# What Drives International Migration Flows? Evidence from 41 Countries 1970-2008 Using Non-Economic Determinants and Gravity Model[†]

Keuntae Kim[‡]

ABSTRACT

This paper quantifies the determinants for international migration flows by using only demographic, social, and geographic information owing to uncertainty and paucity of economic measures. The vast majority of past studies are focused on immigration into wealthy nations. Consequently, our knowledge on the determinants of emigration from developed nations to other countries is significantly limited. Yet, international migration has been increasing substantially and new destinations are emerging. The present study aims to fill the gap in the literature by combining international migration flows data from UN, Eurostat, and OECD from 1970 to 2008. The results from the gravity model and generalized estimating equations (GEE) support various theories of international migration. However, magnitude of each determinant of inflow was not identical in the outflow model, which suggests that migrants from wealth nations to the rest of the world are not constrained by economic resources as much as immigrants to rich countries.

[†]This paper is prepared for presentation at the 2011 Annual Meeting of the Population Association of America, March 31-April 2 Washington, DC.

[‡]Department of Sociology, University of Wisconsin – Madison, 7106 William Sewell Social Science Building, Madison, WI 53706. Tel: (608) 262-9518. Email: ktkim@ssc.wisc.edu.

**INTRODUCTION**

Since the beginning of the modern demography, the association between demographic phenomena and economic changes has been well established (e.g., Malthus 1970[1798]). About half century ago, Philip Hauser (1959) observed that economic data for developed countries are readily available and relatively well measured despite the complexity of the concepts whereas many developing countries do not maintain the statistics of basic economic indicators, and the quality of data is often questionable when it is available. And he proposed to use demographic measures as indirect estimation of the level of economic developments for those countries lacking appropriate statistics since the economy and population are associated very closely, and the latter is much more likely to be available than the former even among the least developed parts of the world. Sen (1993; 1998) argued that demographic variables, especially mortality conditions such as mortality rate or life expectancies, reflect general well-being of the population more accurately than economic measures do. Sen maintains that mortality statistics can address the blind spots of economic data, such as the efficacy of human agencies and organizations (e.g., 1974 Bangladesh famine) and gender/racial inequalities. Economic measures, most notably average personal income (i.e., GDP per capita), may not represent the actual income that people enjoy because distribution of income also matters. Furthermore, quality of life of a person depends not only on income but also on various social and physical conditions (Sen 1998). Easterlin (1978) established that the age structure of the population (i.e., ratio of older men to young men) is associated with various socioeconomic conditions such as unemployment rates, fertility rates, homicide rates, and suicide rates. Although his prediction about economic and social indicators in the 1980s based on the trajectories of the changes in age structure in the

1970s was falsified, the insight on the mechanisms of how demographic structure is interwoven with economy has an important implication for subsequent demographic studies.

Perhaps, the nexus between economic measures and demographic changes might be felt most strongly in the field of international migration. Early migration theorists suggested that pursuing the betterment of economic well-being is the most fundamental motivation for migration (Ravenstein 1885; Lee 1966). Later development of the migration theories pointed out the importance of other determinants, such as relative deprivation in sending communities (Stark and Bloom 1985; Stark and Taylor 1989), dual labor markets in receiving countries (Prior 1979), and penetration of capitalism into peripheral regions (Wallerstein 1974). Once international migration flows are initiated between two particular countries, self-reinforcing effects of migrant networks (Palloni et al. 2001), development of migrant supporting institutions, cumulative causations which make additional migrations more likely, and establishment of a stable migration systems (Massey et al. 1993) help the movements of people and goods perpetuate over time. Nonetheless, all theories of international migration suggest that a model should incorporate characteristics in both of sending and receiving communities as well as the intervening obstacles (Lee 1966). Of the characteristics, economic propensities are most essential and complex in nature and therefore most scarce.

Another significant stumbling block for the estimation of international migration flow is the lack of reliable and consistent data (DeWaard, Kim, and Raymer 2009). As the volume of international migration is growing substantially and greater number of countries are affected by migration regardless of their initial intentions (Castles and Miller 2003; Kim and Cohen Forthcoming), it is important to estimate the determinants of it. Yet, existing literature on the determinants of international migration tend to focus on immigration into a single country or a small number of countries that are similar in terms of socioeconomic conditions and cultural background. Moreover, empirical studies are limited to a few high-income countries that have necessary economic indicators for estimation.

The present paper aims to quantify the determinants of international migration flows by using only non-economic measures and an augmented gravity model. Most of the previous studies on international migration tend to be limited to immigrations to industrialized nations from less developed parts of the world (Kim and Cohen Forthcoming). And our knowledge on the flows in the opposite direction - international migration flows from wealthy nations to the rest of the world - is significantly limited. Furthermore, the small number of studies on the determinants of international migration flows is based on mostly North American and European countries. As growing number of the Central and Easter European and Eastern Asian countries become destinations for immigration in recent years (Castles and Miller 2003), it is important to examine the past empirical evidence by taking into account those newly emerging destination countries. The goal of the study is to fill the gap in the literature by combining international migration flows data from UN, Eurostat, and OECD reported by 41 countries from 1970 to 2008. The final sample yields 77,477 inflows and 50,667 outflows which were reported by lower middle to high income countries. The results from the gravity model and generalized estimating equations (GEE) support various theories of international migration. Consistent with what the gravity model suggests, the size of population in origin and destination have significant and positive effects on the inflows and outflows. Also, the distance between the two country significantly hampers the volume of movements. The effects of social, cultural, and geographical proximity were also significantly associated with inflows and outflows. The results indicate that the magnitude of each determinant is not identical in the inflow and outflow model, which suggests that migrants from developed to less developed countries are not constrained by economic resources as much as immigrants to rich countries.

## PREVIOUS RESEARCH

Empirically, the effect of distance between two countries is negative, significant, and robust across different model specifications (Greenwood and McDowell, 1982; Mayda, 2010). Increases in distance can be a proxy for increases in transportation cost and psychic cost (Greenwood, 1975). Persons tend to have less information about relatively distant places and are less likely to move to a locale about which they have little or no prior information. This argument suggests that if two countries share a border, the cost of moving could be significantly lower than otherwise, while a relatively inaccessible destination, for example, a land-locked country, should have fewer immigrants than countries with oceans or seas as borders, due to the increased cost of over-land transportation (Mayda, 2010).

Language, culture and shared history also affect international migration (Greenwood and McDowell 1982; Clark, Hatton, and Willamson 2007; Neumayer 2005; Mayda 2010; Karemera et al., 2000). For example, Clark and others (2007) found that having an English-speaking origin significantly and positively affected U.S.-bound immigration. Former colonial relationships appear to facilitate both trade and migration. The former colonial power's language is often spoken in the former colony, and the former colonial power may host many people from a former colony—people who can help migrants from the former colony find jobs and assistance in the new environment (Neumayer, 2005). Former colonial links consistently and significantly increased international migration in empirical studies (Pedersen et al., 2008; Mayda, 2010; Neumayer, 2005; Karemera et al., 2000).

Criticizing the simple view of the rural to urban migration, Todaro (1969) argued that migration in less developed countries would be two-stage phenomenon. In the first stage, unskilled rural worker moves to urban area and spends some time in urban traditional sector. In the second, the worker eventually enters the modern industrial sector. Todaro's theory not only suggests that expected earnings – product of earnings multiplied by employment probability - not the absolute wage differentials between rural and urban matter, but also implies that higher ratio of urban to rural population indicates greater economic development. Neumayer (2005) finds that people living in cities are likely to be better informed than rural inhabitants about international migration. Also, migrants go to cities in developing countries to get visas and documents for legal migration or make arrangements for illegal migration (Martin, 2003). Therefore, a higher percentage of an origin country's urban population is expected to become international migrants than the corresponding percentage of the origin's rural population. In a destination country, relatively large urban populations might indicate better job opportunities for newly arrived immigrants and a greater likelihood of getting help from people who came from the same origin. Furthermore, world system theory suggests that global cities in destination countries, such as New York, London, or Tokyo, concentrate wealth and a highly educated workforce and create strong demands for unskilled workers from overseas (Massey et al., 1993).

The age structure of a population may also affect international migration. There are ample empirical evidence suggesting that older persons are less mobile than younger ones (Stark and Bloom 1985). In addition, a high old-age dependency ratio (ODR), defined as the number of people aged 65 or over per persons aged 15 to 64, is closely linked to low fertility and indicates population ageing, and (depending on retirement ages and labor-force participation rates among the elderly) may indicate a shortage in the working-age population and a destination's economic demand for immigrants workers. Currently, most developed countries have a high ODR (Bongaarts 2004) and sometimes express a need for a larger percentage of working-age people. Hence, if all other conditions are equal, an origin with a low ODR would be expected to send more migrants to wealthy destinations than would an origin with a high ODR. Also, all other things being equal, a destination with a high ODR would be expected to attract more immigrants than a destination with a low ODR.

Infant mortality rate (IMR) and life expectancy at birth are demographic indices of quality of life for whole populations because factors affecting the health of an entire population have a significant impact on the mortality of infants (Reidpath and Allotey, 2003). Some scholars prefer the probability of dying before age five (${}_5q_0$) over the IMR not only because it represents cumulative mortality throughout early childhood owing to its relatively low sampling error and robustness to errors in reporting age at death (Cleland, Bicego, and Fegan 1992) but also because indirect estimation technique, i.e., Brass method, is well established (Hill and Pebley 1989). For less developed countries, IMR or life expectancy might be the only available measures of health or quality of life. Thus, ceteris paribus, an origin with a high IMR or a low life expectancy might be expected to send more emigrants to a destination than an origin with a low IMR or a high life expectancy. And, *ceteris paribus*, a destination having a high IMR would be expected to attract fewer immigrants than a destination having a low IMR.

## METHODS

## Data

The data for this article are drawn from various sources. The migration flows are taken from three sources: United Nations (UN), Eurostat, and Organization for Economic Cooperation and Development (OECD)[1]. More detailed information on the compositions of data from each source is provided in the appendix 1. When combining the flow data from three different sources, following rules were applied. First, if flow data for a pair of destination and origin in a year is available in only one source, then it is included in the sample. Second, if data on the same flow (i.e., the same destination and origin country in one year) is available at more than one source, then UN data were chosen first and Eurostat follows. The OECD data were considered as the last priority. The reason for this hierarchical ordering of the each source is two folds: 1) UN and Eurostat take the migration data from the reporting country and do not adjust whereas OECD makes some adjustments to harmonize the data. That is, there are inconsistencies in the reported number of migrants depending on which country reports (see e.g., DeWaard, Kim, and Raymer 2009). Hence, for the same set of destination and origin country in the same year, there is a greater consistency between UN and Eurostat compared to between UN and OECD or between Eurostat and OECD; and 2) because Eurostat provides migration statistics only for their member countries, UN data has greater number of reporting countries and observations than Eurostat. Greater number of observations may not necessarily be tantamount to greater accuracy of the data. Nonetheless, the one of the main goals of the present paper is to estimate international

[1] United Nations data is from International Migration Flows to and from Selected Countries: The 2008 Revision CD-ROM (United Nations 2010), Eurostat data is available at http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database, and OECD data is drawn from http://stats.oecd.org/Index.aspx?DataSetCode=MIG.

migrations by using reports from as many and diverse countries as possible, and therefore, UN data better conforms to the purpose.

Another major data source is the World Population Prospects (WPP): The 2008 Revision United Nations (2009). The WPP provided the total populations each year, the surface areas (in square kilometers), the old age dependency ratios, the life expectancy at birth, the infant mortality rates, the proportions of populations aged 15 to 24, and the proportions of the populations considered urban. From the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII, or the French Research Center in International Economics) came data on distances between geographical regions, official languages, and colonial relationships.

## Measures

The dependent variable in this study is logarithm of the number of migrants from an origin (*i*) to a destination (*j*) in year *t*, $m_{ijt}$. All logs refer to natural logarithms. I excluded migrant-related information involving geographical regions of multiple countries (e.g., African Commonwealth in United Kingdom data). Also, I excluded countries that, in the original data, lacked country codes.[2] For instance, the study excludes Taiwan because the United Nations recognizes the island as a province of China. Similarly countries, such as Czechoslovakia, the USSR, Yugoslavia, Serbia and Montenegro, and the German Democratic Republic, that no longer officially exist due to separation or unification are excluded. The term 'migrants' here refers to foreign-born people who obtained a residence permit or a work permit from the destination.

[2] This paper follows the standard country or area codes and geographical regions for statistical use provided by United Nations Statistics Division. See http://unstats.un.org/unsd/methods/m49/m49regin.htm.

Hence, for example, Australian citizens who had settled abroad and later moved back to Australia were excluded. In addition, some countries such as Germany maintain separate migration-registration systems for foreigners and citizens. I excluded all data for in- and out-migration of countries' own citizens. Although all the source assigns country codes for Hong Kong and Macao and provides separate migration flows for these areas, I treated their migrants as Chinese migrants. Similarly, reported flows for Puerto Rico and Guam are included at the United States.

Followings are the independent variables employed in the analysis. First, the total population of the destination and the population of the origin refer to the number of people, male and female combined, as of July 1 in each year.

Urbanization is the percentage of urban population, constructed by dividing the urban population in the given year by the total population of that year and multiplying by 100.

The old age dependency ratio (ODR) is the ratio of persons aged 65 and over to persons aged 15 to 64. The WPP furnishes only quinquennial estimates for the numerator and the denominator of ODR, and I linearly interpolated annual estimates by assigning one fifth of the 5-year change to each year.

The infant mortality rate (IMR) is the probability (between 0 and 1) that a live birth died before 1 year of age for boys and girls combined. IMR is a proxy for overall living conditions and well-being. The WPP provides only quinquennial IMR estimates for each country, and I linearly interpolated annual estimates. In the WPP, IMR is available only for countries with more than 100,000 inhabitants in 2009. As a result, IMR for small countries was not available and the number of observations of IMR was smaller than the numbers of observations of other demographic variables.

An official or national language is defined as a language spoken by at least 20% of the population of a country (Mayer and Zignago, 2006). If the destination and the origin had a common official language, the independent variable “common official language” is defined to equal 1; otherwise, the variable was 0.

Geographical distance is defined as the distance (in kilometers) between the two capital cities. Distances were calculated from the cities’ longitude and latitude using the great circle formula (Mayer and Zignago, 2006). A country is coded 1 if it is landlocked and, otherwise, 0. If two countries share a common border, the independent variable for having a common border is set to 1 and, otherwise, to 0.

When two countries have had a colonial or post-colonial relationship of colonizer to colonized for a relatively long period of time and when the (possibly former) colonizer substantially participated in the governance of the colonized country (Mayer and Zignago, 2006), the independent variable for colonial relations is set to 1 for colonial relations; and to 0 otherwise.

As Cohen et al. (2008) observed in different data, demographic variables might be so closely correlated one another that multicollinearity problem may arise. To address the concern, I calculated variance inflation factors (VIFs) for all the independent variables in the inflow model and the outflow model. The mean VIF for variables in the inflow model was 1.93, and none of the VIFs for each variable exceeded 5. In the outflow model, the mean VIF for variables was 1.90, and none of the VIFs for each variable was greater than 5. Therefore, multicollinearity seems unlikely to be a concern in this study.

Gravity Model

The gravity model, in its simplest form, views migration as determined by the sizes of the populations of destination and origin and the distance between origin and destination (Lewer and van den Berg 2008). The gravity model predicts that, all other things being equal, countries with large populations send more emigrants to destinations than countries with small populations, and that countries with large populations attract more immigrants. The greater the distance between origin and destination, the smaller the migration predicted. The gravity model can be easily extended by including other determinants (Lewer and van den Berg 2008), and the final model in the study is the following:

$$\begin{aligned}\log(m_{ijt}) = {} & \beta_0 + \beta_1 \log(Pop_{it}) + \beta_2 \log(Pop_{jt}) + \beta_3 \log(Dist_{ij}) + \\ & \beta_4 \log(IMR_{it}) + \beta_5 \log(IMR_{jt}) + \beta_6 \log(Urban_{it}) + \beta_7 \log(Urban_{jt}) + \\ & \beta_8 \log(ODR_{it}) + \beta_9 \log(ODR_{jt}) + \beta_{10}(Contig_{ij}) + \beta_{11}(OffLang_{ij}) + \\ & \beta_{12}(Colonial_{ij}) + \beta_{13}(LandLocked_i) + \beta_{14}(LandLocked_j) + \beta_{15}(\mathrm{SameRegion}_{ij}) + \\ & \beta_{16}(Year) + \varepsilon_{ijt}\end{aligned}$$

in which, $m_{ijt}$ represents the log of migration flows from source country $i$ to destination country $j$ at time $t$, $Pop_{it}$ is the population size of $i$, $Pop_{jt}$ is the population size of $j$, $Dist_{ij}$ is the distance between the capital city of $i$ and capital city of $j$, $IMR_{it}$ is the infant mortality rate of $i$, $IMR_{jt}$ is the infant mortality rate of $j$, $Urban_{it}$ is the percent of percent of urban population of $i$, $Urban_{jt}$ is the percent of percent of urban population of $j$, $ODR_{it}$ is the old age dependency ratio of $i$, $ODR_{jt}$ is the old age dependency ratio of $j$, $Contig_{ij}$ is a dummy variable for whether $i$ and $j$ have a common border, $OffLang_{ij}$ is a dummy variable for whether $i$ and $j$ have a common official language, $Colonial_{ij}$ is a dummy variable for whether $i$ and $j$ have ever been a colonial relationship, $Landlocked_i$ and $Landlocked_j$ are dummies for whether $i$ or $j$ is landlocked, and $SameRegion_{ij}$ is a dummy variable for whether $i$ and $j$ are in the same geographical region.

**RESULTS**

Justification for utilizing only non-economic measures in the model is provided in Figure 1, which illustrated examples of association between economic measures and selected demographic variables. The first panel presents the correlation between logged GDP per capita and logged IMR. It shows a clear negative correlation, and the regress of IMR on GDP per capita suggested that 1% increase in GDP is associated with 0.82% decrease in IMR ($p<0.001$) and 75% of the variations in the IMR can be explained by GDP per capita (adjusted $R^2 = 0.749$). The second panel shows the correlation between GDP per capita and the percent of urban population. The elasticity of urbanization with respect to GDP was 0.36 ($p<0.001$), and the about 55% of the variation in the urbanization can be accounted for by GDP per capita (adjusted $R^2 = 0.547$). Direct test of the Easterlin's (1978) argument on the link between the relative size of younger population to older one and unemployment is presented in the third panel. Although the explanatory power of the age structure for unemployment rate was not substantial (adjusted $R^2 = 0.058$), the coefficient showed an expected direction and was statistically significant ($\beta = -0.637$, $p<0.001$). Thus, Easterlin's thesis was correct in the US and abroad and is the case in more recent periods. The last panel depicts the relationship between level of average schooling and the IMR. The results implied that 1 year increase in mean years of schooling decreases the IMR by

33.1%[3] and 74.3% of the variation in the IMR can be explained by mean years of schooling (adjusted $R^2$ = 0.744). Substituting the IMR with the probability of dying before age five (${}_0q_5$) or the life expectancy at birth showed essentially the same results. Moreover, other covariates in the model were associated with socioeconomic indicators as theories suggest.

[FIGURE 1 ABOUT HERE]

Annual total inflow by destinations and annual total outflow by origins are presented in Figure 2a and Figure 2b. It is notable that there are substantial variations across the countries with respect to the number of immigrants and emigrants. Interestingly, popular destinations for immigrants in Europe – Germany, France, and UK – are receiving fewer people than in the past while immigrants to the new destinations such as Spain, Czech Republic, and Greece are growing in the most recent years. Also, immigration into destinations outside Europe and North America has been increasing over short period. For example, inflows to Japan and Turkey increased about three folds since the mid 1990s, and Republic of Korea is receiving more immigrants than many European countries since early 2000s. Immigration to traditional destinations, such as Australia, Canada, and US[4], is decreasing since 2005 perhaps due to worldwide economic recessions (Fix et al. 2009). Overall trends of immigration across countries clearly indicate the diversifications of destination countries since the mid 1990s (Castles and Miller 2003).

---

[3] Note the difference in the scale: the IMR was in logged form while the mean years of schooling was not. Thus, the regression coefficient, -0.3311, should be interpreted as a semi-elasticity of the IMR with respect to mean years of schooling (Wooldridge, 2006).

[4] The absurd spike in the trends for the US may be due to the Immigration Reform and Citizenship Act (IRCA) 1986, which granted permanent residence permits to qualified illegal immigrants. The UN data does not provide detailed information on how to treat them although those who obtained amnesty can not be considered as flows because they had been residing in the US for several years. For more information on this issue, see Kim and Cohen (Forthcoming).

[FIGURE 2A ABOUT HERE]

[FIGURE 2B ABOUT HERE]

Descriptive statistics are shown in the table 1.[5] The inflow data represented flows from 193 origin countries to 41 nations while the outflow represented from 33 origins to 192 destinations. The difference in the number of reporting countries between the inflow and outflow data results from the fact that not all countries maintain migration statistics for both directions of flows. For example, United States collects only immigration statistics (i.e., inflows) and the number of emigrants from US to other countries is not reported. Similarly, Bosnia and Herzegovina, Canada, France, Greece, Ireland, Israel, and Turkey do not have statistics on emigration.[6] The 41 destination countries received, on average, 1,362 immigrants, and the 33 source countries sent 798 emigrants. Both of the inflow and outflow indicated that there are substantial variations in the number of migrants across countries. Also, immigrants moved 6,768 km (4,205 miles) and emigrants moved (4,043 miles), on average. The infant mortality is considerably higher in origin countries, about six folds, compared to the destination countries in inflows. As expected, the percent of urban population is higher in destinations and old age dependency ratio is higher in destination countries. Only 3% of the destination and origin pairs share the borders. About 13% of the inflows and only 9% of the outflows occurred between the two countries that have the same official language, suggesting that emigrants from relatively developed nations to the rest of world are less constrained by cultural and psychic costs compared to the immigrants to rich countries.

[5] More detailed information on the data sources is provided in the appendix 2.

[6] Spain began to report their emigration since 2002. For more information, see the appendix 1.

[TABLE 1 ABOUT HERE]

The table 2 presents the results from various regression analyses of the determinants for inflow and outflows. Model 1 in the inflow and outflow model presents the results from ordinary least squares (OLS) regression. In order to relax the homoscedasticity assumption in linear regression, I used robust standard errors (sandwich estimator), which is more robust and preferable to model-based estimator (Rabe-Hesketh and Skrondal 2008). To account for the fact that the occasions (year in this paper) are nested in subjects (origin and destination pair), I also used cluster option in the OLS regression. The OLS results indicate that 1% increase in the population of destination increases the inflow by 0.9% while 1% increase in the population of origin increases the inflow by 0.67%.[7] On the other hand, 1% increase in destination and origin increases the outflow by 0.54% and 0.83%, respectively. Furthermore, 1% increase in the distance between the two capital cities is associated with 0.71% and 0.52% decrease in the inflows and outflows, respectively. These are consistent with that conventional gravity model for international migration suggests (Lewer and van den Berg 2008): the large population in origin and destination increases the inflow and outflow while the greater distance between the two decreases them. Interestingly, the standardized coefficients (i.e., beta) of the distance were -0.26 for inflow and -0.22 for outflow. This suggests, in contrast to what existing migration theories generally suggest, that emigration from rich countries to the rest of the world may not be constrained by economic costs, which are represented by distance, as much as immigration into developed countries. Infant mortality rate in both of destination and origin is negatively associated with inflow and outflow. The IMR is used for a proxy for general living conditions

[7] Wooldridge (2006) suggests that the coefficient in the log-log model is the elasticity of dependent variable with respect to the independent variable and interpreted as such.

(Sen 1993; 1998) and the level of economic development (Hauser 1959) in this paper, and it confirms what the neoclassical economic theory of migration (Massey et al. 1993). Another indicator of economic conditions, the percent of urban population, has significant effects on the international flows. In case of inflow, 1% increase of the percent of urban population in the destination is associated with 4.12% increase of inflows. Although the magnitude is much smaller than the percent of urban population in the destination country, urbanization in origin is also positively associated with inflows. Comparing the size of coefficients of urban population in the destination and origin between inflow and outflow yields interesting interpretation. That is, the effect of urbanization in destination is much larger than in origin with respect to inflows and the opposite is the case for the outflow. If we consider that the destinations in the inflow model are rich countries and the origin countries in the outflow model are rich countries, then it may illustrate that wealthy nations are drawing immigrants from lower income countries while the citizens in wealthy nations are emigrating to another affluent countries. Age structure, however, does not conform to the existing theories neatly. 1% increase in ODR in destination (i.e., larger proportion of population aged 65 and older compared to aged 14 to 64) decreases the inflow by 0.46% while 1% increase in ODR in origin increases inflow by 0.09%. The direction of the coefficients should be reversed if the existing literature is correct. That is, a population with a larger proportion of older population should draw more immigrants due to the lack of working age population, all other things being equal. Likewise, the sign of the ODR coefficient for origin should be a negative. The unexpected finding for ODR is also the case for the outflow. Given the fact that statistical significance of ODR between destination and origin differ, it may indicate either that ODR suffers from measurement error or that existing literature on the effect of age structure on migration flows might be incorrect.

Sociocultural determinants have significant effects on the inflow and outflow. Common official language increases inflows by about 91.8% and outflows by about 96%.[8] This result confirms the augmented gravity model, which suggests that common official language and cultural similarity is positively associated with migration flows (Karemera, Oguledo, and Davis 2000). Colonial relationship has even bigger effects on the international migration flows. If the destination and origin had ever been in a colonial relationship, inflows increase by about 120% and outflows increase by 140%. These results are consistent with world systems theory (Wallerstein 1974). Geographical determinants are also important predictors for international migration flows. Contiguity has a positive impact on inflows and landlocked location for origin has a negative impact on inflows. Being in the same region matters only for outflows, which is consistent with the fact that emigrants from the 33 countries moved shorter distance compared to the immigrants.

Models 2 and 3 in Table 2 report the estimated coefficients resulting from GEE estimation for inflow and outflow, respectively, specifying exchangeable and AR(1) as the correlation structure within panels, including demographic, geographic and social independent variables. The population-averaged generalized estimating equation (GEE) estimator permits time independent variables and allows the user to specify panels' within-group correlation structure (Liang and Zeger, 1986; Pedersen et al., 2008; Hardin and Hilbe, 2003). The GEE is equivalent to the random-effect estimator when the distribution of the dependent variable is Gaussian with an identity-link function and when the working correlation structure is

---

[8] The coefficient for a dichotomous variable in log-level model is called semi-elasticity of the dependent variable with respect to the independent variable and should be interpreted as β*100 % (Wooldridge 2006).

exchangeable, but GEE allows the user to adjust standard errors for clustering (Hardin and Hilbe, 2003; Horton and Lipsitz, 1999; Pedersen et al., 2008; Neumayer, 2005).

An advantage of GEE is the gain in efficiency in parameter estimation that results from including a hypothesized structure of the within-panel correlation (Hardin and Hilbe, 2003). Hypothetical correlation structures include independence, exchangeable, autoregressive, stationary, non-stationary, and unstructured (Cui, 2007; Hardin and Hilbe, 2003). The independence structure is equivalent to OLS models because it assumes observations within a panel are independent. The exchangeable correlation structure hypothesizes that observations within a panel have some common correlation. When the variance is Gaussian with an identity link, the exchangeable correlation GEE estimates are equal to random effects linear regression (Hardin and Hilbe, 2003). Autoregressive structure assumes a time-dependence for the association in the repeated observations within the panels. Stationary structure hypothesizes that correlation exists only for small number of time units, and the non-stationary is the same as stationary except that it does not assume constant correlations down the diagonals. Finally, unstructured correlation imposes no assumptions on the correlation matrix.

The results from GEE analysis are similar to the OLS results for most of the independent variables both in the inflow and outflow models. However, for some variables such as percent of urban population in the origin country for the outflow, signs as well as the size of the coefficients changed across models. Nonetheless, most of the variables are consistent across different models. The most consistent predictor is the distance between capital cities as Mayda (2010) and Kim and Cohen (forthcoming) find.

[TABLE 2 ABOUT HERE]

## DISCUSSION

This study investigated determinants of international migration flows on the basis of a large panel-data set and identified differences between inflows and outflows. However, caution should be exercised when interpreting the results.

First, the effects of policy changes on international migration flows are not accounted for in this study. Apparently, states and governments influence migration via their laws and regulations (Vogler and Rotte, 2000; Greenwood and McDowell, 1999), and several past empirical studies attempted to incorporate some form of policy measures (e.g., Ortega and Peri 2009; Mayda, 2010). However, data on this subject are sparse, and predictive models of policy do not seem to be available. Recently, Boussichas and Goujon (2010) proposed an indicator of openness of immigration policies using the OECD flow data and following methods. They fitted a regression of migration inflows on economic, geographic and cultural determinants. Then they derived predicted values from the model. Finally, they argued that the residuals of the regression represent the impact of the migration policy and can be used to build an indicator. Nonetheless, the independent variables in their regression model are limited to GDP per capita, distance, unemployment rate, and language. These are substantially insufficient to explain complex international migration. Moreover, their indicator is available only for OECD member countries and has a significant limitation of applicability. Hence, currently there are no reliable indicators of international migration policies.

Second, the present analysis is focused only on legal migration. Although United Nations, Eurostat, and OECD do not provide information on illegal or unauthorized migrants, illegal migration may be large and heterogeneous in size across countries. The data to overcome this limitation do not exist although there are some indirect estimation techniques for illegal immigrant flows or stocks (e.g., Jandl, 2004). Presumably illegal immigrants would be influenced by the determinants in our models but the dynamics of illegal flows is beyond the scope of this study.

Third, each country has its own definitions regarding international migrants. For example, Denmark considers a person who holds a residence permit or a work permit for at least 3 months to be a migrant whereas Finland defines a migrant as a person who has a residence permit and who intends to stay there for at least 1 year (DeWaard, Kim, and Raymer 2009). The United States and Canada use the place of birth to classify migrants whereas European countries use previous residence or citizenship (Cohen et al., 2008). Given the wide variations in defining migration and migrants, the numbers of migrants reported by the United Nations may include very different groups of people. Although I used the best available data, future research must take these problems into account to get more reliable estimates, and national statistical systems need to be harmonized to generate more comparable data (Poulain et al., 2006). Internationally harmonized time series estimates of migrant stocks by origin and destination are not presently available so migrant stocks are not considered in this analysis.

## CONCLUSION

This study examined determinants of international immigration to 41 wealthy nations—and international emigration from 33 of those 41 wealthy nations—between 1950 and 2008 with a panel-data approach. This study used only demographic, geographic, and social independent variables that are less time-sensitive and less uncertain than economic factors. The overall results were consistent with, amplify, and quantify existing migration theories. This study confirmed and extended Kim and Cohen (Forthcoming) and Cohen et al. (2008)'s suggestion that international migration can be effectively estimated by using time-invariant covariates and GLM methods.

Economic theories of international migration typically postulate that differences in economic factors such as income and employment drive international migration. If IMR can represent the general economic situation in a country and can be projected using demographic methods more accurately than economic factors such as income and employment, then we might be able to project international migration more reliably by incorporating IMR as a predictor.

When the annual inflows were classified by the income class of the origin and the destination (Table 3, left), between 1987 and 1991, about 39% of immigrants to high-income countries came from "lower middle-income" countries while about 17% of immigrants came from the low-income countries. During the same period, about 21% of immigrants to the high-income countries came from upper middle-income countries and about 24% came from high-income countries. Comparing the trends for the inflow throughout the periods under consideration shows an interesting pattern. Even though there might be fluctuations in the number of inflows by year, the proportion of the inflow by the origin and destination countries are quite stable. About a quarter of immigrants to the high-income countries came from another high-income country while about 40% came from lower-middle countries. Contrary to the conventional wisdom, the proportions of low income-countries are less than 20%. In case of upper-middle income destinations, they are receiving the half of the immigrants from lower middle-income countries. This finding provides empirical support for the theory of the "migration hump" (Martin 1993; Martin and Taylor 1996; de Haas 2007, 2010), which postulates that development and migration exhibit an inverted J- or U-shape pattern over time. When annual outflows from the high-income nations were classified by the income class of the destination (Table 3, right), about 50-60% of the migrants were heading to other wealthy nations while only 5-9% were heading to low-income countries. The outflows to high-income destinations were even more pronounced among the upper middle-income countries throughout the periods under consideration. Interestingly, majority of the outflows from lower middle-income countries are heading to other lower middle-income countries. In sum, owing to this complex relationship between the level of development and international migration, there might be counterintuitive results. These uncertain interpretations are post hoc and are offered to stimulate further empirical investigation.

[TABLE 3 ABOUT HERE]

My description of inflows and outflows by separately estimated models is equivalent to a unified model in which every independent variable of the original separate models interacts with an indicator variable that specifies whether each datum and each estimate are for inflows or outflows. Eventually such unified models would incorporate independent variables that describe why some flows are classified as inflows and other flows as outflows. Such a unified gravity-based model should make it possible to extrapolate from data on north-north, north-south, and south-north migration to south-south migration.

**REFERENCES**

Barro, R.J. and J.W. Lee. 2000. "International Data on Educational Attainment: Updates and Implications." CID Working Paper No. 42. Center for International Development. Harvard University, Cambridge, MA.

Bongaarts, J. 2004. "Population Aging and the Rising Cost of Public Pensions." *Population and Development Review* 30(1):1-23.

Boussichas, M. and M. Goujon. 2010. "A Quantitative Indicator of the Immigration Policy's Restrictiveness." *Economics Bulletin* 30(3): 1727-1736.

Castles, S. and M. Miller. 2003. *The Age of Migration*. Macmillan, Basingstoke.

Clark, X., T. J. Hatton, and J. G. Williamson. 2007. "Explaining U.S. immigration, 1971-1998." *The Review of Economics and Statistics* 89(2):359-373.

Cohen, J. E., M. Roig, D.C. Rueman, and C. GoGwilt. 2008. "International Migration Beyond Gravity: A Statistical Model for Use in Population Projections." *Proceedings of National Academy of Science* 105(40):15269-15274.

Cleland, J., G. Bicego, and G. Fegan. 1992. "Socioeconomic Inequalities in Childhood Mortality: The 1970s to The 1980s." *Health Transition Review* 2(1):1-18.

Cui, J. 2007. "QIC Program and Model Selection in GEE." *The Stata Journal* 7(2):209-220.

De Haas, H. 2007 ''Turning the Tide? Why Development Will Not Stop Migration.'' *Development and Change* 38(5):819–841.

-------. 2010. “Migration and Development: A Theoretical Perspective.” *International Migration Review* 44(1):227-264.

DeWaard, J., K. Kim, and J. Raymer. 2009. “International Migration Systems: Evidence from Harmonized Flow Data.” CDE Working Paper No. 2009-07. Center for Demography and Ecology. University of Wisconsin – Madison, Madison, WI.

Easterlin, R. A. 1978. “What Will 1984 Be Like? Socioeconomic Implications of Recent Twists in Age Structure.” *Demography* 15(4):397-432.

Fix, M., D. G. Papademetriou, J. Batalova, A. Terrazas, Y. Y. Lin, and M. Mittelstadt. 2009. “Migration and the Global Recession: A Report Commissioned by the BBC World Service.” Migration Policy Institute. Washington, D.C.

Greenwood, M. J. 1975. “Research on Internal Migration in the United States: A Survey.” *Journal of Economic Literature* 13(2): 397-433.

Greenwood, M. J. and J. M. McDowell. 1999. *Legal U.S. Immigration: Influences on Gender, Age, and Skill Composition*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

-------. 1982. “The Supply of Immigrants to the United States.” Pp. 54-85 In *The Gateway: U.S. Immigration Policies and Issues*, edited by B. R. Chiswick, Washington, DC: American Enterprise Institute.

Hardin, J. W., and J. M. Hilbe. 2003. *Generalized Estimating Equations*. Boca Raton, FL: Chapman & Hall/CRC.

Hauser, P. M. 1959. “Demographic Indicators of Economic Development.” *Economic Development and Cultural Change* 7(2):98-116.

Hill, K. and A. R. Pebley. 1989. “Child Mortality in the Developing World.” *Population and Development Review* 15(4):657-687.

Horton, N. J., and S. R. Lipsitz. 1999. “Review of Software to Fit Generalized Estimating Equation Regression Models.” *The American Statistician* 53:160-169.

Jandl, M. 2004. “The Estimation of Illegal Migration in Europe.” *Migration Studies* XLI(153): 141-155.

Jennissen, R. 2003. “Economic Determinants of Net International Migration in Western Europe.” *European Journal of Population* 19:171-198.

Karemera, D., V. I. Oguledo, and B. Davis. 2000. “A Gravity Model Analysis of International Migration to North America.” *Applied Economics* 32:1745-1755.

Kim, K. and J. E. Cohen. Forthcoming. “Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach Beyond Gravity.” *International Migration Review.*

Lee, E.S. 1966. “A Theory of Migration.” *Demography* 3(1):47-57.

Lewer, J. J. and H. Van den Berg. 2008. “A Gravity Model of Immigration.” *Economic Letters* 99:164-167.

Liang, K. Y., and S. L. Zeger. 1986. “Longitudinal Data Analysis Using Generalized Linear Models.” *Biometrika* 73(1):13-22.

Malthus, T. 1970[1798]. *An Essay on the Principle of Population*. New York: Penguin.

Martin, P. 1993. *Trade and Migration: NAFTA and Agriculture*. Institute for International Economics, Washington, D.C.

Martin, P. and J. E. Taylor. 1996. "The Anatomy of a Migration Hump" Pp. 43-62 in *Development Strategy, Employment, and Migration: Insights from Models*, edited by J. E. Taylor. Paris: OECD.

Martin, P. 2003. "Economic Integration and Migration." Discussion Paper No. 2003/35. United Nations University/World Institute for Development and Economic Research.

Massey, D. S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19(3):431-466.

Mayda, A. M. 2010. "International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows." *Journal of Population Economics* 23:1249-1274.

Mayer, T. and Zignago. 2006. "Notes on CEPII's Distance Measures." <http://www.cepii.fr/distance/noticedist_en.pdf>

Neumayer, E. 2005. "Bogus Refugees? The Determinants of Asylum Migration to Western Europe." *International Studies Quarterly* 49: 389-409.

Ortega, F. and G. Peri. 2009. "The Causes and Effects of International Labor Mobility: Evidence from OECD Countries 1980-2005." NBER Working Paper No. 14833. National Bureau of Economic Research, Cambridge, MA.

Palloni, A., D. S. Massey, M. Ceballos, K. Espinosa, and M. Spittel. 2001. "Social Capital and International Migration: A Test Using Information on Family Networks." *American Journal of Sociology* 106(5): 1262-1298.

Pedersen, P. J., M. Pytlikova, and N. Smith. 2008. "Selection and Network Effects - Migration Flows into OECD Countries 1990-2000." *European Economic Review* 52:1160-1186.

Piore, M. J. 1979. *Birds of Passage: Migrant Labor in Industrial Societies*. Cambridge: Cambridge University Press.

Poulain, M. et al. 2006. *THESIM – Towards Harmonised European Statistics on International Migration.* Louvainla-Neuve, Presses Universitaires de Louvain.

Rabe-Hesketh, S. and A. Skrondal. 2008. *Multilevel and Longitudinal Modeling Using Stata*. College Station, TX: Stata Press.

Ravenstein, E. G. 1885. "The Laws of Migration." *Journal of the Statistical Society of London* 48(2):167-235.

Reidpath, D. D., and P. Allotey. 2003. "Infant Mortality Rate as An Indicator of Population Health." *Journal of Epidemiology of Community Health* 57(5): 344-346.

Sen, A. 1993. "The Economics of Life and Death." *Scientific American* May:40-47.

-------. 1998. "Mortality as an Indicator of Economic Success and Failure." *The Economic Journal* 108(446):1-25.

Stark, O. and D. E. Bloom. 1985. "The New Economics of Labor Migration." *American Economic Review* 75(2): 173-178.

Stark, O. and J. E. Taylor. 1989. "Relative Deprivation and International Migration." *Demography* 26(1):1-14.

Todaro, M. P. 1969. "A Model of Labor Migration and Urban Unemployment in Less Developed Countries." *American Economic Review* 59(1): 138-148.

United Nations. 2009. *International Migration Flows to and from Selected Countries: The 2008 Revision*. United Nations, Department of Economic and Social Affairs, New York, NY.

-------. 2009. *World Population Prospects: The 2008 Revision*. United Nations, Department of Economic and Social Affairs, New York, NY.

Vogler M., and R. Rotte. 2000. "The Effects of Development on Migration: Theoretical Issues and New Empirical Evidence." *Journal of Population Economics* 13:485-508.

Wallerstein, I. 1974. *The Modern World System, Capitalist Agriculture and the Origins of the European World Economy in the Sixteenth Century*. New York: Academic Press.

Wooldridge, J. M. 2006. *Introductory Econometrics: A Modern Approach, Third Edition*. Cincinnati, OH: South Western College Publishing.

**Table 1. Descriptive Statistics for the Variables**

| | Inflow | | | | | Outflow | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | Mean | SD | Min | Max | N | Mean | SD | Min | Max |
| Migrants | 77477 | 1362 | 8160 | 1 | 946167 | 50667 | 798 | 4729 | 1 | 214811 |
| Population (Dest) | 77477 | 40869604 | 69411901 | 244123 | 311666000 | 50667 | 47210065 | 150292876 | 70542 | 1337411125 |
| Population (Origin) | 77477 | 42967657 | 142692296 | 49543 | 1337411125 | 50667 | 19582304 | 26469807 | 244123 | 127450906 |
| Distance | 77272 | 6768 | 4511 | 60 | 19586 | 50524 | 6508 | 4708 | 60 | 19586 |
| IMR (Dest) | 77477 | 0.007 | 0.004 | 0.003 | 0.040 | 50667 | 0.038 | 0.038 | 0.003 | 0.260 |
| IMR (Origin) | 77477 | 0.041 | 0.039 | 0.0029 | 0.26 | 50667 | 0.006 | 0.003 | 0.003 | 0.021 |
| Percent of Urban (Dest) | 77477 | 76 | 9.9 | 47 | 97 | 50667 | 56 | 23 | 0 | 100 |
| Percent of Urban (Origin) | 77472 | 54 | 24 | 0 | 100 | 50667 | 77 | 10 | 51 | 97 |
| Old Age Dependence Ratio (Dest) | 77477 | 21 | 4.10 | 7.30 | 32 | 50667 | 12 | 7.10 | 1.30 | 33.00 |
| Old Age Dependence Ratio (Origin) | 77477 | 11 | 6.90 | 1.30 | 33 | 50667 | 22 | 3.50 | 12.00 | 32.00 |
| Contiguity | 77272 | 0.03 | 0.16 | 0 | 1 | 50524 | 0.03 | 0.18 | 0 | 1 |
| Common Official Language | 77272 | 0.13 | 0.34 | 0 | 1 | 50524 | 0.09 | 0.29 | 0 | 1 |
| Colonial Relationship | 77272 | 0.04 | 0.19 | 0 | 1 | 50524 | 0.03 | 0.17 | 0 | 1 |
| Landlocked (Dest) | 77477 | 0.13 | 0.34 | 0 | 1 | 50187 | 0.17 | 0.37 | 0 | 1 |
| Landlocked (Origin) | 76799 | 0.18 | 0.38 | 0 | 1 | 50667 | 0.16 | 0.37 | 0 | 1 |
| Same Region | 77477 | 0.05 | 0.23 | 0 | 1 | 50667 | 0.07 | 0.25 | 0 | 1 |
| Year | 77477 | 1996 | 9.4 | 1970 | 2008 | 50667 | 1996 | 9.1 | 1970 | 2008 |

*Notes*: Dest means destination. Year is the calendar year.

**Table 2. Regression Estimates of Inflows and Outflows from OLS and GEE Models.**

| | Inflow | | | Outflow | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| | OLS | GEE (EXC) | GEE (AR1) | OLS | GEE (EXC) | GEE (AR1) |
| Population (Destination) | 0.902*** | 0.776*** | 0.871*** | 0.540*** | 0.319*** | 0.541*** |
| | (0.018) | (0.017) | (0.015) | (0.016) | (0.038) | (0.014) |
| Population (Origin) | 0.668*** | 0.618*** | 0.655*** | 0.832*** | 0.622*** | 0.740*** |
| | (0.014) | (0.021) | (0.012) | (0.024) | (0.029) | (0.021) |
| Distance | -0.713*** | -0.724*** | -0.727*** | -0.520*** | -0.434*** | -0.514*** |
| | (0.033) | (0.038) | (0.029) | (0.039) | (0.053) | (0.035) |
| IMR (Destination) | -1.017*** | -0.610*** | -1.013*** | -0.717*** | -0.357*** | -0.668*** |
| | (0.072) | (0.087) | (0.061) | (0.042) | (0.074) | (0.037) |
| IMR (Origin) | -0.437*** | 0.061 | -0.393*** | -0.685*** | -0.587*** | -0.914*** |
| | (0.038) | (0.062) | (0.035) | (0.101) | (0.140) | (0.083) |
| Percent of Urban Population (Destination) | 4.118*** | 2.308*** | 4.243*** | 0.161** | 0.086 | 0.135** |
| | (0.154) | (0.231) | (0.135) | (0.065) | (0.118) | (0.057) |
| Percent of Urban Population (Origin) | 0.261*** | 0.841*** | 0.278*** | 3.688*** | -0.959** | 3.541*** |
| | (0.059) | (0.090) | (0.051) | (0.181) | (0.385) | (0.156) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Old Age Dependency Ratio (Destination) | -0.462*** | 0.165 | -0.284*** | 0.034 | 0.469*** | 0.135** |
| | (0.137) | (0.160) | (0.109) | (0.065) | (0.127) | (0.058) |
| Old Age Dependency Ratio (Origin) | 0.091 | 0.140 | 0.243*** | 0.514*** | 0.129 | 0.535*** |
| | (0.059) | (0.096) | (0.056) | (0.193) | (0.165) | (0.156) |
| Contiguity | 0.333* | 0.626*** | 0.513*** | 0.200 | 0.459* | 0.394* |
| | (0.180) | (0.183) | (0.173) | (0.210) | (0.250) | (0.210) |
| Common Official Language | 0.918*** | 1.374*** | 1.064*** | 0.960*** | 1.437*** | 1.080*** |
| | (0.091) | (0.088) | (0.082) | (0.125) | (0.116) | (0.111) |
| Colonial Relationship | 1.191*** | 1.205*** | 1.251*** | 1.406*** | 1.399*** | 1.359*** |
| | (0.153) | (0.136) | (0.134) | (0.214) | (0.200) | (0.180) |
| Landlocked location (Destination) | 0.088 | -0.283*** | -0.008 | -0.123 | -0.254*** | -0.121* |
| | (0.059) | (0.056) | (0.054) | (0.077) | (0.087) | (0.066) |
| Landlocked location (Origin) | -0.332*** | -0.226*** | -0.261*** | 0.170** | -0.564*** | -0.097 |
| | (0.065) | (0.069) | (0.056) | (0.070) | (0.074) | (0.066) |
| Same Region | 0.156 | 0.285** | 0.094 | 0.426*** | 0.323* | 0.346** |
| | (0.130) | (0.125) | (0.119) | (0.164) | (0.174) | (0.155) |
| Constant | -37.891*** | -29.199*** | -38.835*** | -37.054*** | -10.176*** | -36.172*** |
| | (0.885) | (1.164) | (0.776) | (1.206) | (2.169) | (1.016) |
| Observations | 76,799 | 76,799 | 76,232[a] | 50,187 | 50,187 | 49,705[a] |

| | | |
|---|---|---|
| Adjusted R-squared | 0.620 | 0.596 |

*Notes*: Robust standard errors in parentheses. All regressions include dummy variables for year. Population, distance, IMR, percent of urban population, old age dependency ratio in the natural log form. Contiguity, common official language, colonial relationship, landlocked location, and same region are dichotomous variables.

[a]Cases that are available for one year are excluded because autocorrelation can not be computed.

*** p<0.01, ** p<0.05, * p<0.1

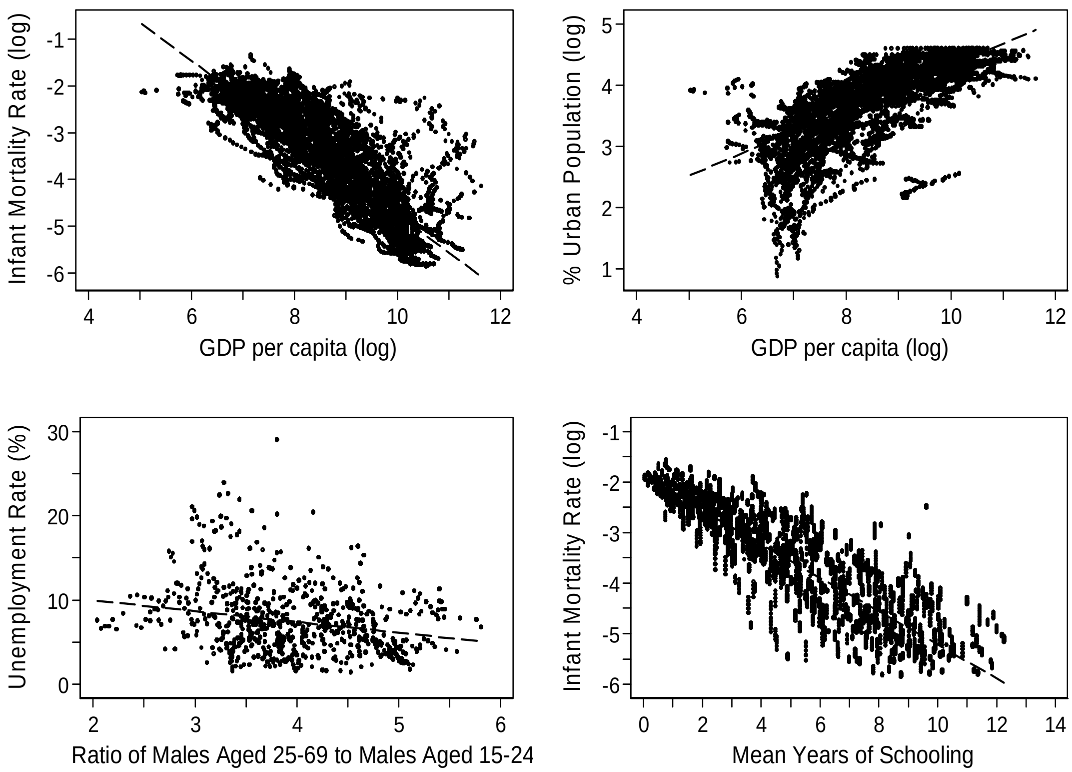
**Table 3. Percentage Distribution of Inflow and Outflow by Period and Income Levels of Origin and Destination Country, 1987-2008**.

| | Inflow | | | | | Outflow | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Origin Country's Income Level | | | | | Destination Country's Income Level | | | |
| Period | Destination's Income Level | Low | Lower Middle | Upper Middle | High | Origin's Income Level | Low | Lower Middle | Upper Middle | High |
| 1987-1991 | Lower Middle | 0.00 | 0.00 | 0.00 | 0.00 | Lower Middle | 0.00 | 0.00 | 0.00 | 0.00 |
| | Upper Middle | 10.51 | 41.55 | 7.03 | 40.92 | Upper Middle | 0.00 | 0.00 | 0.00 | 0.00 |
| | High | 16.53 | 38.78 | 20.72 | 23.97 | High | 5.72 | 27.27 | 9.00 | 58.01 |
| 1992-1996 | Lower Middle | 26.76 | 43.93 | 4.58 | 24.73 | Lower Middle | 7.86 | 76.04 | 1.10 | 15.01 |
| | Upper Middle | 45.51 | 25.34 | 7.79 | 21.36 | Upper Middle | 26.02 | 23.42 | 14.69 | 35.87 |
| | High | 22.01 | 38.48 | 13.11 | 26.40 | High | 8.96 | 29.42 | 12.08 | 49.54 |
| 1997-2001 | Lower Middle | 14.95 | 61.53 | 1.30 | 22.22 | Lower Middle | 2.40 | 52.27 | 1.06 | 44.28 |
| | Upper Middle | 17.19 | 56.66 | 3.83 | 22.31 | Upper Middle | 6.78 | 14.62 | 11.09 | 67.51 |
| | High | 15.96 | 37.43 | 18.10 | 28.52 | High | 7.31 | 19.85 | 18.05 | 54.79 |
| 2002-2008 | Lower Middle | 12.00 | 61.06 | 2.28 | 24.66 | Lower Middle | 0.00 | 36.14 | 38.55 | 25.30 |
| | Upper Middle | 7.41 | 43.41 | 20.72 | 28.46 | Upper Middle | 1.93 | 14.87 | 18.92 | 64.28 |
| | High | 12.62 | 40.56 | 22.23 | 24.59 | High | 4.85 | 25.88 | 17.92 | 51.34 |

*Source*: World Bank Historical Income Classifications.

*Notes*: Zero percent indicates that no data is available for the income level at the given period. Although the analysis sample is available from 1970, because income classifications are available since 1987, this table starts from 1987.

**Figure 1. Examples of the Associations among Socioeconomic Factors and Demographic Variables That are Commonly Used in the Studies of International Migration.**




*Sources*: Infant mortality rate and percent of urban population are from United Nations World Population Prospect (WPP) 2008 Revision. Unemployment rates are drawn from International Labor Organization (ILO) Key Indicators of Labor Market (KILM). Mean years of schooling is drawn from Barro and Lee (2000). Real GDP per capita (PPP adjusted in US Dollars) is taken from Penn World Table 6.3.

*Notes*: Dashed line is the fitted regression line. Log means the natural logarithm.

**Figure 2a. Total Annual Inflows (Immigration) by Destinations.**

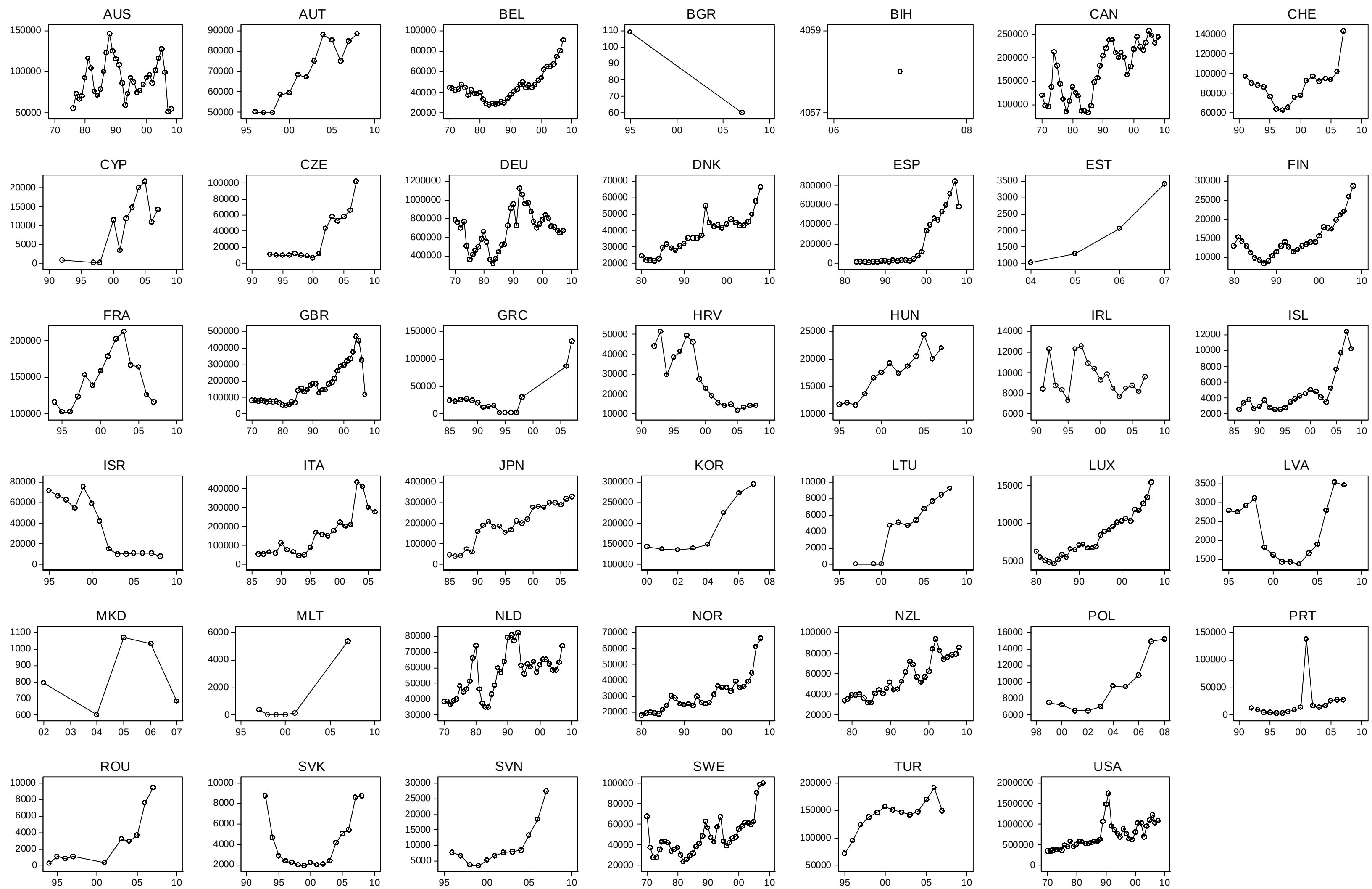
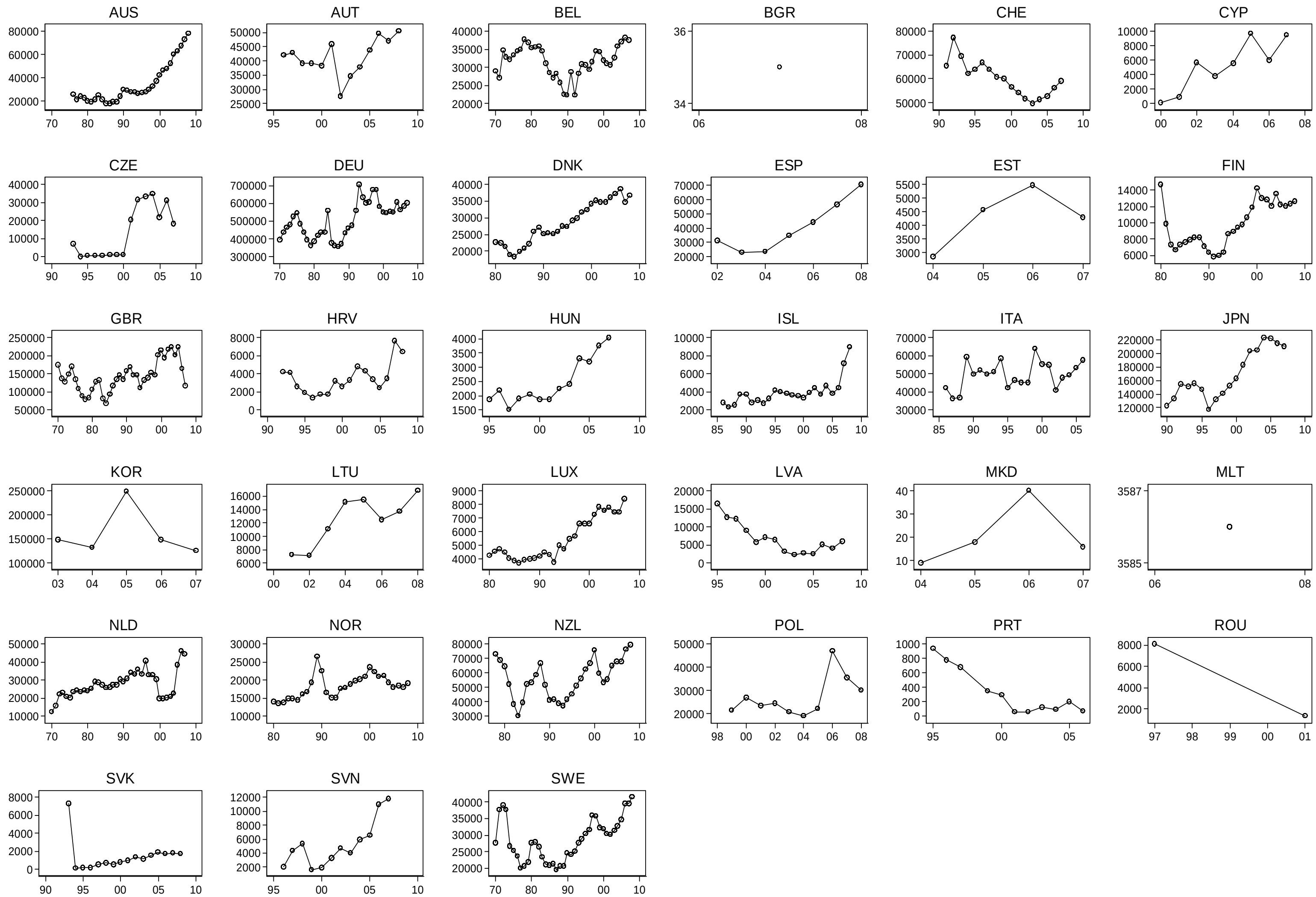


**Figure 2b. Total Annual Outflows (Emigration) by Origins.**



# Appendix 1. Composition of Inflows and Outflows by Reporting Country and Data Sources.

| | Inflow | | | | | | | | | Outflow | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | UN | | | OECD | | | Eurostat | | | UN | | | OECD | | | Eurostat | | |
| Reporting Country | N | Min. year | Max. year | N | Min. year | Max. year | N | Min. year | Max. year | N | Min. year | Max. year | N | Min. year | Max. year | N | Min. year | Max. year |
| Australia | 4,209 | 1976 | 2008 | 1 | 2007 | 2007 | | | | 3,453 | 1976 | 2008 | 1 | 2004 | 2004 | | | |
| Austria | 2,071 | 1996 | 2008 | | | | | | | 1,965 | 1996 | 2008 | | | | | | |
| Belgium | 1,510 | 1970 | 2007 | | | | 125 | 1999 | 2003 | 1,497 | 1970 | 2007 | | | | 44 | 2003 | 2003 |
| Bosnia and Herzegovina | | | | | | | 89 | 2007 | 2007 | | | | | | | | | |
| Bulgaria | | | | | | | 24 | 1995 | 2007 | | | | | | | 9 | 2007 | 2007 |
| Canada | 5,891 | 1970 | 2008 | | | | | | | | | | | | | | | |
| Croatia | 245 | 1992 | 2008 | | | | 139 | 2002 | 2007 | 231 | 1992 | 2008 | | | | 20 | 2003 | 2007 |
| Cyprus | | | | | | | 440 | 1992 | 2007 | | | | | | | 231 | 2000 | 2007 |
| Czech Republic | 1,081 | 1993 | 2007 | | | | 4 | 1997 | 2007 | 837 | 1993 | 2007 | | | | 5 | 2002 | 2007 |
| Denmark | 4,400 | 1980 | 2008 | | | | | | | 4,212 | 1980 | 2008 | | | | | | |
| Estonia | 91 | 2004 | 2007 | | | | | | | 91 | 2004 | 2007 | | | | | | |
| Finland | 3,412 | 1980 | 2008 | | | | | | | 2,718 | 1980 | 2008 | | | | | | |
| France | 1,432 | 1994 | 2007 | 296 | 1995 | 2007 | 66 | 2003 | 2006 | | | | | | | | | |
| Germany | 4,771 | 1970 | 2007 | | | | | | | 4,770 | 1970 | 2007 | | | | | | |
| Greece | | | | 31 | 1998 | 1998 | 1,314 | 1985 | 2007 | | | | | | | | | |
| Hungary | 423 | 1995 | 2007 | 2 | 2006 | 2007 | 596 | 1997 | 2007 | 409 | 1995 | 2007 | 1 | 2007 | 2007 | 347 | 1997 | 2007 |
| Iceland | 1,885 | 1986 | 2008 | | | | | | | 1,460 | 1986 | 2008 | | | | | | |
| Ireland | | | | 8 | 1994 | 1998 | 40 | 1991 | 2007 | | | | | | | | | |
| Israel | 551 | 1995 | 2008 | | | | | | | | | | | | | | | |
| Italy | 2,294 | 1986 | 2006 | | | | | | | 1,962 | 1986 | 2006 | | | | | | |
| Japan | | | | 516 | 1985 | 2007 | | | | | | | 518 | 1990 | 2007 | | | |
| Latvia | 830 | 1995 | 2008 | | | | 2 | 1999 | 2007 | 805 | 1995 | 2008 | | | | | | |
| Lithuania | 547 | 2001 | 2008 | | | | 14 | 1997 | 2007 | 575 | 2001 | 2008 | | | | 3 | 2003 | 2006 |
| Luxembourg | 224 | 1980 | 2007 | 32 | 1996 | 2007 | 1,279 | 1994 | 2007 | 224 | 1980 | 2007 | 90 | 1996 | 2007 | 1,001 | 1994 | 2007 |
| Malta | | | | | | | 92 | 1997 | 2007 | | | | | | | 66 | 2007 | 2007 |
| Netherlands | 4,148 | 1970 | 2007 | | | | | | | 3,972 | 1970 | 2007 | | | | | | |
| New Zealand | 4,249 | 1978 | 2008 | | | | | | | 3,930 | 1978 | 2008 | | | | | | |
| Norway | 4,138 | 1980 | 2008 | | | | | | | 3,805 | 1980 | 2008 | | | | | | |
| Poland | 1,027 | 1999 | 2008 | | | | | | | 756 | 1999 | 2008 | | | | | | |
| Portugal | | | | 162 | 1992 | 2007 | 664 | 1992 | 2006 | | | | 149 | 1995 | 2006 | | | |
| Republic of Korea | | | | 97 | 2000 | 2007 | | | | | | | 122 | 2003 | 2007 | | | |
| Romania | | | | | | | 406 | 1994 | 2007 | | | | | | | 16 | 1997 | 2001 |
| Slovakia | 1,307 | 1993 | 2008 | | | | | | | 562 | 1993 | 2008 | | | | | | |
| Slovenia | 283 | 1996 | 2007 | | | | 568 | 1996 | 2007 | 282 | 1996 | 2007 | | | | 467 | 1997 | 2007 |
| Spain | 1,424 | 1983 | 2008 | | | | 677 | 1992 | 2007 | 554 | 2002 | 2008 | | | | 163 | 2002 | 2007 |
| Sweden | 5,110 | 1970 | 2008 | | | | | | | 4,612 | 1970 | 2008 | | | | | | |
| Switzerland | 2,831 | 1991 | 2007 | 1 | 2007 | 2007 | | | | 2,743 | 1991 | 2007 | | | | | | |
| TFYR Macedonia | | | | | | | 248 | 2002 | 2007 | | | | | | | 43 | 2004 | 2007 |
| Turkey | | | | 1,375 | 1995 | 2007 | 153 | 2005 | 2005 | | | | | | | | | |
| United Kingdom | 224 | 1970 | 2007 | 405 | 1991 | 2000 | 748 | 1985 | 2006 | 224 | 1970 | 2007 | 293 | 1991 | 2000 | 429 | 1985 | 2006 |
| United States of America | 6,255 | 1970 | 2008 | | | | | | | | | | | | | | | |
| Total | 66,863 | | | 2,926 | | | 7,688 | | | 46,649 | | | 1,174 | | | 2,844 | | |

Note: Blank cells indicate that data are not available.
N = number of observations; Min. year = Minimum year; Max. year = Maximum year.
Overlappings of years within a reporting country indicate that different origin or destination countries are avaiable in different data sources.

**Appendix 2. Descriptions and Sources of Variables in the Analysis.**

| Variables | Description | Source |
|---|---|---|
| Population | Total annual population (both sexes combined) | World Population Prospect 2008 Reivision (United Nations 2009) |
| Distance | Distance between capital cities in kilometer | CEPII (Centre D'Etudes Prospectives et D'Informations Internationales): the Research Center in International Economics |
| Infant Mortality Rate | Infant deaths per 1,000 live births | World Population Prospect 2008 Reivision (United Nations 2009) |
| Percent of Urban Population | Percent of population living in urban area | United Nations Population Division Demobase |
| Old Age Dependency Ratio | Ratio of population 65+ per 100 population 15-64 | World Population Prospect 2008 Reivision (United Nations 2009) |
| Contiguity | 1 if origin and destination share the border, 0 otherwise | CEPII (Centre D'Etudes Prospectives et D'Informations Internationales): the Research Center in International Economics |
| Common Official Language | 1 if ogirin and destination use the same official primary language | CEPII (Centre D'Etudes Prospectives et D'Informations Internationales): the Research Center in International Economics |
| Colonial Relationship | 1 if the pair had been ever in colonial relationship | CEPII (Centre D'Etudes Prospectives et D'Informations Internationales): the Research Center in International Economics |
| Landlocked Location | 1 if a country is landlocked | CEPII (Centre D'Etudes Prospectives et D'Informations Internationales): the Research Center in International Economics |
| Same Region | 1 if the pair is the same geographical region defined by UN | World Population Prospect 2008 Reivision (United Nations 2009) |