Hilde Aspenberg Jordal

# Globalization and Democracy

## – a factor based approach

Master's thesis in Economics
Supervisor: Ragnar Torvik
Trondheim, June 2017

Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

**NTNU**
Norwegian University of
Science and Technology

# Abstract

This thesis uses panel data for 217 countries between 1960 and 2007 to examine how relative factor endowments of countries affect whether democratization is an option when a non democratic country starts opening up to globalization. The analysis is based on the theory by Acemoglu and Robinson (2006), and uses a wide specter of econometric methods to see if the data points in the same direction as the model. In accordance with the model, the countries in the data set are divided into two groups: one group containing labor abundant countries, and the other groups containing capital and land abundant countries. The statistical results indicate that globalization is positively associated with democratization among labor abundant countries, but that globalization is negatively associated with democratization in capital and land abundant countries. The result are robust, and thus this thesis concludes that the evidence in support of Acemoglu and Robinsons (2006) argument is present.

# Preface

This Master thesis concludes a 2 year Master's Degree in Economics at the Department of Economics, NTNU. I would like to thank my supervisor Ragnar Torvik for valuable and good feedback, whose door was always open for the questions I had along the way. Furthermore, I would like to thank Marie Skara for proof reading chapter after chapter, and Elisabeth Werenskiold for much needed English rectification. I also like to thank the Institute of Economics for a memorable five years, and all who have made my years as a student fantastic.

Trondheim, 1. June 2017.

Hilde Aspenberg Jordal

# Contents

# List of Figures

# List of Tables

# 1   Introduction

In 1950, there were only 23 democracies in the world. By 2009, the number was $91^1$. During this period, an increase in trade and financial flows made most of the countries in the world more economically integrated. Has globalization promoted democracy around the world? In an April 2002 speech, President Bush said that "trade creates the habits of freedom," and those habits "begin to create the expectations of democracy and demands for better democratic institutions. Societies that are open to commerce across their borders are more open to democracy within their borders."

Typically, increasing wealth is the main part of a democratization story (Doces & Magee, 2015). The main argument that globalization and democratization are interrelated is mainly because increased openness raises per capita incomes, thus leading to democratization. This thesis takes this argument as a starting point, but develops it further, and tests the proposition that the democratization forces associated with trade openness depend on each country's relative factor endowment. This argument is put forward in the Acemoglu and Robinson (2006) model of democracy. This model has not yet been precisely tested. The main difference between this paper and previous work on the same topic, is that this paper focuses on both capital, land and labor abundant countries.

In the absence of international trade, locally abundant factors has lower prices. If a country has a lot of capital or arable land, and this is the country's abundant factor, the returns to this factor are lower than it would be if the economy was open. When opening up, or globalizing, the returns will increase to the world market price, and the ones owning the factor will get relatively richer. If the abundant factor is owned by the rich in the country, they will get even richer, and the income distribution will become increasingly more askew.

If the country is abundant in labor, opening up the economy will increase the returns to labor, and the agents who own labor will get relatively richer. In any country, labor is owned by the laborer who is also the main consumer of the economy. Globalization thereby increases their income, making the income distribution more even.

Acemoglu and Robinson (2006) theorize this by creating a model. They assume that the rich own the land and capital, and the poor own their own labor. The poor are

---

[1]Democracy is here defined as having a polity score of at least 6 in the Polity IV data set.

the majority of the population in the country. They use the median voter theorem to show that if there was a democracy, the redistributive tax rate the median voter wants, is the one implemented.

If a country is non democratic and abundant in land and capital, opening up the economy for international trade will make the rich even richer, thus making democratization less likely. This is because globalization will lead to a relatively poorer median voter, thus a higher redistributive tax rate. The rich will oppose this higher redistributive tax rate and democratic concessions will become less likely. If the non democratic country is abundant in labor, opening up the economy will make the poor richer, thus making democratization more likely. This is because the median voter is relatively richer than before, leading to a lower redistributed tax rate.

This thesis shows that the relative factor abundance conditions the political effects of trade. The statistical estimates indicate that higher levels of trade in labor abundant countries are associated with higher levels of democracy. However, in capital and land abundant countries the statistical results indicate that trade is negatively associated with democracy. The results are robust to at number of robustness tests.

The outline of the thesis is as follows; First, empirical examples and the definitions of globalization and democratization are discussed. Then, previous literature on the topic is reviewed, paying attention to the studies that are most closely related to the factor-based argument. Next, the model by Acemoglu and Robinson (2006) is examined, where the level of democracy depends on the relative factor endowments, and how democracy and trade are connected. Before the regression analyses, chapter 5 addresses the data used in the empirical approach. Then, chapter 6 looks at the empirical approach, including the conditions for the pooled OLS estimation, the fixed effects estimations and the two stage least square estimations. Chapter 7 discusses the results, and includes the instrumental variable tests and the results from the hypothesis. Chapter 8 is a robustness analysis, including an alternative econometric approach, alternative time periods, additional explanatory variables and interaction terms. Finally, the conclusion is presented in chapter 9, and additional information on the data set and results from robustness are found in the appendix.

# 2    Definitions and case studies

Daniel T. Griswold (2004) has commented that a number of economies have, in the past decades, followed the path of economic and trade reform leading to polit-

ical reform. South Korea and Taiwan were in the 1980s governed by authoritarian regimes that left little room for open political protest and dissent. In the 2000s, after years of expanding trade and rising income, both are democracies with full political and civil liberties. Other countries that have followed the path of reform include Chile, Ghana, Hungary, Mexico, Nicaragua, Paraguay, Portugal and Tanzania. This chapter will look at the definitions of globalization and democratization, and some empirical examples.

## 2.1   Globalization

Albrow and King (1990) define globalization as the action or procedure of international integration arising from the interchange of world views, products, ideas and other aspects of culture. The International Monetary Fund (IMF) define four aspects of globalization: trade and transactions, capital and investment movements, migration and movement of people, and the spread of knowledge. The focus of this thesis is economic globalization, which is the historical process of economic integration, the result of human innovation and technological process. Economic globalization refers to the increasing integration of economies around the world, particularly through trade and financial flows (IMF, 2000).

The historical process of globalization can be looked at through waves. The first wave of globalization began in the early 19th century with the end of the Napoleonic Wars in Europe. The representative government took root in the United States, Great Britain, Switzerland, France, several small European states, and Britain's overseas dominions such as Canada, Australia and New Zealand (Griswold, 2004). The second wave began after the Second World War with the democratization of the defeated power, such as West Germany, Italy, Austria, Japan, and in Uruguay, Brazil, Costa Rica and India (Huntington, 1991). The third wave began after the fall of the Berlin Wall, in the late 1980's (Huntington, 1991).

This thesis focuses on the definition that globalization is the increase in the integration of economies around the world, especially focusing on trade.

## 2.2   Democratization

Democratization is the transition to a more democratic political regime, or a substantive political change moving in a democratic direction. Acemoglu and Robinson (2006) consider democratization as the transitory nature of de facto power (i.e. the

actual political power). Acemoglu and Robinson (2006) divide the population into two groups: The citizens (the poor majority) and the elite (the rich minority). In a non democracy, the elites have de jure power (i.e. the legislative power), and if they are unconstrained, they will choose the policies they prefer; for example, low tax rates and no redistribution to the poor. However, in a non democracy sometimes the citizens may pose a revolutionary threat. This is when they have a temporarily de facto power. The citizens could use their de facto power to undertake a revolution and change the system to their benefit. This revolution can create massive losses to the elites, but also significant collateral damage and social losses. The elites would like to prevent this outcome, and they can do it by transferring the formal political power to the majority, the citizens, which is what the democratization process achieves (Acemoglu & Robinson, 2006). This is illustrated by different countries' political histories, such as the British, Argentinian and South Africa, and follows waves of democratization.

The first wave of democratization occurred at the same time as the first wave of globalization, in the 1840s. This began with Britain's unilateral free trade and lasted until the First World War. This wave of democracy lasted until the 1930s, but was interrupted by Benito Mussolini whom rose to power in Italy in 1922. This was at the same time as countries turned away from the protectionism of the 1930s and opened up for more unilateral trade. In contrast to this, periods known as "reverse waves" (Huntington, 1991), also occurred resulting in the retreat of democracy, and increased protectionism. During the first reverse wave, in the 1920s and 1930s, the industrialized countries closed their economies and focused on economic nationalism. The second reverse wave, in the 1960s and early 1970s, saw less developed countries also closing their economies and embracing import substitution policies, as well as showing increasing hostility to foreign investment (Griswold, 2004).

The second wave of democracy took place after the Second World War. The Allies in World War II initiated a second wave of democratization that reached its peak in 1962, with 36 countries that were defined as democracies. This was the results of successive rounds of negotiation in the General Agreement on Tariffs and Trade (GATT) established in 1947. The agreement was signed by 23 states. The second reversal wave from 1960 to 1975 then brought the number of democracies back down to 30 (Huntington, 1991).

The third wave of democracy began in 1974 and included the democratic transition in Latin America in the 1980s, the Asian Pacific countries, such as Philippines, South Korea and Taiwan, from 1986 to 1988, Eastern and Central European countries

following the demise of the Soviet Union in 1989, and sub-Saharan African countries (Huntington, 1991). The third wave of democracy is associated with an increase in the world globalization.

The recent protest movement in the Middle East and North Africa, known as the "Arab Spring", has many similarities with the fall of the Soviet Union, resulting in these movements being denoted by some as the fourth wave of democratization. However, this movement has not been linear. An example of this is when the military in Egypt took control after the fall of the dictator Hosni Mubarak, and consequently stopped any attempts at a democratic transition. Scholars are therefore still not in complete agreement as to whether the "Arab Spring" constitutes a forth wave.

## 2.3   Empirical examples

There are countries where economic freedom and globalization are positively associated with democratization, and some countries where this is not yet the case, but a possibility.

### 2.3.1   Chile

Chile is an example of how globalization was a factor in the democratic process. Chile had a socialist and protectionist government before 1973. The country was experiencing high levels of inflation because of expropriations, price controls and protectionism. By the end of 1973, the nominal average tariff on imports was at 105 % and the inflation was at 150 %. In 1973, Augusto Pinochet lead a coup against the elected socialist government, and turned Chile into an autocracy. The autocratic government welcomed foreign investment, eliminated trade barriers, and privatized state owned companies. Buc (2006) has pointed out that this liberalization of the economy, that lead to a rise in income in the population, was a main factor in facilitating the democratic elections in 1989. Today, Chile is one of the OECD countries, and have one of the freest economies in the world (Buc, 2006).

### 2.3.2   South Korea

South Korea is another example of globalization as a factor in the democratic process. At the beginning of the 1960s, developing countries aimed to bring their economies to the level of the industrialized countries. One of these countries was

South Korea. South Korea was authoritarian in the aftermath of the Korean war, and the GDP per capita was at 1100 $. A coup in 1961, lead the country to be controlled by the military. This lead to a five year plan that had the purpose to free the economy by mobilizing natural resources and establishing a self supporting industrial economy. Cumings (2005) indicates that this five year plan was a factor that lead to the increase in the income of the regular South Korean, and to the democratic transition in 1987. Now, the GDP per capita is 24000 $, they are a member of the OECD and one of the G-20 major economies (Cumings, 2005).

### 2.3.3  Mexico

Mexico is also an example of globalization as a factor in the democratic process. In Mexico, more than a decade of economic and trade reforms constitute the foundation for the historic election in 2000. In this election, the opposition party was elected, ending 71 years of one-party rule by the Institutional Revolutionary party (PRI). Internal economic reforms, entry into the General Agreement of Tariffs and Trade (GATT) in 1986, and the North American Free Trade Agreement in 1994, increased the income of the regular Mexican and undermined the dominance of the PRI, being a factor in the process of a regime change (Griswold, 2004).

### 2.3.4  China

The economic development of China has not yet lead to a democratization process. However, China is one of the most striking examples of the impact of opening an economy to the global market. Over the last half century, China has undergone a shift from a largely agrarian society to an industrial powerhouse. China was labor abundant before the economic reforms in 1978. The reforms lay the ground work for future growth by allowing farmers to sell their produce in local markets and began the shift from collective farming to the household responsibility system (i.e. a system where surplus agricultural production were sold at unregulated prices). A year later a law was introduced, allowing foreign capital to enter China helping the regional economies. However, it was not until the mid 1980s that the government began to gradually ease pricing restrictions and allow companies to retain profits and set up their own wage structures. Although as stated, the economic globalization of China has not lead to democratization, the GDP per capita has increased since the globalization, and only time will show if China will become a democracy (Hirst, 2015).

### 2.3.5  Cuba

The development of the relationship between Cuba and the United States in recent years has lead to an increase in the globalization of Cuba. Cuba has been a self-proclaimed socialist state since 1959, which outlaws political pluralism, represses dissent, and severely restricts freedom of the press, assembly, speech, and association. The government, first lead by Fidel Castro and now lead by his brother Raúl Castro, monopolizes economic activity within centralized and inefficient state enterprises. The increased engagement with the United States did not result in lifting the restrictions by the government (Freedom House, 2017). However, will the globalization in Cuba lead to a democratization process in the coming years in the same way as Chile, South Korea and Mexico?

# 3  Literature

The idea that globalization promotes the spread of democratic ideas goes back to at least Kant (1795). Schumpeter (1950), Lipset (1959) and Hayek (1960) all argued that free trade and capital flows fosters the demand for democracy. This is done by making resource allocation more efficient, thus raising income and increasing economic development. The existing research on globalization and democracy varies remarkable in content and conclusions. Some authors argue that globalization leads to democratization, while others argue the inverse. Several authors conclude that there is a two-way causality, or simultaneity, between the two phenomena. At the same time, various authors also argue that there is no relationship between the two at all.

## 3.1  Globalization leads to democracy

Doces (2006) looks at the relationship between globalization and democracy, where he defines globalization as exports from a less developed country to the United States (U.S.). The hypothesis is based on the Hecksler-Ohlin model, and developed based on the income gain accruing to the abundant factor, labor, in the developing country. He argues that laborer's income gain raises their marginal benefit from political activity and thus leads to more democracy. As the less developed country gets more democratic, it will be rewarded with easier access to the U.S. market. This relationship between democracy and globalization is modeled as a simultane-

ous process. Doces (2006) tests the hypothesis that exports to the U.S. increases democracy in the exporting country with case studies of countries such as South Korea, Bangladesh, Mexico and the Philippines, and combines this with an econometric analysis. Doces (2006) comes to the conclusion that an increase in exports to the U.S. leads to a higher level of democracy in the less developed country. He also uses a Granger causality analysis, and this reveals statistical support in favor of the two-way causality between exports and democracy.

López-Córdova and Meissner (2005) study whether increased international trade leads to democracy. They focus on the endogeneity between the two variables, globalization and democracy, and use an instrumental variable for globalization to account for this. The variable they use is based on the gravity model. Because the gravity model includes time-invariant factors, such as bilateral distance, they conduct a series of cross-section analyses at various points in time. They find a positive impact of globalization, or openness, on democracy from 1895 and onwards. However, this is a significant decreasing effect with time. As an extension they consider the effect of factor endowments and find limited support for the Acemoglu and Robinson (2006) proposition using two measures of factor abundance, including land-to-population and capital-to-worker ratios. More capital-per-worker leads to trade openness, so that globalization is positively associated with democracy.

Other authors such as Rudra (2005), Maxfield (2000), Papaioannou and Siourounis (2008) and Yu (2007) all conclude that globalization has a positive effect on democracy. Rudra (2005) concludes that the relationship between capital flows and democracy is conditional on the amount of social spending. Maxfield (2000, p. 96) focuses on capital flows, arguing that financial liberalization strengthens democracy because it diversifies "economic power" and creates demand for transparency. Papaioannou and Siourounis (2008) limit their initial sample to non democratic countries and conclude that trade openness plays a significant role in driving transitions to democracy. Yu (2007) rejects the hypothesis that democratization leads to openness and conclude in favor of the opposite. He rationalizes his findings by observing that concentrated interests may be better able to secure the imposition of protectionist policies in democratic political systems where they are better represented.

Doces and Magee (2015) test the model by Acemoglu and Robinson (2006), where the relative factor endowments determine whether trade promotes democracy or not. They use 142 countries between 1960 and 2007, and use an instrumental variable approach to the endogeneity problem between globalization and democracy. They

use two instruments to solve this. The first instrument is average level of openness among the country's ten nearest neighbors. The second instrument is the predicted level of the country's bilateral trade flows (as a share of GDP) by using the gravity model. The standard argument about globalization and democracy is that increased openness raises per capita income and thus contributes to democratization. Doces and Magee (2015) further test the proposition that the key democratizing forces associated with trade openness depend on each country's relative factor endowment. They determine if a country is relatively capital or labor abundant, thus providing a more accurate test of the Acemoglu and Robinson argument that the relative factor abundance conditions the effect of trade on democracy. Their statistical estimates indicate that higher levels of trade in labor abundant countries are associated with increased democracy, but capital abundant countries openness is negatively associated with democracy. Their results are not robust. They use a different measure of relative factor abundance than Ahlquist and Wibbels (2012). Ahlquist and Wibbels (2012) rely on Leamer's (1984) definition of relative factor abundance, which they operationalize as the proportion of country $i$'s population divided by its proportion of world gross domestic product (GDP). This measure captures how far ahead or below a country is, relative to world GDP per capita. It says nothing explicit about capital stocks, and furthermore it says nothing about the relative abundance of each country. This measure does position each country relative to the world median, and thus it measures its relative factor abundance. A critique of the paper is that they do not account for the land abundance of the country, only the capital abundance, and thus this does not test the theory of Acemoglu and Robinson (2006) properly.

## 3.2 Democracy leads to globalization

Giavazzi and Tabellini (2005) consider a large country sample and measures of democracy taken from the Polity data set, and they report a positive effect of democracy on trade liberalization. Milner and Kubota (2005) find that democracy promotes trade openness. They do this with an instrumental variable strategy that admits the possibility of causality running in the opposite direction. The instruments for democracy they use are the average age of the party system and the level of secondary school completion. However, they mostly use OLS regression, and argue that the reversed causality running from trade openness to the political regime is unlikely to be important.

## 3.3   Two-way causality

Eichengreen and Leblang (2008) account for the simultaneous relationship between trade and democracy and find that they are mutually reinforcing. They also find support for O'Rourke and Taylor's (2006) proposition that democratization leads to protection rather than trade openness in relatively labor-scarce countries. They base their results on data from 1870 to 2000, which broadens over time. The measure in use for globalization is imports plus exports as a percentage of gross domestic product, and the measure in use for capital flows is IMF's annual report on Exchange Arrangements and Exchange Restrictions, supplemented with historical sources. They measure democracy by the measure created by Cheibub, Przeworski, Limongi Neto, and Alvarez (1996), who argue that a country should be regarded as democratic if governments are chosen in contested elections. They also use POLITY IV as a robustness check. The econometric approach is GMM (Generalized Method of Moments). They use the gravity model as an instrument for trade openness, the number of other countries with open capital markets in place the preceding year as an instrument for openness to capital flows, and the years since independence as an instrument for democracy. They perform different checks for robustness: Alternative instruments, interaction terms, alternative econometric specifications, alternative periods and alternative country samples. Which all lead to similar results, that the relationship between democracy and trade is mutually reinforcing.

## 3.4   Negative or no relationship

Li and Reuveny (2003) find a negative relationship between trade openness and democracy. They measure the degree of globalization as openness to trade, FDI, and portfolio capital flows, as well as the number of countries that constitute a democracy in the country's region, in order to capture the expansion of democratic ideas across borders. They control for the endogeneity between the two variables by using lagged values of the globalization variable. López-Córdova and Meissner (2005) argue that there are several econometric specification problems in Li and Reuveny's (2003) paper. First, if there are serial correlation in the error terms in the regression, which is likely, then using lags in the globalization variable does not solve the endogeneity problem. This is also the case with the democracy variable, where the serial correlation could lead the OLS estimates to be biased. In addition, they use fixed effects and lagged variables in the same model, which is a misspecification of the econometric model. Second, using all globalization variables at the same

time could lead to multicollinearity. Third, including income per capita and GDP growth could lead to another endogeneity problem since income per capita affects GDP growth, and GDP growth affects income per capita. Finally, since there is one channel where globalization or trade could affect democracy through its impact on economic development, or income per capita, including this regressor makes the model not take this channel into account.

Bussmann (2001) uses a simultaneous equation to model the relationship between democracy and globalization. The author does not find a significant impact of globalization on democracy. She uses a two stage least square estimation to take into account the endogenous relationship between the variables. Her use of instruments could be questioned, as she uses instruments that are not exogenous to democracy, which is the dependent variable in her regression.

O'Rourke and Taylor (2006) argue on the basis of the Stolper-Samuelson theorem: In countries where labor is relatively scarce, democratic reforms that raise labor's leverage over policy will encourage protectionism rather than opening to the rest of the world. They use historical data from the pre-1913 wave of globalization. They argue that democratization broadens the extent of the franchise and should encourage trade openness in labor abundant countries, since laborers, who now vote, benefit from trade liberalization. For the same reason democratization will have the opposite effect in labor scarce countries.

Ahlquist and Wibbels (2012) conclude that there is little, if any, relationship among factor endowments, trade and democracy. This paper contradicts the Acemoglu and Robinson (2006) argument. They study the effect of world trade openness on democracy, arguing it is exogenous to democracy. Doces and Magee (2015) argue that this assumption is not tenable. This is because politically powerful countries, like the US, can and do affect the degree of global openness through trade-supporting institutions like the WTO, IMF and World Bank. This can also happen trough the provision of public goods, like safe and open international waters, all of which are necessary for an open trading system. Importantly, most of the trade-supporting countries are democracies. This means that democracy, especially the number in the world system, influences the openness of the world economy. Thus, global trade is unlikely to be exogenous to democracy, as Ahlquist and Wibbels (2012) claim. Doces and Magee (2015) argue that the use of world trade openness further biases Ahlquist and Wibbels (2012) sample, as global trade is dominated by a few mostly wealthy democracies, excluding numerous poor countries that have largely been left out of the process.

Ahlquist and Wibbels (2012) use world trade as a measure of globalization because they believe it provides a more compelling theoretical argument about the relationship between trade and democracy. They argue that it need not matter if a country trades a lot or a little for the distributive implications underpinning most arguments in the literature to bite. Rather, the degree of global openness is the key factor affecting democracy at the country level. Ahlquist and Wibbels (2012) argue that the aspect of globalization driving democracy is the openness of the global economy, not the level of openness within the domestic economy. This argument, according to Doces and Magee (2015), suggests that countries such as China and its neighbor North Korea should be experiencing the same pressure to democratize due to increasing world openness. In contrast, Doces and Magee (2015), think that the connection between openness and democracy via factor returns depends primarily on the degree of openness at the domestic level. Global trade flows will not impact the distribution of income within a country if that country is not open and engaged in globalization. In contrast, a country's distribution of income and the pressures within it for democracy are affected by a change in the country's own level of openness even if the degree of global openness remains the same.

# 4   Theory

This thesis tests the empirical predictions about the effect of globalization on democracy taken from the Acemoglu and Robinson (2006) model of democracy. The model presents a theory of democratization based on the economic conflict between the citizens and the elites. The citizens are the majority of the population, so in a democracy with one vote per person, the citizens would have the majority (i.e. the median voter theorem), and are able to execute the policies they prefer. This involves a higher transfer from higher-income elite to lower-income citizens. The elite prefer a different policy, and have greater control over policies in non democracies, than in democracies. The country starts in a state of non democracy, and during a certain period in time the citizens gain de facto power (the actual political power), but not de jure power (the legislative power). In order to prevent a revolution, where the model assumes that the elite lose all of their assets, the elite are forced to make democratic concessions to the citizens. This means that the citizens get more favorable economic policies, but that they also hold more of the de jure power in the future. An example of this theory is the "Arab Spring", where citizens in many Middle Eastern and African countries gained de facto power that forced the elite to

grant more democracy.

## 4.1   The model by Acemoglu and Robinson (2006)

To connect this idea to globalization the theory is modeled along the lines of the Hecksler-Ohlin model of international trade. This is when the non democratic country is scarce in capital and abundant in labor relative to the rest of the world. However, the model also takes into account that some non democratic countries are also relatively abundant in land. As the Stolper-Samuelson theorem shows, international trade benefits the relatively abundant factor in the country and harms the scarce factor. This means that greater openness helps workers and harms capital owners in labor abundant countries. Openness harms workers and helps capital owners in capital abundant countries. The model first focuses on a single country, which is initially assumed to be closed to international trade. Then the country is opened to international trade, and the model looks at the political conflict between the elite and the citizens.

### 4.1.1   A closed economy

The aggregate production function is:

$$Y = (Y_K + \sigma Y_L)^\theta (Y_N)^{1-\theta} \qquad (1)$$

Where Y is the output of the final good that is consumed, and $Y_K$, $Y_L$ and $Y_N$ are the amounts of the three intermediate goods used in the production of Y. It is assumed that $0 < \theta < 1$ and $\sigma > 0$.

In a closed economy without trade, the intermediate goods are themselves produced by domestic factors of production. The intermediate goods have different subscriptions, which indicate that one is capital intensive, $Y_K$, one is land intensive, $Y_L$, and one is labor intensive, $Y_N$. In an open economy, the intermediate goods are traded internationally. All three goods are produced using their own factors, therefore, domestic production of each intermediate good is given by:

$$Y_K = K, Y_L = L, Y_N = 1 - \delta \qquad (2)$$

Where $\delta$ is the elite in the society, and $1 - \delta$ is the workers. The elite do not own

any labor, but holds a fraction, $\delta$ of the total capital stock, K, and of the total land stock, L. The population is normalized to 1.

When there is no international trade, the country in question has to use its domestic production of capital, land and labor intensive intermediate goods to produce production. Substituting (2) into (1):

$$Y = (K + \sigma L)^\theta (1 - \delta)^{1-\theta}$$

The model assumes that all markets are perfectly competitive, and set the price of final output to be equal to 1 and uses this as a numeraire. The prices of the intermediate goods are denoted as $p_K$, $p_L$ and $p_N$. To determine these prices, the model looks at a cost-minimizing problem of a firm choosing input demand to minimize the cost of production. Formally,

$$\min_{Y_K, Y_L, Y_N} \quad \{p_K Y_K + p_L Y_L + p_N Y_N\}$$
$$\text{subject to} \quad Y = (Y_K + \sigma Y_L)^\theta Y_N^{1-\theta}$$

Where $p_K Y_K + p_L Y_L + p_N Y_N$ is the total cost of using the three intermediate goods. This minimization problem can be solved with the use of a Lagrangian function:

$$\mathcal{L} = p_K Y_K + p_L Y_L + p_N Y_N - \lambda[(Y_K + \sigma Y_L)^\theta Y_N^{1-\theta} - Y]$$

The first order conditions:

$$\lambda\theta(Y_K + \sigma Y_L)^{\theta-1}Y_N^{1-\theta} = p_K$$
$$\lambda\theta\sigma(Y_K + \sigma Y_L)^{\theta-1}Y_N^{1-\theta} = p_L \quad\quad (3)$$
$$\lambda(1-\theta)(Y_K + \sigma Y_L)^{\theta}Y_N^{-\theta} = p_N$$

Which leads to:
$$\frac{p_K}{p_N} = \frac{\theta}{1-\theta}\frac{Y_N}{Y_K + \sigma Y_L} \quad \text{and} \quad \frac{p_K}{p_L} = \frac{1}{\sigma} \quad\quad (4)$$

Which implies that:

$$p_K = \frac{p_L}{\sigma} = \theta\left(\frac{Y_N}{Y_K + \sigma Y_L}\right)^{1-\theta} \quad \text{and} \quad p_N = (1-\theta)\left(\frac{Y_N}{Y_K + \sigma Y_L}\right)^{-\theta} \quad\quad (5)$$

Because one unit of each factor is used to produce its respective good. Factor markets are competitive, where each factor is paid the value of its marginal product, which gives us:

$$w = p_N, \; r = p_K \text{ and } v = p_L$$

Where $w$ is the wage rate, $r$ is return to capital and $v$ is the rental rate of land. In a closed economy, the model uses (2) and (5) together, and get:

$$
\begin{aligned}
r = p_K &= \theta\left(\frac{K + \sigma L}{1 - \delta}\right)^{\theta - 1} \\
v = p_L &= \sigma\theta\left(\frac{K + \sigma L}{1 - \delta}\right)^{\theta - 1} \\
w = p_N &= (1 - \theta)\left(\frac{K + \sigma L}{1 - \delta}\right)^{\theta}
\end{aligned}
\tag{6}
$$

These are the relative prices because they are measured in terms of the final good. The shares of national income accruing to the three factors are given as:

$$
\begin{aligned}
s_K &= \frac{rK}{Y} = \theta\frac{K}{K + \sigma L} \\
s_L &= \frac{vL}{Y} = \theta\frac{\sigma L}{K + \sigma L} \\
s_N &= \frac{wN}{Y} = 1 - \theta
\end{aligned}
\tag{7}
$$

The share of national income accruing to labor is a constant, so if for example capital accumulates and the real wages increase, the share of labor in the national income is constant. At the same time, the share of capital in the national income increases and the share of land decreases.

Total income is $(K + \sigma L)^{\theta}(1 - \delta)^{1-\theta}$ and because the model assumes that the total population is equal to 1, the average income is:

$$\bar{y} = (K + \sigma L)^{\theta}(1 - \delta)^{1-\theta} \tag{8}$$

Taking into account that citizens only have labor income, the income to the citizens are:

$$y^p = (1 - \theta)\left(\frac{K + \sigma L}{1 - \delta}\right)^{\theta} = \frac{(1 - \theta)(K + \sigma L)^{\theta}(1 - \delta)^{1-\theta}}{1 - \delta} = \left(\frac{1 - \theta}{1 - \delta}\right)\bar{y} \tag{9}$$

The model assumes that all members of the elite are homogeneous and own both capital and land. This leads to:

$$y^r = \frac{rK + vL}{\delta} = \frac{\theta}{\delta}(K + \sigma L)^{\theta}(1 - \delta)^{1-\theta} = \frac{\theta}{\delta}\bar{y} \tag{10}$$

This gives the income of a member of the elite, keeping the assumption that they are richer than the average (i.e., $\theta > \delta$).

There is a political conflict between the elite and the citizens, and this is shown in their preferred level of taxation. The utility of an individual $i$ is $(1-\tau)y^i + T$, where $i=p,r$. The government's budget constraint is

$$T = \delta\tau y^r + (1-\delta)\tau y^p - C(\tau)\bar{y} = (\tau - C(\tau))\bar{y}$$

The indirect utility of a poor agent is then $V(y^p|\tau) = (1-\tau)y^p + (\tau - C(\tau))\bar{y}$, and maximizing this leads to:

$$-y^p + (1 - C'(\tau^p))\bar{y} = 0 \quad \text{and} \quad \tau^p > 0 \tag{11}$$

This is because $y^p < \bar{y}$. Including (8) and (9) in (11) leads to:

$$\left(\frac{\theta - \delta}{1 - \delta}\right) = C'(\tau^p) \tag{12}$$

This gives the preferred tax rate for the poor in a closed economy, where both sides are positive because $\theta > \delta$, and because the poor have less income than the rich.

### 4.1.2 An open economy

Now assume that the country joins the world trading system and can trade with other countries in the world without any friction, assumed as one aspect of globalization. Since there is only one final good produced, there is no incentive to trade, but there may be incentives for the country to trade intermediate goods that are inputs to the production. They can be used as inputs in the final production and may be differently endowed. Assume that factors of trade cannot be traded (i.e. there is no capital mobility and no migration). Instead of just using domestic stocks of capital, land and labor, a country can trade these intermediate goods with the rest of the world. Thus, the country attains a different level of production of the final good. Let $\underline{p_N}, \underline{p_K}$ and $\underline{p_L}$ be the world prices of the capital, land, and labor intensive goods. The budget constraint is then:

$$\underline{p_K}Y_K + \underline{p_L}Y_L + \underline{p_N}Y_N = \underline{p_K}K + \underline{p_L}L + \underline{p_N}(1-\delta)$$

The left side of the equation is the total expenditure of intermediate good at world prices, and the right side is the total revenue that this country raises by selling its production of intermediate goods at world prices.

The model assumes that the world prices are determined in some world market equilibrium, such that:

$$\underline{p_K} = \frac{\underline{p_L}}{\sigma} = \theta \Psi^{\theta-1} \quad \text{and} \quad \underline{p_N} = (1-\theta)\Psi^{\theta} \tag{13}$$

Where $\Psi$ is the ratio of the sum of capital and land relative to labor in the world economy. In (6), what mattered for the determination of the domestic price was the ratio of $K + \sigma L$ to $1 - \delta$, the ratio of the sum of capital and land relative to labor. In the world economy, the same ratio matter, but the difference is that the world stock of factors is important, not just the stocks in one country. For example, if all the countries of the world traded with no tariffs or trading frictions:

$$\frac{\underline{p_K}}{\underline{p_N}} = \frac{\theta}{1-\theta}\frac{\sum_j N_j}{\sum_j K_j + \sigma \sum_j L_j}$$

Where $N_j$ is the total labor stock in country j, $K_j$ is the capital stock and $L_j$ is the stock of land. Then $\Psi$ is equal to the sum of the capital-and-land-to-labor ratio across the world:

$$\Psi = \frac{\sum_j N_j}{\sum_j K_j + \sigma \sum_j L_j}$$

If there are tariffs or trading frictions, $\Psi$ will differ from this ratio. The model focuses on the emergence and consolidation of democracy in non democratic societies, and non democratic societies are typically poorer and, therefore, more abundant in labor than capital. It is therefore natural to think that the country in question is relatively scarce in capital, which gives the assumption:

$$\Psi > \frac{K + \sigma L}{1 - \delta} \tag{14}$$

The most important implication of this assumption is shown by comparing (6) and (13), which shows that after trade opening, the price of labor intensive intermediate goods increases in the specific country in question (which is presumed to be a relatively labor abundant country). This country is relatively abundant in labor

compared to the rest of the world economy, which lowers the price of the labor in-
tensive intermediate good when there is no international trade. International trade
increases the price of the labor intensive good to the world level. When these prices
are given, factor rewards in this economy are again given by the relevant value of
the marginal products, now:

$$\underline{w} = \underline{p_N}, \ \underline{r} = \underline{p_K} \quad \text{and} \quad \underline{v} = \underline{p_K} \tag{15}$$

This implies that international trade also increases wages relative to capital and land
returns. These changes in the relative factor prices are the main channel by which
international trade has an impact on whether democracy emerges or consolidates.

After trade opening, the price of the abundant factor increases relative to other
factor prices. If the country in question is a relative poor, the abundant factor is
labor. This implies that the change in factor prices make inequality decline after
trade. Combining (13) and (15), the post-trade factor prices are given by:

$$\underline{w} = (1 - \theta)\Psi^\theta$$
$$\underline{r} = \theta\Psi^{\theta-1} \tag{16}$$
$$\underline{v} = \sigma\theta\Psi^{\theta-1}$$

Equation (14) implies that wages are higher and the returns to land and capital are
lower than under autarky. Using these factor prices, post-trade incomes are:

$$\underline{y^p} = (1 - \theta)\Psi^\theta$$

and

$$\underline{y^r} = \frac{\theta}{\delta}\Psi^{\theta-1}(K + \sigma L)$$

and the average income is:

$$\bar{y} = \Psi^{\theta-1}((1 - \theta)(1 - \delta)\Psi + \theta(K + \sigma L)) \tag{17}$$

This leads to the most preferred tax rate of the citizens as $\underline{\tau^p}$, which satisfies the
first order conditions:

$$\frac{\underline{y}^p}{\bar{y}} = 1 - C'(\underline{\tau}^p) \quad \text{or}$$

$$\frac{1-\theta}{(1-\theta)(1-\delta) + \theta\frac{K+\sigma L}{\Psi}} = 1 - C'(\underline{\tau}^p)$$

(18)

By (14), $\frac{K+\sigma L}{\Psi} < 1 - \delta$, therefore

$$\frac{1-\theta}{(1-\theta)(1-\delta) + \theta\frac{K+\sigma L}{\Psi}} > \frac{1-\theta}{1-\delta}$$

and

$$\underline{\tau}^p < \tau^p$$

Where $\underline{\tau}^p$ is the preferred tax rate for the citizens after trade, and $\tau^p$ is the preferred tax rate before trade. The citizens, who's income comes from supplying labor, prefer to set lower taxes after trade. This implies that after globalization, democracy becomes less redistributive because globalization reduces income equality.

## 4.2   Implications of land abundance

The analysis is based on the assumption that (14) holds, which implies that the country in question is abundant in labor and scarce in capital. This seems like a reasonable assumption for many non democratic countries that are joining the world economy, but there are exceptions such as Chile and Argentina. Their most abundant factor in the beginning of the last century was land. In this case, the opposite of (14) holds since international trade increases the relative income of the elite, and thus making democratization and democratic consolidation less likely. The importance here is that the implications of international trade for the political equilibrium depend on its implications for factor prices.

## 4.3   Hypothesis

Acemoglu and Robinson (2006) state that if a country is abundant in labor and scarce in capital and land, it will have lower wages and higher returns to capital than on average in the world. International trade therefore increases wages and reduces interest rates in such a country. As a result, the income gaps in these countries are reduced. The reduction in the gap between the poor and the rich

implies a reduced political conflict. With the smaller gap, the poor have less reasons to vote for highly redistributive policies, and democracy is less of a threat to the rich. Therefore, international trade reduces the conflict between the rich and the poor. With a reduced conflict the rich are less willing to incur the cost of a coup to revert back to non democracy, and democracy is more likely to consolidate. Globalization has the opposite effect on factor prices in rich countries. With a globalization process wages should fall and returns to capital should increase in the capital rich countries, such as the OECD countries. However, Acemoglu and Robinson (2006) do not think that the likelihood of coups is high in the OECD countries, since the democracies are already fully consolidated and the marginal increase in the democratic redistribution will not push them over the edge.

As a result, I test the hypothesis that not only that openness leads to greater democracy in labor abundant countries, but also that openness leads to less democracy in a capital and land abundant countries.

$H_1$ : Globalization leads to increased democracy in labor abundant countries

$H_2$ : Globalization leads to less democracy in capital and land abundant countries.

# 5   Data

The estimation uses polity data and openness data from 1960 to 2007 in all the 217 countries in the sample. To test the hypotheses, the paper uses a large-n cross-section time-series analysis. The data is collected from different sources and combined to form a panel data set.

## 5.1   Variables

The dependent variable is *Democracy.* The data for this index is from Polity IV data set (Marshall, Gurr, & Jaggers, 2008). This index is used by other authors e.g. López-Córdova and Meissner (2005) and Doces and Magee (2015). The polity score summarizes different indicators of political authority patterns in order to measure three aspects of a country's political system (López-Córdova & Meissner, 2005):

1. *Competitiveness and openness in the process of executive recruitment.*
   Whether the head of government is selected through ancestral succession, or by the ruling political elites decisions, or through free and fair elections among

candidates from different political parties where the citizens has the opportunity to become the executive.

2. *Constraints on the chief executive.*
   Whether the head of the government must take into consideration other point of views during the decision making process, for example the legislature.

3. *Competitiveness and regulation of political participation.*
   Measuring the degree to which political competition is institutionalized and free from governmental control.

A weighted sum of the component variables is taken, and two new variables, one measuring democracy and the other autocracy, are created. The democracy variable takes a value between 0 and 10, and the autocracy variable takes a value between 0 and -10. The polity score combines these two variables, and ranges from -10 to 10.[2]

The reason for the use of the Polity IV measure is that it appears to be the best available data set that measures deep political and institutional change over a long period of time and for a large number of nations. There are other measures of democracy, such as the Freedom House index, but the Polity IV measure is more often used by scholars, and goes further back than the Freedom House index, which only goes back to 1973.

As López-Córdova and Meissner (2005) state: Any attempt to measure democracy is fraught with conceptual and mechanical problems or inaccuracies. They also argue that the Polity IV data set may not capture all aspects of a democratic system; only show the competition on the recruitment of the executive, more open participation in choosing the executive and more checks and balances against an executive.

The main explanatory variable is *Openness*. The data is collected from Penn World Tables 6.3, and is also used by Doces and Magee (2015). The variable is measured as exports plus imports divided by GDP, which is a well known measurement of a country's openness.

To test the model by Acemoglu and Robinson (2006), the regression needs to have a variable that shows capital and land per labor. The variable *capital-labor* is from Penn World Tables 8.0, and is also used by Doces and Magee (2015). The variable is measured as capital stock value (in 2005 $)/labor force. The variable *landprlabor* is collected from the World Bank's World Development Indicator data set. The data

---

[2]Countries are only includes in the Polity IV data set if they achieved independence by 1998 and had a population of 500 000 or more in 1998.

for labor force only runs from 1990, and not from 1960, and therefore it is replaced by a variable which runs through the whole period (1960-2007). This variable is *arable*, which is arable land in hectars per person in each country. This variable is also collected from the World Bank's World Development Indicator data set, and is from The Food and Agriculture Organization (FAO), which is under the UN. The variable is defined as land under temporary crops, temporary meadows for mowing or for pasture, land under market or kitchen gardens, and land which is temporarily fallow. Land abandoned as a result of shifting cultivation is excluded. The reason for choosing arable land, and not agricultural land, is that agricultural land can include areas not suitable for agriculture, such as forests, mountains and inland water bodies (The World Bank, 2017).
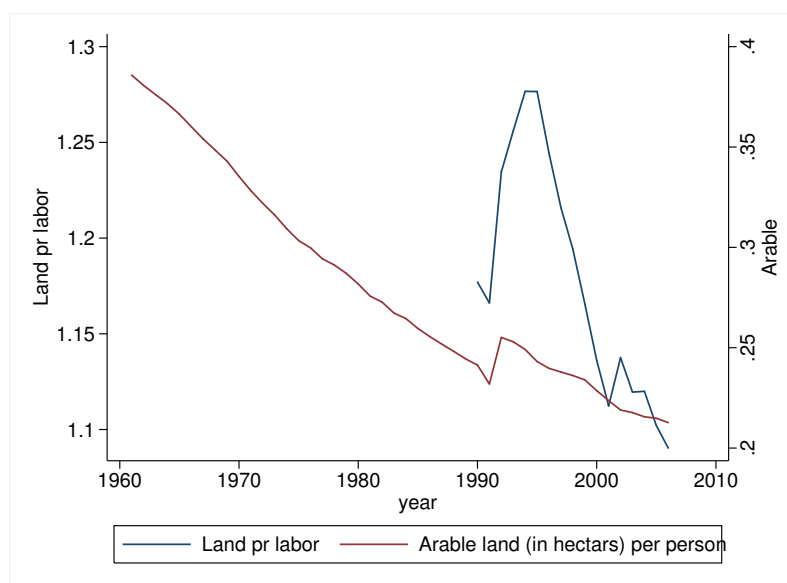


Figure 1: Means of variables *landprlabor* and *arable*

The correlation between *landprlabor* and *arable* is theoretically high, since there is a correlation between number of people and laborers in a country. Figure 6.1 shows the development of the two variables, *arable* starting from 1960, and *landprlabor* starting from 1990. The graph lines show the same conjuncture from 1990, with an increase in 1992-1995, and then a steady decrease. The line of *landprlabor* is higher than the line of *arable* due to the fact that the number of laborer in a country is always less than the number of people, thus resulting in a higher ratio. The correlation coefficient is 0.8474, which is high (Wooldridge, 2013).

The model by Acemoglu and Robinson (2006) assumes that only the rich own capital and land, and to differentiate between the two groups, two dummy variables *above median capital per laborer* and *above median arable land per person*, are created.

These two dummy variables are equal to 1 if the observation is above median in either capital or land, and equal to 0 if the observation is below median. These to variables are combined and form the variable *above*. This variable is equal to 1 if the observation is above median in both capital and land. Another variable is created, called *below*. This variable is equal to 1 when both *above median capital per laborer* and *above median arable per person* are equal to 0. This variable shows labor abundant countries. These two variables separate the difference between the capital and land abundant countries and the labor abundant countries.

Other control variables are *female*, *age 15-65* and *age < 15*. They are all from the World Bank Development Indicator data set and show the percentage of the population that is female, between the age of 15-64 and under the age of 15. The variables *ethnic fractionalization*, *religious fractionalization* and *linguistic fractionalization* are from Ellingsen (2000). They are available until 1994, but since the characteristics of a countries' population are relatively stable over time, the 1994 values are used in each country for more recent years.

Many of the countries have governments that are affected by the fact that the country have been previously colonized, thus the regression include a dummy variable *colony*, which is equal to 1 if the country has been a colony after 1945, and equal to 0 if not. The variable *1/years since independent* is measuring whether the country has been decolonized. Countries that have never been colonized have a value of 0, while countries that still are colonies are equal to 1. The variables are both from Head, Mayer, and Ries (2010).

## 5.2  Descriptive statistics

Table 1 shows the descriptive statistics of the central variables in the analysis. The dependent variable *Democracy* has a mean of 0.13, and a standard deviation of 7.50. In this case, since the variable lies in the interval $[-10, 10]$, the standard deviation is quite high. A high standard deviation indicates that the data points are spread out over a wider range of values, and there is variation on the data set. The main explanatory variable *Openness* has a mean of 74, and a standard deviation of 51.57. Both main variables have a positive development from 1960 to 2007.

Table 1: Descriptive statistics

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Democracy | 6682 | .1302504 | 7.502788 | -10 | 10 |
| Openness | 8257 | 73.99936 | 51.5669 | 0 | 622.6263 |
| Capital-labor ratio | 6952 | 44793.55 | 55703.16 | 499.6375 | 581945.5 |
| Arable-population ratio | 8542 | .2772317 | .3398747 | .0001458 | 3.498432 |
| Female | 9090 | 50.13291 | 2.35323 | 25.14558 | 56.12519 |
| Age 15 - 64 | 9090 | 58.06943 | 6.615945 | 44.85925 | 82.34296 |
| Age < 15 | 9090 | 36.04176 | 9.801915 | 13.18091 | 51.91537 |
| Linguistic fractionalization | 6711 | 2.178836 | 1.52554 | 1 | 8 |
| Ethnic fractionalization | 7165 | 2.18067 | 1.124372 | 1 | 8 |
| Religious fractionalization | 6710 | 1.857675 | .8139136 | 1 | 4.5 |
| 1/Years since independent | 9501 | .1857871 | .3464751 | 0 | 1 |
| Colony after 1945 | 9597 | .7602376 | .4269607 | 0 | 1 |
| Above | 6952 | .2685558 | .4432402 | 0 | 1 |
| Below | 10421 | .1208137 | .03259263 | 0 | 1 |
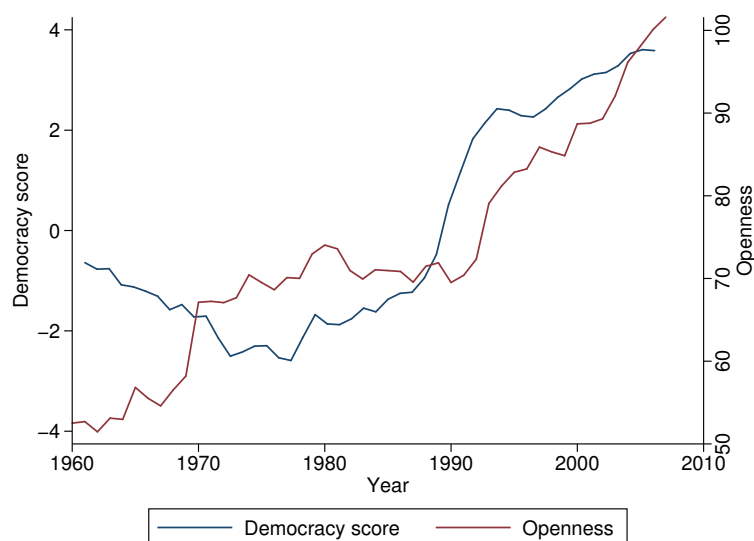


Figure 2: The development of the variables *Democracy* and *Openness* from 1960 to 2007

Figure 2 shows the development of the two main variables from 1960 to 2007. *Democracy* has a negative development until 1977, where it reaches bottom at an average of -2.59. From this point the variable has a positive development. The explanation for this is the fall of the Berlin wall, the end of the Soviet Union, and the democratic

wave in Latin America starting in 1978. *Globalization* has a positive development
from 1960 to 2007, the only exception is a few smaller downturns.
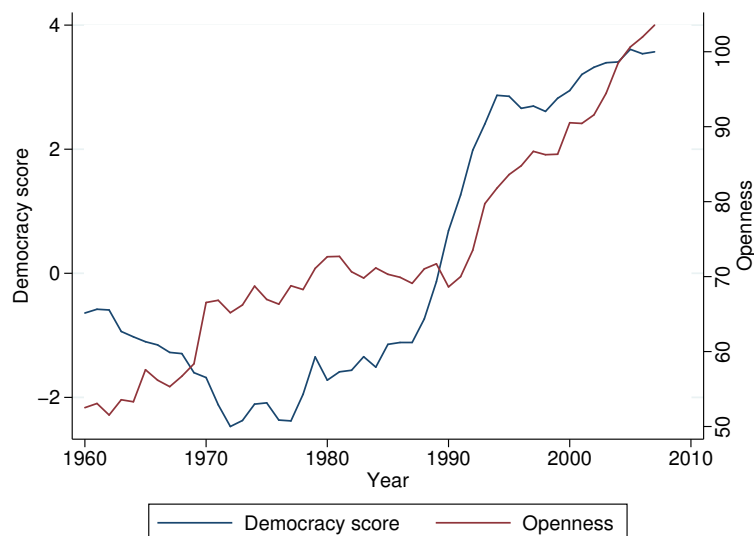


Figure 3: The development of the variables *Democracy* and *Openness* from 1960 to
2007 in labor abundant countries

Figure 3 shows the development of the two main variables from 1960 to 2007, but
only in the labor abundant countries. The main difference between figure 2 and 3
is that the democracy line has a larger difference between the two lines in the years
1970 to 1990, the globalization line is higher than the average, and the democracy
line is lower than the average. Another difference is that the democracy line has a
steeper line from 1987 to 1993.

## 5.3   The advantages of panel data

Panel data is considered to be an efficient analytic method in handling econometric
data. Panel data analysis is often preferred because it allows the inclusion of data
for N cross sections (e.g. countries) and T time periods (e.g. years). The combined
panel data matrix set consists of a time series for each cross sectional member in
the data set, and offers a variety of estimation methods.

A data set that only consists of observations of N individuals at the same point in
time is a cross section data set. Some cross section data sets also exist over time,
thus the number of cross sectional samples are taken at different points in time.
However, these data sets do not constitute a panel data set, because it generally
does not follow the same individual over a period of time. A true panel data set

follows each individual in the panel over a number of periods. In this case, the data set is a panel and follows 217 countries from 1960 to 2007.

If the panel has the same number of time observations for every variable and every individual, it is a balanced panel. In this case, it is an unbalanced panel, since the number of observations are not the same for every variable or every individual. When the variable does not have any information about the observation in that time period, the observation is reported as missing. When the regression is executed by the software program, in this case Stata, the missing values are excluded from the regression.

The basic idea behind panel data analysis is that the individual relationships will all have the same parameters. This is known as the pooling assumption, and comes from the fact that the panel analysis is pooling all the individuals together into one data set and imposing a common set of parameters across them. If the assumption about pooling is correct, then panel data estimation can offer some considerable advantages:

(a) The sample size can be increased considerably by using a panel and hence much better estimates can be obtained.

(b) Under certain circumstances the problem of omitted variables, which might cause biased estimates in a single individual regression, may not occur in panel context.

The disadvantages of panel data can occur when the pooling assumption does not hold, and the individual relationships do not all have the same parameters. Although in this case, which is referred to as heterogeneous panel (because the parameters are different across individuals), the panel data estimation is expected to give some representative average estimate of the individual parameters.

Another problem with panel data is the lack of time series which extend more than 50 years. An efficient solution is to pool the data into a panel of time series from different cross sectional units. This pooling of the data generates differences among the different cross sectional or time series observations that can be captured with the inclusion of dummy variables. The panel data set will in this case include variables that extend 47 years. However, all the variables are from different time series from different cross sectional units, pooled together into one data set.

# 6  Econometric specification

The panel data set is formulated by a sample of 217 cross-sectional units (countries) that are observed at 47 time periods (years). The estimation of the model is executed by different estimation methods, such as the pooled OLS method, the fixed effects method and the two stages least squared method. The linear model is given by:

$$Polity_{it} = \beta_0 + \alpha_i + \beta_1 Openness_{it} + X'\eta + \varepsilon_{it} \tag{19}$$

where the variables *polity* and *openness* have both $i$ and $t$ subscriptions for $i = 1, 2...217$ countries and $t = 1, 2...47$ time periods. In the simpler version $\alpha_i$ is equal to $\alpha$, but to show the heterogeneity in the panel the assumption that $\alpha_i$ is identical for all countries is relaxed. This is to show the difference in the sample and to include the expected differences in their behavior. $\alpha_i$ can now differ for each country in the sample. Here, $\alpha_i$ can for example show the difference between the labor abundant countries and the capital and land abundant countries.

## 6.1  The pooled OLS method

For the pooled ordinary least square (OLS) method to be efficient and unbiased there are several assumptions that has to be fulfilled:

1.  The population model is linear in its parameters. The model in the population can be written as:

    $$Polity_{it} = \beta_0 + \alpha_i + \beta_1 Openness_{it} + X'\eta + \varepsilon_{it}$$

    where $\alpha_i, \beta_0, \beta_1$ and $\eta$ are unknown parameters (constants) of interest and $\varepsilon_{it}$ is an unobserved random error or disturbance term.

2.  There is a random sample of $n$ observations following the populations model in assumption 1. This assures that $cov(\varepsilon_i, \varepsilon_j | X) = 0$, where X is all the explanatory variables, and $i \neq j$ which means that the error term is independent.

3.  No perfect collinearity, which means that none of the explanatory variables can be written as a perfect linear combination of the other explanatory variables.

4.  Zero conditional mean, which means that the conditional assumption for the error term, given explanatory variable must be equal to 0. $E(\varepsilon_i | X) = 0$, which implies that $E(\varepsilon_i) = 0$, and thus $cov(\varepsilon_i | X) = 0$.

5. The error term has the same variance given any values of the explanatory variables. In other words, $var(\varepsilon_i|X) = \sigma^2$.

Under the assumption 1 through 4,

$$E(\hat{\beta}_j) = \beta_j$$

which means that any value of the expected population parameter $\hat{\beta}_j$ is $\beta_j$. In other words, the OLS estimators are unbiased estimators of the population parameters. The same applies to the estimation of the variance. Under the assumption 1-5,

$$E(\hat{\sigma^2}) = \sigma^2$$

the estimation of the variance is unbiased. The Gauss-Markov theorem says that under the assumptions 1 through 5, the estimated coefficients are the best linear unbiased estimators (BLUE) of the coefficients. If one or more of the assumptions are not met, the estimation is no longer unbiased or efficient. Since assumption 4 states that $E(\varepsilon_i|X) = 0$, which also implies that $E(\varepsilon_i|\alpha) = 0$, this means that there can not be any correlation between the error term and any of the explanatory variables or the unobserved heterogeneity term.

The pooled OLS method of estimation presents results under the assumption that there are no differences among the data matrices of the cross-sectional dimension, $i$. In other words, $\alpha_i = \alpha$. Practically, the pooled OLS method implies that there are no differences between the estimated cross-sections. This is useful under the hypothesis that the data set is a priori homogeneous (e.g. the sample are only capital and land abundant, or only labor abundant) (Asteriou & Hall, 2007). This means that OLS can be biased and inefficient if there are a priori differences in the panel set when applying the OLS method. However, if the data set is divided into different groups that are homogeneous, the OLS method is unbiased and efficient.

## 6.2   The fixed effects method

Another possible way of estimating the effect of globalization on democracy can be by the fixed effects (FE) method. This method allows $\alpha_i$ to be arbitrary correlated with the other explanatory variables and not change over time. This means that there are differences between the countries in the sample that could effect the explanatory variables. The way fixed effects works is by transforming the model so that it eliminates the unobserved effect $\alpha_i$. The fixed effect transformation or within transformation is obtained in two steps: First, averaging the equation over time for each $i$ to get the cross section equation:

$$\overline{Polity_i} = \beta_0 + \alpha_i + \beta_1 \overline{Openness_i} + \overline{X'}\eta + \overline{\varepsilon_i} \tag{20}$$

Then equation (2) is subtracted from equation (1), leading to:

$$Polity_{it} - \overline{Polity_i} = (Openness_{it} - \overline{Openness_i})\beta_1 + (X' - \overline{X'})\eta + \varepsilon_{it} - \overline{\varepsilon_i} \tag{21}$$

$$\overline{\overline{Polity_{it}}} = \overline{\overline{Openness_{it}}}\beta_1 + \overline{\overline{X'}}\eta + \overline{\overline{\varepsilon_{it}}} \tag{22}$$

Where $\overline{\overline{Polity_{it}}}$ is the time-demeaned data, and the unobserved effect $\alpha_i$ has been cancelled out. The fixed effects estimator is the pooled OLS estimator in the regression with the time-demeaned variables in (22). Notice that time constant explanatory variables (i.e. variables that are time invariant for all countries, like *colony after 1945* and *region*) are swept away by the fixed effects transformation. This is because the method wants to allow for arbitrary correlation between $\alpha_i$ and the explanatory variables, and there is no way to distinguish between the effect of a time constant observable variable and the effect of the time constant unobservable variable $\alpha_i$.

The fixed effects estimation method has different assumptions than the pooled OLS method. Each explanatory variable has to change over time (at least some of the countries), strict exogenity of the explanatory variables conditionally on $\alpha_i$: $E(\varepsilon_i|X, \alpha_i) = 0$, constant variance of errors conditional on the explanatory variables and $\alpha_i$: $var(\varepsilon_i|X, \alpha_i) = \sigma_\varepsilon^2$, and no serial correlation among errors conditional on the explanatory variables and $\alpha_i$: $cov = (\varepsilon_i\varepsilon_j|X, \alpha_i) = 0$, for $i \neq j$ has to be fulfilled to make the method efficient and unbiased.

The main difference between pooled OLS and fixed effects is that fixed effects tolerates that $\alpha_i$ exists, because it is transformed away, where as pooled OLS does not allow for $\alpha_i$ to exist. One downside of fixed effects is that it loses a lot of variation by disregarding the time variation. The pooled OLS method exploits both the time and country variation. However, the downside with pooled OLS is that the estimation method does not take into account that there should be one regression for each country over time, a restriction called parameter similarity. Both methods are estimated.

## 6.3   The endogeneity problem

An econometric problem arises: Democracy and globalization affect each other simultaneously. Simultaneity means that at least one of the explanatory variables in a multiple linear regression model is determined jointly with the dependent variable. This could lead to a simultaneity bias, the bias that arises from using OLS to estimate an equation in a simultaneous equation (Wooldridge, 2013). This leads to a biased and inefficient estimates. This means that the if the level of globalization in a country affects the level of democracy, and the level of democratization in a country affects the level of globalization, then the two variables are simultaneously determined. The simultaneity can be identified when looking at the opening of Central and Eastern Europe following the demise of the Soviet Union, or the wave of democratization in Latin America from 1978, which Leblang and Eichengreen (2006) conclude in favor of. The reason for this simultaneity is, as argued by Acemoglu and Robinson, that an increase in globalization leads to higher wages for the poor, and thus a higher probability for democracy. When a country becomes more democratic, the country will also open up its markets for international trade. The transparency in the political institutions will spill over into the economic institutions (Giavazzi & Tabellini, 2005).

When there is simultaneity in the model, assumption 4 of the pooled OLS method is not fulfilled; $E(\varepsilon_i|X) \neq 0$. This is because an explanatory variable, *globalization*, is partly determined by the dependent variable, *democracy*. This simultaneity leads to an endogeneity in the equation. Other sources for this endogeneity could be omitted variables or measurement errors. The instrumental variable method provides a general solution to the problem of endogenous regressor.

## 6.4   The two stages least squared method

The idea behind two stages least squared (2SLS) method is that the endogenous variable is estimated in a secondary regression, with instruments that have a high correlation with the endogenous regressor and are exogenous in the main equation (the instrument can not directly determine the dependent variable) (Wooldridge, 2013). The estimations from the secondary regression are then stored and used as a replacement for the endogenous variable in the main regression. This leads to unbiased and efficient estimates when the assumptions for the instruments are fulfilled. The assumptions are not that different from the OLS assumptions. The main difference is that the conditions for the instruments hold.

### 6.4.1   The instruments

An instrumental variable has to be observable, not already included in the other explanatory variables, $X'\eta$, and fulfill two additional assumptions: Firstly, it has to correlate with the endogenous explanatory variable in question - otherwise it has no power. This is referred to as the instrumental relevance. Secondly, a valid instrument must also satisfy the exclusion restriction, which means that it can not have any independent effect on the dependent variable. One must be able to assume that it only affects the dependent variable via the endogenous explanatory variable that it is being used to instrument. The two possible instruments are *gravity* and *regional openness*. Both are used by Doces and Magee (2015). The first instrument is *gravity*, and is also used by López-Córdova and Meissner (2005). The variable is the gravity model for predicted bilateral trade flows, as a share of GDP, between a country $i$ and its trading partner $j$.

$$ln\Big(\frac{exports_{ij} + imports_{ij}}{GDP_i}\Big) = \lambda G + u_{ij} \tag{23}$$

Where G is a vector of explanatory variables including the country's population, land area, distance, landlocked status, island, sharing a border, and sharing a common language. Country $i$'s predicted bilateral trade is then summed up over all of its trade partners to generate a predicted level of openness:

$$\widehat{Open} = \sum_{j \neq i} exp(G\hat{\lambda}) \tag{24}$$

Previous studies of the impact of globalization on democracy have used the gravity model to identify the exogenous component of trade. One of these studies is by Eichengreen and Leblang (2008), and they highlight that this variable satisfies the criterion of instrumental relevance.

The variable also satisfies the exclusion restriction. Eichengreen and Leblang (2008) conduct a separate estimation with $\widehat{open}$ and *democracy*, and the results are that the variable is not significant in this estimation.[3] Therefore, they argue that the variable *open* satisfies the exclusion restriction.

The second instrument for *openness* is the average openness among the country's 10 nearest neighbors. Previous studies, like Doces and Magee (2015), use this variable

---

[3]I also run the same estimation, and the results is that the $\widehat{open}$ variable is not significant in the estimation.

as an instrument as a way to identify the exogenous component of globalization. Doces and Magee (2015) also test with the 20 nearest neighbors, but the results are similar. The variable is argued by Doces and Magee (2015) to both satisfy the criterion of instrumental relevance and exclusion. A regression with the variables *democracy* and *regionopenness*, leads to that the variable *regionopenness* being insignificant. The reduced form equation is then:

$$Openness_{it} = \beta_0 + \beta_1 \widehat{open_{it}} + \beta_1 regionalopenness_{it} + X'\eta + v_{it} \tag{25}$$

Using the results from the reduced form equation (25) into the structural equation:

$$Polity_{it} = \beta_0 + \alpha_i + \beta_1 \widehat{Openness_{it}} + X'\eta + \varepsilon_{it} \tag{26}$$

When using multiple instruments, there are two additional conditions that have to be fulfilled: The rank condition and the order condition. The rank condition has already been mentioned, and says that at least one of the coefficients from the reduced form equation (25) is different from zero. The order condition is that there has to be at least as many instruments as explanatory variables that are endogenous (Woolridge, 2013). These two conditions are fulfilled.

## 6.5 Heteroskedasticity

Another well known econometric problem that could arise is heteroskedasticity. Heteroskedasticity is when the variance of the error term is not the same given different values of the explanatory variable, therefore assumption 5 of the pooled OLS method does not hold; $var(\varepsilon_i|X) = \sigma_i^2$. The Breusch-Pagan test can be used to determine whether there is any heteroskedasticity in the model. If there is heteroskedasticity in the model, this can be compensated by using cluster robust standard errors at country level. In panel data sets heteroskedasticity is common, so to ensure that the heteroskedasticity is taken care of the estimations in the next chapter is done by cluster robust standard errors (Wooldridge, 2013)[4].

# 7 Results

In this chapter, I will look at the results from the different regression methods. Firstly, looking at the sample as a whole, as Doces and Magee (2015) did in their

---

[4]Further information on clustered standard errors in Wooldridge (2013) page 687-688.

article. Secondly, extending the research on this topic to look at labor abundant countries and capital and land abundant countries apart to test the model by Acemoglu and Robinson (2006). The difference between this paper and previous research is that this paper divides more clearly between labor abundant and capital and land abundant countries. The results will, overall, lend support to the model by Acemoglu and Robinson (2006).

## 7.1   Estimation with pooled OLS

The first model is estimated with the pooled OLS method, and the results from the regression are shown in table 2. The estimation of model (1) is performed by using equation (19), and the use of polity data and openness data from 1960 to 2007 in all the countries in the sample. All standard errors in the estimation are robust against arbitrary heteroskedasticity and serial correlation at the country level, i.e. they are clustered at the country level (Wooldridge, 2013).

The variable *openness* is negative and significant at the 10 % level. This shows that an increase in openness or globalization in a country, goes along with a lower level of democracy. This contradicts López-Córdova and Meissner (2008), which shows that an increase in openness leads to higher levels of democracy. The reason for the results could be unobserved heterogeneity and that most of the countries in the data set have high levels of capital and/or land, which could lead to more protection, as the model of Acemoglu and Robinson (2006) argues.

The variable of *percent of population under the age of 15* has a negative and significant effect, which indicate that the younger the population is, the less democratic is the country. Countries with a low democratic score often have a high population growth, thus have a higher number of inhabitants under the age of 15. While countries with a high democratic score often have a low population growth, leading to less inhabitants under the age of 15 (Mutascu, 2009). In the result from the fixed effects method, model (2), the country specific variation is removed, and the variable has a smaller impact on the predicted level of democracy and is not significant.

The variable *regional polity* has a positive and significant effect on democracy, which indicates that if the surrounding countries experience an increase in the level of democracy, it will lead to a higher predicted level of democracy in the specific country. This result could come from the definition of regions, and the fact that countries that are located in Europe tend to have a higher level of democracy (Brinks & Coppedge, 2006). A rise in the neighbouring countries average democracy score

with one unit, leads to a rise of about 0.7 units in the countries' own polity score.

The variable *colony* is positive and significant at the 10 % level. The variable is equal to 1 if the country has been a colony after 1945, and equal to 0 if it has not been a colony after 1945. If the country in question has been a colony after 1945, their predicted level of democracy is 2.033 higher than if the country never has been a colony.

The variable *1/Years since independence* has a positive and significant effect on democracy. The variable is equal to 0 for countries that have not been a colony, and equal to 1 if the country is still a colony in the current year. If the country was a colony, and has become independent, it is divided on the number of years since it became independent. If there still were any colonies, their predicted polity score would have been 4.987 units higher than a non-colony. The effect decreases during the years after independence, thus the effect is highest when the country is newly independent.

The constant is 17.93. As known, the variable *democracy* is only in the interval [-10,10]. This means that the constant is relatively high, especially since the average value of *democracy* is around zero at 0.13. The R-squared is the proportion of the total sample variation in the dependent variable that is explained by the independent variable (Woolridge, 2013). In the pooled OLS estimations the R-squared is equal to 0.541, which means that 54,1 % of the variation of the variable *democracy* is explained by the explanatory variables.

## 7.2   Estimation with Fixed Effects

The second model is estimated with the fixed effects method and shown in table 2. The estimation is performed by estimating equation (22) by the pooled OLS method (Wooldridge, 2013). With a fixed effect estimation method, variables that do not vary over time are omitted. The variables who are omitted are *colony after 1945* and the regions (*Africa, Europe etc.*) The variable *openness* is negative and not significant. The coefficient is small, indicating that the impact on the predicted level of democracy is small.

The variable *linguistic diversity* is positive and significant at the 1 % level. *Linguistic diversity* is number of language groups compromising at least 5 % of the population, and the results implies that an increase of one language group will lead to an increase of almost 3 in the democratic score. Countries with a high average linguistic diversity

Table 2: Estimates from the pooled OLS-, fixed effects- and 2SLS estimation.

| Variables | (1) OLS | (2) FE | (3) 2SLS |
|---|---|---|---|
| Openness | -0.0120* | -0.000634 | -0.00855* |
| Female percent of population | 0.215 | -0.0305 | 0.230*** |
| Age 15 - 64 | -0.264 | -0.0134 | -0.270*** |
| Age<15 | -0.437*** | -0.0154 | -0.442*** |
| Ethnic diversity | -0.0374 | 0.917 | -0.0286 |
| Religious diversity | 0.218 | -1.071 | 0.167 |
| Linguistic diversity | 0.469 | 2.968*** | 0.465*** |
| Regional polity | 0.734*** | 0.942*** | 0.728*** |
| Colony after 1945 | 2.033* | | 1.981*** |
| 1/Years since independence | 4.987*** | 5.050*** | 4.853*** |
| Americas | 1.559 | | 1.507*** |
| Europe | -1.840 | | -1.846*** |
| Africa | -0.549 | | -0.552** |
| Pacific | 2.826 | | 2.824*** |
| Time | 0.0225 | 0.0279 | 0.0198** |
| Constant | 17.93 | -4.570 | 17.69*** |
| Observations | 4,824 | 4,824 | 4,814 |
| R-squared | 0.541 | 0.449 | 0.539 |

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

score are Asian countries such as the Philippines (5.15) and India (6.10), and African countries such as Mozambique (6.03), Mali (8), Kenya (8), Angola (5) and Cote d'Ivoire (5). A reason for these results could be that the regional effect that is omitted by the fixed effects method is picked up by this variable, thus leading to a much higher estimate than with the other estimation methods.

Some of the effect of the omitted variables can also be picked up by the variable *regional polity*. This variable is positive and significant at the 1 % level. The variable is the average democracy score of the countries' ten nearest neighbors, and if this score increases with one unit, the country in question will have an increase of almost one (0.942) in their predicted level of democracy.

The variable *1/years since independence* is positive and significant at the 1 % level. The same logic applies to this interpretation as in pooled OLS. The variable is equal to 0 if the country never was a colony, and is equal to 1 if the country is still a

colony. If the country was a colony, and has become independent, it is divided on the number of years since it became independent. If there still were any colonies, their predicted polity score would have been 5.050 units higher than a non-colony. The effect decreases during the years after independence, thus the effect is highest when the country is newly independent.

The constant is negative and smaller compared to the other estimation methods. The R-squared is smaller than with the other estimation methods. Which could come from the omitting of variables such as *region* and *colony after 1945*.

## 7.3 Estimation with two stage least squares

The third model is estimated with two stage least square method (2SLS) and uses equation (25) as the reduced form equation and equation (26) as the structural equation. The results are shown in table 2. The variable *openness* is negative and significant at the 10 % level. This is a different result than Doces and Magee (2015) who performed a similar regression which resulted in a negative, but insignificant coefficient.

The effect of the variable *female percent of the population* in the population has a positive effect on the predicted level of democracy and is significant at the 1 % level. If the percent of females in the population increase with one percentage point, the increase in the predicted democratic score is 0.23. The variable of *age 15-64* is negative and significant at the 1 % level. If the variable increases with one percentage point, the decrease in the predicted democratic score is 0.27. Hence, if the average age of the population is increasing, the democratic index goes down. It is the same with the variable *age < 15*. The variable is negative and significant at the 1 % level. If the variable increases with one percentage point, the predicted level of democracy goes down with 0.44 units.

The variable *linguistic diversity* is positive and significant at the 1 % level. The results indicate that if there is an increase in the linguistic diversity in a country with one unit, this will lead to an increase in the predicted level of democratic by 0.46 units. The variable *regional polity* is positive and significant at the 1 % level, and has the same effect in the two stage least square estimates as the pooled OLS estimates. If there is an increase in the regional average polity score with one unit, this will lead to an increase in the specific country's predicted democracy score of 0.72 unit. The variable *colony after 1945* is equal to 1 if the country was a colony after 1945. The variable is positive and significant at the 1 % level. If the country was a

colony after 1945 it has a 1.98 units higher predicted democratic score, than if it was independent. The variable *1/years since independence* is positive and significant at the 1 % level, and has the same interpretation as the pooled OLS method results.

The variables *Americas, Europe, Africa, Pacific* are equal to 1 if the country is in that specific region, and equal to 0 if the country is in Asia. All the variables are highly significant. If the country is in the Americas, the results indicate that it has a positive effect on the democratic index, the same are shown for countries in the pacific. If the country is in Europe or Africa, the results indicate that it has a negative effect on the predicted democratic score. The reason that there is such a negative effect on democracy if the country is in Europe could be due to the Soviet Union. The variable *time* counts from 0 to 47, where 0 is 1960 and 47 is 2007. It has a positive and significant effect on the predicted level of democracy. For every year that passes, the democratic index is positively affected.

The constant is positive and highly significant. It is also similar to the constant from the pooled OLS estimation method. The R-squared in the 2SLS estimation is high and shows the explanatory power of the regression. In the 2SLS regression the R-square is 0.539, which show that the model explained 53,9 % of the variation in the dependent variable. Hill, Griffiths, and Lim (2011) discourage the use of measures like R-squared when referring to methods such as 2SLS. They argue that when there are endogenous variables on the right-side of a regression equation, the concept of measuring how well the variation in *democracy* is explained by *openness* breaks down because these kinds of models exhibits feedback (Hill et al., 2011).

## 7.4   Tests of endogeneity and instruments

**Endogeneity in the model - The Hausman test**

If there is an endogeneity problem in the model (1), only the 2SLS method is consistent, although if the variable *openness* is exogenous both approaches are consistent, but the pooled OLS method is more efficient. To check if openness is endogenous this thesis uses the Hausman test (Hausman, 1978):

$H_0$ :   Openness is exogenous, E(openness, $\varepsilon_{it}$ )=0

$H_1$ :   Openness is endogenous, E(openness, $\varepsilon_{it}$ )$\neq$0

The test is based on the difference between the 2SLS method and the pooled OLS method, which should be close to zero under $H_0$, while a statistically significant difference will lead to the conclusion that openness is endogenous. The result do not reject $H_0$ at the 10 % level in the whole sample of countries. However, this does not mean that openness is exogenous nor endogenous. This is the same results as Doces and Magee (2015), and therefore both OLS and 2SLS estimates are included in the tables.

**Weak identification - The Cragg-Donald F-test statistic**

When an instrumental variable exhibits only weak partial correlation with the endogenous regressors the instrument is said to be weak or poor. The consequences of this is that the asymptotic variance of the estimates are large, and much larger than the OLS asymptotic variance, this a source of inconsistency in the estimates, and the distribution of the estimates is poorly approximated by the normal distribution (even with a large N) (Wooldridge, 2010). The Cragg-Donald F-test statistics are able to test whether any relationship between the instruments and the endogenous variable is sufficiently strong for reliable econometric inference (Hill et al., 2011).

$$Cragg - Donald\ F = [(N - G - B)/L]/[r_B^2/(1 - r_B^2)]$$

N is the sample size, B is the number of right-hand-side endogenous variables, G is the number of exogenous variables included in the equation (including intercepts), L is the number of "external" instruments that are not included in the model, and $r_B$ is the minimum correlation. The null hypothesis is that the instruments are weak against the alternative that they are not. In this case, the null hypothesis is rejected at the 5 % level because the F-statistic is higher than the critical value (Stock & Yogo, 2005). The instruments are therefore not weak, but strong.

**Test for over-identifying restrictions**

A model is said to be over-identified if there is more instruments than needed to identify the structural equation. If there is too many instruments, some of the instruments can be discarded and the model will still achieve identification. To test this, Hausman (1978) proposed to build a test of the exogeneity of both the instruments based on the difference between $\tilde{\beta} - \tilde{\tilde{\beta}}$. The betas are the difference between two estimates, one with *gravity* as an instrument and one with *regionalopenness* as an instrument. If the conclusion is that the difference is significant, one of the two instruments or both are endogenous. If the differences is small, both instruments are valid and exogenous. Here the difference are not significant, therefore the instruments are valid and exogenous.

## 7.5  Comparison of the estimates

Comparing the estimates of pooled OLS-, fixed effects- and two stage least squares method can be done by looking at the predicted levels of an average country. This indicates the bias in the estimation methods. The predicted level of an average country is calculated by using the average level of the explanatory variable and the results from the estimation in table 2.

$$\widehat{Democracy}^{OLS} = 1.3712$$

$$\widehat{Democracy}^{FE} = 1.1876$$

$$\widehat{Democracy}^{2SLS} = 0.9666$$

An average country estimated with the pooled OLS method has a democratic score of 1.3712, while with the fixed effects method the democratic score is lower at 1.1876, and in the 2SLS method the score is even lower at 0.9666.

The difference between the predicted level of $\widehat{Democracy}^{OLS}$ and $\widehat{Democracy}^{FE}$ is 0.1836, thus the pooled OLS method have 0.1836 higher estimates than the fixed effects method. The difference between the predicted level of $\widehat{Democracy}^{OLS}$ and $\widehat{Democracy}^{2SLS}$ is 0.4045, thus the pooled OLS method have a higher estimate than the 2SLS method. The difference between $\widehat{Democracy}^{FE}$ and $\widehat{Democracy}^{2SLS}$ is at 0.2209, thus the fixed effects method has a higher estimate than the 2SLS method. If there is endogeneity in the model, the 2SLS method is unbiased and efficient, and thus the bias of the pooled OLS method is 0.4045, and the bias of the fixed effects method is 0.2209.

The comparison can also be done by looking at the predicted level of democracy one standard deviation below and above. This is done to look at the differences in the methods when the values of the explanatory variables are lower or higher. The predicted level of democracy one standard deviation below is:

$$\widehat{Democracy}^{OLS} = 0.1735$$

$$\widehat{Democracy}^{FE} = -9.9678$$

$$\widehat{Democracy}^{2SLS} = -0.2173$$

A country, one standard deviation below the average, estimated with the pooled OLS method has the predicted democratic score of 0.1735, while with the fixed effects method the predicted democratic score is -9.9678. When the interval of the

democracy index is [-10,10], this country is a full autocracy, for example Iran from 1960 until the revolution in 1978. With the 2SLS method the predicted democracy score is 0.2173, lower than the pooled OLS method and the fixed effects method. The difference between the pooled OLS method and the fixed effects method are 0.3909, almost the same as the average country. The predicted level of democracy one standard deviation below shows that countries with a lower level of openness have a lower level of democracy. The predicted level of democracy one standard deviation above is:

$$\widehat{Democracy}^{OLS} = 2.5689$$
$$\widehat{Democracy}^{FE} = 12.3431$$
$$\widehat{Democracy}^{2SLS} = 2.1506$$

A country, one standard deviation above the average, estimated with the pooled OLS method has the predicted democracy score of 2.5689, while with the fixed effects method the predicted democracy score is 12.3431. This is outside of the interval of the democracy variable, showing that a higher level of openness leads to a higher predicted level of democracy. This is true for the other methods as well. The 2SLS method predict a level of democracy at 2.15, close to the pooled OLS methods estimations.

## 7.6 Labor abundant countries

To test the model by Acemoglu and Robinson (2006) the sample of countries are divided into two groups; labor abundant countries and capital-and land abundant countries. As already mentioned, labor abundant countries are defined as countries with below median capital per laborer and below median arable land per person. Capital and land abundant countries are defined as countries with above median capital and above median arable land. This section focuses on the results from the labor abundant countries. The labor abundant countries in the sample only represent 16,6 % of the total sample. The variable *Europa* is omitted from the regression because there are no labor abundant countries in Europe.

### 7.6.1 Estimation with pooled OLS

The fourth model is estimated with the pooled OLS method. The results from the regression is shown in table 3.

The variable *openness* is positive, however, the variable is not significant. If the variable was significant this would indicate that an increase in the level of *openness* with one unit, would lead to an increase in the predicted level of democracy by 0.0154. This result points in the direction of the theory by Acemoglu and Robinson (2006). They argue that if a country is labor abundant, an increase in globalization will lead to an increase in democratization.

The variable *linguistic diversity* is positive and significant at the 1 % level. An increase of one unit in the number of linguistic groups compromising at least 5 % of the population will lead to a 1.442 increase in the predicted level of democracy. The variable *regional polity* is positive and significant at the 5 % level. This implies that an increase of one unit in the ten neighboring countries democracy index, leads to an increase in the country's index by 0.642. The variable *1/Years since independence* is positive and significant at the 1 % level. The effect of the variable on democracy decreases with the amount of years since independence. The variable *Americas* is also positive and significant at the 1 % level. This implies that if the country lies in the Americas, the predicted democratic score is 7.505 higher than if the country was in Asia. The constant is -96.50, which is low taking into account that that the democratic index lies in the interval [-10,10]. The R-squared is still high compared with the R-squared from estimation (1), at 48.7 %.

### 7.6.2   Estimations with fixed effects

The fifth model is estimated with the fixed effects method, and the results are shown in table 3. The fixed effects method omits the variables that does not vary in time (*colony after 1945* and the regions *Americas* etc.). The variable *openness* is positive and significant at the 10 % level. This indicates that an increase in the level of openness with one unit, will lead to a 0.0265 higher predicted level of democracy.

The variables *religious diversity* and *linguistic diversity* are both positive and significant at the 5 % level. These variables shows the number of language groups and religious groups compromising at least 5 % of the population, so an increase in the number of language groups will increase the predicted level of democracy by 5.039, and an increase in the number of religious groups will increase the predicted level of democracy by 5.884.

The variable *regional polity* is positive and significant at the 1 % level. If the 10 nearest neighbors of the country in question have a one unit increase in their polity index, the country has an increase of 0.975.

Table 3: Labor abundant countries

| Variables | (4) OLS | (5) FE | (6) 2SLS |
|---|---|---|---|
| Openness | 0.0154 | 0.0265* | 0.0212** |
| Female percent of population | 0.291 | 1.494 | 0.292 |
| Age 15-64 | 0.813 | 1.251 | 0.818*** |
| Age <15 | 0.625 | 0.925 | 0.637*** |
| Ethnic diversity | -0.910 | -2.113 | -0.921*** |
| Religious diversity | 2.200 | 5.884** | 2.180*** |
| Linguistic diversity | 1.442*** | 5.039** | 1.424*** |
| Regional polity | 0.642** | 0.975*** | 0.640*** |
| Colony after 1945 | 0.444 | | 0.401 |
| 1/Years since independence | 10.93*** | 1.830 | 10.96*** |
| America | 7.505*** | | 7.351*** |
| Africa | -0.950 | | -1.127 |
| Pacific | 0.470 | | 0.227 |
| Time | 0.0759 | -0.0158 | 0.0755*** |
| Constant | -96.50 | -203.0* | -97.36*** |
| Observations | 801 | 801 | 801 |
| R-squared | 0.487 | 0.467 | 0.487 |

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

The constant is negative and significant at the 10 % level. The constant is equal to -203. The R-squared is equal to 0.467, which means that the estimation model (5) explains 46,7 % of the variation in the data (Wooldridge, 2013).

### 7.6.3   Estimation with two stage least squares

The sixth model is estimated with the 2SLS method and shown in table 3. The variable *openness* is positive and significant at the 5 % level. This implies that an increase in openness with one unit leads to a 0.0212 increase in the predicted level of democracy. This result lends support to the theory by Acemoglu and Robinson (2006). They argue that if a country is labor abundant and closed for international trade, a opening towards globalization and international trade will increase the probability of democratization.

The variable *age 15-64* is positive and significant at the 1 % level. An increase of

one percentage point in the population in the age of 15 to 64, will lead to an increase in the predicted level of democracy of 0.818. The variable *age < 15* is positive and significant at the 1 % level. This implies that if the percent of the population under the age of 15 increase by one percentage point, the predicted level of democracy will increase with 0.637. This result contradicts the theory that countries with a younger population often are poor, and therefore have a lower probability of sustaining a democracy. This is called "the youth bulge" by researcher such as Cincotta (2008).

The variable *ethnic diversity* is negative and significant at the 1 % level. The variable is the number of ethnic groups comprising at least 5 % of the population. If there is an increase in the number of ethnic groups by one unit, the results indicate that it will lead to a decrease in the predicted level of democracy by 0.921. Authors, such as Cheibub et al. (1996), as well as Hadenius and Teorell (2007), showed that ethnic diversity or ethnic fractionalization impedes democratization and tends to trigger democratic breakdowns. The results from model 6 lends support to their results. The variable *religious diversity* is positive and significant at the 1 % level. The variable is the number of religious groups compromising at least 5 % of the population. If there is an increase in the number of religious groups by one unit, it will lead to a increase in the predicted democratic score by 2.18. The variable *linguistic diversity* is positive and significant at the 1 % level. The variable is the number of language groups compromising at least 5 % of the population. If there is an increase in the number of language groups by one, it will lead to an increase in the predicted democratic score by 1.424.

The variable *regional polity* is positive and significant at the 1 % level. If there is an increase in the democracy score of the country's ten nearest neighbors, this will lead to an increase in the predicted democracy score of the country in question by 0.640. The variable *1/years since independence* is positive and significant at the 1 % level. The effect of the variable on democracy decreases with the number of years since independence. The variable *Americas* is positive and significant at the 1 % level. This means that if a country lies in the Americas, its predicted democratic score is 7.351 higher than if the country is in Asia, which is the reference category. The variable *time* is positive and significant at the 1 % level. This means that from each year that goes by, from 1960 to 2007, the predicted democracy score increases.

The constant is equal -97.36, which lies outside the democracy index interval [-10,10], and the constant is significant at the 1 % level. The R-squared is equal 0.487, which means that the estimated model explain 48.7 % of the variation in the data (Wooldridge, 2013).

### 7.6.4   Comparison of the estimates

Comparing the estimates of the pooled OLS-, fixed effects, and two stage least squares method can be done by comparing the predicted levels of an average labor abundant country. As before, this indicates the bias in the estimation methods. The predicted level of an average labor abundant country is calculated by using the average level of the explanatory variable and the results from the estimations in table 3.

$$\widehat{Democracy}^{OLS} = 5.9485$$
$$\widehat{Democracy}^{FE} = -1.3078$$
$$\widehat{Democracy}^{2SLS} = 6.1007$$

An average labor abundant country estimated with the pooled OLS method has a democratic score of 5.9485, while with the fixed effects method the democratic score is lower at -1.3078, and in the 2SLS methods the score is higher than the pooled OLS at 6.1007. These predicted levels of democracy are much higher than the levels for the whole sample. This implies that labor abundant countries have a higher democratic score than other countries. This is in line with the theory by Acemoglu and Robinson (2006). The difference between the predicted level of democracy with the pooled OLS method and the fixed effects method is small. Fixed effects predict a 7.25 lower democratic score than the pooled OLS method. The difference between the pooled OLS method and the 2SLS method is that the 2SLS predicts a 0.15 higher democratic score than the pooled OLS.

As before, the comparison can also be done by looking at a country's predicted level of democracy one standard deviation below and above. The predicted level of democracy one standard deviation below is:

$$\widehat{Democracy}^{OLS} = -18.22$$

$$\widehat{Democracy}^{FE} = -39.58$$
$$\widehat{Democracy}^{2SLS} = -18.46$$

A country, one standard deviation below the average, estimated with the pooled OLS method has the predicted democratic score of -18.22, while with the fixed effects method the predicted democratic score is -39.58. With the 2SLS method the predicted democracy score is -18.46. The predicted level with pooled OLS and 2SLS are similar, however the predicted level with the fixed effects method is much lower.

All of the predicted values of democracy are outside of the interval of the democracy index. The predicted level of democracy one standard deviation above is:

$$\widehat{Democracy}^{OLS} = 30.12$$

$$\widehat{Democracy}^{FE} = 36.97$$

$$\widehat{Democracy}^{2SLS} = 30.66$$

A country, one standard deviation above the average, estimated with the pooled OLS method has the predicted democracy score of 30.12, which is outside the interval. However, with the fixed effects method the predicted level is 36.97, also way outside of the democracy index's interval. With the 2SLS method the predicted level is 30.66, almost the same as the pooled OLS prediction.

The overall results from the labor abundant countries is that they have a higher predicted democracy score than the overall sample.

## 7.7   Capital and land abundant countries

This section focuses on the results from the capital- and land abundant countries. As already mentioned, labor abundant countries are defined as countries with below median capital and below median arable land. Capital and land abundant countries are defined as countries with above median capital per laborer and above median arable land per person. The capital and land abundant countries in the sample only represent 24.35 % of the total sample.

### 7.7.1   Estimation with pooled OLS

The seventh model is estimated with the pooled OLS method. The results from the regression are shown in table 4.

The variable *openness* is negative and significant at the 10 % level. This means that an increase in *openness* leads to a decrease in the predicted level of democracy. This result is in accordance with the theory by Acemoglu and Robinson (2006). They argue that if a country is capital and land abundant, where only the rich owns the land and the capital, an opening of the country will indicate that the rich gets more wealthy. This will therefore lead to an even more unequal income distribution. This will make the probability for democratization less likely, since the rich do not want to share their fortune.

The variable *age < 15* is negative and significant at the 5 % level. If the percent of the population under the age of 15 increases with one percentage point, the predicted democratic score decreases with 0.479. The variable *ethnic diversity* is positive and significant at the 5 % level. This implies that an increase in the number of ethnic groups compromising at least 5 % of the population will increase the predicted democracy score by 2.22. The variable *regional polity* is positive and significant at the 1 % level. If there is an increase in the democracy score of the country's ten nearest neighbors, this will increase the country's predicted democracy score of 0.993. The constant is equal 63.56, which is outside the democracy index interval. The R-squared is equal 0.645, which means that the estimated model explain 64,5 % of the variation in the data (Wooldridge, 2013).

Table 4: Capital and land abundant countries

| Variables | (7) OLS | (8) FE | (9) 2SLS |
|---|---|---|---|
| Openness | -0.0225* | -0.00676 | -0.0856*** |
| Female percent of population | -0.682 | -0.837 | -1.004*** |
| Age 15-64 | -0.190 | -1.124** | -0.565*** |
| Age <15 | -0.479** | -1.089** | -0.660*** |
| Ethnic diversity | 2.220** | 3.825** | 1.655*** |
| Religious diversity | -0.264 | -0.792 | 0.179 |
| Linguistic diversity | -1.370 | 4.134*** | -0.281 |
| Relgional polity | 0.993*** | 1.225*** | 1.131*** |
| Colony after 1945 | 0.413 | | 0.0409 |
| 1/Years since independence | 3.550 | -0.419 | 2.253 |
| Americas | -2.127 | | -5.657*** |
| Europe | -3.317 | | -6.979*** |
| Africa | -0.701 | | -1.580 |
| Pacific | 1.495 | | -2.810* |
| Time | -0.0579 | -0.128* | -0.0105 |
| Constant | 63.56 | 133.2 | 112.0*** |
| Observations | 1,175 | 1,175 | 1,166 |
| R-squared | 0.645 | 0.584 | 0.553 |

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

### 7.7.2   Estimation with fixed effects

The eight model is estimated with the fixed effects method, and the results are shown in table 4. The fixed effects method omits the variables that does not vary in time (*colony after 1945* and the regions *Americas* etc.). The variable *openness* is negative, however not significant. This is in line with the theory by Acemoglu and Robinson (2006), even though the variable is not significant.

The variables *age 15-64* and *age<15* are negative and significant at the 5 % level. This results are the opposite of the results in model (5). The results here show that an increase of one percentage point in the population under the age of 15, or in the age 15-64 will decrease the predicted democracy score. The variable *ethnic diversity* is positive and significant at the 5 % level. A one unit increase in the number of ethnic groups compromising at least 5 % of the population will increase the predicted democracy score by 3.825. The variable *linguistic diversity* is positive and significant at the 1 % level. A one unit increase in the number of linguistic groups compromising at least 5 % of the population will increase the predicted democracy score by 4.134. The variable *regional polity* is positive and significant at the 1 % level. If the country in question's ten nearest neighbors have a one unit increase in their democracy score, the country's predicted democracy score will increase with 1.225. The variable *time* is negative and significant at the 10 % level. This is the opposite of the results from model (5), where the variable is negative. This implies that when the years go by the predicted democracy score of capital- and land abundant countries will go down. The R-squared is equal to 0.584, which means that model (8) explains 58.4% of the variation in the data set (Wooldridge, 2013).

### 7.7.3   Estimation with two stage least squares

The ninth model is estimates with the 2SLS method, and the results are shown in table 4. The variable *openness* is negative and significant at the 1 % level. This implies that an increase in *openness* leads to a decrease in the predicted democracy score. This result also agree with the theory by Acemoglu and Robinson (2006).

The variable *female percent of population* is negative and significant at the 1 % level. An increase in the female percent of the population with one percentage point will lead to a decrease in the predicted democracy score. The variables *age 15-64* and *age<15* are negative and significant at the 1 % level. The variable *regional polity* is positive and significant at the 1 % level. An increase in the countries ten nearest

neighbors' democratic score will lead to an increase in the country in question's predicted democratic score.

The variable *Americas* and *Europe* are both negative and significant at the 1 % level. The variable *Pacific* is also negative, but only significant at the 10 % level. This implies that capital- and land abundant countries in Asia have a higher predicted democratic score than countries in the Americas, Europe or the Pacific. The constant is positive and significant at the 1 % level. The R-squared is equal 0.553, which implies that the model explain 55,3% of the variation in the data (Wooldridge, 2013).

### 7.7.4   Comparison of the estimates

Comparing the estimates of pooled OLS-, fixed effects- and two stage least squares method can be done by comparing the predicted level of an average capital- and land abundant country. As before, this shows the bias in the estimation methods. The predicted level of an average capital- and land abundant country is calculated by using the average level of the explanatory variable and the results from the estimations in table 3:

$$\widehat{Democracy}^{OLS} = 1.47$$
$$\widehat{Democracy}^{FE} = -0.54$$
$$\widehat{Democracy}^{2SLS} = 0.93$$

An average capital- and land abundant country estimated with the pooled OLS method has a democratic score of 1.47. With the fixed effects method, the democratic score is lower at -0.54, and with the 2SLS method the score is in between at 0.93. These predicted levels of democracy is on average lower than the levels for the whole sample. This agrees with the theory by Acemoglu and Robinson (2006). The difference between the predicted level of democracy with the pooled OLS method and the 2SLS method is still small, at 0.53. The fixed effects method predicts a lower democratic score than both the pooled OLS and the 2SLS method.

As before, the comparison can also be done by looking at a country's predicted level of democracy one standard deviation below and above. The predicted level of democracy one standard deviation below is:

$$\widehat{Democracy}^{OLS} = 3.06$$

$$\widehat{Democracy}^{FE} = 3.59$$

$$\widehat{Democracy}^{2SLS} = 9.36$$

A country, one standard deviation below the average, estimated with the pooled OLS method has the predicted democratic score of 3.06, while with the fixed effects method the predicted democratic score is 3.59. With the 2SLS method the predicted democracy score is 9.36, which is higher than the average. The predicted level of democracy one standard deviation above is:

$$\widehat{Democracy}^{OLS} = -0.12$$

$$\widehat{Democracy}^{FE} = -4.69$$

$$\widehat{Democracy}^{2SLS} = -7.49$$

A country, one standard deviation above the average, estimated with the pooled OLS method has the predicted democracy score of -0.12, with the fixed effects method the predicted level is -4.69, and with the 2SLS method the predicted level is -7.49. This is the opposite of the predicted levels of the labor abundant countries, one standard deviation below, where the predicted levels are much higher. These results agree with the theory by Acemoglu and Robinson (2006) that countries with an over median capital and land, have a lower probability of consolidating democracy.

The overall results from the capital and land abundant countries points in the direction that they have a lower predicted democracy score than the overall sample.

## 7.8   Hypothesis results

As Acemoglu and Robinson (2006) argues that if a country is abundant in land and capital, globalization will lead to a lower probability for democracy. If a country is abundant in labor, globalization will lead to a higher probability of democracy. As the hypothesis in chapter 4 states:

$H_1$ : Globalization leads to increased democracy in labor abundant countries.

$H_2$ : Globalization leads to less democracy in capital and land abundant countries.

The results from the regression, with different estimation methods, support these hypotheses. The results from this chapter are not the same as those of Doces and Magee (2015), who obtain an insignificant result. The reason is that they do not use the variable *arable* to take into account that a country's land resources could have an impact.

# 8    Robustness

This chapter examines the robustness of the previous findings in several ways: Alternative econometric set-ups, alternative time periods, additional explanatory variables and interaction terms. Overall, the main results from the previous chapter remain.

## 8.1    Alternative econometric approach: GMM

Generalized method of moments (GMM) is an alternative method to estimate equation (25) and (26). If there is endogeneity in the model and heteroskedaticity, GMM is an alternative way to estimate the model.

The method of moments approach to parameter estimation dates back more than 100 years (Stigler, 1986). The notion of a moment (i.e. the expected values of the random variable under examination) is fundamental for describing features of a population. The theory of GMM uses two sets of population moment conditions, in a manner that minimizes the asymptotic variance among method of moments estimators (Wooldridge, 2001). GMM weights the two sample moment conditions to obtain an asymptotically optimal estimator.

Cragg (1983) discovered that one could improve compared to OLS in the presence of heteroskedasticity of unknown form by applying generalized moments of methods. How GMM works, is first by choosing which extra moment conditions to add to those generated by the usual zero correlation assumption. Next, having first executed ordinary least squares, one must obtain the weighting matrix that is a crucial component in an efficient GMM analysis. The weighting matrix is obtained by inverting a consistent estimator of the variance-covariance matrix of the moment conditions. The GMM estimator minimizes a quadratic form in the sample moment conditions, where the weighting matrix appears in the quadratic form (Wooldridge, 2001).

There are some problems with the generalized method of moments. One problem that arises is that the researchers have to choose the additional moment conditions to be added in an ad hoc manner. If there are two researchers, they would probably choose different moment conditions. Thus, this procedure would open one's research to the criticism of searching over different sets of moment conditions until the desired result is achieved. The second issue is that ordinary least squares will be unbiased as well as consistent, whereas GMM is only guaranteed to be consistent.

Given the additional decisions required for using GMM to improve ordinary least squares, it is understandable that most researchers stick with ordinary least squares. If there are concerns about heteroskedasticity, they can use standard errors and test statistics that are robust to heteroskedasticity, as White (1980).

Hansen (1982) showed that every previously suggested instrumental variable estimator, in linear or nonlinear models, with cross-section, time series or panel data, could be become a GMM estimator. Hansen also showed how to choose among the many possible methods of moments estimators in a framework that allows for heteroskedasticity, serial correlation and nonlinearities. The optimal GMM estimator is asymptotically no less efficient than two-stages least squares under homoskedasticity, and is generally better under heteroskedasticity. GMM estimators can often be found more efficient than common methods of moments estimators, such as ordinary least squares and two-stages least squares, when assumptions concerning homoskedasticity fails.

In this thesis, the GMM estimation method is used as an alternative approach to the problems of the model, too see if an alternative approach will give the same results as in chapter 7.

### 8.1.1   Results

The results from the regression with the GMM estimation method are similar to the results from chapter 7, and shown in table 5[5].

In the full sample, the variable *openness* is negative and significant at 10 % level. This is the similar to the results from estimation model (1) with the pooled OLS method, and estimation of model (3) with the 2SLS method. The estimates from estimation model (10) are closer to the 2SLS methods estimations than the pooled OLS methods estimations. The standard errors in (10) are almost the same as the standard errors in the 2SLS results with robust standard errors. The rest of the point estimates and statistical significance are quite similar between the 2SLS and the GMM estimation method. Using GMM do not fundamentally affect the results, and perhaps with other models or data sets it might have offered greater precision. However, both alternative estimation methods seem reasonable.

In the sample of capital and land abundant countries the results are also similar to pooled OLS -and 2SLS method estimations (from table 4). The variable *openness* is

---

[5]See appendix B.1

negative and significant at the 1 % level. The standard errors in (11) are also similar to the standard errors in the 2SLS results with robust standard errors. Likewise, the rest of the point estimates and statistical significance are similar between the 2SLS and the GMM estimation method.

In the sample of labor abundant countries the results are also similar to the pooled OLS and 2SLS methods estimations (from table 3). The variable *openness* is positive and significant at the 1 % level. The standard errors in (12) are also corresponding to the standard errors in (6), which is the 2SLS results with robust standard errors. Similarly, the rest of the point estimates and statistical significance are corresponding to the 2SLS estimation method.

These results point in the same direction as the previous results. Thus, they are in line with the predictions of the model by Acemoglu and Robinson (2006).

## 8.2   Alternative periods

The panel data runs from 1960 to 2007. From 1978 to 1995 the number of democracies in the world increased significantly. This was mainly due to the fall of the Berlin wall, the end of the Soviet Union, and the democratization in Latin America. Countries that had closed economies before 1978, started to open up their economies. These events could be affecting the estimates, and driving the main results. To test if this period is inducing the results, a separate regression that excludes the period from 1978 to 1995 would be useful. In the whole sample, the previous results carry over[6].

In the labor abundant countries the results are similar to the previous results (i.e. table 3), and the results are just as significant as before.

In the capital and land abundant countries the results are similar to the previous results (i.e. table 4). However, not as significant as before. The coefficients are closer to zero, thus making the predicted levels of democracy less extreme. However, this result points in the direction of the model by Acemoglu and Robinson, which says that capital and land abundant countries have a lower probability for democratization.

---

[6]See appendix B.2.

## 8.3 Additional explanatory variables

Including additional explanatory variables could enhance the explanatory power of the model even further (Webster, 2013). To check if the results remain, even if the model is extended, this estimation includes logged GDP per capita (GDPPC), population, the urban percentage of the population, and an oil exporter dummy variable. This strengthens the support for the theory[7].

In the whole sample, the results are similar to the previous results. The variable *openness* is negative and significant in both the pooled OLS method and the 2SLS method estimation. The variable *oil exporter* is negative and significant in the pooled OLS method model and the 2SLS method model. If a country is an oil exporter, the predicted level of democracy is 2.804 lower in the pooled OLS method estimation, and 2.031 in the 2SLS method estimation. Oil exporting countries tend to have a richer elite, which leads to an increased income inequality, and hence a lower predicted democracy score. The variable *urban* is only significant in the 2SLS method estimation. Here, the variable is positive. If the urban percentage of population increases with one percentage point, the predicted democracy score increases with 0.033. The variable *log(GDPPC)* is positive and significant in the pooled OLS method estimation and the 2SLS method estimation. If the GDP per capita in the country in question increases with one percentage point, the predicted change in the democracy score is 0.1434. The variable *log(population)* is insignificant in all three models of estimation. The R-squared is higher when including the four additional explanatory variables.

In labor abundant countries the results are similar to the previous results. However, the variable *openness* is not significant in either of the models. As in the total sample, the variable *oil exporter* is negative. This indicate that if a country is an oil exporter, the predicted democracy score is lower than if the country does not export oil. The variable *log(GNPPC)* is negative and significant at the 1 % level in all models. In the fixed effects method estimation the variable indicates that if there is an increase of one percentage point in the GDP per capita, the predicted level of democracy will increase with 0.587. The R-square is higher in this estimation than the previous estimations.

In capital and land abundant countries the results are also similar to the previous results. The variable *openness* is negative, however it is only significant in the 2SLS method estimation. The regional variables are more extreme than previous results.

---

[7]For regression tables see appendix B.3.

Hence, the variable *Europe* is negative and significant at the 1 % level. If a country is in Europe, the predicted level of democracy is 13.38 levels lower than if the country is in Asia. The variable *log(GDPPC)* is positive and significant in the pooled OLS method estimation and the 2SLS method estimation. If there is an increase in the level of GDP per capita, the predicted level of democracy is increased by 0.419. The variable *log(population)* is negative and significant at the 1 % level in the 2SLS method estimation. An increase in the population by one percentage point, will decrease the predicted level of democracy by 0.3389. The R-squared in the pooled OLS method estimation and the FE method estimation is similar to previous results. However, the R-squared in the 2SLS method estimation is half as high as in previous estimations.

## 8.4  Interaction term

Doces and Magee (2015) argue that it is very difficult to estimate a model where the equation has an endogenous variable that is both in the equation itself and part of an interaction. Thus, they use 2SLS to estimate the equation separately for the groups of countries above the world median capital-labor ratio and below. In the previous results in this paper, the same approach is used. However, in this robustness check it is tested whether an interaction term would have an impact. First, the estimation includes the variables *above* and *below* to examine if the two variables have an impact on the results. The variable *above* is equal to 1 if a country is capital and land abundant. The variable *below* is equal to 1 if a country is labor abundant. Second, the estimation only includes the variable *below* and the interaction term between *openness* and *below*. Third, the estimation only includes the variable *above* and the interaction term between *openness* and *above*. Fourth, the estimation includes all the variables and interaction terms. [8]

The results from including the variables *above* and *below* is shown in table 12 in the appendix. The results are similar to the previous results. The variable *openness* is still negative, the only difference is that now the variable is not significant in the 2SLS method estimation, only in the pooled OLS method estimation. A few variables have also become insignificant. The variable *above* is negative and significant in both the pooled OLS method estimation and the 2SLS method estimation. This indicates that if a country is capital and land abundant their predicted democracy score is 1.595 lower than if was not. The variable *below* is also negative, although, the

---

[8]For estimation result, see appendix B.4.

variable is not significant.

The results from including the variable *below* and the interaction term *openness\*below* are shown in table 13[9]. The results are similar to the previous results. The main difference is that in table 13 the whole sample of countries is included. The variable *openness\*below* is positive and significant in the 2SLS method estimation. If a country is labor abundant the results indicate that an increase in globalization leads to a higher predicted level of democracy, than if a country is not labor abundant. The variable *below* is negative and significant in the 2SLS method estimation, indicating that labor abundant countries tend to have a lower predicted level of democracy.

The results from including the variable *above* and the interaction term *openness\*above* are shown in table 14[10]. The result are similar to the previous result. The main difference is that the variables are not as significant as before. The variable *openness\*above* is negative, however, not significant. The variable *above* is not significant either.

The results from including the variables *above* and *below*, and the interaction terms *openness\*above* and *openness\*below* are shown in table 15[11]. The results are similar to the previous results. The variable *openness* is negative and significant in both the pooled OLS method estimation and the 2SLS method estimation. The interaction term *openness\*above* is negative in the pooled OLS method estimation and positive in the two other estimations. However, the interaction term is not significant in any of the estimations. The interaction term *openness\*below* is positive and significant in the pooled OLS method estimation and the 2SLS method estimation, indicating that an increase in the level of globalization in labor abundant countries, will increase the predicted level of democratization. The interesting results are the variables *above* and *below*. A country could either be a capital and land abundant country, a labor abundant country, or in between. The variable *above* is negative and significant in the 2SLS method estimation. This indicates that a capital and land abundant country has a 1.827 lower predicted level of democracy than the countries in between. The variable *below* is negative and significant in both the pooled OLS method estimation and the 2SLS method estimation. Indicating that a labor abundant country have an even lower democracy score than a capital and land abundant country. The predicted level of democracy in a labor abundant country is 3.350 lower than the countries in between the two variables.

---

[9]See appendix B.4

[10]See appendix B.4

[11]See appendix B.4

When dealing with three endogenous explanatory variables it is important to check the identification of the instruments. This could be done by using The Cragg-Donald F-test statistics, just as in chapter 7.4. If the instruments only exhibit weak partial correlation with the endogenous regressors the instrument is said to be weak or poor. This could lead to large asymptotic variance, which could lead to inconsistency. In the Cragg-Donald test the null hypothesis is that the instruments are weak against the alternative that they are not. For the instruments of the variables *openness\*above* and *openness\*below* the null hypothesis is rejected at the 5 % level, indicating that the instruments are strong (Stock & Yogo, 2005).

# 9   Conclusion

This thesis has tested the Acemoglu and Robinson (2006) hypotheses that relative factor endowments condition the effect of globalization on democracy. In relatively labor abundant countries, globalization increases the production using labor, thus producing democratizing effects. However, in capital and land abundant countries globalization is not helpful for the development of democracy.

In labor abundant countries, opening up to world trade will increase the returns to labor, and the laborer will get relatively richer. The income distribution will become more equal and the elites in power will have less to lose by introducing democracy.

In capital and land abundant countries, opening up to world trade will increase the returns to capital and land, and the rich will get relatively richer. Thus, making the income distribution more askew leads to a lower probability for democratization.

Based on these hypotheses by Acemoglu and Robinson (2006), the empirical approach uses panel data for 217 countries between 1960 and 2007 to examine how a country's relative factor endowments affect whether democratization is an option when a non democratic country starts opening up to globalization.

The main model in the empirical estimation includes 217 countries, and estimates the effect of trade on democracy. The estimation also includes explanatory variables, such as the gender percentage, age and diversity of the population, region and colonial heritage. The empirical approach includes estimations with the pooled OLS method, the fixed effect method and the two stage least square method. The results indicate that an increase in globalization has a negative impact on democracy.

To test the model by Acemoglu and Robinson (2006) the sample of countries are

separated into two groups: Labor abundant countries and capital and land abundant countries. Labor abundant countries are defined as countries with below median capital per laborer and below median arable land per person. Capital and land abundant countries are defined as countries with above median capital per laborer and above median arable land per person. In the empirical estimation with only labor abundant countries, the result indicate that an increase in globalization has a positive impact on democracy. In the empirical estimation with only the capital and land abundant countries, the results indicate that an increase in globalization has a negative impact on democracy

The results from the empirical estimations indicate that labor abundant countries have a higher predicted level of democracy when increasing trade, and that capital and land abundant countries have a lower predicted level of democracy when increasing trade. Thus, the results from the empirical estimation concur with the hypotheses theorized by Acemoglu and Robinson (2006).

To examine the robustness of the empirical estimations alternative econometric set-ups, alternative time periods, additional explanatory variables and interaction terms are tested. The results are similar to the main empirical estimations, and thus the main results are considered robust.

# References

Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy.* Cambridge, MA: Cambridge University Press.

Ahlquist, J. S., & Wibbels, E. (2012). Riding the wave: World trade and factor-based models of democratization. *American Journal of Political Science*, *56*(2), 447–464.

Albrow, M., & King, E. (1990). *Globalization, knowledge and society: readings from international sociology.* Newbury Park, CA: SAGE Publications.

Asteriou, D., & Hall, S. G. (2007). *Applied econometrics: a modern approach, revised edition.* Hampshire, UK: Palgrave Macmillan.

Brinks, D., & Coppedge, M. (2006). Diffusion is no illusion: Neighbor emulation in the third wave of democracy. *Comparative Political Studies*, *39*(4), 463–489.

Buc, H. B. (2006). How chile successfully transformed its economy. *Backgrounder*, *1958*, 1–10.

Bussmann, M. (2001). Examining causality among conflict, democracy, openness, and economic growth. *Unpublished paper, University of Alabama*.

Cheibub, J. A., Przeworski, A., Limongi Neto, F. P., & Alvarez, M. M. (1996). What makes democracies endure? *Journal of democracy*, *7*(1), 39–55.

Cincotta, R. P. (2008). Half a chance: Youth bulges and transitions to liberal democracy. *Environmental Change and Security Program Report*(13), 10–18.

Cragg, J. G. (1983). More efficient estimation in the presence of heteroscedasticity of unknown form. *Econometrica: Journal of the Econometric Society*, *51*(3), 751–763.

Cumings, B. (2005). *Korea's place in the sun: A modern history (updated).* New York, NY: WW Norton and Company.

Doces, J. A. (2006). Globalization and democracy: An empirical analysis using a simultaneous equation approach. *Center for Globalization and Governance*.

Doces, J. A., & Magee, C. S. (2015). Trade and democracy: A factor-based approach. *International Interactions*, *41*(2), 407–425.

Eichengreen, B., & Leblang, D. (2008). Democracy and globalization. *Economics & Politics*, *20*(3), 289–334.

FAO. (2017). World development indicators. *The World Bank*. Retrieved from `http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators#`

FreedomHouse. (2017). *Freedom in the world 2017, cuba profile.* Freedom House. Retrieved 04. may 2017, from `https://freedomhouse.org/report/freedom`

`-world/2017/cuba`

Giavazzi, F., & Tabellini, G. (2005). Economic and political liberalizations. *Journal of monetary economics*, *52*(7), 1297–1330.

Griswold, D. T. (2004). Trading tyranny for freedom: How open markets till the soil for democracy. *Trade Policy Analysis*, *26*, 1–16.

Hadenius, A., & Teorell, J. (2007). Pathways from authoritarianism. *Journal of democracy*, *18*(1), 143–157.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, *50*(4), 1029–1054.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, *46*(6), 1251–1271.

Hayek, F. A. v. (1960). *The constitution of liberty*. Chicago, IL: University of Chicago Press.

Head, K., Mayer, T., & Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of international Economics*, *81*(1), 1–14.

Hill, C. R., Griffiths, W. E., & Lim, G. C. (2011). *Principles of econometrics* (4th ed.). Hoboken, NJ: John Wiley Sons.

Hirst, T. (2015). *A brief history of china's economic growth*. Retrieved 4. may 2017, from `https://www.weforum.org/agenda/2015/07/brief-history-of-china-economic-growth/`

Huntington, S. P. (1991). Democracy's third wave. *Journal of democracy*, *2*(2), 12–34.

IMF. (2000). Globalization: Threat or opportunity? *International Monetary Found briefs*. Retrieved from `http://www.imf.org/external/np/exr/ib/2000/041200to.htm`

Kant, I. (1795). *Perpetual peace: A philosophical sketch*. Indianapolis, IN: Hackett Publishers.

Leamer, E. E. (1984). *Sources of international comparative advantage: Theory and evidence*. Cambridge, MA: MIT press.

Li, Q., & Reuveny, R. (2003). Economic globalization and democracy: An empirical analysis. *British Journal of Political Science*, *33*(01), 29–54.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American political science review*, *53*(01), 69–105.

López-Córdova, J. E., & Meissner, C. M. (2005). The globalization of trade and democracy, 1870-2000. *NBER Working Paper*.

López-Córdova, J. E., & Meissner, C. M. (2008). The impact of international trade

on democracy: A long-run perspective. *World Politics*, *60*(04), 539–575.

Marshall, M. G., Gurr, T. R., & Jaggers, K. (2008). Polity iv project. *Center for Systemic Peace*. Retrieved from `http://www.systemicpeace.org/inscrdata.html`

Maxfield, S. (2000). Capital mobility and democratic stability. *Journal of Democracy*, *11*(4), 95–106.

Milner, H. V., & Kubota, K. (2005). Why the move to free trade? democracy and trade policy in the developing countries. *International organization*, *59*(01), 107–143.

Mutascu, M. I. (2009). Population growth and democracy: An extreme value analyzes in romania's case. *Journal of Applied Quantitative Methods*, *4*(3), 259-269.

O'Rourke, K. H., & Taylor, A. M. (2006). Democracy and protectionism. *National Bureau of Economic Research*.

Papaioannou, E., & Siourounis, G. (2008). Economic and social factors driving the third wave of democratization. *Journal of comparative Economics*, *36*(3), 365–387.

Rudra, N. (2005). Globalization and the strengthening of democracy in the developing world. *American Journal of Political Science*, *49*(4), 704–730.

Schumpeter, J. A. (1950). The march into socialism. *The American Economic Review*, *40*(2), 446–456.

Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900.* Cambridge, MA: Harvard University Press.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80-108.

Webster, A. (2013). *Introductory regression analysis: With computer application for business and economics.* New York, NY: Routledge.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, *48*(4), 817–838.

Wooldridge, J. M. (2001). Applications of generalized method of moments estimation. *The Journal of Economic Perspectives*, *15*(4), 87–100.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* Massachusetts, MA: MIT press.

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach* (5th ed.). South-Western, Cengage Learning.

Yu, M. (2007). Trade globalization and political liberalization: a gravity approach. *TIGER Working Paper*, *104*.

# Appendix

# A   List over countries in data set

| | | |
|---|---|---|
| Afghanistan | Botswana | Cuba |
| Albania | Brazil | Cyprus |
| Algeria | Brunei | Czech Republic |
| American Samoa | Bulgaria | Denmark |
| Angola | Burkina Faso | Djibouti |
| Antigua and Barbuda | Burundi | Dominica |
| Argentina | Czechoslovakia | Dominican Republic |
| Armenia | Cambodia | Ecuador |
| Aruba | Cameroon | Egypt |
| Australia | Canada | El Salvador |
| Austria | Cape Verde | Equatorial Guinea |
| Azerbaijan | Cayman Islands | Eritrea |
| Bahamas | Central African Republic | Estonia |
| Bahrain | Chad | Ethiopia |
| Bangladesh | Channel Islands | Faeroe Islands |
| Barbados | Chile | Fiji |
| Belarus | China | Finland |
| Belgium | Colombia | France |
| Belize | Comoros | French Polynesia |
| Benin | Congo, Dem. Rep. | Germany, East |
| Bermuda | Congo, Republic of | Germany, West |
| Bhutan | Costa Rica | Gabon |
| Bolivia | Cote d'Ivoire | Gambia, The |
| Bosnia and Herzegovina | Croatia | Georgia |

| | | |
|---|---|---|
| Germany | Kenya | Mayotte |
| Ghana | Kiribati | Mexico |
| Gibraltar | Korea, Dem. Rep. | Micronesia, Fed. Sts. |
| Greece | Korea, Republic of | Moldova |
| Greenland | Kosovo | Monaco |
| Grenada | Kuwait | Mongolia |
| Guam | Kyrgyzstan | Montenegro |
| Guatemala | Laos | Morocco |
| Guinea | Latvia | Mozambique |
| Guinea-Bissau | Lebanon | Myanmar |
| Guyana | Lesotho | Namibia |
| Haiti | Liberia | Nepal |
| Honduras | Libya | Netherlands |
| Hong Kong | Liechtenstein | Netherlands Antilles |
| Hungary | Lithuania | New Caledonia |
| Iceland | Luxembourg | New Zealand |
| India | Macao | Nicaragua |
| Indonesia | Macedonia | Niger |
| Iran | Madagascar | Nigeria |
| Iraq | Malawi | Northern Mariana Islands |
| Ireland | Malaysia | Norway |
| Israel | Maldives | Oman |
| Italy | Mali | Pakistan |
| Jamaica | Malta | Palau |
| Japan | Marshall Islands | Panama |
| Jordan | Mauritania | Papua New Guinea |
| Kazakhstan | Mauritius | Paraguay |

| | | |
|---|---|---|
| Peru | South Africa | Turkey |
| Philippines | Spain | Turkmenistan |
| Poland | Sri Lanka | Turks and Caicos Islands |
| Portugal | St. Kitts and Nevis | Tuvalu |
| Puerto Rico | St. Lucia | U.S.S.R. |
| Qatar | St.Vincent and Grenadines | Uganda |
| Romania | | Ukraine |
| Russia | Sudan | United Arab Emirates |
| Rwanda | Suriname | United Kingdom |
| Samoa | Swaziland | United States |
| San Marino | Sweden | Uruguay |
| Sao Tome and Principe | Switzerland | Uzbekistan |
| Saudi Arabia | Syria | Vanuatu |
| Senegal | Taiwan | Venezuela |
| Serbia | Tajikistan | Vietnam |
| Seychelles | Tanzania | Virgin Islands (U.S.) |
| Sierra Leone | Thailand | Western Samoa |
| Singapore | Timor-Leste | Yugoslavia |
| Slovak Republic | Togo | Yemen |
| Slovenia | Tonga | Zambia |
| Solomon Islands | Trinidad and Tobago | Zimbabwe |
| Somalia | Tunisia | |

# B    Robustness check

## B.1    Generalized method of moments

Table 5: Generalized method of moments.

| Variables | (10) GMM | (11) GMM capital/land | (12) GMM labor |
|---|---|---|---|
| Openness | -0.00837* | -0.0880*** | 0.0220** |
| Female percent of population | 0.233*** | -1.047*** | 0.237 |
| Age 15-64 | -0.269*** | -0.585*** | 0.794*** |
| Age <15 | -0.441*** | -0.664*** | 0.621** |
| Ethnic diversity | -0.0279 | 1.588*** | -0.929*** |
| Religious diversity | 0.166 | 0.213 | 2.158*** |
| Linguistic diversity | 0.463*** | -0.211 | 1.393*** |
| Regional polity | 0.729*** | 1.136*** | 0.660*** |
| Colony after 1945 | 1.973*** | -0.00137 | 0.637 |
| 1/Years since independence | 4.858*** | 2.211 | 10.97*** |
| America | 1.497*** | -5.805*** | 7.387*** |
| Europe | -1.846*** | -7.089*** | |
| Africa | -0.549** | -1.716 | -1.126 |
| Pacific | 2.820*** | -2.987* | 0.0493 |
| Time | 0.0198** | -0.00703 | 0.0787*** |
| Constant | 17.46*** | 115.7*** | -92.85*** |
| Observations | 4,814 | 1,166 | 801 |
| R-squared | 0.539 | 0.547 | 0.486 |

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

## B.2 Estimates excluding the time period 1978-1995

Table 6: Estimates from the pooled OLS, the fixed effects and the 2SLS estimation excluding the time period 1978-1995.

| Variables | (13) OLS | (14) FE | (15) 2SLS |
|---|---|---|---|
| Openness | -0.0134** | -0.00447 | -0.0167*** |
| Female percent of population | 0.219 | -0.0262 | 0.204*** |
| Age 15-64 | -0.244 | -0.0404 | -0.248*** |
| Age <15 | -0.381** | -0.0655 | -0.383*** |
| Ethnic diversity | -0.309 | 1.108 | -0.304*** |
| Religious diversity | 0.694 | -2.558 | 0.727*** |
| Linguistic diversity | 0.609** | 2.986** | 0.613*** |
| Regional polity | 0.783*** | 0.897*** | 0.786*** |
| Colony after 1945 | 1.816* | | 1.880*** |
| 1/Years since indpendence | 5.776*** | 5.350*** | 5.864*** |
| America | 1.694 | | 1.702*** |
| Europe | -1.632 | | -1.675*** |
| Africa | -1.077 | | -1.130*** |
| Pacific | 2.207 | | 2.191*** |
| Time | 0.0262 | 0.0297 | 0.0285*** |
| Constant | 14.13 | 1.047 | 15.26** |
| Observations | 2,893 | 2,893 | 2,883 |
| R-squared | 0.541 | 0.473 | 0.537 |

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

Table 7: Labor abundant countries excluding the time period 1978-1995.

| Variables | (16) OLS | (17) FE | (18) 2SLS |
|---|---|---|---|
| Openness | 0.0121 | 0.0404** | 0.0274** |
| Female percent of population | 0.619 | 1.516 | 0.647* |
| Age 15-64 | 0.857 | 0.00987 | 0.890** |
| Age <15 | 0.730 | 0.0426 | 0.770** |
| Ethnic diversity | -0.860 | -17.22*** | -0.881*** |
| Religious diversity | 2.442* | 3.790 | 2.328*** |
| Linguistic diversity | 1.443*** | 3.278 | 1.404*** |
| Regional polity | 0.533* | 1.319*** | 0.511*** |
| Colony after 1945 | -0.637 | | -0.694 |
| 1/Years since independence | 13.32*** | 2.673 | 13.28*** |
| America | 6.393** | | 6.039*** |
| Africa | -2.294 | | -2.621*** |
| Pacific | -0.660 | | -1.036 |
| Time | 0.108* | 0.00489 | 0.105*** |
| Constant | -119.3 | -59.53 | -124.4*** |
| Observations | 481 | 481 | 481 |
| R-squared | 0.493 | 0.488 | 0.488 |

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

Table 8: Capital and land abundant countries excluding the time period 1978-1995.

| Variables | (19) OLS | (20) FE | (21) 2SLS |
|---|---|---|---|
| Openness | -0.0142 | -0.000880 | -0.0710** |
| Female percent of population | -1.476 | -0.425 | -1.745*** |
| Age 15-64 | 0.0953 | -1.267* | -0.161 |
| Age <15 | -0.283 | -1.521** | -0.399*** |
| Ethnic diversity | 0.793 | 6.202** | 0.634 |
| Religious diversity | 0.653 | -2.765 | 1.195** |
| Linguistic diversity | -0.289 | 4.662 | 0.177 |
| Regional polity | 0.931*** | 1.176*** | 1.040*** |
| Colony after 1945 | 0.466 | | 0.417 |
| 1/Years since independence | 8.724* | -0.255 | 6.931** |
| America | 1.105 | | -1.463 |
| Europe | -1.077 | | -3.498* |
| Africa | -2.719 | | -2.495 |
| Pacific | 3.517 | | 0.146 |
| Time | -0.0432 | -0.216** | 0.00312 |
| Constant | 77.06 | 132.4 | 112.4*** |
| Observations | 676 | 676 | 667 |
| R-squared | 0.643 | 0.623 | 0.565 |

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

## B.3   Additional explanatory variables

Table 9: Estimates from the pooled OLS, fixed effects and 2SLS estimation including *oil*, *urban*, *log(GDPPC)* and *log(population)*.

| Variables | (22) OLS | (23) FE | (24) 2SLS |
|---|---|---|---|
| Openness | -0.0149** | -0.00330 | -0.0555*** |
| Female percent of population | 0.277 | 0.190 | 0.294*** |
| Age 15-64 | -0.212 | -0.202 | -0.0873 |
| Age <15 | -0.246 | -0.172 | -0.127** |
| Ethnic diversity | 0.111 | 0.325 | 0.0592 |
| Religious diversity | -0.413 | 0.978 | 0.467 |
| Linguistic diversity | 0.602* | 1.886 | 0.793*** |
| Regional polity | 0.784*** | 1.020*** | 0.858*** |
| Colony after 1945 | 0.721 | | 0.672** |
| 1/Years since independence | 4.517* | 6.625** | 4.243*** |
| America | -0.469 | | -1.240** |
| Europe | -2.526 | | -3.174*** |
| Africa | -0.889 | | -2.021*** |
| Pacific | 2.051 | | -0.446 |
| Time | 0.0237 | -0.0234 | 0.0475*** |
| Oil exporter | -2.804** | -1.106 | -2.031*** |
| Urban | 0.0176 | -0.0200 | 0.0330*** |
| Log (GDPPC) | 1.176* | -0.193 | 1.434*** |
| Log (Population) | 0.349 | 2.881 | -0.333 |
| Constant | -3.266 | -3.321 | -15.53 |
| Observations | 3,152 | 3,152 | 3,142 |
| R-squared | 0.628 | 0.500 | 0.578 |

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

Table 10:   Labor abundant countries including *oil, urban, log(GDPPC)* and *log(population)*.

| Variables | (25) OLS | (26) FE | (27) 2SLS |
|---|---|---|---|
| Openness | 0.00476 | 0.00809 | -0.0130 |
| Female percent of population | -1.542 | 2.301** | -1.282** |
| Age 15-64 | -0.381 | 0.0933 | -0.217 |
| Age < 15 | -0.197 | -0.302 | -0.103 |
| Ethnic diversity | -0.00633 | -2.208*** | -0.206 |
| Religious diversity | -1.267 | 6.796*** | -1.228 |
| Linguistic diversity | 1.819*** | 4.405 | 1.859*** |
| Regional polity | 0.675** | 0.760** | 0.675*** |
| Colony after 1945 | -5.722 | | -5.149** |
| 1/Years since independence | 10.69*** | 10.77** | 9.921*** |
| America | -3.253 | | -2.620 |
| Africa | -6.011 | | -5.200*** |
| Time | 0.260*** | 0.344 | 0.255*** |
| Oil exporter | -2.636** | -0.358 | -2.187*** |
| Urban | 0.0718 | 0.0525 | 0.0640 |
| Log (GDPPC) | -2.849*** | -5.870*** | -2.627*** |
| Log (population) | -0.393 | -9.119 | -0.532 |
| Constant | 122.6 | -60.83 | 96.36* |
| Observations | 466 | 466 | 466 |
| R-squared | 0.647 | 0.605 | 0.644 |

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

Table 11: Capital and land abundant countries including *oil, urban, log(GDPPC)* and *log(population)*.

| Variables | (28) OLS | (29) FE | (30) 2SLS |
|---|---|---|---|
| Openness | -0.0175 | -0.00371 | -0.180*** |
| Female percent of population | -0.916 | -1.147 | -1.899*** |
| Age 15-64 | 0.126 | -1.197* | -0.180 |
| Age <15 | -0.0559 | -1.179* | -0.137 |
| Ethnic diversity | 1.236 | 2.907 | 1.429*** |
| Religious diversity | 0.459 | -0.510 | 0.698 |
| Linguistic diversity | -0.132 | 4.353** | 0.721 |
| Regional polity | 0.992*** | 1.231*** | 1.333*** |
| Colony after 1945 | -0.240 | | -0.207 |
| 1/ Years since independence | -3.807 | -1.965 | -13.35** |
| America | -4.746 | | -10.64*** |
| Europe | -3.777 | | -13.63*** |
| Africa | 1.917 | | -4.393 |
| Pacific | -3.843 | | -13.38*** |
| Time | -0.0265 | -0.166 | 0.111* |
| Oil | -3.533 | 1.034 | -1.217 |
| Urban | 0.0866 | 0.0314 | -0.0570 |
| Log (GDPPC) | 2.537** | -0.385 | 4.190*** |
| Log (population) | -0.336 | 2.251 | -3.389*** |
| Constant | 19.16 | 154.2 | 96.95** |
| Observations | 1,021 | 1,021 | 1,012 |
| R-squared | 0.684 | 0.574 | 0.282 |

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

## B.4   Interaction terms

Table 12: Estimates from the pooled OLS, fixed effects and 2SLS estimation including variables *above* and *below*.

| Variables | (31) OLS | (32) FE | (33) 2SLS |
|---|---|---|---|
| Openness | -0.0153** | -0.00365 | -0.000682 |
| Female percent of population | 0.343 | 0.0178 | 0.415*** |
| Age 15-64 | -0.0550 | -0.0678 | -0.0826 |
| Age <15 | -0.328* | -0.0798 | -0.348*** |
| Ethnic diversity | -0.0316 | 0.968 | -0.0117 |
| Religious diversity | 0.144 | -1.223 | -0.0324 |
| Linguistic diversity | 0.364 | 2.956*** | 0.356*** |
| Regional polity | 0.720*** | 0.928*** | 0.705*** |
| Colony after 1945 | 1.572 | | 1.475*** |
| 1/Years since independence | 4.357** | 4.713*** | 3.858*** |
| America | 2.542 | | 2.488*** |
| Europe | -1.456 | | -1.359*** |
| Africa | -0.460 | | -0.391 |
| Pacific | 2.996 | | 3.257*** |
| Time | 0.00863 | 0.0226 | -0.00112 |
| Above | -1.595** | -0.529 | -1.466*** |
| Below | -0.495 | -0.297 | -0.216 |
| Constant | -2.618 | -0.501 | -4.369 |
| Observations | 4,290 | 4,290 | 4,280 |
| R-squared | 0.571 | 0.453 | 0.561 |

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

Table 13: Estimates from the pooled OLS, fixed effects and 2SLS estimation including variable *below* and interaction term *openness*below*.

| Variables | (34) OLS | (35) FE | (36) 2SLS |
|---|---|---|---|
| Openness | -0.0143** | -0.00258 | -0.0192*** |
| Openness*below | 0.0344 | 0.0162 | 0.0543*** |
| Female percent of population | 0.193 | -0.0293 | 0.174*** |
| Age 15-64 | -0.291 | -0.0257 | -0.300*** |
| Age <15 | -0.452*** | -0.0219 | -0.452*** |
| Ethnic diversity | -0.0421 | 0.913 | -0.0497 |
| Religious diversity | 0.206 | -1.088 | 0.251* |
| Linguistic diversity | 0.468 | 2.910*** | 0.469*** |
| Regional polity | 0.736*** | 0.945*** | 0.743*** |
| Colony after 1945 | 2.019* | | 2.097*** |
| 1/ Years since independence | 5.078*** | 5.085*** | 5.204*** |
| America | 1.442 | | 1.339*** |
| Europe | -1.900 | | -2.050*** |
| Africa | -0.695 | | -0.897*** |
| Pacific | 2.709 | | 2.541*** |
| Time | 0.0221 | 0.0292 | 0.0246*** |
| Below | -1.783 | -1.463 | -3.127*** |
| Constant | 21.32 | -3.340 | 23.09*** |
| Observations | 4,824 | 4,824 | 4,814 |
| R-squared | 0.545 | 0.450 | 0.542 |

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

Table 14: Estimates from the pooled OLS, fixed effects and 2SLS estimation including variable *above* and interaction term *openness\*above*.

| Variables | (37) OLS | (38) FE | (39) 2SLS |
|---|---|---|---|
| Openness | -0.0133* | -0.00400 | 0.00257 |
| Openness*above | -0.00686 | 0.00333 | -0.0102 |
| Female percent of population | 0.322 | 0.0228 | 0.405*** |
| Age 15-64 | -0.0856 | -0.0654 | -0.114* |
| Age <15 | -0.352** | -0.0792 | -0.368*** |
| Ethnic diversity | -0.0122 | 0.969 | -0.000687 |
| Religious diversity | 0.0927 | -1.225 | -0.0807 |
| Linguistic diversity | 0.369 | 3.018*** | 0.359*** |
| Regional polity | 0.723*** | 0.924*** | 0.711*** |
| Colony after 1945 | 1.554 | | 1.436*** |
| 1/Years since independence | 4.369** | 4.695*** | 3.835*** |
| America | 2.474 | | 2.315*** |
| Europe | -1.484 | | -1.483*** |
| Africa | -0.347 | | -0.364 |
| Pacific | 2.933 | | 3.068*** |
| Time | 0.00761 | 0.0218 | -0.00156 |
| Above | -1.087 | -0.784 | -0.795 |
| Constant | 0.881 | -1.041 | -1.445 |
| Observations | 4,290 | 4,290 | 4,280 |
| R-squared | 0.570 | 0.453 | 0.560 |

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

Table 15: Estimates from the pooled OLS, fixed effects and 2SLS estimation including variable *above*, *below*, interaction term *openness\*above* and *openness\*below*.

| Variables | (40) OLS | (41) FE | (42) 2SLS |
|---|---|---|---|
| Openness | -0.0181*** | -0.00681 | -0.0117* |
| Openness*above | -0.00274 | 0.00477 | 0.00422 |
| Openness*below | 0.0379* | 0.0187 | 0.0463*** |
| Female percent of population | 0.323 | 0.0240 | 0.367*** |
| Age 15-64 | -0.0822 | -0.0790 | -0.0869 |
| Age <15 | -0.337* | -0.0867 | -0.340*** |
| Ethnic diversity | -0.0500 | 0.962 | -0.0433 |
| Religious diversity | 0.161 | -1.239 | 0.0761 |
| Linguistic diversity | 0.361 | 2.967*** | 0.354*** |
| Regional polity | 0.728*** | 0.927*** | 0.714*** |
| Colony after 1945 | 1.494 | | 1.449*** |
| 1/Years since independence | 4.472** | 4.727*** | 4.194*** |
| America | 2.221 | | 2.253*** |
| Europe | -1.686 | | -1.571*** |
| Africa | -0.770 | | -0.786** |
| Pacific | 2.569 | | 2.804*** |
| Time | 0.00939 | 0.0231 | 0.00313 |
| Above | -1.466 | -0.895 | -1.827** |
| Below | -2.952* | -1.640 | -3.350*** |
| Constant | 0.798 | 0.366 | -1.169 |
| Observations | 4,290 | 4,290 | 4,280 |
| R-squared | 0.575 | 0.455 | 0.570 |

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