Journal of International Economics and Finance
Vol. 1, Issue 2, 2021: pp. 79-96

INTERNATIONAL MIGRATION IN OECD COUNTRIES: PANEL ECONOMETRIC ANALYSIS OF THE DETERMINANTS

T. Lakshmanasamy*
**ICSSR Senior Fellow and Formerly Professor, Department of Econometrics, University of Madras, Chennai.*
E-mail: tlsamy@yahoo.co.in

Article History
Received : 30 October 2021
Revised : 15 November 2021
Accepted : 8 December 2021
Published : 30 December 2021

To cite this paper
T. Lakshmanasamy (2021). International Migration in OECD Countries: Panel Econometric Analysis of the Determinants. *Journal of International Economics and Finance.* 1(2), 81-98.

Abstract: In the globalised world economy, international migration is essential for the individual as well as national skill development and income growth. The causes and consequences of sizable migration among the OECD countries, which enjoy reasonable growth rates with high standards of living and low unemployment rates, are to be understood in the changing global demographic and economic scenario. This paper estimates the effects of push and pull factors of migration in 11 OECD countries for 15 years from 2002 to 2016 applying the panel data regression methods and specification tests for model fitness. The estimated results show that GDP is a significant pull factor and the unemployment rate is a strong push factor of migration. While high GDP in the origin country discourages migration, high GDP in the destination country has a positive effect on migration. The unemployment rate in the origin country directly influences migration while the same in the destination country is inversely related to the number of out-migrants in the home country. The distance between two countries affects migration negatively, and the origin and destination population size influences the number of outgoing migrants directly.

Keywords: OECD, migration, GDP, unemployment, determinants, panel regression

Introduction

The migration of labour is a key driver in economic development as well as individual development and is one of the high-priority issues for both developing and developed countries. Human migration is defined as the movement by people (as individuals, family units or large groups) from one place to another with the intention of settling, permanently or temporarily in a new location. Demographers classify migration generally into two types: (i) internal migration - a change of residence within national boundaries, such as between states, provinces, cities, or municipalities or from rural to rural, rural to urban, urban to rural and urban to urban localities, and (ii) international migration - a change of residence over national boundaries, moving to another and stay in the host country for some minimum length of time. Weinstein and Pillai (2001) refer to the third type of migration - forced migration when a person is moved against wish (slaves) or when the move is initiated because of external factors (natural disaster or civil war), as a refugee or displaced. Both international and internal migration, permanent, temporary or forced, have everlasting and significant demographic and cultural impacts on the economies and societies of the chosen destination and the place of origin.

There are several reasons and motives, push and pull factors, that cause migration. International migrants leave their origin countries in search of economic opportunities in another country or to rejoin family members who have migrated earlier for jobs or some political issues. Among the individual factors, education is the single most reason for international migration, as students for studies abroad or employment. Other push factors which impel people to relocate are poor medical care, unavailability of jobs, fewer opportunities, primitive conditions, political fear and religious discrimination, natural disasters and lower odds of finding courtship. The pull factors may include higher chances of employment, better living standards, education, health facilities, security, family ties and marriage, lower crime rate, etc.

The migrants do not move to any one place or a country, they are dispersed across regions and countries around the world, with most having moved from middle-income to high-income regions or countries. In 2015, 244 million people, or 3.3 percent of the world's population lived outside their country of origin. According to the UN 2015 Report, the total number of emigrants from Germany has been 4.045 million, followed by the U.S (3.023 million), Italy (2.9 million), Portugal (2.306 million) and France (2.145 million). Among destination countries, the US records the largest number of international migrants (46.6 million), the other top destination countries being Germany (12.0 million) and United Kingdom (8.5 million). The recent Pew Research Center research on Origins and Destinations of World's migrants highlights the differences in the number of outgoing migrants from a country. From the Tables presented below, it is evident that for most of the selected origin countries, the migrant flows have increased over the years (Table 1), their population have increased between 2001 and 2015, and the unemployment rates have decreased till 2007 and then increased up to 2013 and started declining thereafter. Figure 1 portrays that the GDP of all 11 select countries has an upward linear trend over the span of 2000 to 2017.

Table 1: Total Number of Migrants in 11 OECD Countries (in 000s)

Country	*2000*	*2010*	*2017*	*Country*	*2000*	*2010*	*2017*
Australia	380	480	540	Portugal	2000	1940	2270
Canada	1150	1270	1360	Spain	1300	1100	1350
France	1530	1950	2210	Sweden	260	300	350
Germany	3350	3850	4210	UK	3870	4460	4920
Italy	3120	2610	3030	US	1990	2650	3020
New Zealand	490	660	830				

Table 2: Trend in Population in 11 OECD Countries (in 000s)

Country	*2000*	*2005*	*2010*	*2015*	*Country*	*2000*	*2005*	*2010*	*2015*
Australia	19269	20239	22120	23800	Portugal	10399	10566	10652	10418
Canada	31025	32288	34169	35950	Spain	41392	44043	46789	46398
France	59911	61234	63027	64457	Sweden	8898	9039	9390	9764
Germany	81536	81671	80895	81708	UK	59149	60287	63307	65397
Italy	57506	58808	59730	59504	US	284852	295130	308641	319929
New Zealand	3908	4135	4370	4615					

Table 3: Trend in Unemployment Rates in 11 OECD Countries (percent)

Country	*2000*	*2007*	*2013*	*2015*	*Country*	*2000*	*2007*	*2013*	*2015*
Australia	6.7	4.4	5.7	6.1	Portugal	3.8	8.0	16.3	12.4
Canada	7.2	6.0	7.1	6.9	Spain	10.4	8.2	26.1	22.1
France	8.6	7.7	9.9	10.4	Sweden	4.7	6.2	8.1	7.4
Germany	7.8	8.7	5.2	4.6	UK	4.7	5.3	7.5	5.3
Italy	9.6	6.1	12.2	11.9	US	4.7	4.6	7.4	5.3
New Zealand	5.4	3.7	6.3	5.4					

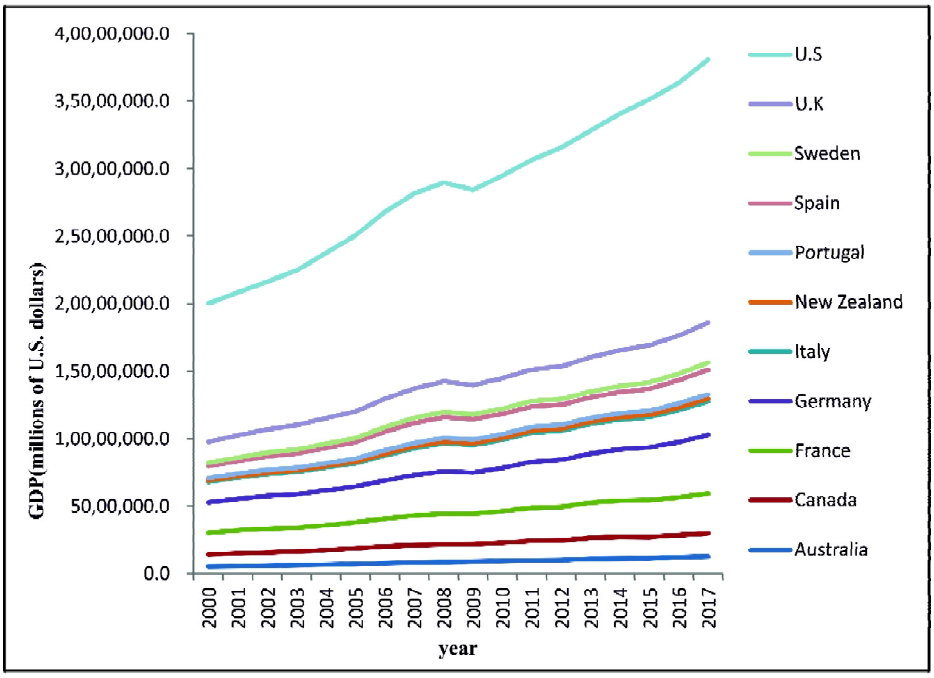

Figure 1: Trend in GDP in 11 OECD Countries (US$ million)

Given that these 11 select OECD countries enjoy positive growth rates and high standards of living, as well as low unemployment rates, a moot question is why so many people of these countries migrate to other countries. This paper tries to identify the determinants of international migration and to measure the impact of these determinants on international migration of these select 11 OECD countries, namely Australia, Canada, France, Germany, Italy, New Zealand, Portugal, Spain, Sweden, United Kingdom and the United States. The data for this study is for 15 years from 2002 to 2016 collected from various sources. With 15 years of data for 11 countries, the total number of observations is 1650. Empirically, this paper follows the panel data regression methods. In this paper, the number of migrants is the response variable. The determinants of international migration are time invariant factors such as the distance between the countries and the explanatory variables such as population, population, unemployment rate, gross domestic product of origin and destination countries. The empirical analysis distinguishes between the role played by the same explanatory variable in the sending region (push factor) and in the destination region (pull factor). The estimates of panel techniques such as pooled ordinary least squares, fixed effects between-group, fixed effects within-group, fixed effects least squares dummy variable method and random effects generalised least squares are compared. Thereafter, Hausman and Breusch-Pagan tests are applied to identify the better fit of the estimation of the determinants of international migration in developed OECD countries.

Review of Literature

The earliest model of migration, the gravity model, is derived from Newton's law of gravity, which states that "any two bodies attract one another with a force that is proportional to the product of their masses and inversely proportional to the square of the distance between them". In the case of migration, the gravity model is utilised to predict the degree of migration interaction between two places. When used geographically, the words 'bodies' and 'masses' are replaced by 'locations' and 'importance' respectively, where importance can be measured in terms of exogenous (economic and non-economic) variables like population, gross domestic product or other appropriate variables. Thus, the gravity model of migration implies that as the importance of one or both of the locations increases, there will also be an increase in movement between them. However, the farther apart the two locations are, the less movement between them, a phenomenon known as distance decay. Greenwood (1985) argues that the omission of distance or spatial structure, in general, may seriously affect empirical studies on migration because the distance is a proxy that captures all costs that cannot be measured but surely affect migration flows.

Ravenstein (1885) is the earliest migration theory based on the observation of migration patterns in Great Britain, and later the United States. He proposed that although most migrants travel short distances, longer-distance migrants prefer to go to centres of commerce or industry. Each stream of migration produces a counter-stream at the destination. Large towns owe more of their growth to migration than to natural development and the volume of migration increases with the development of industry and commerce and as transportation improves; most migration is from agricultural areas to urban centres and the main cause of migration is economic. These observations motivated an abundance of quantitative models of migration flows and the aggregate variables that affect those flows.

The Ravenstein and Newton type gravity models posit that migration between origin place *i* and destination place *j*, m_{ij}, is a positive function of repulsive forces or push factors at *i* (R_i) and attractive forces or pull factors at *j* (A_j) and is inversely related to the 'friction' or distance between *i* and *j* (D_{ij}):

$$m_{ij} = f(R_i, A_j)/g(D_{ij}) \quad (1)$$

In practice, most formulations of the gravity model simply assume that migration between *i* and *j* is directly proportional to the product of the two places' populations and inversely proportional to the intervening distance:

$$m_{ij} = (P_i P_j)/D_{ij} \quad (2)$$

Population and distance represent the standard gravity variables and population size of both origin and destination countries affect positively migration while distance, a proxy to costs of migration, discourages migration.

Stouffer (1940) argues that there is no necessary relationship between mobility and distance. Stouffer extends the Ravenstein gravity model by introducing the notion of intervening opportunities. He proposed that migration over a given distance is held to be directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities or possible alternative migration destinations between i and j. In other words, the number of persons going a given distance is directly proportional to the percentage increase in opportunities at that distance. In this approach, the nature of a particular place may be more important than the distance in determining where the migrant goes, as generally migrant's base migration decisions on the characteristics, social and economic conditions of the origin and destination.

Lee (1966) summarise the factors which enter into the decision to migrate and the process of migration as: (i) factors associated with the area of origin, (ii) factors associated with the area of destination, (iii) intervening obstacles, and (iv) personal factors. Lee hypotheses that both the destination and the origin have characteristics that attract or repel migrants and that perceptions of these characteristics differ between migrants. Lee notes that not only the actual factors at origin and destination result in migration but also the perception of these factors leads to migration. Thus, the pull and push factors of migration depend on the evaluation of the situation at origin and knowledge of the situation at the destination.

Lowry (1966) introduce wages and unemployment rates so as to assess the role played by the economic variables. The basic form of the Lowry model is:

$$m_{ij} = g\left(\frac{u_i}{u_j}, \frac{w_i}{w_j}, \frac{L_i L_j}{D_{ij}}\right) e_{ij} \tag{3}$$

where m_{ij} measures the flows of migrants from region i to region j, g is a gravitational constant, u is the unemployment rate, w is the hourly wage in the manufacturing sector, L is the labour force, D is the (airline) distance between the two regions and e is the error term. In this formulation, the unemployment rate and the wage level play two different roles. An increase (decrease) in the value of one of these variables in the region of origin or destination relative to the value of the same variable in the destination or origin region can discourage (encourage) migration. That is, they may act as push factors when their value encourages people to leave their country of origin or they may act as pull factors in the destination region when they attract people from other regions.

Mukhopadhyay (1980) compare the aggregate 1961 and 1971 Indian censuses data on birthplace migration streams to assess the quantitative magnitudes of various rates of migration for both genders between rural and urban sectors. The analysis shows that in general rural migrants are deterred by distance (as per gravitational theories of migration), whereas urban migrants have a higher propensity to travel a long distance.

Anjomani (2002) analyse US interstate migration by groups of determinants: (i) previous gross migration as a proxy for social networks or availability of information, (ii) economic variables like regional income, employment rate, unemployment rate, local income tax, (iii) amenity variables such as population density, mean temperature, welfare benefits, criminality rate, and (iv) demographic variables population size or growth, mean educational level, median population age.

van Wissen *et al*. (2008) classify the determinants of migration into 6 types of exogenous variables: (i) gravity variables like population size, with positive influence, and physical distance, with negative influence; (ii) economic variables like gross domestic product per capita, newly created businesses, wages, etc, which attract migration, (iii) labour market variables like levels and/or rates of employment and unemployment, changes in working conditions, etc., (iv) real estate variables or housing market variables, possibly high prices of houses and low vacancy rates deterring migration unless anticipated by potential migration, (v) environment variables such as terrain conditions, population density, degree of urbanisation, social behaviour of local inhabitants, climatic conditions, leisure and entertainment activities, which affect quality of life both on short and long-term, and (vi) policy variables like governmental subsidies, local taxes, defense spending, educational offer, urban area plan, or direct measures such as migration incentives and policies also act as a pull factor for migration.

van Leuvensteijn and Parikh (2002) study interregional migration in 16 regions of Germany in order to identify the factors involved, using panel data and allying the least square dummy variable (LSDV) and generalised method of moments (GMM) estimations. The variables include differential unemployment rate, wage differential for blue-collar and white-collar workers, differentials in hospital and hotel beds per inhabitant, differential in per capita rented or owned housing, differential in rental price per square km, distance between the main cities and differential in cost of living index.

Kumar and Aggarwal (2003) examine migration in Assam by pooling of the 1901 to 1991 Indian census data for 10 major states from where people migrated to Assam. The census data provides total migration from other states and other countries along with the distribution of Assam population by language, sex composition, workers versus non-workers, rural-urban migration and district-wise migration in Assam. They also consider the unemployment rate at the destination sector and per capita income and urbanisation rates at the destination and origin regions. The study shows that distance and urbanisation are very significant factors for domestic migration, while prospects of higher earnings are the big attraction for foreign migrants, and for female migration, family-related reasons are the main cause.

Andrienko and Guriev (2004) study the determinants of internal migration in Russia using panel data on gross region-to-region migration flows during 1992-1999 by estimating the effects of economic, political and social factors. The study shows that people relocate from poorer and job scarce regions with worse public good provision to ones that are richer and more prospering both in terms of employment prospects and public goods.

van Wissen *et al.* (2008) analyse out-migration and in-migration in Sweden, Netherlands and the United Kingdom, and in Spain to validate the models. For both out-migration and in-migration, the explanatory variables used in the study are population size, employment, regional mass, unemployment, gross domestic product, accessibility and population density. Internal migration has been modeled as a two-stage process with out-migration followed by destination choice, conditional on out-migration. The internal migration model is estimated as a Poisson regression model for the period 1991-1995 and using the estimated model coefficients, the model is validated by predicting the flows for 1996-1998.

Ghatak, Mulhern, and Watson (2008) analyse interregional migration in Poland using the seemingly unrelated regression equations (SURE) method. The variables considered are GDP per capita, unemployment rate, number of dwellings per 1,000 inhabitants, number of students per 1,000 inhabitants, road distance between capital cities, the density of road length, and rate of infant mortality. Though the Polish regional migration is low by international standards, which is a feature of transitional economies, the migration pattern followed economic patterns being influenced by relative regional economic opportunities and costs. The analysis shows that asymmetry in unemployment and GDP per capita in the origin and destination provinces have a strong effect on internal migration. The GDP per capita in the donor province had a stronger influence than in the destination province. Unemployment has a stronger impact on migration in the destination rather than donor provinces Further, provincial migration has been influenced by distance, regional facilities like road infrastructure, human capital and housing.

Etzo (2011) investigate the determinants of interregional migration flows in Italy after the upsurge of 1996, after two decades of decreasing internal migration rates. The fixed effects vector decomposition estimator (FEVD) and the generalised method of moments (GMM) are applied to a gravity model on bilateral migration flows for the period 1996-2005 using population size, distance between main cities, GDP per capita, unemployment rate, infrastructure index and crime rate in both origin and destination regions. The study shows that GDP per capita and unemployment rate are the key determinants whose changes push migrants out from their regions and direct them to prosperous destinations. The distance is also an important determinant and that migrants respond differently to the same variable in the region of origin rather than in the destination one.

Bunea (2012) examine the potential determinants of internal migration in Romania using county data for the period 2004-2008 applying panel data methodologies. From a static point of view, the fixed effects least square dummy variable regression (LSDV) estimates show significant impacts of population size, real gross product per capita, amenity index, road density and crime rate, and from a dynamic point of view, the two step generalised method of moments (GMM) estimates show significant effects of previous migration ratio, population size and amenity index.

Data and Methodology

This paper uses panel data on international migration between 11 OECD countries for 15 years from 2002 to 2016, a total of 1650 observations. The countries considered are Australia, Canada, France, Germany, Italy, New Zealand, Portugal, Spain, Sweden, United Kingdom and the United States. In the empirical analysis, the number of migrants is the response variable and the independent variables are population, unemployment rate, gross domestic product of origin and destination countries, which act as pull and push factors of migration. In the panel context, the time invariant variable is the distance between the two countries.

The data on the dependent variable is assembled from the OECD International Migration database, which gives the number of migrants from country i to country j in thousand units. The distance between the two countries is taken from the distance table of the distance calculator on the website globefeed.com. The country-wise unemployment rate data is extracted from the ILOSTAT website, where the variable is defined as the number of unemployed persons as a percentage of the total number of employed and unemployed persons (labour force). The statistics on the gross domestic product is collected from the IMF World Economic Outlook database, which is measured in US$ million. The yearly data on population (in thousand units) is derived from the United Nations Population Division database.

In the empirical formulation of the gravity model (Lowry 1966), the number of people moving from region i to region j, mij depends positively on the population size in each region Pi, Pj and negatively on the distance between the two regions Dij. The basic gravity model is specified as:

$$m_{ij} = g\frac{P_i^{\alpha}P_j^{\beta}}{D_{ij}^{\gamma}} \tag{4}$$

where g is a constant. Generalising to include all the possible push and pull factors, an extended version of the gravity model is specified as:

$$m_{ij} = g^{\alpha_0}\frac{P_i^{\alpha_1}P_j^{\alpha_2}}{D_{ij}^{\alpha_3}}\prod_{s=1}^{n}\frac{x_{sj}^{\gamma}}{y_{si}^{\delta}} \tag{5}$$

where y_{si} includes all the possible exogenous variables in the origin region i that acts as push factors of migration while X_{sj} includes all the exogenous variables that can attract (pull) migrants to the destination j (Andrienko and Guriev, 2004).

In logarithmic form, the linear specification of the gravity model is expressed as:

$$lnm_{ij} = \alpha_0 lng + \alpha_1 lnP_i + \alpha_2 lnP_j + \alpha_3 lnD_{ij} + \sum_{s=1}^{n}\left(\gamma_s lnx_{sj} - \delta_s lnx_{si}\right) \tag{6}$$

The bilateral migration flows together with distance and the identification of push and pull factors make the logarithmic equation a spatial interaction model.

In the econometric estimation, the panel data model is specified as:

$$y_{it} = \alpha + \beta x_{it} + {}_i + \gamma_t + u_{it} \tag{7}$$

where γ_t is the solely time-dependent variables that do not vary across cross-sections (individual heterogeneity), i is the time invariant solely cross-section dependent variables (country fixed effects), and uit the idiosyncratic error term which satisfies the usual ordinary least square assumptions.

The pooled OLS regression of panel estimation is specified as:

$$y_{it} = \alpha + \beta x_{it} + {}_i + u_{it} \tag{8}$$

The pooled regression of y_{it} on x_{it} using all the data together assumes that there is no correlation across individuals or across time periods for any individual. If λ_i contains only a constant term, then estimation by OLS provides consistent and efficient estimates of the intercept α and the slope vector β.

The fixed effects within-group regression of panel estimation method removes the effect of those time-invariant characteristics, which are assumed to be unique to the individual and fixed over time so that the net effect of the predictors on the outcome variable can be obtained. The fixed effects within-group panel regression model is specified as:

$$y_{it} = \beta x_{it} + \alpha_i\, i + u_{it} \tag{9}$$

Taking deviations from the group means:

$$(y_{it} - \bar{y}_i) = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i) \tag{10}$$

Then, the regression parameters are estimated using the within-group averages as:

$$\hat{\beta}_W = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it}-\bar{x}_i)(y_{it}-\bar{y}_i)}{\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it}-\bar{x}_i)^2} \tag{11}$$

The fixed effects between group panel regression of estimation method use the variation of the group means around the overall mean:

$$(\bar{y}_i - \bar{y}) = \beta(\bar{x}_i - \bar{x}) + (\bar{u}_i - \bar{u}) \tag{12}$$

The parameters of the fixed effects between group panel regression is estimated as:

$$\hat{\beta}_B = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_i-\bar{x})(\bar{y}_i-\bar{y})}{\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_i-\bar{x})^2} \tag{13}$$

The fixed effects least squares dummy variable (LSDV) panel regression estimation method specifies a set of dummy variables for each cross-section:

$$y_{it} = \beta x_{it} + d_i + u_{it} \tag{14}$$

The random effects generalised least squares panel regression estimation method assumes that the variations across observations are random, and hence the time invariant variables can be included in the estimating equation. The random effects panel regression model is specified as:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} \tag{15}$$

where $\varepsilon_{it} = (\lambda_i + u_{it})$ is the composite error term. In the estimation, as there is serial correlation in errors i.e $Cov(u_{it}, u_{is}) = \sigma^2 \neq 0$, the OLS is inefficient. Therefore, to obtain consistent estimates, the generalised least squares (GLS) method of estimation is to be used.

The pooling of cross section and time series in a panel regression can be tested with the F-test:

H_0: fixed effects intercepts are zero i.e. pooled OLS is applicable.

H_1: fixed effects intercepts are non-zero i.e. fixed effects method is applicable.

The Breusch Pagan Lagrangian Multiplier test for random effects model is based on the OLS residuals. The null hypothesis in the LM test is that variance of errors from regression is dependent on the values of the independent variable i.e. presence of heteroscedasticity or there is no significant difference across units i.e. no panel effect, which means that pooled OLS can be applicable.

H_0:$\sigma_u^2 = 0$ or correlation between ε_{it} and $\varepsilon_{is} = 0$ or pooled OLS method is applicable.

H_1:$\sigma_u^2 \neq 0$ or random effects method is applicable.

LM is distributed as a chi-squared distribution.

The Hausman specification test is used to choose between fixed or random effects. The Hausman test compares the random and fixed effects estimators which should be approximately the same if the zero-correlation OLS assumption [$Cov(\lambda_i, X_{it}) = 0$] holds, but different if the assumption is false:

H_0: the preferred model is random effects

H_1: the fixed effect model.

The Hausman test is specified as:

$$H = (\hat{\beta}_{RE} - \hat{\beta}_{FE})^{\Omega-1}(\hat{\beta}_{RE} - \hat{\beta}_{FE}) \sim X_k^2 \quad (16)$$

where Ω^{-1} is the variance-covariance matrix and the statistic is distributed as chi-square.

Empirical Analysis

The determinants of migration considered in this paper are previous year (one year lag) population size, unemployment rates and gross domestic products for both origin and destination countries. The time-invariant factor affecting migration is the distance between the two countries. Some country-specific variables are language, temperature, etc. which may be termed as unobserved heterogeneity and controlled by the panel. With 11 countries and 15 years, for each country, the migration to the remaining 10 countries are considered. Table 4 presents the definition and Table 5 presents the descriptive statistics of variables used in the empirical analysis. The United Kingdom has the highest average number of outgoing migrants in thousands (10459.09), whereas Sweden has the lowest (865.46). The migrants from New Zealand travel the highest average distance (16116.68kms), while for German migrants, the average distance travelled is the lowest (5502.041kms). The highest mean recorded for the previous year origin population size is 302942.763 thousand for the United States, while New Zealand has the lowest average population size (4272.320 thousand) over the years. In the case of the unemployment rate of the previous year, Spain has the highest average (15.853), whereas New Zealand has the lowest mean (5.247). The highest mean for the previous year origin gross domestic product is US$14327509 million for the United States, whereas New Zealand has the lowest mean (US$127375.025 million).

Table 4: Definition of Variables in Migration Analysis in 11 OECD Countries

Variable	*Definition*
Migrants	Number of migrants in the current time period (t) leaving out of their origin or birth country i to go to their destination countries of choice j (thousands of persons).
Distance	Distance between countries is the straight line distance (flying or air distance) between the two countries based on their centre latitudes and longitudes (kilometres).
OPop	Total number of people residing in the origin country (thousands).
DPop	Total number of people residing in the destination country (thousands).
OUnemp	Origin country unemployment rate - ratio of number of unemployed workers to total labor force.
DUnemp	Destination country unemployment rate - ratio of number of unemployed workers to total labour force.
OGDP	Origin country gross domestic product at constant prices (US$ millions).
DGDP	Destination country gross domestic product at constant prices (US$ millions).

Table 5: Descriptive Statistics of Variables in 11 OECD Countries

Country	*Migrants*	*Distance*	*OGDP*	*DGDP*	*OUnemp*	*DUnemp*	*OPop*	*DPop*
Australia	3055.285 (5735.95)	13841.05 (3311.628)	855536.4 (185017.1)	2712457 (4062779)	5.45 (0.69)	8.28 (3.92)	21405.1 (1475.6)	67038.04 (82847.79)
Canada	2978.56 (4711.28)	7676.9 (3303.60)	1290574 (214976.1)	2668953 (4080032)	7.13 (0.65)	8.11 (4.0)	33432.57 (1559.56)	65835.29 (83415.94)
France	5085.94 (5208.14)	5553.04 (6369.77)	2196240 (337918.3)	2578386 (4098827)	8.85 (0.86)	7.94 (4.0)	62263.72 (1447.49)	62952.17 (83989.87)
Germany	5146.22 (4479.74)	5502.04 (6033.41)	3064541 (501258)	2491556 (4095118)	7.68 (1.11)	8.05 (3.96)	81370.67 (294.74)	61041.48 (83765.86)
Italy	6522.43 (10142.63)	5799.47 (6037.60)	1954614 (223026.1)	2602549 (4096556)	9.0 (2.0)	7.92 (3.96)	59085.16 (734.51)	63270.03 (83981.51)
New Zealand	3887.48 (9153.23)	16116.68 (4696.63)	127375 (26578.41)	2785273 (4022259)	5.25 (1.08)	8.30 (3.91)	4272.32 (215.37)	68751.31 (81694.87)
Portugal	2688.09 (4315.04)	6012.24 (6419.6)	262392.7 (33122.96)	2771771 (4030975)	9.51 (3.79)	7.87 (3.8)	10549.58 (86.48)	68123.59 (82165.6)
Spain	3901.49 (6298.31)	5910.10 (6492.31)	1360869 (221568)	2661923 (4082320)	15.85 (6.45)	7.23 (2.28)	45079.39 (1868.13)	64670.61 (83779.92)
Sweden	865.46 (862.78)	5713.0 (5392.0)	366119.1 (66378.44)	2761398 (4037282)	7.02 (1.2)	8.12 (3.99)	9271.61 (284.15)	68251.38 (82074.25)
UK	10459.09 (9452.79)	5513.86 (6019.77)	2174332 (320951.7)	2580577 (4098751)	6.02 (1.028)	8.22 (3.94)	62112.57 (2139.44)	62967.29 (83988.25)
US	5660.63 (6041.25)	7671.42 (3954.07)	14327509 (2275331)	1365259 (964572.9)	6.47 (1.7)	8.17 (3.99)	302942.8 (11036.85)	38884.27 (25710.3)

Note: Standard deviation in parentheses.

Table 6 presents the inter-correlations among the variables of migration analysis. The highest correlation is realised between the logarithm of previous year origin population size and the logarithm of previous year origin gross domestic product (0.988) and the logarithm of previous year destination population size and the logarithm of previous year destination gross domestic product (0.988). As such the contemporaneous population size and GDP in the origin or the destination are highly related to each other. This may be because higher the population size in a country, higher the number of employees, hence higher will be the cumulative income of the working population which leads to a high gross domestic product in that country. However, the other variables are showing a lesser degree of inter-relatedness among themselves. There is a weak correlation between

Table 6: Pairwise Correlation between Variables in 11 Countries

Variable	*lnMigrants*	*lnDistance*	*lnOPop*	*lnDPop*	*lnOunemp*	*lnDunemp*	*lnOGDP*
lnDistance	-0.452	1	-	-	-	-	-
lnOPop	0.435	-0.206	1	-	-	-	-
lnDpop	0.394	-0.188	-0.099	1	-	-	-
lnOunemp	0.074	-0.265	0.153	-0.007	1	-	-
lnDunemp	-0.020	-0.274	-0.007	0.152	0.050	1	-
lnOGDP	0.428	-0.175	0.988	-0.095	0.120	0.025	1
lnDGDP	0.404	-0.155	-0.095	0.988	0.025	0.120	-0.081

unemployment rates in origin and destination places. The distance is consistently negatively correlated with all other variables.

In the econometric estimation of the determinants of migration, the first task is to identify the appropriate method of estimation, and the second task is to measure the effect of the factors that influence migration. There are various methods of estimating a panel data model, the pooled OLS regression, fixed effects within-group, between-group and least squares dummy variable, and random effects generalised least square methods. Empirically, the estimating extended gravity model is expressed as:

$$lnm_{ijt} = \beta_0 + \beta_1 lnDis_{ij} + \beta_2 lnOGDP_{it-1} + \beta_3 lnDGDP_{jt-1} + \beta_4 lnOUnemp_{it-1} + \beta_5 lnDUnemp_{jt-1} + \beta_6 lnOPop_{it-1} + \beta_7 lnDPop_{jt-1} + \lambda_i + u_{ijt} \quad (17)$$

where the prefixes O and D stand for origin and destination countries respectively. The least squares dummy variables (LSDV) is specified as:

$$lnm_{ijt} = \beta_0 + \beta_1 lnDis_{ij} + \beta_2 lnOGDP_{it-1} + \beta_3 lnDGDP_{jt-1} + \beta_4 lnOUnemp_{it-1} + \beta_5 lnDUnemp_{jt-1} + \beta_6 lnOPop_{it-1} + \beta_7 lnDPop_{jt-1} + \beta_8 d2003 + \beta_9 d2004 + \cdots + \beta_{21} d2016 + {}_i + u_{ijt} \quad (18)$$

Including the time dummies for the years 2003 to 2016 respectively. The application of fixed effects and random effects methods lies with the question of where to include the time-invariant and individual heterogeneity factors before running the regression.

The empirical results of the panel data estimates are presented in Table 4. The coefficient estimates of distance of migration from pooled OLS, fixed effects between-group and random effects GLS methods have a statistically significant negative effect on migration at 1 percent level. The estimates of these three methods indicate that for a one percentage increase in distance between two countries, there will be 0.52 to 0.76 percent decrease in the number of outgoing migrants respectively. Thus, it is evident that the distance between two countries affects the migration negatively, the greater the distance, the lesser will be the number of migrants willing to leave their home country. In the fixed effects within-group and least square dummy variable estimates, the coefficient of distance variable has been omitted due to collinearity problem.

The gross domestic product of origin countries discourages migration, whereas the GDP of destination countries encourage migration significantly. The coefficient estimates are significant at a one percent level in fixed effects within-group, LSDV and random effects methods of estimation. With a growing economy, people have more opportunities, and hence no need to migrate. The discouragement effect of home GDP ranges from 1 to 1.6 percent. The random effects GLS estimate for the coefficient of destination GDP shows the maximum effect on the migration. It has a positive and relatively elastic value of 2.48 and thus implies that with a one percent increase in the destination GDP, the number of migrants will increase by about 2.5 percentage. This shows that migrants prefer to go to a country with better economic growth and living standards acting as a pull factor for migration.

Table 6: Pooled OLS, Fixed Effects LSDV, Within-Group and Between-Group, and Random Effects GLS Estimates of Migration in 11 OECD Countries

Variable	*Pooled OLS*	*Fixed effects within-group*	*Fixed effects between-group*	*Fixed effects LSDV*	*Random effects GLS*
lnDist	-0.633*** (20.26)	-	-0.761*** (6.36)	-	-0.515*** (4.97)
lnOGDP	0.075 (0.37)	-1.209*** (39.34)	-1.615 (1.53)	-0.937*** (29.81)	-1.602*** (7.53)
lnDGDP	1.482*** (7.33)	2.445*** (78.53)	0.011 (0.01)	2.338*** (70.75)	2.479*** (11.65)
lnOUemp	-0.540*** (5.82)	0.245*** (49.62)	-2.038*** (3.81)	0.269*** (54.72)	0.234*** (6.38)
lnDUemp	-0.693*** (7.46)	-0.115*** (23.3)	-1.913*** (3.57)	-0.227*** (39.41)	-0.108*** (2.93)
lnOPop	0.536** (2.40)	-0.427*** (5.54)	2.411** (2.07)	-0.661*** (8.74)	2.213*** (8.61)
lnDPop	-1.026*** (4.59)	-1.819*** (27.19)	0.598 (0.51)	-1.986*** (23.14)	-2.080*** (8.09)
Constant	-1.592** (2.16)	13.435*** (16.88)	12.597** (2.35)	15.462*** (16.92)	-2.303 (1.21)
σ_u	-	3.148	-	3.014	1.134
σ_ε	-	0.319	-	0.315	0.319
ρ	-	0.990	-	0.989	0.926
R-square	0.514	0.034	0.3523	0.0512	0.439
F-value/ χ^2	248.91	6476.57	21.23	2916.36	763.18

Note: Absolute t-values for pooled OLS, fixed effects within-group, between-group and LSDV regressions, and z-values for random effects GLS regression are in parentheses. ****, ** and * Significant at 1, 5 and 10 percent levels respectively.

The unemployment rate of the previous year at the origin country encourages migration significantly according to fixed effects within-group, LSDV and random effects estimates whereas pooled OLS and fixed effects between group estimates indicate discouragement in the migration. Fixed effects between-group estimate is more elastic and other estimates are inelastic. From the results, it can be inferred that the origin unemployment rate acts as a push factor of migration. This is because a high unemployment rate origin country means many people are unable to find any work and will have the incentive to migrate to countries which are having higher options for employment. Thus, the origin unemployment rate directly influences migration from the home country. In contrast, the coefficient estimates of the unemployment rate in the previous year of the destination countries across all methods of estimation are negative and statistically significant at 1 percent level. The coefficient estimates show that for a one percent increase in the unemployment rate in the destination country, the outward migration from the home country will decrease by 0.10 to 1.9 percent. Thus, it is evident that the unemployment rate in the destination country discourages people to leave their origin country as high unemployment rates in the destination country means long spells of search and wait and less probability of finding a suitable job there.

The pooled OLS and fixed effects between-group and random effects coefficient estimates of the previous year origin population are significantly positive, whereas fixed effects within-group and LSDV estimates are significantly positive. A one percent increase in the population of the origin country increases migration by 2.2 to 2.4 percent, acting as a push factor for migration from the origin country. This may be because a large origin population implies lower chances of employment or receiving better opportunities. Compared to the positive effects, the reduction in migration due to the home country population increase is rather less, 0.42 to 0.66 percent only. The coefficient estimates of the population size of destination countries from all the methods (except insignificant positive fixed effects between-group estimates) are negative and significant at a 1 percent level. The estimates indicate that a one percent increase in the size of the destination population reduces the migration by 1 to 2. Thus, the higher the population in the destination country, the lower will be the number of people willing to move from their birth or place i.e. a populated country does not attract migration.

The variance of within-group residuals (σ_u) of the random effects GLS estimation (1.13) is much lower than the fixed effects within-group variance (3.13) showing that much of the variations in international migration in the 11 OECD countries are due to time effects rather than country effects. This is also seen from the constant variance of overall error residuals (σ_ε) across estimation methods. The intraclass correlation (?) between 0.92 and 0.99 by the panel estimates show that 92.6 to 99 percent of the variations in migration is due to differences across the panels. The R2 values are in the range of 0.03 of fixed effects within-group estimation to 0.51 of pooled regression and fixed effects LSDV estimations, with 0.44 of random effects GLS estimation. The F-values, indicating whether all the coefficients in the model are different from zero, show that the model is correctly specified for pooled OLS and all three fixed effects within-group, between-group and LSDV estimation methods, and by the Wald ?2 value for the random effects GLS model. All the test statistics have p-values less than 0.05 showing the significance of the fitted model for estimation.

Given the five estimation methodologies of the migration model, the appropriate estimation method is to be identified. The F-test value for pooled OLS vs fixed effects within-group estimation is 6476.57, which strongly rejects the null hypothesis that all the coefficients in the model excluding the constant are zero and hence the pooled OLS is not applicable. The Breusch Pagan LM test for pooled OLS vs random effects GLS estimation reject the null hypothesis of homoscedasticity and hence the pooled OLS method is not applicable. Since the alternative hypothesis of heteroscedasticity is to be accepted, the random effects GLS method may be a suitable method of estimation for the model. The Breusch Pagan LM test model is specified as $lnm_{it} = \beta x_{it} + u_i + \varepsilon_{it}$, for which the estimated results are shown in Table 7.

The Hausman specification test value between the fixed effects least square estimates and the random effects generalised least squares estimates, presented in Table 8, strongly

Table 7: Breusch-Pagan LM Test for Random Effects Estimation
Dependent variable: ln (migration)

Variable	*Variance*	*Std. dev.*
x	3.087	1.757
u	1.287	1.134
ε	0.102	0.319
Test: var(u) = 0	χ^2 (01) = 8990.61	Prob > χ^2 = 0.00

rejects the null hypothesis that the difference in coefficients is not systematic, and hence the random effects model is not applicable. Though the results of the Hausman test rule is in favour of the fixed effects method, the R^2 value for LSDV is 0.0512 or 5.12 percent is very low (Table 6).

Table 8: Hausman Specification Test for Fixed Effects and Random Effects Estimation

Variable	*LSDV* (β_{FE})	*RE (GLS)* (β_{RE})	*Difference* (β_{FE}- β_{RE})	$\sqrt{diag(_{FE} - _{RE})}$std. error
lnOGDP	-1.340	-1.602	0.262	0.110
lnDGDP	2.568	2.479	0.089	0.131
lnOUemp	0.258	0.234	0.024	0.002
lnDUemp	-0.181	-0.108	-0.073	0.023
lnOPop	0.0408	2.214	-2.173	0.425
lnDPop	-1.545	-2.080	0.536	0.571

Test: Ho: difference in coefficients not systematic

β_{FE} consistent under H_0 and H_1 — β_{RE} inconsistent under H_1 efficient under H_o

$\chi^2 = (\beta_{FE} - \beta_{RE})' (V_{FE} - R_{RE})^{-1} (\beta_{FE} - \beta_{RE}) = 58.54$ — Prob > χ^2 =0.00

From the empirical analysis, it is evident that the random effects GLS estimates for all independent variables are statistically significant. Also, the random effects method has a better explanatory power (R^2=43.89 percent) than those of the fixed effects techniques, Though pooled or LSDV estimations have relatively higher power than the random effects estimator, they do not take into account the differences between cross-sections or over time. The pooled OLS assumes that all the intercept terms for all cross-sections are homogenous, whereas country-wise regression should have different intercept terms for different countries. Therefore, it is reasonable to conclude that the random effects (GLS) estimation of the panel of migration data may be a better estimation strategy as also the random effects model produced the estimates that are in line with the assumptions about how each exogenous variable influences migration.

Conclusion

International migration in OECD countries is sizable and a host of push and pull factors influence migration. This paper tries to identify the factors associated with migration between 11 select OECD countries over 15 years from 2002 to 2016, a panel data of 1650 observations. The exogenous variables considered are the GDP, unemployment rate, population and distance of both origin and destination countries, which are lagged by one year since the migrants would want to decide on the choice of leaving their origin country based on prior knowledge of the conditions in the sending and receiving countries. An extended gravity model has been estimated using panel data estimation methods of pooled ordinary least squares, fixed effects within-group, between-group and least squares dummy variable, and random effects generalised least squares estimations. A host of tests, the F-test, chi-square test, Breusch Pagan LM test and Hausman specification test are applied for model fitness.

In the pooled OLS method, the estimates for distance, origin and destination population sizes, origin and destination unemployment rates and destination gross domestic product are statistically significant, whereas the estimate for origin gross domestic product is not significant. In the fixed effects within-group estimation method, distance not included, all coefficient estimates are statistically significant. In the fixed effects between-group method, the estimates for distance, origin population size and unemployment rate, destination unemployment rate are statistically significant while the other estimates are not. In the fixed effects LSDV method, all estimates of the regression coefficients are statistically significant, the distance being omitted. In the random effects GLS estimation, all coefficients are statistically significant. Based on the tests, the random effects GLS estimation method is identified as a suitable estimation technique for migration in OECD countries.

From the estimated results, this paper has observed that the distance between two countries affects migration negatively. The greater the distance, the lesser will be the number of people willing to migrate from their home country. The GDP, being a measure of economic growth, is a strong pull factor of migration since people migrate due to the urge of residing in a country with better economic growth and living facilities. The empirical results in this paper show that high GDP in the origin country discourages migration, while high GDP in the destination country has a positive effect on migration. The unemployment rate in the origin country directly influences migration while the same in the destination country is inversely related to the number of out-migrants in the home country. This is because the high unemployment rate in any country means that job prospects are less, so the migrants will have the incentive to migrate to countries with higher options for employment. The unemployment rate acts as a push factor for migration. The origin and destination population sizes influence the number of outgoing migrants directly. If the rate of increase in the origin population is higher than that of the destination country, the chances of migration are higher.

References

Andrienko, Y. and S. Guriev (2004). "Determinants of Interregional Mobility in Russia: *Economics of Transition*, 12, 1, 1-27.

Anjomani, A. (2002). "Regional Growth and Interstate Migration", *Socio-Economic Planning Sciences*, 36, 4, 239-265.

Bunea, D. (2012). "Modern Gravity Models of Internal Migration: The Case of Romania", *Theoretical and Applied Economics*, 18, 4, 127-144.

Etzo, I. (2011). "The Determinants of the Recent Interregional Migration Flows in Italy: A Panel Data Analysis", *Journal of Regional Science*, 51, 5, 948-966.

Ghatak, S., A. Mulhern and J. Watson (2008). "Inter-Regional Migration in Transition Economies: The Case of Poland", *Review of Development Economics*, 12, 1, 209-222.

Greenwood, M.J. (1985). "Human Migration: Theory, Models, and Empirical Studies", *Journal of Regional Science*, 25, 4, 521-544.

Kumar, P. and S.C. Aggarwal (2003). "An Econometric Analysis of Causes of Migration in Assam", *Demography India*, 32, 1, 107-122.

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

Lowry, I.S. (1966). *Migration and Metropolitan Growth: Two Analytical Models*, San Francisco: Chandler Publishing Company.

Mukhopadhyay, S. (1980). "Birthplace Migration in India: 1961, 1971", *Indian Economic Review*, 15, 3, 185-200.

Poot, J., O. Alimi, M. Cameron and D.C. Mare (2016). "The Gravity Model of Migration: The Successful Comeback of an Ageing Superstar in Regional Science", IZA Discussion Paper No. 10329, Bonn.

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

Stouffer, S.A. (1940). "Intervening Opportunities: A Theory Relating Mobility and Distance", *American Sociological Review*, 5, 6, 845-867.

van Leuvensteijn, M. and A. Parikh (2002). "How Different Are the Determinants of Population versus Labour Migration in Germany", *Applied Economics Letters*, 9, 11, 699-703.

van Wissen, L., L. van der Gaag, P. Rees and J. Stillwell (2008). "In Search of a Modelling Strategy for Projecting Internal Migration in European Countries: Demographic versus Explanatory Approaches", in J. Poot, B. Waldorf and L. Wissen (eds.): *Migration and Human Capital: Regional and Global Perspectives*. Cheltenham: Edward Elgar, 49-74.

Weinstein, J.A. and V.K. Pillai (2001). *Demography: The Science of Population*, Boston, USA: Allyn and Bacon.