the trade statistics for each census year. The total number of immigrants to the United States grew from 9.7 million in 1970 at an annual growth rate of 3.8% to 36.1 million in 2005, which is reflected in the steady rise of recent immigrants, i.e. those immigrants who arrived up to 5 years prior to the census year. From 1965 to 1970, only 1.7 million people arrived, whereas from 2000 to 2005 that number grew to 8.1 million. Immigration became also more diversified. The number of different countries from which immigrants migrated rose from 67 in 1970 to 138 in 2005. During the same period, the total value of intermediate imports increased from USD 2.3 billion in 1970 to USD 82.7 billion in 2005. This increase corresponds to an annual growth rate of 10.7%, much

¹³See US Department of Homeland Security. 2005 yearbook of immigration statistics. Washington, DC: US Department of Homeland Security. Available at http://www.dhs.gov/xlibrary/assets/statistics/yearbook/2005/OIS_2005_Yearbook.pdf.

¹⁴The former USSR includes the following present-day nations: Azerbaijan, Armenia, Belarus, Estonia, Georgia, Latvia, Lithuania, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan, Turkmenistan, Uzbekistan, and Ukraine.

higher than the growth rate of the immigrant population. Overall, the summary statistics suggest a positive relationship between the number of recent immigrants and intermediate imports.

To infer a causal link between migration and trade, we take advantage of the exogenous variation in migration in the form of a natural experiment. In particular, we use the emigration caused by hurricane Mitch and refugees from other geopolitical conflicts. For this reason, Table 1 splits the sample into three country groups: countries affected by hurricane Mitch, Latin American countries and refugee countries. Hurricane-affected countries are El Salvador, Guatemala, Honduras, Nicaragua and Belize. Latin American countries are all American countries beside the United States and Canada; they act as a control group for migration caused by hurricane Mitch. Refugee countries are all countries that sent refugees to the United States over the 1985-2005 period.

With respect to the different country groups, the total number of immigrants from hurricane Mitch affected countries only account for a small share of US immigration and trade. In 2005, they represented 6.7% of total immigrants, while their contribution to intermediate import flows was 0.18%. When compared to other Latin American countries, their share rises to 16% in terms of immigration and 1.4% in terms of total intermediate imports from Latin America to the United States. Given the low representation of hurricane-affected countries in overall trade, we compare the migration and import decisions of these countries relative to other Latin American countries. Below, we argue in more detail that these countries are more similar in their unobserved trade and migration characteristics, which helps to reduce potential omitted variable biases.

To further address concerns about the generality of the migration and trade effects in the case of Central America, we also focus on the trade-migration relationship for countries that sent political refugees to the United States. As in the case of migration caused by hurricane Mitch, refugees take their migration decisions independently of economic reasons and act as an instrument in our empirical analysis. The advantage of looking at the allocation of political refugees is that these refugees come from a wide range of countries, which trade a lot with the United States. Between 2001 and 2005, refugees from 83 countries settled across the United States and the combined import share of these countries was 72%.¹⁵ However, the correlation between refugees and migrants may be weak because the share of refugees in the total number of immigrants is only 3.3%.

4 Results

This section shows empirically that the migration channel was important for increasing intermediate imports across US states. We start by a simple OLS regression of equation 2. The results are presented in Table 3. Columns (1) to (3) show the estimation results for all countries, where we observe migration and trade data. Columns (4) to (6) contain the results for the restricted sam-

¹⁵Table 2 contains the number of refugees per country of origin over the whole sample period, while figure 3 plots the distribution of refugees across states.

ple, where we only consider Latin American countries. Overall, we find a positive and significant migrant trade effect on intermediate imports across all specifications. The baseline OLS results in columns 1 and 4 suggest that a 10% increase in the number of recent migrants raises intermediate imports by 0.49% in the full and 0.78% in the restricted sample. When testing for significant differences between the restricted and the full sample, the results show that the effect is significantly higher for Latin American countries.

Both the magnitude of the migration elasticity for Latin American countries (0.0784) and for all countries (0.049) are significantly lower than the values found in the literature. Looking at the national level of US imports for the period 1870 to 1910, [Dunlevy and Hutchinson \(1999\)](#) found a value of 0.28. Other studies at the regional level also found significantly higher values. [Combes et al. \(2005\)](#) found a value of 0.12 in the case of France. [Wagner et al. \(2002\)](#) found a value of 0.28 using a variation across Canadian provinces and [Bratti et al. \(2011\)](#) found a value of 0.32 for Italy. There are at least two explanations for the low migration elasticities. First, the literature does not distinguish between intermediate and final goods. The above-mentioned papers focus exclusively on aggregate imports and consider both the preference effect and the trade cost reducing effect of immigrants. Second, this paper focuses on newly arrived immigrants. Thus, it sets their effect on trade apart from the already established trade network effects created by previous immigrants.

In columns (2) and (3) as well as (5) and (6), we add the number of previous immigrants as an additional control variable. It is likely that the migration decision of recent immigrants is not random and influenced by the presence of previous immigrants from the same country of origin. If the omitted variable drives both trade and migration, the previous results will be biased. For example, assume migrants move to states where there are already previous migrants for cultural or economic reasons. If these regions then experience positive demand shocks that lead to more production and employment, it is unclear whether the demand for intermediate imports increases because of the new immigrants or just because of the positive demand shock on the already established trade network. As a result, we expect the trade cost effect through the arrival of recent immigrants to be lower than in the absence of previous immigrants. The resulting bias from the omission of previous immigrants in the regression may overestimate the true effect.

To control for the number of previous immigrants, we follow two distinct approaches because immigrants are likely to change location between the year they arrived in the United States and the year the census took place. In columns (2) and (5) we add the number of immigrants who arrived 10 or more years ago and reported to be living in state i in the year the census was conducted. In columns (3) and (6) we use the number of immigrants before the first year of trade is observed. More precisely, we use the number of immigrants who arrived up until in 1960 and reported to be living in state i in 1970. Previous immigration will be correlated with actual immigration if recent immigrants prefer to settle in states where there exists a large pre-existing community. However, the specification is less demanding because it is based on the distribution of immigrants in the

year 1960.

Table 3 presents the results. When we control for pre-existing immigrants, the migration elasticity on intermediate imports increases slightly. For Latin American countries, the coefficient decreases from 0.078 to 0.062 when using the number of immigrants who arrived ten years ago and from 0.049 to 0.04 for the full sample. In the case of the immigrant distribution in the year 1960, the coefficient increases to 0.07 and 0.012 respectively. However, while these coefficients are not significantly different from the baseline results, they point to a potential upward bias in excluding the previous immigration. Note that the combined effect of previous and recent immigrants is larger and comparable to values found in the literature.

4.1 Instrumental variable results

Note that the OLS results might be biased, for example if people immigrate in order to take advantage of trading opportunities, i.e. trade causes immigration. To resolve the question of endogeneity, we consider the following two approaches. First, we use geopolitical conflicts that forced people to leave their home country. Second, we make use of a natural disaster, Hurricane Mitch. We start by discussing the results based on political refugees before showing the results in the case of hurricane Mitch.

Refugees

The first stage results are presented in columns (1) to (3) of Table 4. First, the number of refugees is positive and significantly correlated with the number of immigrants in all specifications. Note that the positive effect decreases in magnitude once we add the number of previous immigrants as an additional control variable, see columns (2) to (3). The reason is that part of the refugees are settled in states where they have family members living already in the US.

Before turning to the IV baseline results, we say a few words on the validity of our instrumental variable. To confirm the validity of the IV regressions, we include the Cragg-Donald F statistics and the Kleibergen-Paap test p-values at the bottom of our tables. The Kleibergen-Paap test examines whether the excluded instrument is correlated with the endogenous regressors conditional on the control variables. A p-value below 0.1 indicates that we can reject the null hypothesis of no correlation and that the instrument is statistically significant. The Cragg-Donald F statistic provides an indication of the significance of our instrument. If the instrument is only weakly correlated with the endogenous regressor, the IV estimator is not valid. To assess the weakness of our instrument we need to compare these F statistics with the Stock-Yogo critical values for the Cragg-Donald F-statistic with one endogenous regressor (Stock and Yogo (2002)). As a rule of thumb, an F-statistic above 10 indicates that the IV is acceptable.

Columns (4) to (6) in Table 4 contain the results of the second stage IV regression. The coefficient on the number of recent immigrants is significantly positive with an elasticity higher than in the OLS regression for all specifications. However, performing a Hausman test does not reveal any significant difference in the coefficients. Overall, our results suggest that a 10% increase in the number of immigrants in the past five years increases intermediate imports by 1.8%.

There are potential concerns that countries where political refugees are coming from do not form a representative sample for US trade. Table 2 shows that the majority of refugees are either from current or former communist countries (USSR, Vietnam, Cuba) or from countries with whom the US has political tensions (Cuba, Iran, Myanmar). Furthermore, correlation between refugees and migrants may potentially be weak because only a limited number of refugees are allowed to enter the United States every year, i.e. the share of refugees in terms of total number of immigrants is 3-5 percent. For this reason, we consider an alternative instrument, i.e. hurricane Mitch. The argument supporting the validity of the instrument is outlined in section 2 and parallels the argument in the case of political refugees. Hurricane Mitch forced people to migrate who would have not done so otherwise.

Hurricane Mitch

Columns (1) to (3) in Table 5 present the results of the first stage IV regression with different control variables. First, distance has a significant negative effect: immigrants move to states that are closer to their country of origin. The variable instrument, capturing the effect of distance after the hurricane, is significant and positive. While the overall effect of distance on the migration decision of Central Americans is still negative, migrants from Central America moved relatively further away from their country origin after the hurricane. Note that this effect is robust to the inclusion of the previous stock of previous immigrants in the state, see columns (2) and (3).

Focusing on the p-value of the Kleibergen-Paap test, we can reject the null hypothesis of no correlation between the endogenous variable and the instrument at the 1% level in all specifications. The Cragg-Donald F statistic is higher than 10 and therefore increases our confidence in the validity of the instrument.

Columns (4) to (6) in Table 5 contain the second stage IV regression, where the number of recent immigrants is instrumented by the log of distance interacted with a post-time dummy for hurricane Mitch. The coefficient on recent immigrants is significant and positive with an elasticity higher than in the OLS regression for all specifications. Thus, states that received more Central American immigrants after hurricane Mitch face a lower distance elasticity and import more. The estimated immigrant elasticity is 0.13 and comparable in magnitude to the effect when using political refugees as an exogenous migration variation.¹⁶

¹⁶As noted in section 2, the identification assumption is that the only reason why the elasticity of trade with respect to distance between US states and Central American countries changed is because (1) these states received relatively more

Overall, the similarity between the estimated coefficients leads us to conclude that immigrants have a significant and positive impact on intermediate imports. A 10 percent increase in the number of immigrants in the past five years increases intermediate imports by 1.5 to 1.8 percent.

5 Robustness

In this section, we check the robustness of our main findings using an IV strategy, conducting the following two falsification/robustness exercises. First, we address the potential measurement bias that arises from the fact that we construct intermediate import flows via a demand model. To do so, we use recent observed US state import data and compare them to the constructed import flows. Second, given our discussion about the fact that after hurricane Mitch Central American migrants moved predominantly to the United States, we should find that Central American exports to the United States increased more strongly than those to other countries. The following paragraphs address each of these points separately.

5.1 Observed import data

As mentioned above, we do not observe intermediate imports over the period from 1970 to 2005. However, the US Census provides state specific import flow data for the years 2008 to 2012. For these years, we can compare the estimated import flows with the observed import flows and check the robustness of our estimation method. We proceed in two stages. First, we use the import demand model from section 2 together with the national US intermediate import data and generate state specific import data for the years 2008 to 2012. We then compare this estimated import flow with the observed import flow for the same years. If the model works well, we should expect a high correlation between the two flows. Second, we compare the estimated migration elasticity from the observed import flows with the elasticity based on the estimated import flows in the previous section. If the measurement error of the demand model is uncorrelated with immigration, then the two elasticities should not be significantly different from each other.

To construct the import flow data, we apply equation 5 to the trade cost function specified in equation 7 and use the industry and aggregate US trade data for the years 2008 to 2012. We then calculate a correlation coefficient between the two series and obtain a coefficient of 0.85. This suggests that our procedure approximates the true import flows quite well.

We then use the estimation equation 2 and replace the calculated import flow with the observed import flow. We focus on the cross section of the year 2010. The reason for this is that we do not

immigrants from Central America and (2) immigrants have trade enhancing effects by providing information or having preferences for goods of their country of origin.

have trade data over a 5 year interval so as to use the number of immigrants who arrived in the past 5 years as an explanatory variable.¹⁷ The precise regression model for the year 2010 is:

$$\log(X_{ij}) = \beta_1 \log(Imm_{ij}) + \beta_2 \log d_{ij} + \beta_3 b_{ij} + \beta_4 \log PImm_{ij} + f_i + f_j + v_{ij} \quad (9)$$

Table 6 plots the OLS estimates for Latin American countries and for all countries. The results are similar to the OLS estimates in the constructed sample, see Table 3. The level of the coefficients is similar and the elasticity for Latin American countries is higher than for all countries. The only exception is that in the Latin American sub-sample the immigration elasticity is not significant when we include the number of immigrants who migrated 10 or more years before.

Since the hurricane occurred in 1998, our instrumental variable identification strategy for the year 2010 is solely based on the number of refugees who came to the United States between 2005 and 2010. Table 7 shows the results. The first stage results are reported in columns 1 to 3. We find that the number of refugees is positively correlated with the number of immigrants over the past five years in all specifications. The F statistics are larger than 10 and the Kleibergen-Papp P-value are smaller than 0.05, which suggests that our instrument is appropriate. Columns 4 to 6 present the second stage IV. The number of immigrants has a significant impact on the trade level in all but column 5, where we include the number of immigrants who came at least ten years ago to the United States. Overall, the results and the magnitude of the coefficients are similar to our previous results. These findings suggest that our results are robust to the introduced measurement error of the calculated trade flows.

5.2 Effect on the exports of the country of origin

In the previous section, we use the variation in immigrants and Central American exports across US states. Next, we use the variation in the export revenue of Central American countries across all possible destinations in the world. According to our discussion in the previous section, Central Americans mainly responded to the disaster by migrating to the United States. Given this argument, Central American exports to the United States should be higher than those to other countries. We consider the following regression

$$\log(X_{ijt}) = \beta_1 D_{US,2000} + d_{ij} + d_{jt} + \epsilon_{ijt} \quad (10)$$

where the X_{ijt} is total intermediate exports from country i to country j in year t . Exporting countries are El Salvador, Guatemala, Honduras, Nicaragua and Belize. Importing countries are all

¹⁷For example, if we use the number of immigrants who arrived in the past 5 years as an explanatory variable in the year 2012 within the panel structure, then the previous import flow should correspond to the year 2007. This ensures independence between the previous trade flow and the arrival of immigrants in the past 5 years. If the two periods overlap, the number of immigrants and the trade flow are correlated. Note that we do not observe the import data for the year 2007. Once the trade data of the year 2013 become available, we can conduct a panel analysis and use the period 2008 to 2013. We experiment with the period 2008 to 2012 and use the number of immigrants who arrived in the past 4 years as an explanatory variable. The results are quantitatively similar to the ones reported in Table 6.

countries in the world. To focus on long-term effects, we consider 5 year intervals. The years considered are 1975, 1980, 1985, 1990, 1995, 2000 and 2005. $D_{US,2000}$ is a dummy variable that equals 1 if the year is 2000 or 2005 and the importing country is the United States. Otherwise, the dummy takes the value zero. This dummy captures whether exports from hurricane-affected countries to the United States were significantly higher than to other countries after the hurricane in the year 2000. The model includes importing country year fixed effects, d_{it} , as well as exporting-importing country pair fixed effects d_{ik} . d_{it} controls for any exporting country-time specific effects, such as aggregate demand shocks. d_{ik} is a dummy variable that controls for any pair fixed effects, such as distance, preferences or other bilateral characteristics that are constant over time.

Table 8 shows the results. The dummy coefficient is significant, which shows that between 1995 and 2000 exports from hurricane-affected countries increased significantly in comparison to other countries. The average per year effect is: $(1.81/1)^{(1/5)} - 1 = 12.6\%$. While we cannot attribute the increase in exports entirely to the number of emigrants to the US, it is consistent with the idea that migration fosters trade between countries.

6 Conclusion

Migrants carry information about the goods and market conditions of the country of origin as well as the country of destination. By providing this information, migrants reduce bilateral trade costs and increase trade between countries. This paper focuses on the trade cost reducing effect of migrants by looking at the relationship between intermediate import demand and immigration across US States. Using the exogenous migration decision brought about by a natural experiment (hurricane Mitch), we establish the causal relationship between immigrants and imports. Our results indicate a strong positive impact of migration on trade. We find that a 10 percent increase in immigrants raises trade by 1.5 to 1.8 percent.

Policy makers may want to take into account the pro-trade effect of immigrants. By providing information on market conditions in both countries, the country of origin as well as the country of destination, they reduce transaction/trade costs for importers and exporters. As a result, industries can purchase cheaper intermediate inputs from abroad and produce more efficiently.

Taking a broader perspective, immigrants may also have knowledge of production techniques used in their country of origin, which can increase the comparative advantage of industries in their country of destination, see [Bahar and Rapoport \(2014\)](#) for empirical evidence on this issue. All in all, these results suggest that the mobility of people between countries can serve as a key element in enhancing industrial productivity growth. However, more research is needed to assess the long-term impact of immigration on the economy.

References

- ANDERSON, J. E. AND Y. V. YOTOV (2010): "The changing incidence of geography," *American Economic Review*, 100, 2157–2186.
- AUTOR, D. H., D. DORN, AND G. H. HANSON (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, Forthcoming.
- BAHAR, D. AND H. RAPOPORT (2014): "Migration, knowledge diffusion and the comparative advantage of nations," Tech. rep., Harvard University Working Paper.
- BARDHAN, A. D. AND S. GUHATHAKURTA (2004): "Global linkages of subnational regions: coastal exports and international networks," *Contemporary Economic Policy*, 22, 225–236.
- BORJAS, G. J. (1999): "The economic analysis of immigration," *Handbook of labor economics*, 3, 1697–1760.
- BRATTI, M., L. DE BENEDICTIS, AND G. SANTONI (2011): "On the pro-trade effects of immigrants," *Review of World Economics*, 1–38.
- CARD, D. (2001): "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 19, 22–64.
- CLARK, X., T. J. HATTON, AND J. G. WILLIAMSON (2007): "Explaining US immigration, 1971–1998," *The Review of Economics and Statistics*, 89, 359–373.
- COMBES, P.-P., M. LAFOURCADE, AND T. MAYER (2005): "The trade-creating effects of business and social networks: evidence from France," *Journal of International Economics*, 66, 1–29.
- DUNLEVY, J. A. (2006): "The influence of corruption and language on the protrade effect of immigrants: Evidence from the American States," *Review of Economics and Statistics*, 88, 182–186.
- DUNLEVY, J. A. AND W. K. HUTCHINSON (1999): "The impact of immigration on American import trade in the late nineteenth and early twentieth centuries," *Journal of Economic History*, 59, 1043–1062.
- FEENSTRA, R. C. (1996): *US imports, 1972–1994: Data and concordances*, vol. 5515, National Bureau of Economic Research.
- FELBERMAYR, G. J., V. GROSSMANN, AND W. K. KOHLER (2012): "Migration, international trade and capital formation: Cause or effect?" Tech. rep., Discussion Paper Series, Forschungsinstitut zur Zukunft der Arbeit.
- GIRMA, S. AND Z. YU (2002): "The link between immigration and trade: Evidence from the United Kingdom," *Review of World Economics/Weltwirtschaftliches Archiv*, 138, 115–130.

- GOULD, D. M. (1994): "Immigrant links to the home country: empirical implications for US bilateral trade flows," *The Review of Economics and Statistics*, 302–316.
- HANSON, G. H. (2010): "International Migration and the Developing World," *Handbook of Development Economics*, 66, 4363–4414.
- HEAD, K. AND T. MAYER (2013): *Gravity equations: Workhorse, toolkit, and cookbook*, Centre for Economic Policy Research.
- HEAD, K. AND J. RIES (1998): "Immigration and trade creation: econometric evidence from Canada," *Canadian journal of economics*, 47–62.
- HIJMANS, R., N. GARCIA, AND J. WIECZOREK (2010): "GADM: database of global administrative areas," .
- KERWIN, D. (2012): "The faltering us refugee protection system: Legal and policy responses to refugees, asylum-seekers, and others in need of protection," *Refugee Survey Quarterly*, 31, 1–33.
- KUGLER, A. AND M. YUKSEL (2008): "Effects of low-skilled immigration on US natives: evidence from Hurricane Mitch," Tech. rep., National Bureau of Economic Research.
- MARTIN, D. (2004): *The United States Refugee Admissions Program*, Migration Policy Institute.
- MAYDA, A. M. (2006): "Who is against immigration? A cross-country investigation of individual attitudes toward immigrants," *Review of Economics and Statistics*, 88, 510–530.
- MAYER, T. AND S. ZIGNAGO (2011): "Notes on CEPII's distances measures: The GeoDist database," *CEPII Working Paper 201125*.
- PARSONS, C. AND P.-L. VÉZINA (2014): "Migrant Networks and Trade: The Vietnamese Boat People as a Natural Experiment," Tech. rep., Economic Research Institute for ASEAN and East Asia (ERIA).
- PERI, G. AND F. REQUENA-SILVENTE (2010): "The trade creation effect of immigrants: evidence from the remarkable case of Spain," *Canadian Journal of Economics/Revue canadienne d'économique*, 43, 1433–1459.
- RUGGLES, S., M. SOBEK, T. ALEXANDER, C. A. FITCH, R. GOEKEN, P. K. HALL, M. KING, AND C. RONNANDER (2004): "Integrated Public Use Microdata Series (IPUMS): Version 3.0," *Minneapolis, MN: Minnesota Population Center. Available at: <http://www.ipums.org>. her, GA (1966).* .
- SCHOTT, P. K. (2008): "The relative sophistication of Chinese exports," *Economic Policy*, 23, 5–49.
- STOCK, J. H. AND M. YOGO (2002): "Testing for weak instruments in linear IV regression," .
- STUMPNER, S. (2013): "Trade and the Geographic Spread of the Great Recession," Tech. rep., Mimeo, University of Berkeley.

WAGNER, D., K. HEAD, AND J. RIES (2002): "Immigration and the Trade of Provinces," *Scottish Journal of Political Economy*, 49, 507–525.

WOOLDRIDGE, J. M. (2010): *Econometric analysis of cross section and panel data*, MIT press.

WORLDBANK (2001): *Hurricane Mitch: The Gender Effects of Coping with Crisis*, World Bank: Washington DC.

7 Appendix

7.1 Estimation of industry specific trade costs

To consistently estimate the effect of trade costs on trade flows, we follow the literature on the estimation of gravity equations (see [Head and Mayer \(2013\)](#)). A wide variety of trade models yields an expression for trade flows that can be written in log form as follows:

$$\log(X_{ijt}^k) = \log(t_{ijt}^k) + f_{it}^k + f_{jt}^k + e_{ijt}^k \quad (11)$$

In this expression, f_i^k denotes all factors that promote exports of industry k in country i to all destinations, and f_j^k all factors that promote imports. σ represents the trade cost elasticity. We are only interested in estimating the bilateral trade costs independently of the origin and destination specific effects. To do so, we follow [Stumpner \(2013\)](#) and consider the following industry specific trade costs function

$$\log(t_{ijt}^k) = \delta_1 \log(d_{ij}) - \delta_2 \log(d_{ij}) \log(v_t^k) - \delta_3 b_{ij} \quad (12)$$

where d_{ij} is the distance measured from the capital in state i to the capital in country j . v_t^k is the value-to-weight ratio calculated by dividing the total value in dollars by the total weight in tons. b_{ij} indicates whether state i shares a border with country j . We can now plug in the trade cost function, equation 12, into the trade flow equation, equation 11, and estimate the distance, value-to-weight and border elasticities.

$$\log(X_{ijt}^k) = \delta_1 \log(d_{ij}) - \delta_2 \log(d_{ij}) \log(v_t^k) - \delta_3 b_{ij} + f_{it}^k + f_{jt}^k + e_{ijt}^k \quad (13)$$

Since we do not observe regional trade flows, we cannot estimate equation 13 at the state level. Instead, we consider the following two alternatives for constructing trade costs. First, we assume that the inter-state trade cost function from [Stumpner \(2013\)](#) also holds at the international level. He finds $\delta_1 = 0.666$, $\delta_2 = -0.0803$ and $\delta_3 = 0$. Alternatively, we can estimate equation 13 at the international level using aggregate US international trade flows and assuming that the parameters are valid at the regional trade level. To estimate the effect of trade costs on trade flows, we re-write equation 13 at the national level by aggregating trade flows over all US states. US intermediate imports, X_j^k , from country j in year t are now given by:

$$\log(X_{jt}^k) = \delta_1 \log(d_j) + \delta_2 \log(d_j) \log(v_t^k) + \delta_3 b_j + \delta_4 w_{jt} + d_t^k + e_j^k \quad (14)$$

where d_j is the distance from the capital of country j to the capital of the United States. b_j is a border dummy equal to 1 for Mexico and Canada and v_t^k is the value-to-weight in US dollars per ton shipped. In addition, we use industry-time fixed effects to control for industry specific variations. w_{jt} are control variables that are common to all regions within the United States but may be correlated with distance and/or border. The bilateral control variables are a common currency indicator, a trade agreement indicator and an indicator of whether the partner country shares a common language with the United States. Furthermore, we include PPP converted GDP

and population data from the Penn World Table. Results are in Table 9. As expected, they show a strong negative effect of distance on trade flows, which is muted in industries with high value-to-weight ratios. Sharing a border with the United States increases exports to the United States.

Given the estimated parameters, we assume that the estimated aggregate trade cost elasticities ($\hat{\delta}_1 - \hat{\delta}_3$) hold at the state level. As a result, we calculate industry specific trade costs using:

$$\log t_{ijt}^k = 0.691 \log d_{ij} - 0.154 \log d_{ij} \log v_t^k - 1.897 b_{ij}$$

where d_{ij} represents the distance from the capital of state i to the capital of country j . b_{ij} equals one if the state i shares a land border with Canada or Mexico and v_t^k remains the average value-to-weight ratio of industry k at year t .

7.2 Calculation of the value-to-weight ratio

We calculate the average value-to-weight ratio per SIC 87 industry as follows. The data used are drawn from [Feenstra \(1996\)](#) for 1972 and 1980 and [Schott \(2008\)](#) for the years 1990, 2000 and 2005. They record the customs value of all US imports by exporting country and year from 1990 to 2005 according to SIC 87 codes and the corresponding value and weight of trade shipped by air and shipped by sea. For the years 1972 and 1980, we convert the seven-digit Tariff Schedule of the United States (TSUSA) to SIC87 codes following the correspondence table of [Feenstra \(1996\)](#). Next, we compute the total import value and the average weight for each industry code and for each transportation mode, i.e. shipped by air or by sea. To compute the average value-to-weight ratio, we divide the total value shipped by its weight and convert it into tons for each mode of transportation and industry pair. We then weight the importance of the transportation mode by its share in the total per industry import value.

7.3 Calculation of intermediate imports

We classify goods according to the Broad Economic Categories (BEC) industry classification scheme of the United Nations. Intermediate goods are defined as the sum of the categories: Processed food and beverages (12), Industrial supplies (2), Capital goods (4) and Parts of transport equipment (53). To map the import data into the BEC categories, we use the World Bank concordance table between SITC Revision 2 codes and BEC categories. To convert the SITC Revision 2 codes into SIC 1987 codes, we follow the procedure outlined in [Feenstra \(1996\)](#). Finally, intermediate imports are simply the total value of all intermediate imports per SIC code.

7.4 Immigration response to hurricane Mitch

In this paragraph, we test whether hurricane Mitch increased immigration. We apply the difference in differences estimation procedure. Consider the following regression model

$$\log Imm_{ijt} = \beta_0 + \beta_1 X_{ij} + \beta_2 D_{ijt} + d_{ij} + b_{ij} + f_{it} + f_{jt} + \epsilon_{ijt}$$

where Imm_{ijt} is the number of immigrants from country j living in state i at time t . X_{ij} is a dummy variable equal to 1 if immigrants are from the hurricane-affected countries. d_{ij} controls for distance and b_{ij} for border effects between state i and country j . f_{it} are state-year and d_{jt} country of origin year fixed effects. D_{ijt} is a dummy variable equal to 1 if country j was affected by hurricane Mitch and the observation period is year 2000 or 2005. ϵ_{ijt} is an error term. The variable of interest is β_2 and measures the effect of hurricane Mitch on immigration from the affected countries, i.e. El Salvador, Guatemala, Honduras and Nicaragua, to the United States. The control group consists of Latin American immigrants, i.e. immigrants from the following countries: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, Mexico, Panama, Paraguay, Peru, Uruguay and Venezuela. We distinguish between two definitions of immigrants: standard and new. A standard immigrant is a foreign born individual who migrated to the United States at any point in time, while a new immigrant is a foreign born individual who migrated to the United States within the last 5 years.

Table 10 shows the results. The treatment dummy is significant and negative, implying that the level of migration from hurricane-affected countries is significantly lower than from other Latin American countries. The coefficient on the difference in differences dummy is statistically significant. The number of immigrants and the number of newly arrived immigrants increased significantly after hurricane Mitch. Column (1) shows that while the number of immigrants from affected countries is 74.3 percent lower than from other Latin American countries, the number of immigrants from those countries surged by 53 percent after the hurricane. Column (2) presents the results for the new immigrants who arrived during the last 5 years. The estimates show that the number of new immigrants from affected countries increased more strongly than those from other Latin American countries. After the hurricane, the number of new immigrants per state posted a stronger rise, climbing by an additional 75.3 percent. Overall, the evidence in Table 10 supports the view that hurricane Mitch caused significantly more immigration from the affected countries to the United States than from other Latin American countries.

8 Figures

Figure 1: The log of the number of immigrants who immigrated from hurricane Mitch affected countries (Nicaragua, El Salvador, Honduras, Guatemala and Belize) over the period from 1998 to 2005.

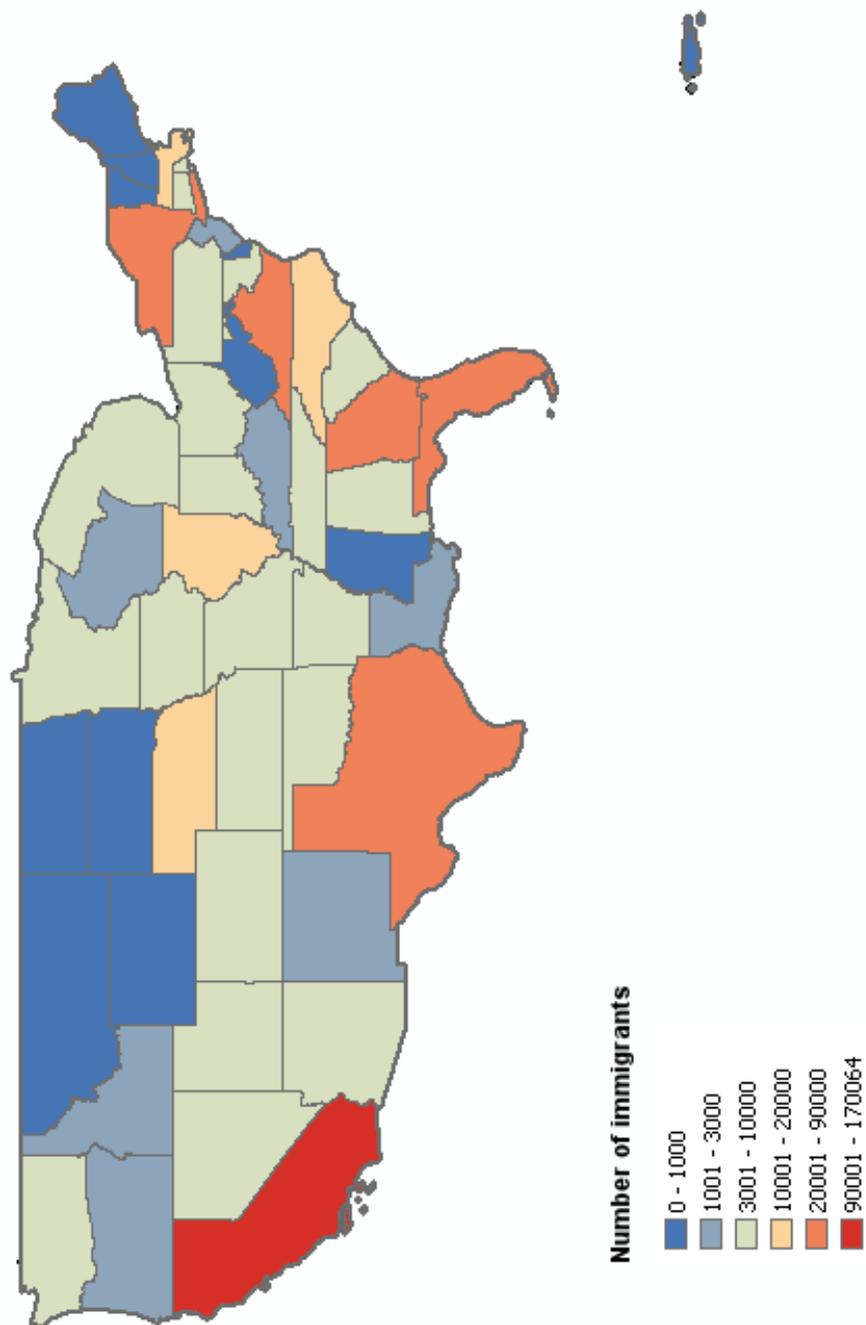
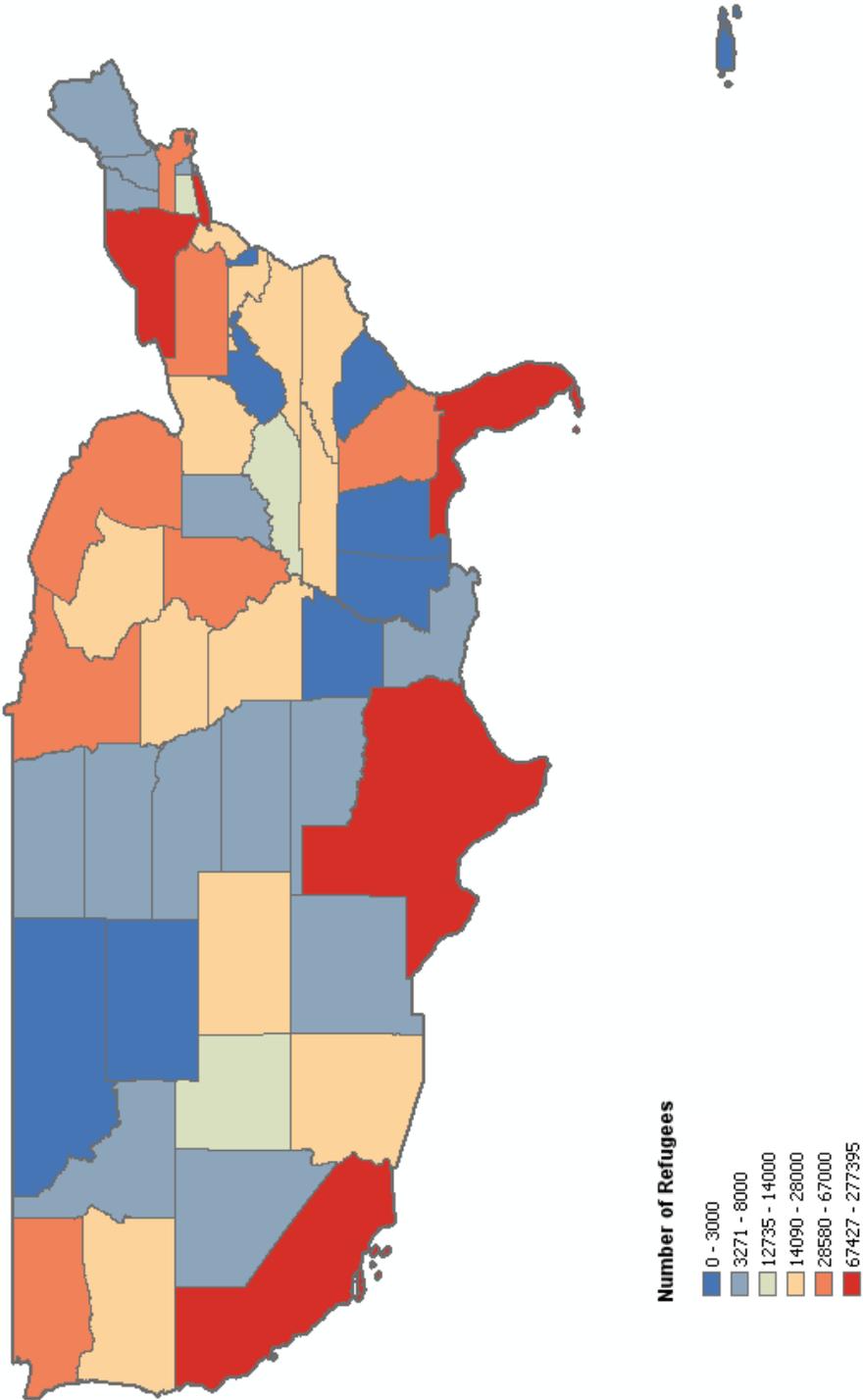


Figure 3: The log of the number of refugees who immigrated to the United States over the period from 1985 to 2005.



9 Tables

Table 1: Summary Statistics

		1970	1980	1990	2000	2005
Total	Nr. of immigrants (thousands)	9,752	13,700	23,100	32,300	36,100
	Nr of immigrants last 5 years (thousands)	1,708	3,054	5,508	7,783	8,056
	Value of Imports (USD millions)	2330	8510	24500	66100	82700
	Number of different nationalities among immigrants	67	148	149	127	138
Latin American countries	Nr of immigrants (thousands)	1,189	3,089	7,566	13,100	15,500
	Nr of immigrants last 5 years (thousands)	360	997	2,292	3,682	4,125
	Share of immigrants (%)	12.20	22.55	32.74	40.47	42.88
	Share of immigrants last 5 years (%)	0.21	0.33	0.42	0.47	0.51
	Value of Imports (USD millions)	133	621	2270	9550	12700
	Share of US Imports	5.71	7.30	9.27	14.45	15.36
Hurricane affected countries	Nr of immigrants (thousands)	80	253	1,051	1,871	2,260
	Nr of immigrants last 5 years (thousands)	29	105	410	439	538
	Share of immigrants (%)	0.82	1.85	4.55	5.79	6.26
	Share of immigrants last 5 years (%)	1.72	3.43	7.44	5.65	6.68
	Value of Imports (USD millions)	3.62	13.3	56.2	101	145
	Share of US Imports	0.16	0.16	0.23	0.15	0.18
Refugee countries	Nr of immigrants (thousands)			8,664	12,567	25,993
	Nr of refugees last 5 years (thousands)			288	425	266
	Share of immigrants (%)			37.51	38.91	72.00
	Share of refugees in immigrants last 5 years (%)			5.23	5.46	3.30
	Value of Imports (USD millions)			4340	15300	38600
	Share of US Imports			17.71	23.15	46.67
	Number of different nationalities among refugees			56	67	83

Note: Affected countries: Belize, El Salvadore, Guatemala, Honduras and Nicaragua. Latin American countries: Argentina, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvadore, Guatemala, Honduras, Mexico, Nicaragua, Panama, Peru, Uruguay and Venezuela. Value of exports and imports in current US dollars. Number of immigrants in absolute numbers. 1980, 1990 and 2000 are the respective Census years from the 5% sample. 1970 from the 1% sample and 2005 from the American Community Survey.

Table 2: Refugee-Sending Countries over the 5 years prior to the census years: 1985 to 1990, 1995 to 2000 and 2001 to 2010

Refugees	Country	Refugees	Country	Refugees	Country
367995	USSR	267	Yemen	24	Senegal
284999	Vietnam	252	Central African Republic	21	Lesotho
207944	Cuba	249	Cote d'Ivoire	20	Seychelles
168788	Yugoslavia	239	Syria	19	Bangladesh
113607	Somalia	216	Lebanon	15	Germany
108372	Laos	213	Jordan	14	Benin
85652	Iran	201	Sri Lanka	14	Morocco
77766	Iraq	187	Chad	11	Argentina
66419	Myanmar	179	Malaysia	11	Spain
55781	Cambodia	178	South Africa	10	United Arab Emirates
46875	Liberia	162	Gambia, The	10	Saudi Arabia
40797	Sudan	161	Nepal	9	Madagascar
39682	Ethiopia	151	Egypt	9	Venezuela
32096	Afghanistan	148	Indonesia	8	Burkina Faso
30873	Bhutan	147	Cameroon	8	Greece
21088	Poland	123	Algeria	8	Italy
20596	Romania	109	Kuwait	7	United Kingdom
12856	Haiti	95	North Korea	6	Bahrain
12658	Sierra Leone	93	Ghana	6	Brunei
10909	Burundi	91	Guinea	5	Mali
9833	Thailand	88	Equatorial Guinea	4	Antigua and Barbuda
9683	DR Congo	83	Djibouti	4	Maldives
5806	Czechoslovakia	74	Zimbabwe	4	Mexico
4207	Hungary	61	Philippines	4	Panama
2506	Colombia	61	Turkey	3	Canada
2476	Rwanda	51	Namibia	3	Iceland
1801	Kenya	50	Honduras	2	Belgium
1511	Nigeria	50	Mozambique	2	Bahamas
1333	Togo	46	India	2	Guinea-Bissau
1258	Republic of Congo	40	Zambia	2	Peru
801	Bulgaria	37	El Salvador	2	Paraguay
666	Israel	36	Gabon	1	Switzerland
631	Nicaragua	34	Botswana	1	Guatemala
611	Albania	34	Costa Rica	1	Jamaica
427	Angola	34	Ecuador	1	Japan
404	Mauritius	31	Austria	1	Mongolia
387	Pakistan	31	Tunisia	1	Niger
373	China	31	Tanzania	1	Palau
344	Uganda	25	Libya	1	Portugal
269	Mauritania	24	Malawi	1	Taiwan

Table 3: Gravity Equation with OLS estimates for the full sample including all countries and the restricted sample with Latin American countries only

Dependent variable	log(intermediate imports)					
	All countries			Latin American countries		
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of immigrants past 5 years)	0.0490*** [0.00495]	0.0399*** [0.00560]	0.0123* [0.00666]	0.0784*** [0.0189]	0.0619* [0.0354]	0.0723** [0.0313]
log(Distance)	-0.718*** [0.0645]	-0.696*** [0.0653]	-0.670*** [0.0654]	-1.086*** [0.195]	-0.872*** [0.219]	-1.074*** [0.201]
Border	0.369* [0.209]	0.349* [0.209]	0.341 [0.209]	0.28 [0.223]	0.199 [0.227]	0.284 [0.223]
log(No. of immigrants 10 years ago)		0.0208*** [0.00597]		0.0461* [0.0279]		
log(No. of immigrants in 1960)			0.0577*** [0.00700]			0.00969 [0.0395]
State x year fixed effects	yes	yes	yes	yes	yes	yes
Country of origin x year fixed effects	yes	yes	yes	yes	yes	yes
Observations	15633	15633	15633	2659	2659	2659
R-squared	0.923	0.942	0.932	0.933	0.956	0.945

Note: The dependent variables are total intermediate imports from country j to state i in period t . Distances are in kilometers from the capital city in country i to the capital city in state j . All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 4: Gravity Equation with first and second stage IV estimates for all countries

Dependent variable	log(No. of immigrants past 5 years)	log(intermediate imports)				
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of refugees in past 5 years)	0.434*** [0.0220]	0.270*** [0.0250]	0.425*** [0.0220]			
log(No. of immigrants past 5 years)				0.160** [0.0634]	0.155*** [0.0489]	0.133** [0.0622]
log(Distance)	-1.62*** [0.199]	-1.01*** [0.191]	-1.56*** [0.199]	-0.462 [0.285]	-0.24 [0.307]	-0.251 [0.285]
log(No. of immigrants 10 years ago)		0.493*** [0.0287]			0.0989 [0.0924]	
log(No. of immigrants in 1960)			0.014*** [0.003]			0.0115** [0.00493]
State x year fixed effects	yes	yes	yes	yes	yes	yes
Country of origin x year fixed effects	yes	yes	yes	yes	yes	yes
Observations	1674	1674	1674	1674	1674	1674
R-squared	0.744	0.825	0.747	0.962	0.968	0.962
Cragg-Donald F	28.72	42.91	31.6	28.72	42.91	31.6
Kleibergen-Paap p-val	0.0195	0.0109	0.0128			

Note: The dependent variables are total intermediate imports from country j to state i in period t . The number of immigrants is instrumented by the number of refugees who arrived in the past 5 years in state i from country j . Distances are in kilometers from the capital city in country i to the capital city in state j . All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 5: Gravity Equation with first and second stage IV estimates for Latin American countries

Dependent variable	log(No. of immigrants past 5 years)	log(intermediate imports)				
	(1)	(2)	(3)	(4)	(5)	(6)
log(Distance) x dPost	0.432*** [0.035]	0.163*** [0.0335]	0.150*** [0.0355]			
log(No. of immigrants past 5 years)				0.156** [0.079]	0.131*** [0.060]	0.132*** [0.064]
log(Distance)	-1.846*** [0.200]	-0.704*** [0.196]	-1.553*** [0.200]	-0.647* [0.340]	-0.578* [0.340]	-0.521 [0.320]
Border	2.038*** [0.241]	0.758*** [0.212]	1.908*** [0.238]	0.821** [0.377]	0.465 [0.353]	0.748** [0.365]
log(No. of immigrants 10 years ago)		0.559*** [0.021]			0.021* [0.012]	
log(No. of immigrants in 1960)			0.022*** [0.002]			0.007* [0.004]
State x year fixed effects	yes	yes	yes	yes	yes	yes
Country of origin x year fixed effects	yes	yes	yes	yes	yes	yes
Observations	1946	1946	1946	1946	1946	1946
R-squared	0.809	0.88	0.814	0.948	0.953	0.948
Cragg-Donald F	15.24	17.9	15.1			
Kleibergen-Paap p-val	0.0352	0.0215	0.0318			

Note: The dependent variables are total intermediate imports from country j to state i in period t . The number of immigrants is instrumented by distance interacted with a time dummy for the period after Hurricane Mitch. Distances are in kilometers from the capital city in country i to the capital city in state j . All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 6: Gravity Equation with OLS estimates for the 2010 sample including all countries and the restricted sample with Latin American countries only

Dependent variable	log(intermediate imports)					
	Latin American countries			All countries		
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of immigrants past 5 years)	0.0638* [0.0349]	0.0183 [0.0378]	0.0623* [0.0349]	0.0352** [0.0152]	0.0203** [0.009]	0.0366** [0.0152]
log(Distance)	-2.25*** [0.227]	-2.08*** [0.233]	-2.19*** [0.230]	-1.15*** [0.0769]	-1.103*** [0.0776]	-1.05*** [0.0817]
Border	0.311 [1.197]	0.212 [1.190]	0.0464 [1.211]	1.85*** [0.687]	1.82*** [0.696]	1.81*** [0.686]
log(No. of immigrants 10 years ago)		0.121*** [0.038]			0.0403** [0.0165]	
log(No. of immigrants in 1960)			0.048 [0.038]			0.054 [0.039]
State fixed effects	yes	yes	yes	yes	yes	yes
Country of origin fixed effects	yes	yes	yes	yes	yes	yes
Observations	710	710	710	3913	3913	3913
R-squared	0.798	0.799	0.798	0.787	0.789	0.787

Note: The dependent variables are total intermediate imports from country j to state i in period t . Distances are in kilometers from the capital city in country i to the capital city in state j . All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 7: Gravity Equation with first and second stage IV estimates for all countries in the year 2010

Dependent variable	log(No. of immigrants past 5 years)	log(intermediate imports)				
	(1)	(2)	(3)	(4)	(5)	(6)
log(No. of refugees in past 5 years)	0.319*** [0.0372]	0.229*** [0.0391]	0.266*** [0.0500]			
log(No. of immigrants past 5 years)				0.274** [0.114]	0.211** [0.102]	0.225** [0.108]
log(Distance)	-1.988*** [0.398]	-1.191*** [0.372]	-1.873*** [0.227]	-1.017 [0.794]	-1.422** [0.690]	-1.220* [0.676]
log(No. of immigrants 10 years ago)		0.388*** [0.0393]			0.0354 [0.140]	
log(No. of immigrants in 1960)			0.0745*** [0.0186]			0.0789 [0.0817]
State fixed effects	yes	yes	yes	yes	yes	yes
Country of origin fixed effects	yes	yes	yes	yes	yes	yes
Observations	649	649	649	649	649	649
R-squared	0.816	0.868	0.832	0.769	0.777	0.766
Cragg-Donald F	22.14	23.1	24.55			
Kleibergen-Paap p-val	0.051	0.048	0.027			

Note: The dependent variables are total intermediate imports from country j to state i in period t. The number of immigrants is instrumented by the number of refugees who arrived in the past 5 years in state i from country j. Distances are in kilometers from the capital city in country i to the capital city in state j. All regressions include country of origin - as well as state-year fixed effects. Robust standard errors in parentheses (clustered by country state): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 8: Effect of Hurricane Mitch on total exports from hurricane-affected countries.

Dependent variable	log(Total exports)
Dummy_US_2000	0.811** [0.337]
Observations	948
R-squared	0.897

Note: Regression includes exporting country, importing country, time, importer-time and exporter-importer fixed effects. ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 9: Effect of transport costs on industry trade flows in gravity model.

Dependent variable	log(intermediate imports)
log(Distance)	-0.691*** [0.107]
log(Distance)*log(Value-to-weight)	0.154*** [0.022]
Border	1.897*** [0.430]
Trade agreement	0.420*** [0.062]
Common language	0.235*** [0.035]
Common currency	0.093* [0.057]
log(GDP of exporter)	0.990*** [0.011]
log(Population of exporter)	0.024* [0.012]
Industry FE	Yes
Observations	61796
R-squared	0.449

Note: Regression includes exporting country, importing country, time, exporter-time and exporter-importer fixed effects. ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.

Table 10: Difference in Differences estimation: Number of migrants after Hurricans Mitch

Dependent variable	log(number of total immigrants)	log(number of immigrants last 5 years)
	(1)	(2)
Treatment effect	-0.743*** [0.254]	-0.760*** [0.202]
Difference in Difference	0.530* [0.283]	0.753* [0.385]
log(Distance)	-0.991*** [0.0461]	-1.237*** [0.0376]
Border	1.111*** [0.127]	1.351*** [0.117]
Constant	13.26*** [0.427]	15.64*** [0.346]
Observations	15,633	15,633
R-squared	0.723	0.817

Note: All regressions include country of origin, host country and time fixed effects. Robust standard errors in brackets (clustered by country time): ***, **, * indicate the statistically significant difference from zero at the 1%, 5% and 10% levels respectively.