CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Migration and Development: A Meta-Analysis

Master's thesis

Author: BSc. Miroslava María Palecek Rodríguez

Study program: Economics and Finance

Supervisor: doc. PhDr. Tomáš Havránek, Ph.D.

Year of defense: 2020

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, July 29, 2020

Miroslava María Palecek Rodríguez

Abstract

The current literature on international migration is diverse, and there is an ongoing debate as to the size and magnitude of the development-migration nexus, and no consensus about this effect has been reached. In this thesis, I explore quantitatively the effect of GDP (as a measure of development) on migration using a meta-analysis approach by synthesizing the empirical findings on this effect, adjusting for the biases, and controlling for the design of the studies. To examine the phenomenon in a systematic way, I collected 179 regression coefficients from 40 different articles, where the results suggest a weak presence of publication selection. Nevertheless, when correcting for publication bias, the effect of development on migration is rather small. Additionally, to explain the inherent model uncertainty, the Bayesian model averaging (BMA) was conducted. The results suggest that studies controlling for the variables of direct foreign investment and age results in a larger effect of development on migration and that the presence of country-level differences boosts migration inflows, particularly in OECD countries.

JEL Classification	F12, F21, F23, H25, H71, H87
Keywords	Meta-analysis, migration, development, bayesian model averaging
Title	Migration and Development: A Meta-Analysis

Acknowledgments

The author is grateful especially to doc. PhDr. Tomáš Havránek, Ph.D., doc. PhDr. Zuzana Havránková Ph.D., and Mgr. Dominika Ehrenbergerová (Koulcunová) for their help and comments.
Dedicated to the memory of my father. And I would also like to thank my family, for all of the love, support, and encouragement during my studies.

The author acknowledges that this thesis is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 681228.

Bibliographic Record

Palecek Rodríguez, Miroslava María: *Migration and Development: A Meta-Analysis.* Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2020, pages 78. Advisor: doc. PhDr. Tomáš Havránek, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

AIC Akaike Information Criterion

BIC Bayesian Informetion Criterion

BMA Bayesian Model Averaging

FAT Funnel Asymmetry Test

FDI Foreign Direct Investment

FE Fixed-Effects

FMA Frequentist Model Averaging

GDP Gross Domestic Product

LDC Least Developed Countries

MA Meta-analysis

MCMC Markov Chain Mote Carlo

ME Mixed-Effects

MMA Mallows Model Averaging

MRA Meta-Regression Analysis

OECD Organisation for Economic Co-operation and Development

OLS Ordinary Least Squares

PCC Partial Correlation Coefficient

PET Precision Effect Test

PIP Posterior Inclusion Probability

PM Posterior Mean

PMP Posterior Model Probabilities

PPP Purchasing Power Parities

RE Random-Effects

SD Standard Deviation

SE Standard Error

IV Instrumental Variable

WAAP Weighted Average of the Adequately Powered

WLS Weighted Least Squares

Master's Thesis Proposal

Author	BSc. Miroslava María Palecek Rodríguez
Supervisor	doc. PhDr. Tomáš Havránek, Ph.D.
Proposed topic	Migration and Development:
A Meta-Analysis	

Motivation Migration is one of the most determining features to explain the changes in population, when individuals aspire to improve their well-being and the society itself. According to the International Organization for Migration (IOM), a 'migrant' is "a person who moves away from his or her place of usual residence, whether within a country or across an international border, temporarily or permanently, and for a variety of reasons." Some of the verified causes of migration are demographic growth, precarious living conditions, lack of minimum resources for subsistence, absence of a labor market, and private or political violence.

Nowadays migration is one of the most significant problems that the world faces given the demographic and cultural phenomena that tend to increase due to political, social, and economic situations, especially in developing countries. This question is important to future migration policies that will offer better resolutions for the mobility transition, which claims that the type of migration that occurs within a country depends on how developed it is or what type of society it has Zelinsky (1971). In order to understand how this phenomenon behaves, current worldwide researchers into the phenomenon of migration attempt to explain the mobility transition of societies which influence the receiving and sending of the migrant population.

Studies of migration are based on four theories of the global phenomenon. The hypothesis of the 'Neoclassical economics', which is the first theory developed for the study of migration, assumes a world system (capitalist) where the decision to migrate is individual-based either on a macro-level: as a consequence of the structural determinants, that is, the geographical diversities between the labor demand and the labor supply of the destination society, or at a micro-level: individuals analyze whether it is convenient or not to migrate based on the inequalities in wages, or employment rates among countries, with this argument explaining why people move from one place to another looking for areas where the salary is higher than in their place of origin. Thus, the fundamental factor of migrating in this theory is based on economic benefits, which maximize the individuals' income, particularly wages. The 'Dual labor market' theory suggests that pull factors in developed countries drive international migration, where primary and secondary sectors categorize the labor markets. Furthermore, the 'Worlds systems' theory argues that migration is a consequence of the global market increase in capitalism in developing countries, where regional development leads to internal migration or to a country's development in the international arena.

In addition, the most recent migration theories responded to the neoclassical theory as a result of the developing nature of the world, called the 'New economics of migration theory,' which introduced a different approach. In order for migration to happen, migrants are considered to make collective or family decisions in order to maximize their income and employment opportunities, minimizing risks by diversifying them. Thus, migration can modify a society' s income distribution, which increases the number of people who want to migrate. As such, migrant decisions are not based purely on individual utility-maximizing calculations but are instead a household response to both income risk, and the failures of a variety of markets: labor market, credit market, or insurance market Massey *et al.* (1993).

In recent decades, the concern about the examination of the migration process has increased due to the rapid growth that has occurred in this sector. Therefore, an improvement is required to evaluate, determine, and compare migration policies, and the governance of the countries which are affected by immigration, and emigration, since international migration is a highly relevant multidimensional reality for the development of origin, transit, and destination countries (Addis Ababa Action Agenda, 2015). Faced with new migrations, scholars of the phenomenon recognize the impossibility of establishing a dominant theory, or discipline, that explains and facilitates its understanding. My thesis will focus on analyzing the migration-development nexus in the literature utilizing a statistical method called Meta-Analysis, which combines data from multiple studies to provide a comprehensive review of the current knowledge. There have been no previous meta-analyses performed that compare the estimates from previous researches and assess the evidence to evaluate the relationship between emigration and immigration, and the implication of their variation on the development of a country.

Hypotheses

Hypothesis #1: Economic growth (GDP) stimulates migration.

Hypothesis #2: The higher the average per capita income and unemployment rate, the greater the immigration.

Hypothesis #3: The literature on migration is biased toward the development-migration nexus.

Methodology This study aims to analyze multiple studies on the effect of development on migration patterns, thus measuring the size of the effect of this phenomenon, and identifying the causes of the variation across empirical literature. Since most difficulties for the study of migrations lie in their extreme diversity, in terms of forms, types, processes, actors, motivations, and socio-economic and cultural contexts, this makes the problems that theories encounter to explain such complexity understandable. Using the meta-analysis approach will help to address this issue by providing a quantitative effect and examining the heterogeneity and bias among studies. It "refers to the statistical analysis of an extensive collection of results from individual studies to integrate the findings. It connotes a rigorous alternative to the casual, narrative discussions of research studies that typify our attempt to make sense of the rapidly expanding research literature", Glass (1976).

Following the guidelines for research reporting in meta-regression analysis Stanley & Doucouliagos (2012), the collection of several studies' estimates is conducted in order to quantify the effect of interest; this is the migration-development nexus. Every estimate of the size effect will be weighted as an average of the observed effects from the studies included in the systematic review. Thus, each study's weight will be distributed for its precision. Hence, the meta-analysis quantifies the effect of the combination of several studies that may result in it being less influenced by local results of a single study. After assembling the final data set of the estimates that will be used in the meta-regression, the following two steps are implemented: 1) the examination of publication bias, and 2) the examination of heterogeneity.

The first step is to "filter out systematic biases, largely due to miss-specification and selection, already contained in economics research" (Stanley & Doucouliagos, 2012, p.16), as such as results having to be statistically significant, or when researchers indicate favorable results based on their study. In the case of there being a presence of publication bias, the reported effect will be correlated with its standard error, that is, the estimate will depend on its standard error. In order to test the presence, or absence, of publication bias within the reviews, funnel charts and related statistical methods can be used.

For the second step, as the estimates arise from different data sets, studies show differences regarding their characteristics. Therefore, the estimated regression consisting of bias might be heteroscedastic, so, to correct this, weighted least squares (WLS) is applied, that is, where the inverse of each estimate is employed as weights. Moreover, omitting essential variables in the models' regression of the measurement of development on migration will lead to model uncertainty. To address this issue, the Bayesian model averaging (BMA) approach will be used, as well as frequentist checks to corroborate the robustness of the BMA results.

Expected Contribution The purpose of this thesis firstly is to analyze the migration-development nexus in the growth literature and attempt to provide a comprehensive review of the current knowledge. Secondly, I will quantitatively explore the effect of migration on economic performance, which is traditionally perceived as unsettled even though the latest studies seem to suggest a positive relationship. Although the stock of literature on the effect is vast, researchers lack consensus regarding the size and magnitude of the effect. Using meta-analysis techniques, I will synthesize the empirical findings on this effect, adjusting for the biases and without controlling the design of the studies, thus constructing the first quantitative surveys on these effects.

Outline

1. Introduction
2. Literature review
3. Meta-Analysis
4. Data collection
5. Estimation of the mean effect
6. Examination of publication bias: publication selection, funnel asymmetry test and precision effect test (FAT-PET)
7. Examination of heterogeneity: meta-regressions and Bayesian Model Averaging (BMA)
8. Results
9. Conclusions
10. References

Core bibliography

Cazachevici, A., Havranek, T., & Horvath, R. (2019). Remittances and Economic Growth: A Quantitative Survey.

Clemens, M. A., 2014. Does development reduce migration?. In International Handbook on migration and Economic development. Edward Elgar Publishing.

Doucouliagos, H., Stanley, T.D., 2009. Publication selection bias in minimum-wage research? A meta-regression analysis. Br. J. Ind. Relat. 47 (2), 406-428.

Glass, G. V., 1976. Primary, secondary, and meta-analysis of research. Educational researcher, Sage Publications Sage CA: Los Angeles, CA, 5, 3-8.

Gechert, S., Havranek, T., Irsova, Z., & Kolcunova, D., 2019. Death to the Cobb-Douglas Production Function? A Quantitative Survey of the Capital-Labor Substitution Elasticity.

Massey, D. S.; Arango, J.; Hugo, G.; Kouaouci, A.; Pellegrino, A. & Taylor, J. E., 1993. Theories of international migration: A review and appraisal Population and development review, JSTOR, 431-466.

Stanley, T.D., 2005. Beyond publication bias. J. Econ. Surv. 19 (3), 309-345.

Stanley, T.D., Doucouliagos, H., Jarrell, S.B., 2008. Meta-regression analysis as the socioeconomics of economics research. J. Socio-Econ. 37 (1), 276-292.

Stanley, T.D., Jarrell, S.B., 1989. Meta-regression analysis: a quantitative method of literature surveys. J. Econ. Surv. 3 (2), 161-170.

Taylor, E. J., 1999. The new economics of labour migration and the role of remittances in the migration process International migration, 37, 63-88.

UNESCO., 2017. Migration as a Development challange, Analysis of root causes and policy implications. United Nations Educational Scientific Cultural Organization. Management of Scoial Transformations Programme.

Vogler, Michael and Ralph Rotte., 2000. The Effects of Development on Migration: Theoretical Issues and New Empirical Evidence, Journal of Population Economics.

Zelinsky, Wilbur., 1971. The Hypothesis of the Mobility Transition, Geographical Review.

Chapter 1

Introduction

Migration is an essential component in the global economy that influences the demographic, sociological, economic, and geographical scope of the countries facing this phenomenon. International migration affects the short- and long-term development of both the sending and receiving countries, so it has become a globally discussed debate that seeks ways to discourage and minimize the continuous growth: The total number of individuals who do not live in the country where they were born is approximately 3.5% of the world's population with most migrants living in developed countries IOM (2019).

The existing theoretical and empirical literature on the impact of migration systems is based on grand theories, which do not come from a homogeneous model since several different factors are considered. The variations between the countries and regions considered, the policies that regulate migratory flows, and the different mechanisms for obtaining it make the mobility transition complex to quantify. The empirical research on emigration determinants has provided numerous studies which suggest that when the countries of origin are in the initial stages of development, and even if the income differentials in the destination countries decrease, economic progress will increase migration Vogler & Rotte (2000).

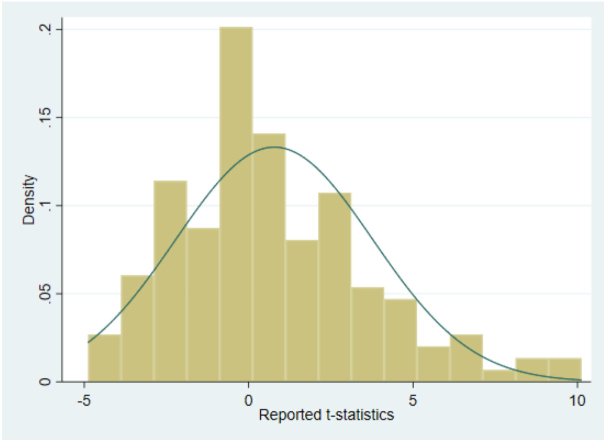
Consequently, a more precise and broad understanding of the correlation between migration and development is needed, which is usually shown as an inverted U-shape relationship Zelinsky (1971), implying that an increase in the economic development of a country will create favorable conditions, leading to emigration decreasing. The evidence from previous research on the migration-development nexus has not reached any conclusive results. According to the Addis Ababa Agenda, the correlation between migration and development outcomes can be either positive or negative. On the one hand, emigration can have a positive effect through the remittances that are sent from immigrants to developing countries. However, on the other hand, it can produce a negative impact on the sending countries by generating labor shortages, as highly educated and skilled migrants tend to be more likely to leave their country.

During the past two decades, research has estimated the economic impact of immigration, but due to differences in data, study design, and the variation across studies, it is not easy to make significant comparisons. Meta-analysis is a useful tool that synthesizes controversial results serving the purposes mentioned above. In a quick overview of the literature, Figure 1.1 depicts how the reported t-statistics of development, particularly the GDP on migration, vary widely from values between -5 and 10 in primary studies. One of the fundamental control variables that measures development on migration is GDP. According to the migration transition hypothesis, Sanderson & Kentor (2009), McKenzie *et al.* (2014), Bahna (2008), and Jennissen (2003) determine that GDP has a positive and robust effect on international migration. As GDP per capita increases, emigration will decrease. On the other hand, Bertoli & Moraga (2013), and Cristina (2008) maintain that GDP harms migration, especially in the push factors (severe conditions in the country of origin that influence migrants to want to emigrate).

Some recent meta-analyses measuring effect of migration are Longhi *et al.* (2008) on labor market impacts and Ozgen *et al.* (2010) on income growth and convergence. However, to the best of my knowledge, there has been no meta-analysis carried out to measure the effect of development on migration. Hence this leaves the following questions: How does development affect migration? Does development reduce migration? What endogenous and exogenous factors weigh in the decision to migrate? Are the results across empirical literature influenced by publication selection? Are the reported results subject to heterogeneity? Does the regression deal with causality? To what extent does the direction of the effect go from development to migration and not vice versa? To address these questions, meta-analysis techniques are employed.

The present study aims to analyze for the first time the impact of devel-

Figure 1.1: Reported t-statistics of the effect of development on migration



Note: The figure shows the histogram of the reported t-statistics from the dataset of the effect of GDP on migration in individual studies.

opment on migration. To measure development, the gross domestic product (GDP) will be used, which helps to determine the size of a country's economy. A meta-analysis will be employed of several econometric researches, which have included GDP as an explanatory variable in the estimated regression models of migration. It is important to point out that this is the first study in which the migration-development nexus has been examined from this perspective. Overall, I collected 179 observations from 40 different studies. In this way, meta-analysis will help obtain more precise estimates of the migration effect by combining the results of the studies and measuring why the estimates vary. Nevertheless, estimates need to be comparable; this means they must have a standard metric such as standard errors or elasticities. To standardize the measured effect, the partial correlation coefficient will be used as the primary studies diverge on estimating the migration effect.

The thesis is structured as follows: Chapter 2 presents the concept of international migration and its link with economic development. Overall, this chapter gives a theoretical overview of the topic of migration. It describes and compares the most critical migration theories and hypotheses of the last decades. Additionally, it reviews the empirical literature through the controversies of previous results. Chapter 3 provides the methodology employed, describing the meta-analytical techniques. It presents the estimation of the overall mean effect, the consequences of publication bias and heterogeneity, and how to deal with it. The first is addressed through graphical tools such as funnel asymmetry and precision asymmetry effect tests. The second is to deal with model uncertainty, which is inherent to the meta-analysis technique, that is, the heterogeneity problem. Bayesian Model Averaging (BMA), and Frequentist Model Averaging (FMA) will help to correct this issue by determining which conditional variables explain the best baseline model based on their 'importance.' Chapter 4 describes a brief overview of the data and the strategy conducted to collect the primary studies. Chapter 5 presents the results of the meta-regression analysis summarizing the main findings. Chapter 6 summarizes the main concluding remarks.

Chapter 2

Literature review

This chapter presents the concept of international migration and its link with economic development. Overall, this chapter looks to give a brief description of several more important hypotheses that have been studied throughout the last decades. These hypotheses relate economic variables with the migration variables in several contexts. Furthermore, in this chapter the theoretical background is defined as is the empirical literature within the related topic.

2.1 What is the concept of Migration?

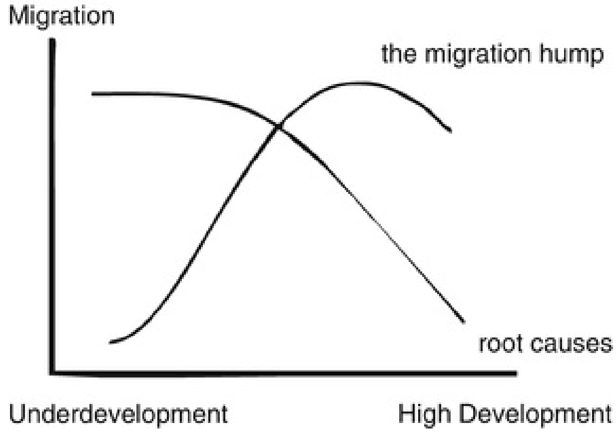
Migration is a historical phenomenon of human mobility associated with structural changes in development. The concept of migration is defined as the displacement of an individual from one place to another. According to the Cambridge Dictionary, emigration is “the process of leaving a country to live permanently in another country”; whereas immigration is “the arrival of people to a place or country, to which they do not belong originally”. These movements each have different factors that motivate people to migrate, such as political, economic, social, cultural causes, and war.

In recent years there has been widespread concern about the international mobility of the population, and the impact on the development of the origin and destination countries. The migratory phenomenon is one of the most significant problems facing many parts of the world, and is ever-increasing, particularly in developing countries. Migration, together with births and deaths, helps constitute the three principal components that explain population changes. The global estimate of international migrants has reached 272 million people (UN, 2019). Even larger numbers migrate within a country's borders (740 million internal migrants in 2009, according to UN 2009). *Why do we care?* Migration is considered to be an important transfer of manpower that facilitates the economic growth; moreover, it addresses labor market imbalances Jauer *et al.* (2014), supplies both high- and low-skilled occupations (OECD, 2012), and decreases the average age of working population Gagnon (2014). From 2000 to 2014, according to McKinsey (2016), immigrants contributed up to 80 percent of the labor-force growth in major destination countries. Such countries of destination (or host countries) have several features in common: in comparison to the country of origin (or a home country), they are more developed, they tend to have better labor, commuting and salary opportunities, and they provide better social security and have higher political stability.

One of the significant contemporary challenges is how to manage migration. The debates range from those who advocate the preeminence of national security, controls, and the closing of borders, to those who advocate privileging human security and the free movement of people who voluntarily make their migration decisions. Therefore, determining the effect that development has on migration through a systematic review would favor achieving a balance between the free movement of migrants or the controlling of this. The present study focuses on the impact of migration on development, measured through GDP, which is a standard indicator of economic growth and how productive and developed an economy is. Economic development is the application of capital to raise human productivity, generate wealth, and increase national income Massey *et al.* (1993). It refers to the ability of a country to create wealth and build the quality of life of its inhabitants. Thus, it not only applies to the productivity capacity of a country, but also to how these resources are used.

Nevertheless, Figure 2.1 shows how a positive relationship between migration and development, in short-term migration, may increase if economic growth is present. However, sooner or later as a country develops, the "migration hump" appears when the country of origin enjoys more favorable economic circumstances causing the inhabitants to stay or to return to their place of birth.

Figure 2.1: Migration and development relationship



Source: Martin & Taylor (1996).

Note: The figure depicts the so-called root causes of migration, suggesting that emigration may accelerate in the short term as economic development increases.

2.2 Theoretical Background

The theoretical channels through which the migration occurs are explained by four major theories. The oldest and most recognized is the 'Neo-classical economics theory'. As neo-classical microeconomists argue: migrants evaluate the costs and benefits of moving to a desired location, establishing migration as an investment structure. The arguments of this theory at the macro-economic level perceive migration by the differences in wage levels from labor supply and labor demand between countries. Hence, if there were no differences in wages, labor migration would cease. Second, the micro-economic level proposes that wage differences and employment rates produce international migration, as individuals seek higher earnings and better living conditions than in their home country.

The 'Pull-push model', advocated by Lee (1966), Ravenstein (1885), or Passaris (1989), suggests an existence of so-called push factors, troublesome forces that make people move out of their homes (such as unemployment, poverty, or lack of political freedoms, repression, persecution, low wages), and so-called pull factors, the alluring elements of the host country (such as higher standards of living or better education). The consequence of this population movement creates a balance within the labor markets: in countries with limited capital, job supply decreases and wages increase; in richer countries, job supply increases and wages fall.

On the other hand, the 'New economics of labor migration', a recent theory, suggests that migration happens as an effort of households to overcome home market failures (such weak credit and insurance markets, Taylor (1999)) that constrain local production. Migrants then send remittances from the host countries back to their family home, thus diversifying the national income. Third, the 'Dual labor market' theory suggests that migration is driven by vacancies in labor-intensive segments of host countries Piore (1979) and migrants optimize their position between the two available labor markets at hand: the one that pays well and is stable, and the one that does not and is not.

Finally, Wallerstein (2011) advocates the 'World-systems' theory, which examines migration as a natural consequence of globalization and market penetration across borders: modernization suggests higher productivity which brings higher profit, and thus wealth. When migrants move from one country to another, they are believed to stimulate the economic growth. On the other hand, increased social diversity and other changes injected to the labor markets could equally hamper such growth Bove & Elia (2017).

2.3 Empirical literature

During the last decade, the discussion on the relationship between migration and development has been a global concern - as a country encounters economic development, emigration will first rise and then decrease (Zelinsky (1971)). While countries are in poverty, the ability to emigrate is limited. As the country begins to develop, emigration rates begin to rise until they reach their maximum point at which they are in a better industrial state. Finally, from that point the migration begins to decrease.

After going through an extensive review of the existing literature on migration and development, I find that there is some debate about the correlation between these two aspects. While some studies claim that development can have a positive impact, others show that there is a detrimental effect of the imbalances in development. There can be many reasons for these differences in results, such as limited data, the quality of the data, the sample selection, the time period, the type of methodology, the control variables selected, and the omitted variable bias. Such analysis, I presume, is highly relevant to the public policy. Moreover, national immigration policies also play a significant role in determining the size of these flows Hatton & Williamson (2005). If migration is well addressed via adequate migration policies, policymakers have the ability to improve well-being and economic growth of both the home and the host country. Host countries and international organizations have looked into where more effective controls and coordination of migratory flows are required. Particularly in European development policies, if socio-economic and political conditions improve through development cooperation, migration flows will diminish and could be prevented.

Across the literature, the authors use different techniques and estimation methodologies to obtain immigration inflows. More often than not, the studies use panel data to estimate the development-migration nexus.Rotte & Vogler (1998), provides the advantage of using a panel structure of the data on international migration estimation, given that panel data exploits both time-series and cross-country variation. Empirical research on the drivers of migration has shown a sophisticated and complex relationship between migration and development. Many authors suggest that the relationship between migration and development must be narrow and focused on improving the living conditions of both the migrant, and the family or community left behind. Some studies try to focus on the explanation of development through economic and non-economic drivers of migration; these are the observable characteristics such as nationality, wealth, and education, among others. However, only some categories have economic prominence; for this purpose, the present work is essential to examine which of these variables are more relevant as a determinant of this effect.

One of the main explanatory variables to help measure development on migration is GDP. Sanderson & Kentor (2009), McKenzie *et al.* (2014), Bahna (2008), and Jennissen (2003) find that GDP has a strong and positive effect on international migration, in that net migration will increase and then, as GDP per capita starts to rise, emigration will decrease, according to migration transition hypothesis. In contrast, Mayda (2010), Simionescu *et al.* (2016), Bertoli & Moraga (2013), and Cristina (2008) argue that GDP is a declining factor for net migration, which is not consistent with the theory particularly relating to the impact of the push factors, thus leading to insignificant results. Besides this, the discrepancies of GDP generate flows of less-educated immigrants into the USA. However, in some developing countries, they see the massive arrival of foreigners as a decisive factor for their finances, since it produces a reduction in the unemployment rate as an influx of migrants of working age helps to counteract the labor shortage, resulting in a demographic transition and boosting GDP growth, especially in advanced economies. Many studies suggest that the 'migratory crisis' that Europe is currently experiencing is not likely to be a crisis but could actually be an economic opportunity.

Income is considered one of the essential features to measure the impact of migration. According to Pedersen *et al.* (2008), Ortega & Peri (2013), Ederveen *et al.* (2007) and Pytlikova (2006), income has a strong positive effect on migration, that is, a higher income per capita in a given destination country induces an increase in immigration flows. On the other hand, Murat (2020) and Dao *et al.* (2018) argue that the country of origin's average income has no robust correlation with immigration. Lastly, Cristina (2008) considers that individuals decide to migrate with their predicted income as a determinant factor, according to Todaro's model.

Another variable included in the regression model is the stock of foreign direct investment (FDI), Sanderson & Kentor (2008), and Buch *et al.* (2006) which indicates that FDI increases net migration over time, especially in the least developed countries and in the primary sector, while in the secondary sector it has a detrimental result. Next, Vogler & Rotte (2000) indicates that trade is a critical determinant for net migration, diminishing it by their integration. While, Rotte *et al.* (1997) suggests that trade, at least in the medium-run, will not lessen migration. Labor mobility strengthens beneficially from outmigration, according to Sanderson & Kentor (2008), and Hanson & McIntosh (2010). Moreover, Ederveen *et al.* (2007) finds that it is crucial to include the female labor determinants to lessen the adverse effects of low migration mobility.

Population growth also plays a vital role in the estimation of migration, according to Hatton & Williamson (2003), and Hanson & McIntosh (2010), the acceleration of this process will lead to more international migration. Whereas,Vogler & Rotte (2000) contradicts this result, by implying that it does not have a direct effect and certainly does not increase the emigration. Besides, education is also one of the critical drivers of migration. As emigration rises with development in the country of origin, it means that the education level will increase promoting individuals to emigrate abroad, according to Dao *et al.* (2018), Dreher & Poutvaara (2006), Docquier *et al.* (2014), Jennissen (2003). Meanwhile, Pedersen *et al.* (2008) predict that even if the level of education is low, there will be an increment in migration to the OECD countries from developing countries.

Among the variables that affect migration, distance seems to also be one of the influential ones. It has a negative and significant effect, according to Mayda (2010), Pedersen *et al.* (2008), and Foo (2017). Moreover, Ederveen *et al.* (2007), and Jennissen (2003) explain that unemployment harms the country of origin. As a result, an individual would leave their country when the unemployment rate increases. Particularly in the Baltic and Central European countries, migrant flows increase with employment opportunities in the destination countries.

The network effect in destination countries has shown positive and substantial evidence for inducing and reinforcing the probability of migration, as it is one of the factors that facilitates this process by having a connection in the destination, which may reduce costs and the uncertainty of possible future migrants. Particularly immigrants from Central European countries and Mexico to the USA, according to Hanson & McIntosh (2010), Rotte *et al.* (1997), Pedersen *et al.* (2008), and Docquier *et al.* (2014). Whereas, Pytlikova (2006) found that the Baltic countries do not rely on network effects. Last but not least, the political situation has been studied as a relevant factor that causes migration from developing countries, since governance affects individuals; there is a direct relationship, Rowlands (1999).

Chapter 3

Methodology

This chapter describes the methodology of meta-analysis and the model of Bayesian model averaging (BMA) applied in this thesis. Firstly, it gives a concise description of meta-analysis methodology, as well as the origins of the methodology. Secondly, it presents how estimation of the mean effect will be used in this case of study, in particular, it presents the partial correlation coefficient (PCC) and how it serves to measure the migration. Finally, it presents the essential element of the interpretation of the Bayesian model averaging in our case of study, that is, the examination of publication bias and heterogeneity, and how to deal with it.

3.1 Meta-analysis

The current literature on international migration is just as diverse as the migration itself. Meta-analysis is a way of quantitatively analyzing the inference from the literature on a certain effect, where no clear message on this effect is present. To the best of my knowledge, there is no study that evaluates the important relationship between migration and economic growth systematically, using meta-analysis tools that exploit the work on this relationship contained in previously published literature. Furthermore, the changes injected into the labor markets, such as possible irregularities (or improvements, for that matter) in human capital, due to migration, are considered important by many (Miyagiwa (1991); Clemens *et al.* (2014)), but remain unexplored via meta-analysis as well. Meta-analysis could not only suggest which drivers (such as level of education, unemployment rate, distance to the destination country, marital status, income, political situation of the origin country, inflation, trade, publication bias, models and estimation methods used in the primary studies, and so on) could be relevant to the estimated effect in question, but which of these could also provide guidance on the magnitude and direction of these effects.

Meta-analysis is a tool that is commonly used to review the quantitative research literature about some observed effect with the purpose of integrating the research conclusions to the studied literature. The results discovered from a particular topic may indicate contradictory opinions across the literature. Meta-analysis suggests a reasonable and more precise interpretation. As Stanley (2012, p. 3) stated: "meta-regression analysis is a systematic and comprehensive review of all existing, yet comparable, empirical evidence . . . [which] . . . allows the systematic reviewer to model and estimate any explanatory or biasing factor for which information or a proxy is available and thereby filters out their influence on our scientific knowledge." Pearson made the first remark on this type of examination in 1904 where his strategy was to group the results. However, Glass (1976) was the first to formally come up with the concept of meta-analysis, which he defined as "The statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings . . . to make sense of the rapidly expanding research literature". He suggested the foundations of this methodology in 1976.

Meta-regression analysis (MRA) is a practical technique (a particular tool of meta-analysis) intended to summarize and explain a vast number of empirical results. Those empirical results must be obtained through the means of statistical techniques that are applied to integrate the findings of a variation of results collected through several studies. In each study, the size effect or treatment effect is determined to obtain the summary size effect and to assign a weight to each study. The highest weights are generally given to those cases in which the investigations are the most accurate. This variation can arise either as a consequence of sampling error or by the variability in the effects population. Meta-analysis (in particular meta-regression analysis) is an excellent tool to address the discrepancies mentioned above because it is capable of:

1. Detecting the so-called publication bias.

2. Analyzing the excess between-study variation.

3.2 Estimation of the mean effect

Following the guidelines for research reporting in meta-regression analysis, the first step is to search several primary studies across the literature on the effect of interest according to the respective search protocols. For further details see Stanley & Doucouliagos (2012). After careful examination, I do consider the literature on the relationship between migration and economic development to provide such comparable estimates, yet only recalculated to partial correlation coefficients (PCC), which according to Doucouliagos (2011), standardize the estimated effects. The resulting estimates retain the ordinality of the original measure but are, however, still conflicted on the idea of how large the effect should be. Several recent papers, such as Doucouliagos (2011), Havranek *et al.* (2016) or Cazachevici *et al.* (2020), use this type of standardizing measure for a practical significance. Thus, calculating partial correlation coefficients (PCC) with the equation below:

$$PCC_{ij} = \frac{t_{ij}}{\sqrt{t_{ij}^2 + df_{ij}}} \tag{3.1}$$

where:

- PCC_{ij} is the partial correlation coefficient of the regression i in the study j.
- t_{ij} denotes the t-statistic of the corresponding regression coefficient i in the study j.
- df_{ij} denotes the degrees of freedom of this t-statistic of the regression i in the study j.

The PCC_{ij} is the correlation between migration and development ceteris paribus and the impact when all other variables are partialled out. Partial correlation coefficient takes values at the interval $[-1, 1]$. So the closer that $|PCC_{ij}|$ is to 1, the larger is the effect.

Their respective standard error $SE_{PCC_{ij}}$ is calculated for every PCC_{ij} as follows:

$$SE_{PCC_{ij}} = \frac{PCC_{ij}}{t_{ij}} \tag{3.2}$$

According to the guidelines for interpreting partial correlation in economics provided by Doucouliagos (2011), absolute value partial correlations coefficients that are greater than 0.327 are considered to have a strong relationship between the economic variables, values between 0.173 and 0.327 can be regarded as a medium effect, and values smaller than 0.070 can be reasonably considered to have a small effect.

3.3 Estimating the effect of development on migration

Most of the primary studies estimate the effect of migration with the following regression:

$$M_{it} = \alpha + \beta GDP_{it_{-1}} + \gamma X_{it} + \epsilon_{ij} \tag{3.3}$$

where

- M_{it} denotes the measure of migration in country i at time t.
- $GDP_{it_{-1}}$ real GDP rate in country i at time $t-1$.
- X_{it} represents a vector of explanatory variables that measure migration in country i at time t.
- ϵ_{it} accounts for the error term in country i at time t.

Some common explanatory variables that measure migration are age, sex, level of education, income, FDI, inflation, unemployment rate, and Gini coefficient, among others. Equation (3.3) reflects the effect of migration on either the host country or place of origin with a panel data specification.

3.4 Examination of publication bias

For the sake of completeness in any research activity, the more information that is possessed on a certain topic, the more precise the analysis will be. As great as it would be to be able to review all human knowledge on a certain topic, the truth is that this is not feasible. The main reason for this is that not all knowledge has been published. So, it is to be expected that some bias in the literature on certain phenomena will be found, because, as it has been stated 'the literature is incomplete'.

3.4.1 Publication selection

The first step in conducting a meta-analysis is to examine publication bias by using the funnel plot graph, the funnel asymmetry test and the precision effect test (FAT-PET). Publication selection bias occurs when there are no common standards for the decisions on which reports will be published. In a perfect world, the quality of a study should determine its publication. However, referees tend to accept more papers with values that reaffirm the previous empirical literature. So, this leads to the rejection of authors who have results that do not agree with the current theory. This kind of favored decision which builds on the magnitude and/or direction of the estimated effect is usually referred to as type I publication bias. Publication selection bias is likely to occur when there is a strong preference in the literature (by whoever makes the meta-analysis) for a certain type of results. This problem of publication selection is commonly known as the "file drawer problem" Rosenthal (1979). This problem is well represent by the following cite "Everyone may have a preference for statistical significant findings" of Card & Krueger (1995).

In the present case of study, an example could be that the latest literature on the migration-development nexus favors positive values of the estimated relationship (see IOM, 2018, for example). A negative result is one that fails to reject the null hypothesis of the nonexistence of the studied effect. This kind of refusal towards negative results is commonly known as type II publication bias. Even if a researcher can often clearly argue why the estimate with a non-intuitive or unexpected sign is not to be used and is better discarded, the 'wrong' estimate may appear in the literature from time to time merely due to the laws of chance. As 'wrong' negative effects get discarded, positive effects will be over-represented in the research record and the overall picture from the literature will become biased Stanley (2005). It is possible to consider these types (I and II) of bias as quantitative bias. Meanwhile, there exists bias which originates from qualitative features such as:

- Preference towards certain languages.
- Preference towards papers with sponsoring.
- Inefficient design and ambiguous reporting method of single studies.
- Inclusion of several results.

- Errata choice of regression models used in meta-regression analysis.
- Deficient design, review, and execution of meta-analysis.

among others.

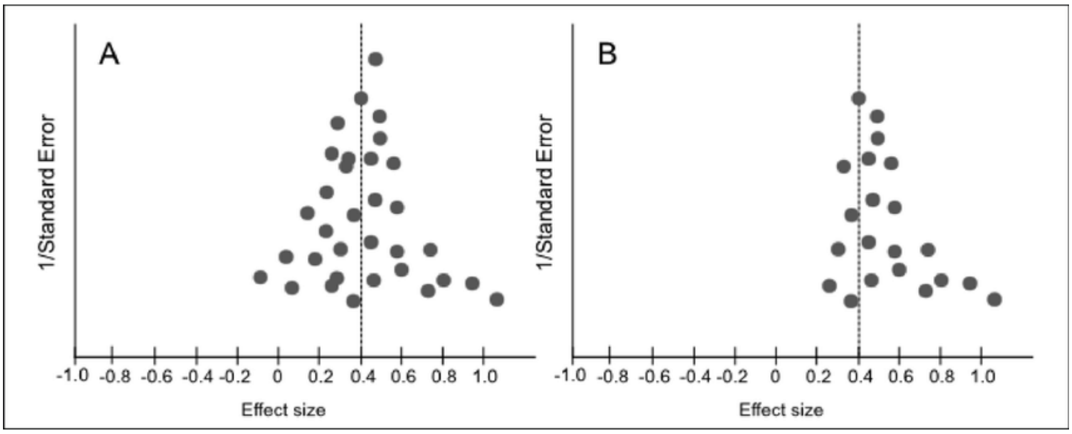
3.4.2 Identifying publication bias

One of the most used methods of spotting publication bias is the funnel plot. The funnel plot method consists of constructing a funnel plot and detecting the bias from the geometric properties of this plot Egger *et al.* (1997). A funnel plot is a scatterplot where the horizontal axis measures the magnitude of the estimated effect, partial correlation coefficient (PCC), while the vertical axis measures the estimate's precision, represented by the inverse of the standard error ($1/SE_{PCC}$). If the publications are unbiased, then the funnel should look symmetrical respect 'true effect' with the most precise estimates concentrated around the 'true effect'. The estimates placed at the bottom should be more dispersed, while those at the top will be compressed as they are more precise. In the case of an absence of publication bias, both the small studies and the larger ones will be distributed symmetrically on both sides of the overall estimate of the effect. Thus, the magnitude of more prominent and accurate studies will be closer to the total value, while smaller or imprecise studies will show more dispersion. Several factors, however, can lead to a funnel plot ending up skewed, such as heterogeneity, differences in methodology, quality of data, or even chance. "In practice, it is difficult to distinguish between these potential reasons for funnel plot asymmetry, or indeed to distinguish any of them from chance" Langan *et al.* (2012).

For practical reasons it is a common hypothesis that the estimates with the highest precision value are closer to the 'true effect'. Thus, the symmetry around the 'true effect' could be estimated from the clustering of the more precise estimates. Also, from this hypothesis it is deduced that the higher the precision, the closer to the 'true effect', so the plot will have the shape of an inversed funnel. It is worth mentioning that measuring the precision as the inverses of the standards deviation is just one way to do it. The precision can also be measured by the quantity of studies taken in the primary studies to calculate those estimates Stanley (2005). The main reason we do not favor this method is because many of the studies, in the case of study presented in this work, do not report how many observations they have used in their regression.

As it is shown in Figure 3.1, in funnel plot A the dispersion of estimates distributes symmetrically around the 'true effect'. Thus, this demonstrates that if the funnel plot is symmetrical, then the estimates do not have a publication bias. Meanwhile in funnel plot B, the dispersion of estimates is distributed skewed around the 'true effect'. Which means that, in this example, the publication estimates are biased. However, in both plots it is shown that if the standard deviation is low, that is, that the value in the plot with respect to the y-axis is higher, then the dispersion respects that the 'true effect' is low, which means that the estimate is more precise.

Figure 3.1: Symmetrical and asymmetrical funnel plot



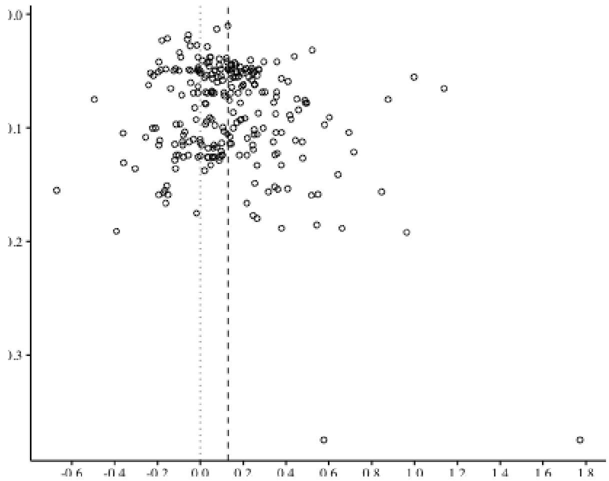
Source: Impellizzeri & Bizzini (2012).

Note: The figure shows an example of a symmetric (A) and an asymmetric (B) funnel plot, where each dot represents a single study, the x-axis depicts the result of the study, the y-axis depicts the effect estimate of the standard error where larger studies with higher precision are placed towards the top and studies with lower precision are placed towards the bottom.

As was mentioned earlier, it is assumed that more precise estimates distribute around the 'true effect'. So, a way to estimate the value of the 'true effect' is by calculating a measure of centrality of the magnitude with the estimated effect as the mean or median.

It is possible to examine the correlation between the effect and its standard error more rigorously Stanley *et al.* (2008). That is to say, that the formulation of the method is of an analytic nature instead of the visual nature of the funnel

Figure 3.2: Estimates distributes around the "true effect".



Source: Impellizzeri & Bizzini (2012).

Note: The figure shows an example of more precise estimates, that is, studies with higher power.

plot. The analytic technique that is proposed is based on regression analysis. The essential assumption of this technique is that: if it is the case that there is a of lack of publication bias, then the estimates should be randomly distributed around the 'true effect' and they should be independent of their standard error. So, the corresponding equation should be:

$$PCC_{ij} = \beta_0 + \beta_1 SE_{PCC_{ij}} + u_{ij} \tag{3.4}$$

where

- PCC_{ij} denotes the partial correlation coefficient of the i-th effect estimated in the j-th study.
- $SE_{PCC_{ij}}$ denotes the standard error of the partial correlation coefficient of the i-th effect estimated in the j-th study.
- u_{ij} is the error term of the i-th effect estimated in the j-th study.

The intercept of the equation, β_0, is the 'true' underlying effect beyond publication bias. The slope of the equation, β_1, represents publication bias. Therefore, if publication bias is present and $\beta_1 \neq 0$, correlation should be observed between the partial correlation coefficient (PCC) and their standard error - either because researchers discard negative estimates or because researchers compensate large standard errors with large estimates of partial correlation coefficient (PCCs).

A common assumption that appears in regression analysis about the family of error terms $\{u_{ij}\}$ is that they are independent and identically distributed with conditional distribution $u_{ij} \mid SE_{PCC_{ij}} \sim N(0, \sigma^2)$. However, is not a reasonable hypothesis in the case of study presented in this work. Naively, there is no reason to expect that the residuals for the regressed data linked to the same study will be independent, nor reason either to expect them to be identically distributed if they belong to different studies. But even in this case of uncertainty about the correlation between the errors, it is possible to expect that the estimates will be unbiased. The main difficulty is that the default standard errors tend to overstate precision. This may flow to an incorrect inference based on the statistical tests. The common practice to solve this issue is to cluster the regression. This is done in a way that it is possible to fit the regressions so that the standard errors will cluster.

3.4.3 Funnel asymmetry test and precision effect test

The FAT-PET approach makes a more rigorous test to detect publication selection bias. It is a good linear approximation, considering the following regression between studies effect and standard errors:

$$\text{effect}_{ij} = \beta_0 + \beta_1 SE_{PCC_{ij}} + \epsilon_{ij} \tag{3.5}$$

where

- β_0 stands for the constant term known as the 'real' population value.
- $SE_{PCC_{ij}}$ is the standard error of the i-th estimate in the j-th study.
- β_1 measures the publication bias.
- ϵ_i is the error term of the i-th estimate in the j-th study.

The Funnel Asymmetry Test (FAT) examines the null hypothesis $H_0 : \beta_1 = 0$, that is, whether the funnel plot is asymmetric. The Precision Effect Test (PET) applies the same mathematical equation as FAT, which examines the null hypothesis of $H_0 : \beta_0 = 0$, that is, whether there is a genuine empirical effect of the estimates once the publication bias is adapted and corrected.

Therefore, if the null hypothesis of FAT and PET is rejected, bias and the true effect exist. However, it is known that the Funnel Asymmetry Test (FAT) has low power, while the Precision Effect Test (PET) is strong enough for large studies. Thus, it is suggested to average only the largest studies, that is, the top 10%. This rule of thumb is stated by Stanley (2005).

Besides this, there are also non-linear techniques; this is when publication selection does not come from a linear function of the standard error to identify and correct for publication bias estimating the underlying effect that arises from the meta-analysis to approximate the 'true' effect. Such techniques are the following:

- Ioannidis *et al.* (2017): The method is based on 'adequate power.' The reported estimates are weighted by their average, where results are said to be appropriately powered if the median statistical power is greater than 80%.
- Furukawa (2019): The estimation technique focuses on the 'stem-based' method, which helps select in a more robust way, the appropriate number of the most precise studies through an algorithm that minimizes the mean standard error. This non-parametric technique optimizes the bias-variance trade-off, that is, the more precise the studies, the less they suffer from publication selection, as they are unlikely to be omitted. Moreover, those studies become irrelevant either if the estimates are statistically significantly low or have high negative values.
- Andrews & Kasy (2019): This non-parametric approach is based on 'bias-corrected' estimators and confidence sets. It considers that some studies are most likely to be published, especially those that have statistically significant results. In order to identify and correct for publication bias, the conditional probability of publication selection must be previously known, given the distribution of their p-values.
- Bom & Rachinger (2019): This approach suggests an 'endogenous kink meta-regression model' through estimating Monte Carlo simulations, which helps detect at which level there is an absence of publication selection, where the value of the standard error will be cut-off.

The present study utilizes the methodologies of Ioannidis *et al.* (2017), and the Stanley (2005) rule of thumb mentioned above, for practical purposes, and the results were similar with the use of other non-linear techniques.

3.5 Examination of heterogeneity

Once publication bias is addressed, it is possible to examine other potential drivers of the estimates' magnitude. These drivers often include different aspects of study design, such as the estimation technique, characteristics of the data set, choice of a country or a region, time period of the data, control variables that are considered in the model of a primary study or even publication characteristics of a primary study, such as the number of citations or the impact factor of a publication outlet. The number of these potential explanatory variables can get quite large (in large studies it may be around 40, see for example, Gechert *et al.* (2019)) which brings inherent model uncertainty into picture. Hence, if heterogeneity is not taken into account, it can lead the estimates to be biased. When a relevant variable is not included, it results in over-estimating or under-estimating the effect of the explanatory variables. In either case, the omitted variable must be correlated with the dependent variable, or the omitted variable must be correlated with one or more independent variables.

3.5.1 Meta-regressions

Meta-regression is a sensitivity analysis technique to examine heterogeneity generated by the relationship between the variables and study effect sizes. It is defined as the statistical dissimilarity across various studies, which implies that any average effect size will not fully indicate the economic event under examination. However, some primary studies are more precise than others, therefore those that are more precise (that is, they have broad information) will have a higher weight assigned by computing a weighted mean. In order to adjust it, the random-effects model will be the solution that assumes variability across populations. Alternatively, in the absence of heterogeneity, when the studies are similar to each other, the fixed-effects method is applied. It allows the 'true effect' to diverge from study to study where the distribution of the samples of the true effect sizes is assumed to be random. Thus, characterizing the mean of the true effect population, rather than just one common effect as would be seen in the fixed-effects model.

In order to correct the heteroskedasticity originated from the primary studies, weighted least squares (WLS) is employed, which divides each term of the equation (3.5) by their standard deviation $SE_{PCC_{ij}}$ (precision weight,) generating the equation:

$$\frac{\text{effect}_{ij}}{SE_{PCC_{ij}}} = \frac{\beta_0}{SE_{PCC_{ij}}} + \beta_1 + \frac{\epsilon_i j}{SE_{PCC_{ij}}} \tag{3.6}$$

$$\implies t_{ij} = \beta_1 + \beta_0 \frac{1}{SE_{PCC_{ij}}} + \epsilon_{ij} \tag{3.7}$$

where t_{ij} denotes the t-statistic related to the individual effect of the i-th estimate in the j-th study. When the standard error is used as a weight, the inverse of it results in a measure of heteroskedasticity. If the estimated β_1 is statistically significant, then it means the estimates present publication bias.

Additionally, to weighted least squares (WLS), the following methodologies will be applied in order to check the robustness of the results.

- Fixed effects model (FE): Fixed-effects model are more robust and general, they assume estimates dependency within at least one of the explanatory variables, and that there is a single common effect size for all the studies (one 'true effect size'), and the differences observed are due to change. Thus, allowing for correlation between the study-level effects and the conditional variables. It accounts for unobserved heterogeneity where the overall effect can be estimated as a weighted average of the individual effect of each study, the weights being the inverse of the variance of the corresponding estimate, which may result as a biased estimate.
- Mixed effects model (ME):): In contrast, mixed-effects models vary the true effect size across the studies. It accounts for within and between-study variation effects (excess heterogeneity) and within-study dependence. Hence, the error term is depicted as a study-level random-effect that is uncorrelated with the estimated standard error coefficient.

However, examining heterogeneity among primary studies from the Equation 3.6 mentioned above, will serve to detect which variables can select the best model of the estimated effect of development on migration based on the empirical literature where the estimation of migration varies within the variables chosen. Thus, by adding conditional variables (moderator variables), the equation will be as follows:

$$t_{ij} = \beta_0 \frac{1}{SE_{PCC_{ij}}} + \beta_1 + \sum_{k=1}^{K} \gamma k \frac{1}{SE_{PCC_{ij}}} \zeta k_{is} + \epsilon_{ij} \frac{1}{SE_{PCC_{ij}}} \tag{3.8}$$

where

- t_{ij} denotes the t-statistic of the corresponding regression coefficient i in the study j.
- k is the number of conditional variables.
- γk is the corresponding regression coefficient i of the respective conditional variables in the study j.
- ζk_{is} represents the conditional variable
- ϵ_i is the error term of the i-th estimate in the j-th study.

3.5.2 Bayesian moving average model

Bayesian Model Averaging (BMA) is an approach that addresses model uncertainty in terms of probability. Many potential explanatory variables can be included in the model, but usually we are not sure which ones are the ones that have more impact on the result of the model. Bayesian model averaging (BMA) was enforced by Hoeting *et al.* (1999), which essentially allows that if the probability of an event to occur is previously known, its value when we have new information can be modified. The theory does not say much about why, for instance, a particular identification strategy should lead to a systematically different partial correlation coefficient (PCC) estimate. Therefore, it is necessary to estimate the models with all the combinations of the potential explanatory variables and weight the models by the goodness of fit (using Bayesian model averaging). VanderPlas (2014), sums it well: "Bayesian approaches (...) are often conceptually more straightforward, and pose results in a way that is much closer to the questions a scientist wishes to answer: this is, how do these particular data constrain the unknowns in a certain model?". Bayesian model averaging (BMA) allows us to concentrate on the problem of heterogeneity (while still addressing the above mentioned problems) by estimating all the feasible models, in total 2^K models, where K denotes the total amount of regressors in the model (see more on BMA in Zeugner & Feldkircher (2015), for example). The first thing to determine in the BMA procedure is the probability distribution of that magnitude with the external information available, which is called a priori probability. Thus, given data D, BMA inferences are made using posterior distribution of the quantity of interest:

$$p(\Delta|D) = \sum_{k=1}^{K} p(\Delta|D, M_k)p(M_k|D) \tag{3.9}$$

where

- $p(\Delta|D)$ is the probability density function
- M_k is the set of given models from $M_1, \ldots, M_k$
- $p(\Delta|D, M_k)$ is the posterior distribution of an unknown quantity of interest Δ
- $p(M_k|D)$ is the probability that M_k is the correct model

Hence, the Bayesian model averaging (BMA) posterior distribution of Δ is a weighted average of them under each model weighted by their posterior model probabilities. Next, the information provided by the data observed in the study is quantified using what is called the likelihood function. The likelihood function is calculated by integrating over the unknown parameters. Thus, the posterior probability of the model M_k is given by:

$$p(M_k|D) = \frac{p(D|M_k)p(M_k)}{\sum_{l=1}^{K} p(\Delta|D, M_l)p(M_l|D)} \tag{3.10}$$

where

$$p(D|M_k) = \int \ldots \int p(D|\theta_k, M_k)p(\theta_k|M_k)d\theta_k \tag{3.11}$$

- θ_k is the vector of parameters
- $p(D|\theta_k, M_k)$ is the prior density of θ_k
- $p(D|\theta_k, M_k)$ is the prior probability of θ_k under the model M_k

For further details see Raftery *et al.* (1997). Finally the Bayesian model averaging estimate of θ is obtained by employing the above-mentioned equation:

$$\hat{\theta} = \sum_{l=1}^{K} \hat{\theta}_k p(M_k|D) \tag{3.12}$$

where $\hat{\theta}$ is the posterior mean under the model M_k. Thus, Bayesian model averaging (BMA) weights posterior distributions for any statistic θ.

The Bayesian model averaging (BMA) helps to set which co-variances are convenient to explain the variation of the explanatory variables, where the best models with the highest posterior probabilities (PMP) fit the data accurately, and the posterior inclusion probability (PIP, which is the sum of all the PMP's for all models), helps to detect which variables should be included in the model. In order to obtain all the posterior results, the list of all the potential variables is needed. When there is a large number of variables, the Bayesian model averaging (BMA) applies the Metropolis-Hastings algorithm which use the method of MCMC (Markov Chain Monte Carlo) approximation, where the distribution is set to the desired posterior distribution, which is a random walk through all the possible combinations of the model, selecting the one with the highest PMP, see Zeugner & Feldkircher (2015) for a detailed review.

3.5.3 Frequentist Model Averaging

Frequentist Model Averaging also addresses model uncertainty by a weighted average over the explanatory variables from the best model. It centers on the estimators where the parameters are fixed and unknown, unlike Bayesian model averaging (BMA), it does not require 'prior distributions' on the parameters instead choosing the right' weights. It uses orthogonalization where each regression will be distributed a weight of the co-variate space to lower the number of models from 2^K to K.

Following the baseline of BMA (Bayesian model averaging) from the previous section, the estimators can be constructed as:

$$\hat{\beta} = \sum_{l=1}^{K} w_l \hat{\beta}_l \tag{3.13}$$

where

- $\hat{\beta}_l$ is the estimator of parameters under the model M_k

- w_l are the weights from $l = 1, \ldots, K$

And such weights, that is, w_l can be determined according to the information criteria: AIC or BIC as follows Buckland *et al.* (1997):

$$w_l = \frac{exp(-I_l/2)}{\sum_{l=1}^{K} exp(-I_l/2)} \tag{3.14}$$

where I_l is the information criterion (the AIC or the BIC) for the model M_k.

Additionally, Hansen (2007), suggested the Mallows model averaging (MMA) estimator, where asymptotically optimality weights are chosen by minimizing the Mallow's criterion, that is, minimizing the squared error in the data fit. Thus, FMA (frequentist model averaging) is based on the goodness of fit and parsimony of the underlying included models. As a disadvantage, FMA restricts the computation of larger models as it does not account for the estimated model probabilities.

Chapter 4

Data description

The first step in meta-analysis is to collect relevant studies, for this purpose, addressing the impact of migration on development as an empirical research proposition. The search strategy was conducted through Google Scholar, and restricted to academic papers only using the combinations of the following keywords: *migration* and *development*, *GDP*, and *labor mobility*. I went through the first 500 of them; however, many of these did not provide the direct effect of development on migration, or were theoretical researches. One fundamental problem facing the migration estimation is the limited accuracy and the hardship of obtaining long-term data on international migration. Moreover, it tends to present inaccurate reporting systems and does not have standard techniques for estimation. As a consequence, the data set was sparse. To expand it, it was essential to perform the so-called snowballing approach, that is, to examine the references of the primary studies previously collected that were not found in the first query. Thus, further studies were obtained additionally, including until the end of March 2020.

The principal selection criteria for identifying relevant studies to include in the meta-analysis were:

1. The reported variables must at least contain the level of GDP as part of the regression coefficients to estimate the effect on migration.

2. The selection of only articles written in English.

3. 3. The publication must be an official one, thus controlling the quality and having been previously subjected to careful analysis, as the inclu-

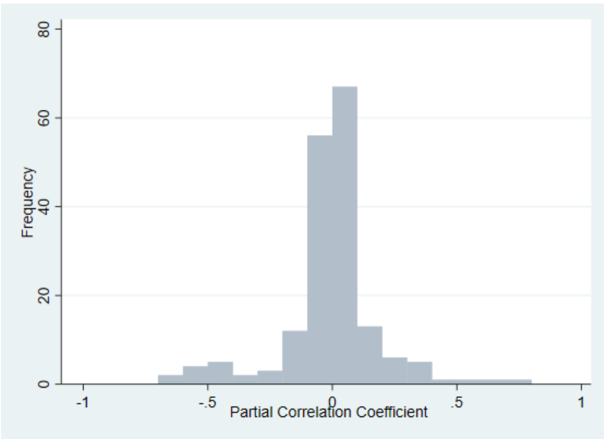
sion of low-quality estimates might taint the meta-analysis Stanley & Doucouliagos (2012).

4. The coefficients are captured either by standard errors or t-statistic.

5. The size of the effect is evaluated by the development gap, that is, how developed a country is, one way to determine this is through the GDP that measures the wealth of a country, reflecting the monetary value of all final goods and services produced by a country or region in a given period.

The final data set consists of 179 observations (effect sizes) from 40 different studies; see Appendix A for the list of the primary studies included in the data set. I have recalculated it to partial correlation coefficients with their respective standard errors and 29 potential explanatory variables. Figure 4.1 shows the distribution of the partial correlation coefficient (PCC) estimate of the effect size of GDP on migration reported in the primary studies. The corresponding mean reported is -0.01, implying that an increase in one point in the GDP coefficient estimate will result in 0.01% decrease in the annual immigration rate. In addition, it can be noted that zero effects are reported, demonstrating indeed that the empirical literature is distorted by selection bias.

Figure 4.1: Development effect on migration



Note: The figure represents the distribution of development on migration, particularly the effect that GDP has on immigration flows. The histogram contains values of the partial correlation coefficient (PCC) within the interval [-1,1].

There were still some outliers on the final data set, hence it was essential to winsorize the partial correlation coefficients with their standard errors at a 5% level on both distributions' side without biasing the results. It was applied to reduce the effect of possibly spurious outliers without affecting the main findings. Besides, the whole sample is restricted to English literature as "trials published in languages other than English tend to be of lower quality and produce more favorable treatment effects" Haidich (2010). Furthermore, articles that determine the direct effect of migration on GDP have been published in recent years, 70% of them published after 2005, the oldest one being published in 1997, and the most recent being published in 2020, thus encompassing 23 years of research.

In this meta-analysis, primary studies use immigration data mostly from European countries, the USA, and from the Organisation for Economic Cooperation and Development (OECD) regions, Figure 4.2 shows with more precision the distribution of what the data consists of, in regards to being an estimate for a developing country, a developed country or if it's for a collection for both. Besides this, studies use immigrant and emigrant populations data by country of origin or destination. On the whole, most of them considered the destination, see Figure 4.3. One of the reasons for this is that developed countries have a more accurate and extended database of international migration, which depicts global migration patterns and their characteristics, not only in receiving countries, but also in origin countries. The majority of primary studies use panel data; that is, 90 sources exploit the usage of both (cross-sectional and time-series data), 65 works with cross-sectional data, and only 24 with time-series analysis, see Figure 4.4 for a more detailed estimated distribution.

However, the heterogeneity of the estimates varies within and between individual studies. The boxplot is a useful tool to detect this issue graphically, helping to observe the tendency of the data set: range (minimum and maximum), first quartile (25%), third quartile (75%) and the median characterized by the vertical line in the center (see Figure 4.5).

Nevertheless, the tables below show the number of observations, with mean estimates and their 95% confidence interval, for specific groups such as data characteristics, regional database, non-economic and economic control variables for the development effect on migration across studies. It presents the standard

Figure 4.2: Stylized facts: Type of countries' development

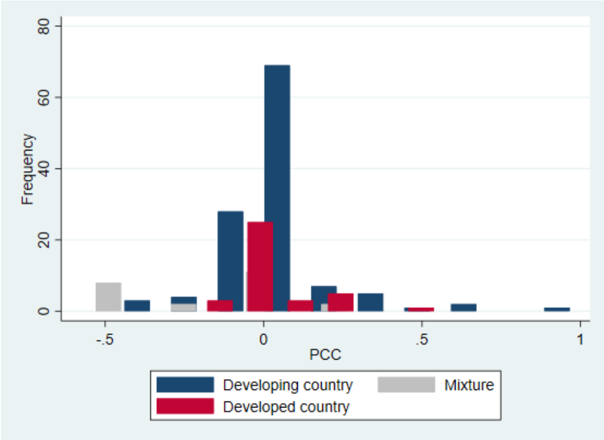


Figure 4.3: Stylized facts: Type of country

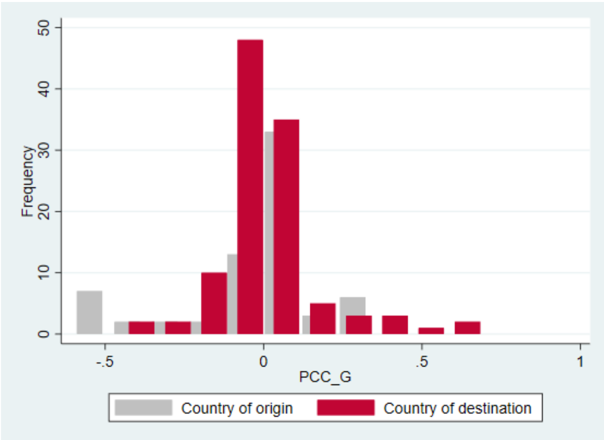


Figure 4.4: Stylized facts: Data specification

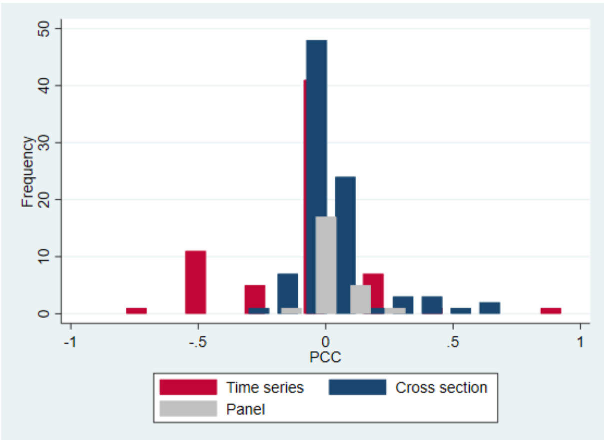
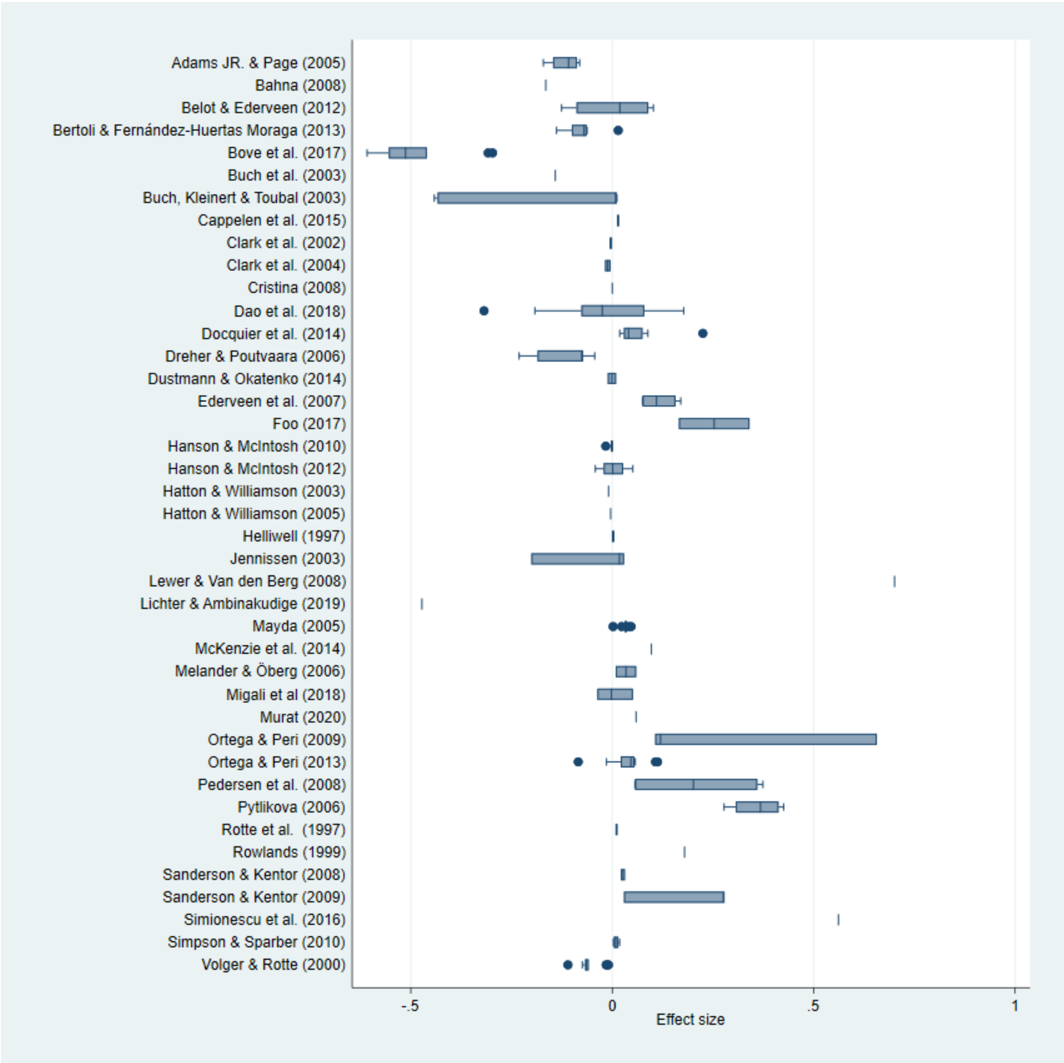


Figure 4.5: Estimates distribution of development on migration effect



Note:The figure depicts a boxplot of the estimates of development on migration effect reported for primary studies. The vertical line represents the median. It suggests large cross-study heterogeneity.

Table 4.1: Development effect on migration acrosss data characteristics

	Unweighted				Weighted		
	No. of est.	Mean	95%conf. int.		No. of est.	Mean	95% conf. int.
panel =1	90	0.023	-0.003	0.050	0.075	0.045	0.105
time series =1	24	0.037	-0.003	0.078	0.043	-0.001	0.088
cross-sectional =1	65	-0.078	-0.132	-0.024	-0.097	-0.155	-0.039

Note: The table presents the mean estimates of the effect of GDP on migration for specific data characteristics used in the primary studies. The confidence intervals around the mean are constructed using standard errors clustered at study-level. On the right-hand side (the weighted results) are estimated by using the inverse of the number of estimates reported per study.

errors clustered at the study level with previously winsorizing for potential outliers at [5%,95%] confidence interval. It shows the unweighted and weighted method. It is weighted by the inverse of the number of estimates reported per study, where each one obtains a weight in the same significance, see Havranek & Irsova (2017). The results presented above, depict that each primary study varies indeed due to the dissimilarities across the data.

The Table 4.1 depicts the differences across data estimation characteristics depending on the type of data. It can be noted that using a cross-sectional database of immigration inflows has a negative effect compared to a panel or time-series data analysis. The mean estimate yields the highest for time-series analysis for the unweighted case, while the panel data analysis mean estimate is the highest for the weighted specification.

According to the results of the Table 4.2 the effects vary according to the geographic area. The table depicts a positive average impact overall for OECD countries for both unweighted and weighted cases. Additionally, if the studies rely on developing databases, they also have a positive average effect but are slightly lower than for the OECD, whereas the lowest average that yields a negative impact overall is for European countries. The analysis of developed countries also gives a negative mean effect, but it is significantly lower than for the European countries. And the mean estimate has a much lower negative mean impact for the United States.

Table 4.2: Development effect on migration acrosss country regions

	Unweighted				Weighted		
	No. of est.	Mean	95% conf. int.		No. of est.	Mean	95% conf. int.
OECD=1	65	0.058	0.025	0.091	0.125	0.089	0.162
EUR=1	28	-0.084	-0.154	-0.014	-0.098	-0.199	0.001
USA=1	23	-0.035	-0.073	0.001	-0.004	-0.046	0.038
developing=0	156	0.011	-0.011	0.033	0.018	-0.010	0.047
developed=1	23	-0.165	-0.277	-0.053	-0.014	-0.102	0.072

Note: The table presents the mean estimates of the GDP's effect on migration for a particular country group used in the primary studies - including if it is considered to be a developing country or developed country. The confidence intervals around the mean are constructed using standard errors clustered at study-level. On the right-hand side (the weighted results) are estimated by using the inverse of the number of estimates reported per study.

Table 4.3 presents the non-economic drivers of the development effect on migration of dummy variables used in the regression. It is observed that countries sharing a colonial tie (colonial link) have the highest positive average mean among all the variables, either for the unweighted or weighted technique. Meaning that where the countries share a link in the past this has a strong effect; the effect of migrating will be higher. It is also important to note that destination is also a critical variable, having a negative average mean in the unweighted case, whereas it has a positive mean effect in the weighted one. Thus the distance between the home country and the country of destination matters. The level of education of the country of origin is associated with a negative mean effect in the unweighted case while it has a positive mean effect in the weighted one. The political situation is also a crucial variable, meaning that migrants care about the policy structure of the desired destination country, supporting a more open one. The results show that it has a negative mean effect for both (weighted and unweighted).

Lastly, Table 4.4 presents the estimates of the effect of development on migration drivers of the economic control variables from the estimated regression in the literature. It can be noted that FDI has a positive and the highest impact from explanatory variables across the research. In addition, income is the explanatory variable that is most commonly used to measure the effects of

Table 4.3: Development effect on migration acrosss non-economic dummy variables

		Unweighted			Weighted		
	No. of est.	Mean	95% conf.	int.	Mean	95% conf.	int.
age=1	62	0.018	-0.008	0.044	0.018	-0.002	0.039
married=1	11	0.001	-0.053	0.056	0.003	-0.036	0.037
population=1	148	-0.021	-0.050	0.007	-0.002	-0.033	0.028
education=1	93	-0.019	-0.062	0.022	0.024	-0.017	0.067
distance=1	130	-0.025	-0.055	0.004	0.011	-0.017	0.040
colonial_link=1	61	0.039	0.009	0.069	0.083	0.047	0.119
networks=1	88	-0.060	-0.098	-0.021	-0.026	-0.063	0.010
political_situation=1	77	-0.066	-0.112	-0.021	-0.030	-0.075	0.014

Note: The table presents the mean estimates of the effect of GDP on migration of the non-economic variables used in the primary studies. The confidence intervals around the mean are constructed using standard errors clustered at study-level. On the right-hand side (the weighted results) are estimated by using the inverse of the number of estimates reported per study.

Table 4.4: Development effect on migration acrosss economic dummy variables

		Unweighted			Weighted		
	No. of est.	Mean	95% conf.	int.	Mean	95% conf.	int.
income=1	104	-0.010	-0.046	0.026	0.036	0.002	0.071
FDI=1	21	0.042	-0.051	0.135	0.066	-0.016	0.150
Gini=1	23	-0.006	-0.066	0.054	-0.061	-0.151	0.028
unemployment=1	57	-0.007	-0.052	0.037	-0.008	-0.061	0.043
labor=1	75	0.006	-0.027	0.040	0.021	-0.024	0.068
trade	41	0.014	-0.038	0.067	0.069	0.013	0.125

Note: The table presents the mean estimates of the effect of GDP on migration of the economic variables used in the primary studies. The confidence intervals around the mean are constructed using standard errors clustered at study-level. On the right-hand side (the weighted results) are estimated by using the inverse of the number of estimates reported per study.

migration. Still, it was a negative mean effect in the unweighted case, while weighting by the inverse of the number of estimates results in a positive average mean. It can be seen that trade and labor have a significantly lower positive average mean but remain positive for the unweighted and weighted case. And unemployment has the lowest average negative mean effect among the economic variables used in the regressions in primary studies.

Table 4.5 provides and explains the explanatory variables included in the present work, which reflect the essential features measuring the effect of migration, helping to control heterogeneity between estimates within primary studies. The potential variables are divided into the following classification: non-economic drivers of migration, economic drivers of migration, data and estimation characteristics, the different country groups selected in the sample, and the publication characteristics. The correlation matrix of these variables is presented in Appendix C.

Table 4.5: Description and summary statistics of explanatory variables

Variable	Definition	Mean	SD
TSTAT	the t-statistic estimate of the effect size	0.826	9.49
PCC	the partial correlation coefficient of the estimate	-0.0104	0.202
Precision	precision of the estimated PCC (the inverse of the standard error)	54.069	53.89
Non-economic drivers of migration			
Age	dummy =1 if the regression estimates the age share of the population, 0 otherwise	0.346	0.477
Marriage	dummy =1 if the regression estimates current marital status of the individual, 0 otherwise	0.061	0.24
Population	dummy =1 if the regression estimates share of population size, 0 otherwise	0.826	0.379
Education	dummy =1 if the regression estimates level of education of migrants by years of schooling, 0 otherwise	0.519	0.501
Distance	dummy =1 if the regression estimates geographical distance between the country of origin and contry of destination, 0 otherwise	0.726	0.447
Colonial link	dummy =1 if the regression estimates colonial relationship between the country of origin and country of destination, 0 otherwise	0.341	0.475
Networks	dummy =1 if the regression estimates presence of some family or friends between the country of origin and country of destination, 0 otherwise	0.491	0.501
Political situation	dummy =1 if the regression estimates domestic political structure and policies of the home country(country of origin), 0 otherwise	0.43	0.496
Economic drivers of migration			
Income	dummy =1 if the regression estimates income of individual household per capita reported in country of origin, 0 otherwise	0.581	0.495
FDI	dummy =1 if the regression estimates long-term accumulation of the stocks of foreign direct investment, 0 otherwise	0.117	0.323
Gini coef.	dummy =1 if the regression estimates the measure of inequality level is the proxy for the spurce country poverty rate, 0 otherwise	0.128	0.335
Unemployment	dummy =1 if the regression estimates level of unemployment rate in the country of origin (the difficulty of finding a job), 0 otherwise	0.318	0.467
Labor	dummy =1 if the regression estimates labor mobility between the country of origin and country of destination, 0 otherwise	0.419	0.495
Trade	dummy=1 if the regression estimateslevel of international trade values (imports and exports) for country pairs	0.229	0.421
Data and estimation characteristics			
Panel	dummy =1 if data-set is panel, 0 otherwise	0.503	0.501
Time-Series	dummy =1 if data-set is time-series, 0 otherwise	0.134	0.341
Cross-section	dummy =1 if data-set is cross-section, 0 otherwise	0.363	0.482
Time-span	logarithm of the number o fyears used in the sample	2.874	0.603
No.of obs	logarithm of the number of observations used in the sample	7.157	1.786
No. of variables	logarithm of the number of variables used in the sample	2.027	0.631
Endogeneity	dummy =1 if primary studies accounts for endogeneity, 0 otherwise	0.368	0.483
OLS	dummy =1 if the estimation method accounted for OLS, 0 otherwise	0.782	0.414
FE	dummy =1 if the estimation method accounted for fixed effects, 0 otherwise	0.759	0.428
Country of origin	dummy=1 if the data-set accounts from country of origin (home country), 0 otherwise	0.379	0.486
Developing	dummy=2 if the estimation is from developing country, =1 if the estimation is a combination of developing and developed, countries, =0 if the estimation is from a developed country	0.466	0.714
Regions			
OECD	dummy=1 if the data-set is from the OECD region, 0 otherwise	0.363	0.482
EUR	dummy=1 if the data-set is from european countries, 0 otherwise	0.156	0.364
USA	dummy=1 if the data-set is from the United States, 0 otherwise	0.128	0.335
Publication characteristics			
Citations	logarithm of the number of citations in Google Scholar	4.754	1.197
Impact	recursive impact factor from RePEc, 0 in case the journal was not included	0.389	0.382
Pub. year	logarithm of the year the article was published	7.605	0.003

Note: The tables presents the definitions and summary statistics of the regression variables. Mean= simple arithmetic estimate means. SD=Standard Deviation.

Chapter 5

Results

5.1 Mean effect estimation

As previously mentioned in the methodology section, the estimation of the effect of development on migration was conducted by the use of the partial correlation coefficient (PCC), since the primary studies diverge on the estimation of migration effect. The average simple mean for the partial correlation coefficient (PCC) is -0.01, indicating that there might not be an effect of development on migration. However, since the coefficients suffer from some limitations, like selection bias, and they (coefficients) do not consider the precision of the estimate itself, that is, that the estimates will have different variances, it was indeed crucial to perform the fixed-effects and random-effects models. First, the fixed-effects model obtained an average of 0.027; that is, weighting the partial correlation coefficients (PCC) by the inverse of their variance of the corresponding estimate, considers within heterogeneity. Moreover, the random-effects estimated an average effect of -0.005, accounts for between-study heterogeneity. According to the preliminary and relative guidelines for interpreting partial correlation in economics Doucouliagos (2011), it implies that the results suggest no effect.

Table 5.1 presents results that may be biased. Either as a consequence of publication selection bias or selective reporting, the estimates may suffer for the interpretation of the statistical inference. And the estimates also do not account for the heterogeneity derived from the different methodologies of individual studies.

Table 5.1: Partial correlation coefficients of the effect of development on migration

	PCC	95% Conf. Interval	
Average mean	-0.011	0.013	-0.036
Fixed-effects	0.027	0.025	0.029
Random-effects	-0.005	-0.019	0.009
Number of obs	179		

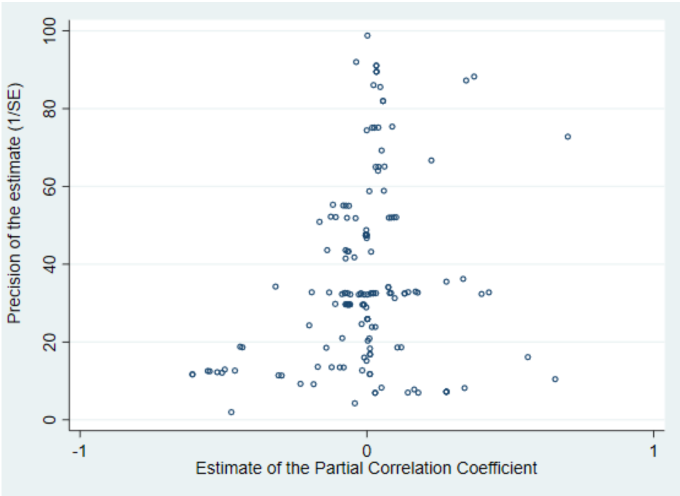
Note: The table presents the partial correlation coefficient (PCC) estimated on the impact of the migration-development nexus. The average mean represents a simple arithmetic mean of the effect size of GDP on migration. The fixed-effects estimate depicts the partial correlation coefficient (PCC) weighted by the inverse of their variance. The random-effects estimate depicts the partial correlation coefficient (PCC) weighted by the inverse of their variance, considering the heterogeneity among primary studies.

5.2 Publication bias

There are several ways to detect the so-called publication bias. The easiest and standard method to detect publication bias in the primary studies is the funnel plot. It plots the size effects of the estimates and the measure of the estimates' precision. Figure 5.1 depicts the funnel plot of the collected estimates of the development effect on migration, transformed to partial correlation coefficient (PCC). The horizontal axis lines represent the corresponding mean and median of all the estimates, and the vertical line depicts the precision of the partial correlation coefficients (PCC), obtained by the inverse of the standard errors. The figure shows that, to some extent, the right-hand side seems to be denser; thus it may indicate that primary studies tend to have a preference in reporting the positive effect of development on migration. Also, most of the estimates appear to be somewhat clustered at the bottom; this may be caused because the researchers prefer to publish significant results.

Moreover, those with extreme values are omitted for the purpose of a better visual inspection. Nevertheless, there are three outliers where the partial correlation coefficient is almost 0.25, and the estimate's precision is circa 80 which indicates a systematic difference between those studies. This may perhaps be from an inappropriate effect measure on the primary studies. However, the estimates that are included in the funnel plot suggest the presence of publication bias in the sample, mainly when the scatter is asymmetrical, but "symmetry may be more in the eye of the beholder than the actual research record itself" Stanley (2005). Therefore, it is necessary to conduct tests more rigorous than funnel plots, with a more formal statistical analysis.

Figure 5.1: Funnel plot of the development effect on migration



Note: The figure depicts a funnel plot of the GDP effect estimates on migration reported on primary studies, transformed into partial correlation coefficients (PCC) winsorized at 5% level on both distrbutions' side.

5.2.1 Identifying publication selection: funnel asymmetry test and precision effect test

A more formal technique of asymmetry of the funnel plot is known as regression-based funnel asymmetry tests (FAT-PET). The regression equation previously mentioned in chapter three is Equation 3.4 , which will help determine and correct in the case of the presence of publication bias and the underlying effect of development on migration. In the case where there is no publication bias, the partial correlation coefficients (PCC_{ij}) will be randomly distributed around the 'true' effect and are independent of their standard errors ($SE_{PCC_{ij}}$). Thus it is estimated through Equation 3.4 by using several methods. The results are provided in the table below.

Table 5.2: Tests for publication bias on the effect of development on migration

	(1) OLS	(2) FE	(3) BE	(4) Precision	(5) Study	(6) IV
Constant (effect beyond bias)	-1.313	1.886	-0.260	-0.457	-0.220	-1.464
	(0.77)	(1.50)	(1.26)	(1.19)	(0.79)	(1.05)
SE(publication bias)	0.044***	-0.016	0.041*	0.036***	0.040*	0.047**
	(0.01)	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)
Observations	179	179	179	179	179	179
Number of studies	40	40	40	40	40	40

Note: The table above represents the results from the regression $PCC_{ij} = \beta_0 + \beta_1 SE_{PCCij} + u_{ij}$ *where the dependent variable is the partial correlation coefficient (PCC). In the specification (1) OLS (ordinary least squares) is used. Following in specification (2), FE study-level fixed-effects. The next specification (3) is BE study-level between-effects. Followed by specification (4) where Precision is estimated by WLS (weighted least squares) where the weight is taken as the inverse of the estimates' standard error. Next, the specification (5) Study is also estimated by WLS where the weight is the inverse of the number of estimates. Lastly, specification (6) IV (instrumental variable) where the instrument is taken as the inverse of the square root of the number of observations. The standard errors are reported in paretheses clustered at study level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level.*

The Table 5.2 explains the variation across the studies in the effect sizes, that is, correcting for heteroskedasticity. In the first column, a simple OLS regression model estimates the partial correlation coefficient (PCC) on its standard error to the baseline with unweighted regressions. The β_1 coefficient measures the publication bias is positive and significant at 1% level, thus implying a strong selective reporting bias, and the estimated constant, that is, β_0 (the 'true' effect corrected for publication bias) is negative but insignificant. However, the database is heterogeneous due to different variables, methodologies, and time periods considered in the primary studies. Thus, the second column reports fixed effects, which appears to be more convenient. It helps detect discrepancies at a study-level specification, differing in sign from a positive to negative 'true' effect (β_1). This specification is the only one where the publication bias remains insignificant and has a positive magnitude.

Furthermore, between-effects estimation is also provided, contributing to a more robustness check, where the mean is strongly affected but remains a positive 'true' effect significant at 10% level. The fourth column presents the estimates regressed by weighted least squares (WLS), where the precision variable is used as a weight, that is, by the inverse of their standard errors, $(1/SE_{PCC_{ij}})$ statistically significant at 1% level and the mean remains almost unaffected. The fifth column depicts IV (instrumental variable) regression, calculated by the inverse of the square root of the number of observations across the literature, where it can be noted that the estimate for publication bias is positive and significant at 10% level. Lastly, the sixth column presents the estimates weighted by the inverse of the number of estimates reported per study in the regression. The results suggest that the estimates may be biased towards the positive effects of development on migration effect at 1% significance level, which means that the estimates report an even larger and selective reporting bias and insignificant negative size effect.

To summarize, the results mentioned above depict the presence of publication bias in the partial correlation coefficient (PCC), thus indicating to have a positive and, in amongst most of the different methodologies, significant publication selection; thus, a negative 'true' mean. Consequently, the null hypothesis, H_0, from Equation 3.5 is rejected, thus empirical literature tends to report positive effects of GDP on migration, that is, have a presence of bias towards positive and significant estimates. Furthermore, corresponding to the guidelines for interpreting partial correlations Doucouliagos (2011), the β_0 is insignificant in all the model specifications, which means that the effect correcting for publication selection bias varies within the technique employed, but in most of them is having an adverse effect, while the β_1 coefficient suggests a small and positive effect under the between-effect estimation, which has a negative magnitude.

Moreover, the regressions estimated in Table 5.2 assume a linear relationship between the publication bias and the standard error. In case the assumption does not hold, the estimates will still be biased. Thus it is essential to perform non-linear techniques to correct for publication selection. The first method is the weighted average of the adequately powered estimates (WAAP), in which the average development effect corrected is 0.021. This correction takes into account just the estimates that are considered to have sufficient power as previously mentioned in the methodology section; this means the ones with statistical power greater than 80%, see Ioannidis *et al.* (2017). Another technique is the so-called "top 10", which gives a mean corrected effect of 0.031, which is quite close to the results from WAAP. Therefore, according to the guidelines for interpreting partial correlations Doucouliagos (2011) these results suggest that there is little evidence of the effect of development, as GDP, on migration, once publication bias is well addressed by non-linear techniques.

5.2.2 Identifying heterogeneity

Once publication bias is identified, the next step is to address the heterogeneity of development on migration effect estimates that arose from the literature. Since primary studies vary across the literature, depending on the study characteristics: sample selection, time-period, conditional variables, or estimation techniques, it is essential to detect the crucial variables that reflect the estimated effect of development on migration, identifying the differences among the estimated coefficients. The best model that estimates migration's effect in the present study has 29 potential explanatory variables, thus computing all the feasible models, which will result in 2^{29} model combinations. It is important to mention that the total conditional variables were 31. However, two variables were excluded from the Bayesian model averaging (BMA) regression model since they were correlated with each other (panel and time-series in case the data set presented those characteristics). The correlation matrix of all the conditional variables included is presented in Appendix C. In practice, computing all those possible models will result in model uncertainty, and it will be computationally unfeasible. Hence, the Bayesian model averaging (BMA) is performed to address this issue, which averages the 'best models' according to the posterior model probability (PMP). The BMS package in R will be employed to estimate the Bayesian model averaging (BMA) analysis. This relies on a Markov Chain Monte Carlo (MCMC) sampler, which combines the essential posteriors distribution results using the Metropolis-Hastings algorithm. The Bayesian model selection (BMS) arguments were selected as follows:

- *burn*: the number of burn-in draws, in this case, was set to 1 000 000 (a positive integer).
- *iter*: the number of iterations to be sampled, in this case, was set to 2 000 000 draws.

- *nmodel*: the number of best models for which information is stored set to 5 000 models.
- *mcmc*: the argument stands for the model sampler to be used, where mcmc="bd" corresponds to the birth-death MCMC algorithm.
- *g*: the g-prior for the regression coefficients was set to g="UIP" corresponds to the number of observations, in this case, 179 observations.
- *mprior*: the argument denotes the model prior choice, set to mprior="uniform." Thus, employing the uniform model prior, see Zeugner & Feldkircher (2015) for a detailed review.

Figure 5.2 depicts the Bayesian model averaging (BMA) results, which are estimated on the winsorized sample at 5% level (see Data description). The horizontal axis of the plot displays all the 29 potential variables included in the regression model. The variables are distributed according to their relevance in the model, that is, from the highest posterior inclusion probability (PIP) to their lowest PIP. Graphically they will be depicted from the left to the right cells. The vertical axis depicts their posterior model probabilities (PMP), from the top to the bottom of the cells. The blue color corresponds to a positive coefficient of the explanatory variable included in the model (darker in greyscale), and the red color (lighter in greyscale) illustrates a negative impact on the estimated development effect on migration. Moreover, the white color corresponds to the variable not included in the regression model, meaning this is a zero coefficient.

Table 5.3 presents the numerical results of the Bayesian model averaging BMA analysis, reflected in the first three columns. It is essential to mention that the BMA baseline is weighted by the inverse of the number of estimates reported per study. It depicts the most prominent explanatory variables that reflect the effect of GDP on migration, and the following columns show the posterior mean (post mean), posterior standard deviation (post SD), and the posterior inclusion probabilities (PIP) for each of the variables. Followed by the next three columns, which report the frequentist model averaging (FMA) for more robustness checks. It shows the coefficient estimate (coef.), standard error (std. err.), and p-value. The posterior inclusion probability (PIP) indicates whether a conditional variable must be included or not in the 'true model' of the estimated effect of development on migration, that is, how significant are the regressors' variables from the baseline model. Thus, to interpret the posterior inclusion probability (PIP) of the estimated magnitudes of the regression parameters, according to the approach of Eicher *et al.* (2011) values of the PIP between 0.5 and 0.75 are considered to have a weak significance, between 0.75 and 0.95 are substantial, between 0.95 and 0.99 have strong significance and values higher than 0.99 are decisive. Thus, concluding the following remarks:

- *Marital-status*: The inclusion of the variable civil status into the regression baseline model appears to be a decisive factor, with a PIP value of 0.999. The results suggest a negative impact, implying, on average, a 0.489 effect size coefficient. Since the dummy variable =1, if the individual set to migrate is currently married, it would seem to be a 'right' outcome, as if an emigrant is married, they will be less likely to migrate. Nevertheless, the marital-status variable is not often used in measuring the effect of development on migration; with just 11 out of 179 observations including this conditional moderator; overall, only 6% of the estimates fall into this category. Thus, this result seems not be reliable enough. Additionally, after performing the frequentist check, it supports the findings, as the estimated regression coefficient is positive, contradicting BMA results.

- *Age*: The migrating individual age has a positive average effect of 0.245 influencing the mobility transition, and the PIP implies a decisive significance value of 0.997. Most of the models trying to explain migration's effect consider this explanatory variable to have a significant impact on migration patterns over time; additionally, the migration behavior during the life cycle of migrants usually starts from age 18, then reaches the maximum between the ages of 25 and 30, and then decreases. Hence, to confirm the BMA results, another robustness check is performed, that is, the frequentist check OLS, with a significant and positive response effect.

- *OECD*: The inclusion for country groups in the regression, particularly the Organization for Economic Co-operation and Development (OECD) countries, has a high PIP value of 0.997, being a decisive variable for measuring GDP effect on migration. It produces a positive effect size value becoming, on average, 0.282 higher. Generally, OECD countries have a more extensive and precise inflows of migration movements database and support of free-market economies.

- *FDI*: The variables in primary studies considering the foreign direct investment (FDI) flows yields a positive and decisive estimated effect with a PIP value of 0.995. This finding yields, on average, 0.342 value to have larger coefficients. Thus, indicating that long-term investment is positively related to the presence of migrants in terms of economic contributions. This means that the presence of immigration boosts investment flows between countries of origin and destination countries, lowering information barriers.

- *Fixed-effects*: The inclusion probabilities for the control variables accounting for fixed-effects have a strong impact of 0.997. When weighted, the reported estimate yields, on average, a negative 0.222 effect size estimates. Thus, to check the BMA results, according to the frequentist check performed by OLS, the results seems to be consistent as both imply a negative response variable. Hence, it means that relying on fixed-effects regression models reduces the omitted variable problem (raised from the variations from independent and dependent variables). If not conducted right, this may result in a negative effect in terms of the outcome.

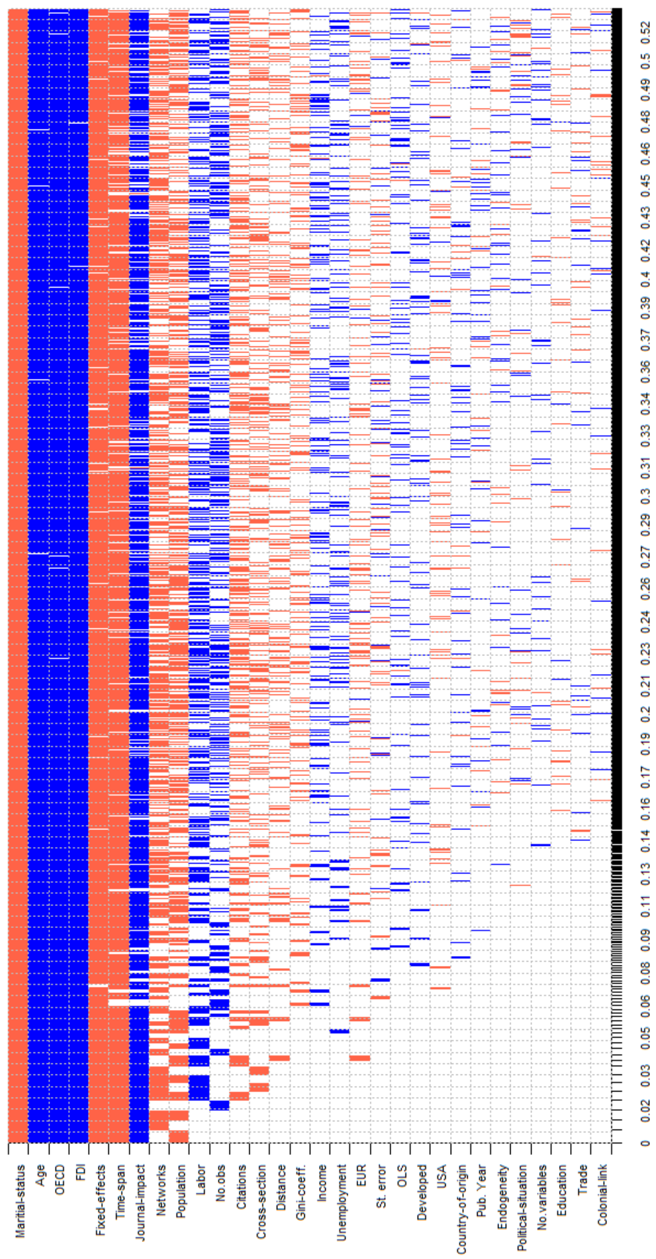
- *Time-span*: : The variables indicating the time-span, which is obtained by the logarithm of the number of years used in the sample, resulted in being a substantial PIP of 0.919. This might be because most of the migration data is taken from a national census, which is conducted at least every five or ten years, where the data is sparse or not reliable. Moreover, it generates, on average, a negative estimate of 0.155 according to BMA results. Thus, it contradicts the frequentist check yielding from negative to positive effect size; implying the results are not reliable enough.

- *Journal-impact*: The inclusion of the journal recursive impact factor, on average suggests the production of effect size estimates higher than 0.091. The PIP has a substantial value of 0.850, providing evidence that the journal's quality impacts the baseline model effect size. Thus, it helps as an evaluation tool that reflects the prestige associated with its ranking across different countries.

- *Networks*: The variables indicating the network link's inclusion are weak, with a PIP value of 0.559, and yields to a negative coefficient. Hence, some family or friends' presence in the destination country might predict some negative externalities not previously considered by migrants

The remaining explanatory variables included in the regression baseline specification for the BMA depicted a low posterior inclusion probability (PIP). Hence, they might not be considered essential to measure the effect size of development on migration estimates and its variation in those variables' responses. Thus, to check the BMA results' reliability, the frequentist OLS check was performed, but only for those variables where the PIP value was higher than 0.5. The regression coefficients' signs and their magnitude in most of the variables confirm the BMA analysis, apart from marital-status and journal impact, differing in the coefficients' sign but remaining significant.

Figure 5.2: Model inclusion in Bayesian Model Averaging



Note: The figure plot depicts the response variable of the effect size estimate of development on migration. The data is winsorized at 5% level on both distributions' side. All regressions are weighted by the inverse of the number of estimates reported per study. The columns denote individual models. The vertical axis depicts variables sorted from descending order by the posterior inclusion probability (PIP). Blue color (darker in greyscale)=the variable is included, and the estimated effect is positive. Red color (lighter in greyscale)=the variable is included, and the estimated effect is negative. White color=the variable is not included in the model. The horizontal axis depicts the cumulative posterior model probabilities (PMP).

Table 5.3: Heterogeinety of the effect of development on migration

	BMA			Frequentist check (OLS)		
	Post Mean	Post SD	PIP	Coefficient	Std. Err	pvalue
St. error(publication bias)	-0.007	0.316	0.122	-0.003	0.0059	0.611
Age	0.245	0.067	**0.997**	0.023	0.003	0.000
Maritial-status	-0.489	0.130	**0.999**	0.044	0.009	0.000
Population	-0.086	0.103	0.499			
Colonial-link	-0.001	0.015	0.038			
Networks	-0.072	0.075	**0.559**	-0.037	0.004	0.000
Political-situation	0.000	0.018	0.049			
Education	-0.001	0.012	0.041			
Labor	0.049	0.066	0.427			
Unemployment	0.013	0.043	0.131			
Income	0.016	0.044	0.162			
Trade	0.000	0.015	0.037			
FDI	0.342	0.103	**0.995**	0.005	0.005	0.298
Gini-coeff.	-0.025	0.066	0.176			
Country-of-origin	0.003	0.025	0.066			
Developed	0.004	0.018	0.080			
OECD	0.282	0.062	**0.997**	0.061	0.008	0.000
EUR	-0.017	0.061	0.123			
USA	-0.007	0.031	0.078			
Distance	-0.019	0.051	0.175			
Cross-section	-0.026	0.054	0.235			
Time-span	-0.155	0.070	**0.919**	0.012	0.003	0.000
No.obs	0.015	0.024	0.360			
No.variables	0.003	0.028	0.045			
Pub. Year	0.001	0.019	0.054			
OLS	0.006	0.027	0.085			
Fixed-effects	-0.222	0.079	**0.977**	-0.007	0.004	0.111
Endogeneity	0.001	0.006	0.051			
Citations	-0.038	0.063	0.337			
Journal-impact	0.091	0.050	**0.850**	-0.006	0.003	0.065
Constant	-0.157	NA	1.000	0.120	0.014	0.000

Note: The table presents the Bayesian model averaging (BMA) results, PIP=Posterior Inclusion Probability. PIPs above 0.5 are highlighted in bold. SD= Standard Deviation. The results for the Frequentist check include PIP values of the explanatory variables higher than 0.5, according to BMA. Std. Err=Standard Error clustered at study level.

Chapter 6

Conclusion

The present thesis focuses on understanding the impact of development on migration patterns, which shows a U-inverted shape relationship. In order to assess the economic development of a country, the observed GDP per capita was considered as an explanatory variable. The existing theoretical and empirical literature on migration systems has not reached any conclusive results since it differs in the direction of the effect, and results vary both within and across studies. Clemens *et al.* (2014) suggests that more empirical research needs to be developed to have a better and more precise understanding of this relationship. It is also essential to determine the reasons why the results vary so much between the primary studies. For these reasons, I was interested in conducting the first meta-analysis of the migration-development nexus to have more quantitative findings.

Meta-analysis technique is a powerful tool that helps to correct publication bias inherent in the migratory movement's literature, and it measures the effect size and the causes of heterogeneity across the primary studies rising due to differences in data, study design, and variability across these studies. The final collected data set consists of 179 effect sizes from 40 different articles. Since the coefficient estimates were not directly comparable because they use diverse proxy variables, they were transformed into partial correlation coefficients (PCC). Thus, according toDoucouliagos (2011), the preliminary guidelines for interpreting partial correlation coefficients (PCC), the overall effect of GDP on migration is negative but with no effect after combining the findings from the primary studies.

However, the empirical literature on immigration flows suffers from publication bias. According to the present work results, the funnel plot depicts evidence of the slight preference for reporting positive effect results of the migration-development nexus. Moreover, more formal techniques than a graphical approach were conducted to confirm the publication selection. The funnel asymmetry and precision effect tests indicated publication bias as reported by several test specifications. In general, publication bias turned out to be positive, except for in the study level fixed-effects (FE) test, which reported a negative but not significant effect. It should be noted that there is no clear evidence of the existence of a direct impact of development on migration. Since the reported β_0 gives insignificant results on the entire test specifications (the 'true' effect corrected for publication bias). There is usually no relationship in which publication selection is a linear function of the standard errors; thus, additional non-linear techniques were employed, yielding similar outcomes, that is, a positive effect of GDP on the migration phenomenon. Therefore, it can be concluded that there is presence of publication bias and over-estimation in the primary studies size effects, thus corroborating hypothesis #3 towards positively biased literature of the development-migration nexus. This conclusion reinforces the migration transition theory, and the results of Sanderson & Kentor (2009), McKenzie *et al.* (2014), Bahna (2008), and Jennissen (2003) suggesting that GDP has a positive effect on international migration but from the present analysis, this is rather small. Nevertheless, there is no proof to indicate that there is an actual direct effect of GDP on migration.

Furthermore, to explain the heterogeneity across studies from the primary literature, the Bayesian model averaging (BMA) analysis was applied. The BMA baseline was weighted by the inverse of estimates reported per study addressing the uncertainty's model specification. Hence, the main result suggests that controlling for the variables of age and foreign direct investment is fundamental when precisely estimating the effect of GDP on migration, which will boost migration flows. Moreover, the presence of country-level dummies produces a decisive factor and higher effect size estimates. Specifically, in samples for the OECD region, developed countries will see higher immigration rates as individuals seek better living conditions. It also matters that primary studies control for fixed-effects that yields to lower the effect on the GDP variable's impact on migration regressions. Lastly, the control variables indicating the data structure, such as journal recursive impact factor, tend to report more significant positive effects. At the same time, the length of the time-span used in the sample has a negative size effect in the estimates.

To conclude, even if the theoretical literature argues that rising incomes in developing countries will diminish immigration, it is observed that development in terms of the GDP does not stimulate migration significantly which, as a consequence, rejects hypothesis #1 (economic growth stimulates migration). Furthermore, income per capita and unemployment rates turn out to be insignificant factors in the BMA baseline model, thus rejecting hypothesis #2 (the higher the average per capita income and unemployment rate, the greater the immigration rate). The migration-development nexus might not be so straightforward and perhaps the direction of the effect of development on migration is vice versa. Moreover, the mechanisms regarding the impact of development on migratory movements still need to be explored. There must be other composition factors that weigh into the decision to migrate, explaining development in terms of the human development index (HDI), literacy rates, and migrants' remittances, which could have a significant role since they affect economic growth. Further empirical research is required to produce more comprehensive results, leaving an open question, particularly for policy makers, which may well boost the economic growth of both the countries of origin and the destination countries creating new migratory and policy challenges.

Bibliography

Andrews, I. & M. Kasy (2019): "Identification of and correction for publication bias." *American Economic Review* **109(8)**: pp. 2766–94.

Bahna, M. (2008): "Predictions of migration from the new member states after their accession into the european union: Successes and failures." *International Migration Review* **42(4)**: pp. 844–860.

Bertoli, S. & J. F.-H. Moraga (2013): "Multilateral resistance to migration." *Journal of development economics* **102**: pp. 79–100.

Bom, P. R. D. & H. Rachinger (2019): "A kinked meta-regression model for publication bias correction." *Research Synthesis Methods* **10(4)**: pp. 497–514.

Bove, V. & L. Elia (2017): "Migration, diversity, and economic growth." *World Development* **89**: pp. 227–239.

Buch, C. M., J. Kleinert, & F. Toubal (2006): "Where enterprises lead, people follow? links between migration and fdi in germany." *European Economic Review* **50(8)**: pp. 2017–2036.

Buckland, S. T., K. P. Burnham, & N. H. Augustin (1997): "Model selection: an integral part of inference." *Biometrics* pp. 603–618.

Card, D. & A. B. Krueger (1995): "Time-Series Minimum-Wage Studies: A Meta-analysis." *American Economic Review* **85(2)**: pp. 238–243.

Cazachevici, A., T. Havranek, & R. Horvath (2020): "Remittances and economic growth: A meta-analysis." *World Development* **134**: p. 105021.

Clemens, M. A., Ç. Özden, & H. Rapoport (2014): "Migration and development research is moving far beyond remittances." *World Development* **64**: pp. 121–124.

CRISTINA, A. D. (2008): "What sways the decision to migrate? an empirical analysis of the argentinean case." *Revista de Economia y Estadistica* **46(1)**: pp. 7–29.

DAO, T. H., F. DOCQUIER, C. PARSONS, & G. PERI (2018): "Migration and development: Dissecting the anatomy of the mobility transition." *Journal of Development Economics* **132**: pp. 88–101.

DOCQUIER, F., G. PERI, & I. RUYSSEN (2014): "The cross-country determinants of potential and actual migration." *International Migration Review* **48(1_suppl)**: pp. 37–99.

DOUCOULIAGOS, H. C. (2011): "How large is large? preliminary and relative guidelines for interpreting partial correlations in economics." .

DREHER, A. & P. POUTVAARA (2006): "Student flows and migration: An empirical analysis." *Institute of Labor Economics (IZA)* .

EDERVEEN, S., R. NAHUIS, & A. PARIKH (2007): "Labour mobility and regional disparities: the role of female labour participation." *Journal of Population Economics* **20(4)**: pp. 895–913.

EGGER, M., G. D. SMITH, M. SCHNEIDER, & C. MINDER (1997): "Bias in meta-analysis detected by a simple, graphical test." *BMJ* **315(7109)**: pp. 629–634.

EICHER, T. S., C. PAPAGEORGIOU, & A. E. RAFTERY (2011): "Default priors and predictive performance in bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

FOO, G. (2017): "Quantifying the malaysian brain drain and an investigation of its key determinants." *Malaysian Journal of Economic Studies* **48(2)**: pp. 93–116.

FURUKAWA, C. (2019): "Publication Bias under Aggregation Frictions: Theory, Evidence, and a New Correction Method." *EconStor Preprints 194798*, ZBW - Leibniz Information Centre for Economics.

GAGNON, J. (2014): "Demographic change and the future of the labour force in the eu27, other oecd countries and selected large emerging economies." .

Gechert, S., T. Havránek, Z. Havránková, & D. Kolcunova (2019): "Death to the cobb-douglas production function." *Technical report*, IMK Working Paper.

Glass, G. V. (1976): "Primary, secondary, and meta-analysis of research." *Educational researcher* **5(10)**: pp. 3–8.

Haidich, A.-B. (2010): "Meta-analysis in medical research." *Hippokratia* **14(Suppl 1)**: p. 29.

Hansen, B. E. (2007): "Least squares model averaging." *Econometrica* **75(4)**: pp. 1175–1189.

Hanson, G. H. & C. McIntosh (2010): "The great mexican emigration." *The Review of Economics and Statistics* **92(4)**: pp. 798–810.

Hatton, T. J. & J. G. Williamson (2003): "Demographic and economic pressure on emigration out of africa." *Scandinavian Journal of Economics* **105(3)**: pp. 465–486.

Hatton, T. J. & J. G. Williamson (2005): "What fundamentals drive world migration?" In "Poverty, international migration and asylum," pp. 15–38. Springer.

Havranek, T., R. Horvath, & A. Zeynalov (2016): "Natural resources and economic growth: A meta-analysis." *World Development* **88**: pp. 134–151.

Havranek, T. & Z. Irsova (2017): "Do Borders Really Slash Trade? A Meta-Analysis." *IMF Economic Review* **65(2)**: pp. 365–396.

Hoeting, J. A., D. Madigan, A. E. Raftery, & C. T. Volinsky (1999): "Bayesian model averaging: a tutorial." *Statistical science* pp. 382–401.

Impellizzeri, F. & M. Bizzini (2012): "Systematic review and meta-analysis: A primer." *International journal of sports physical therapy* **7**: pp. 493–503.

Ioannidis, J. P., T. D. Stanley, & H. Doucouliagos (2017): "The power of bias in economics research." .

IOM, U. (2019): "World migration report 2020." `https://publications.iom.int/system/files/pdf/wmr_2020.pdf`.

Jauer, J., T. Liebig, J. P. Martin, & P. A. Puhani (2014): "Migration as an adjustment mechanism in the crisis? a comparison of europe and the united states." .

Jennissen, R. (2003): "Economic determinants of net international migration in western europe." *European Journal of Population/Revue Européenne de Démographie* **19(2)**: pp. 171–198.

Langan, D., J. P. Higgins, W. Gregory, & A. J. Sutton (2012): "Graphical augmentations to the funnel plot assess the impact of additional evidence on a meta-analysis." *Journal of clinical epidemiology* **65(5)**: pp. 511–519.

Lee, E. S. (1966): "A theory of migration." *Demography* **3(1)**: pp. 47–57.

Longhi, S., P. Nijkamp, & J. Poot (2008): "Meta-analysis of empirical evidence on the labour market impacts of immigration." *Available at SSRN 1136223* .

Martin, P. L. & J. E. Taylor (1996): *The anatomy of a migration hump.* Washington, D.C: Paris: OECD Publications and Information Center.

Massey, D. S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, & J. E. Taylor (1993): "Theories of international migration: A review and appraisal." *Population and development review* pp. 431–466.

Mayda, A. M. (2010): "International migration: a panel data analysis of the determinants of bilateral flows." *Journal of Population Economics* **23(4)**: pp. 1249–1274.

McKenzie, D., C. Theoharides, & D. Yang (2014): "Distortions in the international migrant labor market: evidence from filipino migration and wage responses to destination country economic shocks." *American Economic Journal: Applied Economics* **6(2)**: pp. 49–75.

McKinsey (2016): *People on the move: Global migration's impact and opportunity.*

Miyagiwa, K. (1991): "Scale economies in education and the brain drain problem." *International Economic Review* pp. 743–759.

Murat, M. (2020): "Foreign aid, bilateral asylum immigration and development." *Journal of Population Economics* **33(1)**: pp. 79–114.

Ortega, F. & G. Peri (2013): "The effect of income and immigration policies on international migration." *Migration Studies* **1(1)**: pp. 47–74.

Ozgen, C., P. Nijkamp, & J. Poot (2010): "The effect of migration on income growth and convergence: Meta-analytic evidence." *Papers in Regional Science* **(3)**.

Passaris, C. (1989): "Immigration and the evolution of economic theory." *International migration (Geneva, Switzerland)* **27(4)**: pp. 525–542.

Pedersen, P., M. Pytlikova, & N. Smith (2008): "Selection or network effects? migration flows into 27 oecd countries, 1990-2000." *European Economic Review* **52**: pp. 1160–1186.

Piore, M. J. (1979): *Birds of passage: Migrant labor and industrial societies.*

Pytlikova, M. (2006): "Where did central and eastern european emigrants go and why." *Aarhus: University of Aarhus, School of Business, CIM and Department of Economics.* .

Raftery, A. E., D. Madigan, & J. A. Hoeting (1997): "Bayesian model averaging for linear regression models." *Journal of the American Statistical Association* **92(437)**: pp. 179–191.

Ravenstein, E. G. (1885): "The laws of migration." *Journal of the statistical society of London* **48(2)**: pp. 167–235.

Rosenthal, R. (1979): "The file drawer problem and tolerance for null results." *Psychological bulletin* **86(3)**: p. 638.

Rotte, R. & M. Vogler (1998): "Determinants of international migration: Empirical evidence for migration from developing countries to germany." *Institute of Labor Economics (IZA)* .

Rotte, R., M. Vogler, & K. F. Zimmermann (1997): "South–north refugee migration: Lessons for development cooperation." *Review of Development Economics* **1(1)**: pp. 99–115.

Rowlands, D. (1999): "Domestic governance and international migration." *World Development* **27(8)**: pp. 1477–1491.

Sanderson, M. R. & J. Kentor (2008): "Foreign direct investment and international migration: A cross-national analysis of less-developed countries, 1985â€"2000." *International Sociology* **23(4)**: pp. 514–539.

Sanderson, M. R. & J. D. Kentor (2009): "Globalization, development and international migration: a cross-national analysis of less-developed countries, 1970–2000." *Social Forces* **88(1)**: pp. 301–336.

Simionescu, M., D. Ciuiu, Y. Bilan, W. Strielkowski *et al.* (2016): "Gdp and net migration in some eastern and south-eastern countries of europe. a panel data and bayesian approach." *Montenegrin Journal of Economics* **12(2)**: pp. 161–175.

Stanley, T. & H. Doucouliagos (2012): *Meta-Regression Analysis in Economics and Business*. Taylor & Francis.

Stanley, T. D. (2005): "Beyond publication bias." *Journal of economic surveys* **19(3)**: pp. 309–345.

Stanley, T. D., C. Doucouliagos, & S. B. Jarrell (2008): "Meta-regression analysis as the socio-economics of economics research." *The Journal of Socio-Economics* **37(1)**: pp. 276–292.

Taylor, E. J. (1999): "The new economics of labour migration and the role of remittances in the migration process." *International migration* **37(1)**: pp. 63–88.

VanderPlas, J. (2014): "Frequentism and bayesianism: a python-driven primer." *arXiv preprint arXiv:1411.5018* .

Vogler, M. & R. Rotte (2000): "The effects of development on migration: Theoretical issues and new empirical evidence." *Journal of Population Economics* **13(3)**: pp. 485–508.

Wallerstein, I. (2011): *The modern world-system II: Mercantilism and the consolidation of the European world-economy, 1600–1750*, volume 2. Univ of California Press.

Zelinsky, W. (1971): "The hypothesis of the mobility transition." *Geographical Review* **61(2)**: pp. 219–249.

Zeugner, S. & M. Feldkircher (2015): "Bayesian model averaging employing fixed and flexible priors: The bms package for r." *Journal of Statistical Software* **68**.

Appendix A

Studies includies in the dataset

Table A.1: List of primary studies

Adams JR. & Page (2005)	Helliwell (1997)
Bahna (2008)	Jennissen (2003)
Belot & Ederveen (2012)	Lewer & Van den Berg (2008)
Bertoli & Fernandez-Huertas Moraga (2013)	Lichter & Ambinakudige (2019)
Bove et al. (2017)	Mayda (2005)
Buch et al. (2003)	McKenzie et al. (2014)
Cappelen et al. (2015)	Melander & Oberg (2006)
Clark et al. (2002)	Migali et al. (2018)
Clark et al. (2004)	Murat (2020)
Cristina (2008)	Ortega & Peri (2009)
Dao et al. (2018)	Ortega & Peri (2013)
Docquier et al. (2014)	Pedersen et al. (2008)
Dreher & Poutvaara (2006)	Pytlikova (2006)
Dustmann & Okatenko (2014)	Rotte et al. (1997)
Ederveen et al. (2007)	Rowlands (1999)
Foo (2017)	Sanderson & Kentor (2008)
Hanson & McIntosh (2010)	Sanderson & Kentor (2009)
Hanson & McIntosh (2012)	Simionescu et al. (2016)
Hatton & Williamson (2003)	Simpson & Sparber (2010)
Hatton & Williamson (2005)	Volger & Rotte (2000)

Appendix B

Bayesian model averaging (BMA) diagnostics

Figure B.1: Posterior and Prior Model Probabilities in Bayesian Model Averaging

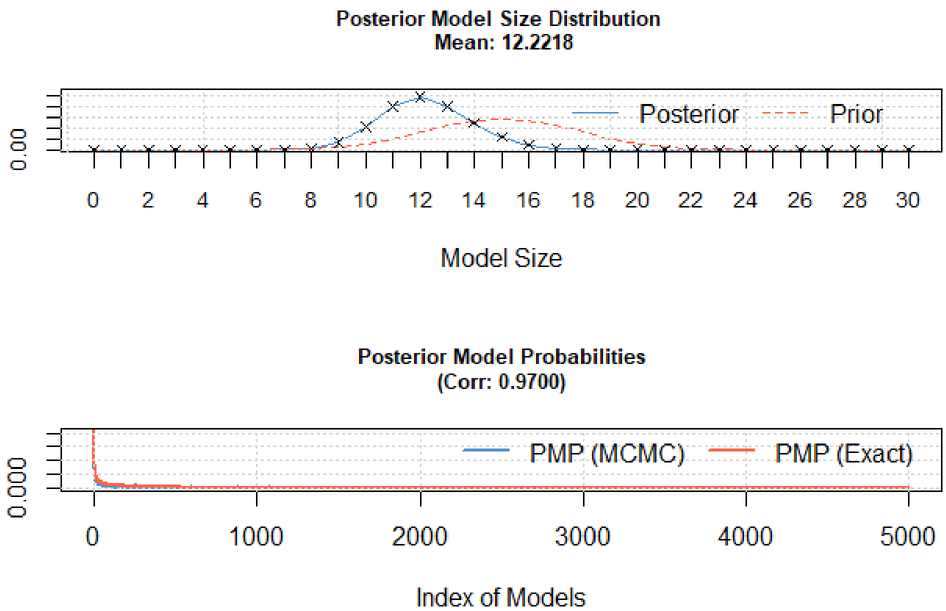
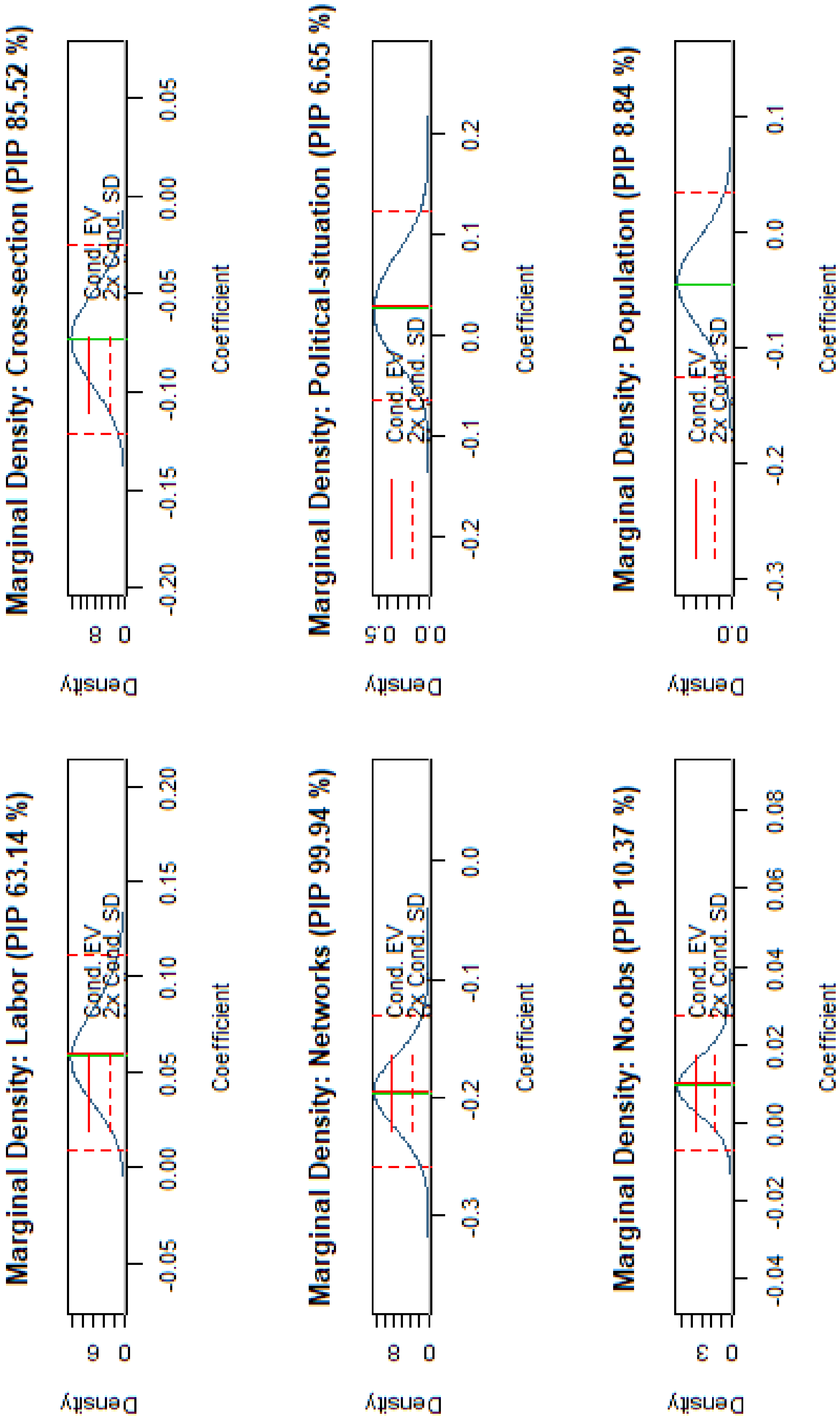


Figure B.2: Marginal densities for Posterior Inclusion Probabilities (PIP) higher than 0.5 in Bayesian Model Averaging



Appendix C

Correlation matrix

Figure C.1: Correlation matrix of the explanatory variables included in the regression baseline model