‘SHOULD I STAY OR SHOULD I GO?’

– THE CLASH (1982)

A panel analysis on the determinants of bilateral migration to the Netherlands

Abstract

This thesis looks at the determinants of bilateral migration to the Netherlands. Little research is done on this specific topic regarding the Netherlands. The research builds upon the famous Gravity equation. After the use of certain statistical analyses, this thesis finds that the Dutch GDP per capita, the population in the country of origin, a colonial history, a political conflict and the stock of migrants push and pull migrants to the Netherlands, while the distance, the unemployment in the Netherlands and stricter migration policies tend to reduce this inflow.

Master thesis International Economics, Erasmus School of Economics

Author:
Stefan Twigt, 341364
Supervisor:
Prof. dr. J.-M.A.R.G. Viaene
Co-reader:
Prof. dr. H.D. Webbink

Contents

Foreword

First, I would like to thank my supervisor, Prof. Dr. Viaene, for his help, useful feedback and flexibility throughout the process. Second, I would like to thank the co-reader, Prof. Dr. Webbink, for the time he is going to invest in this thesis.

Next to that, I would also like to thank my friends for their support and encouragement over the course of my study. In particular I would like to thank Olivier Hagenbeek for his support during the process of this thesis.

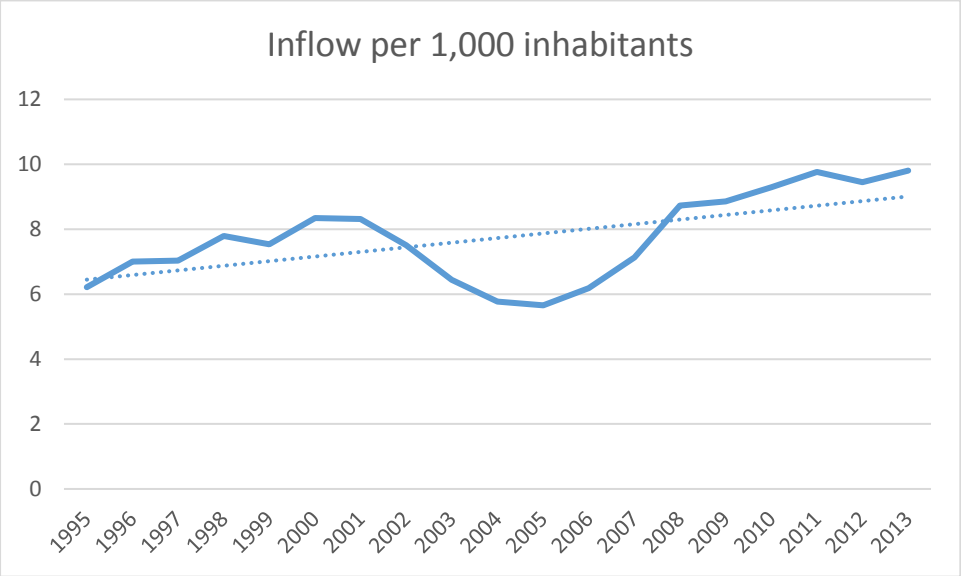
Finally, I want to thank my grandmother, Sjaan Twigt – Versnel, for her support and encouragement over the years of study. She was always there to help me. Unfortunately she cannot witness my graduation. I would therefore like to dedicate this thesis to her.

1. Introduction

Nowadays one cannot turn on the news or open a newspaper without noticing large headlines about the migration crisis. Momentarily, Europe is trying to cope with a large number of migrants who want to find a better life there.

1.1. Overview of the Problem

Ayo (2015) of Foreign Policy Magazine shows that during the first third of 2015 Germany was quite popular with migrants. Ayo also shows that the Netherlands had approximately 3,000 applications. However during August 2015, the amount really soared. In recent news, it was shown that the Netherlands will receive more than 9,000 migrants in 2015 and 2016 (Peeperkorn, 2015). As shown by Statistics Netherlands, the population growth has declined in the Netherland due to a lower natural growth[1], but the growth of population growth was higher than the years before, because of a positive migration surplus[2] (CBS, 2015). CBS also shows that immigration to the Netherlands has risen in the last couple of years, with a downfall in 2002 till 2005. This can be seen in Figure 1.1 below. Since 2005 the number of immigrants has grown almost every year.



Source: CBS
Graph 1.1 – Inflow of migrants in the Netherlands (1995-2013)

[1] The difference between fertility rate and death rate
[2] The difference between migration and emigration

But why do migrants choose for the Netherlands? There are certain pull and push factors. The most common pull factors are the standards of living, economic progress, more job certainty and a higher wage. However, there are also push factors in the country of origin. The most common are unemployment, war and poor human rights (Embrace, 2015). This thesis will research what the effect is of certain pull and push factors. Resulting in the following research question; *what determines migration to the Netherlands?*

1.2. Motivation

When examining the existing literature, little research has been done into the factors that determine migration to the Netherlands. This makes this study an addition to the existing literature on the topic of migration. Most researchers have only looked at the effects of migration and not at its determinants. Migration is also a current issue, the last couple of months most broadcasts in the news are about migration and European leaders have to make decisions on how to cope with the increasing growth of migration. The lack of research and because it is a current issue are the motivation for this thesis. Furthermore this thesis will expand the existing literature and finally find the drivers behind migration to the Netherlands. It will show policy makers what determines migration, and they could use it to see how they can increase or decline the inflow. The research will be done with the use of statistical methods (more below). That way one can see what the significant determinants are, how to anticipate on it, and see what the consequences are.

1.3. Analysis

The main method used in this thesis is the gravity equation. With this equation one can determine what influences the flow of migrants to the Netherlands. Using panel data regression analysis this thesis finds that the random effects model and the Mundlak model are the most suitable to use. After careful selection of the variables in the main analysis it shows that the Dutch GDP per capita, the population in the country of origin, a colonial history, a political conflict and the stock of migrants push and pull migrants to the Netherlands, while the distance, the unemployment in the Netherlands and stricter migration policies tend to reduce this inflow.

1.4. Contribution to the Literature

As mentioned above, not much research has been done regarding bilateral migration to the Netherlands. Most researchers only include the Netherlands in a set of countries. This thesis

will contribute to the existing literature by looking at only the Netherlands as a destination country. Furthermore, this is one of the first studies to look at the marginal effects of income in the origin country and the stock of migrants. Most papers that include the gravity model for migration do not look at the marginal effect of these variables.

1.5. Structure

This thesis is organized as follows: the first section will introduce the subject and show why it is relevant for both research and social purposes. The second section of the thesis will look at what migration is and will delve into the theory behind the gravity model. The third section provides a literature study. The fourth section will discuss the datasets which will be used and how the data is transformed. Section five will focus on the methodology. The sixth section will show the results and provide an interpretation of the results. Finally, the discussion and the conclusion follow in section seven.

2. Theoretical Framework

This section will focus on the theory behind migration and the different application of the model, namely the gravity equation. It will first start with a broad definition of migration, followed by an explanation of the gravity equation. For the gravity equation, this section will first look at the most common form, that of international trade, which forms the basis of the gravity equation for migration.

2.1. Migration

First, the definition of migration. As stated by Lee (1966), migration is defined as short term, long term or permanent change of residence, without any restriction upon distance. Migration can be voluntary as well as involuntary, and internal (inside a country) as well as external (between countries). Migration always involves an origin, a destination, a set of obstacles, and a distance. These are, however, not the same for all migrants, as they differ in personal factors. Every origin and destination have a set of personal factors for every individual. In Figure 2.1 below these are given as a "+" for positive factors that pull migrants or keep individuals, a "-"as negative factors that push individuals to migrate or repel migrants, and 0 as a factor for which an individual or migrant is indifferent. The obstacles and distance are also taken into account. An example of a personal factor could be a tax rate on cars. This is pull factor (+) for people who value the environment, because a high tax will mean that more people will use environment friendlier alternatives. While people with cars who need to travel far for work might be pushed (-) out and migrate. A person without a car will be indifferent (0). Obstacles in this model are certain restriction on the migrants. An example of an obstacle could be certain immigration policies. Distance is also different for some people; some people can travel a lot easier and further than others because of their savings.

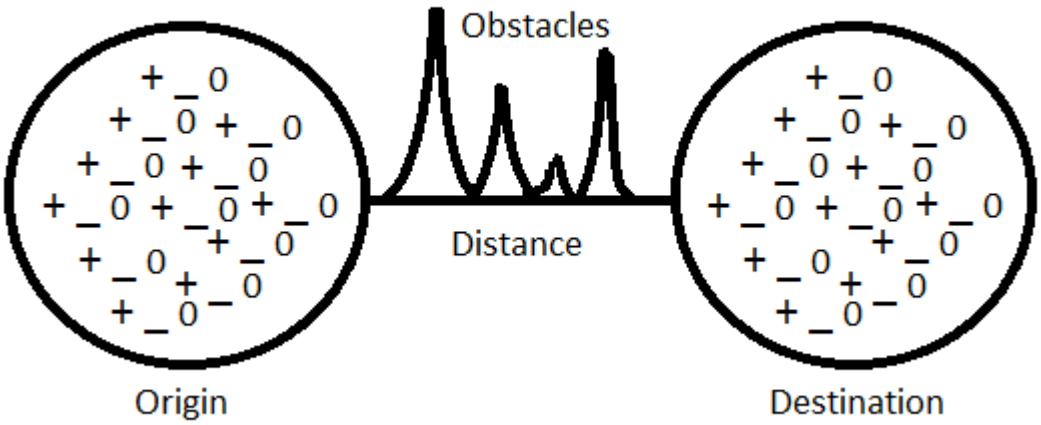


Figure 2.1- A simple figure of migration
Source: Lee, 1966[3]

One of the pioneers when it comes to the theory of migration was Ravenstein (Lee, 1966). Ravenstein (1885 and 1889) developed a theory for migration in the 1880s. This theory formed the basis for the modern migration theory. In his 1885 paper he stated five laws for migration; his 1889 paper added two more. These laws are summarized below.

1. Migration and distance: migrants tend to travel only short distances and migrants who travel long distances generally tend to migrate to big cities. He also stated that distance has a negative effect on migration, hence transportation costs are present.
2. Migration by stages: displacements of population produces migration in the direction of big industries. The population gaps will be filled with migrants from nearby cities or countries. The gaps they produce will again be filled by migrants from countries or cities nearby. Thus going in stages.
3. Stream and counter stream: every current of migration produces a counter-current. This can be explained that some migrants only migrate for a short time to gain wealth after which they migrate back to their origin country. Some will even take back children which were born in the destination country (Lee, 1966).

[3] Lee's original figure only contained push and pull factors, origin and destination, and the obstacles. It lacked distance which should also be considered for an individual. Since this is also taken into account in the decisions of the migrants, the laws of migration as conducted by Ravenstein below show this.

4. Urban-rural differences: those who live in towns are less likely to migrate then those in rural parts of the country. One could also interpret this as people who are living in rich countries being less likely to migrate then those living in poor countries.
5. Differences in female and male migrants: female migrants seem to travel only for short distances and within the country of origin, while male migrants appear to travel longer distances and also travel international.
6. Migration and technology: comparing different countries and cities, Ravenstein found that increases in technology, and innovations led to an increase in migration.
7. Economic motives: Ravenstein found several push and pull factors within the economic motives. Bad laws, high taxes, unattractive climates, unpleasant social surroundings, and constraints (like transportation), produce migration. These can be seen as the push factors. While, according to Ravenstein, the economic state of a country attracts large streams of migration. This can be seen as the pull factor.

These seven laws can thus be seen as the basis for migration theory. While Ravenstein mostly focussed on internal migration, these laws can also be applied for international migration. Especially the first, third, fourth, sixth and seventh law seem to apply for international migration.

2.2. The Gravity Equation

It is 1666: an apple falls from a tree while a certain man observes this. That man was Newton, who then used this phenomenon to develop the famous Newton's law of gravity. This law of gravity has been used first in international economics by Tinbergen in 1962. Tinbergen used the original gravity model

$$F = G\frac{M_1 M_2}{D^2} \tag{2.1}$$

to explain international trade by transforming it in to

$$F_{ij} = G\frac{{M_i}^{\beta_1}{M_j}^{\beta_2}}{{D_{ij}}^{\beta_3}} \tag{2.2}$$

, where F_{ij} stands for the volume of trade between countries i and j, G is a proportionality constant, M_i is the mass of the origin country, M_j is the mass of the destination country (in most applications this is reflected by a country's GDP), D_{ij} is the distance between countries, β_1 reflects the push flows, β_2 reflects the pull flows, and β_3 reflects the distance decay. This

basic gravity model is one of the most common models that has been used in the field of bilateral trade due to the fact that it can easily be augmented with other variables. Here, one could think of a shared border, languages, regional agreements and such others (Feenstra, 2004).

When looking at the literature on the gravity equation, two main branches can be observed; one that looks at the theoretical model, and one that looks at the empirical side. Tinbergen (1962) was the first empirical user of the model. At the same time, Poyhonen (1963) wrote a paper using a similar analysis of trade. Both find supportive evidence of the model with high fits. However, both researchers only use the empirical analysis, without a supportive theoretical analysis. Anderson (1979) was one of the first who finds a supportive theoretical analysis. In his renowned paper he provides a link between the economic theory and the equation. Over the years, different authors such as Bergstrand (1985) also found additional support to the model and its rationale.

2.3. Applying Gravity to Migration

The gravity equation for trade has largely been used in the economic literature to explain the flows between countries regarding their trade volume. If the gravity equation is suitable for trade flow, this could also hold for flows regarding factor movements. Anderson (2010) first observes the empirical model to fit well. He then builds the basics for the theoretical model. With the use of standard labour choice models, he shows that there is a solid theoretical model for the gravity equation for migration.

Over time, the gravity model has also been used for migration flows. Most researchers only stated the model without further introduction to it. Although these models showed a good fit, none of them tried to explain the model using underlying deeper theories for migration.

Ravenstein (1889) laid the basic theory for migration. He also laid the basics for the gravity model to be used. Looking at these laws, Ravenstein explained in words the foundation of the gravity model. Combining law one and seven results in the basis for a gravity model for migration. Like most gravity equations, migration will be influenced by push and pull factors, and distance. However, one still needs to know what the suitable push and pull factors are. Borjas (1989) proposed different models of immigration. He showed in theory that migrations depends on three equations, one for the wage in the origin country, one for the wage in the

destination country and one for the costs of moving from one country to the other. Combining the laws of Ravenstein with the theory of Borjas, the push and pull factors can be explained by the wages in both countries. However, Borjas also proposed that a model should include variables representing characteristics of origin and destination countries. Karemera et al. (2000) are among the first to develop the gravity equation using the theory proposed by Borjas. They state the following model:

$$M_{ij} = a_0 \frac{{S_i}^{a_1} {D_j}^{a_2}}{{R_{ij}}^{a_3}} \tag{2.3}$$

The migration flow depends on a supply function (S), representing the push factors, and demand function (D) representing the pull factors. Migration flows are negatively influenced by factors that restrain migration such as transport costs but also the characteristics of the country that aid or restrict migration(R). The supply and demand functions and the R-function are as follows:

$$S_i = b_0 Y_i^{b_1} N_i^{b_2} \tag{2.4}$$

$$D_j = c_0 Y_j^{c_1} N_j^{c_2} \tag{2.5}$$

$$R_{ij} = d_0 C_{ij}^{d_1} Z(\cdot) \tag{2.6}$$

where Y stands for the income of a country and N for the population of the country. In equation (2.5) C stands for the transportation costs and $Z(\cdot)$ stands for a function that contains the aiding and restricting factors and characteristics of a country. Substitution of (2.4), (2.5) and (2.6) into (2.3) and taking the logs of both side of (2.3), and of course substituting the terms by their equivalents gives the following basic migration model as proposed by Karemera et al.:

$$m_{ij} = \beta_0 + \beta_1 y_i + \beta_2 y_j + \beta_3 n_i + \beta_4 n_j + \beta_5 c_{ij} + z(\cdot) \tag{2.7}$$[4]

In other words, migration flows are influenced by the incomes and populations of both the origin and destination country, the transport costs, and the other costs and characteristics of the countries. In this simple model, $z(\cdot)$ can be seen as an error function. However it can also contain certain control variables needed for the empirical analysis. Borjas (1989) argues that

[4] The lower case letters indicate the natural logarithms of the original values, e.g. $\ln Y = y$

indicators of the economy ought to be included, such as inflation rates and unemployment rates. Moreover, Karemera et al. (2000) state that other authors suggested the use of certain variables measuring the democratic state of the country. Theoretically speaking, the models always include transportation costs. However, one of the main problems in economic literature is calculating the costs of transportation. These costs are not always available, so a common proxy used to measure this is the shortest air distance between the origin and destination country (Borjas, 1989). Hence, the most common equation for empirical use in the literature is the following model:

$$m_{ij} = \beta_0 + \beta_1 y_i + \beta_2 y_j + \beta_3 n_i + \beta_4 n_j + \beta_5 d_{ij} + \varepsilon_{ij} \quad (2.8)$$

The literature section below will delve deeper into the different variables and findings by other researchers.

3. Literature Review

In this section different research will be consulted to see the effects of certain variables and the use of them. Next to that, the hypotheses regarding the determinants will also be constructed. There is not a vast amount of literature on international migration; The empirical literature is quite small when looking at gravity equations on migration. Besides that, most of this small amount of literature concerns worldwide migration and migration in the USA. When looking at the Netherlands there is almost no literature of migration to the Netherlands.

3.1. Income

In the basic model, income is denoted as y. Economic theory suggests that one should look at the wages as an indication for income. An individual will compare his wage in his origin country and compare it with his expected wage in the destination country, if the difference between destination and origin is positive, he might have an incentive to migrate (Borjas 1989; Anderson 2011). One problem regarding the use of wages is that it is difficult to measure the wages in a country. Many researchers therefore look at the per capita income or income of the whole country. Karemera et al. (2000) look at data for the U.S. and Canada regarding international migration towards those countries. The researchers regress migration flows on GDP. A significant result in the U.S. data for income of the origin country is found. If the income is higher in the country of origin this reduces migration. However, for the Canadian data no significant effects are found. Leblang et al. (2006) look at 28 destination countries and their migration inflows from 144 origin countries. In their research they look at the ratio of income instead of splitting it. The ratio is the income of the origin country divided by the destination country. When the ratio of origin to destination country is increased, it decreases the migration flow. However, the results are insignificant which is unexpected. Clark et al. (2007) also look at the ratio of GDP per capita of the origin country and the destination country. When the per capita GDP ratio increases the migration flows decrease. Mayda (2010) looks at the effect of per worker GDP on the migration rate of 14 OECD countries and their inflow. She finds a significant and positive effect of an increase in per worker GDP. Ortega and Peri (2013) look at data consisting of 15 OECD destination countries and 120 origin countries over a large time span. A significant effect of the income per capita

at the destination country is found. This shows that migrant flows are very responsive to a higher income per capita at the destination country.

Greenwood and McDowell (1991) look at the determinants of migration to the USA and Canada. For their income variable they look at average manufacturing costs. This measures the earning differences between the countries of origin and the destination countries. There is a significant negative effect of per capita income differences between the countries. Migrant flows decrease if the origin country has a higher per capita income.

Most researchers look at the per capita GDP as a measure on income. This could either be in the form of a ratio between the origin and the host country, or separately. In line with equation (7), this thesis will look at the effects of both incomes separately. Following the literature, the following hypotheses are stated:

Hypothesis 1: an increase in the income per capita of the origin country has a negative effect

Hypothesis 2: an increase in the income per capita of the destination country has a positive effect

3.2. Population

Unlike a measure for income, a measure for population is consistent in the literature. Most literature looks at the total population in both the origin and destination countries. Karemera et al. (2000) look at the total population. They find a positive effect of the population increase in the origin country on migration flows and a negative effect of an increase in population in the destination country in their analysis. This is an intuitive and commonly found effect of population. Lewer and Van den Berg (2008) use a different approach. They look at the effect of the product of population in both the destination as the origin country. According to them, population size matters because a larger population in the destination country indicates a larger labour market, and a larger population in the origin country results in a larger fraction of migrants from that country. A positive effect is found for both. However, this is because they use the product of both. Other authors like Karemera et al. (2000) keep it separate, because one could also argue that a higher population in the destination country could indicate that the labour market could be saturated and thus slow migration down. This is supported by Ramos and Suriñach (2013). Their findings are in line with Karemera et al. who look at EU countries and their neighbouring countries (Ramos and Suriñach, 2013).

Following the literature the following hypotheses are formulated:

Hypothesis 3: an increase in the population of the origin country has a positive effect

Hypothesis 4: an increase in the population of the destination country has a negative effect

3.3. Transportation costs

Following Ravenstein's laws as stated above, transportation costs, or as Anderson (2011) called them 'iceberg costs', are measured by the distance between the origin and the host country. According to the first law, distance will decrease the inflow of migrants. There is a wide variety of measures for distance. One could measure the nearest air distance between countries, the nearest distance by sea or the nearest distance by ground[5].

Greenwood and McDowell (1991) measure distance by the airline mileage between the principal city in the country of origin and the nearest major city in the destination country. They do notice that the costs of transportation have been declined over time. Due to technological changes, travelling can be less time consuming than it used to be. Because of this they included a time trend. They find a significant negative effect for the distance between countries. Karemera et al. (2000) measure distance by the nearest air distance. No significant result is found. A different measure is used by Leblang et al. (2006) and Mayda (2010). Both papers measure distance by looking at the great circle distance between the country capitals. An expected and significant sign is found. Lewer and Van den Berg (2008) look at the geological distance between the capitals of the destination and origin country and find a negative and significant effect. Clark et al. (2007) find the same results.

Following the literature, the following hypothesis is formulated:

Hypothesis 5: an increase in the distance between the countries has a negative effect

3.4. Control Variables

Over the years, authors have used different sets of control variables depending on their main interest of research. They either look at migration flows to a specific country or look at it worldwide. Most control variables are country specific variables and variables that serve as a push of pull factor for migrants.

[5] Air distance is the distance by plane, sea distance is the distance by boat and ground distance is the distance from one capital to another over land.

In the existing literature a common variable is the stock of migrants in the destination country. If a country already has a large stock of migrants, this will increase the migration flows. This is due to the fact that future migrants are likely to receive information about the destination country from the migrants who are already in that country. In the literature, most authors find a positive and significant effect (Leblang et al., 2006; Clark et al., 2007; Lewer and Van den Berg, 2008). Mayda (2010) uses the lag of migration flows instead of stock, however the motivation and the effects are the same. Most of the literature considers a dummy variable for common language. If countries share the same language or have an education that includes that language this should increase migration (Greenwood and McDowell, 1991; Lewer and Van den Berg, 2008; Tranos et al., 2015). In the literature the effect is thus positive, however not all find a significant effect (Karemema et al., 2000; Clark et al., 2007; Mayda, 2010; Ortega and Peri, 2013). Other control variables that are widely used are variables that look at the political rights in the destination and origin countries. If the origin country has bad rights for its inhabitants some are more likely to migrate. If the destination country has better rights compared to the origin countries, the flows increase. The literature suggests a strong significant and economics effect (Greenwood and McDowell, 1991; Karemera et al., 2000; Leblang et al., 2006). Lewer and Van den Berg (2008) find a significant effect but the effect is rather small, however this could also be due to the fact that they use an extra variable that looks at the protected property rights. Greenwood and McDowell (1991) also use instability measured by a crisis (for example a war), this effect is also significant and increases migration. Another variable that is widely used, is a variable that measures immigration policies in the destination country, some countries have a strict policy regarding immigrants (for example the U.S. border control with Mexico). If a country has a strict policy regarding immigrants this will reduce the inflow of migrants. In the literature the effect is negative and significant (Greenwood and McDowell, 1991; Karemera et al., 2000; Ortega and Peri, 2013). Ortega and Peri (2013) also use an extra variable to determine the effect of an immigration policy. They used the Maastricht treaty, which was ratified by most EU countries and introduced free labour mobility for member states. A positive significant effect is found. Another variable that is used by some studies is the unemployment rate in the destination and origin country. An increase in unemployment in the origin country could influence the decision of workers to migrate. On the other hand, if the unemployment ratio increases in the destination country, this could decrease the migration flows towards the destination country. Most studies do not

find a significant effect (Karemera et al., 2000; Leblang et al. 2006; Mayda, 2010). Studies regarding worldwide immigration also include variables that measure if the origin and destination country share a border and if the origin country is a former colony of the destination. Sharing a border makes it easier to travel and most people are familiar with the neighbouring countries. This would thus increase migration flows. However, the literature is not consistent about the results: some find a negative effect and others a positive effect. Most papers, however, do not find a significant effect (Leblang et al., 2006; Lewer and Van den Berg, 2008; Mayda, 2010; Ortega and Peri, 2013). Tranos et al. (2015) find a positive and significant effect. Regarding countries that have a colonial background with the destination country; these countries could share certain cultural aspects, this would make it easier to migrate to the destination countries. The evidence however is quite ambiguous. Leblang et al. (2006) and Mayda (2010) find no significant effect. Lewer and Van Den Berg (2008), Ortega and Peri (2013), and Tranos et al. (2015) do find a significant and positive effect.

Some others use special variables not included in other research. For instance Karemera et al. (2000) look at the credit worthiness of the countries and the inflation of the countries. Higher credit worthiness could indicate a better financial performance and better future economic opportunities, so migrants might move to countries that score better. Regarding inflation, increasing rates in the origin country could be an indication of domestic malaise, thus leading to higher migration flows to the destination countries. The higher the inflation in the destination country the less attractive it is to migrate there. Another indicator for economic opportunities is the use of GDP growth, if the destination country has a higher growth this will increase migration flows. Growth in the country of origin decreases migration on the other hand. However there is little to no evidence that this is the case.

3.5. Application to the Netherlands

The literature that focuses on the Netherlands is scarce. Little research is done that focuses specific on the Netherlands only. The Netherlands were only included in panel research that looked at migration on a worldwide scale. However, some variables that are used in other research not focussing on the Netherlands but on worldwide migration can be used. These variables include being part of the EU, the stock of migrants, a colonial history with the origin countries, the political rights in the origin countries, the unemployment rate in both countries, inflation in both countries, GDP growth in both countries and immigration policies. Shared

borders will not be included, because this could lead to multicollinearity, due to the facts that the country that share a border with the Netherlands are also part of the EU. Common Language will also be dropped, because countries that share the same language are often former colonies of the Netherlands, and might have led to multicollinearity.

A popular belief is that many immigrants and refugees come to welfare state countries because of the unemployment benefits. Researchers have shown that this so-called "welfare migration" is not backed-up by empirical research. They show that the flows of migration are not related to unemployment benefit expenditures of the government (Giulietti et al., 2011). This thesis will also use it applied to the Netherlands to see if regarding the Netherlands this effect also holds.

Recent research by the *Sociaal en Cultureel Planbureau* (Social Policy Institute) (SCP) show that the public opinion in 2015 and 2016 has shifted. In their research the SCP shows that the Dutch population has growing concerns regarding refugees and immigrants (Den Ridder et al., 2016). This is backed up by earlier research by the OECD. In 2010 the OECD concluded that in most countries there is a growing negative opinion about migrants (OECD, 2010). A shift in the public opinion could be measured by the rise of right winged parties. In 2000 former politician Pim Fortuyn gained fame with his opinion of foreign people. Later in 2004 the Netherlands saw the rise in other right winged parties, like the PVV, and some parties also took different opinions about migrants. Especially now a days with the European migrant crisis, far right winged parties are on the rise, gaining ground in most European countries (Payne, 2016). It is therefore interesting to include variables to measure the public opinion in this analysis, how this is measured will be explained in section 4.6 below.

Following the above the following hypotheses are formulated:

Hypothesis 6 :An increase in inflation in the country of origin has a positive effect
Hypothesis 7: An increase in inflation in the country of destination has a negative effect
Hypothesis 8: An increase in GDP growth in the country of origin has a negative effect
Hypothesis 9: An increase in GDP growth in the country of destination has a positive effect
Hypothesis 10: An increase in unemployment in the country of origin has a positive effect
Hypothesis 11: An increase in unemployment in the country of destination has a negative effect

Hypothesis 12: Sharing a colonial history has a positive effect

Hypothesis 13: Being part of the EU has a positive effect

Hypothesis 14: A political conflict in the country of origin has a positive effect

Hypothesis 15: Stricter migration policies in the country of destination has a negative effect

Hypothesis 16: An increase in umeployment benefit expenditures has a positive effect

Hypothesis 17: An increase in the stock of migrants has a positive effect

Hypothesis 18: A right-winged public opinion has a negative effect

For the convenience the above hypotheses are summarized in Table 3.1 below.

Sign	Variable in country of origin	Variable in country of destination
Positive	Population, Inflation rate, Unemployment rate, Colonial history, EU membership, Political conflict	GDP per capita, GDP growth, Unemployment benefits, Migrant stock
Negative	GDP per capita, Distance, GDP growth	Population, Inflation rate, Unemployment rate, Migration policy, Public opinion

Table 3.1 - Summary of the hypotheses

4. Data

This section will look at the data that has been used in this thesis. First the dependent variable will be discussed, followed by the variables in the main theoretical model and lastly the control variables. Table 4.1 defines the variables and displays their respective sources.

Variable	Measure	Source
MP	The bilateral inflow per 1,000 inhabitants	Statline CBS
POP	The *de facto* definition of population	World Bank
GDP	Per capita GDP in 2005 $ divided by midyear population	World Bank
DIST	The distance in kilometres	CEPII
INFL	Percentage growth rate of the GDP implicit deflator	World Bank
UNEMP	ILO percentage of total labour force estimates	World Bank via ILO
GROWTH	Per capita growth rate of GDP in 2005 $	World Bank
COLONY	Dummy variable if the country has a Dutch colonial history	CEPII
CONFLICT	Dummy variable if a country has a political conflict	Major Episodes of Political Violence
POLCIY	Measures how less or more restrictive a policy has become	Determinant of International Migration
EU	Dummy variable if a country is a EU member state or not	EU website of member states
PUBLIC	Dummy that measures the rise of right winged ideas	Own construction
GOV	Dummy variable that measures if the Dutch cabinet is right winged or not	Own construction
STOCK	Can either be the lag of the inflow or the stock of foreign population	Statline CBS/OECD

Table 4.1 –Summarized data

4.1. Bilateral inflow

Bilateral migration is measured as the inflow from one country to another. In this case the inflow from a selected country to the Netherlands. Statistics Netherlands (CBS) has measured the inflow of migrants from different countries to the Netherlands over the period of 1995 – 2013. The selected variable *immigration* contains all persons who enter the Netherlands that plan on staying for an indefinite time. These persons are of all ages and of all marital status and sorted per country of origin.

4.2. Population, GDP per capita, and Distance

For the variable *population* of both the destination and the origin country the World Bank measures are used. Population is measured by the *de facto* definition of population. Which takes all residents regardless of their legal status and citizenship, apart from refugees into account. GDP per capita is used as a substitute for the wages. *GDP per capita* is measured as the GDP using current 2005 U.S. dollars and dividing it by the midyear population. Distance comes from the CEPII. *Distance* is measured as the distance in kilometres between the countries of origin and the Netherlands.

4.3. Control Variables

Inflation is measured in percentages by the annual growth rate of the GDP implicit deflator. This is the ratio of the GDP in current local currency to GDP in constant local currency. It shows the rate of price changes as a whole in an economy. The *GDP growth* is measured by taking percentage growth in the total GDP in current 2005 U.S. dollars at time t compared to time t-1. The *unemployment rate* is measured as a percentage of the total labour force. The rates used are the ILO estimates from come from the ILO's Key Indicators of the Labour Market database. These estimates are strictly selected by certain criteria and methods, which ensure comparability over time and across countries. *Colonial history* is a dummy variable that indicates if a country has a colonial history with the Netherlands or not. The variable *political conflict* is the total summed magnitudes of all major episodes of political violence. This variable consists of the magnitudes of all civil, ethnical and international violence, and all civil, ethnical and international warfare. The total *unemployment benefits* are measured by taking the total expenditures by the government for people who are unemployed. The government expenditures are measured by the total amount of money the government spends. For the *policy* variable was created by performing certain transformations. The variable originates from the Determinant of International Migration (DEMIG) dataset which measures the magnitude of the policy. This can be either a fine-tuning, a minor, a mid-level or a major change. Furthermore, the change in restrictiveness is measured, this can either be less restrictive, no change or more restrictive (De Haas et al, 2014). The construction of the variable will be discussed in the section below. The *stock of migrants* consist of the stock of foreign population grouped by the nationality of the country of origin.

4.4. Initial Dataset

The above described datasets combined, results in the dataset that is used in this thesis. The dataset consists of 164 countries (for the full list of countries see Appendix Table A.1.), which consists of a time span of 1995 till 2013. However for the statistical analysis some data needs to be transformed. First the migration inflows and the stock of migrants are divided by the population per 1,000 inhabitants, in this way the absorption rate of migrants in the Netherlands is measured. Second, the natural logarithms are taken of the variables for the migration inflows per 1,000 inhabitants, GDP per capita, population, and distance. Third, the expenditures on unemployment benefits are divided by the total government expenditures, that way it is relative to government spending. The lag of the logarithm of migration inflows is used as a measure for the stock of migrants; this is in line with research of Mayda (Mayda, 2010).

To make a variable for political conflict, a dummy is created to measure if the country of origin is in a conflict or not. If there is any sign of conflict the dummy is set to 1, otherwise it is 0. To get the variable for migration policies, the magnitude of change is changed in a categorical variable on the scale of 0 till 3[6]. The change in restrictiveness is changed with -1 for less restiveness, 0 for no change, and +1 for more restrictiveness. Afterwards the scale of the magnitude is multiplied with the change in restrictiveness and summed per year. The policy change variable is thus measured from a negative value to a positive value. The more negative the less restrictive the total policy changes were, the more positive the variable is, the more restrictive the total changes were. Following Madya (2010) this variable will be used to measure the policy in the Netherlands. It will increase with 1 (decrease with 1) when the migration policy becomes more (less) restrictive. Using the list of countries that are part of the European Union (EU) from the website of the EU a dummy is created that is 1 if the country of origin is part of the EU and 0 otherwise. Next to the economic variables, this thesis will also look at some political variables for the Netherlands. This thesis will look at the public opinion, a dummy is introduced which measures the time that the former Pim Fortuyn gained public support in his political campaign. The dummy is set from 2000 in which the politician gained support and the public opinion on foreigners shifted. Another measure for the Dutch public opinion is to look at the composition of the cabinet. The Netherlands mostly has a

[6] Where 0 = fine-tuning, 1 = a minor change, 2= a mid-level change and 3 is a major change.

combination of left, right and middle parties. From 1995 till 2013 there has never been a fully left party so the variable is coded as right-wing cabinet (1) or a middle-wing cabinet (0).

4.5. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Id	3,116	82.5	47.349	1	164
Year	3,116	2004	5.478	1995	2013
lnMP	3,025	-4.997	1.979	-9.729	0.200
lnPOPi	3,097	16.056	1.592	11.547	21.029
lnPOPnl	3,116	16.600	0.025	16.554	16.637
lnGDPi	2,981	7.896	1.612	4.228	11.364
lnGDPnl	3,116	10.606	0.086	10.422	10.718
lnDIST	3,097	8.365	0.920	5.153	9.845
INFLi	2,989	15.385	123.114	-31.566	5399.507
INFLnl	3,116	2.011	1.128	0.142	4.423
UNEMPi	3,059	8.798	6.218	0	39.3
UNEMPnl	3,116	4.295	1.396	2.1	7.2
GROWTHi	2,821	2.755	6.493	-62.214	141.642
GROWTHnl	2.952	1.393	2.190	-3.794	3.974
BENEFITS	3,116	1.236	0.567	0.579	2.584
COLONY	3,116	0.049	0.215	0	1
EU	3,116	0.110	0.313	0	1
CONFLICT	3,087	0.172	0.377	0	1
PUBLIC	3,116	0.737	0.440	0	1
GOV	3,116	0.368	0.482	0	1
POLICY	3,116	1.789	1.609	-1	5
STOCK	2,824	0.255	0.840	0	9.981

Table 4.2 - Descriptive statistics

Table 4.2 above shows the descriptive variables, this will give a better look at the data. As one can see, the sample contains 164 countries over 19 years. On average there is an inflow of 0.007 (exp(-4.99)) migrants per 1,000 inhabitants over these years. However looking at the standard ratio there is a large variation in this amount. The population of the countries of origin tends to fluctuate some over time, with an average of around 0.94 million population (exp(16.056)) and varying from approximately 100,000 inhabitants till 1.35 billion inhabitants. For the Netherlands the population does not fluctuate much. The population growth is not that large in the Netherlands over the 19 years. Just like the population, GDP per capita shows a similar patron. The origin countries seem to fluctuate more, while the Netherlands seems to have a steady GDP per capita over time, ranging from 33,500 till 45,200, with an average of 40,500. On average the origin countries have a much smaller GDP per capita (2,750) than the Netherlands. Looking at the distance the average is 4,300 kilometers and varies from 173 kilometers till 18.870 kilometers. The inflation in the country of origin seems to fluctuate a

lot, which is expected because some countries (e.g. Zimbabwe) in the sample have had large inflation rates. The Netherlands appears to be around an average of 2%, mostly based on European Union Policy. Unemployment rate seems to be higher in countries that migrate then in the Netherlands based on their averages and standard deviation. Economics growth in the origin countries has a large fluctuation and varies from -62% till 142%. This is not unusual because some countries had huge growth as upcoming economies, like India and China. The Netherlands seem to have an economic growth between -3.8% and 4%, with an average of 1.4%. The unemployment benefits in the Netherlands are on average 1.2% of the total expenditures. Overtime this mostly is between 0.6% and 2.6% of total expenditures. Over time 11% of the observed countries is part or became part of the EU, 5% of the sample has a colonial history with the Netherlands. And there were 536 (17.20% of the 3,116 observations) cases of a form of conflict on average in the origin countries. The public opinion is hard to interpret as it is a dummy that shows the rise of extreme right. The distribution of the government shows that it is mostly middle winged (37% of the 19 years a right winged government was in office in the Netherlands). Policy regarding migration seems to become stricter over time, the average is positive and is varies from -1 till 5. This is skewed to left.) Looking at the stock of migrants, the data shows that this seems to vary a lot. Some countries have a larger amount of migrants in the Netherlands then others.

5. Methodology

This section will look at the statistical methods used in the results. First the main equation used in this thesis will be looked upon. After which this thesis will delve deeper in the underlining analysis regarding panel data.

5.1. Model

In order to compute a statistical analysis the main equation needs to be rewritten in the form of a simplistic OLS regression:

$$Y_{i,t} = X'_{i,t}\beta + \varepsilon_{i,t} \qquad (5.1)$$

where Y is the dependent variable, X is a matrix containing the independent variables, β is a vector that measures the effect of the independent variables on the dependent variable, and ε is the error term. i is the subscript for individuals and t is the time subscript (Hill et al., 2012).

Using the variables from section 4, the main regression in this thesis is as follow:

$$\ln \frac{M_{i,nl,t}}{P_{nl,t}}$$
$$= \beta_0 + \beta_1 \ln GDP_{i,t} + \ln GDP_{i,t} + \beta_3 \ln POP_{i,t} + \beta_4 \ln POP_{nl,t} + \beta_5 \ln DIST_i$$
$$+ \beta_6 INF_{i,t} + \beta_7 INF_{nl,t} + \beta_8 UNEMP_{i,t} + \beta_9 UNEMP_{nl,t} + \beta_{10} GROWTH_{i,t}$$
$$+ \beta_{11} GROWTH_{nl,t} + \beta_{12} BENEFITS_{nl,t} + \beta_{13} COLONY_i + \beta_{14} EU_{i,t} + \beta_{15} CONFLICT_{i,t}$$
$$+ \beta_{16} PUBLIC_{nl,t} + \beta_{17} GOV_{nl,t} + \beta_{18} POLICY_{nl,t} + \beta_{19} STOCK_{i,nl,t} + \varepsilon_{i,nl,t}$$

(5.2)

where i is the country of origin, nl is the country of destination, in this thesis the Netherlands, and t is the time. $\frac{M_{i,nl,t}}{P_{nl,t}}$ is the bilateral inflow of migrants per 1,000 Dutch inhabitants from country i to the Netherlands. GDP is the GDP per capita, POP is population size and DIST is the distance from country i to the Netherlands. INF is the inflation rate in percentages, UNEMP the unemployment rate in percentages and GROWTH the growth in GDP per capita. BENEFITS are the expenditures of the Dutch government on unemployment benefits in percentages of total expenditures of the government. COLONY is dummy that is 1 if a country of origin has a colonial history with the Netherlands, EU is a dummy that is 1 if a country is a member of the EU and CONFLICT is a dummy that is 1 if the country has a conflict that is civil, ethnical or international. PUBLIC and GOV are dummies that measure the public opinion that is equal to 1, respectively if the type of cabinet that takes office is right winged and when the more extreme right winged parties rose in popularity. POLICY measures the restrictive changes in migration policy in the Netherlands as described in the section 4.6. STOCK is the stock of foreign based population in the Netherlands per thousand inhabitants. This variable measures the family migration effect, and the information streams between migrants. Furthermore, $\beta_k, k \in (1,19)$ measures the effect of the independent variables on the dependent variable, with $\beta_2, \beta_3, \beta_6, \beta_8, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{15}, \beta_{19} > 0$, and $\beta_1, \beta_4, \beta_5, \beta_7, \beta_9, \beta_{10}, \beta_{16}, \beta_{17}, \beta_{18} < 0$.

Because the literature is quite divided in the use of the family and information effects for migrants, some use the stock, while others use the lag of the dependent variable, this thesis will run two regression forms in this thesis. First the model of equation (9), second the same model only with the lag of the dependent variable instead of the stock.

5.2. Panel Analysis

When looking at equation (5.2) one can see that this regression includes multiple countries over a certain time frame, which indicates that a regular OLS regression is not suitable for this thesis. This subsection will delve deeper in to the underlying assumption of panel analysis.

Panel regression come in many different forms and are no exemption in international economic research when one looks at multiple countries. The most common regressions in panel research are the pooled, fixed and random effects models. These different forms will be looked more closely into.

A pooled model pools all the individuals (in this case countries) together. This means that there are no individual effect taken into account.

$$y_{i,t} = \beta_0 + \beta_1 x_{i,t} + \varepsilon_{i,t} \quad (5.3)$$

where:

$y_{i,t}$ = the dependent variable for country *i* at time *t*

β_0 = the "pooled" constant

$x_{i,t}$ = the independent variable per country *i* at time *t*

β_1 = the effect of the independent variable on the dependent variable

$\varepsilon_{i,t}$ = the error term

When a model that is not pooled is assumed, one can choose either between a fixed effects model or a random effects model. The fixed effect (FE) model is as follow:

$$y_{i,t} = \beta_{0,i} + \beta_1 x_{i,t} + \varepsilon_{i,t} \quad (5.4)$$

where:

$y_{i,t}$ = the dependent variable for country *i* at time *t*

$\beta_{0,i}$ = the constant per country *i* (which differs between countries), the so-called "fixed effect"

$x_{i,t}$ = the independent variable per country *i* at time *t*

β_1 = the effect of the independent variable on the dependent variable

$\varepsilon_{i,t}$ = the error term

When a fixed effects model is used, one assumes no correlation between the error terms of countries. Independent variables of the countries are allowed to be correlated with other independent variables of countries. It is also assumed that the error term has an expected value of zero and a constant variance of σ^2.

The random effects (RE) model is as follow:

$$y_{i,t} = \beta_0 + \beta_1 x_{i,t} + u_i + \varepsilon_{i,t} \quad (5.5)$$

where:

$y_{i,t}$ = the dependent variable for country *i* at time *t*

β_0 = the average constant of all countries

$x_{i,t}$ = the independent variable per country *i* at time *t*

β_1 = the effect of the independent variable on the dependent variable

u_i = the random effect

$\varepsilon_{i,t}$ = the error term

When the random effects model is used, the same assumptions as the fixed effect model hold expect for one. The random effects model assumes no correlation between the country based independent variables.

Due to the fact that the gravity model contains country specific variables such as the distance, and the colonial and EU dummies, fixed effects models will not be suitable. Fixed effects models already correct for these variables and would thus result into multicoliniarity. Because of the use of a lagged dependent variable in the second regression, a random effects model could result in a pooled model. Therefore a dynamic model needs to be used. However, this only works for fixed effects models. So in order to look at the effect of the lagged term a pooled model with the term included needs to be considered. Another option is the use of a linear model with a AR(1) term, also known as the xtregar command in Stata. This method is also used by Tranos et al. (2015). This model uses a lagged term of the dependent variable, hence one can interpret its effect.

Before the different models are looked upon, some tests need to be performed. First, one needs to make sure that the dependent variable is not showing a random walk. So a unit root test needs to be conducted. This will be done by performing the Augmented Dickey-Fuller (ADF) test by Fisher (Hill et al., 2012). First a plot will be made of the dependent variable to determine which type of ADF test is needed. There are three options, namely:

- No constant and no trend (fluctuation around zero)
- A constant, no trend (no fluctuation around zero)
- A constant and a trend (fluctuation around a linear trend)

When the null hypothesis is rejected, there is a unit root, hence the research can be continued. If the null hypothesis is not rejected one needs to take the first differences of the dependent variable.

After this a test for autocorrelation is conducted. The test used for this is the Wooldridge test by Drukker. The null hypothesis of this test is that there is no autocorrelation. If the null hypothesis is rejected, two models can be used depending on which type of panel regression is being run (Drukker, 2003). In case of a pooled model one can use the Driscroll Kraay standard error, otherwise one needs to use the robust standard errors for the random effects model. With the second regression the same will go up as stated above, except that one can use the xtregar regression as mentioned above. So first one needs to check which model is most suitable.

To test this the Breusch and Pagan Lagrange Multiplier (BP LM) test needs to be performed. The null hypothesis states that the model is pooled, while the alternative hypothesis states that the model is random (Breusch and Pagan, 1979).

5.3. Mundlak Model

One of the assumption in the RE model is that the random effect u in equation (5.5) has no correlation with the country independent variables. This is however not always the case. Sometimes a RE model needs to be used while a FE model is more suitable. This is also one of the problems in this thesis. A better use is the FE model, however this thesis includes certain country specific variables which tend to be correlated, such as the variables for distance and colonial history. This means that the RE model needs to be used in the analysis, but it may not give the most suitable estimators. A way to correct for this has been proposed by Mundlak

(Mundlak, 1978). Mundlak proposed that one should use the following expression for the adjusted RE model:

$$y_{i,t} = X'_{i,t}\beta + Z'_i\delta + c_i + \varepsilon_{i,t} \quad (5.6.)$$

where X is the matrix of all variables but does not contain a constant. Z is the matrix of the variables with no variance or a small variance over time, like the distance or the colonial history dummies. The variable c is a function that contains the random effect and the constant. This function is as follow:

$$c_i = \alpha + \overline{X'_i}\theta + w_i \quad (5.7)$$

where α is the constant, $\bar{X}$ is the average of the explanatory variables which have a variance over time and w is the random effect. Substituting equation (5.7) into (5.6) gives us the following equation

$$y_{i,t} = \alpha + X'_{i,t}\beta + Z'_i\delta + \overline{X'_i}\theta + w_i + \varepsilon_{i,t} \quad (5.8)$$

The β in this expression gives us the fixed effects coefficients for the time varying variables which are better that the random effects coefficient expression [5.5] gives us. The δ is the coefficient for the country specific variables for which a FE model would correct. One can now interpret the outcome better.

6. Results

To the extent of answering the research question, this section will focus on the results found by the use of statistical analyses of the data. First the data will be described, followed by some tests that are needed for the choice of models. After that the regression analysis is conducted.

6.1. The Regression Analysis

Before this thesis continues to the models, a test for unit roots needs to be conducted to see if the model doesn't follow a random walk. Analysis of the dependent variable shows in Table 4.2 that the dependent variables fluctuates around a mean of -4.99. So there seems to be a drift. An Augmented Dickey Fuller test with a drift included on the dependent variable needs to be performed. The results are stated below in Table 6.1.

Test	Statistic
Inverse Chi-squared[324] (P)	1201.296***
Inverse normal (Z)	-22.972***
Inverse logit t[814] L*	-25.237***
Modifief inverse Chi-squared (Pm)	34.463***

*Notes: sign: ***<0.01; **<0.05; *<0.10*
Table 6.1 - ADF test with a drift

Multiple tests are included for robustness. All show the same result, the null hypothesis of unit roots is rejected. There is a strong significant evidence that the dependent variables is stationary and thus does not show a random walk. The regression analyses can be performed.

As can be seen in Table 6.2, the coefficient for lnGDPnl is significant at the 10% level, but has a negative sign. This is not intuitive, because it should increase migration. The coefficient for lnPOPnl is very high and significant (16.95 in the pooled model and 15.63 in the RE model). This would mean that a 1% increase in the Dutch population will result in a 16% increase in inflow of migrants per 1,000 inhabitants. lnDIST has a negative but insignificant effect, which means that it has no effect on the inflow. This is also not in line with the literature. This is likely due to high correlation among the independent variables.

lnMP	POOLED	RE
lnGDPi	0.283*** (0.018)	0.201*** (0.05)
lnGDPnl	-5.227* (2.741)	-4.182*** (1.029)
lnPOPi	0.644*** (0.018)	0.483*** (0.056)
lnPOPnl	16.949 (10.624)	15.632*** (3.969)
lnDIST	-0.008 (0.039)	-0.001 (0.11)
INFLi	0.001* (0.001)	0.001* (0.001)
INFLnl	0.063 (0.044)	0.071*** (0.017)
UNEMPi	-0.005 (0.005)	0.013** (0.005)
UNEMPnl	-0.125 (0.101)	-0.1*** (0.038)
GROWTHi	0.002 (0.004)	0.001 (0.002)
GROWTHnl	0.034 (0.029)	0.022** (0.011)
BENEFITS	-0.127 (0.306)	-0.137 (0.114)
COLONY	1.14*** (0.107)	1.094*** (0.422)
EU	0.976*** (0.108)	1.298*** (0.083)
CONFLICT	0.529*** (0.072)	0.208*** (0.04)
GOV	-0.071 (0.09)	-0.061* (0.034)
PUBLIC	-0.074 (0.184)	-0.123* (0.069)
POLICY	-0.051** (0.023)	-0.065*** (0.009)
STOCK	0.514*** (0.032)	0.408*** (0.039)
_cons	-243.016 (154.473)	-229.253*** (57.68)
Obs	2,508	2,508
R^2	0.599	0.578
BP-test	181.16***	
Wooldridge	47.079***	
BP-LM	13694.5***	

*Notes: ***<0.01; **<0.05; *<0.10*

Tabel 6.2 - Pooled and Random regression

To see if the variables in the proposed model are correct and not correlated among each other, a correlation matrix is used. This correlation matrix is attached in the appendix in Table A.2. In this matrix the Pearson correlations are given. A high chance of multicollianiarity is assumed when the correlation is significant and above the threshold of 0.60. As one can see lnPOPnl is highly and significantly correlated with lnGDPnl, GROWTHnl, BENEFITS and PUBLIC. BENEFITS has a significant correlation of 0.9 with lnGDPnl. This could indicate that these variables are not suitable in the model and will cause a form of multicollinearity. Also notable is that the dummy EU is highly significantly correlated with lnDIST, this could indicate the unintuitive effect of lnDIST. Although the correlations are not that high but above 0.3, one could argue that GOV is also a variable that is endogenous with POLICY. A policy implementation is most likely the result of a sitting government, thus GOV can also be a variable that is not suitable in the model. The test for autocorrelation and the heteroskedisticity test for the pooled model show significance. This means that clustered variables need to be used. The test BP-LM test is also significant and shows that the RE model needs to be used.

Looking at the correlation matrix, the following regression equation will be used for in the analysis:

$$\begin{aligned}\ln\frac{M}{P}_{i,nl,t} = \quad & \beta_0 + \beta_1 \ln GDP_{i,t} + \beta_2 \ln GDP_{nl,t} + \beta_3 \ln DIST_i + \beta_4 \ln POP_{i,t} + \beta_5 INF_{i,t} \\ & + \beta_6 INF_{nl,t} + \beta_7 UNEMP_{i,t} + \beta_8 UNEMP_{nl,t} + \beta_9 GROWTH_{i,t} \\ & + \beta_{10} GROWTH_{nl,t} + \beta_{11} COLONY_i + \beta_{12} CONFLICT_{i,t} + \beta_{13} PUBLIC_{nl,t} \\ & + \beta_{14} POLICY_{nl,t} + \beta_{15} STOCK_{i,nl,t} + \varepsilon_{i,nl,t}\end{aligned} \tag{6.1}$$

First five variables are selected for a model to begin with. These variables are lnGDPi, lnGDPnl and lnDIST from the original Gravity Equation and CONFLICT and STOCK. This restricted model is regressed, this is shown in Table A.3. In the restricted pooled model, lnGDPi is positive and has a significant effect, this is not in line with the hypothesis, but it is in line with some literature. Mayda also finds these effects and argues that is could be due to the fact that poorer people will migrate if they become richer at a certain point or when they have a certain skill level. People who are poor will not migrate if they notice a growth in income. To check if this is true a regression will be performed with the squared of lnGDPi later on. lnGDPnl has

the expected sign, an increase in GDP per capita in the Netherlands results in more migrants, however it is insignificant. lnDIST has a negative and significant effect, this is in line with the hypothesis. CONFLICT and STOCK both have a strong and significant effect, which is in line with the hypotheses. When including time fixed effects it should noticed that lnGDPnl is now significant in the pooled model, but the other variables are still the same. In the RE model nothing has changed with the inclusion of time fixed effects. Country fixed effects can only be included in the pooled model, due to the underlying assumptions of the RE model. With country fixed effects and both country and time fixed effects lnGDPi seems to be insignificant, but this could be due to the fact that it is correlated with the country fixed effects.

In the appendix the tables per hypothesis are included. First, the inclusion of lnPOPi as a variable in Table A.4. In the pooled model a significant effect of a change in the population in the country of origin is seen. The inclusion of year fixed effects and random fixed effects show no difference compared to the basic model. In Table A.5 inflation of both the Netherlands and origin country is included. Without fixed effects one can see a significant effect of inflation in both types of countries. However, an increase in the Dutch inflation seems to increase the inflow. Both economics effects are also very low. It seems to be the case that migrants do not take inflation into account. Another explanation for the Netherlands is that in the Netherlands the inflation has small fluctuations mostly the inflation is fixed in a bounded interval. Including time fixed effect the Dutch inflation is insignificant, whereas with inclusion of country fixed effect alone and with time fixed effects seems to make the inflation in the host country insignificant. This could still be due to the fact that this is correlated with the country effects. When including the unemployment rates of both countries (Table A.6) one can see that in the pooled model with or without the time fixed effects the result is insignificant. With country fixed effect the results are significant for the unemployment in the country of origin. With both time and country fixed effects however the sign for the unemployment in the Netherlands is positive and significant. But lnPOPi is now negative, but they are likely to be correlated with the country fixed effects. While with the RE model they are insignificant with the time fixed effects. It should be noticed that time fixed effects seems to be highly correlated with some variables in the analysis. When including the growth rates (Table A.7) for both countries one can see that this effect is insignificant for most regressions, however in the RE model both growth rates are significant with an unexpected sign, the same goes for

the pooled model with country fixed effects. It could be argued that this is due to the fact that not all migrants take growth rates into account and the effect seems to be very small (a 1% increase decreases the inflow by 0.03%). There is little literature to confirm this, or to show what the effect is of economic growth on migration. Perhaps the unexpected signs could be due to the fact that growth is endogenous. Growth could be influenced by migration. Including COLONY (Table A.8) shows to have a significant and expected effect in most regressions. Except in the country fixed effect pooled model and the country and time fixed effect model, here the sign is unexpected. However this is a country specific variable and could thus be correlated with the country effects. In Table A.9 the public opinion is included, this shows no significant result in all regression expect for the country fixed effect pooled model and the RE model without the fixed effects. This could be due to the fact that most migrants do not know much about Dutch politics.

Table A.10 shows the results of these regressions with all the variables included. One can see that the inclusion of the migration policies have a significant and expected effect in the models without fixed effects, while in the time fixed effect models these are either insignificant of show an unexpected sign. However as noticed before, the time effects seem to have a high correlation with most variables, so these are not the most suitable models. The most suitable models are the pooled and RE model without fixed effects. The country fixed effects are correlated with certain country specific variables such as distance and colonial history. Distance is one of the key variables in the model so one cannot gain an insight from the country fixed effects models. Regarding the time fixed effects, most variables seem to have correlation with the years so these are also not very suitable. Perhaps the Mundlak model will be the most suitable to use. A closer look at this model will be given in section 6.4.

The final model including the variables that were added in Tables A.3 till A.11 will be looked upon. Table 6.2 shows the full models and shows the outcome of several autocorrelation and heteroskedasticity tests. Looking at the full pooled model without any fixed effects one can see that the signs of the initial five variables are as expected. Regarding the significance, lnGDPnl seems to have no effect. The unexpected significance, could be due to the fact that GDP per capita is used as an indicator of wages. lnPOPi shows that an increase in population in the country of origin increases migration flows, this is as expected because a larger population indicates a larger labour market that is saturated, and thus less available jobs.

INFLi has an expected positive effect and is significant, while INFLnl has an unexpected positive effect and is significant. These results are therefore difficult to explain. At first hand, the seemingly unexpected results could be due to the facts that most migrants do not take inflation in to account in their decision to migrate as mentioned before. Another explanation for the unexpected sign in the Netherlands, is that the Netherlands mostly has a small fluctuation in inflation rate. The effect in the Netherlands seems to have a small economic effect. As stated by Karemera et al. (2000): *"… may also represent substitution elasticities with respect to composite economic events that are not linked to emigration".* So the inflation measures could also be linked to other events that are not part of this analysis. The estimated coefficient of UNEMPi shows an unexpected sign but is significant. An explanation could be that unemployed migrants do not have the means to migrate. Another explanation could be that the pooled model is not suitable. A rise in UNEMPnl has an expected negative and significant sign. Larger unemployment rates in the destination country do influence a decision to migrate, the chances of not finding a job could be high. GROWTHi has an insignificant effect. This could be due to the fact that GDP per capita is already taken into account. Another explanation could be that migrants do not know much about growth. The same applies to the Netherlands where it is insignificant. As stated above, it could also be due to endogeneity, migration could influence growth. Small variations in growth could also cause this result. Another explanation could be that migrants compare both growth rates and choose the one that is higher. This shows in Table 5.2 where the average of GROWTHi is larger than GROWTHnl. COLONY has a significant and expected positive sign. Most of the migrants from these countries already know the culture and the language, so it is easier to adapt in the Netherlands. PUBLIC has no significant effect, perhaps because not many migrants know about Dutch politics. POLICY has a negative expected and significant sign. The stricter a migration policy is, the less migrants come to the Netherlands because it is more difficult to get a residence permit.

lnMP	Pooled full	Pooled full(robust)	Pooled full lnGDP²i(robust)	Random full	Random full(robust)	Random full lnGDP²i(robust)
lnGDPi	0.323*** (0.018)	0.323*** (0.059)	-0.263 (0.676)	0.269*** (0.053)	0.269*** (0.097)	-2.041*** (0.293)
lnGDPnl	-0.02 (0.741)	-0.02 (0.502)	0.087 (0.509)	0.899*** (0.311)	0.899** (0.447)	0.888*** (0.307)
lnDIST	-0.212*** (0.032)	-0.212* (0.116)	-0.167 (0.127)	-0.209* (0.116)	-0.209 (0.136)	-0.02 (0.117)
lnPOPi	0.654*** (0.019)	0.654*** (0.075)	0.656*** (0.076)	0.397*** (0.059)	0.397*** (0.102)	0.477*** (0.059)
lnGDP²i			0.037 (0.042)			0.15*** (0.019)
INFLi	0.001* (0.001)	0.001*** (0.001)	0.001*** (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
INFLnl	0.073** (0.036)	0.073*** (0.014)	0.074*** (0.014)	0.076*** (0.014)	0.076*** (0.012)	0.078*** (0.014)
UNEMPi	-0.009** (0.005)	-0.009 (0.013)	-0.003 (0.014)	0.015*** (0.005)	0.015 (0.01)	0.021*** (0.005)
UNEMPnl	-0.04 (0.03)	-0.04*** (0.015)	-0.039*** (0.015)	-0.023* (0.012)	-0.023* (0.013)	-0.024** (0.012)
GROWTHi	-0.001 (0.004)	-0.001 (0.009)	0.002 (0.009)	-0.001 (0.002)	-0.001 (0.003)	0.001 (0.002)
GROWTHnl	-0.011 (0.016)	-0.011 (0.009)	-0.012 (0.009)	-0.021*** (0.006)	-0.021*** (0.005)	-0.022*** (0.006)
COLONY	1.2*** (0.108)	1.2** (0.586)	1.212** (0.593)	1.089** (0.447)	1.089* (0.555)	1.156*** (0.444)
CONFLICT	0.466*** (0.072)	0.466** (0.216)	0.455** (0.217)	0.211*** (0.042)	0.211** (0.093)	0.183*** (0.041)
PUBLIC	0.019 (0.134)	0.019 (0.061)	0.008 (0.06)	-0.032 (0.052)	-0.032 (0.048)	-0.05 (0.051)
POLICY	-0.051** (0.022)	-0.051*** (0.012)	-0.052*** (0.012)	-0.064*** (0.009)	-0.064*** (0.009)	-0.067*** (0.009)
STOCK	0.518*** (0.033)	0.518*** (0.112)	0.53*** (0.11)	0.553*** (0.04)	0.553*** (0.164)	0.536*** (0.039)
_cons	-16.072** (7.873)	-16.072*** (5.496)	-15.429*** (5.542)	-21.518*** (3.224)	-21.518*** (4.723)	-15.784*** (3.264)
Country FE	NO	NO	NO	NO	NO	NO
Time FE	NO	NO	NO	NO	NO	NO
Obs	2,508	2,508	2,508	2,508	2,508	2,508
R²	0.585	0.585	0.589	0.549	0.549	0.543
Wooldridge	49.438***					
BP tests	135.18***					
BP-LM	13173.76***					

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table 6.2 - Regression outcome

As mentioned above in the analysis, perhaps the pooled model is not suitable for this analysis. To check this the BP-LM test has been performed. The statistic is shown in the Table above and it is highly significant. This means that the RE model is more suitable and the pooled model is not the right model.

The restricted random effect model shows the same significant results as the restricted pooled model, only lnDIST is less significant, but this is only marginal. The sign of lnGDPi is smaller, the signs of lnGDPnl and STOCK are larger, and the signs of lnGDPi, lnDIST, lnPOPi and CONFLICT are smaller. This is likely due to the fact that a more suitable model which takes the countries on their own is used. The same solution as mentioned above regarding the square of lnGDPi can be used here to see the effect of being richer and more likely to migrate. Looking at the full model one can see that most variables have the same significance and the same sign as in the pooled model. lnGDPnl is significant, lnDIST is less significant. INFLi is not significant, while INFLnl is more significant. UNEMPi does have the expected sign with the RE model.

Testing for heteroskedasticity and autocorrelation is somewhat difficult on a random effects model, because the proper test are not widely available. Therefore these test will be conducted on the pooled model as it should give an indication for the RE model as well. As shown in the Table 6.2 the Breusch Pagan test for heteroskedasticity is significant and thus the model shows heteroskedasticity. The Wooldridge test for autocorrelation also shows significance and thus the model also shows signs of autocorrelation. To correct for this, robust cluster standard errors will be used for the RE model. As shown in Table 6.2 the model still shows the same significant results for most variables. lnGDPnl is less significant and lnDIST is not significant. UNEMPi is also not significant anymore and COLONY and CONFLICT are now less significant but still significant at a certain level below the 10% or 5%. So the model performs quite well.

The proposed effect of the square of GDP per capita in the country of origin will now be looked upon. Table 6.2 shows these results for the full models of both the pooled and random effects. In the pooled model one can see that including the square of lnGDPi gives the expected sign and shows a positive sign for the square. This indicates that richer migrants tend to migrate more. Another explanation is that migrants who surpass the threshold can finally migrate.

GDP per capita instead of wages is used, so caution in the interpretation is advised. The signs of the other variables stay the same, although most variables are now insignificant compared to the robust pooled model. In the RE model only lnDIST, INFLi, GROWTHi and PUBLIC are now insignificant.

Because the RE model is the most suitable, one can also try and interpret from which threshold of expected income the migrants in the origin will migrate when their expected income increases. To calculate this, the derivative with respect to lnGDPi need to be taken and solved taking the first order condition for lnGDPi. In other words:

$$\frac{\delta \ln \frac{M}{P}}{\delta \ln GDP_i} = -2.041 + 2 * 0.150 \ln GDP_i = 0$$

$$\ln GDP_i = \frac{2.041}{2 * 0.150} = 6.803$$

$$e^{\ln GDP_i} = GDP_i = e^{6.803} = 900.85$$

So, when a migrant's income becomes more than 900 dollars, he or she will migrate to the Netherlands.

6.2. Using the Lagged Dependent Variable instead of the Stock

As mentioned earlier, most authors do not use the stock of migrants as a variable to measure information and family effects, but rather use the bilateral inflow of one period earlier. There is, like mentioned earlier in this thesis, one problem with the use of the lag of a dependent variable. The inclusion of a lag of the dependent variable will result in a dynamic model. Random effects models will therefore not be suitable, because of correlations with the random effect of the error term. That is why only the pooled estimation of the model, or a random model with an autocorrelation coefficient (also known as xtregar) can be looked upon. Robust errors will be included in the pooled model with the use of the Driscoll-Kraay standard errors.

Table 6.3 shows the regressions. In the pooled model one can see that, compared to the earlier pooled model, lnGDPnl has a significant effect and the sign is correct. Most variables have a smaller effect on the inflows of migration, while the effect of GROWTHi is more negative and significant. The sign of INFLi is intuitive, but it is insignificant and very small. The sign for inflation in both countries is still not in line with the expectations, however this could

still be due to the facts that most migrants do not take inflation in to account when they decide to migrate. Or just as with growth, they compare the inflations of both countries as the average inflation is the Netherlands is lower than that of the origin countries. UNEMPi is now insignificant. GROWTHnl has the expected sign, and is significant. COLONY has an insignificant effect. POLICY is now more significant. The inclusion of the lag dependent variable as an independent variable has increased the R^2 drastic, from around the 0.59 to 0.96. As one can see there is still a form of autocorrelation and heteroskedasticity as the statistics are highly significant. Using robust errors that correct for heteroskedasticity and autocorrelation, also known as Driscroll-Kraay errors does not change much in the outcome. The only difference is that lnGDPi, lnGDPnl, INFLnl and GROWTHnl are less significant, COLONY is now significant and PUBLIC is insignificant.

Using random effects with an autocorrelation term, one can notice some slight changes compared to the model with STOCK included and the pooled DK model. Compared to the DK model above most variables have a larger sign, while INFLi, UNEMPi and GROWTHnl change their sign. A few other variables have a smaller effect but these are very small, only lnDIST seems to have a more negative effect. The coefficient of the autocorrelation term has a smaller effect compared to the lag of the dependent variable. Regarding the significance, lnGDPi, lnGDPnl, INFLnl and UNEMPnl are more significant, UNEMPi and COLONY are now significant, while GROWTHi became insignificant. CONFLICT is the only variable that became less significant. Compared to the original RE STOCK model some variables are slightly different in the sign, but this is marginal. lnGDPnl, lnDIST, lnPOPi, INFLnl and COLONY have a larger effect, while CONFLICT seems to have a smaller effect. Here the effect of the autocorrelation term is larger than the effect of STOCK in the original RE model. Regarding the significance lnGDPnl, UNEMPi and COLONY are more significant, while CONFLICT is less significant, lnDIST is now significant and UNEMPnl is insignificant.

The square of lnGDPi in both the pooled DK model and the RE model with an autocorrelation term is also included in Table 6.3. As one can see including of $lnGDP^2i$ shows that being more rich will decrease the inflow, which is not in line with Mayda (2010), however both variables are insignificant. Compared to the model with STOCK instead of the lag, the lag lowers most variables, mostly because the lag is also determined by the lag of most dependent variables.

lnMP	Pooled	Pooled (DK errors)	Pooled lnGDP²i (DK)	Random AR	Random AR lnGDP²i
lnGDPi	0.026***	0.026*	0.065	0.311***	-1.806***
	(0.006)	(0.014)	(0.056)	(0.061)	(0.404)
lnGDPnl	1.307***	1.307*	1.3*	1.039***	1.125***
	(0.233)	(0.696)	(0.69)	(0.367)	(0.364)
lnDIST	-0.029***	-0.029***	-0.031***	-0.329***	-0.182
	(0.01)	(0.009)	(0.011)	(0.119)	(0.122)
lnPOPi	0.043***	0.043***	0.043***	0.728***	0.749***
	(0.008)	(0.007)	(0.007)	(0.063)	(0.063)
lnGDP²i			-0.003		0.135***
			(0.004)		(0.026)
INFLi	-0.001	-0.001	-0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
INFLnl	0.048***	0.048*	0.048*	0.048***	0.049***
	(0.012)	(0.026)	(0.026)	(0.011)	(0.011)
UNEMPi	-0.002	-0.002	-0.002	0.012**	0.017***
	(0.002)	(0.002)	(0.002)	(0.006)	(0.006)
UNEMPnl	0.012	0.012	0.012	-0.019	-0.017
	(0.01)	(0.015)	(0.015)	(0.012)	(0.012)
GROWTHi	-0.004***	-0.004***	-0.004***	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
GROWTHnl	0.02***	0.02*	0.02*	-0.014***	-0.016***
	(0.005)	(0.01)	(0.01)	(0.005)	(0.005)
COLONY	0.034	0.034*	0.033	1.417***	1.458***
	(0.036)	(0.021)	(0.021)	(0.453)	(0.452)
CONFLICT	0.052**	0.052***	0.053***	0.077*	0.065
	(0.024)	(0.019)	(0.018)	(0.045)	(0.045)
PUBLIC	-0.095**	-0.095	-0.094	-0.038	-0.049
	(0.043)	(0.116)	(0.115)	(0.046)	(0.046)
POLICY	-0.048***	-0.048***	-0.048***	-0.062***	-0.064***
	(0.007)	(0.015)	(0.015)	(0.009)	(0.009)
laglnMP	0.941***	0.941***	0.941***		
	(0.006)	(0.006)	(0.007)		
_cons	-14.772***	-14.772**	-14.822**	-27.532***	-22.108***
	(2.478)	(7.363)	(7.397)	(3.765)	(3.871)
RHO				0.698	0.694
Country FE	NO	NO	NO	NO	NO
Time FE	NO	NO	NO	NO	NO
Obs	2,664	2,664	2,664	2,690	2,830
R^2	0.958	0.958	0.958	0.528	0.512
Wooldridge	131.628***				
BP test	1070.97***				
BP-LM	N/A				

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table 6.3 - Regression outcome

Most variables that were insignificant are now significant and GROWTHi has the correct sign. However caution should be used in interpretation because the pooled model is not the suitable model. In the RE model lnGDPi has the expected sign and is significant, the same goes for the squared variable. More rich migrants tend to migrate when they have more income. lnDIST has a smaller effect and is insignificant, and CONFLICT also has no significant effect.

When comparing it to the model in the previous subsection the signs stay the same, some variables have a larger effect and some a smaller effect, but in the end the model shows almost the same results, only UNEMPnl and CONFLICT are insignificant. The autocorrelation term has a larger effect than STOCK.

The point on which migrants tend to migrate to the Netherlands can now be calculated. This goes the same as in the previous section.

$$\frac{\delta \ln \frac{M}{P}}{\delta \ln GDP_i} = -1.806 + 2 * 0.135 \ln GDP_i = 0$$

$$\ln GDP_i = \frac{1.806}{2 * 0.135} = 6.689$$

$$e^{\ln GDP_i} = GDP_i = e^{6.689} = 803.52$$

So when a migrant expects to earn more than 803 dollars of GDP per capita. This is slightly lower than the previous findings but it could be due to the fact that the lag variable if less suitable.

As one can see comparing the regressions, STOCK has a smaller effect, and the lag of the dependent variable has a much larger effect. However, the preference goes to the use of the stock of migrants over the use of the lag, due to the fact that the lag is highly correlated is with the inflow and it can only be used with the pooled model. If an event happens in a country that trigger migration, it is very likely that the next year this will also influence the inflow (e.g. the migration crisis in Europe). This is also explained by the lags of the independent variables. That way, the use of the stock instead of the lag is preferred.

6.3. The Mundlak Model

As stated in section 5.3. this thesis is dealing with a model that is perhaps more suitable for a fixed effects model than a random effects model, however due to the use of variables which have a small or no variation over time one is bounded to the RE model. As proposed by Mundlak (1978) one can combine a fixed and random model into the Mundlak model.

lnMP	Random full(robust)	Random full lnGDP²i(robust)	Mundlak	Mundlak incl GDPi squared	Mundlak incl STOCK squared
lnGDPi	0.269***	-2.041***	0.198**	-2.24***	0.105
	(0.097)	(0.293)	(0.077)	(0.327)	(0.075)
lnGDPnl	0.899**	0.888***	1.326***	1.069***	1.234***
	(0.447)	(0.307)	(0.328)	(0.326)	(0.321)
lnGDP²i		0.15***		0.165***	
		(0.019)		(0.022)	
lnDIST	-0.209	-0.02	-0.235*	-0.253*	0.048
	(0.136)	(0.117)	(0.133)	(0.143)	(0.136)
lnPOPi	0.397***	0.477***	0.166**	0.296***	0.069
	(0.102)	(0.059)	(0.07)	(0.071)	(0.074)
INFLi	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
INFLnl	0.076***	0.078***	0.077***	0.079***	0.077***
	(0.012)	(0.014)	(0.014)	(0.014)	(0.014)
UNEMPi	0.015	0.021***	0.018***	0.024***	0.018***
	(0.01)	(0.005)	(0.006)	(0.006)	(0.006)
UNEMPnl	-0.023*	-0.024**	-0.014	-0.02*	-0.015
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)
GROWTHi	-0.001	0.001	-0.001	0.001	-0.002
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
GROWTHnl	-0.021***	-0.022***	-0.023***	-0.024***	-0.021***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
COLONY	1.089*	1.156***	0.8*	0.917*	0.473
	(0.555)	(0.444)	(0.471)	(0.468)	(0.454)
CONFLICT	0.211**	0.183***	0.185***	0.16***	0.157***
	(0.093)	(0.041)	(0.042)	(0.041)	(0.04)
PUBLIC	-0.032	-0.05	-0.039	-0.054	-0.034
	(0.048)	(0.051)	(0.052)	(0.051)	(0.05)
POLICY	-0.064***	-0.067***	-0.066***	-0.068***	-0.064***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
STOCK	0.553***	0.536***	0.529***	0.509***	1.512***
	(0.164)	(0.039)	(0.041)	(0.041)	(0.092)
STOCK²					-0.115***
					(0.01)
_cons	-21.518***	-15.784***	522.694	548.015	1752.29
	(4.723)	(3.264)	(1229.719)	(1226.431)	(1190.968)
R^2	0.549	0.543	0.517	0.565	0.582
obs	2,508	2,508	2,508	2,508	2,508

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table 6.4 - Regression outcome

Table 6.4 shows the results of the Mundlak model. As can be seen, the results are quite in line with the above RE model. The effect of lnGDPi, lnPOPi and COLONY is lower than in the RE models, while the effect of lnGDPnl is higher. lnDIST is quite the same, however it is now significant. An increase in lnPOPi shows a lower effect. Most other variables have the same effect and significance as in the RE models. However UNEMPi now has a significant effect, while UNEMPnl is insignificant. It should be noted that the variables lnDIST, lnPOPi and COLONY have their RE estimates, all other variables are FE estimates. When including the square of lnGDPi not much changes, only UNEMPnl is now significant. The estimated expected income in the Mundlak model is now:

$$\frac{\delta \ln \frac{M}{P}}{\delta \ln GDP_i} = -2.240 + 2 * 0.165 \ln GDP_i = 0$$

$$\ln GDP_i = \frac{2.240}{2 * 0.165} = 6.788$$

$$e^{\ln GDP_i} = GDP_i = e^{6.788} = 887.03$$

So when a migrant expects to earn more than 887 dollars of GDP per capita. This is close to the calculations above.

Next to the marginal effect of income, another interesting marginal effect is that of the stock of bilateral migration in the Netherlands. Included in Table 6.3 is a Mundlak model with the square of STOCK. Because the Mundlak model and the RE model showed almost the same marginal effects for the square of GDP, only the Mundlak model is looked at, for this effect. The marginal effect of STOCK is as follow:

$$\frac{\delta \ln \frac{M}{P}}{\delta STOCK_i} = 1.512 - 2 * 0.115 STOCK_i = 0$$

$$STOCK_i = \frac{1.512}{2 * 0.115} = 6.574$$

In other words, when there is a stock 6.6 migrants per 1,000 inhabitants, migrants from that country will stop migrating to the Netherlands. An explanation for this could be that when it is full of migrants from their own country they fear a certain competition. Before the

threshold the increase in migrants could be explained due to migrants reuniting with their family or getting information that the Netherlands is a suitable country to migrate to.

Overall one can argue that the Mundlak model seems to be the better model for the analysis. Because of the inclusion of both fixed and random effects this model performs better compared to the other models.

6.4. The Hypotheses; to Reject or not to Reject

To conclude this section the hypotheses stated in the literature review are compared with the results found in this section. As stated in the above subsection the robust RE model of section 6.1 and the Mundlak model of section 6.3 are the most suitable models. Table 6.5 compares the found results with hypotheses and shows which are rejected and which are not.

Variable (country)	Hypothesis	Random Effects (+/-/0)	Mundlak (+/-/0)	Reject (yes/no)
GDP per cap (origin) **GDP per cap squared**	- 	+/- +	+/- +	Yes/No
GDP per cap (NL)	+	+	+	No
Population (origin)	+	+	+	No
Population (NL)	-	Dropped	Dropped	Dropped
Distance	-	0	-	Ambiguous
Inflation rate (origin)	+	0	0	Yes
Inflation rate (NL)	-	+	+	Yes
Unemployment rate (origin)	+	0/+	+	Ambiguous/No
Unemployment rate (NL)	-	-	0/-	Ambiguous/No
Growth rate (origin)	-	0	0	Yes
Growth rate (NL)	+	-	-	Yes
Unemployment benefits	+	Dropped	Dropped	Dropped
Colonial history	+	+	+	No
EU membership	+	Dropped	Dropped	Dropped
Political conflict	+	+	+	No
Public opinion	-	0	0	Yes
Migration policies	-	-	-	No
Stock of migrants	+	+	+	No

Tabel 6.5 – Hypotheses and outcome

As can be seen in Table 6.5 the hypothesis are not rejected for GDP per capita in the Netherlands, Population in the origin countries, colonial history, political conflict, migration policies and the stock of migrants. The hypothesis are rejected for the inflation rates and the growth rates for both countries and the public opinion. GDP per capita in the origin country is rejected but when controlled for the squared effect, the effect is as hypothesized. The effect of distance is ambiguous with or

without the inclusion of the square of GDP per capita, the same goes for the unemployment rates, although the hypotheses are not rejected when the squared effect of GDP per capita is included. Due to multicoliniarity problems the Dutch population, unemployment benefits and EU membership are dropped.

7. Conclusion

This final section will conclude this thesis. First, the most important parts of this thesis are summarized, followed by the results and the research question will be answered. Second, the policy implications are given. What can the government learn from this thesis for their policies regarding migration? Third, the drawbacks are discussed of this research. Finally, this thesis will be finished with suggestions for further research.

7.1. Summary

This thesis looked at different push and pull factors and how they determine migration towards the Netherlands. Section two looked at the seven laws for migration as stated by Ravenstein. Many of these laws are universally accepted as those that drive migration. A good summary of the main theory of migration is found in figure 2.1. Migration always has a set of push and pull factors that influence a migrant's decision. These decisions are also influenced by obstacles and the distance. The second part of section two looked at the main theory behind the model. The gravity equation, widely used in economic literature to describe trade flows between nations. Mostly it is used to determine what influences trade, however it is also a suitable model with a theoretical background as found by Anderson to determine bilateral migration. Section three looked at the findings of different studies over the years. What was surprising is that there is not much research regarding this topic, especially for the Netherlands. Most research only looks at international migrations among nations (in which they included the Netherlands) or it is based on the US and Canada. This thesis is one of the first in its field to look the Netherlands in specific and to look at the marginal effect of income in the origin country and the stock of bilateral migrants. As can be seen in section four, six databases are consulted for data on this topic, which formed the basis for the combined dataset to research what determines bilateral migration to the Netherland. Section five discussed the main assumption behind the models used in this thesis. Just as assumed the most proper model that needed to be used was the Random Effects model for panel analysis, but a different model for Random Effects; namely the Mundlak model, was also considered. Section six, the key section of this thesis, showed the results regarding the analysis. After a thorough analysis on the variables, the most suitable for the regression analysis were selected. It was then shown that the population in the Netherlands, the composition of the Dutch cabinet and the dummy for the EU membership were highly correlated with other

explanatory variables, and were thus dropped. Looking at different regressions and using different tests, the random effects model was the most suitable to use. Running multiple regression this thesis showed what the effect was of the variables on the inflow of migrants to the Netherlands.

One can see that a rise in the GDP per capita, which is an indication for income, resulted in an increase in the flow of migrants. This was not in line with the expectations. However, one could argue that it could be due to the fact that richer migrants tend to migrate if their income rises. It could also be that these migrants can save enough to travel and find a better life. Poorer migrants do not usually have the resources to migrate. After including the square of GDP per capita it was shown that migrants that earn approximately 900 or more dollars will be more likely to come to the Netherlands. As expected, an increase in the GDP per capita in the Netherlands will result in an increase of the inflow of migrants. Strong elasticities for this variable were found. A rise in the population of the country of origin resulted in an increase of migrants. A larger population usually indicates a more saturated labour market. In the model that does not correct for the richer migrants, that distance negatively influences inflows, but correcting for the richer migrants shows no effect depending on the model. This is most likely due to the resources they have. A rise in inflation in the country of origin results in an increase of migrants to the Netherlands in some of the models. However, in other it has no effect, thus this result is quite ambiguous. Inflation in the Netherlands also gave a surprising outcome as it increases migration. However these two results could be due to the fact that migrants almost have no idea about inflation in the country of destination or that they compare both inflation rates with each other. The elasticities were also quite small. Unemployment in the country of origin showed a positive effect. This however was not significant in the initial model, so the effect is quite ambiguous. It is likely that this has an influence on the decision to migrate. Unemployment in the Netherlands does have a negative effect. If the unemployment rate rises, migrants are less likely to migrate to the Netherlands, as it would reduce their chances on finding a job. However, the effect differs when the square of GDP per capita is included. The same assumption of inflation also holds for economic growth in the country of origin. This has no effect, while growth in the Netherlands results in a decrease. It could be that this is because the GDP per capita is included, or that migrants do not take economic growth into account in their decision to migrate, or that they compare

both rates. If the country of origin has a colonial history with the Netherlands this will result in an increase in migrants. More migrants come from these countries. However this is only a small fraction of the data. The effect of the public opinion does not influence the inflow, one could argue that this could be because most migrants have little information about local politics. If a country has a political conflict, this will increase the inflow. Stricter migration policies also show an effect in the reduction of migrants. Finally there is a family and information effect. A rise in the stock of migrants results in a larger inflow. However, this effect decreases the more migrants are living in the Netherlands. When there are approximately 6.6 migrants per 1,000 inhabitants the inflow reduces. This is most likely the effect of competition among migrants.

To conclude and answer the main research question; the Dutch GDP per capita, the population in the country of origin, a colonial history, a political conflict and the stock of migrants increase in the inflow, while the unemployment rate in the Netherlands and stricter migration policies tend to reduce the inflow.

7.2. Policy implications

What can the government learn from this thesis? As is shown above, this thesis also included some policy variables. If the government wants to reduce the inflow of migrants they should make stricter policies. This thesis showed that this is an effective way of reducing the inflow of migrants. This is not always easy due to agreements like the European agreement of free movement of labour and could result in some international friction between countries. Helping solve a conflict in the country of origin and investing in this also could stop the inflow. It is also shown that the bilateral migration inflow will slow itself down, because there is a diminishing effect. After a certain amount of bilateral migrants the inflows will reduce.

7.3. Limitations

As is the case in all research, this thesis also has some limitations. One of the mayor limitations is that a strongly balanced dataset is used, which means that there are some variables missing. This could have biased some results. One should notice that this is also one of the earliest researches on migration to the Netherlands, there is not much research to compare with. When compared, the authors only include the Netherlands as a small part of their research. Some countries that are included could also have distorted variables as they could have

reported variables more rosy, but this is not something that could be corrected for. It should also be stated that GDP per capita is not quite the same as income, on the contrary it is the only measure that is available. Other variables could also have measured differently, which could result in different results, this problem is always the case with statistics. Caution in interpreting the results is therefore advised. Economic growth could also be an endogenous variable as it could be influenced by migration. Another limitation is that the random effects model has some drawbacks, as not all test could be conducted. There are no tests that look at autocorrelation and heteroskedasticity for the random effects model yet. Robust errors like the Driscoll-Kraay errors are not suitable for random effects models. Another problem is that fixed effects seems to be better for these researches, however due to certain variables this could not have been performed.

7.4. Suggestions for Further Research

There are some suggestions for further research. A first suggestions could be to perform a more in-depth research. By looking at the decision to migrate to a country. One could interview different migrants on their decisions of travelling to the Netherlands, and then perform a binary regression analysis on their choice. This would be time consuming, but gives a more individual view on migration decisions. This thesis only looked at the general influences. Another suggestion is that in future research one could try and measure average wages in both the host and origin country instead of GDP per capita. One could also try and see if there is a difference between high-skilled and low-skilled migrants or in the wake of recent events, if the model also holds for refugees. Regarding some of the variables like inflation, growth and unemployment rates, further research could include the difference between these variables as it could show that migrants tend to compare countries.

8. Literature

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American Economic Review*, *69*(1), 106-116.

Anderson, J. E. (2010). *The gravity model* (No. w16576). National Bureau of Economic Research.

Ayo, E. (2015, September 3). *Europe's migrant crisis by the numbers*. Retrieved from Foreign Policy: http://foreignpolicy.com/2015/09/03/europes-migrant-crisis-by-the-numbers/

Bergstrand, J. H. (1985). The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The review of economics and statistics*, 474-481.

Borjas, G. J. (1989). Economic theory and international migration. *International migration review*, 457-485.

Breusch, T. S., and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, 1287-1294.

Centraal Bureau voor Statistiek (2015). Statline Database [Data file]. Retrieved from http://statline.cbs.nl/Statweb/.

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

DEMIG (2015) *DEMIG POLICY, version 1.3, Online Edition*. Oxford: International Migration Institute, University of Oxford. www.migrationdeterminants.eu

Drukker, D. M. (2003). Testing for serial correlation in linear panel-data models.*Stata Journal*, *3*(2), 168-177.

Embrace. (2013). *Why do people migrate?* Retrieved from Embrace: https://www.embraceni.org/migration/why-do-people-migrate/

European Union. (2016). EU member countries - European Union website, the official EU website - European Commission. Retrieved August 16, 2016, from https://europa.eu/european-union/about-eu/countries/member-countries_en

Feenstra, R. C. (2004). Increasing returns and the gravity equation. *Advanced international trade: theory and evidence. Princeton University Press, Princeton*.

Greenwood, M. J. (1997). Internal migration in developed countries. *Handbook of Population and Family Economics*, *1*, 647-720.

Greenwood, M. J., and McDowell, J. M. (1991). Differential economic opportunity, transferability of skills, and immigration to the United States and Canada. *The Review of Economics and Statistics*, 612-623.

Giulietti, C., Guzi, M., Kahanec, M., & Zimmermann, K. F. (2013). Unemployment benefits and immigration: evidence from the EU. *International Journal of Manpower*, *34*(1), 24-38.

Haas, H. de, Natter, K., and Vezzoli, S. (2014). Compiling and coding migration policies: Insights from the DEMIG POLICY database. *International Migration Institute, DEMIG project paper*, *16*, 43.

Hill, R. C., Griffiths, W. E., and Lim, G. C. (2012). *Principles of Econometrics*. Hoboken, NJ: Wiley.

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

Leblang, D. A., Fitzgerald, J., and Teets, J. (2009). Defying the law of gravity: The political economy of international migration. *Available at SSRN 1421326*.

Lee, E. S. (1966). A theory of migration. *Demography*, *3*(1), 47-57.

Lewer, J. J., and Van den Berg, H. (2008). A gravity model of immigration.*Economics letters*, *99*(1), 164-167.

Marshall, M. G. (2015). COUP D'ÉTAT EVENTS, 1946-2014 CODEBOOK Monty G. Marshall and Donna Ramsey Marshall. *Center for Systemic Peace*.

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

Mayer, T. and Zignago, S. (2011). Notes on CEPII's distances measures: the GeoDist Database. CEPII Working Paper 2011-25

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: Journal of the Econometric Society*, 69-85.

OECD (2010). Public Opinion and Immigration: Individual Attitudes, Interest Groups and the Media. *International Migration Outlook*: SOPEMI 2010. Eds. Garson, J-P., J-C. Dumont, G. Lemaitre, 115-155. Paris: OECD.

OECD (2016). International Migration Database [Data file].Retrieved from https://stats.oecd.org/Index.aspx?DataSetCode=MIG.

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

Payne, A. (2016, March 20). *MAPPED: The growth of the far-right in Europe*. Retrieved August 16, 2016, from Business Insider, http://uk.businessinsider.com/map-shows-far-right-growth-across-europe-2016-3.

Peeperkorn, M. (2015, September 7). *EU: 7.214 vluchtelingen extra naar Nederland.* Retrieved from de Volkskrant: http://www.volkskrant.nl/buitenland/eu-7-214-vluchtelingen-extra-naar-nederland~a4136900/

Pöyhönen, P. (1963). A tentative model for the volume of trade between countries. *Weltwirtschaftliches Archiv*, 93-100.

Ramos, R., and Suriñach, J. (2013). *A Gravity Model of Migration between ENC and EU* (No. 7700). Institute for the Study of Labor (IZA).

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

Ravenstein, E. G. (1889). The laws of migration. *Journal of the Royal Statistical Society*, *52*(2), 241-305.

Ridder, J. den, Mensink, W., Dekker, P., and Schrijver, E. (2016). Continu onderzoek burgerperspectieven. *Kwartaalbericht 2016*.

Tinbergen, J. (1962). Shaping the world economy; suggestions for an international economic policy. *Books (Jan Tinbergen)*.

Tranos, E., Gheasi, M., and Nijkamp, P. (2015). International migration: a global complex network. *Environment and Planning B: Planning and Design*, *42*(1), 4-22.

The World Bank, World Development Indicators (2016). WDI 2016 [Data file]. Retrieved from http://databank.worldbank.org/data/download/WDI_excel.zip.

9. Appendix

Countries			
Afghanistan	Denmark	Laos, People's Democratic Republic	Rwanda
Albania	Djibouti	Latvia	Saudi Arabia
Algeria	Dominican Republic	Lebanon	Senegal
Angola	Ecuador	Lesotho	Sierra Leone
Argentina	Egypt	Liberia	Serbia
Armenia	El Salvador	Libya	Singapore
Australia	Equatorial Guinea	Lithuania	Slovak
Austria	Eritrea	Luxembourg	Slovenia
Azerbaijan	Estonia	Macedonia	Solomon Islands
Bahrain	Ethiopia	Madagascar	Somalia
Bangladesh	Fiji	Malawi	South Africa
Belarus	Finland	Malaysia	Spain
Belgium	France	Mali	Sri Lanka
Benin	Gabon	Mauritania	Sudan
Bhutan	Gambia	Mauritius	Suriname
Bolivia	Georgia	Mexico	Swaziland
Bosnia and Herzegovina	Germany	Moldova	Sweden
Botswana	Ghana	Mongolia	Switzerland
Brazil	Greece	Montenegro	Syria
Bulgaria	Guatemala	Morocco	Taiwan
Burkina Faso	Guinea	Mozambique	Tajikistan
Burundi	Guinea Bissau	Myanmar	Tanzania
Cambodia	Guyana	Namibia	Thailand
Cameroon	Haiti	Nepal	Timor-Leste
Canada	Honduras	New Zealand	Togo
Cape Verde	Hungary	Nicaragua	Trinidad and Tobago
Central African Republic	India	Niger	Tunesia
Chad	Indonesia	Nigeria	Turkey
Chile	Iran	Norway	Turkmenistan
China	Iraq	Oman	Uganda
Colombia	Ireland	Pakistan	Ukraine
Comoros	Israel	Panama	United Arab Emirates
Congo	Italy	Papua New Guinea	United Kingdom
Costa Rica	Jamaica	Paraguay	United States
Ivory Coast	Japan	Peru	Uruguay
Croatia	Jordan	Philippines	Uzbekistan
Cuba	Kazakhstan	Poland	Venezuela
Cyprus	Kenya	Portugal	Vietnam
Czech Republic	Korea, Republic of	Quatar	Yemen
Korea, People's Democratic Republic	Kuwait	Romenia	Zambia
Congo (Demogratic Republic)	Kyrgyzstan	Russian Federation	Zimbabwe

Table A.1. - List of countries

	lnMP	lnPOPi	lnPOPnl	lnGDPi	lnGDPnl	lnDIST	INFLi
lnMP	1						
lnPOPi	0.5848*	1					
lnPOPnl	0.0557*	0.0536*	1				
lnGDPi	0.3101*	-0.0562*	0.0879*	1			
lnGDPnl	0.0519*	0.0496*	0.9316*	0.0837*	1		
lnDIST	-0.2416*	0.0440	-0.0000	-0.4779*	0.0000	1	
INFLi	0.0171	0.0241	-0.0987*	-0.0643*	-0.1023*	0.0151	1
INFLnl	0.0023	-0.0293	-0.4753*	-0.0476*	-0.3115*	0.0000	0.0166
UNEMPi	-0.0814*	-0.1654*	-0.0368	0.0228	-0.0448	-0.1596*	0.0001
UNEMPnl	-0.0330	-0.0051	-0.1859*	-0.0086	-0.4563*	0.0000	0.0645*
GROWTHi	-0.0346	0.0026	-0.0362	-0.0455	-0.0311	0.0026	-0.0112
GROWTHnl	-0.0445	-0.0299	-0.6070*	-0.0437	-0.3980	0.0000	0.0464
BENEFITS	-0.0516*	-0.0419	-0.8285*	-0.0684*	-0.9076*	0.0000	0.1014*
COLONY	0.1282*	-0.0883*	0.0000	0.0917*	-0.0000	-0.0629*	-0.0086
EU	0.3329*	0.0210	0.0813*	0.5080*	0.0775*	-0.6629*	-0.0378
CONFLICT	0.2382*	0.3434*	-0.0682*	-0.2336*	-0.0695*	0.1258*	0.0944*
PUBLIC	0.0494*	0.0412	0.8346*	0.0676*	0.8476*	-0.0000	-0.0858*
GOV	0.0043	0.0137	0.3487*	0.0189	0.2769*	-0.0000	-0.0445
POLICY	-0.0061	0.0203	0.4550*	0.0353	0.5921*	-0.0000	-0.0845*
STOCK	0.4600*	0.2630*	-0.0116	0.1960*	-0.0145	-0.3417*	-0.0031
	INFLnl	**UNEMPi**	**UNEMPnl**	**GROWTHi**	**GROWTHnl**	**BENEFITS**	**COLONY**
INFLnl	1						
UNEMPi	0.0263	1					
UNEMPnl	-0.4747*	0.0056	1				
GROWTHi	0.0164	0.0187	0.0540*	1.0000			
GROWTHnl	0.3319*	0.0004	-0.0568*	0.1322*	1.0000		
BENEFITS	0.1435*	0.0315	0.6258*	0.0699*	0.4934*	1.0000	
COLONY	-0.0000	0.0460	0.0000	-0.0063	0.0000	0.0000	1.0000
EU	-0.0628*	-0.0184	0.0054	-0.0562*	-0.0342	-0.0575*	0.1016*
CONFLICT	0.0205	-0.0550*	0.0328	-0.0339	0.0359	0.0668*	-0.0006
PUBLIC	-0.0702*	-0.0232	-0.4817*	-0.0140	-0.4930*	-0.8616*	0.0000
GOV	-0.2031*	0.0052	-0.1378*	0.0535*	-0.0892*	-0.3636*	0.0000
POLICY	0.0400	-0.0258	-0.5115*	0.0391	0.0711*	-0.5910*	-0.0000
STOCK	-0.0008	0.0357	0.0166	0.0095	0.0029	0.0163	0.0230
	EU	**CONFLICT**	**PUBLIC**	**GOV**	**POLICY**	**STOCK**	
EU	1						
CONFLICT	-0.1609*	1					
PUBLIC	0.0583*	-0.0636*	1				
GOV	0.0192	-0.0215	0.4564*	1			
POLICY	0.0370	-0.0373	0.5905*	0.3713*	1		
STOCK	0.2603*	0.0503*	-0.0157	-0.0119	-0.0167	1.0000	

Table A.2. - Correlation Table

*Note: * means a significant effect at 0.01*

lnMP	1	2	3	4	5	6
lnGDPi	0.329*** (0.021)	0.331*** (0.021)	0.035 (0.066)	0.064 (0.067)	0.140*** (0.052)	0.167*** (0.052)
lnGDPnl	0.549 (0.351)	1.700** (0.740)	1.110*** (0.149)	2.473*** (0.258)	0.951*** (0.138)	2.281*** (0.25)
lnDIST	-0.126*** (0.038)	-0.125*** (0.038)	-4.887*** (0.446)	-4.936*** (0.427)	-0.333** (0.132)	-0.306** (0.128)
CONFLICT	1.382*** (0.082)	1.390*** (0.082)	0.168*** (0.041)	0.19*** (0.039)	0.192*** (0.041)	0.214*** (0.040)
STOCK	0.829*** (0.037)	0.830*** (0.037)	0.469*** (0.038)	0.467*** (0.037)	0.498*** (0.038)	0.498*** (0.036)
_cons	-12.663*** (3.732)	-24.96*** (7.818)	26.986*** (4.346)	12.831*** (4.768)	-13.586*** (1.619)	-28.120*** (2.666)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.347	0.349	0.933	0.939	0.278	0.286
obs	2,668	2,668	2,668	2,668	2,668	2,668

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.3 – Regression outcome restricted model

lnMP	1	2	3	4	5	6
lnGDPi	0.334*** (0.017)	0.336*** (0.017)	0.023 (0.066)	-0.068 (0.068)	0.173*** (0.049)	0.214*** (0.049)
lnGDPnl	-0.106 (0.287)	0.902 (0.603)	2.012*** (0.189)	4.198*** (0.312)	0.554*** (0.147)	1.751*** (0.26)
lnDIST	-0.212*** (0.031)	-0.211*** (0.031)	-6.316*** (0.479)	-6.704*** (0.46)	-0.34*** (0.115)	-0.303*** (0.113)
lnPOPi	0.632*** (0.017)	0.631*** (0.017)	-0.915*** (0.12)	-1.244*** (0.131)	0.361*** (0.056)	0.373*** (0.056)
CONFLICT	0.514*** (0.071)	0.522*** (0.071)	0.142*** (0.041)	0.153*** (0.039)	0.199*** (0.042)	0.220*** (0.04)
STOCK	0.509*** (0.031)	0.51*** (0.031)	0.452*** (0.038)	0.434*** (0.036)	0.495*** (0.037)	0.497*** (0.036)
_cons	-15.024*** (3.046)	-25.818*** (6.364)	45.361*** (4.93)	31.781*** (5.094)	-15.368*** (1.561)	-28.898*** (2.634)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.565	0.569	0.934	0.941	0.530	0.541
obs	2,668	2,668	2,668	2,668	2,668	2,668

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.4 – Regression outcome with population of the origin country added

lnMP	1	2	3	4	5	6
lnGDPi	0.337***	0.337***	0.099	-0.059	0.241***	0.224***
	(0.017)	(0.017)	(0.067)	(0.068)	(0.050)	(0.049)
lnGDPnl	0.333	0.129	1.94***	1.782*	0.731***	0.374
	(0.304)	(2.482)	(0.19)	(0.93)	(0.147)	(0.960)
lnDIST	-0.214***	-0.212***	-0.757***	-6.627***	-0.286**	-0.292***
	(0.031)	(0.031)	(0.126)	(0.470)	(0.115)	(0.111)
lnPOPi	0.631***	0.631***	-6.023***	-1.256***	0.412***	0.381***
	(0.017)	(0.017)	(0.49)	(0.132)	(0.056)	(0.055)
INFLi	0.000**	0.000**	0.000	0.000	0.000*	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
INFLnl	0.093***	-0.194	0.04***	-0.614**	0.072***	-0.348
	(0.023)	(0.679)	(0.01)	(0.253)	(0.009)	(0.262)
CONFLICT	0.503***	0.510***	0.139***	0.141***	0.193***	0.210***
	(0.071)	(0.071)	(0.041)	(0.039)	(0.041)	(0.040)
STOCK	0.509***	0.509***	0.460***	0.434***	0.502***	0.498***
	(0.031)	(0.031)	(0.038)	(0.036)	(0.037)	(0.036)
_cons	-19.876***	-17.366	40.331***	57.661***	-19.201***	-14.151
	(3.241)	(27.259)	(5.080)	(11.238)	(1.620)	(10.569)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.569	0.57	0.935	0.941	0.549	0.545
obs	2,663	2,663	2,663	2,663	2,663	2,663

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.5 – Regression outcome with inflation added

lnMP	1	2	3	4	5	6
lnGDPi	0.336***	0.336***	0.126*	-0.021	0.279***	0.235***
	(0.017)	(0.017)	(0.071)	(0.068)	(0.051)	(0.049)
lnGDPnl	-0.109	0.623	2.272***	4.318***	0.494**	1.604***
	(0.429)	(0.844)	(0.268)	(0.391)	(0.205)	(0.348)
lnDIST	-0.214***	-0.213***	-0.870***	-6.635***	-0.230**	-0.265**
	(0.031)	(0.031)	(0.133)	(0.466)	(0.117)	(0.112)
lnPOPi	0.628***	0.627***	-6.107***	-1.289***	0.420***	0.370***
	(0.018)	(0.018)	(0.486)	(0.131)	(0.059)	(0.057)
INFLi	0.000**	0.000**	0.000	0.000	0.000*	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
INFLnl	0.058*	-0.076	0.047***	-0.036	0.050***	-0.073
	(0.033)	(0.174)	(0.013)	(0.064)	(0.013)	(0.066)
UNEMPi	-0.005	-0.006	0.024***	0.018***	0.017***	0.012**
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
UNEMPnl	-0.042	0.017	0.015	0.088***	-0.024**	0.041
	(0.028)	(0.087)	(0.012)	(0.032)	(0.011)	(0.033)
CONFLICT	0.502***	0.510***	0.125***	0.133***	0.186***	0.205***
	(0.071)	(0.071)	(0.040)	(0.039)	(0.041)	(0.040)
STOCK	0.511***	0.512***	0.462***	0.437***	0.509***	0.501***
	(0.031)	(0.031)	(0.038)	(0.036)	(0.037)	(0.036)
_cons	-14.821***	-22.742**	39.037***	29.666***	-17.578***	-28.059***
	(4.684)	(9.127)	(5.106)	(5.635)	(2.104)	(3.631)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.564	0.564	0.945	0.945	0.534	0.53
Obs	2,639	2,639	2,639	2,639	2,639	2,639

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.6 – Regression outcome with unemployment rates added

lnMP	1	2	3	4	5	6
lnGDPi	0.336*** (0.018)	0.336*** (0.018)	0.073 (0.077)	0.013 (0.076)	0.288*** (0.053)	0.271*** (0.052)
lnGDPnl	-0.491 (0.461)	0.209 (1.526)	2.055*** (0.275)	3.412*** (0.586)	0.006 (0.209)	1.133* (0.581)
lnDIST	-0.208*** (0.033)	-0.207*** (0.033)	-6.462*** (0.48)	-6.69*** (0.467)	-0.209* (0.117)	-0.222* (0.115)
lnPOPi	0.625*** (0.019)	0.625*** (0.019)	-1.168*** (0.141)	-1.338*** (0.138)	0.398*** (0.059)	0.374*** (0.058)
INFLi	0.001* (0.001)	0.001* (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
INFLnl	0.071** (0.034)	-0.046 (0.177)	0.069*** (0.013)	-0.039 (0.064)	0.066*** (0.013)	-0.018 (0.066)
UNEMPi	-0.007 (0.005)	-0.008* (0.005)	0.021*** (0.006)	0.018*** (0.006)	0.015*** (0.005)	0.013** (0.005)
UNEMPnl	-0.029 (0.03)	0.011 (0.083)	0.04*** (0.013)	0.079*** (0.03)	-0.011 (0.012)	0.031 (0.031)
GROWTHi	0.001 (0.004)	0.001 (0.004)	-0.002 (0.002)	-0.001 (0.002)	-0.002 (0.002)	-0.001 (0.002)
GROWTHnl	-0.02 (0.013)	0.012 (0.071)	-0.042*** (0.005)	0.003 (0.026)	-0.03*** (0.005)	0.02 (0.027)
CONFLICT	0.518*** (0.074)	0.523*** (0.074)	0.138*** (0.041)	0.146*** (0.04)	0.206*** (0.042)	0.218*** (0.041)
STOCK	0.534*** (0.034)	0.534*** (0.034)	0.502*** (0.041)	0.482*** (0.04)	0.565*** (0.04)	0.554*** (0.039)
_cons	-10.815** (5.001)	-18.302 (16.688)	49.714*** (5.129)	40.494*** (7.577)	-12.321*** (2.159)	-23.737*** (6.314)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.566	0.569	0.945	0.944	0.531	0.530
Obs	2,508	2,508	2,508	2,508	2,508	2,508

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.7 – Regression outcome with economic growth rates added

lnMP	1	2	3	4	5	6
lnGDPi	0.323*** (0.018)	0.323*** (0.018)	0.073 (0.077)	0.013 (0.076)	0.285*** (0.053)	0.267*** (0.052)
lnGDPnl	-0.483 (0.45)	0.238 (1.49)	2.055*** (0.275)	3.412*** (0.586)	-0.014 (0.208)	1.117* (0.581)
lnDIST	-0.213*** (0.032)	-0.212*** (0.032)	-6.462*** (0.48)	-6.69*** (0.467)	-0.199* (0.115)	-0.212* (0.114)
lnPOPi	0.653*** (0.019)	0.653*** (0.019)	-1.168*** (0.141)	-1.338*** (0.138)	0.421*** (0.059)	0.395*** (0.058)
INFLi	0.001* (0.001)	0.001* (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
INFLnl	0.07** (0.033)	-0.045 (0.172)	0.069*** (0.013)	-0.039 (0.064)	0.066*** (0.013)	-0.018 (0.066)
UNEMPi	-0.009** (0.005)	-0.009** (0.005)	0.021*** (0.006)	0.018*** (0.006)	0.015*** (0.005)	0.013** (0.005)
UNEMPnl	-0.029 (0.029)	0.011 (0.081)	0.04*** (0.013)	0.079*** (0.03)	-0.011 (0.012)	0.031 (0.031)
GROWTHi	-0.001 (0.004)	0.001 (0.004)	-0.002 (0.002)	-0.001 (0.002)	-0.002 (0.002)	-0.001 (0.002)
GROWTHnl	-0.021 (0.013)	0.012 (0.07)	-0.042*** (0.005)	0.003 (0.026)	-0.029*** (0.005)	0.02 (0.027)
COLONY	1.202*** (0.108)	1.2*** (0.108)	-1.652*** (0.494)	-2.095*** (0.482)	1.097** (0.442)	1.09** (0.437)
CONFLICT	0.463*** (0.072)	0.468*** (0.072)	0.138*** (0.041)	0.146*** (0.04)	0.206*** (0.042)	0.218*** (0.041)
STOCK	0.519*** (0.033)	0.519*** (0.033)	0.502*** (0.041)	0.482*** (0.04)	0.564*** (0.04)	0.553*** (0.039)
_cons	-11.25** (4.883)	-18.957 (16.29)	49.714*** (5.129)	40.494*** (7.577)	-12.589*** (2.153)	-23.998*** (6.314)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.585	0.585	0.941	0.944	0.552	0.550
Obs	2,508	2,508	2,508	2,508	2,508	2,508

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.8 – Regression outcome with colonial history added

lnMP	1	2	3	4	5	6
lnGDPi	0.323*** (0.018)	0.323*** (0.018)	0.061 (0.077)	0.013 (0.076)	0.28*** (0.053)	0.267*** (0.052)
lnGDPnl	-0.114 (0.74)	2.371 (17.276)	2.943*** (0.362)	4.571 (6.374)	0.718** (0.314)	-0.238 (6.624)
lnDIST	-0.213*** (0.032)	-0.212*** (0.032)	-6.48*** (0.479)	-6.69*** (0.467)	-0.202* (0.115)	-0.212* (0.114)
lnPOPi	0.653*** (0.019)	0.653*** (0.019)	-1.188*** (0.14)	-1.338*** (0.138)	0.417*** (0.059)	0.395*** (0.058)
INFLi	0.001* (0.001)	0.001* (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
INFLnl	0.08** (0.036)	-0.022 (0.155)	0.091*** (0.014)	-0.026 (0.058)	0.085*** (0.014)	-0.033 (0.06)
UNEMPi	-0.009** (0.005)	-0.009** (0.005)	0.022*** (0.006)	0.018*** (0.006)	0.016*** (0.005)	0.013** (0.005)
UNEMPnl	-0.028 (0.029)	0.029 (0.112)	0.043*** (0.013)	0.089** (0.042)	-0.009 (0.012)	0.019 (0.044)
GROWTHi	-0.001 (0.004)	0.001 (0.004)	-0.002 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)
GROWTHnl	-0.025* (0.014)	0.012 (0.07)	-0.052*** (0.006)	0.003 (0.026)	-0.037*** (0.006)	0.02 (0.027)
COLONY	1.201*** (0.108)	1.2*** (0.108)	-1.701*** (0.493)	-2.095*** (0.482)	1.095** (0.444)	1.09** (0.437)
CONFLICT	0.463*** (0.072)	0.468*** (0.072)	0.137*** (0.041)	0.146*** (0.04)	0.206*** (0.042)	0.218*** (0.041)
PUBLIC	-0.08 (0.127)	-0.467 (4.065)	-0.182*** (0.049)	-0.254 (1.497)	-0.155*** (0.05)	0.296 (1.558)
STOCK	0.519*** (0.033)	0.519*** (0.033)	0.498*** (0.041)	0.482*** (0.04)	0.562*** (0.04)	0.553*** (0.039)
_cons	-15.129* (7.87)	-41.39 (181.069)	40.936*** (5.621)	28.299 (66.787)	-20.162*** (3.257)	-9.764 (69.39)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.585	0.585	0.941	0.944	0.550	0.550
Obs	2,508	2,508	2,508	2,508	2,508	2,508

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.9 – Regression outcome with public opinion added

lnMP	1	2	3	4	5	6
lnGDPi	0.323***	0.323***	0.031	0.013	0.269***	0.267***
	(0.018)	(0.018)	(0.076)	(0.076)	(0.053)	(0.052)
lnGDPnl	-0.02	3.041	3.241***	3.194	0.899***	0.653
	(0.741)	(10.908)	(0.357)	(4.021)	(0.311)	(4.182)
lnDIST	-0.212***	-0.212***	-6.601***	-6.69***	-0.209*	-0.212*
	(0.032)	(0.032)	(0.471)	(0.467)	(0.116)	(0.114)
lnPOPi	0.654***	0.653***	-1.27***	-1.338***	0.397***	0.395***
	(0.019)	(0.019)	(0.138)	(0.138)	(0.059)	(0.058)
INFLi	0.001*	0.001*	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
INFLnl	0.073**	-0.019	0.082***	-0.032	0.076***	-0.029
	(0.036)	(0.114)	(0.014)	(0.042)	(0.014)	(0.044)
UNEMPi	-0.009**	-0.009**	0.021***	0.018***	0.015***	0.013**
	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)
UNEMPnl	-0.04	0.032	0.03**	0.083**	-0.023*	0.023
	(0.03)	(0.099)	(0.012)	(0.037)	(0.012)	(0.038)
GROWTHi	-0.001	0.001	-0.002	-0.001	-0.001	-0.001
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)
GROWTHnl	-0.011	0.011	-0.034***	0.005	-0.021***	0.019
	(0.016)	(0.06)	(0.006)	(0.022)	(0.006)	(0.023)
COLONY	1.2***	1.2***	-1.907***	-2.095***	1.089**	1.09**
	(0.108)	(0.108)	(0.486)	(0.482)	(0.447)	(0.437)
CONFLICT	0.466***	0.468***	0.141***	0.146***	0.211***	0.218***
	(0.072)	(0.072)	(0.041)	(0.04)	(0.042)	(0.041)
PUBLIC	0.019	-0.63	-0.047	0.083	-0.032	0.079
	(0.134)	(2.6)	(0.05)	(0.957)	(0.052)	(0.997)
POLICY	-0.051**	0.005	-0.071***	-0.009	-0.064***	0.006
	(0.022)	(0.078)	(0.008)	(0.029)	(0.009)	(0.03)
STOCK	0.518***	0.519***	0.485***	0.482***	0.553***	0.553***
	(0.033)	(0.033)	(0.04)	(0.04)	(0.04)	(0.039)
_cons	-16.072**	-48.404	40.479***	42.72	-21.518***	-19.08
	(7.873)	(114.385)	(5.529)	(42.343)	(3.224)	(43.835)
Year	NO	YES	NO	YES	NO	YES
Country	NO	NO	YES	YES	NO	NO
R^2	0.586	0.585	0.943	0.944	0.547	0.550
obs	2,508	2,508	2,508	2,508	2,508	2,508

*Notes: ***<0.01; **<0.05; *<0.10; Standard errors in parentheses*

Table A.10 – Regression outcome with migration policies added