An Empirical Analysis of the Determinants of International Migration to Canada

by

Alieh Ghaedi

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

September 2012

Acknowledgments

I would like to sincerely thank my supervisors, Professor Charles Beach and Professor Beverly Lapham, for their constructive feedback and generous guidance. Even though the duration of my work coincided with Professor Beach's retirement, I am very thankful that he was willing to supervise me with his invaluable suggestions and attention to detail.

Abstract

In this paper, I empirically analyze the economic and non-economic determinants of migration flows into Canada. The empirical specification is based on a modified version of the gravity model of migration which includes economic and political characteristics specific to both the origin country and Canada. My panel dataset consists of data collected on annual migration flows of forty source countries to Canada and data on various factors of both origin and source countries, for the period 1980 to 2010. The analysis examines which set of economic, political and demographic factors impact migration flows. Among the major determinants are population density, GDP per capita and the unemployment rate of the origin country as well as the unemployment rate of Canada. A set of factors reflecting cultural and historical ties between the origin country and Canada are also found to impair or promote migration flows to Canada. Moreover, the results indicate that immigration to Canada varies by region of the source country.

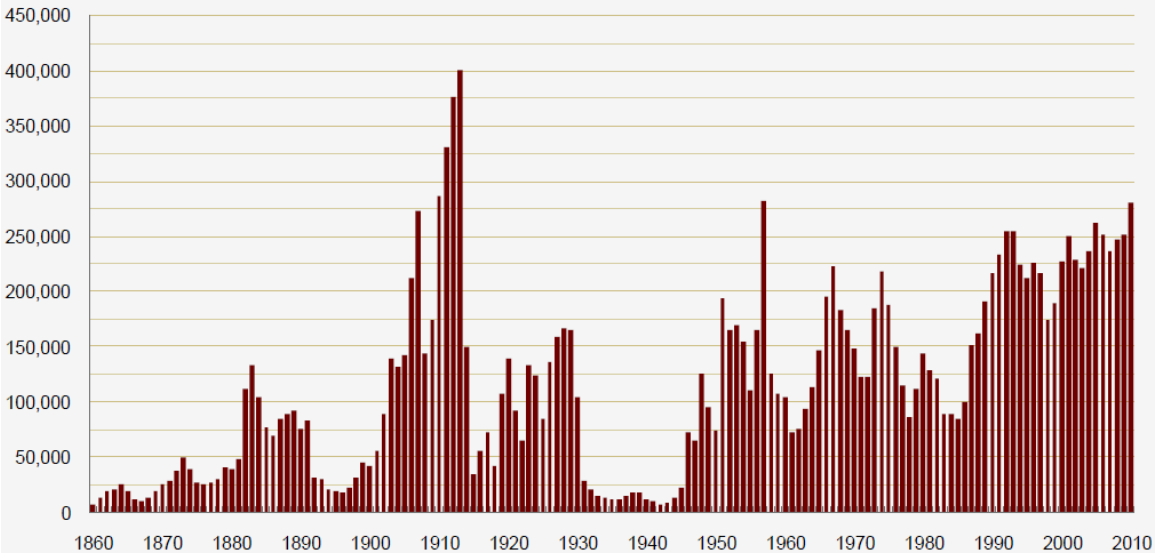
Table of Contents

1. Introduction

In this paper I assess the economic and non-economic determinants of migration to Canada using annual data on immigration flows from forty source countries over a period of thirty-one years. An objective of this paper is to highlight the significance of not only economic factors but also political and social factors of both the origin and destination countries in the migration decision. Moreover, I can assess whether factors that were significant in the previous studies preserve their importance or not.

Immigration is a phenomenon that has experienced different trends in various parts of the world. There are a wide range of different reasons that motivate people to migrate from their place of birth to another country perhaps with a different culture. One of the most popular destination regions for immigrants has been North America, both in the past and in the present. Having immigrated to Canada myself, I also have a personal interest in studying this topic as I can very closely relate to many of the factors that motivate people to migrate. Figure 1 presents a snapshot of the number of immigrants to Canada for the group of permanent residents from 1860 to 2010.

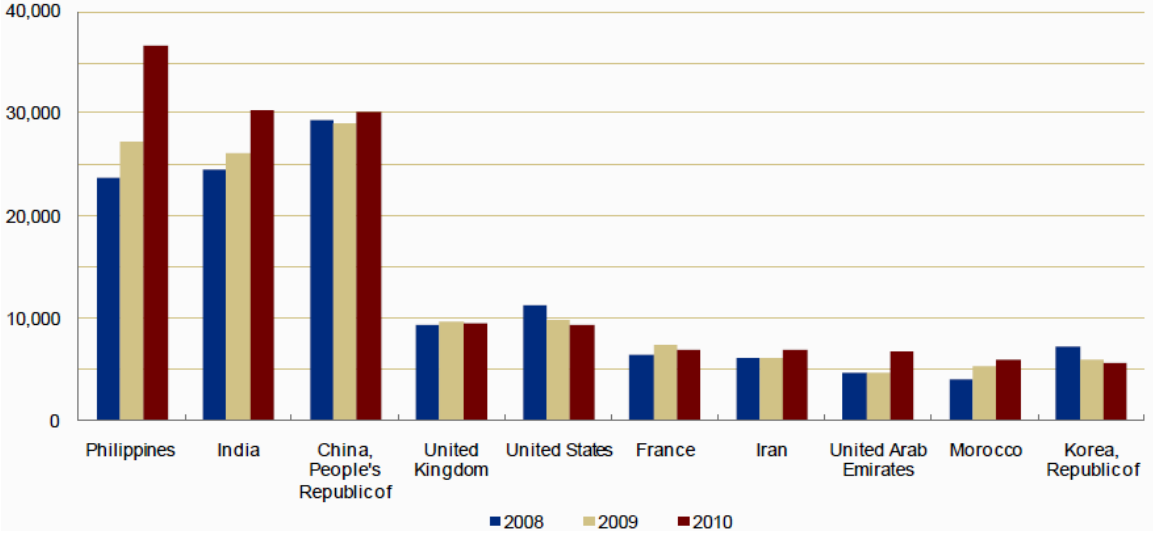
Figure 1: Immigration flows to Canada of permanent residents, 1860 to 2010



Source: Citizenship and Immigration Canada: Facts and Figures 2010, immigration overview of permanent and temporary residents

In the first half of the twentieth century and particularly in the 1950s and 1960s the majority of the immigration to North America was from high-wage countries. For example, in 1960 two-thirds of all immigration flows to North America were from European countries; this share was significantly reduced to one-ninth by 1985 (Karemera, Oguledo and Davis, 2000). Since then the majority of immigration to North America has been from the lower-wage countries of Asia, Latin America and the Caribbean. This change in the composition of immigrants' source countries is in contrast to the immigration acts of both Canada and the U.S., which promotes family reunion by enabling immigrants to sponsor their family members (Karemera et al., 2000). It is expected that family reunion would lead to new immigrants coming from the same origin countries as their family members who have already immigrated to Canada. Figure 2 provides a detailed depiction of flows of permanent residents to Canada by top source countries from 2008 to 2010. As Figure 2 shows, from 2008 to 2010 the top three countries of origin for immigration to Canada were from Asia rather than Europe.

Figure 2: Permanent residents in Canada by top source countries, 2008 to 2010



Source: Citizenship and Immigration Canada: Facts and Figures 2010, immigration overview of permanent and temporary residents

Since the year 2002, Canada's immigration has closely followed the regulations set forth by the Immigration and Refugee Protection Act (IRPA). The latter act replaced the 1976 Immigration Act. The IRPA divided permanent immigrants to three categories: family reunions, people contributing to economic development, and refugees. The family group includes people who are sponsored by a family member or a close relative who is a Canadian citizen. The second group can be divided into subcategories; it includes people who are selected based on the merits of their particular skill and ability such as skilled workers; business people who come through investment; and provincial nominees and caregivers (Facts and Figures 2010 CIC).

Demographic projections for OECD countries indicate that many of them would face the issue of a declining working force and an ageing population, which would exert more pressure on the welfare system as a larger proportion of national income has to be devoted to pension and public health payments. One suggested way to address this issue is to attract young immigrants from around the world (Pederson, Pytlikova, and Smith,

2004). However, people who are against increased immigration flows turn their focus to the consequences of migratory inflows on the destination country's labor market and public finances (Pederson et al., 2004).

In this paper, I empirically investigate the main factors affecting immigration flows. For the empirical specification, I use a modified version of the gravity model and use my constructed panel dataset for estimation. The gravity model of trade which follows Newton's theory of gravity suggests that trade between two countries is proportional to the product of their gross domestic product (GDP) and the distance between the two countries. The main idea is that countries with larger economies are more likely to import and export due to more variety of goods and services they have (Tansey and Touray, 2010). The gravity model of trade has been extended in applications to other contexts such as immigration. Further details of the gravity model of migration used in this paper are given in the theoretical section.

The remainder of this paper is organized as follows. Section 2 provides an overview of some of the main findings in the migration literature. Section 3 presents a simple theoretical framework for international migration. Section 4 presents a detailed description of the dataset and all variables used. In section 5, I provide the econometric specifications and empirical results of my paper. Section 6 presents further empirical refinements. Section 7 provides a summary of the main findings and concluding remarks.

2. Literature Review

Numerous works have been written in the field of international migration and different aspects of it have been studied. Some of the existing works in the literature have investigated the economic and non-economic determinants of migration flows to different countries. Whereas some other works have been done to determine the impact of international migration on the destination country's labor market and wage responses.

In its simplest form, migration can be thought of as a process of movement of labor from a source to a receiving country, which could be as a result of many different factors. Various theories and approaches have been used to model migration and here I briefly describe a subset of those theories.

2.1 The classical theory of migration and the human capital framework

The classical economic theories of migration focus on wage differentials between countries as the main driver of international migration. The classical view is also encountered in empirical studies as the "human capital" framework. According to this framework, an individual's decision to migrate depends on whether the expected future benefit net of migration costs is positive or not (Sjaastad, 1962). However, in reality the set of factors shaping a person's incentive to migrate are far beyond mere earning differences. Following that, some extensions to the classical theory of migration have been made, for example looking at migration as a decision of the whole household, which incorporates every member's benefit or loss as a result of the movement (Pederson et al., 2004).

2.2 The labor market model of migration

The labor market model of migration merely focuses on the economic incentives for migration and considers movement as a means of increasing earning potential. The size and composition of migration flows are determined as a result of the interaction of various supply and demand factors. A thorough insight into the benefits and costs associated with facilitating migration flows requires identification of the forces and restraints that affect movements of people (Mayda, 2009).

A number of works have analyzed the impact of migratory inflows and trade on a country's labor market. Borjas, Freeman and Katz (1991) provide an empirical analysis of the drop in the wages and employment rates of unskilled workers in the United States in the 1980s. Their estimates indicate that about thirty to fifty percent of the decline in the weekly wage of high school dropouts was due to immigration of less-skilled workers and import of goods to the U.S.

2.3 The self-selection theory

A prominent idea in the literature is that the characteristics of the foreign-born workers are determined through the selection effect; this says that immigrants are a selected group with distinctive levels of ability and motivation.

The more recent need for skilled workers in the OECD countries has led to the formulation of immigration policies tailored towards selecting highly skilled people. One of the theories that has been used to analyze the determinants of immigration flows of people with particular skills is the self-selection-theory. A commonly encountered example of the application of this theory is the study done by Borjas (1987) in which he analyzes the skills of immigrants by mainly focusing on the United States as the

destination country. A highly criticized prediction of Borjas (1987) is his negative selection theory. This theory predicts that immigrants from a source country with high income inequality are likely to be negatively selected, implying that their skill levels are generally lower than the average skill levels of the receiving and source country. One possible intuitive reasoning for this theory might be that higher inequality in a source country provides more incentive for low-skilled people to migrate and motivates them to be willing to face the obstacles of migration in the hopes of higher earnings in the destination country. This motivation is likely to be much weaker for groups of people who are financially well-off in the source countries. A second prediction that follows from Borjas's (1987) negative self-selection theory is that countries with higher income inequality will have lower emigration rates. An intuitive explanation of this prediction could be that poor people in a source country have more incentive to migrate but at the same time are financially constrained and may not afford to migrate; this could lead to lower migration rates of countries with higher inequality. Borjas (1987) empirically tests his theory by using data on immigration into the U.S. He uses U.S. immigration rates and immigrant earnings as dependent variables and the relative skill differences of workers from different source countries as the explanatory variables. The empirical findings of Borjas (1987) support his negative self-selection theory regarding the relationship between inequality rates and the skills of immigrants.

Quite often supporters of the positive selection theory have criticized Borjas's negative self-selection theory. Positive self-selection occurs when highly educated and highly skilled workers constitute the immigrants and the main reason being that the source country in effect taxes highly skilled workers and insures low skilled workers to a

higher extent than the destination country. Borjas (1987) argues that positive self-selection will occur only if two conditions are met. First, there has to be a positive correlation between expected earnings of an immigrant at home and in the U.S., and secondly, income inequality has to be greater in the U.S. than in the origin country. Some authors such as Chiswick (1999) use an extended version of the human capital theory of migration to show that higher income inequality in the source country does not lead to negative-self selection and instead only leads to attenuation of the positive self-selection effect. Chiswick (1999) argues that the presence of direct migration costs is the main reason for the latter prediction, implying that the opportunity cost of migration must be lower for highly skilled workers, which reinforces positive self-selection even in the presence of high levels of income inequality of the origin country.

A common weakness in many studies that have tried to test the self-selection theory is that they usually use data from the destination country. This could bias the results as it suffers from the issue of sample-selection due to factors such as migration policy and network effects that are specific to each receiving country (Liebig and Sousa-Poza, 2004). A distinctive way to address this shortcoming is proposed by Liebig and Sousa-Poza (2004). They suggest that in testing the self-selection theory the focus should be on the intention and propensity to migrate rather than the actual migration flows. Consequently, they use the survey collected by the 1995 International Social Survey Programme (ISSP), which covers 23 countries for a sample of 28,000 individuals. The key question asked in the survey is "Would you be willing to move to another country to improve your work or living conditions?" (Liebig and Sousa-Poza, 2004, p.129). The empirical findings of Liebig and Sousa-Poza only partially confirm Borjas's (1987)

model. They support the theory that the relative income inequality between the source and the destination country does have an impact on the skill composition of the immigrants, but they argue that it only attenuates the positive self-selection effect rather than leading to negative self-selection. Moreover, their empirical results indicate that higher earnings inequality of the origin country is associated with higher emigration rates and that positive self-selection is generally expected.

2.4 Investigation of the determinants of international migration

In addition to analyzing the impact of immigration on labor markets and the composition of the immigrants, another area in this literature seeks to investigate the determinants of international migration. Below, I provide an overview of some of the findings.

Mayda (2009) performs a panel data analysis of the determinants of international migration in which she empirically explores the determinants of migration inflows into 14 OCED countries by the country of origin, covering the period 1980 to 1995. Mayda (2009) uses a wide range of economic, geographical, demographic and cultural factors as well as changes in the destination countries' migration policies as explanatory variables. Consistent with the previous findings, Mayda (2009) finds that gross domestic product per capita in the destination country (a pull factor) is statistically significant with a positive impact on immigration rates. However, in contradiction to the theoretical framework, the estimated coefficient for income levels in the country of origin (a push factor) is not negative in most of the specifications. Distance between the origin and destination countries is found to be statistically significant in determining emigration rates with a negative estimated coefficient. Moreover, the share of the origin country's

population who is young is also another important determinant of emigration rates. Mayda (2009) also assesses the impact of immigration policies of a destination country on the immigration inflows to that country; in this context a binding immigration policy is defined as one that sets a limit on the immigration inflows that is less than the number of people that would have entered in the absence of the policy. The results also indicate that a binding immigration policy of the host country matters; the estimated impact of pull factors become more positive and that of push factors more negative when the immigration policy of a host country becomes less restrictive.

Other authors have done studies in which the focus has been on one or two destination countries over time rather than on a multitude of receiving countries. Karemera, Oguledo and Davis (2000) analyze international migration flows to North America by using a combination of cross-section and time-series (panel) dataset. Their dataset covers the migration flows of a set of 70 countries for the period 1976-1986 to Canada and the United States. They use a modified version of the gravity model and include immigration regulations of Canada and U.S. with a set of variables representing the characteristics of the origin and destination countries. The authors find that the origin country's population and the destination country's income level are two important determinants of migration flows to North America. Moreover, political restrictions and lack of civil freedom in origin countries are found to weaken migration to North America.

Kim and Cohen (2010) also use a panel dataset to analyze the determinants of migration flows to and from industrialized countries. Unlike most of the previous work in this field, Kim and Cohen (2010) turn their focus to non-economic factors such as

demographic, geographic and social determinants of migration flows. They analyze migratory inflows to seventeen industrialized countries and outflows from thirteen of the countries between 1950 and 2007. The findings of that paper indicate that the main determinants of migration inflows are distance between the two countries, land area of the destination country, and population and infant mortality rate (IMR) of both the origin and destination countries, whereas social and historical factors are found to be less influential. Similarly, Kim and Cohen (2010) found the determinants of migratory outflows to also be mainly demographic; the most influential ones being the population of both the origin and destination countries, the IMR of the destination country and the distance between the capital cities of the two countries. They also found the young proportion of the population of the destination country to have a negative impact on inflows and that of the origin country to have a positive impact on emigration. Moreover the results indicate that the level of urbanization in both the origin and destination countries has a positive impact on migration.

As the gravity model was initially proposed and used to analyze international trade, Lewer and Berg (2007) provide empirical work in which the gravity model is applied to both international trade and immigration. They use data on legal immigration flows to sixteen OECD countries from almost all source countries in the world for the period 1991 to 2000. For most of the variables included such as the relative gross domestic product, population of the two countries, bilateral distance, common language and colonial links, the estimated coefficients have the same sign in both the gravity model of trade and immigration. Lewer and Berg (2007) also include the existing stock of source country natives already living in the destination country as an explanatory variable. Intuitively,

one expects the psychic costs of adapting to a new country to be reduced when there are compatriots living in the destination country. Their results indicate that the current stock of immigrants from a source country is statistically significant with a positive impact on migration flows. Moreover, Lewer and Berg (2007) also include the ratio of the destination to origin country's government adherence to property rights and rule of law as explanatory variables. The estimated coefficients for both are positive indicating the positive impact that expected institutional improvement has on the propensity to emigrate.

Given the applicability of the gravity model to different contexts, I have chosen to develop an extended version of the gravity model to provide the framework for my empirical analysis.

3. Theoretical Framework

A number of techniques have been used to model international migration, some of which include the gravity model, linear regression models and Markov chain models (Cohen and Kim, 2010). The earnings differential, between the origin and destination country is commonly encountered in the migration literature as a significant driver of the migration decision and is the cornerstone of the labor market model of immigration. From that perspective the migration of workers from one region to another is considered as a human capital investment in which workers would choose the region that is likely to give them the highest possible lifetime earnings net of economic and psychological costs incurred through the moving process (Borjas, 2001).The theory suggests that economic opportunities of the origin and destination countries along with migration costs are

important determinants of migration flows. Consequently, one expects higher wage differentials and lower migrations costs to increase migration flows (Borjas, 2001).

Similar to many economic outcomes, migration flows are determined by the interaction of a variety of supply and demand factors. The supply side represents all economic and non-economic factors that influence a potential migrant's decision to move. The demand side is determined by the immigration policies and the demand for immigrants in the destination country (Mayda, 2009). The immigration policy of a destination country can be considered to be an outcome of a variety of factors such as individual attitudes towards immigrants, government structure and preferences of policymakers (Mayda, 2009). The following theoretical section on wage functions and the probability of immigration follows directly the work of Mayda (2009).

The supply side reveals a person's decision to migrate; the determination of the probability of migration entails a comparison of the current wage that a person is earning in the source country with the alternative wage that he could earn in the destination country. The wage of a person in the source and destination country is a function of that individual's skill level. The wage of individual k in the origin country (i) is:

$$w_{ik} = a_i + b_i s_k + e_{ik} \qquad (1)$$

where w_{ik} is the wage earned by individual k in the origin country i, s denotes the skill level of individual k, a and b represent the intercept and coefficients in the equation and e is the disturbance term. Similarly, the wage of individual k if he migrates to destination country j is:

$$w_{jk} = a_j + b_j s_k + e_{jk} \qquad (2)$$

Since macro data are generally used for migration empirical studies, it would be ideal to express the above wage equations in terms of the mean wage of the population of the origin country at home and abroad:

$$w_{ik} = \theta_i + v_{ik} \quad \text{where } v_{ik} \sim N(0, \sigma_i^2) \tag{3}$$

$$w_{jk} = \theta_j^i + v_{jk} \quad \text{where } v_{jk} \sim N(0, \sigma_j^2) \tag{4}$$

where $\theta_i = a_i + b_i\overline{s_i}$, $\theta_j^i = a_j + b_j\overline{s_i}$ and $\overline{s_i}$ is the mean skill level of the country of origin's population. It should also be noted that $\theta_j = a_j + b_j\overline{s_j}$ is different from θ_j^i. θ_j represents the mean wage of the destination country's population which depends on the mean skill level of its population, but θ_j^i denotes the mean wage of the origin country's population if they migrate to country j (Mayda, 2009).

We will assume that there are only two goods produced in the world and individual preferences are characterized by a Cobb-Douglas utility function:

$$U(x_1, x_2) = Bx_1^{\alpha}x_2^{1-\alpha} \ , \quad 0 < \alpha < 1, \ B > 0 \tag{5}$$

The above utility function leads to the following well-known Cobb-Douglas Marshallian demand functions:

$$x_1 = \frac{\alpha M}{p_1} \text{ and } x_2 = \frac{(1-\alpha) M}{p_2} \tag{6}$$

where M denotes the income of each individual and p denotes the price of good j, j =1,2. We can furthermore assume that each country is small relative to the rest of the world and thus the goods' prices are determined at the world level and are equal across all countries. Despite the existence of free trade among countries, the rate of return to labor is not the same in all countries due to factors such as productivity differences. By substituting the

Marshallian demand functions into the utility function we obtain the following indirect utility function:

$$V(p_1, p_2, M) = B\, M \left(\frac{\alpha}{p_1}\right)^{\alpha} \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} \quad (7)$$

A person in the origin country i will decide to migrate to a destination country j if the utility level of migration is higher than the utility obtained if the person were to stay in the origin country. In other words, a person will migrate if the expected level of income in the destination country net of migration costs is greater than the income level in the home country. An index, I, can be used to denote a measure of the net-benefit of migrating relative to staying in the home country for a risk-neutral individual k:

$$I_k = \gamma \left(w_{jk} - w_{ik} - C_k\right) + (1-\gamma)(-w_{ik} - C_k) \quad (8)$$

which can be reduced to:

$$I_k = \gamma w_{jk} - w_{ik} - C_k \quad (9)$$

in which γ represents the probability that a migrant from country i would be able to stay in country j and work there. Thus a person who migrates is giving up the current wage earned in the home country and also incurs the costs of migration captured by C_k .

An immigrant who has already incurred the costs of immigration might not be able to stay in the destination country due to various reasons such as an increase in the immigration policy restrictions of the host country. Thus the probability that a person from the origin country's population would migrate is given by :

$$P = \Pr(I_k > 0) = \Pr\left[\, \gamma(\theta_j^i + v_{jk}) - (\theta_i + v_{ik}) - C_k > 0\right] \quad (10)$$

The above probability is in fact the supply immigration rate which is the size of the immigration flow from an origin country i to a destination country j divided by the

population of the origin country i. Countries could set immigration quotas in which they impose quantitative restrictions on the immigration flows to their country. An immigration quota implies a country's demand for immigrants and, if binding, it will result in a discrepancy between the actual immigration flows (observed in the data) from an origin country and the supply immigration rate.

The labor market model of immigration has some straightforward implications for empirical purposes. First, migration flows or more specifically the migration rate, would be the dependent variable; and secondly, the distance between the two countries is a proxy for migration costs and earnings of workers in each country should be included as explanatory variables. For the empirical purposes of my paper, I am combining the latter framework with those of the gravity model of trade applied to the concept of migration.

The gravity model is an equation that is derived from a set of supply and demand interactions (Karemera et al., 2000). The simplest form of the gravity model of migration considers the populations of origin and destination countries, national income of each country as well as the bilateral distance between the two as the main determinants of migration flows (Cohen and Kim, 2010). The set of potential supply and demand factors likely to affect the migration flow between origin country i and destination country j are given by the following expressions:

$$S_i = a_0 \, Y_i^{a1} N_i^{a2} \tag{11}$$

$$D_j = b_0 \, Y_j^{b1} N_j^{b2} \tag{12}$$

where S_i represents push factors of the origin country. Here Y_i denotes the origin country's income and N_i denotes its population. D_j represents the set of pull factors of the

destination country, in which Y and N again indicate its income and population respectively. Each exponent in the above equations, represents the migration elasticity for each corresponding variable. Combination of the supply and demand equations yields the basic gravity equation of migration (Karemera et al., 2000):

$$F_{ij} = k_0 \frac{S_i^{k1} D_j^{k2}}{R_{ij}^{k3}} \tag{13}$$

in which R_{ij} represents various factors that either facilitate or restrain immigration, such as transportation costs. In the above equation k_0 represents a_0 and b_0 combined, $k1$ is a combined migration elasticity of the different components in the supply expression and $k2$ is a combined migration elasticity for the different components of the demand expression. By taking logs of both sides of the above equation and substituting each term by its log equivalent (lower case of variables for logs), we obtain the basic migration model as the following:

$$m_{ij} = \beta_0 + \beta_1 n_i + \beta_2 n_j + \beta_3 y_i + \beta_4 y_j + \beta_5 d_{ij} + z(.) \tag{14}$$

where $m_{ij}$is the log of migration flow between the origin country i and destination country j, n and y represent log of population and log of income for each country as already defined, d_{ij} and $z(.)$together represent R_{ij} as mentioned above. Here, d_{ij} denotes the log of distance between the two countries and is used as a proxy to capture the effect of transportation costs, and $z(.)$represents various characteristics of both the origin and destination countries that would ease or restrain immigration. In the very basic form, $z(.)$can be considered as representing the error term, but for empirical purposes its components should be identified and included in the model (Karemera et al., 2000).

The general prediction of this simple gravity model of migration is that the higher is the population of a country, the higher would be the emigration rates from that country and immigration flows are higher to countries with higher populations. A high urban population in a destination country is likely to be linked with more employment opportunities for new immigrants. Bilateral distance is considered to represent a large portion of migration costs, and thus both theory and the empirical works done in the literature imply that the estimated coefficient of distance would be negative. In the empirical works, distance between the origin and destination countries is thought of as a proxy for both financial and psychological costs of migration, with very distant destination countries leading to higher economic and emotional burden on the migrants (Kim and Cohen, 2010). It should be noted that the consistency of the latter argument with empirical findings is contingent on the time period being covered by the study, as the significance of transportation costs in the migration decision is highly impacted by the degree of technological advances. This point will be further emphasized in the empirical section of the paper.

The existing works in the literature have expanded the basic gravity model of migration by adding various types of relevant explanatory variables. In addition to gross domestic product as a measure of income, it has been proposed that other economic variables such as the unemployment rate and inflation rate of the origin and destination countries also be included. Furthermore it has been suggested that non-economic factors such as the political status and civil rights should also be accounted for (Karemera et al., 2000).

Another set of factors used in the literature is geographical and historical factors. For example, one factor that is thought to affect migration costs is the existence of a common border between the country of origin and destination as, for example, one expects a land-locked country to receive fewer immigrants than a country bordered with an ocean. Moreover, having a common language, culture and a shared history such as a colonial relationship in the past are all expected to have an influence on the migration flows between two countries. For example, most of the empirical works in this area have shown that past colonial relationships tend to ease the process of migration (Kim and Cohen, 2010). Furthermore, the variable population can be explored in more detail. The age decomposition of the population of each country may also affect immigration flows since younger people are more likely to migrate. One such decomposition would be to find a measure of the population ageing and working-age population of both the origin and destination countries. Other demographic factors that may have a significant impact on immigration flows would be measures of the overall health and quality of life in a country. Two such measures would be the infant mortality rate and life expectancy at birth which might be the only available measures of health for many developing countries (Kim and Cohen, 2010).

4. Data and Variable Description

For the empirical purposes of my paper I have collected a panel dataset for analyzing immigration flows to Canada from forty source countries. The period covered in this paper is annually from 1980 to 2010 for most variables used. As the theme of my paper is to investigate both economic and non-economic determinants of migration flows to

Canada, data collection for the relevant explanatory variables mandated the use of multiple sources to account for different groups of relevant explanatory variables.

4.1 Dependent variable

The dependent variable implied by the theoretical background in this paper is the natural logarithm of immigration flows to Canada from various source countries. This variable was obtained from Citizenship and Immigration Canada (CIC) as it was found to be the most comprehensive and reliable source of this data. Immigrants can be broadly divided into permanent residents, foreign workers and foreign students, the latter two groups being part of the temporary residents. As my theme is inclined towards a permanent and long-term analysis, I use immigration flows for permanent residents. CIC has different ways of defining a source country and thus has data available for permanent residents by country of last permanent residence, country of citizenship, and country of birth. As the migration flow model already described is aiming to capture push and pull factors as well as attributes of the origin and destination countries, I have considered it more appropriate to use data for Canada's permanent residents by country of last permanent residence. For example, a potential migrant does not necessarily live in his country of birth or citizenship before deciding to migrate and it would be more realistic to assume that the situation of the country of his residence would impact the migration decision. Data on immigration flows to Canada was readily available for as many as 200 countries, but due to lack of availability of data for the explanatory variables, I have focused on 40 source countries which is also a more appropriate number for the scope of this paper. For a complete list of the source countries used, please refer to the appendix.

4.2 Explanatory variables

I have divided the explanatory variables into the following categories: gravitational and demographic variables, economic variables, domestic politics and indicators of human rights, and historical and cultural factors.

4.2.1 Gravitational and demographic factors

A common practice in empirical studies is to use the distance between the two countries as a proxy for transportation costs, as the latter are not readily available (Karemera et al., 2000). I have obtained data on bilateral distance between each source country and Canada from the gravity dataset prepared by the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII). There are two measures of distance available in the dataset, a simple distance and a weighted distance. The exact definition of all the explanatory variables is shown in Table 1 below.

An important demographic variable which is implied by both the labor market model and the gravity model of migration is total population of the source and destination countries. As suggested by Kim and Cohen (2010), a measure representing the age structure of the population is also constructed. This measure is called the potential support ratio (PSR) and is constructed as:

PSR= (number of people aged 15-64) / (number of people aged 65 or over)

This measure indicates the number of people between the ages of 15 to 64 for every person who is 65 or older. Consequently, a low PSR is an indication of population ageing and might signal a shortage in the working age population of a country. Thus we would expect a source country with a high PSR to have a higher immigration outflow than one

with a low PSR and a destination country with a low PSR to attract more immigrants than one with a high PSR (Kim and Cohen, 2010).

As already mentioned, two other demographic factors that have been used are the infant mortality rate (IMR) and life expectancy at birth. These two factors can potentially be measures of the overall health and quality of life in a country. It is expected that the higher is the IMR and lower the life expectancy, the higher would be the migrations from an origin country. A destination country with a high IMR and a low life expectancy is expected to attract fewer immigrants. All different measures of population, IMR, life expectancy at birth are obtained from the World Bank's World Development Indicators (WDI) for the source countries and Canada for the period of 1980 to 2010.

4.2.2 Economic factors

A set of variables are used to capture the economic situation of the origin and destination countries, which would represent push and pull factors of migration. In the migration literature, national income is generally used as a proxy for the earnings of workers in a country. My dataset includes two measures: gross domestic product (GDP) per capita and a measure of total GDP.

Two other economic variables are also included as explanatory variables. One is the unemployment rate and the other is the inflation rate. A high unemployment and inflation rate in the source country are expected to be associated with higher migration rates from that country, whereas a low unemployment and inflation rate in a destination country are expected to attract more immigrants. It should be noted that intuitively enhanced economic opportunities can play a dual role in the migration decision of the origin country. On the one hand, improved domestic economic conditions in the source country

would imply higher national income and perhaps weaker economic incentives for migration, but on the other hand it could also imply enhanced affordability for immigration (Karemera et al., 2000).

Another factor that might play a role in the determination of immigration flows is the overall level of education or literacy rates in a country. Unfortunately, due to lack of availability of data, such a measure could not be used in this study and instead I have obtained data on primary school enrollment.

All the economic variables are obtained from the World Bank's World Development Indicators (WDI) for the source countries as well for Canada, covering the period 1980 to 2010. The data on the unemployment rate, the inflation rate and primary school enrollment suffered from missing values for a few countries. The average of preceding and proceeding entries were used to fill the missing values.

4.2.3 Domestic politics and human rights indicators

A properly specified model of migration mandates the use of non-economic factors as well. To this end I have collected data on variables that represent a country's political orientation and freedom status of its residents. The annual publications of the Freedom House provide indicators for political rights (pr) and civil liberties (cl) covering the period of January 1st to December 31st of each year. Each indicator is measured on a scale of 1 to 7 in which 1 denotes the most free country and 7 the least free. It is expected that a country with a poor rating for political rights would create more incentive for its residents to emigrate. The effect of civil liberties might be ambiguous at first as one expects a low rating of civil rights to create incentives for people to emigrate from that country, but at

the same time poor civil rights could also imply more restrictions for moving out of the country.

Data on these two indicators are collected for all source countries (except Hong Kong) as well as Canada for the period 1980 to 2010. However, due to collinearity, Canada's data on these two variables are omitted in the empirical analysis, as the indices do not vary over the period covered.

I have also used another source for obtaining data on human rights indicators of the countries in my dataset. The Cingranelli-Richards (CIRI) Human Rights Dataset contains standard-based quantitative information on government respect for 15 internationally recognized human rights. I have used some of the indicators of that dataset as explanatory variables for the migration analysis. One of the variables used is physical integrity. The physical integrity rights index is an additive index which is constructed from indicators on torture, extrajudicial killing, political imprisonment and disappearance indicators; it ranges from 0 to 8 with the latter representing full government respect for the rights. The empowerment rights index is also an additive index which is created from indicators on the freedom of foreign movement, domestic movement, freedom of speech, freedom of assembly and association, worker's rights, electoral self-determination and freedom of religion; it ranges from 0 to 14 with the latter denoting full government respect for the rights. Data on the human rights indicators are collected for all source countries (no data on Hong Kong) and Canada for the period of 1981 to 2010. Data for 1980 has been interpolated based on the following year's data.

4.2.4 Historical, cultural and location factors

A set of dummy variables are also included to take into consideration the impact of cultural and historical commonalities on immigration flows. A location variable commonly used is the population density (total population/area) of the source and destination countries. The aim is to estimate the effect of changes in population concentration per square kilometer on the tendency to migrate (Karemera et al., 2000). Data on this variable was also obtained from the World Bank's World Development Indicators for all source countries and Canada for the period 1980 to 2010.

A language proficiency dummy variable is used which equals 1 if the source country has either English or French as its official language and 0 otherwise. It is expected that sharing a common language and cultural similarities would encourage immigration.

A dummy variable is used to indicate whether the source country and Canada have ever had any colonial relationship for a relatively long period of time and with a significant participation in the governance of the colonized country or not, equaling 1 if they did and 0 otherwise. If a colonial relationship was of the type that has led to stronger political ties and similar institutions, then we would expect it to promote migration. A dummy variable is also used to denote whether a country is landlocked or not, 1 if the country is landlocked and 0 otherwise. Dummy variables on official language, colonial link, being landlocked or not are all obtained from the CEPII gravity dataset.

I have categorized the source countries in the dataset into nine regional groups: North America, Africa, Europe, Latin America and the Caribbean, South Asia, East Asia, West Asia, South East Asia, and Pacific which are represented by the eight regional dummy

variables in the empirical analysis. These dummy variables are included to capture attributes specific to each region that could impact migration.

It is also common practice to use a dummy variable indicating whether the source and the destination country share a common border or not. In this case, the only country that is contiguous with Canada is the United States and thus such a dummy variable would basically capture all characteristics specific to the U.S. that impact immigration to Canada. Since the regional dummy variable for North America only includes the U.S., I have not included a separate variable for common border.

Table 1 lists the variables used along with their definitions and sources. Tables 2 and 3 provide some basic summary statistics of the main variables in my panel dataset. Further detailed summary statistics for immigration flows and GDP are also provided in the appendix in order to give a better snapshot of the dataset used.

Table 1: Definition of Variables

Variable Description	Source
Immigration flows (migrants): Annual number of immigrants (permanent residents) entering Canada by country of last permanent residence.	(1)
Simple distance to Canada(dist): Distance between each source country and Canada calculated following the great circle formula (by Eric Weisstein), which uses latitudes and longitudes of the most important cities/agglomerations (in terms of population), also incorporate internal distances based on areas.	(2)
Weighted distance to Canada: Distance between each source country and Canada using city-level data to assess the geographic distribution of population (in 2004) inside each nation. Distance is calculated based on bilateral distances between the biggest cities of those two countries, inter-city distances being weighted by the share of the city in the country's population.	(2)
Population, total (pop): Total population counts all residents regardless of legal status or citizenship, except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin.	(3)
Population ages 15-64, total: Total population between the ages 15 to 64 is the number of people who could potentially be economically active.	(4)
Population ages 65 and above, total: Total population 65 years of age or older.	(4)
Population density(popden): Population density is midyear population divided by land area in square km.	(3)
Inflation, consumer prices (annual %) (infl): Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly.	(3)
Unemployment, total (% of total labor force) (unemp): Unemployment refers to the share of the labor force that is without work but available for and seeking employment.	(3)
School enrollment, primary (% gross) (pschool): Gross enrollment ratio is the ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the level of education shown.	(3)
GDP (constant 2000 US$): GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. Dollar figures for GDP are converted from domestic currencies using 2000 official exchange rates.	(3)
GDP per capita (constant 2000 US$) : GDP per capita is gross domestic product divided by midyear population.	(3)
GDP per capita, PPP (constant 2005 international $)(pcgdp): Purchasing Power Parity (PPP) GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the U.S.	(3)
Life expectancy at birth (life): It indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.	(4)
Mortality rate, infant (per 1,000 live births) (imr): Infant mortality rate is the number of infants dying before reaching one year of age, per 1,000 live births in a given year.	(4)
Physical Integrity Rights Index (phyint): This is an additive index constructed from the Torture, Extrajudicial Killing, Political Imprisonment, and Disappearance indicators. It ranges from 0 (no government respect for these four rights) to 8 (full government respect for these four rights).	(5)
Empowerment Rights Index (empinx): This is an additive index constructed from the Foreign Movement, Domestic Movement, Freedom of Speech, Freedom of Assembly & Association, Workers' Rights, Electoral Self-Determination, and Freedom of Religion indicators. It ranges from 0 (no government respect for these seven rights) to 14 (full government respect for these seven rights).	(5)

Source: (1) Citizenship and Immigration Canada's database; (2) Centre d'Etudes Prospectives et d'Informations Internationales (CEPII); (3) World Bank's World Development Indicators database (WDI) database; (4) World Bank's Health Nutrition and Population Statistics database (5) Cingranelli-Richards (CIRI) Human Rights Dataset.

Table 2: Summary Statistics for Countries of Origin

Variable	Obs	Mean	Std. Dev.	Min	Max
Annual immigration flows	1240	2436.762	5257.537	25	44225
Total GDP (constant 2000 USD)	1240	5.55E+11	1.49E+12	4.51E+09	1.17E+13
PPP measure of GDP/capita	1240	16723.97	11439.12	523.9503	51966.4
Total population	1240	6.62E+07	1.88E+08	1078200	1.34E+09
Population density	1240	400.3556	1216.398	1.9124	7252.857
Unemployment rate	1237	7.703021	4.339136	0.9	28.1
Primary school enrollment	1196	104.0941	10.47745	11.7438	154.5384
Inflation, annual (%)	1203	38.13427	381.8255	-9.6286	11749.64
Civil liberty index	1209	2.470637	1.508038	1	7
Political rights index	1209	2.177006	1.632179	1	7
Physical integrity index	1209	5.486352	2.302465	0	8
Empowerment rights index	1209	10.78577	3.492366	0	14
Potential support ratio*	1240	8.601613	4.298702	2.8190	18.7074
Infant mortality rate	1209	20.57452	22.11745	2.1	113.6
Life expectancy at birth	1240	73.38314	5.596043	51.9636	82.9327
Distance to Canada (kms)	1240	7868.293	3471.735	548.3946	15586.66

*Potential support ratio=(population, ages 15-64) / (population, ages 65 or over).

Table 3: Summary Statistics for Canada

Variable	Obs	Mean	Std. Dev.	Min	Max
Total GDP	1240	6.28E+11	1.52E+11	4.12E+11	8.73E+11
PPP measure of GDP/capita	1240	29165.71	4374.058	22635.37	36124.31
Total population	1240	2.93E+07	2816808	2.46E+07	3.41E+07
Population density	1240	3.2171	0.3098	2.704456	3.7528
Unemployment rate	1240	8.6226	1.6950	6	12
Primary school enrollment	1240	101.0576	2.5259	96.9419	105.4663
Inflation	1240	3.5495	2.8806	0.1853	12.4624
Civil liberty index	1240	1	0	1	1
Political rights index	1240	1	0	1	1
Physical rights index	1240	7.3226	0.7359	6	8
Empowerment rights index	1240	13.6774	0.6422	12	14
Potential support ratio*	1240	5.8679	0.6568	4.92171	7.2149
Infant mortality rate	1240	6.5	1.5168	5.2	10.4
Life expectancy	1240	78.2078	1.7563	75.0781	80.9649

*Potential support ratio=(population, ages 15-64) / (population, ages 65 or over).

5. Empirical Model and Results

Following the migration model developed in Section 3 in combination with the collected variables as described in Section 4, I now build an econometric model that parameterizes migration flows over time and across countries.

5.1 Econometric Specification

As outlined in Section 4, there are a number of economic and non-economic factors pertaining to both the source and destination countries that are likely to affect migration flows. Before delving into the model specification, a point should be made regarding the GDP per capita of the origin country. Adverse economic conditions in a source country as reflected by a low GDP per capita indicates a strong incentive for immigration, but at the same time implies less affordability to emigrate. Similarly, in a source country with good economic conditions, there is a higher affordability to emigrate but less incentives due to more satisfactory economic conditions (Karemera et al., 2000). As suggested in the literature, the latter argument implies an inverse U-shaped (concave) relationship between immigration flows and GDP per capita of the origin country. Thus in the specification, I have entered both the log of GDP per capita and the square of the log of GDP per capita for the origin country. The full econometric specification that captures the economic and non-economic factors discussed above that might affect migration behavior is given by the following equation:

$$\log(migrants_{ijt}) = \beta_0 + \beta_1 \log(pcgdp_{it}) + \beta_2[\log(pcgdp_{it})]^2 + \beta_3 \log(pcgdp_{jt}) + \beta_4 \log(dist_{ij}) + \beta_5 \log(pop_{it}) + \beta_6 \log(pop_{jt}) + \beta_7 psr_{it} + \beta_8 psr_{jt} + \beta_9 imr_{it} + \beta_{10} imr_{jt} + \beta_{11} life_{it} + \beta_{12} life_j + \beta_{13} \log(popden_{it}) + \beta_{14} \log(popden_{jt}) + \beta_{15} unemp_{it} + \beta_{16} unemp_{jt} + \beta_{17} infl_{it} + \beta_{18} infl_{jt} + \beta_{19} pschool_{it} + \beta_{20} pschool_{jt} + \beta_{21} cl_{it} + \beta_{22} cl_{jt} + \beta_{23} pr_{it} + \beta_{24} pr_{jt} + \beta_{25} phyint_{it} + \beta_{26} phyint_{jt} + \beta_{27} empinx_{it} + \beta_{28} empinx_{jt} + \beta_{29} colony_{ij} + \beta_{30} comlang_{ij} + \beta_{31} landlocked_i + \beta_{32} Afr + \beta_{33} Eur + \beta_{34} LAC + \beta_{35} Pac + \beta_{36} SEA + \beta_{37} EA + \beta_{38} SA + \beta_{39} WA + \varepsilon_{ijt} \quad (15)$$

where the subscript i indicates the origin country and j indicates the destination country and because I am focusing on one destination country in my dataset j refers to Canada; the subscript t denotes time. The dependent variable is the natural logarithm of the number of people who migrated from the origin country i to Canada in year t. As a consequence, the estimated coefficients of the logged explanatory variables represent elasticities and the estimated coefficients on dummy variables capture proportional effects. The notation $pcgdp$ indicates the GDP per capita, $dist$ is the simple distance between each source country and Canada, the exact description of the calculation is shown in Table 1. Pop is the total population, psr is potential support ratio as explained in Section 4, imr is infant mortaliy rate, life is the life expectancy, $popden$ is population density, $unemp$ is the unemployment rate, $infl$ is the inflation rate, $pschool$ is primary school enrollment, cl is the civil liberty index, pr is the political rights index, $phyint$ is the physical integrity rights index which is an additive index explained in Table 1, and $empinx$ is the empowerment rights index explained in Table 1 too. $Colony$ is the dummy variable which equals 1 if the origin country and Canada were in any type of colonial relationship for a relatively long period of time at some point in the past and 0 otherwise;

$comlang$ is the dummy variable for whether the origin country and Canada have a common official language or not, and $landlocked$ is the dummy variable which equals 1 if the origin country is landlocked and 0 otherwise. I have divided the origin countries in my dataset into nine different regions and thus I have included eight regional dummy variables in the empirical specification. Afr is the dummy variable for Africa, Eur is for Europe, LAC denotes Latin America and the Caribbean, Pac denotes Pacific, SEA is South-East Asia, EA is East Asia, SA is South Asia and WA is West Asia.

5.2 OLS estimation results

Table 4 shows the Ordinary Least Squares (OLS) estimation results of the full specification along with some restricted specifications. Before exploring the estimation results, it should be mentioned that the political rights and civil liberties indices for Canada have been omitted from the specifications due to collinearity as the indices do not change over time. Moreover, in every specification that has Canada's total population, the population density of Canada has been omitted due to collinearity. Also, it should be noted that only the first model has inflation as an explanatory variable; it has been omitted from the rest of models for a couple of reasons. Firstly, World Bank's data on inflation has quite a few missing values for some years for some countries and thus some values were filled based on the preceding and proceeding values, thus the quality and reliability of data on inflation is somewhat questionable. Moreover, for some countries there are almost no entries for inflation and thus those countries would be omitted from the estimation and this leads to some unnecessary loss of information. In the first model, inflation of origin country is not significant and that of Canada is significant but its estimated coefficient is positive which is not consistent with the general expectation.

Furthermore, as will be discussed later, inclusion of inflation also impacts the significance and estimated coefficients of some other variables.

Table 4: Ordinary Least Squares (OLS) Estimation Results

Dependent Variable:	Log (annual immigration flow)				
Specification	(1)	(2)	(3)	(4)	(5)
log GDP/capita(origin)	4.509***	3.965***	3.724***	5.388***	-4.724***
	(0.790)	(0.627)	(0.627)	(0.671)	(0.857)
Square of log GDP/capita (origin)	-0.253***	-0.230***	-0.215***	-0.306***	0.248***
	(0.0442)	(0.0361)	(0.0361)	(0.0387)	(0.0489)
log GDP/capita (Canada)	-0.356	-2.798*	-2.776*	-2.762	-4.774*
	(1.843)	(1.619)	(1.628)	(1.755)	(2.641)
log (distance to Canada)	-0.325**	0.0233	0.168	-0.738***	-0.352***
	(0.137)	(0.128)	(0.122)	(0.125)	(0.0725)
log population(origin)	0.110***	0.128***		0.157***	
	(0.0347)	(0.0345)		(0.0373)	
log population(Canada)	1.142	7.995		9.695	
	(8.375)	(7.823)		(8.477)	
log potential support ratio(origin)	-0.382***	-0.192*	-0.134	-0.396***	0.172
	(0.118)	(0.115)	(0.114)	(0.123)	(0.137)
log potential support ratio(Canada)	-9.417*	-4.581	-4.380	-3.504	-12.12**
	(5.308)	(4.921)	(4.948)	(5.332)	(6.164)
log infant mortality rate (origin)	0.698***	0.509***	0.698***	0.698***	0.729***
	(0.115)	(0.112)	(0.100)	(0.120)	(0.145)
log infant mortality rate (Canada)	1.721**	1.774**	1.787**	1.609*	3.041**
	(0.827)	(0.823)	(0.828)	(0.892)	(1.309)
log life expectancy (origin)	1.580*	1.824**	2.851***	3.023***	6.909***
	(0.879)	(0.876)	(0.835)	(0.945)	(1.332)
log life expectancy (Canada)	-10.74	-10.70	-10.80	-12.93	27.79
	(19.70)	(19.57)	(19.68)	(21.20)	(26.15)
log population density (origin)	0.388***	0.410***	0.417***		0.448***
	(0.0292)	(0.0292)	(0.0293)		(0.0270)
Unemployment rate (origin)	0.0157**	0.0154**	0.0185***	0.0115*	0.0244**
	(0.00646)	(0.00637)	(0.00635)	(0.00690)	(0.00966)
Unemployment rate (Canada)	-0.0444	-0.136***	-0.135***	-0.136***	-0.132**
	(0.0515)	(0.0394)	(0.0396)	(0.0427)	(0.0559)
Inflation (origin)	2.34e-06				
	(5.38e-05)				
Inflation (Canada)	0.0539**				
	(0.0215)				
Primary school enrollment(origin)	0.0156***	0.0108***	0.0138***	0.00923***	
	(0.00294)	(0.00281)	(0.00271)	(0.00305)	

Primary school enrollment(Canada)	0.0645***	0.0794***	0.0797***	0.0770***	
	(0.0248)	(0.0241)	(0.0242)	(0.0261)	
Civil liberty (origin)	-0.123***	-0.144***	-0.141***	-0.151***	
	(0.0428)	(0.0427)	(0.0429)	(0.0463)	
Political rights (origin)	0.110***	0.143***	0.146***	0.0912**	
	(0.0353)	(0.0344)	(0.0346)	(0.0370)	
Physical integrity index (origin)	-0.153***	-0.157***	-0.179***	-0.166***	
	(0.0170)	(0.0168)	(0.0157)	(0.0182)	
Physical integrity index (Canada)	0.0110	0.0120	0.0126	0.0104	
	(0.0477)	(0.0475)	(0.0478)	(0.0515)	
Empowerment rights index (origin)	-0.0226	-0.0176	-0.00918	-0.0222	
	(0.0143)	(0.0141)	(0.0140)	(0.0153)	
Empowerment rights index (Canada)	-0.0798	-0.0822	-0.0837	-0.0788	
	(0.0564)	(0.0561)	(0.0564)	(0.0608)	
Colony dummy	1.282***	1.261***	1.566***	1.106***	
	(0.138)	(0.139)	(0.112)	(0.150)	
Common language dummy	1.003***	1.040***	0.903***	1.412***	
	(0.0838)	(0.0837)	(0.0755)	(0.0860)	
Landlocked dummy	-0.727***	-0.854***	-0.964***	-0.897***	
	(0.101)	(0.0989)	(0.0949)	(0.107)	
Africa dummy	-1.319**	-2.442***	-3.312***	0.670	
	(0.544)	(0.522)	(0.469)	(0.512)	
Europe dummy	-1.564***	-2.386***	-3.192***	0.289	
	(0.417)	(0.401)	(0.339)	(0.383)	
Latin American and Caribbean dummy	-2.250***	-2.907***	-3.933***	-0.610	
	(0.456)	(0.449)	(0.355)	(0.453)	
Pacific dummy	0.150	-0.994**	-1.775***	0.607	
	(0.503)	(0.477)	(0.431)	(0.502)	
South East Asia dummy	-1.363**	-2.726***	-3.618***	0.827	
	(0.592)	(0.561)	(0.509)	(0.542)	
East Asia dummy	-0.173	-1.404***	-2.028***	2.360***	
	(0.511)	(0.480)	(0.452)	(0.432)	
South Asia dummy	-1.484**	-2.798***	-3.533***	0.449	
	(0.603)	(0.574)	(0.542)	(0.569)	
West Asia dummy	-2.019***	-2.969***	-3.910***	0.0374	
	(0.503)	(0.486)	(0.416)	(0.473)	
log population density (Canada)			8.665		-5.461
			(7.866)		(8.581)
Constant	17.84	-84.14	40.31	-111.6	-50.92
	(93.17)	(82.33)	(65.60)	(89.20)	(86.43)
Observations	1,126	1,163	1,163	1,163	1,206
R-squared	0.804	0.795	0.792	0.759	0.295

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The first specification in Table 4 includes all potential explanatory variables; from the second model onward inflation has been omitted. In the second specification only inflation of both Canada and the origin country has been omitted in order to highlight any major differences arising due to the inflation rate. In the immigration empirical model, each variable is bilateral, and some authors have shown that when a variable only applies to either the source or origin country it may lead to standard error clustering (Lewer and Berg, 2007). Furthermore, it is possible for the correlation between population density and total population to affect the results. Thus in the third specification of Table 4 , I have omitted total population of both Canada and the origin country along with the inflation rate. The fourth specification of Table 4 shows the results when population density is omitted and instead we only include total population. Specifications four and five show the erroneous estimation results that we would get when the model is not correctly specified.

Gravitational and demographic factors:

An important variable in the gravity model of trade and migration is the bilateral distance between the two countries. All the previous works in the literature have shown that distance has a negative impact on trade and migration flows between two countries. Intuitively the idea has been that the longer is the distance between two countries, the higher would be the transportation costs and more obstacles are likely to be encountered in the process of obtaining information about the migration process to the destination country. However it should be noted that this general expectation about the effect of distance on migration flows goes back to a time when transportation was a costly and cumbersome process and when this industry had not experienced major technological

breakthroughs. In the first specification, distance is significant at the 1% level of significance. However when we move to the second specification in which inflation is omitted, distance loses its significance, this could be due to low quality of data available on inflation which may have led to questionable results in the first specification. Also, in the third specification in which I omit total population and only include population densities, distance is not significant. An implication is that the results regarding distance shown in the fourth and fifth specifications do not convey the true picture and could be just as a result of model misspecification. Thus the OLS estimation results show that, although distance has been almost always shown to be significant with a negative coefficient, that may be closely related to the time period being covered and the level of technological progress in that period. My dataset covers the period starting from 1980 and since then there have been many inventions and improvements in the transportation industry. Intuitively when a family is deciding to emigrate, spending ten versus five hours in the airplane would not make a huge difference, neither in the cost nor in the hassles of moving an entire family. Thus, it may make sense for distance to lose its significance with the facilitation of travelling long distances.

It is expected that the higher is the total population of the origin country the higher would be the immigration rates from that country. Consistent with previous works in the literature, I find that total population of the source country is significant at the 1% level in all three specifications in which it is included. On average for every 1% increase in the population of the source country, immigration flows to Canada increase by around 0.12%. However, the total population of Canada is not significant in any of the specifications.

A measure of the age structure of the population is the potential support ratio (psr) which, as explained earlier, shows the number of people aged 15 to 64 for every person who is 65 years or older. Since it is a measure of population ageing, it is expected that a low psr for a destination country would be associated with a higher demand for immigrants. In the first and last specifications of Table 4, the potential support ratio of Canada is significant with an expected negative coefficient. Intuitively, we would expect a high psr for the country of origin to be associated with a higher immigration flow as it would indicate more working age population who would have a higher incentive to immigrate. In three of the specifications, psr of the origin country is significant but in contradiction to what is expected it has a negative coefficient.

Two other demographic factors included in the model are the infant mortality rate (imr) and life expectancy at birth. We expect that the higher is the infant mortality rate of the origin country the higher would be people's incentive to emigrate out of that country. Consistent with our expectation, imr of the origin country is consistently significant at the 1% level in all five models with a positive estimated coefficient. On average every 1% increase in the infant mortality rate of the origin country is associated with a 0.70% increase in the migration flows to Canada. However, the imr of Canada also has a positive coefficient, which is opposite to what is generally expected; but its significance in all models is less than that of the origin country. The other measure is life expectancy at birth; we would expect a low life expectancy in the country of origin to be a sign of poor economic conditions. Life expectancy of the origin country is consistently significant in all equations with a positive estimated coefficient. The life expectancy of

Canada is not significant as we would not expect it to be a very important determinant of immigration flows to a country.

Economic Factors:

Most of the existing works done in the literature have used GDP per capita of a country as a proxy for the wage levels of a country and as a measure of the overall economic condition of that country. I collected two different measures of GDP per capita: 2000 constant US dollars GDP per capita and purchasing power parity (ppp) measure of GDP per capita (constant 2005 international dollars). Both measures give similar estimation results and as the purchasing power parity measure is commonly used in the literature, Table 4 uses this measure of GDP per capita to represent a country's level of economic development. As explained earlier, because of the dual role played by GDP per capita, we would expect a concave relationship between GDP per capita and immigration flows. As expected, in specifications one to four of Table 4, GDP per capita of the origin country is consistently significant at the 1% level with an estimated positive coefficient and its square is also significant with a negative coefficient, thus showing a concave relationship. The only specification that does not show a concave relationship is the last one in which GDP per capita of the origin country shows a convex relationship. This is perhaps as a result of misspecification since many important explanatory variables are omitted in the last specification. In a country that is very poor, there are incentives to emigrate out of that country but very little economic affordability, and thus low immigration flows occur from that country. As a country starts growing economically, people can better afford the immigration costs and thus there is a higher emigration rate from such a country. Once a country has become highly developed, its residents have

much less incentive to emigrate and thus the immigration rate from that country falls. The empirical results of my panel dataset support the concave relationship hypothesis. The results of GDP per capita for Canada are not consistent with what is expected. It is significant at 10% in only three specifications but its estimated coefficient is negative which is a surprising result.

The unemployment rate is an economic factor that has been generally referred to in the literature as one of the main incentives for migration. The labor market model of migration also suggests that people migrate in seek of a better job. Unemployment of the origin country is consistently significant in all five equations with the expected positive coefficient. The higher is the unemployment rate of an origin country, the higher would be the economic incentives for migration, and thus it is associated with higher migration flows. We expect Canada's unemployment rate to inversely impact immigration to that country as people would have less motivation and interest to move to a country with a high unemployment rate. Canada's unemployment is statistically significant in all specifications except the first one, which might be a sign that the first specification is not correct. As expected in all other specifications, the unemployment of Canada inversely impacts its immigration inflows.

The inflation rate is an economic factor that has not been widely used in the migration literature. In the first specification of Table 4, inflation of the origin country is not significant and that of Canada is significant but with a positive estimated coefficient which is not consistent with what is generally expected. As explained earlier, this may also be due to the low quality of data on inflation. Moreover, when we omit inflation from the model, distance loses its significance which could show that the low quality data

on inflation also negatively impacted other estimation results; thus it is omitted from the other specifications.

A measure of the overall level of literacy in a country has also been used as a potential determinant of migration flows; however due to lack of availability of data I could not get such a measure and instead use the primary school enrollment. A higher primary school enrollment would generally be associated with higher literacy rates and we can expect immigration rates to be higher among more educated societies. Primary school enrollment of the origin country is statistically significant at the 1% level in all specifications in which it is included. As expected, it has a positive estimated coefficient which implies that the higher is the primary school enrollment rate in a source country the higher would be the migration rates from that country. Moreover, the primary school enrollment of Canada is also significant at the 1% level in all four specifications and the estimated coefficient is positive too. Intuitively it makes sense for people to be attracted to a destination country that has a society with a high literacy rate, as the demand for jobs and employment opportunities are likely to be higher in a more educated society.

Domestic politics and human rights indicators:

The two measures of freedom obtained from the Freedom House are the two indices for civil liberty and political rights. The two measures range from 1 to 7, with 1 indicating the most free country and 7 the least free. The lower is the civil liberty status of a source country, the higher would be the incentives for its residents to migrate, but at the same time the less may be their freedom to migrate. The civil liberty index for the source country is significant at the 1% level in all specifications in which it is included with a negative estimated coefficient. The interpretation is that the lower is the level of

civil liberty (higher index) in a source country the less freedom people have to be able to migrate and thus a higher civil liberty index is associated with lower migration outflows.

The political rights index for the source country is also significant in all four specifications with an estimated positive coefficient. The idea is that the higher is the political rights index which indicates less political freedom, the higher would be people's incentives to migrate and thus it is associated with higher migration flows from that source country.

I also have another set of variables indicating the human rights status in a country. I have included the two additive indices in my specification. The physical integrity index is a combination of indices on torture, political imprisonment, extrajudicial killing and disappearance. The physical integrity index for the origin country is statistically significant at the 1% level in four specifications in which it is included with a negative estimated coefficient. The higher is this index the higher would a government's respect for these four rights and thus there would be less political incentive for people to migrate as the source country already has a good level of political freedom. The physical integrity index for Canada is not significant; this could be a reflection that what is really important in attracting migrants to a country is its economic status rather than its political status.

The other measure is the empowerment rights index which is constructed from the combination of indices on foreign movement, domestic movement, freedom of speech, freedom of assembly and association, workers' rights, electoral self-determination, and freedom of religion indicators. In the specifications shown in Table 4, this index is not significant, neither for Canada nor for the source country. However, when I omit the physical integrity rights index for both Canada and the source country from the model,

the empowerment rights index for the source country becomes significant at the 5% level with a negative estimated coefficient. This result could be due to the collinearity between the physical rights index and the empowerment rights index.

Historical, cultural and location factors:

Population density is often included as an explanatory variable in the migration equation. The idea is to capture the effect of changes in the population per square kilometer on the propensity to migrate. As Canada's population has not experienced a rapid change in the period covered by the dataset, in a specification we can either include Canada's total population or its population density, but both cannot be included simultaneously. In the third and fifth specifications I have included Canada's population density but it is not significant in explaining variation in migration flows. However, population density of the origin country is an important determinant of migration flows, and in all four specifications it is significant at the 1% level with an expected positive estimated coefficient. The higher is the concentration of people per square kilometer the higher would be people's incentive to migrate out of that country. The result of the third specification of Table 4 indicates that for every one percentage increase in the population density of the origin country, the migration flows from that country on average increases by 0.417%.

I have also included a set of dummy variables to capture cultural and historical links between the destination and source countries. The dummy variables on colony and common language are both consistently significant at the 1% level in all four specifications in which they are included with a positive estimated coefficient. As expected, a source country that had been in a colonial relationship with Canada has a

higher migration flow to Canada than a country that has had no colonial relationship. Similarly a source country that shares an official language with Canada has higher migration flows than one that does not share any language with Canada. These findings are consistent with the idea that having some cultural and historical commonalities would facilitate the immigration process. The dummy variable for being landlocked is also significant at the 1% level in all specifications of Table 4 with a negative estimated coefficient. Movement out of a country that is landlocked is expected to be more restricted than one that is not and thus being landlocked is associated with less migration flows.

I have also included a set of regional dummy variables to capture characteristics affecting migration to Canada which are specific to different regions of the world. I have omitted North America as the reference category for my regional dummies. Since this category includes only the United States, it captures any features specific to the U.S. that impact its migration flows to Canada. The eight regional dummy variables are mostly significant at the 1% level in all specifications except the fourth one. In specification four, population density is omitted and only total population is included. Omission of population density makes all regional dummies except East Asian dummy insignificant and this could further confirm the misspecification of the fourth model in Table 4. The estimated coefficients for all regional dummies in all other specifications are negative which indicates that migration flows from other regions in the world is lower relative to North America. In my dataset, being in a region other than the United States negatively impacts migration flows to Canada. This could be as a result of the proximity of the U.S. to Canada and more importantly due to the cultural ties between the two countries.

Overall, the OLS results show that a set of economic, demographic and political factors are likely to affect immigration flows. Among the main factors of the origin country affecting immigration flows out of the country are: GDP per capita, unemployment rate, population density, and total population. Moreover, the origin country's level of civil liberty and political rights are also found to significantly impact immigration flows. Canada's unemployment rate and primary school enrollment are also important determinants of immigration flows to it. A set of dummy variables representing cultural and historical ties between the source country and Canada as well as dummies capturing regional characteristics of the source country are statistically significant in promoting or impairing immigration flows. Given the OLS findings, now I investigate their robustness by using some estimation techniques specific to panel datasets.

5.3 Panel data estimation methods

A panel dataset enables one to better control for variables that are not observable or measurable and whose effects are captured by the error term of the regression. In this section I will present the results of the panel dataset estimation techniques: fixed effects and random effects models. We can consider the following model:

$$Y_{i\,t} = x_{it}'\beta + \varepsilon_{it} = x_{it}'\beta + \eta_i + u_{it} \qquad (16)$$

where x_{it} is a $k \times 1$ vector of regressors and ε_{it} is the error term which can be divided into two unobservable components. The unobservable η_i is called the time-invariant component or the unobserved heterogeneity. The other component u_{it} is called the time-variant unobservable or transitory shock.

An important issue in the estimation of panel datasets is the possibility of correlation between the explanatory variables and the unobserved heterogeneity (η_i). This has given rise to the two different models for estimating panel datasets. The fixed effects approach allows the explanatory variables to be correlated with the unobserved heterogeneity, but does not make any assumption on their joint distribution. The random effects approach is based on some assumption about the joint distribution of x_{it} and $\eta_{i.}$ For empirical purposes, the random effects model in Stata assumes the correlation between x_{it} and η_i to be zero.

The fixed effects model allows us to analyze the relationship between the dependent and explanatory variables within some possible group of observations; in this paper the group of observations is each source country. Each source country has some characteristics specific to it that may or may not be correlated with explanatory variables. Through fixed effects model we can remove the impact of country-specific characteristics that are time-invariant from the explanatory variables so that the estimated coefficients represent the net effect of explanatory variables on our dependent variable (Torres-Reyna, n.d.). The fixed effects model controls for the time-invariant differences between source countries so that the estimates will not be biased because of omitted time-invariant characteristics. One drawback of the fixed effects model is that it cannot be used to analyze the effect of time-invariant factors themselves on the dependent variable. Basically any time-invariant characteristic of an entity is perfectly collinear with any source country dummy (Torres-Reyna, n.d.). The xtreg and the fe commands of Stata are used in this paper. This set of Stata commands is basically a within-groups estimator in which first the mean of each variable is found and then the mean is subtracted from the

original model in levels. Thus this command addresses the unobserved heterogeneity of each source country by eliminating it from the model; any time-invariant explanatory variable would thus be omitted from the model. Another method which will give the same estimated coefficients is to use the OLS regression (reg) command in Stata with a dummy variable created for each source country and no constant to avoid collinearity. Even in this second method, any time-invariant variable should be omitted as it will be perfectly collinear with the source country-specific dummy variable.

The other estimation method used in this paper is a random effects model under which it is assumed that the variation across source countries is random and uncorrelated with the explanatory variables. Unlike the fixed effects model, in the random effects model we can include time-invariant factors as explanatory variables. The key idea in the random effects model is to estimate variances of the two components of the error term and then use these estimates to construct efficient weights and transform the model using these weights (Torres-Reyna, n.d.). The random effects and fixed effects estimation results of my panel dataset are shown in Table 5.

Table 5: Results of Panel Data Estimation Techniques

Dependent Variable:	Log (annual immigration flow)					
Model	Fixed Effects		Random Effects1		Random Effects2	
log GDP/capita (origin)	3.333**	(1.255)	3.875***	(1.049)	4.243***	(1.071)
Square of log GDP/capita (origin)	-0.185**	(0.076)	-0.218***	(0.065)	-0.242***	(0.067)
log GDP/capita (Canada)	-3.093***	(0.964)	-3.115***	(0.985)	-3.157***	(0.995)
log (distance to Canada)					0.579	(0.668)
log potential support ratio(origin)	-1.314***	(0.455)	-1.128***	(0.418)	-0.966**	(0.403)
log potential support ratio(Can)	-5.212	(4.921)	-4.565	(5.002)	-4.132	(4.980)
log infant mortality rate (origin)	1.443***	(0.387)	1.483***	(0.366)	1.513***	(0.372)
log infant mortality rate (Canada)	2.060**	(0.942)	1.977**	(0.965)	1.933**	(0.970)
log life expectancy (origin)	5.782**	(2.137)	6.986***	(1.982)	7.084***	(2.008)
log life expectancy (Canada)	-4.616	(10.86)	-5.480	(10.91)	-6.017	(10.90)
log population density (origin)	1.828***	(0.644)	0.954***	(0.238)	0.703***	(0.184)
log population density (Canada)	6.704	(6.747)	8.364	(6.763)	9.515	(6.662)
Unemployment rate (origin)	0.0416***	(0.0105)	0.0425***	(0.0103)	0.0426***	(0.011)
Unemployment rate (Canada)	-0.149***	(0.0305)	-0.151***	(0.0312)	-0.154***	(0.031)
Primary school enrollment (origin)	-0.00121	(0.0053)	-0.000384	(0.0052)	0.00013	(0.005)
Primary school enrollment (Can)	0.0775***	(0.0100)	0.0771***	(0.0103)	0.0767***	(0.010)
Civil liberty (origin)	-0.0960	(0.0620)	-0.0984	(0.0614)	-0.102	(0.064)
Political rights (origin)	0.103**	(0.0412)	0.101**	(0.0441)	0.0991**	(0.045)
Physical integrity index (origin)	-0.0727**	(0.0353)	-0.0755**	(0.0366)	-0.0794**	(0.036)
Physical integrity index (Canada)	0.0104	(0.0144)	0.00915	(0.0146)	0.00796	(0.015)
Empowerment right index (origin)	-0.0104	(0.0246)	-0.0139	(0.0225)	-0.0143	(0.022)
Empowerment right index (Can)	-0.091***	(0.0239)	-0.088***	(0.0241)	-0.087***	(0.024)
Colony dummy					1.771***	(0.374)
Common language dummy					0.462	(0.361)
Landlocked dummy					-0.766**	(0.366)
Africa dummy					-5.001**	(2.094)
Europe dummy					-4.925***	(1.687)
Latin America & Caribbean					-5.203***	(1.547)
Pacific dummy					-2.320	(2.040)
South East Asia dummy					-4.839**	(2.453)
East Asia dummy					-3.596	(2.251)
Southern Asia dummy					-5.381**	(2.534)
West Asia dummy					-5.355***	(2.034)
Constant	2.503	(40.04)	-0.552	(39.76)	-1.365	(40.71)
Observations	1,163		1,163		1,163	
Within R^2	0.4364		0.4300		0.4254	
Overall R^2	0.2558		0.2717		0.7209	
Rho	0.9517		0.8713		0.6570	

Notes: Robust standard errors in parentheses. Rho indicates the fraction of variance due to unobserved heterogeneity (i.e. due to difference across panels). *** p<0.01, ** p<0.05, * p<0.1

As mentioned earlier, in the fixed effects model time-invariant variables cannot be included, whereas in the random effects model we can estimate their impact. Before exploring the results it should be mentioned that I have used the robust standard errors for

the estimations. Stata has a built-in command called "xttest3" which is a modified Wald test for group-wise heteroskedasticity in a fixed effects regression model. The null hypothesis of the test is that there is homoskedasticity of the error term. The test results indicate a χ^2= 875.03 and a p-value=0.000. Thus we reject the null hypothesis in favor of the alternative hypothesis which indicates that there is indeed heteroskedasticity present. The way to address this is to use robust standard errors as presented. I have used two specifications for the random effects model. In the second specification of the random effects model, I have added several dummy variables capturing cultural and historical ties between the origin country and Canada as well as dummy variables representing regional characteristics of the origin countries.

In both fixed effects and random effects models, GDP per capita of the source country exhibits a concave relationship to the immigration flows which is consistent with the inverse U-shaped relationship stated in the previous section. However, the GDP per capita of Canada is significant in both fixed and random effects models, but with an unexpected negative coefficient. We generally expect the destination country's level of economic development to attract immigrants rather than repel them. This is a puzzling result which was not encountered at this level of significance in the Ordinary Least Squares (OLS) estimation results.

Consistent with the OLS results, the potential support ratio of the origin country is significant with a negative estimated coefficient and that of Canada is not significant. The infant mortality rate and life expectancy of the origin country are also significant with the expected signs. As in the OLS results, the population density of the origin country is statistically significant at 1% in both fixed and random effect models with a positive

impact on the annual immigration outflows. The unemployment rate of Canada and the origin countries are consistently significant in both models with the expected signs of estimated coefficients. The primary school enrollment of Canada is significant with a positive impact on the immigration inflows; however unlike the OLS results the primary school enrollment of the origin country is not significant. Similar to the OLS results, the political rights index and the physical integrity index of the origin country are both significant with the expected signs. However, the civil liberty index of the origin country which was significant in the OLS results does not preserve its significance in the fixed and random effect models. Moreover, in the random effects model dummy variables for colony and for whether an origin country is landlocked or not are both significant with the expected signs. Most of the regional dummy variables are also significant with negative estimated coefficients which is consistent with our expectation, given that the U.S. is the reference category.

Most of the panel data estimation results are consistent with those of OLS in Table 4; this provides support for the robustness of the OLS findings. Amongst the main factors affecting migration flows are origin country's GDP per capita with a concave relationship to immigration flows, origin country's unemployment rate and population density, and origin country's political rights and physical integrity indices. Similar to OLS results, Canada's unemployment rate and primary school enrollment are also important factors affecting migration inflows.

5.4 Hausman specification test

The test of specification that is generally used to choose between the fixed and random effects models is the Hausman test. This is basically a test of whether the unobserved heterogeneity (η_i) is correlated with the explanatory variables or not. The null hypothesis is that there is no correlation between the explanatory variables and the time-invariant component of the error term. Under the null hypothesis both the fixed and random effects models are consistent, but the random effect model is the efficient one. The alternative hypothesis is that the unobserved heterogeneity is correlated with the explanatory variables, and thus under the alternative hypothesis the random effects model is inconsistent and only the fixed effects model is consistent. I have performed the Hausman test using Stata's Hausman command along with the sigmamore option. The sigmamore option specifies the two variance-covariance matrices used to calculate the test statistic to be based on the common variance matrix obtained from the efficient model.

The Hausman test results in a χ^2= 37.54 and a p-value= 0.0003. Thus the results of the test show that we can reject the null hypothesis at the 1% level. The result indicates that there is evidence for correlation between the unobserved heterogeneity and the regressors and thus the fixed effect model results are consistent. Comparing the OLS results with those of the panel estimation methods shows that most of the OLS findings are robust.

6. Further Empirical Tests: using lags of all time-variant regressors

This section presents the results of a refinement in the empirical specification. Mayda (2009) suggests that reverse causality and more generally the issue of endogeneity in the time series dimension of the analysis is a possibility. One way to deal with the possibility of endogeneity is to relate the annual immigration flows to the first lag of all time-varying explanatory variables. Assuming strict exogeneity of such variables may be somehow unrealistic, as it may be more realistic to assume that they are predetermined (Mayda, 2009).

There could also be an intuitive reason for using the lags of time-varying explanatory variables. It may not be realistic to think of a family's decision to migrate as an instantaneous one. In other words, although it is true that a set of economic and non-economic factors affect a person's decision to migrate and his choice of destination country, we should allow for at least a short time span between the occurrence of such factors and the actual immigration. One way to incorporate this timing issue is to relate the current migration flows to past year's economic and non-economic factors rather than current values. It is realistic to think that when a person observes the current situation in the origin country and potential destination countries and decides to migrate, the actual immigration occurs next year rather than right away. Thus, including lags of time-varying factors could be supported both on econometric and intuitive grounds.

Table 6 shows the OLS results of the first three specifications of Table 1 repeated with the one-year lag of all time-varying variables used. Most of the estimation results using the one-year lag of all potentially time-varying variables are consistent with those of the previous OLS results in Table 4. The GDP per capita of the origin country shows a

concave relationship with the immigration flow as was found in the original OLS results. However, a major improvement compared to the previous results is that the GDP per capita of Canada is statistically significant in all three specifications with the expected positive impact on the immigration flows to Canada. Moreover, in Table 4 the potential support ratio of Canada was only significant at the 10% level in two of the specifications; however, here it is significant at the 1% level in all three specifications with the expected negative coefficient. We generally expect countries with a low potential support ratio to have a high need for a working age population and thus we expect them to have a higher demand for immigrants. Life expectancy of the origin country is only significant in the third specification whereas in the original OLS results it was significant in all specifications. Population density and unemployment rate of the origin country both exhibit significance at the 1% level in all three specifications with the expected positive coefficient. However, the Canadian unemployment rate and primary school enrollment are no longer significant when we do the OLS estimation with the lagged values. All the other variables exhibit the same pattern as they did in the original OLS estimation.

Table 6: OLS results with one-year lag of time-varying regressors

Dependent Variable:	Log (annual immigration flow)					
Specification	(1)		(2)		(3)	
L.log GDP/capita(origin)	3.911***	(0.793)	3.738***	(0.625)	3.469***	(0.625)
L.Square of log GDP/capita(origin)	-0.22***	(0.0445)	-0.217***	(0.036)	-0.20***	(0.0360)
L.log GDP/capita (Canada)	5.463***	(2.042)	2.794*	(1.655)	2.817*	(1.666)
log (distance to Canada)	-0.341**	(0.137)	-0.0220	(0.127)	0.127	(0.121)
L.log population(origin)	0.113***	(0.0348)	0.132***	(0.034)		
L.log population(Canada)	-21.26**	(9.058)	-12.76	(7.960)		
L.log potential support ratio(origin)	-0.39***	(0.118)	-0.227**	(0.115)	-0.167	(0.115)
L.log potential support ratio(Can)	-28.9***	(6.755)	-21.52***	(5.683)	-21.3***	(5.719)
L.log infant mortality rate (origin)	0.614***	(0.116)	0.434***	(0.112)	0.632***	(0.100)
L.log infant mortality rate (Canada)	5.573***	(1.168)	4.862***	(1.109)	4.875***	(1.116)
L.log life expectancy (origin)	0.968	(0.875)	1.053	(0.868)	2.117**	(0.828)
L.log life expectancy (Canada)	-11.84	(19.40)	-10.32	(19.17)	-10.40	(19.30)
L.log population density (origin)	0.391***	(0.0291)	0.410***	(0.029)	0.416***	(0.0291)
L.unemployment rate (origin)	0.020***	(0.0065)	0.019***	(0.006)	0.023***	(0.0063)
L.Unemployment rate (Canada)	0.0733	(0.0601)	-0.0273	(0.041)	-0.0259	(0.0417)
L.Inflation (origin)	3.09e-05	(5.2e-05)				
L.Inflation (Canada)	0.0502**	(0.0234)				
L.Primary school enroll(origin)	0.014***	(0.0029)	0.0099***	(0.003)	0.013***	(0.0027)
L.Primary school enrollment(Can)	-0.00672	(0.0279)	0.0189	(0.026)	0.0192	(0.0257)
L.Civil liberty (origin)	-0.115***	(0.0424)	-0.133***	(0.042)	-0.13***	(0.0424)
L.Political rights (origin)	0.103***	(0.0349)	0.134***	(0.034)	0.136***	(0.0340)
L.Physical integrity index (origin)	-0.168***	(0.0169)	-0.171***	(0.017)	-0.19***	(0.0156)
L.Physical integrity index (Canada)	-0.0627	(0.0564)	-0.0317	(0.054)	-0.0309	(0.0542)
L.Empowerment right (origin)	-0.0251*	(0.0143)	-0.0201	(0.014)	-0.0113	(0.0139)
L.Empowerment right index (Can)	0.00785	(0.0595)	-0.0166	(0.058)	-0.0182	(0.0586)
Colony dummy	1.305***	(0.138)	1.274***	(0.138)	1.589***	(0.112)
Common language dummy	0.979***	(0.0837)	1.021***	(0.083)	0.881***	(0.0752)
Landlocked dummy	-0.739***	(0.100)	-0.846***	(0.098)	-0.96***	(0.0942)
Africa dummy	-1.264**	(0.544)	-2.282***	(0.518)	-3.17***	(0.467)
Europe dummy	-1.553***	(0.416)	-2.299***	(0.399)	-3.13***	(0.337)
Latin America &Caribbean	-2.147***	(0.456)	-2.742***	(0.446)	-3.79***	(0.354)
Pacific dummy	0.242	(0.504)	-0.812*	(0.475)	-1.62***	(0.429)
South East Asia dummy	-1.314**	(0.592)	-2.554***	(0.557)	-3.47***	(0.507)
East Asia dummy	-0.124	(0.511)	-1.256***	(0.477)	-1.89***	(0.450)
South Asia dummy	-1.461**	(0.605)	-2.639***	(0.571)	-3.39***	(0.540)
West Asia dummy	-1.982***	(0.502)	-2.848***	(0.482)	-3.82***	(0.413)
L.log population density (Canada)					-12.07	(8.008)
Constant	386.8***	(115.1)	247.2***	(91.90)	38.91	(64.76)
Observations	1,088		1,125		1,125	
R-squared	0.812		0.805		0.802	

Notes: L.before a variable indicates the first lag. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

7. Conclusions and Extensions

In this paper, I empirically analyze the economic and non-economic determinants of immigration to Canada. A modified version of the gravity model of migration is developed in order to incorporate demographic, economic and political characteristics of both origin and destination countries. The economic development of the origin country captured by GDP per capita is found to be a significant factor, conveying an inverse U-shaped relationship with immigration flows. The significance of the origin country's GDP per capita is robust to the different empirical methods used. GDP per capita of the destination country which is considered as a pull factor is also a significant determinant of immigration flows in the model with lags of time-varying factors. I find the unemployment rate, total population and population density of the origin country as significant determinants of migration flows in all specifications. Moreover, the unemployment rate and primary school enrollment of Canada are also important factors that affect migration flows to Canada.

Perhaps an important contribution of this paper is the inclusion of non-economic factors in the migration model and the new results found regarding the distance between the two countries. In all except two OLS specifications, distance between the origin country and Canada is not a statistically significant factor. A potential explanation for this loss of significance compared to previous studies is the technological advances and decline in travel costs in the transportation industry. As movement of people and goods have been greatly facilitated through different modes of transportation, it does make sense for the significance of distance to fade away.

The civil liberty and political rights indices are significant factors affecting migration flows. A low level of civil freedom in the origin country is found to impair migration . Moreover, an origin country with less political freedom creates higher incentives for its residents to immigrate and thus promotes migration. The empirical results indicate that in addition to political variables, human rights indicators also have a considerable impact on the size and component of migration flows.

Besides population, other demographic variables that are significant determinants of migration flows are infant mortality rate, potential support ratio and life expectancy of origin and destination countries. A low potential support ratio is indicative of population ageing and may act as a signal for a country's demand for immigrants. In the model with lags of time-varying factors, the potential support ratio of Canada is consistently significant in all specifications with the expected negative impact on migration flows.

The dummy variables for colonial relationship, being landlocked and having a common official language with Canada are all consistently significant in explaining variations in migration flows to Canada, with expected signs. This indicates that factors which capture cultural and historical ties are significant determinants that may promote or impair migration. The significance of all regional dummy variables indicate that international migration to Canada varies geographically by the region of the source country.

To conclude, the present study uses the time-series and cross-section features of the collected panel dataset to empirically analyze the significant determinants of migration flows to Canada. The inclusion of carefully chosen political and human right indicators is perhaps a progress in this literature. Some of the findings are consistent with the previous

results and some findings convey sensible changes in the significance of factors such as distance.

In further work, one might carry a similar study for another destination country or a group of destination countries in order to examine whether the same result would hold with regards to the significance of distance. In addition, one might use an extended set of source countries in order to check the robustness of the above empirical findings.

References

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review, 77*(4), 531-553.

Borjas, G. J. (2001). Economics of immigration. *International Encyclopedia of the Social & Behavioral Sciences*, 9803-9809.

Borjas, G. J., Freeman, R. B., & Katz, L. F. (1991). On the labor market effects of immigration and trade. *National Bureau of Economic Research (NBER) working paper No. w3761*.

Chiswick, B. (1999). Are immigrants favorably self-selected? *American Economic Review, 89*(2), 181-185.

Facts and figures 2010-Immigration overview: Permanent and temporary residents. (September 2011). Citizenship and Immigration Canada. Retrieved from http://www.cic.gc.ca/english/resources/statistics/facts2010/index.asp

Karemera, D., Oguledo, V. I., & Davis, B. (2000). A gravity model analysis of international migration to North America. *Applied Economics, 32* (13), 1745-1755.

Kim, K., & Cohen, J.E. (2010). Determinants of international migration flows to and from industrialized countries: A panel data approach beyond gravity. *International migration review, 44* (4), 899-932.

Lewer, J. J., & Berg, H. V. (2007). A gravity model of immigration. *Economic Letters, 99*, 164-167.

Liebig, T., & Sousa-Poza, A. (2004). Migration, self-selection and income inequality: An international analysis. *Kyklos, 57*(1), 125-146.

Mayda, M. (2009). International migration: A panel data analysis of the determinants of

bilateral flows. *Journal of Population Economics*, *23*, 1249-1274.

Pedersen, P. J., Pytlikova, M., & Smith, N. (2004). Selection or network effects? Migration flows into 27 OECD countries, 1990-2000. IZA Discussion paper NO.1104.

Sjaastad, L.A. (1962). The cost and returns of human migration. *Journal of Political Economy*, *70*, 80-93.

Tansey, M. M., & Touray, A. (2010). The gravity model of trade applied to Africa. *The International Business and Economic Research Journal*, *9*(3), 127-130.

Torres-Reyna, O. Panel data analysis: Fixed and random effects (using Stata 10.x). Princeton University. Retrieved from http://dss.princeton.edu/training/ on June 16th, 2012.

Appendix

Table A1: **Annual immigration flows** to Canada by source country for the period 1980 to 2010 (measured in number of people):

country	Obs	Mean	Std.Dev.	Min	Max
Argentina	31	661.9355	381.4483	220	1780
Australia	31	812.9032	257.5292	335	1200
Austria	31	186.6129	62.37504	105	365
Belgium	31	457.9032	137.6939	215	745
Bolivia	31	96.45161	61.7008	30	250
Brazil	31	818.7097	631.6671	160	2600
Chile	31	660.6452	392.6167	250	1740
China	31	18074.68	13730.69	1880	42290
Colombia	31	1843.548	2015.648	210	6035
Costa Rica	31	155.9677	79.55522	35	320
Denmark	31	130	56.02083	60	295
Egypt, Arab Rep.	31	1692.097	845.0343	395	4305
El Salvador	31	1687.742	1715.693	110	7045
Finland	31	98.70968	34.32608	60	195
France	31	3646.613	1579.909	1380	7295
Greece	31	438.7097	252.9195	140	1090
Hong Kong	31	12639.03	13338.76	790	44225
Ireland	31	476.2903	316.1565	155	1335
Israel	31	1918.871	677.6414	425	2855
Italy	31	738.7097	391.512	345	2045
Jamaica	31	3181.774	1233.112	1685	6005
Japan	31	891.7742	422.5489	205	1645
Korea, Rep.	31	4082.097	2448.508	805	9610
Malaysia	31	695.8065	430.345	200	1925
Netherlands	31	833.3871	358.4745	470	1865
Norway	31	101.6129	36.63757	25	195
Pakistan	31	6132.097	5062.523	480	15350
Paraguay	31	84.35484	31.79995	35	145
Philippines	31	12710.16	7422.246	3080	36580
Portugal	31	1593.548	1773.123	280	5975
Singapore	31	672.4194	408.0336	165	1695
Spain	31	191.9355	93.71657	85	440
Sweden	31	218.871	47.18244	140	325
Thailand	31	396.6129	390.0243	70	1930
Trinidad and Tobago	31	1563.387	1038.262	595	4335
Turkey	31	955.4839	503.945	210	2060
United Kingdom	31	7695.806	4048.719	3900	21160
United States	31	7487.258	1905.961	4775	11215

Uruguay	31	152.4194	78.42056	75	430
Venezuela	31	593.5484	398.1084	135	1385
Total	1240	2436.762	5257.537	25	44225

Table A2: **Annual immigration flows from the 40 source countries to Canada by year** (measured in number of people):

Year	Obs	Mean	Std.Dev.	Min	Max
1980	40	1776.25	3367.259	35	18250
1981	40	1924.375	3808.241	35	21160
1982	40	1676.125	3039.489	45	16450
1983	40	1122.375	1816.843	45	7380
1984	40	1063.125	1804.34	25	7695
1985	40	1010.75	1715.16	45	7380
1986	40	1173.75	1786.118	50	7280
1987	40	1962.375	3236.581	80	16200
1988	40	2101.875	4056.437	60	23270
1989	40	2291.875	3780.271	65	19940
1990	40	2659.625	5206.034	70	30050
1991	40	2741.5	4591.322	60	22555
1992	40	3127.875	6620.938	60	39350
1993	40	3122.25	6565.911	100	36650
1994	40	2997	7618.763	75	44225
1995	40	2627.75	5743.758	50	31765
1996	40	2719.875	5698.749	50	29990
1997	40	2489	4891.063	45	22250
1998	40	1960.25	3689.042	35	19795
1999	40	2277.375	4980.972	40	29145
2000	40	2685.625	6275.7	30	36745
2001	40	2988	6963.752	55	40365
2002	40	2603.75	5812.862	55	33305
2003	40	2708.125	6185.766	60	36250
2004	40	2855.75	6286.43	80	36425
2005	40	3235.375	7361.342	60	42290
2006	40	3011.25	6170.68	75	33075
2007	40	2904.75	5421.203	90	27015
2008	40	3171.625	6105.158	90	29335
2009	40	3133.5	6332.464	85	29050
2010	40	3416.5	7406.868	100	36580
Total	1240	2436.762	5257.537	25	44225

Table A3: **PPP measure of GDP per capita** (constant 2005 international $) of source countries

Country	Obs	Mean	Std.Dev.	Min	Max
Argentina	31	9942.323	1657.849	7458.294	14362.75
Australia	31	26525.42	5037.681	19706	34410.71
Austria	31	28137.01	4984.516	20910.71	36177.36
Belgium	31	27055.97	4383.487	20735.33	33592.73
Bolivia	31	3485.544	385.3495	2935.906	4349.923
Brazil	31	7839.434	838.7998	6556.751	10055.92
Chile	31	8829.857	2974.337	4628.076	13595.9
China	31	2384.235	1799.45	523.9503	6816.287
Colombia	31	6501.644	963.3459	5196.198	8479.323
Costa Rica	31	7437.811	1581.598	5466.292	10391.86
Denmark	31	28029.45	4307.474	20611.35	34595.28
Egypt, Arab Rep.	31	3703.207	870.9319	2403.684	5543.542
El Salvador	31	4617.477	933.3632	3500.786	6150.745
Finland	31	24512.8	4913.777	17857.73	33500.76
France	31	25540.4	3411.746	20284.56	30554.43
Greece	31	19600.69	3447.807	16068.94	26258
Hong Kong	31	27091.71	8060.61	13945.11	41712.86
Ireland	31	25209.92	10126.58	13164.84	41136.97
Israel	31	20118.79	3460.059	15027.84	26022.68
Italy	31	24653.11	3383.167	18814.44	29007.91
Jamaica	31	6233.968	813.4265	4808.933	7251.543
Japan	31	26280.6	3982.9	18777.9	31659.86
Korea, Rep.	31	15364.91	6727.733	5543.572	27026.79
Malaysia	31	8635.958	2723.762	4866.91	13213.91
Netherlands	31	29289.4	5484.092	21574.43	38105.94
Norway	31	38109.52	7769.93	26204.78	49175.28
Pakistan	31	1770.526	331.4805	1224.007	2410.858
Paraguay	31	4005.399	220.1101	3689.434	4647.702
Philippines	31	2746.571	331.2785	2308.517	3560.487
Portugal	31	17551.24	3746.507	11927.02	22067.94
Singapore	31	31697.84	11351.26	15066.85	51966.4
Spain	31	21733.78	4580.464	15214.32	28521.62
Sweden	31	26546.89	4635.864	20296.03	34782.18
Thailand	31	4846.573	1747.464	2220.6	7672.913
Trinidad and Tobago	31	14958.69	4472.197	10525.78	24080.34
Turkey	31	8960.389	2011.945	5959.152	12546.67
United Kingdom	31	25392.68	5658.813	17164.3	34321.35
United States	31	34833.57	6048.132	25148.49	43710.28
Uruguay	31	8648.672	1637.193	6184.916	12655.19

Venezuela	31	10134.78	861.4039	7872.955	11877.7
Total	1240	16723.97	11439.12	523.9503	51966.4

Table A4: **Annual permanent resident landings** in Canada by category and source area, 2001-2010

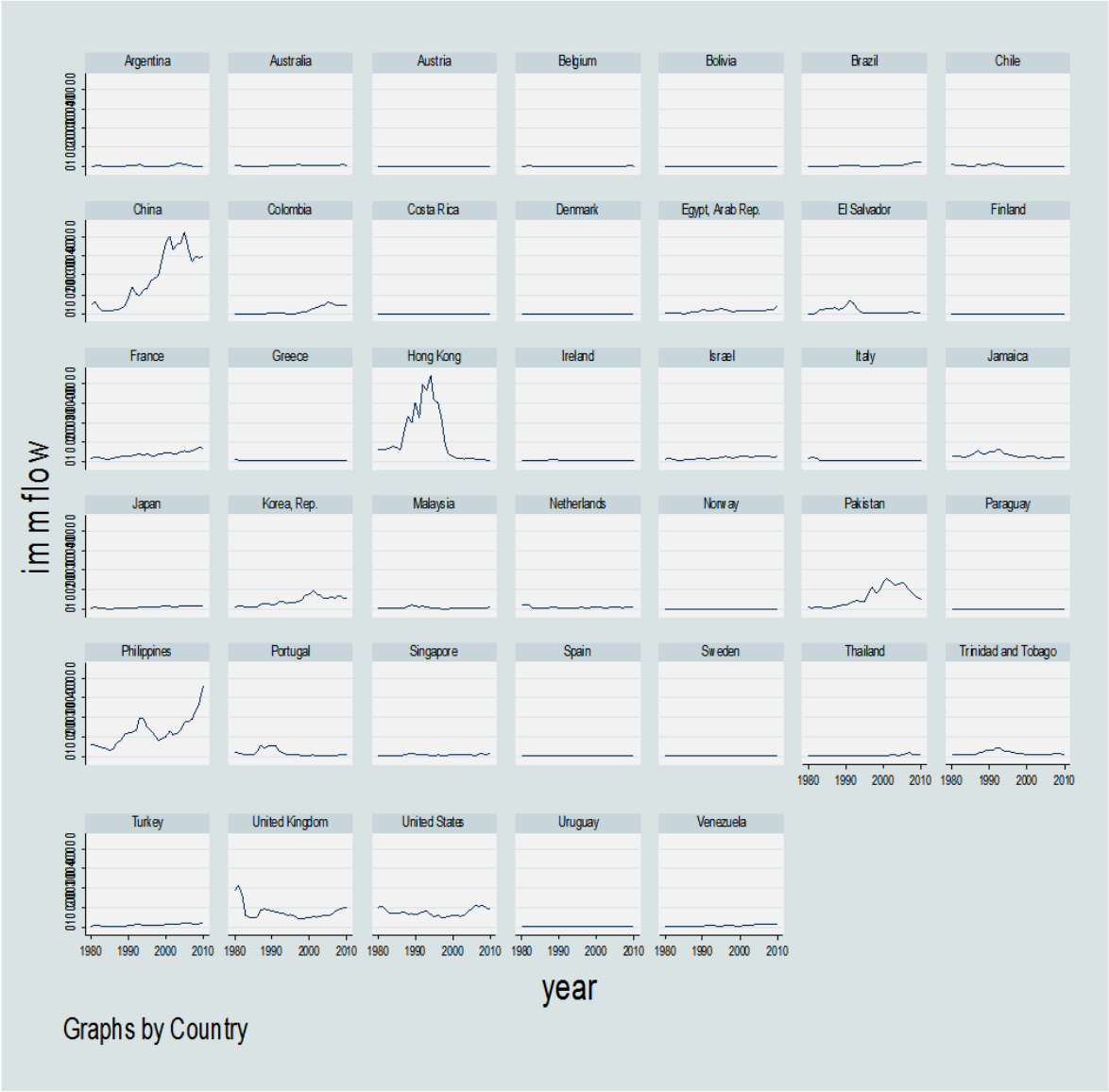
Source area	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
					Number					
Africa and the Middle East	7,808	6,389	7,462	7,904	7,624	8,352	8,966	9,315	9,685	9,437
Asia and Pacific	35,261	36,651	38,685	34,075	36,540	43,008	37,562	34,619	34,423	29,752
South and Central America	10,039	7,832	7,387	7,448	7,171	7,043	7,372	7,711	7,734	8,974
United States	3,614	2,783	2,988	3,707	4,149	4,468	3,952	4,688	4,455	3,987
Europe and the United Kingdom	10,055	8,629	8,589	9,117	7,866	7,626	8,362	9,225	8,895	8,062
Source area not stated	18	8	10	24	24	20	28	24	12	8
Family class	**66,795**	**62,292**	**65,121**	**62,275**	**63,374**	**70,517**	**66,242**	**65,582**	**65,204**	**60,220**
Africa and the Middle East	30,706	30,604	25,384	27,591	28,650	31,346	28,175	31,225	33,873	44,585
Asia and Pacific	87,727	71,212	62,242	66,479	87,740	69,423	62,299	73,304	73,153	96,284
South and Central America	7,474	8,040	7,313	8,454	8,205	7,191	9,466	11,392	12,390	12,148
United States	2,240	1,938	1,703	2,977	3,804	4,498	4,785	4,926	3,695	3,329
Europe and the United Kingdom	27,551	26,059	24,403	28,242	27,912	25,790	26,520	28,224	30,380	30,566
Source area not stated	19	10	2	4	1	2	0	0	0	1
Economic immigrants	**155,717**	**137,863**	**121,047**	**133,747**	**156,312**	**138,250**	**131,245**	**149,071**	**153,491**	**186,913**
Africa and the Middle East	9,662	8,820	9,535	12,592	11,443	10,223	9,250	9,013	10,833	11,035
Asia and Pacific	9,858	10,196	10,165	12,159	11,849	10,573	9,569	6,399	6,648	6,759
South and Central America	2,657	2,841	3,712	4,596	7,637	7,602	6,077	4,699	3,680	4,442
United States	55	33	45	132	772	1,246	831	482	537	1,190
Europe and the United Kingdom	5,683	3,220	2,523	3,177	4,007	2,764	2,153	1,223	1,133	1,228
Source area not stated	4	3	3	31	68	92	74	42	19	42
Refugees	**27,919**	**25,113**	**25,983**	**32,687**	**35,776**	**32,500**	**27,954**	**21,858**	**22,850**	**24,696**
Africa and the Middle East	61	527	1,293	1,444	1,562	1,937	2,171	1,760	1,762	1,634
Asia and Pacific	99	998	2,636	1,858	1,919	3,465	3,224	3,159	2,948	2,209
South and Central America	41	756	1,937	1,756	1,629	2,468	2,975	2,691	2,972	2,790
United States	0	540	1,277	691	538	731	881	1,120	1,036	736
Europe and the United Kingdom	5	959	2,054	1,366	1,121	1,764	2,036	1,977	1,902	1,462
Source area not stated	0	0	0	0	8	8	25	28	6	14
Other immigrants	**206**	**3,780**	**9,197**	**7,115**	**6,777**	**10,373**	**11,312**	**10,735**	**10,626**	**8,845**
Africa and the Middle East	48,237	46,340	43,674	49,531	49,279	51,858	48,562	51,313	56,153	66,691
Asia and Pacific	132,945	119,057	113,728	114,571	138,048	126,469	112,654	117,481	117,172	135,004
South and Central America	20,211	19,469	20,349	22,254	24,642	24,304	25,890	26,493	26,776	28,354
United States	5,909	5,294	6,013	7,507	9,263	10,943	10,449	11,216	9,723	9,242
Europe and the United Kingdom	43,294	38,867	37,569	41,902	40,906	37,944	39,071	40,649	42,310	41,318
Source area not stated	41	21	15	59	101	122	127	94	37	65
Category not stated	1	0	1	0	2	2	1	2	1	7
Total	**250,638**	**229,048**	**221,349**	**235,824**	**262,241**	**251,642**	**236,754**	**247,248**	**252,172**	**280,681**

Source: Citizenship and Immigration Canada: Facts and Figures 2010, immigration overview of permanent and temporary residents

Table A5: **OLS estimation results with total measure of GDP**

Dependent Variable: log (annual immigration flow)						
Specification	(1)		(2)		(3)	
log total GDP (origin)	2.676***	(0.545)	2.415***	(0.544)	2.498***	(0.539)
Square of log total GDP (origin)	-0.05***	(0.0106)	-0.05***	(0.0105)	-0.05***	(0.0105)
log total GDP (Canada)	-0.690	(1.848)	-2.975*	(1.620)	-2.933*	(1.619)
log (distance to Canada)	-0.356**	(0.138)	0.00484	(0.130)	0.0160	(0.129)
log population(origin)	0.0198	(0.0805)	0.0760	(0.072)		
log population(Canada)	2.572	(9.006)	9.948	(8.170)		
log potential support ratio(origin)	-0.43***	(0.123)	-0.241**	(0.122)	-0.216*	(0.119)
log potential support ratio(Canada)	-8.996*	(5.331)	-5.486	(4.964)	-5.420	(4.964)
log infant mortality rate (origin)	0.865***	(0.122)	0.67***	(0.117)	0.751***	(0.089)
log infant mortality rate (Canada)	1.788**	(0.830)	1.929**	(0.831)	1.918**	(0.831)
log life expectancy (origin)	2.279***	(0.879)	2.30***	(0.884)	2.573***	(0.845)
log life expectancy (Canada)	-10.33	(19.72)	-8.801	(19.68)	-9.084	(19.68)
log population density (origin)	0.433***	(0.0290)	0.456***	(0.029)	0.457***	(0.029)
unemployment rate (origin)	0.034***	(0.0063)	0.033***	(0.006)	0.034***	(0.006)
Unemployment rate (Canada)	-0.0520	(0.0523)	-0.14***	(0.039)	-0.14***	(0.039)
Inflation (origin)	1.36e-06	(5.4e-05)				
Inflation (Canada)	0.0527**	(0.0217)				
Primary school (origin)	0.016***	(0.0029)	0.011***	(0.003)	0.012***	(0.003)
Primary school enrollment(Canada)	0.066***	(0.0249)	0.081***	(0.024)	0.081***	(0.024)
Civil liberty (origin)	-0.12***	(0.0430)	-0.15***	(0.043)	-0.14***	(0.043)
Political rights (origin)	0.118***	(0.0354)	0.15***	(0.035)	0.156***	(0.035)
Physical integrity index (origin)	-0.13***	(0.0175)	-0.14***	(0.017)	-0.15***	(0.017)
Physical integrity index (Canada)	0.0106	(0.0479)	0.0129	(0.048)	0.0134	(0.048)
Empowerment rights index (origin)	-0.0271*	(0.0146)	-0.0175	(0.015)	-0.0189	(0.014)
Empowerment rights index (Canada)	-0.0800	(0.0566)	-0.0855	(0.057)	-0.0848	(0.057)
Colony dummy	1.635***	(0.144)	1.622***	(0.146)	1.670***	(0.139)
Common language dummy	0.848***	(0.0819)	0.872***	(0.082)	0.844***	(0.077)
Landlocked dummy	-0.57***	(0.110)	-0.69***	(0.110)	-0.69***	(0.110)
Africa dummy	-1.84***	(0.553)	-2.84***	(0.532)	-2.91***	(0.527)
Europe dummy	-2.35***	(0.442)	-3.17***	(0.426)	-3.24***	(0.421)
Latin American and Caribbean	-2.51***	(0.462)	-3.14***	(0.457)	-3.26***	(0.445)
Pacific dummy	-0.285	(0.517)	-1.42***	(0.489)	-1.48***	(0.485)
South East Asia dummy	-1.65***	(0.597)	-2.93***	(0.564)	-2.97***	(0.563)
East Asia dummy	-0.439	(0.508)	-1.69***	(0.478)	-1.71***	(0.478)
South Asia dummy	-2.08***	(0.611)	-3.21***	(0.581)	-3.22***	(0.581)
West Asia dummy	-2.50***	(0.513)	-3.43***	(0.498)	-3.55***	(0.484)
Log population density (Canada)					10.10	(8.170)
Constant	-10.89	(94.28)	-88.33	(83.29)	68.84	(62.31)
R-squared	0.803		0.791		0.791	

Figure A1: **Immigration flows to Canada** by source country, 1980-2010



Created in Stata using the xtline command for my panel dataset