Universidad Nacional Autónoma de México

Facultad de Ciencias

# Econometric Models for Sovereign Credit Ratings of Emerging Markets

**REPORTE DE ESTANCIA DE INVESTIGACION**

QUE PARA OBTENER EL TÍTULO DE:

Actuario

P R E S E N T A :

Arturo García Ramos

TUTORES

Maxym Chaban, PhD / Act. Edgar Diaz Ordoñez

Ciudad Universitaria, Cd. Mx., 2019

1. Datos del Alumno

García Ramos Arturo

55 55 07 36 85

Universidad Nacional Autónoma de México

Facultad de Ciencias

Actuaría

312275693

2. Datos del tutor 1

Act Díaz Ordóñez Edgar

3. Datos del tutor 2

Dr Maxym Chaban (Universidad de Saskatchewan

4. Datos del Sinodal 1

Mat Rodrigo Quijón Hipólito

5. Datos del Sinodal 2

Fis Jimmy Hernández Morales

6. Datos del Sinodal 3

M. en E. Leslie Alejandra Jiménez Rosas

7. Datos del Sinodal 4

Act Monse Esquivel López

8. Datos del Trabajo Escrito

Econometric Models for Sovereign Credit Ratings of Emerging Markets

| | |
|---|---|
| Maxym Chaban, Ph.D.<br>Department of Economics<br>University of Saskatchewan | Act. Edgar Diaz Ordoñez<br>Facultad de Ciencias<br>UNAM |

*"It is not about winning or losing, but at the end of the day, it is all about winning or losing."*

*Garry Kasparov*

# AGRADECIMIENTOS

A mi madre **Laura**, y a mi padre **Alberto**, por su amor incondicional, por su sabiduría, por impulsarme a ser una mejor persona, por todas las oportunidades que me han dado, y por su ejemplo. Todos mis éxitos, pasados, presentes y futuros son fruto de su trabajo y esfuerzo como padres. Todo esto, y lo que viene, siempre será para ustedes. Pá, má, los amo.

A mi hermano **Alberto** por ser la mejor compañía y por tenerme paciencia, me siento orgulloso del hombre en el que te estás convirtiendo.

A **Alondra**, por la Tabla 1.1, por su cariño, por acompañarme a pesar de la distancia, por ser un chaparrín y por todo lo vivido en tan poco tiempo.

A mi tía **Alma**, **Lupe** e **Irma** por ser como mi segundas madres, por cuidarme, aconsejarme, procurarme y quererme como a uno de los suyos.

A el profesor **Maxym Chaban** y a la **Universidad de Saskatchewan**, por darme la oportunidad de una vida, por apoyarme, abrirme las puertas a un nivel educativo de primer nivel y dejar que terminara mi carrera universitaria con broche de oro.

A **Edgar Díaz** por ser mi mentor,y brindarme sus consejos, disposición, atención e impulasrme a ampliar mis metas y a mis sinodales **Leslie, Monse, Jimmy y Rodrigo**, por acceder a trabajar conmigo a pesar del aviso tan pronto. Todos son jóvenes catedráticos, símbolo del altísimo nivel académico que goza la Facultad de Ciencias. Las futuras generaciones serán afortunadas de preparase con ustedes y fue un privilegio haber podido aprender de cada uno.

Y por último, a la **Universidad Nacional Autónoma de México**, por enseñarme lo que significa defender los colores, por darme el futbol americano que tanto contribuyó en mi formación de caracter, por permitirme estudiar en las aulas de la gloriosa Prepa 6 y la heróica Facultad de Ciencias, por tener el mejor nivel académico de México, por mandarme a titular a Saskatchewan, por CU, por la Biblioteca Central, por los Pumas (del americano), por ser la Máxima Casa de Estudios, y por ser la institución más preciada de una país que no te merece. Con orgullo, mi sangre será azul y mi piel dorada, hasta el útlimo día.

# MÉXICO... PUMAS... UNIVERSIDAD... ¡GOYA!

# ACKNOWLEDGMENTS

To my mother **Laura**, and my father **Alberto**, for your unconditional love, their wisdom, for urging me to be a better person, for all the opportunities you have given me and for your example. All my successes, past, present and future are the result of your work and effort as parents. All this, and what is coming, will always be for you. Mom, Dad, I love you.

To my brother **Alberto** for being the best company and for being patient with me, I am proud of the man you are becoming.

To **Alondra**, for Table 1.1, for your love, for accompanying me despite the distance, for being a shortie and for everything we have lived in such a short time.

To my aunt **Alma**, **Lupe** and **Irma** for being like my second mothers, for taking care of me, advising me, procuring and loving me as one of their own.

To professor **Maxym Chaban** and to the **University of Saskatchewan**, for giving me the opportunity of a lifetime, for supporting me, opening for me the doors to a first level educational level and allowing to finish my undergraduate studies in great manner. I'll will be grateful forever.

To **Edgar Díaz** for being my mentor, and to give me his advice, willingness, attention and impel me to expand my goals and to my synodals **Leslie, Monse, Jimmy and Rodrigo**, for agreeing to work with me despite the soon warning. All are young professors, a symbol of the highest academic level enjoyed by the Faculty of Science. Future generations will be fortunate to prepare with you and it was a privilege to have learned from each one of you.

And finally, to the **National Autonomous University of Mexico**, for teaching me what it means to defend colors, for giving me football that contributed so much to my character formation, for allowing me to study in the classrooms of the glorious Prepa 6 and the heroic Faculty of Sciences, for having the best academic level in Mexico, for sending me to Saskatchewan, for CU, for the Central Library, for the Pumas (of the American), for being the Highest House of Studies, and for being the most precious institution in a country that does not deserve you. With pride, my blood will be blue and my skin golden, until the last day.

**MÉXICO... PUMAS... UNIVERSIDAD... ¡GOYA!**

# PREAMBLE

This report was written during a visiting research stay done during May and August of 2019 in the University of Saskatchewan under the supervision of Dr. Maxym Chaban of the Department of Economics and Act. Edgar Díaz Ordoñez of the Faculty of Science, UNAM. Our work intends to understand which are the factors and weights that determine sovereign ratings of emerging markets economies through two econometric approaches, linear regression for panel data and ordinal logistic regression. Sovereign ratings are opinions on the issuers capacity and willingness, in this case a country, to meet its financial commitments as they come due. We use economic, external and fiscal indicators and include variables that have not been analyzed previously, like the Exchange Regime and the General Government Gross Debt. In the same spirit we use a larger panel in contrast with existing literature. A larger panel provides more robustness and allow us to offset the model's sensibility to the unspecified changes in the methodology of rating agencies to asses sovereign ratings. We focus on the estimation process and we explain the theoretical framework for each model. We compare the various alternatives and do statistical tests to validate which model is more appropriate and why. We analyze the results by identifying the significant macroeconomic variables based on the level of significance and the Akaiken Information Criterion (AIC) to understand which are the ones that influence the regression the most. We will identify the differences among both techniques and identify if results vary much from one to another.

**Chapter I** gives a brief introduction into what are sovereign ratings, the role rating agencies have in the modern financial ecosystem , and a literature review. **Chapter II** explains the sample selected and the intuitive relationships between the selected explanatory variables and the ratings. **Chapter III** covers the theoretical framework and results for the linear regression model, using OLS for the pooled sample, Fixed Effects and Random Effects. **Chapter IV** covers the same regression approach, but now treating the response variable as a discrete ordered variable. We use the proportional odds model for the pooled sample, and utilise as well Fixed and Random Effects. **Chapter V** makes the final analysis and comparison between the two approaches, we identify the most significant economic variables and establish further work. All figures, tables, definitions and references are displayed in bold typography.

# INDEX

# CHAPTER I

# SOVEREIGN CREDIT RATINGS

## 1.1 What are Credit Ratings

Credit rating agencies have become an essential component in the contemporary financial system. Their opinions are for investors as vital as the transmission to a motor vehicle. Credit ratings are opinions about **credit risk**. They reflect the ability of an issuer, such as a corporation, a state or local government to meet its credit obligations in time and form. Issued by rating agencies, each one has its own methodology that incorporate financial, economic and management or governance factors in measuring creditworthiness.

Investors and other market participants may use the ratings as a screening device to match the relative credit risk of an issuer or individual debt issue with their own risk tolerance or credit risk guidelines in making investment and business decisions. Agencies communicate their opinions based on a letter based scale, where the highest rating is commonly regarded as 'AAA', and 'D' or 'SD' as the poorest grade. **Table (1.1)** shows the rating scale system Standard and Poors uses for sovereigns and corporate issues. Rating agencies may change their rating on an issuer or issue if significant changes occur and may be broadly related to overall shifts in the economy or business environment or more narrowly focused on circumstances affecting a specific industry, entity, or individual debt issue.

**Definition 1.1.1. Sovereign Credit Ratings** : Sovereign credit ratings are opinions on the issuer's (nation, country or independent state) capacity and willingness to meet its financial commitments as they come due. In contrast with corporate ratings, the rating agencies include macroeconomic, financial and governance factors to evaluate the sovereign's performance and measure the risk of the country defaulting or not. Governments generally seek credit ratings to ease their own access (and the access of other issuers domiciled within their borders) to international capital markets, where many investors, particularly U.S. investors, prefer rated securities over unrated securities of apparently similar credit risk.

The number of countries that have been graded has grown remarkably in the past 20 years. In 1995, forty nine countries had been graded, mostly strong and developed nations from Europe, North America and Asia. According to Standard & Poor's Global Ratings [1] the agency maintained 132 sovereign ratings, 68% of the world's countries. For the purpose of

[1] 2017 Annual Sovereign Default Study And Rating Transitions. `https://www.spratings.com/documents/20184/774196/2017+Annual+Sovereign+Default+Study+And+Rating+Transitions.pdf`

***Table 1.1:*** *Standard and Poor's Sovereign rating scale*

| | Rating | Definition |
|---|---|---|
| INVESTMENT GRADE | AAA | Highest credit rating assigned by S & P. |
| | AA | Very strong capacity to meet its financial commitments. It differs from the highest-rated obligors only to a small degree. |
| | A | Strong capacity of payment but is somewhat susceptible to the adverse effects of changes in circumstances and economic conditions. |
| | BBB | Adequate capacity to meet its financial commitments, however, adverse economic conditions more likely to weaken its credit profile |
| SPECULATIVE GRADE | BB | Sovereign faces major ongoing uncertainties and exposure to adverse business, financial, or economic conditions could lead to the obligors inadequate capacity to meet its financial commitments. |
| | B | An obligor rated B is more vulnerable than the obligors rated BB, but the obligor currently has the capacity to meet its financial commitments. |
| | CCC | Dependent upon favourable business, financial, and economic conditions to meet obligations. |
| | CC | Highly vulnerable. The CC rating is used when a default has not yet occurred but default is a virtual certainty, regardless of the anticipated time to default. |
| | SD and D | A 'D' rating is assigned when S&P believes that the default will be a general default and that the obligor will fail to pay all or substantially all of its obligations as they come due. An SD (selective default) rating is assigned when the obligor has selectively defaulted on a specific issue or class of obligations but it will continue to meet its payment obligations on others in a timely manner. |

Additional (+) or (-) are allocated from the 'AA' rating and below to show relative standing, with 'A+' being better than 'A' or 'A-'

this study, we will focus on long-term bonds issued on Foreign Currency. Since we are focusing on developing economies, which generally seek for foreign capitals for financing structural projects, long-term debt is usually the benchmark in international markets, and it is also the reference in the literature.

## 1.2 Rating Agencies

Sovereign Ratings are issued by certified institutions which specialize in analysing the financial and economic prospects as well as the terms and conditions of the debt security of a sovereign or corporate entity. This institutions are known as **Credit Rating Agencies** (CRAs). The issuer shares with the rating agency information and historical data, which sometimes it is not entirely public, and in exchange for a fee, the rating agency will evaluate and measure the risk of default on the issuers obligation. After the initial assessment, the agency will carry out continuous surveillance and if needed will make corrections to the rating, based on new economic developments or structural changes on the issuer's management or composition.

The main rating agencies are

- Moody's (Est. 1909)
- Standard and Poor's (Est. 1906)
- Fitch (Est. 1914)

Colloquially known as "The Big 3", the New York-based corporations monopolize almost the entire market. They have been surrounded by controversy [11], [12], [18], [32] regarding their performance in financial crisis, conflicts of interest accusation since they became publicly traded companies and their role in general in the modern financial ecosystem.

Rating Agencies are criticized for reacting to events rather than anticipating them [35]. One example is Mexico's economic crisis of 1994-1995. The December devaluation of the peso rocked hard the financial markets, but S& P had a rating of 'BB+' with a positive outlook on the sovereign's note. CRAs played a crucial role in the Housing crisis in 2007-2008. All the major agencies failed to estimate the default risk on mortgage-backed securities (MBS) and collateralized debt obligations (CDO). The 2007 crisis is considered the worst financial cataclysm since the Great Depression. After the dust had settled, people started wondering how the global economic came so close to an imminent collapse, which had disastrous consequences practically in all the world, even a decade later.

Ana K. Barnett-Hart, a Harvard undergraduate student made a recompilation of the events and analysis of the securities that triggered the catastrophe, and noted that the failure of the rating agencies to detect the imminent default of CDO contributed largely to the "snow ball" effect that whipped out one trillion US dollars from financial markets. The most disturbing and comprehensible fact was that most of the securities marked down where given a 'AAA' rating, the highest possible, and it remained like that for days as the meltdown started to unravel. Investors came to rely almost exclusively on ratings to assess CDO investments. Moody's and Standard and Poor's accounted for almost 90% percent of the CDO deals, while Fitch's role was substantially lower. The combined revenue of the "Big Three" agencies doubled from 2002 from less than 3 billion USD to more than 6 billion in 2007 prior to the collapse [7] p. 17.

*" Problems with CDO ratings rapidly developed as the rating agencies came under enormous pressure to quickly crank out CDO ratings and the market exploded faster than the number of knowledge of analysts. "*

Her work was groundbreaking and earned her an award for the best undergraduate paper at Harvard that year. The thesis explained everything with so much detail and clarity, it helped lawmakers to comprehend what precisely had happened, and was used in Congressional Hearings by the US Senate [11] when interrogating the major banks and rating agencies in 2008. It also served as inspiration for Michael Lewis's book "The Big Short", which narrates how just a few hedge fund managers bet on the collapse of the housing market, and how almost the entirety of the market was unaware of the looming catastrophe. Lewis personally credited the student for her work, and admitted that if it wasn't for her, not even he could have comprehended entirely the events and factors that lead to the debacle. On a curious note, Barnett-Hart went after graduation to work for Morgan Stanley and later to Goldman Sachs, two of the investment banks that where implicated as responsible for the collapse and came under public scrutiny precisely because of her work.

In the aftermath of the crisis, the rating agencies came under fire again when in 2010 and 2011 the agencies relegated Greece, Portugal and Ireland to the Speculative Grade, an action which EU officials say accelerated a burgeoning European sovereign-debt crisis [18]. In January 2012, amid continued eurozone instability, S& P downgraded France and Austria from the "excellence grade ", stripping them from the 'AAA' rating. The same thing happened later that year, when Standard and Poor's downgraded the U.S. sovereign rating from 'AAA' to 'AA+' alleging large amounts of debt and an unbalanced budget. The downgrade brought hard criticism over the agency once again, since many officials claimed it was a retaliation for the recent scrutiny CRAs came under because of the 2007-2008 fiasco. Immediate decline on stock and fixed income markets followed, with the major U.S. stocks index falling 5%-7%, and the U.S. Treasury bonds rose (paradoxically since it was the subject of the downgrade), signalling investors flight to safe assets. deHaan showed that the fixed income market participants reduced their use of credit rating agencies after the crisis, and that the performance of the ratings improved during and after the crisis, consistent with the CRAs positively responding to public criticism and regulatory pressures [12]. One of the main issues with the ratings, beside of ethical arguments or conflict of interests, is the fact that the models and criteria are not publicly known. CRAs don't communicate explicitly how their models work, how they are calibrated or what factors are the ones that matter the most. They communicate the change of ratings with some insight and commentaries on the reasoning preceding the action, yet the criteria is not completely clear. Since market participants don't rely blindly on credit ratings like before, and the lack of not knowing what the rating agencies value more when taking an action is always present, it is of interest for investors and governments to identify these unknown factors to avoid short term impact if a downgrade or negative outlook on their sovereign notes takes place, and to lower costs of financing if the grade gets better.

For better or for worse, Rating Agencies are still well cemented as key players in financial markets, and ratings are sometimes the most pragmatic way to understand and compare opportunity costs and risks for foreign investors . On the contrary position, without a rating, it is practically impossible for an issuer to be taken seriously or to attract external capitals. Their work saves millions in information costs and resources for investors, and although their reputation has been damaged, their performance may have improved as an effort of regaining the market's confidence. Developing economies however, are more dependent on the CRAs opinion if they want to have a broader access into international capital markets and to catch the attention of institutional investors.

## 1.3 Emerging Markets: In Search for Happiness (or Investment Grade)

**Emerging markets** are characterized primarily by the high degree of volatility and their transitional character, with transitions occurring in economic, political, social and demographic dimensions. EM's capability to attract private capital flows and to integrate themselves into the globalized financial markets depends at some degree on the ratings given by the agencies [31]. High ratings help mitigate borrowing costs and enhance economic growth, at least in theory. Institutional investors still rely heavily on the agencies opinion when it comes to explore attractive returns in the fixed income market, specially if the risk tolerance for the seeker is bounded. As an example, pension funds around the globe follow strict rules, given the nature of their operations, that limits their investments of credit instruments solely to assets that have been graded by rating agencies. The Mexican pension funds (AFORES) are responsible for approximately 200 billion US dollars of worker's savings as of June 2019. The National Commission for Retirement Savings (CONSAR) stipulate on their regulation that each fund can only acquire debt instruments that have been previously rated by at least two rating agencies, and that it has to oblige with certain limit thresholds on their portfolio depending on the rating of the security itself **consar**.

The need of being able to replicate the models or at least identify the indicators used by the agencies is of big importance to developing economies, since these countries want increase their role in international financial markets, but not at exorbitant costs. Here's where the **Investment Grade** comes in place. The term "investment-grade" historically referred to bonds and other debt securities that bank regulators and market participants viewed as suitable investments for financial institutions. Now the term is broadly used to describe issuers and issues with relatively high levels of creditworthiness and credit quality. In contrast, the term "non-investment-grade," or "speculative-grade," generally refers to debt securities where the issuer currently has the ability to repay but faces significant uncertainties, such as adverse business or financial circumstances that could affect credit risk. In S&P long-term rating scale, issuers and debt issues that receive a rating of 'BBB-' or above are generally considered by regulators and market participants to be investment grade, while those that receive a rating lower than 'BBB-' are generally considered to be speculative grade. [2]

Empirical work by Jaramillo and Tejada show that transitioning to a Investment Grade rating is more valuable than any other movement among the scale. Econometric results indicate that reaching investment grade lowers sovereign spreads by **36%**, compared to a 5-10 % reduction in spreads following rating upgrades within the investment grade classes, and no impact for movements within the speculative grade asset class, *ceteris parabus* [26]. Translated into basis points, this implies that the spreads of a 'BBB-' rated country woul be a 160 basis points lower than those of a 'BB+' rated country with spreads of 440 points. Developing economies get a huge break when achieving the covet Investment Grade, but are under additional pressure to maintain there sovereign ratings there, since a downgrade can lead to meaningful adverse impact in spreads, exchange rate and local issuers. Markets rely on the information provided by rating agencies, but are closely monitoring other indicators looking for possible delays in rating changes. That implies that investment grade countries have a little more leeway, but will nonetheless be punished some variables or factors changes dramatically.

## 1.4 Literature Review

Various studies have tried to identify the factors and weights CRAs model to assign sovereign ratings.Cantor and Packer [9] is one of the earliest works attempting to asses the relevance of six macroeconomic variables using ratings by Standard & Poor's and Moody's. This study is a cross-section analysis of 49 countries, both developed and emerging economies, utilizing linear regression to identify their significance. The variables were per capita income, GDP growth, inflation, external debt, level of economic development and default history. They found that sovereign yields tend to rise when ratings worsen, reflecting the rise in default premium . Reisen, Von Maltzan, and Larraín analysed the relationship between bond spreads and ratings, and how one influence the other and vice versa using simultaneous equations. Similar to the work done by Jaramillo and Tejada, which analyses the impact of having the investment grade as previously mentioned, both works found that the level of external debt is significantly more important to developing countries. As well, the growth rate plays a central role [26], [35].

Afonso was one of the pioneers in the use of panel data and logistic and exponential transformations for the ratings. Empirical results show that negative actions increase negative impacts on international markets, and boosts downhill cycles for the sovereign in the short term. Also he found that the economic factors in determining the sovereign ratings are distinct in developing and developed countries. They also propose using exponential and logarithmic scale for the ratings as an alternative to the classic linear scale . The more relevant economic variables in his study are GDP per Capita, External Debt to Exports, level of economic development, default history, real growth rate and inflation rate [2]. Gültekin-Karaka, Hisarcıklılar, and Öztürk used Moody's data for 93 countries from 1999 to 2010 using ordered probit with two random effects model, one for developed and one for undeveloped economies. Their results show that rating agencies are biased toward advanced economies, and lower ratings go to low income countries. Their model incorporates economic and external indicators, but also include dummy variables for the level of income and if the country is an OECD member or not [20].

Lastly, in one of the more recent works, Erdem and Varli use panel data OLS with AR(1) disturbances and ordered probit with random effects. Using quarterly data ranging from 2002 to 2011 for 8 developing economies, they found that the significant indicators are GDP per Capita , Budget Balance / GDP, Governance Indicators and Reserves. They found an accuracy rate for predictions of 93% within a four notch interval [14].

[2] Guide To Credit Rating Essentials. https://www.spratings.com/documents/20184/760102/SPRS_Understanding-Ratings_GRE.pdf/298e606f-ce5b-4ece-9076-66810cd9b6aa

# CHAPTER II

# DATA AND SAMPLE SELECTION

***Table 2.1:*** *Emerging-Markets Economies: Ratings and Region*

| Country | Region | Highest Rating | Lowest Rating |
|---|---|---|---|
| Argentina | Lat Am | BB- | SD |
| Brazil | Lat Am | BBB- | B |
| Colombia | Lat Am | BBB- | BB |
| Mexico | Lat Am | BBB+ | BB |
| Venezuela | Lat Am | BB- | SD |
| Egypt | ME / AF | BBB- | B |
| South Africa | ME / AF | BBB+ | BB |
| Turkey | ME / AF | BBB | B- |
| Russia | CEE | BBB+ | SD |
| China | Asia Pacific | AA- | BBB |
| India | Asia Pacific | BB- | BB+ |
| Indonesia | Asia Pacific | BBB- | CCC+ |
| Malaysia | Asia Pacific | A- | BBB- |
| Pakistan | Asia Pacific | B+ | CCC+ |
| Phillipines | Asia Pacific | BBB | BB- |
| Thailand | Asia Pacific | BBB+ | BBB- |

The sample that will be used to specify the models, consists of 17 annual observations from 2000 to 2016 of 16 countries cataloged as "Emerging-Markets Economies" by Standard & Poor's with the indicative ratings acting as response variables and 9 macroeconomic fundamentals as explanatory variables. The sovereigns subject to the study are:

Each country belongs to one of the following regions:

- Latin America with 5 countries (Lat AM)
- Middle Eats and Africa with 3 countries ( ME / AF)
- Central and Eastern Europe with 1 countries(CEE)
- Asia Pacific with 7 countries

There is no strong link among the seventeen sovereigns, but many of them share particular characteristics. For example, they tend to have more modest wealth and weaker institutional frameworks than more advanced economies [38]. We will use sovereign ratings from Standard and Poor's since it is more flexible and easier to obtain historical series than the other major rating agencies. Cantor and Packer [10] found that there is some discrepancies among the ratings of Moody's and S&P especially on low end ratings, suggesting that such disagreements owe much to the subjectivity of many aspects of sovereign risk measurement and the relative youth of the sovereign rating business. Nonetheless, this paper was published in 1995 and the number of sovereigns has almost tripled since then.

**Definition 2.0.1. Panel Data**: Panel data consists of observations of individual $i = 1, \dots, N$ along time $t = 1, \dots, T$ on the response and explanatory variables. This format combines the times series with cross-sectional structure, providing more information about the individuals, but also presenting high time correlation. If the regression consist of $K$ regressors, then the

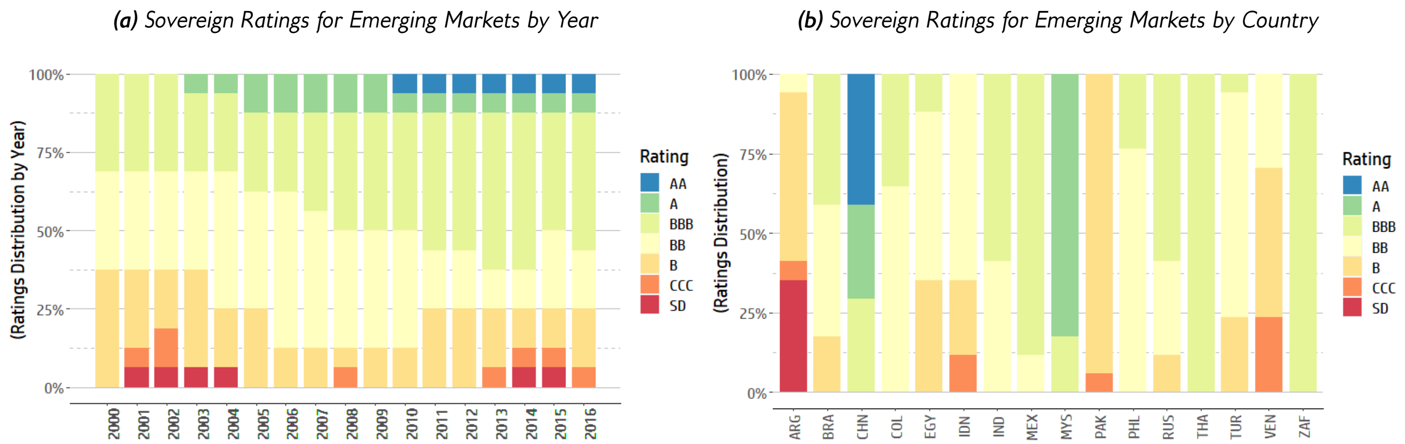



***Figure 2.1:*** *Sovereign Ratings Distribution for the Selected Sample*

matrix of covariables $X$ will be of dimension $NT \times K$. This structure is also referred as longitudinal data. We will model the sovereigns ratings through a selection of 9 variables that we believe are considered in the models of the rating agencies and that have been previously used in the literature [9], [14], [35]. Early work consisted in analyzing with cross-sectional observations. Afonso [2] was one of the pioneers in the use of panel data and logistic and exponential transformations for the ratings.

The main difference between those studies and this one is that we propose a large panel data, considering annual observations of 17 years, for 16 countries. We believe this will provide with more robustness and degrees of freedom when it comes to obtain consistent estimates. We also propose different macroeconomic variables, such as the Exchange regime, the General Government's gross debt and the International Reserves. There are no observations missing for either the macroeconomic variables or the ratings, hence we have what is considered a **Balanced Panel**. Originally, the following countries were considered to be part of the sample: Saudi Arabia, Qatar, Poland and Hungary, however there is no public available information on some external indicators, so they were excluded since they couldn't be integrated into the models.

## 2.1 Response Variable

Rating agencies allocate complementary symbols to the rating, (+) and (-) in the case of S&P, to distinguish economies among others with the same grade, which provides a more refined credit profile of the sovereign. Although in most of the literature, the rating is taken with the additional classifier, for the purpose of this study, the ratings will be considered only at the indicative level. The reasoning behind this is that, according to the public methodology S&P uses to rate sovereigns[1], there's in most cases, an additional adjustment, based mostly of qualitative attributes specific to each nation, that leads to a final rating one or two notches above, or below, the initial assessment. The ratings interpretation for this work will be built on exclusively of quantitative public information, such as economic, external and fiscal indicators, so restricting the levels of the ratings, is expected to provide more precision to the models' results.

**Definition 2.1.1. Response Variable** We will consider as the response variable $y_{it}$ the sovereign rating the $i$-th country had at $t$ year's end, so movements amidst the annual interval are not taken into account.

[1]`https://www.standardandpoors.com/en_US/web/guest/article/-/view/sourceId/8950072`

The distributions of the ratings in the sample can be appreciated in **Figure (2.1)**. We can observe the change in the proportions of the ratings in the sample as time went by in **Figure (2.1a)** The 'AAA' rating is omitted since no EM country has achieved it. From 2000 to 2008, the sovereign's majority did not have the **Investment Grade**, ('BBB' or higher), but the proportion kept increasing every year, except from 2014 to 2015. The year with the most lower ratings is 2002, with Indonesia and Venezuela rated 'CCC' and Argentina on Default, while the span from 2005 to 2013 is the most stable, with only two 'CCC' ratings (Pakistan in 2008 and Argentina in 2013) and no Default.

In contrast, **Figure (2.1b)** shows changes in the sovereign ratings by country. There has not been a 'AAA ' sovereign rating since 1982 (Venezuela). The highest rating achieved during this span of time is 'AA-', by China, while 3 countries fell in SD: Argentina, Russia and Venezuela, with Argentina defaulting twice (2001, 2014). Argentina spent 6 years with the 'SD' rating, the most among the sample, while China is ranked as the economy with the best grade for the most time, with a rating of 'AA' for 7 years. Thailand and South Africa stayed with the same grade for the whole interval, whereas Argentina tested four different ratings and Indonesia, Russia and Venezuela tested three distinct categories. Twelve countries acquired an **Investment Grade** rating at least once, Argentina, Indonesia, Pakistan and Venezuela being the exceptions. The 'CCC'/ 'CC' rating denotes imminent default risk in the short term, and S&P has a special methodology for this cases [2], so it seems proper to concatenate all these observations, along with the default rating and incorporate them into one category.

**Definition 2.1.2. Transition Probability** The one-step transition probability is the probability of transitioning from one state to another in a single step,

$$p_{ij} = \mathbb{P}[X_n = j \,|X_{n-1} = i]$$

[2] https://www.standardandpoors.com/en_US/web/guest/article/-/view/sourceId/7554329

***Table 2.2:*** *Transition Probabilities for the Ratings*

| Rating | SD | B | BB | BBB | A | AA |
|---|---|---|---|---|---|---|
| AA | | | | | | 100 |
| A | | | | | 94.44 | 5.56 |
| BBB | | | 3.33 | 94.44 | 2.22 | |
| BB | 1.25 | 3.75 | 86.25 | 8.75 | | |
| B | 10.20 | 77.55 | 12.24 | | | |
| SD | 61.54 | 38.46 | | | | |
| Total | 5.15 | 19.12 | 30.51 | 35.66 | 6.99 | 2.57 |

**Definition 2.1.3. Absorbing State** : A state $i$ is called absorbing if it is impossible to leave this state. Therefore, the state i is absorbing if and only if

$$p_{ii} = 1 \quad \text{and} \quad p_{ij} = 0 \quad \forall i \neq j$$

Transition probabilities are displayed in **Table (2.2)**. The overall frequencies are at the bottom row, which indicates how many observations for each category the sample has. If the country has the 'B' rating, it can either go under or move to the next level with almost the same probability. If the observation comes from the 'BB' category, then there are higher chances of improving to the Investment Grade with a 'BBB' rating, than falling back. The latter is also the most frequent rating among the Emerging-Markets economies. The 'AA' rating has a perfect probability of staying there, which is unrealistic, since it could be interpreted as once a sovereign gets that rating it will stay there indefinitely. This is caused by the sample and the fact that only a handful of developing economies in the world have high ratings on their sovereigns, i.e., it is **not an absorbing state**. That's why we considered including Qatar, Saudi Arabia and Poland, which are still relatively small economies compared to their peers in their regions, but have strong economic prospects. Including them would have bring more observations to the high end categories, however since they fail to communicate some data to the public, it makes not sense to consider them since those observations would be dropped out anyway.

## 2.2 Explanatory Variables

As previously noted, rating agencies considers several economic, political, social and monetary factors that lead to a sovereign rating. However, some of these aspects are not quantifiable and are based purely on a very particular subjective analysis of each economy. Some studies [2], [14] incorporate governance factors into the modeling, however we will focus exclusively into quantitative factors, which in certain way reflect as well part of the social status the country is going through at the moment. Despite this, each rating incorporates historical information on the nation's economic and financial performance, and some of these indicators are of public domain.

Therefore, a sample of nine quantitative factors was selected as explanatory variables, hoping to identify and prove their weights and significance when it comes to assigning a rating to a sovereign. These indicators are cataloged regarding their economic, fiscal and external nature, as reported by S&P when announcing a rating revision. The name and units of each variable are shown on **Table (2.3)**. This indicators can be classified as Economic Indicators, External and Indicators and Fiscal Indicators. **Figures (2.2), (2.3) and (2.4)** display the distribution of each macroeconomic variable by country, sorted out by regions as in **Table (2.1)** . The following description explains the variables to be used for the econometric models, as well serves as a justification of the intuitive relationship between the indicator and the capacity of each country to fulfil its credit obligations. All the historical series were obtained through the International Monetary Fund's[3] **World Economic Outlook** (April 2019) and the **World Development Indicators** database from the World Bank[4] The Exchange Regime, series were obtained from a working paper by Ilzetzki, Reinhart, and Rogoff [25]. Their work provides a compilation of exchange rate regimes for 194 countries from 1946 to 2016 and studies the evolution of monetary policies in the history of modern economy.

### 2.2.1 Economic Indicators

A wealthy, diversified, resilient and adaptable economy suggests a better debt-bearing ability, and a stronger and efficient public policy performance leads to a potential higher rating [38].

- **GDP growth (GDPG)** Gross Domestic Product annual rate of change (%)GDP is the sum of gross value added by all resident producers in the economy. It is commonly used to evaluate economic growth and dynamism. A high rate of growth would suggest a greater long-term cushion to face obligations and provide a better credit profile for the country.

- **GDP per Capita (GDPPC)** Gross Domestic Product in current US dollar divided by midyear population. It is the most relevant measure the income of a nation. With higher revenues, the country has a better tax collection potential and improve its ability to pay its compromises.

- **Inflation (INF)** Consumer Price Index annual percentage change (%). A relatively high inflationary rate could suggest structural problems or inefficient monetary policy. High inflation almost surely creates significant social discontent and this could lead to drastic political changes, which would generate uncertainty and put additional pressure on the sovereign's credit outlook.

### 2.2.2 External Indicators

One of the crucial aspects to be considered in this study, is the external position and relationship each sovereign has with respect to the rest of the world. As developing economies, each country aspires to access international markets and increase

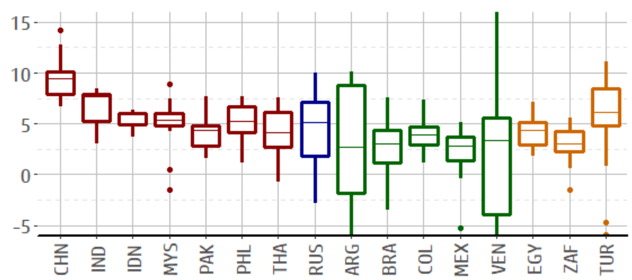
[3] `https://www.imf.org/external/pubs/ft/weo/2019/01/weodata/index.aspx`

[4] `https://data.worldbank.org`

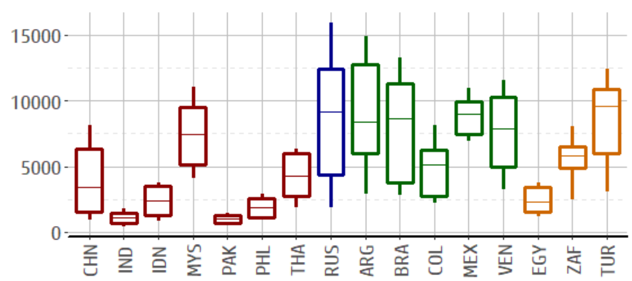
***Table 2.3:*** *Summary of the Explanatory Varaibles*

| Variable Name | Unit of Measure | Source |
|---|---|---|
| GDP per Capita | Current US Dollars | IMF, World Bank |
| GDP Growth | Annual real growth rate (%) | IMF, World Bank |
| Inflation | Annual change rate (%) | IMF, World Bank |
| Current Account Balance | (%) of GDP | IMF, World Bank |
| External Debt | (%) of exports | IMF, World Bank |
| Reserves | (%) of external debt stocks | IMF, World Bank |
| Net Lendind/ Borrowing | (%) of GDP | IMF, World Bank |
| Gross Debt | (%) of GDP | IMF, World Bank |
| Exchange Regime | Dummy variable (0 for Floating, 1 for Fixed) | Ilzetzki , Reinhart & Ragoff |

***(a)*** *GDP Growth*




***(b)*** *GDPP per CAPITA*




***(c)*** *inflation*

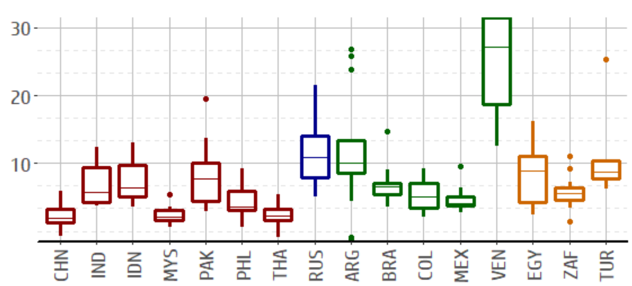



***Figure 2.2:*** *Economic Indicators*

dynamism and flows into their coffers, leading to a greater capability of improving the quality of life of its governed. A higher rating could mitigate costs of debt and enhance sovereign engagement with the global economy.

- **Current Account Balance(CAB)**The current account as a % of GDP is the record of all transactions in the balance of payments covering the exports and imports of goods and services, payments of income, and current transfers between residents of a country and non-residents. Nations with chronic current account deficits often come under increased investor scrutiny during periods of heightened uncertainty, which pressures the nations availability to repay its obligations.

- **External Debt to Exports(EDE)** Total external debt stocks as a percentage of exports of goods, services and primary income. Total external debt stocks is debt owed to non-residents repayable in currency, goods, or services. It is the sum of public, publicly guaranteed, and private non-guaranteed long-term debt, use of IMF credit, and short-term debt. Higher than average ratios could point out the country is accumulating more debt that it can cover with its receipts or indicate that the economy is not receiving enough to justify the level of indebtedness, significantly reducing the sovereign's capability to meet its obligations at all. This is one of the key indicators, given that the subject of study is the rating of a country's capacity to honour its responsibilities to foreign investors.

- **Total Reserves to External Debt(RES)** International reserves as a percentage of External Debt Stocks. A country's capability of obliging it's responsibilities is naturally correlated to its liquidity capacity and the foreign ex-

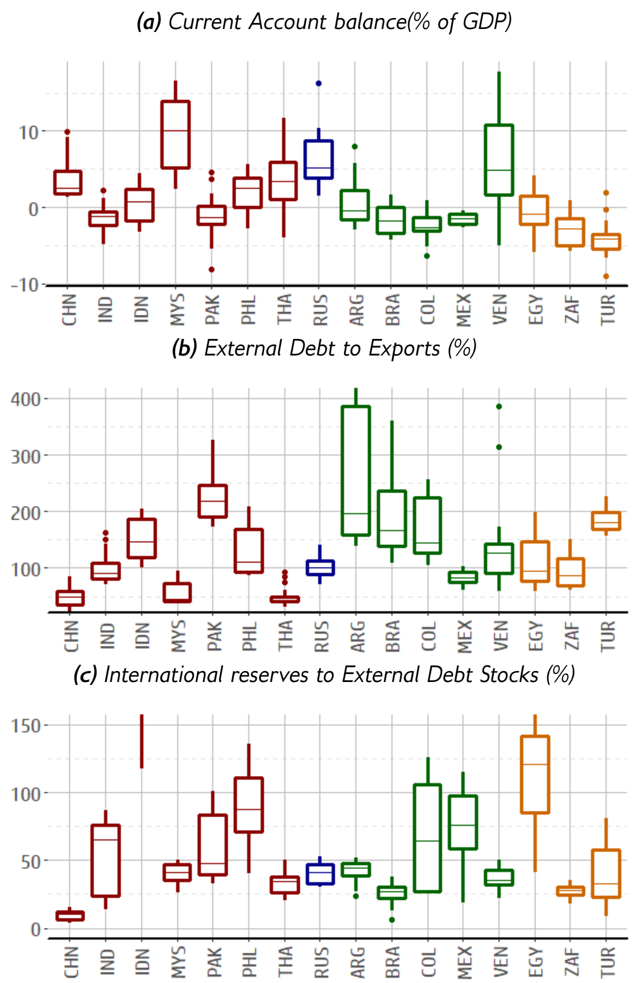



***Figure 2.3:*** *External Indicators*

change reserves it has saved. Relative to the amount of external debt stocks, it might reflect if the nation is being prudent in the amount of debt it is acquiring and how is this money being spent, or if the authorities are being responsible and allocating enough foreign resources in case of a contingency

### 2.2.3 Fiscal Indicators

An efficient and flexible fiscal performance, as well as a reasonable debt burden, reduces the country's default risk. Economic growth shouldn't be the only factor to consider, but also the efficient, moderate and responsible administration of wealth and risks.

- **Budget Balance(BUD)**Defined as the difference between revenue and total expenditure of the General Government as a (%) of GDP. This balance may be viewed as an indicator of the financial impact of general government activity on the rest of the economy. A large deficit absorbs private domestic savings and suggests that a government lacks the ability or will to tax its citizenry to cover current expenses or to service its debt.

- **General Government Gross Debt (GD)** Gross debt as a (%) of GDP consists of all the General Government's liabilities that require payment or payments of interest and/or principal by the debtor to the creditor at a date or dates in the future. A higher level of debt may be positively correlated to a higher risk of default. An economy is more resilient to downward economic cycles if it is exposed to a lesser degree of debt than its peers.

- **Exchange Regime: Floating or Fixed(FIX)** One of the most important tasks for any given Central Bank is to determine the Exchange Regime for its local currency. Although decisions on the currency are taken by the Central Bank of each country, we will consider this variables as a Fiscal Indicator instead of a Monetary Indicator just for simplicity. This decision impacts international debt, reserves and inflation directly. It is also always taken into consideration when studying the Central Bank's policies and its credibility. Floating regimes give monetary institutions more autonomy and a greater flexibility to guide the national economy. However, it conveys exposure to high volatility periods that can damage its prospects. Fixed exchange rates intend to control more effectively inflation and sustain less impact in imports / exports because of currency movements. If not followed with prudent monetary and fiscal policies, it can lead to a loss of credibility and cause uncertainty, affecting the Central Bank's prospects anyhow. For the sake of this study, the Exchange Regime will be modelled with a dummy variable.

Free Floating and Managed-Floating regimes will be considered as a *Floating Regime*. Any kind of Crawling Pegs, Pegged within a Band or Crawling Band will be taken as a *Fixed Regime*, since the intervention of monetary authorities is frequent.

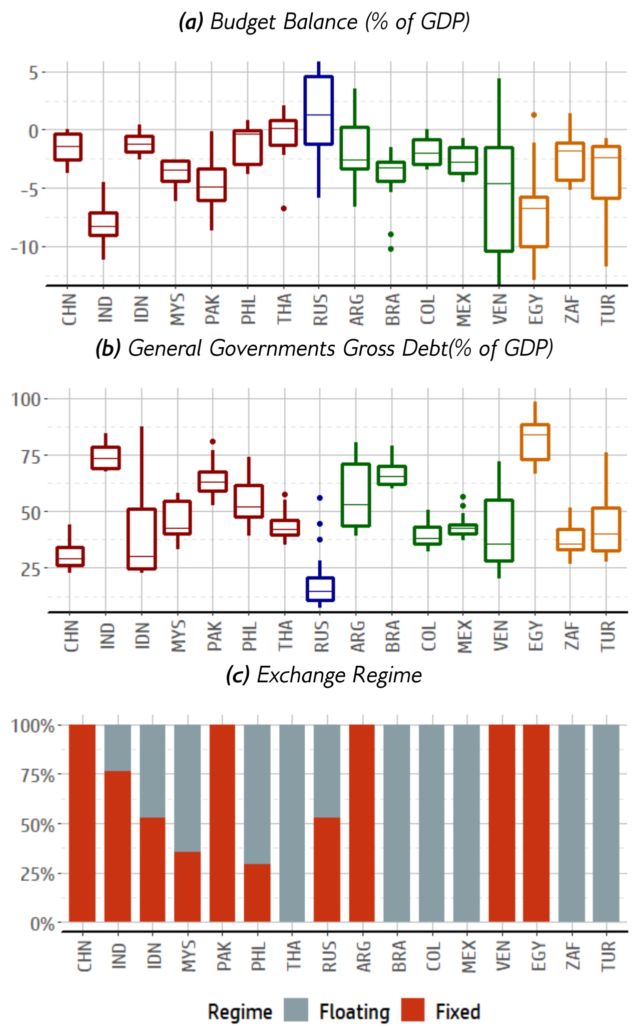



***Figure 2.4:*** *Fiscal Indicators*

**Table (2.4)** shows a statistical summary for the explanatory variables, done with STATA's command `xtsum`. The **overall** row consider all the pooled observations. The **between** row calculates the means by individuals ,

$$\bar{x}_i = \frac{\sum_t x_i}{N_i}$$

***Table 2.4:*** *Summary table for the explanatory variables*

| Variable | | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| GDPG | overall | 4.30 | 3.97 | -17.04 | 18.29 |
| | between | | 1.93 | 1.31 | 9.42 |
| | within | | 3.50 | -14.05 | 21.28 |
| GDPPC | overall | 5,335 | 3,775 | 462 | 15,942 |
| | between | | 2,991 | 976 | 9,146 |
| | within | | 2,414 | -1,491 | 12, 551 |
| INF | overall | 9.31 | 18.24 | -1.07 | 254.39 |
| | between | | 10.14 | 2.23 | 44.22 |
| | within | | 15.36 | -22.38 | 219.48 |
| CAB | overall | 1.03 | 4.85 | -8.94 | 17.76 |
| | between | | 3.84 | -4.08 | 9.44 |
| | within | | 3.11 | -9.98 | 12.74 |
| EDE | overall | 130.93 | 77.64 | 23.66 | 462.15 |
| | between | | 62.07 | 49.14 | 253.65 |
| | within | | 49.01 | 15.61 | 375.68 |
| RES | overall | 65.24 | 73.05 | 3.94 | 547.94 |
| | between | | 65.32 | 9.68 | 289.72 |
| | within | | 36.36 | -106.57 | 323.45 |
| BUD | overall | -3.06 | 3.58 | -15.55 | 7.81 |
| | between | | 2.46 | -8.23 | 1.28 |
| | within | | 2.66 | -13.72 | 6.28 |
| GD | overall | 49.85 | 20.88 | 7.45 | 152.25 |
| | between | | 16.84 | 19.59 | 82.21 |
| | within | | 13.01 | 22.33 | 135.64 |

and then the Standard Deviation of that row is the standard deviation of the individual means to the overall mean,

$$S_{\text{between}} = \textbf{SD}(\bar{x}_i) \qquad i = 1, \ldots, 16$$

Minimums and maximums are the average of the 16 individuals minimums and maximums. The **within** Standard Deviation is defined as the average of the individual standard deviations,

$$S_{\text{Within}} = \textbf{Mean}(\hat{s}_i)$$

Maximum and minimums in the **within** row don't have a significant meaning because of STATA's parametrization.

We can detect some preliminary relationships between the ratings and their macroeconomic performance by analysing the individual graphics and the statistical summary. For example, Malaysia has a higher distribution of its GDP per Capita relative to its peers in the Asia Pacific region, and might reflect why it has been labelled with the **Investment Grade** for the whole sample years. On the other hand, Pakistan has the lowest median of income overall, perhaps one of the reasons it hasn't been able to obtain a rating better than 'B'. Inflation (**Figure 2.2c**) may indicate political and social instability.

All the three economies with 'SD' rating (Argentina, Russia and Venezuela) show accumulation on high inflationary rates, with Venezuela off the chart. Turkey's struggles early in the new millenium account for some major inflation, but has manifested constant rates around 8% since. Thailand shows low and constant inflation, mostly below 8%, which demonstrates stability, and it may ground to their ratings, with few movements and ratings of 'BBB' and 'A', respectively. Colombia and Mexico display similar inflationary rates, nearly all around 3 and 6 percent. They have both been graded 'BB' and 'BBB', but Mexico has spent more time over at the Investment Grade category, probably because it has greater levels of income, among other reasons.

General Government's Gross Debt as a percentage of GDP can be appreciated on **Figure 2.4b**. Right of the bat, Egypt displays the highest median of the sample, with a gross debt hovering over 70 and 90 percent of its Gross Domestic Product. It has one of the poorest levels of income, and it is at the bottom of its region. The country was rattled by civil conflicts and political uncertainty in 2011, and has been downgraded from 'BBB' at the beginning of the sample, to 'B' in the last observation. Brazil and India show both the greatest level of gross debt of its region, between 60 and 80 percent of GDP, yet India has a lower income and improved its rating from 'BB' to 'BBB', while Brazil spent more time over at the **Speculative Grade** (below 'BBB'). As well, Brazil shows a greater distribution of External Debt stocks than most of Middle Eastern and Asian countries, however, it has one of the lowest level of International Reserves.

Despite having the best ratings in the sample, China has a Fixed Exchange Regime and the its Central Bank is known for intervening constantly to favour the Yuan. The country has been accused of manipulating its currency to favour its exports level. The real benefits and consequences of this practice are not perfectly clear, but it is for sure that the Government can act in any moment to improve its outlook [36].

On the other hand, Argentina has a Fixed Exchange regime as well, but is one of the worst performers in the sample. During the economic crisis the South American country suffered early in 2001, the Argentinian government cancelled a decade-long plan that which had coupled the Argentine peso to the US Dollar on a one-to-one basis since 1991. Following an cut-off of support by the IMF and the access to foreign capital, Argentina defaulted on 93bn USD sovereign debt. After four presidents tried and failed to grasp control of the economy, the acting-government implemented the Law of Public Emergency and Reform of the Exchange Rate Regime marking the end of the Convertibility Plan and the hard-peg with the US Dollar [17]. The crisis punished hardly the economy's growth prospects and send the country into a slow and painful recovery, which followed into a subsequent second default a decade later.

The Philippines, Malaysia, Russia, India and Indonesia are countries that transitioned from one exchange regimes to another mid-sample, with hope of improving its access to international capital markets. However, Indonesia was the only country which changed from a Floating regime to a Fixed one. We will see how the Exchange regime influences sovereign ratings and see how does the discrepancy in the extreme categories of the sample plays into the modelling.

Analysing individually the economies' indicators may provide some insight in line in regards of the explanatory variables influence the sovereign rating. Still, a model that incorporates all these factors and observations is needed to understand what are the rating agencies considering the most and considering the most when rating developing economies.

# CHAPTER III

# LINEAR REGRESSION FOR PANEL DATA

## 3.1 Introduction and Scale Transformation

The first modelling approach will be of **Linear Regression for Panel Data**. Panel data models help control individual heterogeneity, taking into account each unit's (country, state, patient, employee, household) specific characteristics that are not observable. In this study, all the individuals have notorious and important differences among their historical and cultural history, the strength in their financial institutions, religious affiliations and political regimes. Its reasonable to try to include them in the modelling and to compare what happens if we consider them or not.

Rating agencies supposedly incorporate all these intangibles into their analysis, reflecting their opinion in the rating grade. However, this study relies solely on quantitative factors, since the main objective is to compare two techniques of classification and to comprehend what are the relative weights and contribution of each indicator to the response. This would allow governments and decision makers to identify what are the rating agencies valuing the most. Panel data models let us measure these effects that are simply undetectable in cross-section or time-series models, making inference and interpretations richer. Moreover, panel data models give additional information, more variability, less collinearity among the variables, more degrees of freedom and more efficiency. Compared to the cross-section or time-series models, panel data can produce more reliable estimates, given some extra assumptions that will have to be tested in order to achieve this [5].

The following equation describes the first modelling proposal:

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + u_i + \varepsilon_{it} \tag{3.1.1}$$

where $i = 1, \ldots, N$ denotes the cross-section units, in this case each sovereign, and $t = 1, \ldots, T$ denotes the time-series dimension, 17 years in our case. The explanatory variables, $Xk$, were already described in **Chapter (II)**. $\varepsilon_{it}$ is the usual random disturbance of the regression, and varies over time

***Table 3.1:*** *Linear scale transformation*

| | Rating Grade | Scale |
|---|---|---|
| Investment grade | AA | 6 |
| | A | 5 |
| | BBB | 4 |
| Speculative grade | BB | 3 |
| | B | 2 |
| | CCC /CC / SD | 1 |

and individual.

The biggest obstacle to apply the regression models is to transform the ratings from a categorical variable to a continuous one. Cantor and Packer suggested in "Determinants and impact of sovereign credit ratings" that treating the rating as a cardinal value works better because of the large categories and the relative few sovereign rating assignments. In similar study of corporate ratings, Ederington suggests that the with larger sample sizes, inferences drawn from ordered probits are likely to be similar to, and perhaps slightly more accurate than, those drawn from least squares regressions [13]. In contrast, in their study of corporate bond ratings, Kaplan and Urwitz argue that linear least squares estimators perform better out of sample than those estimators derived from ordered probits [27].Nonetheless, ordered logit models will be examined further, as they represent the proper theoretical treatment of a credit ratings, which is roughly speaking, a qualitative opinion.

Table (3.1) displays the scale transformation to the ratings in order to apply the linear regression models. The linear continuous scale was chosen for simplicity, although there has been studies like [14], [35] that propose using logistic, logarithmic or quasi-logarithmic scales instead. However, those studies include in their scale the complementary symbols (+) or (-), and in some cases even the perspective outlook for the ratings. These provides more cardinal values and fits better non-lineal scales, but leaves few observations on each level. As well, it makes the job harder for classification techniques, making the models incomparable with the regression.

The time-invariant, *unobservable* individual specific effect is represented by $u_i$. It accounts for any individual-characteristic that is not included in classic regression. The models that deal with $u_i$ in a regression context are the **Pooled Regression**, the **Fixed Effects Regression** , and **Random Effects Regression**. Each one determines key assumptions over the $u_i$ component in order to obtain consistent estimates of the regression model. There are a series of statistical tests to validate the model's assumptions. The three models will be tried out, and after validating and analysing the results, we will determine which approach is the fittest.

# 3.2 Pooled Regression

The Pooled regression model assumes that $u_i = 0$. If true, **Ordinary Least Squares** will provide consistent and efficient estimates for the unknown parameters. Equation 3.1.1 becomes a common multiple linear regression model

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \varepsilon_{it} \tag{3.2.1}$$

This model makes the classical assumptions of the linear regression model:

- The random disturbances $\varepsilon_{it}$ are not correlated with any regressors and $\mathbb{E}[\varepsilon_{it}] = 0$.
- The random disturbances have the same variance $\sigma^2$ and are independent (homoskedasticity and non-autocorrelation).
- There is not exact linear relationship among the explanatory variables (no multicollinearity)

As typical model 3.2.1 is, it is not necessarily the right approach. Just by understanding the way the data is composed, it can be presumed *a priori* that there may be some issues regarding the constant and independent variation assumption. Hence, the OLS estimators are no longer the **Best Linear Unbiased Estimators (BLUE)**. Nevertheless, this model will help estimate the other two approaches and serves as a base of comparison between one and another.

## 3.3 Fixed Effects Model

This model assumes that differences across individuals can be captured in the constant term. A **fixed group effect** model examines individual differences in intercepts, assuming constant variance across individuals. Each $u_i$ is treated as an unknown parameter to be estimated. Since the specific characteristics of each individual are time invariant, and considered part of the intercept, then $u_i$ is allowed to be correlated with the explanatory variables.

The fixed effects model follows the next equation

$$y_{it} = (\beta_0 + u_i) + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + \varepsilon_{it} \tag{3.3.1}$$

This maintains the assumption that the random disturbances are not correlated with the regressors, and allows us to explicitly calculate the unobserved effect. We can test the significance of the estimates and determine how much the countries specific attributes influence the regression.

This model is usually refered as the **Least Squares Dummy Variable (LSDV)** model, as each fixed parameter can be estimated using a dummy variable with OLS, say:

$$Z_i = \begin{cases} 1 & \text{if the observation comes from the } i^{th} \text{ country} \\ 0 & \text{any other case} \end{cases}$$

However, when the number of individuals is large, trying to estimate this model directly using OLS becomes inconvenient since there are too many parameters to estimate and there's a major loss in the degrees of freedom [6]. If model (3.3.1) is true, then it incorporates *N* new variables, and if the experiment consists of thousands of patients or households, then it becomes practically impossible to compute estimates.

An alternative is to make the following transformation. From equation (3.3.1), we can average over time, for each $i$ country, and obtain

$$\bar{y}_i = (\beta_0 + u_i) + \beta_1 \bar{x}_{1i} + \ldots + \beta_k \bar{x}_{ki} + \bar{\varepsilon}_i \tag{3.3.2}$$

Substratcting equation (3.3.2) from equation (3.3.1) gives

$$y_{it} - \bar{y}_i = \beta_1(x_{it} - \bar{x}_{1i}) + \ldots + \beta_k(x_{kit} - \bar{x}_{ki}) + (\varepsilon_{it} - \bar{\varepsilon}_i) \tag{3.3.3}$$

or

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it} + \ldots + \beta_k \ddot{x}_{kit} + \ddot{\varepsilon}_{it} \tag{3.3.4}$$

The transformation in equation 3.3.4 is known as the **Within Transformation**. One can notice that the individual effect $u_i$ is now gone. There are not additional dummy variables , and deviations the group's mean are used instead. The OLS estimated coefficients in the within transformation are exactly the same as in the LSDV model [39]. Despite that, the within transformation brings up a new issue. There is no room available for the intercept and time invariant variables , such as sex, race, since the deviation from the mean is always zero. One has to be careful in considering the degrees of freedom of the regression, since it cause incorrect computations of the mean squared error (MSE) standard errors of the estimates (SEE) or the squared root of mean squared error (RMSE). The appropriate degrees of freedom of the fixed effect model are $(NT - N - K)$.

### 3.3.1 F - Test for Fixed Effects

It is of interest to test the joint significance of the individual effects coefficients:

$$H_0 : u_1 = u_2 = \ldots = u_{N-1} = 0 \tag{3.3.5}$$

The alternative hypothesis is at least one coefficient is not zero. Failing to reject $H_0$ in equation (3.3.5) would indicate that the Pooled Regression model should be used instead of the FIxed Effects model. The hypothesis is tested using an *F test*, which is based on loss of goodnes-of-fit. The test contrasts the Pooled and the LSDV (or Within) regression, examining the extent that the goodness-of-fit measures ($R^2$ or the residuals sums of squares).

The F statistic is given by

$$F_0 = \frac{(R^2_{LSDV} - R^2_{Pooled})/(N-1)}{(1 - R^2_{LSDV})/(NT - N - K)} \tag{3.3.6}$$

and

$$F_0 \sim F(N-1, NT-N-K)$$

under the null hypothesis. If there's enough statistical evidence that at least one intercept is non-zero, then the Fixed Effects model could be preferred over the Pooled Regression since it handles heterogeneity better.

## 3.4 Random Effects Model

The **Random Effects** model assumes that the individual effect from equation 3.1.1 are strictly **uncorrelated** with any regressors, and then estimates error variance specific to groups. This view would be appropriate if it is believed that sampled cross-sectional units were drawn from a large population.The component $u_i$ is the random **heterogeneity** specific to the *i*th observation and is constant through time.

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + \eta_{it} \tag{3.4.1}$$

Where the **composite error** $\eta_{it} = u_i + \varepsilon_{it}$ . The individual specific effects are now considered as random component of the composite error term. The usual assumptions are made about the idiosyncratic error term:

- $\varepsilon_{it} \sim IID(0, \sigma_\varepsilon)$
- $\mathbb{E}[\varepsilon_{it}\varepsilon_{js}|X] = 0 \qquad i \neq j \quad \vee \quad t \neq s$

In addition, the next restrictions have to hold in order so the **RE** model to be valid:

- $u_i \sim IID(0, \sigma_u)$
- $\mathbb{E}[u_i \varepsilon_{it}|X] = 0, \quad \forall\, i, j, t$
- $\mathbb{E}[u_i|X] = 0 \quad \forall i$

The individual effect must not be correlated with any regressor, and has to be independent to the idiosyncratic disturbance for all individuals and time periods. The intercept and slopes of the regression are the same for all countries, and their differences relies in their individual specific errors. The individual effect is characterized as random and inference pertains to the population from which the sample was randomly drawn.

From 3.4.1 and the assumptions stated above,

$$\begin{aligned} Cov(\eta_{it}, \eta_{js}) &= Cov(\varepsilon_{it} + u_i, \varepsilon_{js} + u_j) \\ &= \begin{cases} \sigma_u^2 + \sigma_\varepsilon^2 & i = j \quad \wedge \quad t = s \\ \sigma_u^2 & i = j \quad \wedge \quad t \neq s \end{cases} \end{aligned}$$

With a correlation coefficient between $\eta_{it}$ and $\eta_{js}$

$$\rho = \begin{cases} 1 & i = j \quad t = s \\ \sigma_u^2/(\sigma_u^2 + \sigma_\varepsilon^2) & i = j \quad t \neq s \end{cases} \tag{3.4.2}$$

The covariance structure for the *i*th individual is:

$$\Sigma = \begin{pmatrix} \sigma_u^2 + \sigma_\varepsilon^2 & \sigma_u^2 & \ldots & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 & \ldots & \sigma_u^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 \end{pmatrix} \tag{3.4.3}$$

And so, the global disturbances covariance matrix for model 3.4.1 is

$$\Omega = \begin{pmatrix} \Sigma & 0 & \dots & 0 \\ 0 & \Sigma & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & 0 & \Sigma \end{pmatrix} \tag{3.4.4}$$

The RE model is estimated by Generalized Least Squares **(GLS)** when the covariance structure is known, and by Feasible Generalized Least Squares **(FGLS)** when it is not, which is the usual case.

To apply FGLS, estimates for $\sigma_u^2$ and $\sigma_\varepsilon^2$ are needed. A consistent and efficient estimate for the latter, can be obtained from the LSDV transformation in (3.3.1)

$$\hat{\sigma_\varepsilon^2} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(\varepsilon_{it} - \bar{\varepsilon})^2}{NT - N - K} \tag{3.4.5}$$

A consistent estimate for the individual effects disturbance $\sigma_u^2$ can be obtained from the **Between Transformation**

$$\bar{y}_i = \beta_0 + \beta_1 \bar{x}_i + \ldots + \hat{\varepsilon} \tag{3.4.6}$$

where

$$\hat{\sigma}_u^2 = \hat{\sigma}_{between}^2 - \frac{\hat{\sigma}_\varepsilon^2}{T} \tag{3.4.7}$$

and

$$\hat{\sigma}_{between}^2 = \frac{SSE}{N - K - 1} \tag{3.4.8}$$

Estimates are obtained via OLS applied to

$$y_{it} - \hat{\theta}\bar{y}_i = \beta_0(1 - \hat{\theta}) + \beta(x_{it} - \hat{\theta}\bar{x}_i) + \varepsilon_{it} - \hat{\theta}\bar{\varepsilon} \tag{3.4.9}$$

where

$$\hat{\theta} = 1 - \sqrt{\frac{\hat{\sigma}_\varepsilon^2}{T\hat{\sigma}_{between}^2}} = 1 - \sqrt{\frac{\hat{\sigma}_\varepsilon^2}{T\hat{\sigma}_u^2 + \hat{\sigma}_\varepsilon^2}} \tag{3.4.10}$$

Estimators from equation (3.4.9) are efficient and consistent under the assumption that the individual effects errors are strictly heterogeneous to the regressors [40]. This assumption will be tested in order to compare later on the Fixed Effects vs the Random Effects model. One important remark can be noted from equation (3.4.9)is that if $\theta = 0$ then we get the pooled regression from model (3.1.1). And if $\theta = 1$ then the model turns in the within transformation in (3.3.3). That is the reason why model (3.4.9) is known as the **quasi demeaned transformation**. Instead of taking the total deviation from the mean like the fixed effects model, RE considers just partial weight on the grouped mean.

Note that $\theta = 0 \Leftrightarrow \hat{\sigma}_u^2 = 0$, i.e., there's not any individual disturbance besides the idiosyncratic error, which is in line with what the pooled regression assumes, that the unobserved individual effect is non existent. Furthermore, $\theta = 1 \Leftrightarrow \hat{\sigma}_\varepsilon^2 = 0$. This could be interpreted as if the only random disturbance comes from the individual effect, so the RE and FE models could be indistinguishable.

### 3.4.1 LM for Random Effects

The **Breusch-Pagan Lagrange Multiplier** test examines whether the random disturbances from the individual effects are significant or not, i.e.

$$H_0 : \hat{\sigma}_u^2 = 0$$

The LM statistic is

$$LM = \frac{NT}{2(T-1)}\left(\frac{\sum_{i=1}^{N}(\sum_{t=1}^{T}\varepsilon_{it})^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}^2} - 1\right)^2 \tag{3.4.11}$$

where $\varepsilon_{it}$ are the residuals from the pooled OLS regression. Under the null hypothesis, $LM \sim \chi^2_{(1)}$.

Failing to reject the null hypothesis would indicate that the presence of an additional disturbance coming from the specific characteristics of each country is not relevant, hence the pooled regression is a better choice.

## 3.5 Hausman's Specification Test: RE vs FE

Which model works better? The FGLS estimators from the RE model are consistent, efficient, and the model doesn't suffers from the great lost of degrees of freedom from the LSDV regression. In addition, this model allows the presence of time invariant variables, whereas the FE model don't. However, the Random Effects model could result inconsistent if the assumption of no-correlation among the regressors and the random individual disturbances is violated.

The **Hausman's Specification Test** [22] compares the RE and FE estimators under the assumption that $\mathbb{E}[u_{it}|X] = 0$ ( no correlation ). Under this hypothesis, the RE and FE are consistent, but only the Random Effects estimators are efficient (in fact, they are BLUE). Under the opposing hypothesis, the LSDV estimates would still be consistent, but the random error's estimators would not, hence, the Fixed Effect model is preferable.

The Hausman's test relies on the differences of the estimators for each model, $\hat{\beta}_{LSDV}$ and $\hat{\beta}_{RE}$.The covariance matrix of the difference of both estimators:

$$\begin{aligned} Var(\hat{\beta}_{LSDV} - \hat{\beta}_{RE}) &= Var(\hat{\beta}_{LSDV}) - Var(\hat{\beta}_{RE}) \\ &\quad - 2Cov(\hat{\beta}_{LSDV} - \hat{\beta}_{RE}) \end{aligned}$$

> *Hausman's essential result is that the covariance of an efficient estimator with its difference from an inefficient estimator is zero [19]*

This implies that:

$$\begin{aligned} Cov(\hat{\beta}_{LSDV} - \hat{\beta}_{RE}, \hat{\beta}_{RE}) &= Cov(\hat{\beta}_{LSDV}, \hat{\beta}_{RE}) - Var(\hat{\beta}_{RE}) \\ &= 0 \end{aligned}$$

$\Rightarrow$

$$Cov(\hat{\beta}_{LSDV}, \hat{\beta}_{RE}) = Var(\hat{\beta}_{RE})$$

Therefore,

$$Var(\hat{\beta}_{LSDV} - \hat{\beta}_{RE}) = Var(\hat{\beta}_{LSDV}) - Var(\hat{\beta}_{RE})$$

which is the covariance matrix for the difference of estimators, $\Psi$. The statistic is given by:

$$W = (\beta_{LSDV} - \hat{\beta}_{RE})'\hat{\Psi}(\beta_{LSDV} - \hat{\beta}_{RE}) \tag{3.5.1}$$

For $\hat{\Psi}$, the estimates of the covariance matrix from both models are used. Under the null, $W \sim \chi^2_{(k-1)}$. If $H_0$ is rejected, the it is concluded that the individual errors from the RE are correlated with one or more of the regressors, hence the estimates are inconsistent.

## 3.6 Chow Test for Poolability

Poolability asks if slopes are the same across group or over time. One simple version of poolability is an extension of the Chow test. The null hypothesis of this test is the slope of a regressor is the same regardless of individual for all $k$ regressors [33].

$$H_0 = \beta_{ik} = \beta_k$$

The F statistic is given as

$$F = \frac{(e'e - \sum e_i'e_i)/(n-1)(k+1)}{\sum e_i'e_i/n(T-k-1)} \sim F((N-1)(k+1), N(T-k-1)) \qquad (3.6.1)$$

where $e'e$ is the SSE of the pooled OLS and $e_i'e_i$ is the SSE of the pooled OLS for group $i$. If the null hypothesis is rejected, the panel data are not poolable; each individual has its own slopes for all regressors.

The Chow test assumes that individual error variance components follow a normal distribution $u_i \sim N(0, s^2 I_{nT})$. If the assumption does no hold, the Chow test may not properly examine the null hypothesis [5].

## 3.7 Results and Analysis

**Table 3.2** shows all the parameter estimates and relevant statistics from the panel data models. The Newey - West robust standard errors, shown in parenthesis, account for heteroskedasticity and autocorrelation among the residuals.

**Figure 3.1** displays the Predicted vs Observed ratings after applying the three regressions. This plot is known as a *violin plot*. The area around the box-plot represents an estimated density of the fitted values for each model. Along with the box-plot, that displays the quantiles 25%, 50% and 75%, the intention of the graph is to understand how well the regression models predict ratings, comparing it to the actual observed value that it came from. The observed ratings are displayed on the X-axis, while the fitted values of each regression are displayed on the Y-axis.

First, an individual analysis for the regression models. Then, we will compare the three of them and will define which one makes the best fit.

### 3.7.1 Pooled Estimation

The second column on Table 3.2 display the estimates for the pooled regression . The regression is significant, but only four of the nine parameter estimates are significant at the 0.15% level. The Durbin-Watson statistic for first order autocorrelation in the residuals shows significant evidence of positive autocorrelation. As well, the null hypothesis of the White test for homoskedasticity was rejected.

Only the Inflation estimate became non-significant at $\alpha = 0.05$ level in contrast with the regular covariance matrix estimates. The signs are in line with the preliminary intuition made when describing the variables. Two of the three External Indicators resulted significant at the highest level, as well as the Fixed Exchange Regime dummy variable. None of the Fiscal Indicators estimates is significant, and only the level of income of each sovereign, GDP per Capita, is significant at 0.1%. External Debt to Exports and International Reserves have similar magnitude in their coefficient. If the ratio of External Debt to Exports increases by 100 percentage points, the sovereign rating will **decrease** closely by 0.5 units, maintaining all the other indicators constant. On the contrary, a 100 percent points increase in the proportion of International Reserves to the external debt stocks, will **improve** the ratings by half a unit, *ceteris parabus*. Since the GDP per Capita was transformed to a logarithmic scale, is coefficient indicates that

> *An increase of 100% in GDP per Capita is associated with an expected increase of 0.185 units in the rating scale*

That increase seems relatively small since we are doubling the level of income. The estimated densities of the predicted values are the red "violins" in **Figure 3.1**. At a first glance, the regression seems to fit the best at the 'BB' and 'BBB' rating. Nonetheless, it is where the majority of the sampled observations are. The **Exchange Regime** dummy variable coefficient resulted significant, signalling that having a floating exchange

***Table 3.2:*** *Parameter estimates for the Panel Data models*

| | Pooled OLS | Fixed Effects Model | Random Effects Model |
|---|---|---|---|
| GDP Growth | 0.0017 | -0.0075 | -0.0064 |
| | (0.16) | (0.01) | (0.02) |
| GDP per Capita | 0.1850*** | 0.7292*** | 0.3267*** |
| | (0.09) | (0.09) | (0.11) |
| Inflation | -0.0107 | -0.0065*** | -0.0075*** |
| | (0.01) | (0.001) | (0.003) |
| Current Account Balance | 0.0025 | 0.0142 | -0.0108 |
| | (0.02) | (0.01) | (0.02) |
| External Debt / Exports | -0.0053*** | -0.0008 | -0.0025*** |
| | (0.0009) | (0.001) | (0.0009) |
| External Reserves | 0.0057*** | 0.0005 | 0.0042*** |
| | (0.001) | (0.001) | (0.001) |
| Budget Balance | -0.0106 | 0.0625*** | 0.0342* |
| | (0.02) | (0.003) | (0.02) |
| Gross Debt | -0.0016 | -0.0031 | -0.0017 |
| | (0.003) | (0.003) | (0.003) |
| Fixed Exchange Regime | -0.4862*** | -0.092 | -0.1372 |
| | (0.17) | (0.10) | (0.17) |
| Intercept | 2.4358*** | *(NA)* | 0.9928 |
| | (0.86) | | (1.02) |
| F Statistic$^a$ | 48.34*** | 74.23*** | 212.34*** |
| $R^2_{adj}$ | 0.6287 | 0.8664 | 0.4286 |
| RMSE | 0.6658 | 0.3801 | 0.5146 |

*** Significant at 0.05 level ** Significant at 0.10 level $^a$ Wald Statistic for the RE model

regime improves the rating. However, all the observations of the 'AA' rating come from China, that has a **Fixed** Exchange Regime. The predicted values of the 'SD' and 'A' ratings are the ones with the most variability.

The regression predicts a negative value on the continuous scale, for an observation that belongs to the 'SD' grade, that goes off the chart. This dot represents Venezuela in 2016. Although it doesn't make sense theoretically (ratings cannot go below the lowest possible grade), in a strict empirical point of view, that observation is an outlier, and it reflects the precarious and extreme conditions the South American country has been facing in the last couple of years. Even compared with low-performing nations, Venezuela is doing far worse, at least by what the model is saying. However, keep in mind these particular ratings are evaluated with a different methodology, so it shouldn't cause any surprise that the regression doesn't fit well around them.

On the the other hand, all the predicted points of the highest observed rating are way below their theoretical value. The regression is underestimating completely the sovereigns with the best grade, 'AA', although some observations that have the previous rating, 'A', do go above and reach the next grade. The small sample of the high rating observations may be playing a part in this initial adjustment. Assuming the individual characteristics have no part in estimating the linear regression model, lead to statistical inconsistency and violation of one of the core assumptions. Overall, the low grade observed points are overrated by the regression, while the high ones are being under appreciated. The pooled regression shows high variability, only a few parameters are significant and presents the expected heteroskedastic and autocorrelated residuals.

## 3.7.2 FE Estimation

Estimates are shown on the third column of **Table 3.2**. Only three parameters are statically significant: two economic indicators and one fiscal indicator. GDP per Capita shows great significance, as well as the Inflation rate of change. The General Government Budget Surplus is significant, and has a positive


***Figure 3.1:*** *Observed Ratings vs Predicted for the Panel Data Models*

sign, which is in line with the initial intuition. The estimate of the Budget surplus has the second biggest magnitude, behind the GDP per Capita. The Inflation Rate is significant, but has a low impact in the regression. This may be due the fact that there are countries that present inflation rates beyond 100%, like Turkey or Venezuela.

*An increase of 100% in GDP per capita is associated with an expected increase of 0.72 units in the rating scale, holding all constant.*

*Holding everything constant, an increase of 10 percentage points in the the Budget Balance, relative to GDP, translates in an increase of 0.6 units in the rating.*

*An increase of 20 percentage points, roughly one standard deviation, on the Inflation rate, reduces the rating by 0.13 units, ceteris paribus.*

Though the direct impact of inflation on the rating may seem despicable, a high inflation rate may put in the crossroads the nation's economic prospects, as well as a conflictive political and social climate. The $R^2_{adj}$ is close to one and the RMSE is lower than in the pooled regression, although we have to keep in mind that the FE model has $N-1$ more parameters, so that's probably artificially inflating the regression coefficient.

Nevertheless, we can see in Figure 3.1 that the predicted points (blue violins) are scattered more consistently around the theoretical value, meaning there's less variability in some cases. The estimated kernel densities are centred around the observed value, except for the 'SD' and 'AA' category. In the 'BB' and 'BBB' rating, there are just a few observations that miss their observed value for more than a unit, and the majority of the fitted values hover around the exact value by $\pm 0.5$ units, which is relatively better than the previous model.A majority of the predicted points in the lowest rating tend to go one level higher than the observed, but there's no negative value like in the pooled regression. On the contrary, All the 'AA' predicted values are under six, so the regression is underrating this observations. The estimate for the dummy variable of Fixed Exchange Regime is not significant. This may be due the fact that the Exchange Regime variable is "almost" time invariant. Changes in the exchange regime are rare, and most of the countries had the same regime for the whole sample. The within transformation eliminates all the time invariant variables, so it may be the cause of the lack of significance.

#### Fixed Effects Coefficients

The coefficients of the individual effects $u_i$ can be recovered as follows:

$$\hat{u}_i = (\bar{y}_i - \bar{\bar{y}}) - \hat{\beta}_{LSDV}(\bar{x}_i - \bar{\bar{x}})$$

where $\bar{y}_i$ is the average over time for the $i$th individual, and $\bar{\bar{y}} = \sum_i \sum_t y_{it}/NT$ is the average over time and individual, and $\hat{\beta}_{LSDV}$ are the estimates from model (3.3.4)

**Table (3.3)** shows the fixed effects estimates sorted in ascending order. All the coefficients have a negative sign, which indicates that the regression's hyperplane is "lowered" at different levels. The significance of the parameters is reduced as they decrease in value. Newey-West robust standard errors are reported. The fixed effects coefficients allow a fitter regression and provide a measurement of the unobserved characteristics. In some way, they reflect social, political, cultural and governance factors that simply can't be measured quantitatively.

Argentina and Venezuela, two countries with history of defaulting, have the lowest individual effects. Russia also defaulted, and has the fourth lowest intercept. Indonesia and Pakistan both tested at some point in time the 'CCC' or 'CC' rating, and their individual effects coefficients are statistically significant. Egypt experienced a social revolution around 2011 with considerable economic and financial consequences [1], but according to the model, its credit profile was somewhat less impacted by the unobserved characteristics, than Mexico or Colombia, which have not experienced any violent or conflictive social or political shift. Brazil is Latin America's biggest economy. However, its individual characteristics seem to affect its rating to a greater degree, than Pakistan and its own. The Brazilian government was involved in a grand scheme of corruption with publicly renounced Odebrecht scandal. State-owned oil giant Petrobras was related as well and it is right at the middle of the turmoil.

Being able to measure in some way the intangible individual characteristics of each nation is of great advantage in trying to understand the rating agencies' opinion. Even if we could measure all the intangibles, and could create a model that incorporates every single of the quantitative and qualitative factors into account, there could still be a margin of error, something that is missing.It could be considered as if the coefficients express the *Rating's Agencies bias* towards each country. However, it could be just due the unobserved effects on the sample. That's an outstanding take the Fixed Effect model brings to the table. It allows us to identify the bias in Standard & Poor''s sovereign ratings, at least from the missing factors that are not incorporated in this regression.

### 3.7.3 FE vs Pooled Regression

Recalling from **Section 3.3.1**, we want to test the joint significance of the individual intercepts, to validate the FE model over the ordinary linear regression. Using the results from Table 3.2, we get that:

***Table 3.3:*** *Individual effects coefficients*

| | Intercept | Robust Std. Er. |
|---|---|---|
| Argentina | (4.82) *** | 1.03 |
| Venezuela | (3.91) *** | 0.99 |
| Turkey | (3.29) *** | 0.93 |
| Russia | (3.18) *** | 0.92 |
| Indonesia | (2.95) *** | 0.85 |
| Brazil | (2.84) *** | 0.93 |
| Pakistan | (2.74) *** | 0.80 |
| Colombia | (2.60) *** | 0.86 |
| Mexico | (2.46) *** | 0.91 |
| Egypt | (2.26) *** | 0.88 |
| Phillipines | (2.08) *** | 0.80 |
| Thailand | (1.97) *** | 0.83 |
| S. Africa | (1.96) *** | 0.85 |
| Malaysia | (1.45) | 0.97 |
| China | (0.80) | 0.81 |
| India | (0.78) | 0.79 |

*** Significant at 0.05% ** Significant at 0.10%

$$\begin{aligned} F_0 &= \frac{(R^2_{LSDV} - R^2_{Pooled})/(N-1)}{(1-R^2_{LSDV})/(NT-N-K)} \\ &= \frac{(0.8782 - 0.6402)/(15)}{(1-0.8782)/(247)} \\ &= 32.195 \end{aligned}$$

The test statistic is large enough for us to reject $H_0$, so it can be concluded that at least there's one $u_i$ coefficient statistically distinct from zero, so the Fixed Effects models provides a more reliable linear adjustment than the Pooled Regression.

The **Fixed Effects** regression proved to be suitable for this study. It lowered the variability, but most importantly, gave us a sense of the magnitude of the intangible factors that are not incorporated in the regression. The only disadvantage is that the model doesn't allows time invariant variables,

### 3.7.4 RE Estimation

From Section 3.4 we get that the estimates for the parameters are calculated using Generalized Least Squares. Since the covariance matrix is unknown, $\theta$ is estimated using (3.4.5) and (3.4.8) from the LSDV and between transformation, respectively:

$$\begin{aligned} \hat{\sigma}^2_{between} &= \frac{SSE_{between}}{N-K-1} \\ &= \frac{0.2981}{16-9-1} \\ &= 0.0497 \end{aligned}$$

$$\begin{aligned} \hat{\sigma^2_\varepsilon} &= \frac{SSE_{LSDV}}{NT-N-K} \\ &= \frac{39.2999}{16(17-1)-9} \\ &= 0.1591 \end{aligned}$$

$$\Rightarrow$$

$$\begin{aligned} \hat{\theta} &= 1 - \sqrt{\frac{\hat{\sigma^2_\varepsilon}}{T\hat{\sigma}^2_{between}}} \\ &= 1 - \sqrt{\frac{0.0497}{T(0.1591)}} \\ &= 0.5660 \end{aligned}$$

The $\rho$ coefficient in (3.4.2) represents the ratio of individual specific error variance to the composite disturbance. In this case $\rho = 0.2022$. This can be interpreted as if the individual disturbance only accounts for 20.2% of the total constant variance in the model.

Table (3.2) shows the estimates of the Random Effects regression in the fourth column. Five estimates are significant, two economic indicators (GDP per Capita and Inflation), two external indicators (External Debt and Reserves) and one Fiscal indicator (Budget Surplus). The significant External indicators impact the regression distinctly, showing that the level of International Reserves to External Debt Stocks are more important than the External Debt to Exports ratio. According to this model, it is more valuable to the ratings how well the sovereign is prepared to face contingencies on its international obligations, than the amount of obligations itself.

> *Doubling GDP per Capita is associated with an expected increase of 0.33 on the rating scale*

All the other interpretations are really similar to the ones done in the FE case. The corresponding fitted values for the Random Effects regression are represented by the yellow "violins". The model seems to fit best the observations coming from the 'BB' and the 'BBB' rating. Besides that, the predicted values from the two low ratings are concentrated predominately under their point of observation. The 'BBB' rating's kernel density is just centred off 5 not by much, so the regression is making a good adjustment there, and has a lower variability than the 'BB' predictions. However, the following observations coming from the next categories are mostly overrated. All the 'AA' ratings are above by almost one unit, which technically means the 'AAA' rating. We could say that this model penalizes countries with low ratings, but is lax with sovereigns that have the Investment Grade. There's a lot of variability among the 'SD' rating, and fitted values that have a negative value appear again, like in the pooled regression. But in that case, the majority of the points were overrated, opposed to the RE regression. The following test will help us statically compare the Random Effects Regression and its assumptions against the Pooled OLS regresion.

#### RE vs Pooled

In order to validate the RE model's fit over the Pooled regression, we need the residuals from Section 3.7.1. The Breush-Pagan Lagrange Multiplier estimator from 3.4.11 estimator is

$$LM = 176.52$$

that leads to a $p - value \approx 0$. With this result, it can be concluded that there's presence of an additional disturbances proceeding from the individual unobserved effects, and these random errors account for 20% of the composite error terms on the regression. The LM statistic validates the presence of individual heterogeneity, proving that, at least in a pure statistical point of view, the Random Effects model makes a better fit in the regression.

### 3.7.5 Testing FE vs RE

So far the unobserved individual effects have been treated in different ways: either they are omitted from the regression (Pooled model), they are considered as an extra parameter that is estimated, and its effect makes an adjusts in the regression (Fixed Effects), or are thought to be an complementary random error (Random Effects). There's enough statistical evidence from the F-test in section 3.3.1 and from the Lagrange Multiplier in section 3.4.1 to conclude that the intangible country-specific properties do play a role in estimating the regression and help accomplish a better adjustment. Hausman's Specification test from Section 3.5 compares the estimates from the *consistent* model, the FE model, versus the *efficient* model, the RE model. The test statistic from from Equation (3.5.1) is

$$\hat{W} = 108.95 \quad \sim \chi^2_{(8)}$$

$\hat{W}$ is big, with a $p - value \approx 0$, hence we reject $H_0$, that the individual random components are not correlated with any regressor.

Evidence suggests that the individual disturbances of the composite error are related strongly with one of the explanatory variables, making the estimates of the RE regression inconsistent. However, the estimates of the Within model 3.3.3) are always consistent, even if $H_0$ is rejected. The Random Effects model works better when the sample is assumed to be randomly taken. This is not the case in this study, so it is not so valid to consider that the unobserved effects are completely random. Therefore, based on the results of the test, the Fixed Effects regression is more appropriate in modelling the ratings, since the key assumption of the Random Effects model did not hold.

### 3.7.6 Poolability Test

Lastly we will test if the assumption of having the same vector of coefficients for the regressors is doable or not using the Chow test from Section 3.6. Using the residuals from the pooled model, we get that the Chow statistic is

$$F_{Chow} = 32.20 \sim F(150, 112)$$

The *p-value* associated to the test is $\approx 0$, so we conclude that poolability may not be viable, however, we will discuss this result further.

## 3.8 Discussion

Now that the proper individual analysis and results for the models has been presented, we will determine which explanatory variables made the larger impact in the regression overall, which model is the most reliable and what considerations we have to keep in mind when interpreting the results.

### 3.8.1 Relevant Regressors

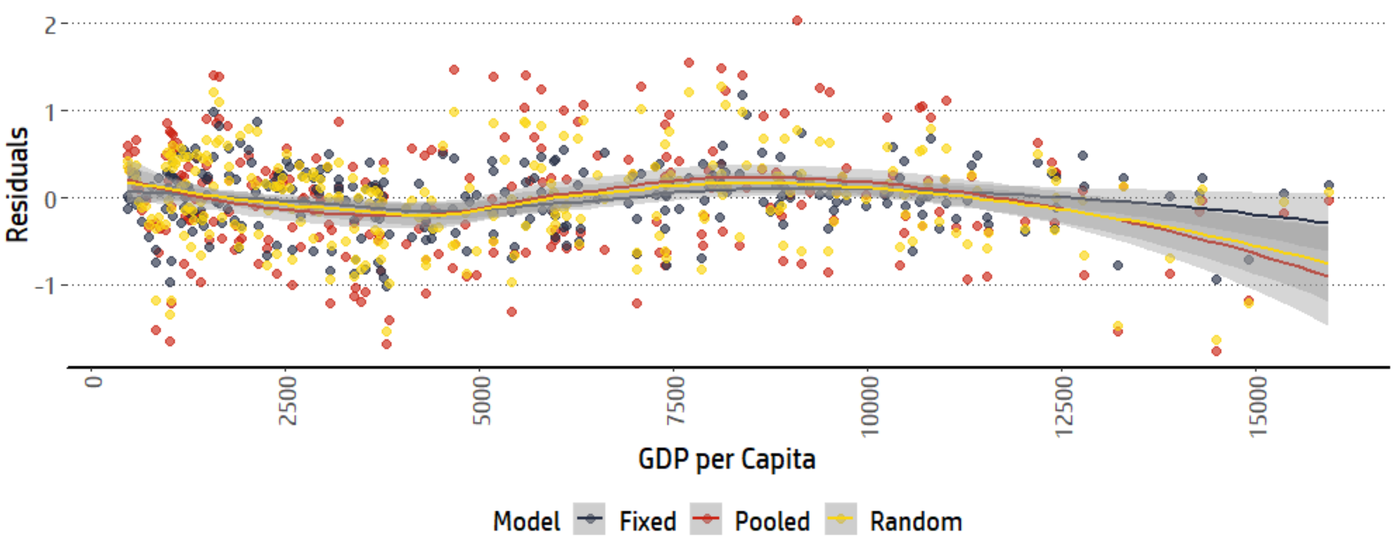
GDP per Capita (logaritmic scale) is the only variable that has a significant coefficient in the three adjustments. Inflation's coefficient, as well as External Debt to Exports, External Reserves and Budget Balance are significant in two. The Exchange regime is significant only in the pooled model. We could say that the Economic Indicators in the Panel Data regressions, are the ones with the most statistical significance, then the External Indicators and finally the Fiscal Indicators. It can interpreted as if the rating opinion values more the wealth and stability of the Emerging Markets nations. A richer country has a better chance of paying off its obligations. We can conclude that, considering the sample, the factor that influenced more sovereign rating was **GDP per Capita**.

Figure (3.2) shows the residuals from the regressions vs the GDP per Capita in Current US Dollars. The Random Effects and the Pooled Regression show greater variability, and have a downwards tendency with high-income observations. The Fixed Effects model shows constant variability to a lesser degree. The External Indicators that resulted significant in two model were the External Debt to Exports and the International Reserves to external Debt Stocks. The signs are in line with the initial intuition, and the pooled regression is the model that presents the highest magnitude on the estimate for the indicators. A higher level of External obligations relative to what the country is gaining from trade, creates a higher risk and puts pressure on the sovereign rating, impacting it negatively.

On the hand, if the country is well prepared and its foreign currencies reserves are sufficient in case of contingency, relative to the level of what it owes to foreign investors, then this will benefit its sovereign rating. The model that assigned the least weight to the External Indicators was the Fixed Effects regression. The estimates have the lowest magnitude of the three, though the three models showed practically the same level of error for the estimates. The Current Account Balance was not significant in any of the panel data models. It has an estimate of $\approx 0.01$ in the FE and Random Effects models, but shows relative big robust standard errors. This may indicate that the level of trade deficit (or surplus) is not congruent with the rating each sovereigns has, so in the end it results not significant to the regression.

The Budget Surplus was significant in the Fixed Effects and Random Effects regression, where it showed the highest magnitude of the three models. The pooled regression gave a negative estimate for this indicator, but its influence in that particular regression was not significant. The estimate in the Random Effectes regressions has important variability, but shows the proper sign. The difference between the revenues and the

*Figure 3.2: GDP per Capita vs Residuals from the Panel Data Models*




total expenses of the General Government relative to GDP, impacts the regression accordingly. A fiscal surplus indicates a higher sovereign rating, and inversely if the Government presents deficit. A standard deviation increase (3.58%) in the Budget Balance translates into an expected increase of 0.22 on the rating scale, holding everything constant, for the FE model.

The Exchange Regime was only significant in the Pooled regression. With a negative sign, if the country has a Fixed Exchange Regime, then its rating decreases by almost 0.5 units. However it resulted non-significant in the Fixed Effects regression because of how this models treats time-invariant variables, and the Random Effects estimate has an error higher that the coefficients value. Gross Debt of the General Government was not significant in any model. The value of the estimate is despicable in all the cases, so it may not be influencing the regression. Apparently, the rating agencies value the foreign debt burden more over the total debt owed by the General Goverment, which is understandable since the ratings we are considering are the ones issued to foreign currency obligations.

### 3.8.2 Fittest Model

Now that the three proposals have been tried, and the corresponding individual comparisons have been analysed, it is fair to say that the **Fixed Effects** model is the fittest choice. The Fixed Effects model shows the less variability among the residuals, has the best goodnes-of-fit measures, and its main assumption was statistically more significant in the F-test (Section 3.3.1) and in Hausman's Specification test (Section 3.5). Furthermore, just by the way each model is intended to work, the Fixed Effects model make more sense in the sake of this experiment. The Random Effects model works well and is efficient when the individuals are assumed being randomly sampled, whereas in here, the sample was selected on purpose. We don't suffer from the disadvantage of loosing considerable degrees of freedom using the LSDV transformation, since the panel data used in this experimentation consists of just 16 individuals.

As discussed in Section (3.2), it is more reasonable to assume that the unobserved characteristics of each sovereign nation are correlated to the explanatory variables, that assuming that they are completely random and have no strong relationship. On top of that, we can **measure** the individual coefficients

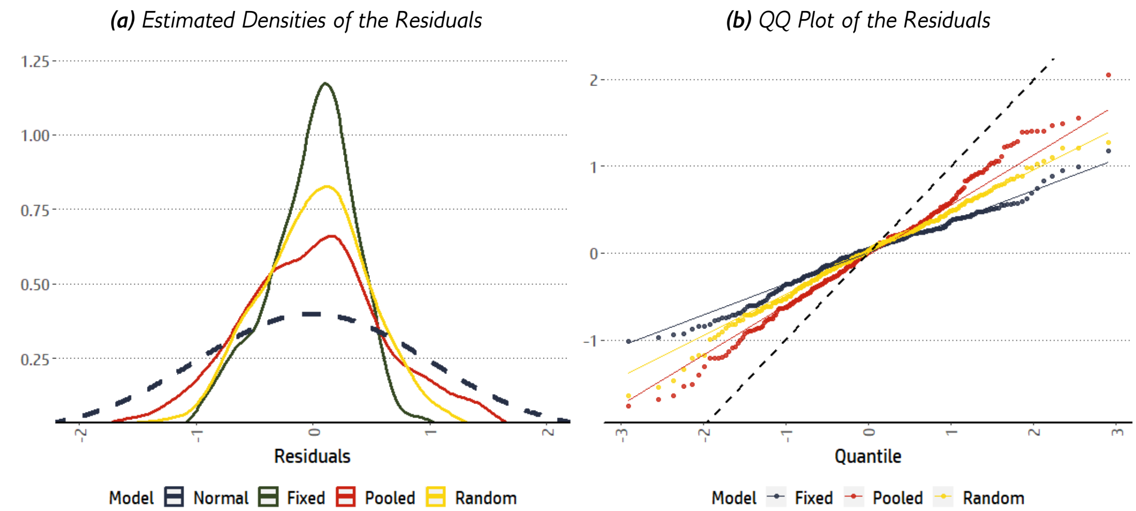



***Figure 3.3:*** *Diagnosis Plots of the Residuals for the Panel Data Models*

and understand the impact the have in the regression. One possibility, in this experiment, is to think of the intercepts estimates as the bias the rating agencies opinion regarding the sovereign's credit profile. If all the possible factors are taken into account, there's still subjectivity present in the rating's agencies final opinion. If the individual effects resulted to be all of the same proportion, then we would assume the note on the sovereigns where being equally rated and the governance and cultural factors specific to each country would not affect its rating. Then the Pooled regression or the Random Effects could have done a better adjustment. However, this was not the case.

Beyond the discussion if the rating agencies actions are strictly ethical or transparent, it does makes sense to study each nation carefully and to have some kind of discrimination as an investor. We noted on Table (3.3) that the most relevant (and significant) individual effect impacting the regression was the intercept coefficient of Argentina. The South American country is the only country that has defaulted twice in the time sample. This means that its rating implies a higher defaulting chance than, for the sake of comparison, Brazil or Colombia, who belong to the same region. This could also be interpreted as if rating agencies punish countries that have defaulted in the past.

### 3.8.3 Considerations

Panel data regression served as a good first approach in order to understand the factors and weights the proposed indicators have and to measure their significance. Still, there are areas of opportunity. **Figure (3.3a)** shows the estimated densities of the residuals for each adjustment. The dotted line is the Normal density curve. We can see that the residuals are skewed, and have different levels of dispersion. Rejecting the null hypothesis that the residuals follow a normal distribution could lead into inconsistencies among the comparisons tests. The FE residuals have the lowest level of variability, still it is not safe to say that they are normally distributed. **Figure (3.3b)** shows the QQ-Plot for the residuals of each model. The dotted line shows the theoretical quantiles, and all the models seem to miss them. Running normality tests showed mixed results, but the model that consistently have **non-normal** residuals is the Fixed Effect Model. So model that apparently worked the best is the one that violates one the linear regression assumption, though not one of the most important. Thus, conclusions done from the comparative tests, like the Hausman's specification test, have less sustenance. Also, we expected this adjustment to be better, since we **forced** the independent variable was forced to be linear. Linearly scaling the ratings, as an easy adjustment, brings non-realistic economic assumptions, since we assume that moving from one rating to another takes the same effort, independently of the rating each country has. Meaning, one country that has the 'SD' rating can get to the next category just a country with the 'BBB' can to the next. We mentioned, however, that since we consider the ratings only at the indicative level, then there is no point in using a different rating scale, such as logarithmic or exponential transformations.

The Chow test resulted significant, however we can notice as well in **Figure (3.3b)** that the residuals from the pooled model are hardly normally distributed, so inferences from that test are also questionable. It is natural to think that the test could result highly significant, since all the countries have large differences in the economic prospects and in the ratings, but it was the purpose of this study to understand the factors in a pooled manner. Analysing individually would leave us with separate correlated regressions for each country, and it could be harder to make interpretations if we were interested in analysing beyond one country at a time. We believe that the violation of this assumption is not serious, and it is in line with has been observed in the literature.

Treating the ratings as continuous instead of as a categorical variable allowed direct interpretation and clarity. Panel Data models provide great simplicity, computationally and analytically. It is the traditional and most common used tool for econometrical analysis. In this case it might have kept short because of the nature of the endogenous variable, and the strong assumptions that had to be done in order to get results. Additional proposals for further work are including in the regression time effects as well as individual effects. The year when the observations was taken may influence the rating since Emerging Markets economies are more susceptible to economic cycles. There are other alternatives that allow this coefficients to vary over time or over individual. Linear Regression treated the rating as cardinal values. Now, we'll try alternatives to model the rating considering the ratings ordinal and categorical nature, and we will compare the results from both techniques and to see if they are coherent or vary a a lot with what we found on this chapter.

# CHAPTER IV

# ORDINAL LOGISTIC REGRESSION

As discussed in the previous chapter, we assumed that the distance underlying the ratings was of one cardinal unit. This may not be a realistic economical nor statistical assumption, since it could be interpreted as if going from 'B' to 'BB' could require the same amount of "effort" that going from 'AA' to 'AAA'. We will try the **proportional odds model** to treat the ratings as a categorical variables instead of a continuous one. This particular model is the most widely used and can be implemented in `STATA`, `SAS` or `R`. It also has an extension to panel data. As we did with the linear regression, we will first treat the data ignoring its longitudinal form, and then we will deal with the heterogeneity via Random Effects and Fixed Effects.

## 4.1 Proportional Odds Model

Given the ordinal nature of the credit ratings, in particular, the sovereigns we are studying, there are several models that have been developed for categorical outcomes. Here's the motivation that leads to the categorical modelling:

Suppose there is an underlying continuous variable, $Y^*$, that discretizes a response variable such that

$$Y_i = \begin{cases} 1 & \tau_0 \leq Y_i^* < \tau_1 \\ 2 & \tau_1 \leq Y_i^* < \tau_2 \\ \vdots & \\ J & \tau_{j-1} \leq Y_i^* < \tau_j \end{cases} \tag{4.1.1}$$

$Y^*$ is a **latent**, unobservable variable, such that the response variable $Y$ takes values on a discrete scale, depending on the range of $Y^*$. $\tau_j$ are "cutpoints" or thresholds on the continuous unobserved scale, so if $Y^*$ falls in the *jth* interval, then $Y$ takes the *Jth* category, where $\tau_0 = -\infty$ and $\tau_j = \infty$

The latent variable can be modelled as

$$y_i^* = X_i\beta + \varepsilon_i \tag{4.1.2}$$

Where $i = 1, \ldots, N$. If we want to know the probability of and observation from the response variable $Y$ falling under the *mth* category, we get from (4.1.1) and (4.1.2) that,

$$\begin{aligned}\mathbb{P}[y_i = j|X] &= \mathbb{P}[\tau_{j-1} \leq y_i^* < \tau_j|X] \\ &= \mathbb{P}[\tau_{j-1} \leq X_i\beta + \varepsilon_i < \tau_j|X] \\ &= F(\tau_j - X\beta) - F(\tau_{j-1} - X\beta) \qquad (4.1.3)\end{aligned}$$

with $F$ being the cumulative probability distributions of the random disturbances $\varepsilon_i$. The most common model is the **proportional odds model**, which is the most widely used in statistical software

This model is an extension of the **binary logistic regression**, since it allows more than two categories. The model assumes the **logistic distribution** as the link function between the regressors and the ordinal outcome ,but instead, it uses the **cumulative probability** of belonging to one category or not:

$$\begin{aligned}C_j &= \log\left(\frac{\mathbb{P}(Y \leq j\,|\mathbf{X})}{\mathbb{P}(Y > j\,|\mathbf{X})}\right) \\ &= \log\left(\frac{\pi_1 + \pi_2 + \ldots + \pi_j}{\pi_{j+1} + \ldots + \pi_{J-1}}\right) \\ &= \tau_j - \beta\mathbf{X} \qquad (4.1.4)\end{aligned}$$

for $j = 1, \ldots, J-1$. $C_j$ represents the **cumulative logit** for a response variable $Y$ that has $J$ possible outcomes.

Each cumulative logit has its own intercept, $\tau_j$. From (4.1.4) we can recover the cumulative probability assuming the logistic distribution as

$$\mathbb{P}[Y \leq j] = \frac{\exp(\tau_j - X\beta)}{1 + \exp(\tau_j - X\beta)} \qquad (4.1.5)$$

The proportional odds model assumes that each logit has its own intercept, but share the same vector of regression coefficients $\beta$. Thus, for a fixed *j*, the response curve is a logistic regression curve for a binary response with outcomes $Y \leq j$ or $Y > j$. So we get *J-1* parallel lines, translated along the x-axis. For every $j < k$, we get that curve of the event $Y < k$ is the same curve as $Y < j$ translated by $\frac{\tau_k - \tau_j}{\beta}$.

An odds ratio of cumulative probabilities is called a *cumulative odds ratio*. The odds of making response $\leq j$ at $X = X_1$ are $\exp[\beta(X_1 - X_2)]$ times the odds at $X = X_2$. The log cumulative ratio is proportional to the distance between $X_1$ and $X_2$. The same proportionality applies to each logit, hence, the name of the model: **proportional odds model**.

Model (4.1.4) subtracts $\beta\mathbf{X}$ rather than adding it. This allows for the same sign of the coefficient $\beta$ to have the usual meaning: if $\beta_k > 0$ then a unit increase in $x_k$ the the cumulative logit decreases. Hence, the corresponding cumulative probability decreases. Then, relatively less probability mass falls at the end of the response scale, and Y is less liely to fall at the low end and more likely to fall at the high end of the scale.

Let $(y_1, y_2, \ldots, y_n)$ be a sample of the response for subject $i$. Then, the likelihood function is

$$\begin{aligned}\prod_{i=1}^{n}\left[\prod_{j=1}^{J}\pi_j^{y_{ij}}\right] &= \prod_{i=1}^{n}\left[\prod_{j=1}^{J}(\mathbb{P}[Y \leq j|X_i] - \mathbb{P}[Y \leq j-1|X_i])^{y_{ij}}\right] \\ &= \prod_{i=1}^{n}\left[\prod_{j=1}^{J}\left(\frac{\exp(\eta_{ij})}{1+\exp(\eta_{ij})} - \frac{\exp(\eta_{ij-1})}{1+\exp(\eta_{ij-1})}\right)^{y_{ij}}\right]\end{aligned}$$

where $\eta_{ij} = \tau_j + \beta X_i$ and $\eta_{ij-1} = \tau_{j-1} + \beta X_i$

Maximum likelihood estimates can be found by differentiating with respect to each of the unknown parameters, setting each of the *J+p* equations equal to zero and solving for $\hat{\beta}$. McCullagh states that the Newton-Raphson method with Fisher scoring converges rapidly even when the initial estimates are poor [29]. Since the same coefficients $\beta$ are assumed for all the cumulative logits, this effects are also independent of the cutpoints $\tau_j$ that chop the latent variable $Y^*$. The effect parameters are invariant to the choice of categories for $Y$ [4]. This feature makes it possible to compare estimates from studies using different response scales. The **deviance** works as a common goodness-of-fit measure, and the classic inferences using asymptotic assumptions can be performed with Wald statistics.

Hosmer and and Fagerland [15] developed a **goodness-of-fit** test for the proportional odds model. It is based on the Homer-Lemeshow test for binary logistic regression, which compares the sum of observed frequencies $O_{jk}$ versus the sum of predicted probabilities, partitioning the data in $G$ (usually 10) groups, for $K$ categories. For polytomous response variables, the Pearson chi-square statistic is computed from a $2K \times G$ table of observed and expected frequencies, so that the test's statistic.

$$X^2_{HL} = \sum_{j=1}^{g} \sum_{k=1}^{K} \frac{(O_{jk} - E_{jk})^2}{E_{jk}} \quad \sim \chi^2$$

Large values for the statistic (small p-values) would indicate evidence of lack-of-fit of the proportional odds model. However, the test performs poorly if too many continuous predictors are present, since the data is too spared. In addition, the test has less power than the Lipzitz test or the Pulkstenis and Robinson test [24]. SAS allows the implementation of the test with an extra command, *LACKFIT*, to the `LOGISTIC` procedure, and is implemented in R as well. Hosmer and Fagerland suggest using the three tests simultaneously, but the Lipzitz test is not always computable for small samples, and the PR tests requires categorical and continuous explanatory variables.

### 4.1.1 Marginal Effects: Changes in probabilities

A marginal effect measures the change in the probability of an outcome for a change in $x_k$, holding all the other independent variables constant at specific values. In the proportional odds model, the marginal change in the probability of outcome $j$ is computed as

$$\frac{\partial \, \mathbb{P}[Y = j|\mathbf{X}]}{\partial x_k} = \frac{\partial F(\tau_j - \mathbf{X}\beta)}{\partial x_k} - \frac{\partial F(\tau_{j-1} - \mathbf{X}\beta)}{\partial x_k} \qquad (4.1.6)$$

which is the slope of the curve relating to the *k-th* explanatory variable, holding the rest constant. The value of the marginal change depends on the value of $x_k$ where the change is evaluated. **Average Marginal Effects** (AME) compute the marginal effect of $x_k$ for each observation at its observed values $x_i$, and then computes the average of these effects.

$$\textbf{AME} = N^{-1} \sum_{i=1}^{N} \frac{\partial \, \mathbb{P}[Y = j|\mathbf{X}_i]}{\partial x_k} \qquad (4.1.7)$$

The AME is the mean of the marginal effect computed at the observed values for all observations in the estimation sample.

### 4.1.2 Proportionality Assumption

McCullagh introduced the model in 1980 and it is regarded as the most widely used strategy for ordinal categorical regression. However, it relays heavily on the assumption that the vector of regression coefficients $\beta$ is independent from the level of response. Other models such as the continuous ratio model or the adjacent category model also assume proportionality for different values on the explanatory variables. In the end, it is of interest to know if model (4.1.4) holds, or a more complex form is needed, like:

$$C_j = \begin{cases} \tau_j - \beta\mathbf{X} & \text{(Proportional Assumption)} \\ \tau_j - \beta_j\mathbf{X} & \text{(Dependent coefficients)} \end{cases} \qquad (4.1.8)$$

Where $C_j$ is the cumulative logit. The alternative assumes we have $J-1$ logistic lines, each with its own intercept and set of effects. Model (4.1.4) assumes that ALL the estimates are the same, for ALL different categories for each one of the predictors whereas the model with dependent parameters in (4.1.8) allows variation. Nonetheless, it is less parsimonious because there is now **p(c-1)** parameters to estimate.

Peterson and Harrell developed score and likelihood ratio tests for testing the parallelism among the logits. The score test is used in the SAS `LOGISTIC` procedure, but it can be somewhat unreliable under certain conditions. According to the authors, the test suffers if the number of observations at one of the levels of $Y$ is small relative to the total sample size, or there are continuous predictors. In addition, if the data is not sparse, the performance of the test tend to bee too liberal: small *p-values* and type I errors [34].

The more complex model has the structural problem that cumulative probabilities can be out of order at some settings of the predictors. Because of this, it is often not feasible to maximize the likelihood function for the alternative model. Thus, the score test comparing the models is more widely applicable than a likelihood-ratio test, or a Wald test, because the score teste evaluate the rate of change of the log likelihood only at the null hypothesis, under which $\beta_1 = \beta_2 = \ldots = \beta_{c-1}$ **agrestiAOCD**.

If the augmented model can be fitted, we can test the proportional assumption via the **Brant Test**. Brant proposed to test the difference between $(\tilde{\beta}_j, \tilde{\beta}_l) \quad j, l \in \{1, 2, \ldots, J-1\}$ of the regression applied to the response

$$Z_j = \begin{cases} 1 & y > j \\ 0 & y \leq j \end{cases}$$

with success probability $\dot{\pi}_j = \mathbb{P}[Z_j = 1] = 1 - \mathbb{P}[\pi_j \leq j]$ , satisfying

$$\log\left(\frac{\dot{\pi}_j}{1 - \dot{\pi}_j)}\right) = \tau_j - \beta_j \mathbf{X} \tag{4.1.9}$$

A Wald-type statistic of the form

$$X^2 = (D\tilde{\beta})'[D\hat{V}(\tilde{\beta})D')]^{-1}(D\tilde{\beta})$$

where $\tilde{\beta}$ is a matrix containing the individual coefficients from the individual maximum likelihood estimates for each level, and $D$ is a contrast matrix of $(J-2)p \times (J-1)p$ dimensions. The statistic will be asymptotically $\chi^2$ with $(J-2)p$ degrees of freedom under the null hypothesis. If the omnibus test is rejected, then the individual differences are statistically significant, proving a violation of the proportionality assumption [8].

Two drawbacks come from this test: there is a major loss of degrees of freedom if either the number of categories or the number of explanatory variables is large, and inspection of individual components of $X^2$ may not provide clear indication as to the nature of the discrepancy. Brant states that the logits on (4.1.9) may not be seen as scientific meaningful models, but as discretional alternatives helpful in validating the simpler proportional odds model.

# 4.2 Random Effects

Model (4.1.8) assumes independence among the observation, ignoring the longitudinal structure of the data. This may cause the standard errors to be significantly smaller, making inferences on the estimates less substantive. We will treat with the heterogeneity on the sample using a **Random Effects Cumulative Logit Model**.

$$\log\left(\frac{\mathbb{P}(Y \leq j\,|\mathbf{X})}{\mathbb{P}(Y > j\,|\mathbf{X})}\right) = u_i + \tau_m + \beta\mathbf{X} \tag{4.2.1}$$

Model (4.2.1) adds a random effect $u_i$ to the intercept term $\alpha_j$, for $i = 1, \ldots, N$ individuals or clusters, and $j = 1, \ldots, J$ categories of the response variable. It uses the same random effect for each cumulative probability. A subject with a relatively large $u_i$ has relatively large cumulative probabilities, and hence a relatively high chance of occurring at the low end of the ordinal scale. (4.2.1) can also be expressedmin terms of a latent linear response, where observed ordinal responses come from an underlying continuous latent variable $Y^*$, such that

$$y^*_{it} = Xit\beta + u_i + \varepsilon_{it} \tag{4.2.2}$$

and

$$Y_{it} = \begin{cases} 1 & \tau_0 \leq Y^*_{it} < \tau_1 \\ 2 & \tau_1 \leq Y^*_{it} < \tau_2 \\ \vdots & \\ J & \tau_{j-1} \leq Y^*_{it} < \tau_j \end{cases} \tag{4.2.3}$$

As in **Section 4.1** $\tau_j$ are cutpoints that map the latent variable into the original response $Y_{it}$. The Cumulative Logit Random Effects model makes the same assumption as in the linear RE model, that the additional random component is **uncorrelated** with the explanatory variables. However, for the ordinal response model, we assume that

$$u_i \sim N(0, \sigma_u^2)$$

$\sigma_u^2$ is unknown and is a parameter to be estimated. As the variance of the random effects increases, the correlation between two observation of the same cluster (or individual) also tends to increase. For each $t < s$, we get that

$$\text{Corr}(y^*_{it}, y^*_{is}) = \rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma^2} \tag{4.2.4}$$

This equals the proportion of the total residual variance that is due to the variability $\sigma_u^2$ in the random effect. The correlation is positive and increases as $\sigma_u^2$ increases, for fixed $\sigma^2$. Since we are using the logistic distribution to identify the distributions of the disturbances on $\varepsilon$, we get that $\sigma = \pi/\sqrt{3}$. In the case where $\sigma_u^2 = 0$, the correlation disappears, the observations behave as if they were independent and we get the pooled model (4.1.4). Lastly, $u_i$ are assumed to be independent from $\varepsilon_{it}$. Maximum marginal likelihood methods are used to estimate the parameters. For this solution, Gauss-Hermite quadrature is utilized to numerically integrate over the distribution of random effects. Essentially, the normal density is approximated by a discrete histogram with bars centered at the quadrature points. The approximation improves as the $q$ of quadrature points increases. Similarly, as $q$ increases, subsequent approximations for the ML parameter estimates and their SE values improve.[23]

### 4.2.1 Test for Random Effects

For testing $H_0 : \sigma_u = 0$, the asymptotic null distribution of the likelihood-ratio statistic has probability 0.5 at 0 and 0.5 following the shape of a $\chi^2$ with one degree of freedom. The test statistic value of 0 occurs when $\hat{\sigma_u} = 0$, in which case the maximum of the likelihood function is identical under $H_0$. When $\hat{\sigma_u} > 0$ and the observed test equals t, the *p-value* for this large sample test is $0.5\,\mathbb{P}[\chi^2_{(1)} > t]$, half the *p-value* that applies for $\chi^2_{(1)}$ asymptotic tests.

Failing to reject the null hypothesis could signal that the pooled model is more desirable since it makes less restrictive assumptions than the RE model.

## 4.3 Fixed Effects

Instead of assuming that the individual unobserved characteristics for each country are normally distributed, we can consider them as fixed parameters to be estimated in the proportional odds model. This technique, however, is not commonly used because of the **incidental parameter problem**. Estimating an additional set of intercepts to the regression decreases parsimony in the model, and when $T$ is fixed, the estimates become inconsistent as $N \to \infty$. In the linear regression model, the within transformation would help mitigate this eventuality, but we cannot apply the same procedure in this case. We still can compute the regression using a set of dummy variables, although one country will have to be omitted to avoid perfect multicollinearity. The model has the same form as in (4.2.1), so will have a set of 16 intercepts for each individual country, and 5 intercepts to account for the categories on the response variable.

Interpretations of the Fixed Effects parameters are not as immediate, since the conclusions have to be done with respect to the country that was dropped. Nevertheless, we can still compare the magnitudes and significance of the estimates.

### 4.3.1 Test for Fixed Effects

We can perform a Wald test to verify the joint significance of the Fixed Effects coefficients. Let $\hat{\beta}$ be the estimated coefficient vector and $\hat{\Psi}$ the estimated variance-covariance matrix. Let $Q\hat{\beta} = q$ denote the set of *q* linear hypotheses to be testes jointly. The Wald statistic is

$$W = (Q\hat{\beta} - q)'(Q\hat{\Psi}Q')^{-1}(Q\hat{\beta} - q)$$

The Wald statistic follows a $\chi^2$ distribution with $q$ degrees of freedom under the null hypothesis that

$$H_0 = u_1 = u_2 = \cdots = u_{N-1} = 0$$

Rejection of $H_0$ could indicate that Fixed Effects coefficients account for the unobserved heterogeneity and the FE model is preferable over the pooled proportional odds model.

## 4.4 Results and Analysis

A stepwise selection method based on the Akaiken Information Criterion (AIC) was applied to the regression to simplify it and obtain a more parsimonious model. In addition, a reduced model allows us to test the proportionality assumption using the Brant test. The variable selection method concluded that the variables to be dropped are GDP Growth, Current Account Balance and Gross Debt. This variables were also not significant at any level in the linear regression models for panel data, so it seems proper to remove them from this analysis. Two observations were dropped that were being perfectly predicted (probability of one in any of the categories). This observations are Venezuela on 2016 and 2015. The extraordinary inflationary rates, above 100% annually were creating predicted probabilities of 1, causing undefined odds ratio.

**Table 4.1** displays the Maximum Likelihood estimates for the parameters in terms of Odds Ratio, $e^{\beta_k}$. Cutpoints are expressed in the regular form, and don't have a meaningful interpretation besides adjusting the regression for the different categories. Average Marginal effects on the probability for all the outcomes are shown on **Table 4.2**. Marginal effects are interpreted as the average change in the probability of falling in one of the categories of the rating, for that specific predictor, holding the rest of the variables at constant values.

### 4.4.1 Pooled Estimation

The Pooled model is significant in all the variables. The Hosmer and Fagerland's statistic for lack-of-fit resulted not significant, but we have to consider that almost all the predictors are continuous and that the table of Observed vs Expected values is to sparse, so the test may perform poorly. All the coefficients in the pooled proportional odds model are significant, but as we mentioned in section (4.2), standard errors are misleading since we are assuming each observation is independently sampled. Nevertheless, at a first glance it seems that it makes a better fit to its panel data alternatives.

The variable with the strongest magnitude in the regression is the GDP per Capita. Then the Exchange Regime and Inflation have relative high impact compared to the External Indicators, External Debt to Exports and International reserves. Since GDP per Capita is transformed into logarithmic scale, then we can make the next interpretation of the odds ratio coefficient:

*Increasing GDP per Capita by 100% increases the odds of hav-*

***Table 4.1:*** *Parameter estimates for the Proportional Odds Model*

| | Pooled POLR | FE POLR | RE POLR |
|---|---|---|---|
| GDP per Capita | 2.2761*** | 81.99*** | 103.85*** |
| | (0.41) | (45.77) | (67.49) |
| Inflation | 0.8584*** | 0.9491** | 0.9600 |
| | (0.02) | (0.03) | (0.03) |
| External Debt / Exports | 0.9795*** | 1.0023 | 1.0031 |
| | (0.003) | (0.004) | (0.005) |
| External Reserves | 1.0143*** | 1.0237*** | 1.0102** |
| | (0.003) | (0.005) | (0.005) |
| Budget Balance | 0.9213*** | 1.3566*** | 1.4784*** |
| | (0.04) | (0.1062) | (0.1312) |
| Fixed Exchange Regime | 0.2892*** | 6.4312*** | 1.7058 |
| | (0.09) | (3.6858) | (1.14) |
| Cutpoint $\tau_1$ | -2.94 | 22.63 | 26.34 |
| Cutpoint $\tau_2$ | 1.01 | 28.28 | 31.83 |
| Cutpoint $\tau_3$ | 4.089 | 34.27 | 38.05 |
| Cutpoint $\tau_4$ | 7.89 | 42.57 | 51.58 |
| Cutpoint $\tau_5$ | 10.03 | 46.93 | 57.85 |
| LR test $^a$ | 326.58*** | 520.90*** | 74.74*** |
| Mcfadden's pseudo $R^2_{adj}$ | 0.4059 | 0.6573 | *(NA)* |
| H & F `Lack-of-Fit` | 37.57 | 10.68 | *(NA)* |

SE in parenthesis   *** Significant at 0.05 level   ** Significant at 0.10 level   $^a$ Wald Statistic for the RE model

*ing a higher rating by a factor of 2.28, keeping all the other variables constant.*

Like in the pooled model of Chapter III, doubling the level of income has a relative small impact in the odds of having a higher rating. For the rest of the variables, we will make the interpretation in terms of percentage change in the odds, meaning that a $\delta$ increase in the predictor will be associated a $100(\delta\beta - 1)$ percentage change in the odds of having a higher rating.

A standard deviation increase (18.4 percentage points ) in the **inflation rate** decreases the odds of having a higher rating by 93.8%. The coefficient for Budget Balance has the opposite sign, meaning that a standard deviation increase in the Budget Balance (3.58 percentage points) relative to GDP **decreases** the odds of getting a higher rating by 25.4%, holding all the other variables constant. The significance of this parameter is the lowest one, but the same situation occurred in the linear regression model. Estimates for the External Indicators have the expected sign and seem to have high impact as well. A standard deviation (75.5 percentage points) increase in External Debt to Exports, the odds of a better rating diminish by 79%, holding all the other variables constant. On the contrary, a standard deviation (73.2 percentage points ) increase in the level of International Reserves increases the odds of having a better rating by 182%. Increasing the percentage of Reserves has more than double the impact of increasing the same percentage in the level of external debt, so according to this model, rating agencies value more the precaution each country takes than the quantity of debt they acquire. Finally, having a Fix Exchange regime instead of a floating one decreases the odds of having a higher rating by 71.1 %.

We can also appreciate the impact of each variable to the marginal change in probability of an observation falling under each rating. The average predicted probabilities are listed on the last column of Table 4.2. The rating with the highest predicted probability is the 'BBB' rating. These probabilities are similar to the frequencies of the ratings, shown in (Chapter Data sample). There's a 54.6 probability of having a Speculative Grade ('BB' or lower) than the Investment Grade. The highest AME for GDP per Capita (GDPPC) is on the 'BBB' rating. Increasing the level of income translates into a better chance of getting an Investment Grade Rating, but once you

*Table 4.2: Average Marginal Effects for the Pooled, Fixed Effects and Random Effects Proportional Odds Model*

| POOLED | GDPPC | INF | EDE | RES | BUD | ER | OVERALL |
|---|---|---|---|---|---|---|---|
| $\mathbb{P}$[SD] | -1.60 | 0.30 | 0.04 | -0.03 | 0.16 | 2.13 | 5.10 |
| $\mathbb{P}$[B] | -4.87 | 0.91 | 0.01 | -0.08 | 0.49 | 8.13 | 18.44 |
| $\mathbb{P}$[BB] | -3.11 | 0.58 | 0.08 | -0.05 | 0.31 | 5.93 | 31.02 |
| $\mathbb{P}$[BBB] | 5.27 | -0.98 | -0.01 | 0.09 | -0.53 | -10.81 | 36.17 |
| $\mathbb{P}$[A] | 3.00 | -0.56 | -0.08 | 0.05 | -0.30 | -3.48 | 6.78 |
| $\mathbb{P}$[AA] | 1.31 | -0.24 | -0.03 | 0.02 | -0.13 | -1.79 | 2.47 |
| FIXED EFFECTS | GDPPC | INF | EDE | RES | BUD | ER | OVERALL |
| $\mathbb{P}$[SD] | -10.82 | 0.13 | -0.006 | -0.06 | -0.75 | -5.34 | 4.50 |
| $\mathbb{P}$[B] | -5.98 | 0.07 | -0.003 | -0.03 | -0.41 | -1.70 | 19.23 |
| $\mathbb{P}$[BB] | -9.71 | 0.12 | -0.005 | -0.05 | -0.67 | -2.77 | 30.73 |
| $\mathbb{P}$[BBB] | 17.96 | -0.21 | 0.01 | 0.09 | 1.24 | 5.71 | 35.85 |
| $\mathbb{P}$[A] | 3.68 | -0.04 | 0.0002 | 0.02 | 0.25 | 1.74 | 7.14 |
| $\mathbb{P}$[AA] | 4.88 | -0.06 | 0.0003 | 0.03 | 0.34 | 2.33 | 2.53 |
| RANDOM EFFECTS | GDPPC | INF | EDE | RES | BUD | ER | OVERALL |
| $\mathbb{P}$[SD] | -8.30 | 0.07 | -0.006 | -0.02 | -0.70 | -1.05 | 9.69 |
| $\mathbb{P}$[B] | -6.87 | 0.06 | -0.005 | -0.01 | -0.58 | -0.84 | 14.75 |
| $\mathbb{P}$[BB] | -6.37 | 0.06 | -0.004 | -0.01 | -0.54 | -0.75 | 24.70 |
| $\mathbb{P}$[BBB] | 14.12 | -0.12 | 0.009 | 0.03 | 1.19 | 1.73 | 43.78 |
| $\mathbb{P}$[A] | 5.75 | -0.05 | 0.004 | 0.01 | 0.48 | 0.71 | 5.69 |
| $\mathbb{P}$[AA] | 1.67 | -0.01 | 0.001 | 0.004 | 0.14 | 0.21 | 1.40 |

* All probabilities are in a $\times 10^2$ scale

have it, getting up the scale becomes not that easy, since the average marginal effect decreases in the 'A' and 'AA' rating. On the other hand, on average and holding all the variables constant, increasing GDP per Capita diminish the chances of having a Speculative Grade rating, but not on the Default Rating.

Having a Fixed Exchange Regime (ER) decrease on average the probability of having the Investment Grade, than a Floating Exchange Regime. The AME for the 'BBB' rating on the probability is of 11, so having tight and restrictive monetary policy reduces the chances of having a high-end rating. It also translate into an average marginal effect of 0.08 in the outcome of a 'B' rating. As we suspected, Rating Agencies favour countries that allow their exchange regime to float freely or have a diminished involvement on the market.

The AMEs for the Inflation Rate (INF) are shown on the second column of Table (4.2). On average, higher inflation rates increase the probability of having a 'B' by 0.0091. Simultaneously, it decreases the probability of having a 'BBB' rating by almost the exact factor, holding all variables constant. The average marginal effects for the External Debt (EDE) and the Reserves (RES) are less than 0.001, nonetheless they are significant. On average, increasing the level of external debt to exports has the highest impact on the 'BB' category and 'A' categories, keeping the rest constant. Whereas the AME for Reserves is greatest for the 'BBB' rating. The marginal effects for the Budgetary Balance are on the contrary of what we could've expected. Since the estimate of the coefficient resulted with a negative sign, it results that on average, a Budgetary Surplus is associated with a 0.0053 decrease in the probability of having a 'BBB' rating. It doesn't make much sense, but we have to keep in mind that the parameter was nearly not significant and that standard errors are not that reliable, so the interpretation may not be fully sustained. Overall the average marginal effects and the change in odds ratios are consistent and balanced, but seem to have low magnitudes, as we will know compare this results when the heterogeneity of the sample data is considered in the regression instead of ignoring it as the Pooled model does.

### 4.4.2 FE Estimation

The Fixed Effects model for linear regression made in our judgement the best adjustment and comprehensible fit. However, it doesn't seems to be the case for the proportional odds model. Although the model is significant, the extra set of intercepts seems to cause too much distortion and is making the estimates have not much sense. Several issues arrived with the full set of dummy variables for all the countries, and a few of the dummies had to be dropped in order to get consistent results. First, the dummy variables for South Africa and Thailand had to be excluded since this countries present no variation in the response variable. The dummy variables of China and Philippines had to be dropped because 34 observations were being perfectly predicted and issues with multicollinearity were present. Cutpoints have significantly a larger scale, yet the impact may be offset by the dummy variables, which are shown in Table **(4.3)** . The estimate coefficient for the External Debt to Exports have a very low impact on the regression and is not statistically significant. Also, the sign for the External Debt is positive, meaning more debt equals to better odds of having a higher rating. This also happens with the Fixed Exchange Regime, but this estimate is in fact significant. According to this model, having a Fixed Exchange Regime improves your chances of having a Higher rating than having a Floating Exchange Regime, holding all the other variables constant.

The odds ratio impact for the level of income seems more reasonable. Increasing **GDP per Capita** from $3,000 USD to $6,000 increases the odds of having a higher rating by a factor of 82, which is immensely higher than in the Pooled model. Colombia went in 2000 from a GDP per Capita of $2,455 to $8,103 in 2013, and that leap earned it the Investment Grade with a 'BBB' from 'BB'. Similarly, Turkey went from having a GDP per Capita of $3,053 in 2001 to $10,817 in 2016, jumping from the 'B' rating, coming out of an economic crisis to the 'BBB' rating.

The **Budget Balance** has the correct sign and a higher significance than in the pooled model. Holding the rest constant, a standar deviation increase (3.58%) in the Budget Balance improves the chances of a higher rating by 193 %. The Phillipines reduced its deficit from 3.6 % of GDP in 2001 to a surplus of 0.9 % of GDP in 2014 and it improved its rating from 'BB' to 'BBB' **The Inflation** rate is influencing the regression in the correct direction, so a standard deviation increase (8.4%) reduces the odds of having a higher rating by 35 %. Brazil acquired the Investment Grade in 2008 with a 'BBB' rating with an inflation rate of 3.6% and a constant rate of growth, but went back to the Speculative grade in 2014 when it fell in recession and had annual inflation rate of 9.03 %. A standard deviation increase in the level of **International Reserves** (75.5%) relative to exports increases the chances of a higher rating by 454%. India had a significant growth period from 2000 to 2007, and increased its level International reserves from 40% to 135 %, and received the Investment Grade with a 'BBB' rating from 'BB'.

Nonetheless, the impact of the **Exchange regime** seems to be non-realistic and may be influenced by the observations of China. According to the model, keeping everything constant and having a Fixed Exchange regime increases the chances of having a higher rating by 543%. China has a very tight and strict monetary policy, and it has been stated that it manipulates its currency to favour its cause, but it has the highest ratings of the sample, so it may be leveraging the regression to that side.

AMEs for GDP per Capita have a strange pattern, since the probability seems to be impact negatively the most for the low ratings in the Speculative Range, but influence the most in changes to the Investment Grade and the 'BBB' rating. An increase in income is related to a higher average increase in probability for the 'BBB' rating than for the 'A' rating. The marginal effects of the External Debt to Exports and Reserves are minimal. There seems to be something odd with the observation from the 'A' rating overall, since for all the explanatory variables, the AME is significantly smaller than the rest of the categories. Even for the 'AA' which only has 7 observations, the average marginal effects seem to be more in line with the previous pooled adjustment and the overall frequency of the observations than the ones from the lower category.

#### Fixed Effects Parameters

***Table 4.3:*** *Individual effects coefficients*

| | Intercept | Std. Er. |
|---|---|---|
| Argentina | -18.03 *** | 2.21 |
| Venezuela | -12.73 *** | 1.69 |
| Turkey | -8.01 *** | 1.30 |
| Russia | -7.57 *** | 1.26 |
| Indonesia | -6.54 ** | 1.05 |
| Brazil | -6.43 *** | 1.19 |
| Pakistan | -6.22 *** | 1.19 |
| Egypt | -5.09 *** | 1.11 |
| Colombia | -4.24 *** | 0.97 |
| Mexico | -3.60 *** | 1.11 |
| India | 3.63 *** | 1.04 |
| Malaysia | 4.35 *** | 1.07 |

*** Significant at 0.05%

We can test the joint significance of the FE coefficients int the proportional odds linear regression. In this case, we got 12 coefficients since we had to drop out four countries to compute the regression. Interpretations have to be done relative to the four countries that where dropped, Thailand, South Africa, China and the Philippines. The dummy variables for these countries were excluded since Thailand and South Africa have the same rating for the whole sample, which causes problems with the estimation of their ML coefficient. The dummy variable for China was avoided as well since 34 observations were being perfectly predicted. In the same manner, including the dummy variable for the Philippines presented an enormous standard error in the estimate, possibly related to multicollinearity with the Fixed Echange Regime variable, so it was dropped.

### 4.4.3 Pooled vs FE

The Wald Statistic to test the Fixed Effects assumptions is $W = 76.3$ which follows a Chi squared distribution with 12 degrees of freedom. The p-value associated with the test is $\approx 0$, so there's is enough evidence that the FE coefficients play a significant role in the regression. There are not many conclusions we can derive, but the order and magnitude of the coefficients is almost the same as in the linear regression model with Fixed Effects.

Overall, the Fixed Effects model in the proportional odds case fits well as in the panel data model for linear regression. We got only one non-significant parameter (External Debt) and the impact on large changes in the predictors reflects larger changes in the odds, compared to the Pooled model. Treating the country-specific effects as parameters to be estimated may be conflictive when trying to obtain the ML estimates, so we were forced to simplify the model and drop four dummies to avoid multicollinearity issues and perfect predictions. Few observations in the highest category and the lack of variation in the response variable for 2 countries and overall the oversaturation in the regression denied us the chance to compare the individual effect coefficients to the panel data model, but grossly speaking we can say that this model made a better job in comparison to the Pooled Proportional Odds model. However, we have to keep in mind that we watered down the regression by dropping models to get consistent results. We will now analyse the results from the Random Effects Model.

### 4.4.4 RE Estimation

The Random Effects model makes the assumption that the unobserved individual effects are uncorrelated with the predictors, and that they follow a normal distribution. Because it is a more complex model to compute, there are not additional tests to compare the Fixed Effects models and the Random Effects model as in the linear regression with the Hausman's test.

There are also no available methods to test goodness-of-fit, at least not in software at our disposal.

First, we will analyse the coefficients in **Table (4.1)**. The **GDP per Capita** estimate has the highest magnitude of the three models. Increasing the level of income from \$3,000 current USD to \$8,154.85 keeping everything constant increases the odds of having a higher rating by a factor of 103.85 . This parameter is statistically significant, so if the RE proportional odds is the true model, then almost all of the predictability could be coming from the level of income of each country. The other Economic Indicator, **Inflation rate**, is not significant but has the proper sign, meaning that increasing by a standard deviation (8.4 %) the inflationary rate, keeping everything constant, will decrease the odds of having a higher rating by 33.5% percent inflation bye diminish the probability of having a better grade.

**External Debt to Exports** is not significant again and has the wrong sign, so an increase in this variable reduces the odds of having a higher rating. For the level of International Reserves, a increase of 73.1 % holding everything constant will improve the odds of a higher rating by 109 %. This percentage change is higher than in the Fixed Effects model, but not higher than in the Pooled model. The **Budget Balance**, nonetheless, do has the expected significant impact on the regression, so that increasing the country's fiscal balance by a standard deviation (3.5%) would affect positively the odds of getting a higher rating by 297%, holding the other variables constant. This percent change is the largest seen in all the models on variables that are not GDP per Capita. Besides from the level of income, according to the Random Effects model the highest impact in the chances of higher ratings come from a balanced fiscal budget. The estimate for the Exchange Regime is significant but has the incorrect sign, indicating that having a **Fixed Exchange** regime could increase the odds of improving the rating by 77%.

Average Marginal Effects seem to behave more reasonably than in the Fixed Effects adjustment. On average, increase on GDP per Capita increases the probability of having a 'BBB' by a factor of 14, holding everything else constant. The marginal effect for this rating is the one with the highest magnitude, so it could seem that increases on the level of income increases the chances of having an Investment Grade rating, but that is not enough if the country aspires for a higher rating. Also, an increase in GDP per capita on average decreases the probability of having a 'BB' by 6.3, 'B' by 7.4 and 'SD' by 7.9.

On average, an increase in inflation is associated with a progressive increase in probability for the Speculative Grade ratings, with the 'SD' having a factor of 0.1. The marginal effects are almost imperceptible if we increase or decrease the External Indicators. A positive fiscal balance increases on average the probability of receiving a 'BBB' by a factor of 1.2, *ceteris parabus*. As we mentioned, since the coefficient for the Fixed Exchange Regime has the opposite expected sign, the probability of having one of the Investment Grade ratings is on average higher than having the low-end ratings. Overall, given all the predictors at constant values, the rating with the highest probability is the 'BBB' rating, then the 'BB' rating and then the 'B' rating. Also, the probability associated with the 'SD' rating is the highest of all the three proportional odds models.

### 4.4.5 Pooled vs RE

The likelihood ratio test for the variance term $\sigma_u^2$ on the assumption that the unobserved characteristics follow a normal distributions was performed, and showed a statistic of $LR = 161.29 \sim \chi^2_{(1)}$. The p-value associated with the statistic was virtually zero, so there's enough evidence to conclude that there's a positive correlation between observations $y_{it}, y_{is}$. The estimate for the variance is

$$\hat{\sigma}_u^2 = 59.61$$

with a 95% confidence Interval of [23.69, 149.98]. Given that value, we can also estimate the positive correlation as

$$\begin{aligned}\hat{\rho} &= \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \sigma} \\ &= \frac{59.61}{59.61 + \pi^2/3} \\ &= 0.9477 \end{aligned} \tag{4.4.1}$$

This test supports the idea that the Random Effects model is preferred over the Pooled proportional odds model, since the first one actually takes into account the additional heterogeneity provided by the way the data was structured.

### 4.4.6 Validating the Proportional Odds Assumption

As discussed in Section 4.1.2, we would like to know if the assumption that the cumulative logit shares the same $\beta$ coefficients holds. We get that the statistic from the Brant test is

$$\begin{aligned} X^2 &= (D\tilde{\beta})'[D\hat{V}(\tilde{\beta})D')]^{-1}(D\tilde{\beta}) \\ &= 135.7701 \end{aligned} \tag{4.4.2}$$

The Wald-like statistic follows a chi squared distribution with 24 degrees of freedom. The associated p-value is minimal, so according to the test, there's significant discrepancies among the model that assumes the effects to be the same independently of the category on the response, and the model that allows a different set of coefficients for each logit. To apply the test we used the estimates of the Pooled Regression, but there's a high chance the results of the test could have been the same if we instead used the estimates of a different model.

But how critical the violation of the proportional assumption really is? It is really not clear how to proceed after the rejection of the null hypothesis. A statistical rejection need not imply that the ordered logit estimates are poor estimates of the true response probabilities. If we specify the unrestricted model,

$$C_j = \tau_j - \beta_j \mathbf{X}$$

We are just left with a bunch of $J_1$ unconnected binary response models of cumulative logits, and it is not clear what we would learn in the end. Alternative approaches would be the **adjacent categories model**, which uses the probability of an observation belonging to a *j-th* observation relative to the $j+1$ category. More complex models could include the partial proportional odds, where some parameters are fixed and some are allowed to fluctuate with depending on the category.

The biggest drawback of the alternative models is that there are almost no extensions to panel data. In the end, the proportional odds model allowed us to compare the cases when we include in the modelling the unobserved country-specific characteristics and when we don't. We couldn't be able to compute the alternative models using a panel data structure, or we could do so but in the end we could still be short, since we are not taking advantage of the information. The test suggests using a different approach, but because of the nature of our experiment and since we want to compare this results with the linear regression models as well, we believe the POLR model is still the best candidate to model the ratings.

## 4.5 Discussion

### 4.5.1 Predictions

We can obtain the category predictions for the Pooled model and the Fixed Effects model given the probabilities obtained from fitting the models. The confusion matrices are shown in **Table (4.4)**. Naturally, the prediction is assigned to the category with the highest probability depending on the model. Since the Random Effects model is fitted with the `xtologit` command of STATA and can not predict automatically, we will leave this predictions as further work. We can estimate the accuracy of prediction for each model, summing all the elements in the diagonal of each confusion matrix and divided by the total number of observations used. We get that the accuracy for the Pooled model is

$$\text{Accuracy}_{Pooled} = 61.48\%$$

and for the Fixed Effects model

$$\text{Accuracy}_{FE} = 79.26\%$$

. **Figure (4.1)** displays the observations against their predicted values for the Pooled model (4.1a) and the Fixed Effects model (4.1b). Both models apparently struggle in the extreme categories, Default and 'AA'. The FE model sends the majority of the observations (9 out of 3) of the 'SD' rating to the adjacent category, nonetheless, it doesn't predicts any observation as 'BB', like the pooled model. The FE predicts correctly better in the 'B', 'BB' and 'BBB' ratings. The pooled model struggles in the 'A' rating and underrates 14 of 19 observations to 'BBB', while the Fixed Effects model only misses in 3 observations. Finally, the FE predicts correctly only 3 out of 4 observations of the highest observed rating, 'AA', whereas the pooled model only predicts correctly one observation.

The Proportional Odds Model allows us to identify clearly which category the regression is predicting, compared to the panel data regression which calculates on a continuous scale and it is not clear what level each prediction is taking. This is a huge uptake since we can clearly determine that in prediction terms, the Fixed Effects also outperforms the pooled model, signaling that ignoring the panel data structure of the data provides less accuracy than including it using dummy variables. However, like in the case of the linear regression, the extra set of parameters naturally improves artificially the accuracy. Intuitively, the Random Effects probably fits better than the pooled model, since the estimated coefficients are similar to the FE model.

***Table 4.4:*** *Confusion matrices for the Pooled and Fixed Effects Proportional Odds Predictions*

| Pooled Model | SD | B | BB | BBB | A | AA | Predicted |
|---|---|---|---|---|---|---|---|
| AA | 0 | 0 | 0 | 0 | 3 | 1 | 4 |
| A | 0 | 0 | 0 | 2 | 2 | 4 | 8 |
| BBB | 0 | 0 | 19 | 79 | 14 | 2 | 114 |
| BB | 1 | 21 | 53 | 16 | 0 | 0 | 91 |
| B | 7 | 27 | 11 | 0 | 0 | 0 | 45 |
| SD | 4 | 4 | 0 | 0 | 0 | 0 | 8 |
| FE Model | SD | B | BB | BBB | A | AA | Predicted |
| AA | 0 | 0 | 0 | 0 | 2 | 3 | 5 |
| A | 0 | 0 | 0 | 1 | 16 | 4 | 21 |
| BBB | 0 | 0 | 11 | 87 | 1 | 0 | 98 |
| BB | 0 | 9 | 64 | 9 | 0 | 0 | 82 |
| B | 9 | 41 | 8 | 0 | 0 | 0 | 58 |
| SD | 3 | 2 | 0 | 0 | 0 | 0 | 5 |
| Observed | 12 | 52 | 83 | 97 | 19 | 7 | 270 |

## 4.5.2 Relevant Predictors

We tried to simplify the model from the previous chapter by removing three non-significant variables, Gross Debt of the General Government, Current Account Balance and GDP Growth Rate. Still, the models showed that the most important factor is the **GDP per Capita**. It looks like if the level of income plays the biggest role in determining if weather a country gets an Investment Grade or not. The AMEs for the 'BBB' rating were in the three cases the ones with the highest impact. We get the same result for each model, GDP per Capita influences the most if a country can achieve the lowest of the Investment Grade ratings, or if its debt will remain at the mercy of speculators. The estimate in the pooled model seems rather small compared to the Fixed Effects model and the Random Effects model, since the magnitude indicates that almost tripling the level of income only affects the odds by a factor of 2, compared to 81 and 103 with the other alternatives. The other relevant indicator was the **Budget Balance of the General Government**, since it was significant in the three models, and showed also relevant average marginal effects. It may be that prudent fiscal policies and a balanced budget signals a better creditworthiness to the rating agencies, so it influences positively the grade the country gets. However, there are countries which presented and increasing deficit overall and still man-

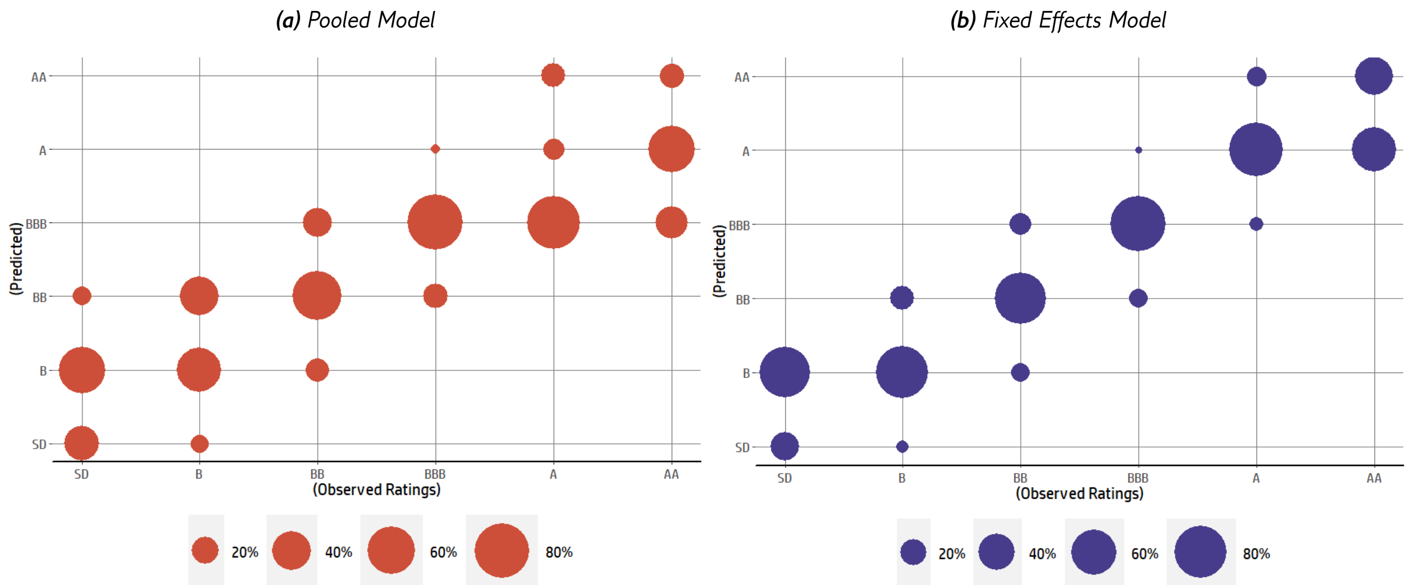



***Figure 4.1:*** *Observed vs Predicted ratings for the Proportional Odds Models*

aged to obtain higher ratings. South Africa increased its fiscal deficit from 0.12% in 2005 to 4.75 % in 2015, but managed to stay in the Investment Grade with a 'BBB' rating. Same story with Thailand, which presented a surplus of 1.98% in 2003, but suddenly fell with a deficit of 2.21% six years later in 2009.

The **Fixed Exchange Regime** was significant in the pooled and FE model, but not in the Random Effects one. However, it has the opposite sign as expected when the regression incorporates the unobserved heterogeneity, signifying that a Fixed Exchange Regime increases the odds of a higher rating, instead of a Floating regime which allows more fluctuations by the market. This may be caused as mentioned previously by the fact that the only 'AA' observations come from China. Other Emerging Markets countries where considered to have a more robust sample, like Poland, Qatar and Saudi Arabia. This sovereigns have high ratings, however, they could not be included since they don't have public data available for the External Indicators. The **External Debt to Exports** didn't resulted significant in the panel data models. The magnitude of their impact seems to be minimal and only in the pooled regression the estimates resulted significant. Also in this model, the sign for the coefficient was correct, but not in the others. From these approaches, it seems like the External Debt play a lesser role than the economic indicators. However, the **International Reserves** was significant in the three regressions, and a high increase in the percentage level of reserves improved massively the odds of having a higher rating. This indicates that according to these models, S& P values more how well the country is prepared and saves to pay the debt than the actual level of liabilities. So, as long as the country has enough money to pay, it seems as it doesn't matter how much leveraged it is.

An interesting remark is than on Table (4.2), there is a clear distinction between the marginal effects for categories with the Investment Grade and with the Speculative Grade. For almost all the explanatory variable in the three models, the AME changes its direction from the 'BB' rating to the 'BBB', which is the line between Investment and Speculative Grade. It may indicate that given our sample, the regression clearly differentiates the characteristics and impact of the predictors among the 'good' and the 'bad' observations. It could be relevant to experiment how well a model fits only considering a binary response, whether the country has the Investment

Grade or not, and if the magnitudes and impacts of the variables are similar, or they have different magnitudes and order. By the results obtained from the different tests, we can conclude that the GDP per Capita is still the most relevant factor regarding the sovereign ratings for Emerging Markets. At least, we can tell that it strongly determines whether a country gets the Investment Grade or not.

### 4.5.3 Fittest Model

Trying to choose the adequate model is trickier in this case than in the linear regression case. On one hand, the Fixed Effects and the Random Effects models, seem to outperform the pooled model, because of the magnitudes and significance of the parameters, as well because the statistical tests for each one vs one comparisons resulted significant in both cases. Nonetheless, we don't have a way of comparing the FE and RE models to one another since we don't have a computationally extension of the Hausman's test from the panel data regression models. The Pooled models doesn't incorporate any of the unobserved characteristics to the ordinal regression. Clearly this is not feasible, and leads to incorrect standard errors and estimates are not ideal since we are using Maximum Likelihood Individually the panel data have some drawbacks.The Random Effects model makes the stronger assumption of a random normally distributed intercept, which seems unrealistic. That assumption could make more sense if the experiment consisted of testing different drug reactions to a sample taken from a large population. In our case, it is not that realistic since the population of Emerging Markets with sovereign ratings is limited. However, the Fixed Effects model suffered from multicollinearity if all the dummy variables were included, so it had to be watered down to obtain reasonable good results.

Hence, we conclude that because we have several limitations and are not capable of testing the FE and RE models, both seem to perform well in the Proportional Odds Logistic Regression. Both the odds ratios coefficients and the Average Marginal probabilities are closely similar, and there are only a few differences. The accuracy of predictions for the RE model could be tentatively similar. In the end, both models suffer from limitations, but at least they treat the data correctly and are preferred over the pooled model. It is up to determine which drawback is less important, the fact that we cannot include all the dummy variables to avoid perfect predictions and multicollinearity, or to assume normality in a random intercept which is not correlated to the predictors

### 4.5.4 Considerations

The proportional odds model resulted practical and showed a good accuracy level. We mentioned at the beginning of the chapter that treating the ratings as a categorical response was the proper way of the modeling, since in the linear regression case we assumed the ratings to be a continuous. Still, we encountered with almost the same results, in a more complex modeling technique. In order to get a fitter ordinal logistic model, we necessarily need to simplify it either by reducing the number of categories in the response variable, or reducing the number of predictors. The average marginal effects showed that the largest differences come between the 'BBB' and 'BB' rating. Maybe one alternative could be to reduce the model just to a dichotomous response: Investment Grade vs Speculative Grade. Binary logistic models are well studied and implementation has become very popular. Nonetheless, we could be scarifying a lot of information regarding the ratings.

On the other hand, literature suggests that the ordinal response model don't fit well with continuous variables that are widely sparse [21]. One possible alternative could have been categorizing the predictors, turning them into categorical variables, allowing the model to work with more clustered data, and make a simpler fit. That could've also allowed interaction terms and we could have analyzed for deeper associations among the indicators and their effects on the ratings. Nonetheless, we could have to find optimal threshold to discretize the continuous variables, and that might also mean into scarifying variability and leading to over fitting of the model. Treating the ratings as a categorical response necessarily requires a simpler and more parsimonious model. In order to achieve that, we need to sacrifice some information, and in the end, it could make the comparisons between the linear regression and the POLR not doable.

# CHAPTER V

# SUMMARY AND CONCLUSIONS

We wanted to identify what are the indicators that influence sovereign ratings and to understand what credit rating agencies value the most. We tested two different techniques, linear regression and ordinal logistic regression, with similarities and discrepancies in some aspects. Both models proved to work better and showed statistical evidence that including individual heterogeneity is better, i.e., pooling the observations is not preferable. The **Fixed Effects** approach is a better alternative than the Random Effects one given the nature of the long panel dataset. It is economically more realistic to treat the individual unobserved characteristics as parameters to be estimated than random intercepts. We also can notice that the magnitudes and impact of the coefficients were practically the same for almost all the variables, independently of linear regression or proportional odds model. The later may seem preferable since it deals with the ratings correctly, as a categorical ordinal variable. We also gained more information since we were able to determine that the model identifies the line between Investment Grade and Speculative Grade, as in [26]. As well, we get precise predictions, in contrast with linear regression and the continuous scale. However, in order to get better results, we had to simplify the model by dropping explanatory variables. In most of the literature that include panel data modeling, ordered probit regression is used. We proved that using the ordinal logistic regression is just as valid, and got results that are closely similar.

We were able to identify the most relevant macroeconomic factors that influence in all cases the regression. Using a stepwise selection method based on Akaiken's Information Criterion (AIC) and the level of significance, without a doubt we can confirm that the most significant indicator in sovereign ratings for emerging markets is **GDP per Capita**. The estimate of the coefficient was significant in all of the six total models, and it was the variable with highest impact in Fixed Effects model for linear regression and the proportional odds model. As we expected, low income countries are punished the most and are assigned lower ratings. Countries should take close looks into the level of income and boost growth if they want to improve their sovereign ratings. The **Inflation Rate** was constantly significant and large changes can lead to downgrades, according to the models. The levels of **External Debt** were significant in the linear regression approach, but not in the POLR model. **Foreign Reserves**, however seemed to have a bigger impact in magnitude and significance. Rating agencies apparently value more hoe well a country is hedged against contingencies, than the quantity of debt it is acquiring. **Budget Balance** is one of the most important, with significance in most cases and high impact on the regression. A controlled and reasonable deficit help to mitigate downgrades and boost the chances of having a better rating. The **Exchange Regime** showed mix results, and was significant in only half of the models. We mentioned earlier that one factor that might influence the impact of this variable is the fact that China, which is the country with the best performing ratings, has a Fixed Exchange Regime. We cannot conclude decisively if the variable had a significant effect, or it suffered from the sample.

These results are in line with what's been done previously in the literature. Our main goal was to identify this economic factors, and to test the differences among the two econometric modelling techniques. Emerging markets are interested in improving their role in international markets and lower borrowing costs to finance domestic projects. Governments should put extra attention to the level of indebtedness and to prepare in case of contingencies. A prudent fiscal balance is almost surely going to improve their chances of having a better rating, but overall the level of income will be the key factor into where the CRAs allocate their sovereign ratings.

Further work could include using a broader data sample to model the ratings using a more refined scale, considering the additional classifiers (+) and (-). This could also allow two use a logarithmic or exponential scale that takes advantage of the fact that the transitions from 'BB+' to 'BBB-', the line of the Investment Grade, are more significant than movements in between levels. Also, adding governance factors and more variables that demonstrate more qualitative factors could improve the regression in both cases. If predictions and classifications is of interest, neuronal networks models or classification trees would work, now that the most relevant indicators have been identified. Since there is a lot of speculation to when a downgrade is coming and the short term impact on the markets is considerable, a model that is able to predict the timing of the downgrade could proof useful. Papers have studied the impact and differences among the rating agencies opinion, so it could also be interesting if the indicative factors change significantly if instead of using Standard and Poor's ratings, we used Moody's or Fitch's.

Sovereign ratings are an essential component of the contemporary financial ecosystem, and emerging markets tend to depend greater into the opinions of CRAs than developed economies. This work intends to incorporate into the literature a fresh review of classical econometric models, using a distinct sample. As long as there are countries that are hoping to improve the quality of life of its constituents, econometricians will contribute by constructing models that can light the way; simplifying and signalling areas of opportunity so that everybody can improve.

# PROLOGUE

This work intends to serve as a future guide for undergraduate students as to how to constitute a dissertation for Visiting Students that work in foreign universities. It also intends to serve as an example of how to oblige with the mathematical rigor that the School of Science demands, but also how to analyze and make economical and social interpretations of those results, enriching the conclusions and refortifying the reach of the investigation. In the same spirit, I would like to help students that haven't written anything in English, as it is vital and brings big rewards academically and professionally.

I would like to encourage current students to pursue this kind of opportunities since it opens up new frontiers for academic growth. Always take the initiative and ask and seek, there are a lot of kind and supportive persons willing to help in the University.

# BIBLIOGRAPHY


[1] P. Abdou, Z. Zaazou, *et al.*, "The egyptian revolution and post socio-economic impact," *Topics in Middle Eastern and African Economies*, vol. 15, no. 1, 2013.

[2] A. Afonso, "Understanding the determinants of sovereign debt ratings: Evidence for the two leading agencies," *Journal of Economics and Finance*, vol. 27, no. 1, pp. 56–74, 2003.

[3] A. Agresti, *Analysis of ordinal categorical data*. John Wiley & Sons, 2010, vol. 656.

[4] A. Agresti and M. Kateri, *Categorical data analysis*, Second Edition. Springer, 2011.

[5] B. Baltagi, *Econometric analysis of panel data*. John Wiley & Sons, 2008.

[6] ——, *Econometrics*, 4th. Springer Science & Business Media, 2007.

[7] A. K. Barnett-Hart, "The story of the cdo market meltdown: An empirical analysis," PhD thesis, Harvard University, 2009.

[8] R. Brant, "Assessing proportionality in the proportional odds model for ordinal logistic regression," *Biometrics*, vol. 46, no. 4, pp. 1171–1178, Dec. 1990. [Online]. Available: `https://www.jstor.org/stable/2532457`.

[9] R. Cantor and F. Packer, "Determinants and impact of sovereign credit ratings," *Economic policy review*, vol. 2, no. 2, 1996.

[10] ——, "Sovereign credit ratings," *Current issues in economics and finance*, vol. 1, no. 3, 1995.

[11] F. C. I. Commission, *The financial crisis inquiry report: Final report of the national commission on the causes of the financial and economic crisis in the united states*, 2010.

[12] E. deHaan, "The financial crisis and corporate credit ratings," *The Accounting Review*, vol. 92, no. 4, pp. 161–189, 2017.

[13] L. H. Ederington, "Classification models and bond ratings," *The Financial Review*, vol. 20, no. 4, p. 237, 1985, issn: 0732-8516.

[14] O. Erdem and Y. Varli, "Understanding the sovereign credit ratings of emerging markets," *Emerging Markets Review*, vol. 20, pp. 42–57, 2014.

[15] M. W. Fagerland and D. W. Hosmer, "A goodness-of-fit test for the proportional odds regression model," *Statistics in medicine*, vol. 32, no. 13, 2013, issn: 0277-6715.

[16] J. J. Faraway, *Extending the linear model with R : generalized linear, mixed effects and nonparametric regression models*. Boca Raton : Chapman & Hall/CRC, 2006.

[17] M. Feldstein, “Argentina’s fall: Lessons from the latest financial crisis,” eng, *Foreign Affairs*, vol. 81, no. 2, pp. 8–14, 2002, issn: 00157120.

[18] N. Gaillard, “Credit rating agencies and the eurozone crisis: What is the value of sovereign ratings?” *VoxEU. org*, vol. 9, 2013. [Online]. Available: `https://voxeu.org/article/credit-rating-agencies-and-eurozone-crisis-what-value-sovereign-ratings`.

[19] W. H. Greene, *Econometric analysis*, 4th. Pearson, 2000.

[20] D. Gültekin-Karaka, M. Hisarcıklılar, and H. Öztürk, “Sovereign risk ratings: Biased toward developed countries?” *Emerging Markets Finance & Trade*, vol. 47, pp. 69–87, 2011, issn: 1540496X, 15580938. [Online]. Available: `http://www.jstor.org/stable/23047443`.

[21] F. E. Harrell Jr., *Regression modeling strategies : with applications to linear models, logistic and ordinal regression, and survival nalysis*, eng, Second Edition, ser. Springer Series in Statistics. Springer, 2015.

[22] J. A. Hausman, “Specification tests in econometrics,” *Econometrica*, vol. 46, no. 6, pp. 1251–1271, 1978, issn: 00129682, 14680262. [Online]. Available: `http://www.jstor.org/stable/1913827`.

[23] D. Hedeker and R. D. Gibbons, “A random-effects ordinal regression model for multilevel analysis,” *Biometrics*, pp. 933–944, 1994.

[24] D. W. Hosmer and S. Lemeshow, *Applied logistic regression.* John Wiley & Sons, 2000, p. 398.

[25] E. Ilzetzki, C. M. Reinhart, and K. S. Rogoff, “Exchange arrangements entering the 21st century: Which anchor will hold?” National Bureau of Economic Research, Working Paper 23134, Feb. 2017. doi: `10.3386/w23134`. [Online]. Available: `http://www.nber.org/papers/w23134`.

[26] L. Jaramillo and C. M. Tejada, “Sovereign credit ratings and spreads in emerging markets; does investment grade matter?” eng, *International Monetary Fund Working Paper*, 2011. [Online]. Available: `https://ideas.repec.org/p/imf/imfwpa/11-44.html`.

[27] R. S. Kaplan and G. Urwitz, “Statistical models of bond ratings: A methodological inquiry,” *The Journal of Business*, vol. 52, no. 2, pp. 231–261, 1979, issn: 00219398, 15375374. [Online]. Available: `http://www.jstor.org/stable/2352195`.

[28] J. S. Long and J. Freese, *Regression models for categorical dependent variables using Stata.* Stata press, 2006.

[29] P. McCullagh, “Regression models for ordinal data,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 42, no. 2, pp. 109–127, 1980, issn: 0035-9246.

[30] P. Mccullagh, “Regression models for ordinal data,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 42, no. 2, pp. 109–127, 1980, issn: 0035-9246.

[31] M. A. Mody, “What is an emerging market?” *International Monetary Fund Working Paper*, no. 4-177, 2004.

[32] R. Neate, *Ratings agencies suffer 'conflict of interest', says former moody's boss*, theguardian.com, Ed., Aug. 2011. [Online]. Available: `https://www.theguardian.com/business/2011/aug/22/ratings-agencies-conflict-of-interest#note#`.

[33] H. M. Park, "Practical guides to panel data modeling: A step by step analysis using stata," *Public Management and Policy Analysis Program, Graduate School of International Relations, International University of Japan.*, pp. 1–52, 2011.

[34] B. Peterson and F. E. Harrell Jr, "Partial proportional odds models for ordinal response variables," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 39, no. 2, pp. 205–217, 1990.

[35] H. Reisen, J. Von Maltzan, and G. Larraín, "Emerging market risk and sovereign credit ratings," *En: X Seminario Regional de Política Fiscal: compendio de documentos-Santiago: CEPAL, 1998-p. D1-D21*, 1998.

[36] R. W. Staiger and A. O. Sykes, "Currency manipulation and world trade," *World Trade Review*, vol. 9, no. 4, pp. 583–627, 2010. doi: 10.1017/S1474745610000340.

[37] W. N. Venables, *Modern applied statistics with S-PLUS*, eng, Third Edition, ser. Statistics and computing. New York: Springer, 1999, isbn: 0387988254.

[38] L. R. Wittie. (May 2018). 2017 annual sovereign default study and rating transitions. eng, Standard and Poors Global Ratings, [Online]. Available: https://www.standardandpoors.com/en_US/web/guest/article/-/view/sourceId/8950072.

[39] J. M. Wooldridge, *Econometric analysis of cross section and panel data*. Cambridge and London: MIT Press, 2002, isbn: 0262232197.

[40] ——, *Introductory econometrics: A modern approach*. Nelson Education, 2015.