Responding to the Eurozone Crisis - Applying the Shadow Rating Approach to Determine Economic Capital for Sovereign Exposures

A study on modeling sovereign credit risk within the Basel IRB framework

Simon Koskinen Rosemarin

Master's Thesis

Centre for Mathematical Sciences
Mathematical Statistics

January 21, 2014

LUNDS UNIVERSITET
Lunds Tekniska Högskola

Abstract

The recent European sovereign-debt crisis has made it clear that exposures towards sovereigns contain credit risk. However, according to the Basel framework's standardized approach banks are not required to hold any regulatory capital for highly rated sovereigns. In response, this thesis develops a shadow rating approach model for sovereign probability of default estimation, subsequently determining economic capital for sovereign exposures within a foundation internal ratings-based framework. Furthermore, the empirical Bayes estimator is utilized for low-default portfolio probability of default calibration. The model is tested on five homogeneous sub-segments in addition to the entire dataset at hand. Empirical findings suggest that the full dataset performs adequately overall. Nonetheless, model performance is superior for accurately constructed sub-segments. In addition, economic, monetary and political indicators as well as banking sector health are found to best replicate S&P's sovereign long-term issuer credit ratings.

Acknowledgements

Firstly, I would like to express my gratitude to senior capital and risk analyst Henrik Rahm at Nordea Markets for his creative solutions, practical insights, immeasurable patience and lightning fast responses to queries. My grateful thanks are also extended to associate professor Magnus Wiktorsson at Centre for Mathematical Sciences part of Lund University, for fruitful discussions and constructive feedback throughout the thesis process.

Finally, I wish to thank my family, Arno, Maaret, and Isak for their unlimited support and encouragement throughout my university studies.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

1 Introduction

1.1 Background

In the last decade, credit risk has increasingly become one of the major research fields in quantitative finance. The amount of traded credit derivatives has risen substantially since the mid-1990s, when interest grew within academia in terms of developing techniques for modeling and managing credit risk. Consequently, with the increasingly complex financial markets and the implications of the regulatory Basel Framework, sophisticated credit risk models have become essential to financial institutions (Trueck and Rachev 2009, p. XI).

As of late, financial institutions and regulatory bodies have paid considerable attention to sovereign credit risk measurement. The recent European sovereign-debt crisis has made it clear that exposures towards sovereigns contain credit risk. The main area of focus has in particular been GIIIPS exposures, i.e. towards Greece, Italy, Iceland, Ireland, Portugal and Spain. Prior to the crisis, investments in sovereign bonds were perceived as more or less risk-free, and it is only recently defaults to "advanced economies" were thought possible. The crisis has indeed contributed to a loss of confidence in sovereign debt amongst investors, and strongly impacted its risk-free status on financial markets.

Historically, financial theory has been based upon the existence of a risk-free asset, dating back to the groundbreaking capital-asset pricing model and modern portfolio theory introduced by Harry Markowitz in the 1950s. Nonetheless, the cruel reality is that there simply does not exist assets that are completely exempted from risk.

The fact that exposures were seen as free from default-risk is reflected in the first regulatory Basel Framework from 1988, Basel I, which set a zero percentage risk weighting for exposures to OECD sovereigns. In other words, banks were not required to hold any capital for exposures to OECD sovereigns. This was later modified in Basel II from 2004, in which financial institutions were required to hold capital in proportion to their risk exposures, and the use of internal rating models was introduced (Benzin et al. 2003, pp. 3-7).

Even though sovereign-debt crises are not uncommon in modern history, sovereign defaults remain a rarity in comparison to the number of corporate defaults, which is depicted in figure 1. Considering the lack of observed sovereign defaults historically, the task of implementing and integrating sound sovereign credit risk management into financial institutions becomes a complex matter.

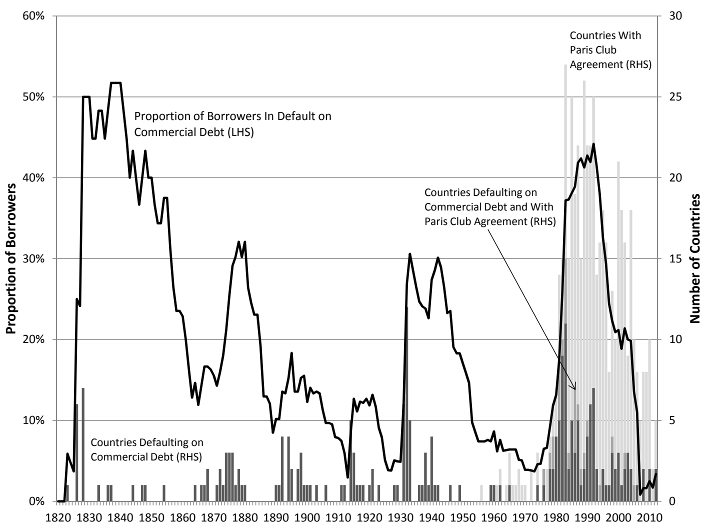

Figure 1: Historical sovereign default frequency (Tomz and Wright 2012)

The main area of research regarding sovereign-debt crises, prior to the Eurozone crisis, was focused on emerging markets (Joffe 2012, p. 357). Nevertheless, the recently increased awareness of sovereign credit risk has stimulated assessment and development of Basel II internal ratings-based modeling practices for Eurozone sovereigns.

1.2 Problem Discussion and Purpose

The issue of estimating economic capital towards sovereign exposures relates to the fact that banks no longer prefer to apply Basel framework's standardized approach, which exhaustively relies on external ratings published by credit rating agencies (CRA). Similarly, calculating capital for credit risk exposures towards externally unrated sovereigns and municipals is a topic that requires addressing. Financial institutions preferably solve these difficulties by developing models according to the internal ratings-based methodology.

With that said, the purpose of the thesis is to estimate economic capital (EC) for sovereign exposures within a foundation internal ratings-based (FIRB) framework. In detail, the objective is to develop a model estimating a sovereign's probability of default (PD), based on both quantitative and qualitative factors according to the shadow rating approach. In terms of selecting quantitative and qualitative factors, current sovereign credit risk methodologies of CRA Standard & Poor's (S&P's) as well as financial service company Bloomberg will act as benchmark methodologies.

Additionally, it is of interest to explore and discuss which practices are most suited for sovereign credit risk measurement, and also see if there exist other academically respected practices for sovereign exposures which may not be as suitable from an economic capital perspective.

This thesis consists firstly of an overview of sovereign credit risk and characteristics of sovereign entities, the Basel framework for IRB models, sovereign credit risk modeling aspects as well as PD calibration of low-default portfolios (LDP). Next, benchmark methodologies are examined in sufficient detail, and the rating model development methodology thoroughly presented. Ultimately, the model's features and performance are extensively analyzed and discussed.

1.3 Delimitations

Firstly, given the time constraints involved and the range a master's thesis is expected to cover, all current academic research within sovereign credit risk measurements will not be explored in-depth. Therefore, the shadow rating approach is selected, since it is considered as best practice for LDPs according to banks and CRAs.

Secondly, the model developed holds under the Basel FIRB framework, in which EC calculations are mainly based on the estimation of one parameter, namely PD. Therefore, it is obvious to limit the study to PD estimation. Other components required to compute EC will be set to the FIRB framework's set of predefined values.

Thirdly, taking the recent European sovereign-debt crisis into account, it is of primary interest to evaluate sovereigns within the EU and sovereigns in close geographical proximity. Hence, data from a selection of relevant sovereigns is used.

Additionally, only brief testing is applied when validating the model, since a thorough model validation process is not the main purpose of this thesis, and may well serve as a separate thesis topic.

1.4 Previous Research

The origin to bankruptcy risk modeling traces back to Edward Altman's pioneering Z-Score model published in 1968, which predicted corporate defaults using financial statement data. CRAs Moody's and S&P's later commercialized Altman's methodologies, acknowledging its robustness and validity for assessing corporate PDs (Joffe 2012, p. 354).

In regards to the history of sovereign risk modeling, Gray (2009, pp. 118-121) has compiled a brief overview of the models developed by researchers over the last decades. Gray (2009) initially refers to the work of Kindleberger (1978), which analyzes the history of financial crises and spillovers over the past 300 years. Kindleberger (1978) makes use of a model developed by Hyman Minsky, incorporating ideas of profound economists such as John Stuart Mill, Alfred Marshall and Irving Fisher. In the Minsky model, both investments in assets, such as real estate and stocks, and borrowers' indebtedness increase in economic upturns, eventually leading to financial bubbles. If the bubble grows sufficiently large and causes a change in investors' behaviour, a dumping of real or financial assets could possibly trigger a financial crises.

In addition, Gray (2009) arranges historical crisis models into three model generations. The first-generation models, developed in the late 1970s and early 1980s, focus on fundamental economic factors and implications of sovereigns with fixed exchange rates, i.e. when a currency's value is fixed towards the development of another currency. The main idea was that governments are not able to finance a fiscal deficit by printing money, while simultaneously upholding a credible fixed exchange rate. Fundamental economic factors also played a major part in the second model generation, which entered the frame in mid 1990s after the exchange rate mechanism crisis in 1992 and the Mexican crisis in 1994. One major feature was the fact that the models allowed for different outcomes, i.e. the models had multiple equilibria. The third-generation models, created in the wake of the Asian crises in $1997-1998$, emphasized the importance of banking, corporate, and government balance sheets. The models sparked researchers to further investigate factors such as solvency and liquidity issues within a sector, the role and consequences of government bail-out guarantees to banks, as well as currency mismatches (Gray 2009, pp. 119-120).

In the 1990s efforts were made to develop models which could signal forthcoming currency and banking crises, so called early warning systems. Macroeconomic factors were found to be significant explanatory variables when predicting the likelihood of a crisis. However, early warning systems performed poorly for out-of-sample crises, due to the use of backward-looking accounting variables, and the difficulty of defining a crisis event. Additionally, researchers have applied contingent claims option analysis when modeling sovereign risk, integrating them with macroeconomic models such as dynamic stochastic general equilibrium or macroeconomic monetary policy models. Gray (2009) also mentions financial contagion as an area which needs to be taken into account, but has only received notable attention within the last decade (Gray 2009, pp. 120-121).

As of late, academia has turned its attention to estimating sovereign PD based on market pricing. More specifically, several efforts have been put forward to deriving PDs from sovereign bonds or credit default swap (CDS) spreads. Remolona et al. (2007) and Longstaff et al. (2011) have presented two papers on the subject, both commonly cited amongst modern literature.

2 Theoretical Framework

2.1 Sovereign Credit Risk

The following section aims to provide a deeper knowledge within the area of sovereign credit risk. Furthermore, sovereign default will be defined, specific characteristics and drivers of sovereign credit risk as well as impact and costs of sovereign default will be presented. A brief analysis and comparison of external sovereign market data will be presented as well.

2.1.1 Defining Sovereign Credit Risk

Banks face several risks, one of which is credit risk. Credit risk is imminent if for instance a bank grants a loan to a counterparty, and the counterparty fails to meet its obligations following the loan. In terms of assessing sovereign credit risk, the first step towards a comprehension of the field is defining the event of sovereign default, which is not entirely uncomplicated. A common conception of default involves a sovereign failing to meet an interest or principal payment. A sovereign may also be perceived as in default if short-term or long-term debt is unserviced (Text Medic 2013a).

While there is no universally accepted definition, a well-cited document from CRA Standard and Poor's (2013b, p. 28) defines sovereign default in the following way:

> Standard & Poor's generally defines default as the failure to meet a principal or interest payment on the due date contained in the original terms of a debt issue. Questions can arise, however, when applying this definition in different situations and to different types of sovereign obligations. Standard & Poor's considers a sovereign to be in default under any of the following circumstances:
> (i) For local- and foreign-currency bonds, notes, and bills issued by the central government and held outside the public sector of the country, a sovereign default occurs when the central government either fails to pay scheduled debt service on the due date or tenders an exchange offer of new debt with less-favorable terms than the original issue.
> (ii) For local currency issued by the central bank, a sovereign default takes place when notes are converted into a new currency of less-than-equivalent face value.
> (iii) For private-sector bank loans incurred by the central government, a sovereign default occurs when the central government either fails to pay scheduled debt service on the due date or negotiates with the bank creditors a rescheduling of principal or interest at less-favorable terms than in the original loan.

Jedidi (2013, p. 5) discusses S&P's definition, emphasizing that the measure of a government's creditworthiness is dependent on the *ability* and *willingness* to uphold its commitment at due date. A challenging aspect when modeling sovereign credit risk is quantifying these two essential aspects. The ability to pay, whether or not a sovereign is able to meet its obligations, is commonly measured through quantitative macroeconomic stability factors, such as terms of GDP or government revenues. The willingness to pay, i.e. having sufficient funds but opting not to repay debt, is however more difficult to quantify. Therefore, qualitative factors such as historical debt payment culture as well as transparency and reliability of economic indicators are used. Macroeconomic factors such as GDP growth and inflation also influence the willingness to pay. Essentially, a sovereign's decision whether or not to pay is determined by the sanctions from a default in comparison to the amount possibly saved from debt payments. Additionally, Jedidi (2013) does point out that S&P's definition of default only incorporates sovereigns that have market access, and are able to issue government bonds.

2.1.2 Impact and Costs of Sovereign Default

Although there does not exist one general default definition, the common denominator lies in the repercussions of a default. Just as in the case of corporate defaults, sovereign financial distress results in a full or partial repudiation of its obligations. However, contrary to a typical corporate case, which outright defaults, sovereigns instead restructure debt and financial commitments (Text Medic 2013a). Through restructuring sovereigns reduce debt repayments costs, but simultaneously succumb to negative impacts on international trades, increased regulatory monitoring, reduced possibility of external financing as well as reputation effects (Remolona et al. 2007, p. 31). Menéndez (2012, p. 2) cites that sovereigns subject to high levels of interest rates may be forced into restructuring debt.

Furthermore, Ang and Longstaff (2011, p. 4) list properties where the consequences of sovereign default differs from corporate default. Firstly, corporations which cancel debt repayments would be forced to render assets to bond holders in collateral in the event that bondholders decide to litigate. However, in the case of a sovereign, most of its assets are domestically located within country boundaries, and in event of default it is not feasible for a sovereign to hand over these assets. Secondly, in contrast to corporate defaults, there does not currently exist an internationally acknowledged mechanism for managing sovereign defaults. Ang and Longstaff (2011) state that the restructuring of international sovereign debt crises of the 20th Century have all been ad-hoc responses from large international organs, such as the International Monetary Fund as well as commercial banks and governments.

Over the past 200 years, 107 sovereigns have defaulted on at least one occasion, and most frequently so in emerging markets. However, since the latest emerging markets debt crises in Latin America in the 1980s, there has been a significant decrease in default rate amongst EM sovereigns. This fact is evident in figure 1 of historical sovereign default frequency in the introduction chapter. A possible explanation lies in the fact that EM sovereigns generally have prospered. Nevertheless, the most interesting conclusion may stem from the fact that it is no longer preferable to opt to default, in contrast to a few decades back, mainly due to the reputation risk involved. A history of having a number of defaults leads to persistent wider spreads in the long term (Text Medic 2013a).

2.1.3 Drivers of Sovereign Credit Risk

Santis (2012, pp. 4-5) has created an overview of the major factors underlying sovereign risk, namely aggregate risk, country-specific risk and contagion risk, depicted in figure 2. Aggregate risk, constituted by global factors, is more specifically related to monetary policy, global uncertainty and risk aversion. Text Medic (2013a) states that the G10 monetary policy represents a sufficiently good measure for monetary policy, while the volatility index VIX is suitable as an indicator for uncertainty and risk aversion. Moreover, the three main factors affecting country-specific risk are variations in sovereign default risk, a sovereign's ability to raise funds in its primary market as well as issues regarding liquidity in secondary markets. Contagion risk involves the vulnerability to a contagious spread from one market to another (Santis 2012, pp. 4-5).

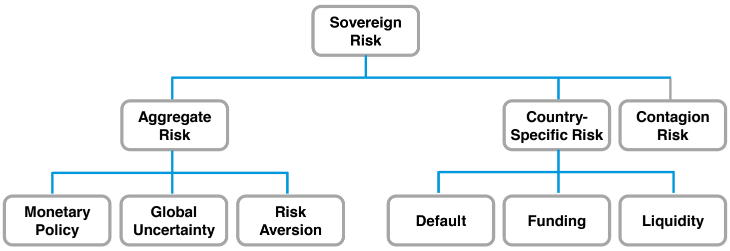

Figure 2: Determinants of sovereign risk (Santis 2012, p. 32)

Country-specifics are good up to a point, but do not change much between different periods, while aggregate factors fluctuate more frequently and represent the primary drivers of short-term sovereign spreads, dominating country-specific factors in the short run (Text Medic 2013a). Similarly, Longstaff et al. (2011, p. 77) suggest that global macroeconomic factors heavily influence sovereign default risk. The study of Ang and Longstaff (2011, p. 19) on system sovereign credit risk also strongly suggests that the systematic risk component is highly correlated with financial markets, rather than country-specific macroeconomic fundamentals.

Nonetheless, the challenging aspect is identifying quantitative or qualitative variables which correspond to each of the three main risk factors: aggregate, country-specific, and contagion risk (Santis 2012, pp. 4-5). Rosenberg and Singenellore (2013, p. 11) have compiled a list of drivers of sovereign spreads, listed in table 1.

Country-Specific	**Global Risk Aversion / G-10 Policies**
Government Debt/GDP Ratio	U.S. Fed Funds Rate
Budget Deficit/GDP Ratio	U.S. 10-Year Yield
Debt Service	VIX Index
Relative GDP Growth	Baa Corporate/Treasury Yield Spread
Current Account Balance/GDP Ratio	MOVE Index
External Debt/GDP Ratio	U.S. High-Yield/Treasury Spread
Banking System Fragility (Bank Share Prices)	G-10 FX Volatility
FX Reserves	Bloomberg U.S. Financial Conditions Index
Default History	U.S. Economic Policy Uncertainty Index
Per Capita GDP	Euro Area Economic Policy Uncertainty Index
Rating Agencies Rating Outlook	Contagion - Correlation of Market Returns
Carry/Risk Ratio	
Relative Equity Market/Local Bond Market Performance	
CDS Spread	
Inflation Performance	
Political Risk	
Liquidity - Bid/Ask Spread	

Table 1: Drivers of sovereigns spreads (Rosenberg and Singenellore 2013, p. 11)

Banking system fragility is listed as one of the drivers in table 1. It is paramount to note that sovereign risk and the banking sector susceptibility are interconnected, since negative feedback loops exist and may abound. This phenomenon is depicted in figure 3. For instance, a decrease in asset prices may deteriorate markets and bank balance sheets, possibly triggering government bailouts and a rise in government budget deficit, subsequently resulting in widening sovereign spreads. Reversely, widening spreads impact the banking sector, since banks own a lot of sovereign debt. As a result, a decrease in bank lending may occur, leading to weaker GDP growth, which widens sovereign spreads even further (Text Medic 2013a).

In connection to listed of country-specific drivers by Rosenberg and Singenellore (2013), Joy (2012, pp. 4-6) examines how macroeconomic factors influence the

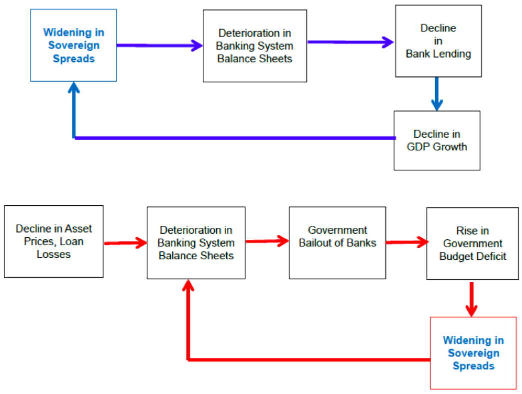

Figure 3: Sovereign risk and banking system interconnection (Rosenberg and Singenellore 2013, p. 14)

probability of a sovereign defaulting on its sovereign debt. Joy (2012) finds that high external debt interest payments and large budget deficits constitute the main macroeconomic factors that drive sovereign default. On a one-year-ahead predictive power basis, these two measures represent solid signals for default.

What is more, Joy (2012, pp. 4-6) claims that the most significant macroeconomic tipping points of sovereign default is not dependent on fiscal solvency, but liquidity. For instance, large general government budget deficits represent a typical fiscal illiquidity tipping point for default. However, the research over the past decade has produced severely inconsistent results. For instance, the findings of Kraay and Nehru (2006) point to external liquidity, external solvency and GDP growth as primary characteristics for sovereign PD estimation. Conflictingly, Tomz and Wright (2007) conclude that growth is not essential, stating that sovereigns have defaulted in the past when having a domestically healthy economy. Additionally, while Manasse et al. (2003) find a vast amount of factors driving sovereigns into debt crises, Bandiera et al. (2010) identify the level of indebtedness as the only macroeconomic characteristic needed to satisfactory predict default. The lack of consensus regarding key determinants of sovereign default indicate that there is a large uncertainty regarding the optimal empirical model within the field.

2.1.4 External Sovereign Market Data

There exist several external market data sources from which sovereign credit risk can be measured, the most notable include credit ratings and CDS spreads, but also bond spreads to some extent.

Bai and Wei (2012, p. 2) pay attention to the advantages of sovereign CDS spreads in comparison to bond spreads. Firstly, sovereign CDS contracts are traded in foreign currencies, and as a result investors are protected against foreign exchange risk and inflation risk. Reversely, government issued bonds are denominated in domestic currency, and therefore contain foreign exchange risk and inflation risk. Furthermore, O'Kane (2012, pp. 5-6) notes that CDS and bond spreads may not follow identical spread paths as a result of the two financial instruments' differing currency denotations. In addition, Longstaff et al. (2011, pp. 76-78) state differences in cash flows between bonds and CDS contracts induce spread divergence. For instance, CDS contracts do not require a cash flow at initiation.

Secondly, Bai and Wei (2012, p. 2) mention that CDS contracts are traded on the OTC (over-the-counter) credit derivative market, where governments cannot influence market prices in the same way as the market for bonds. In bond markets, governments are able to manipulate according to their preferences through buyback schemes, timing of issuance or issuance amount. Contrarily, governments cannot trade CDS protection on their own bonds, due to the counterparty risk involved. Thus, sovereign CDS contracts represent a better indicator for the market's perception of a sovereign's credit quality.

Thirdly, Longstaff et al. (2011, pp. 75-76) acknowledge that sovereign CDS markets generally are more liquid than sovereign bond markets, implying that CDS spreads are more accurately estimated and quicker to incorporate information. However, Joffe (2012, p. 355) states that liquidity still remains an issue within sovereign CDS contracts, finding that the majority of the contracts were traded fewer than five times per day in 2009 and 2010. Furthermore, Arezki et al. (2010, pp. 36-37) highlight that increases of a sovereign's CDS spread lead to spill-over effects to other closely related sovereigns. For instance, Greece's credit rating downgrade in 2010 not only influenced the cost of insuring Greek debt, but also the debt of other euro area sovereigns even though their credit ratings were unchanged.

Moreover, Huberdeau (2013, p. 1) criticizes CDS spreads, arguing that they have pro-cyclical impact. The property is exemplified as a scenario where a deteriorating credit rating generates a higher interest rate, which in turn affects a sovereign's debt sustainability and ability to obtain new loans. Huberdeau (2013) also makes a valid point stating that CDS pricing is heavily reliant on agency ratings, adding to the pro-cyclical characteristic. Likewise, alongside traditional credit ratings it is becoming more common that rating agencies issue "CDS-implied ratings", which further intensifies pro-cyclicality.

Taking the criticism of CDS markets into consideration, Flannery et al. (2010, pp. 2085-2087) state that credit ratings have remained the primary source as indicator of credit risk. “Regulators and investors should not replace one broken system (credit ratings) with another broken system (CDS)”, Flannery et al. (2010) acknowledge. Assessing sovereign credit ratings more in detail, there is a general conception amongst academic research that credit ratings do not provide fully accurate estimates for sovereign credit risk. Altman and Rijken (2006, p. 54) cite that rating agencies prefer not to have ratings fluctuating frequently. Swift reversals of ratings tend to negatively affect agencies’ reputation, even if the rating alterations represent the actual changes in creditworthiness. Nevertheless, Altman and Rijken (2006) list several sound arguments for rating stability, both from an investor and regulatory point of view.

2.2 Basel Framework

The Basel framework is a regulatory standard on banking regulation, and all EU member states hold under the framework. The following section will review the Basel Framework, and in particular with regards to IRB modeling, PD and EC for sovereigns. Trueck and Rachev (2009) have compiled a thorough overview of the Basel II Framework, which will act as a basis. As a side note, the Basel III Framework which has recently been introduced and is currently under implementation at financial institutions, will not be reviewed in this chapter.

The Basel II framework consists of three pillars, visualized in figure 4, which together intend to provide stability to the financial system.

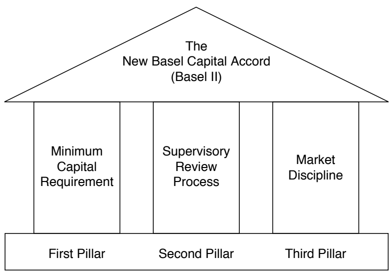

Figure 4: The three pillars of the Basel II framework (Trueck and Rachev 2009, p. 33)

The first pillar contains guidelines on minimum capital requirements for market, credit and operational risk, and is of main interest when estimating EC. Furthermore, the pillar presents two central approaches to measuring credit risk, namely the standardized (STD) approach and the IRB approach, visualized in figure 5. Additionally, the IRB approach offers two sub-approaches, the FIRB approach and the advanced IRB approach (AIRB) (Trueck and Rachev 2009, p. 34).

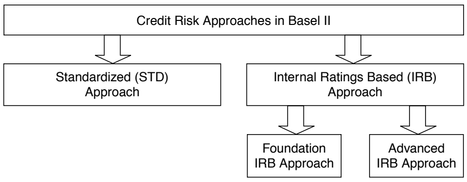

Figure 5: Credit risk approaches in Basel II (Trueck and Rachev 2009, p. 34)

2.2.1 Standardized Approach

Out of the two main approaches the STD approach is the less complex one, as it relies exhaustively on long-term external ratings published by CRAs. Under the STD approach, financial institutions allocate a risk weight to each asset and off-balance-sheet position by using the measure of risk-weighted assets (RWA). The RWA is calculated as $RWA = E \cdot r$, where E is the value of the exposure and r is the risk weight of the exposure. The risk weights for sovereigns, with notation of S&P's, are listed in table 2 (Trueck and Rachev 2009, p. 37).

External Rating	Risk Weights
AAA to AA-	0%
A+ to A-	20%
BBB+ to BBB-	50%
BB+ to B-	100%
Below B-	150%
Unrated	100%

Table 2: Risk weights for sovereigns, notation of S&P's (Trueck and Rachev 2009, p. 33)

Furthermore, the first pillar defines the ratio of regulatory capital to RWA, with a minimum requirement of 8%. In other words, a bank must hold at least 8% of its capital in relation to its RWA:

$$\text{Capital Ratio} = \frac{\text{Total Regulatory Capital}}{\text{RWA}} \geq 8\% \qquad (2.1)$$

For example, an exposure to a 'B'-rated sovereign, which is risk weighted at 100%, implicates a capital charge equal to a minimum 8% of the exposure's full value. That is, a bank must for this particular exposure hold at least 8% of the value of the exposure as regulatory capital. Correspondingly, an exposure with a risk weight of 50% would result in a capital charge equal to at least 4% (Trueck and Rachev 2009, pp. 33-34).

The STD approach has received vast criticism for providing regulatory incentives for banks to accrue large sovereign exposures, and in particular for highly-rated sovereigns. This stems from the fact that exposures to sovereigns with a rating between 'AAA' and 'AA-' are zero risk weighted according to the STD approach. However, the IRB approach does not imply a zero risk weighting when calculating capital charges (Hannoun 2011, pp. 11).

2.2.2 Internal Ratings-Based Approach

In contrast to the STD approach, the IRB approach allows banks to assess credit risk exposures by using their own internal estimates, subject to approval by supervisory minimum requirements. The approach is based on four input parameters, which Basel Committee on Banking Supervision (2005) defines as follows:

Probability of Default (PD): The probability of an obligor defaulting on a contractual payment on a one year horizon.

Exposure at Default (EAD): The nominal amount of the exposure.

Loss Given Default (LGD): The actual loss that the bank faces if an obligor defaults. LGD is commonly denoted as a percentage of EAD.

Maturity (M): The contractual maturity of the exposure.

Moreover, Basel Committee on Banking Supervision (2005) states that expected loss (EL), denoted as a %-figure of EAD, can be calculated from a bottom-up perspective, using the two risk parameters PD and LGD:

$$\mathrm{EL} = \mathrm{PD} \cdot \mathrm{LGD} \tag{2.2}$$

Banks use the measure of EL to forecast the average level of credit losses, which is marked by the dashed line in figure 6, while losses above EL are called unexpected losses (UL). In order to protect banks' debt holders against peak losses exceeding EL, capital is needed as a loss-absorbing buffer. Sufficient capital ensures that the level of UL only exceeds the level of capital by only a very low, fixed probability (Basel Committee on Banking Supervision 2005, pp. 2-3).

There are two different options within the IRB approach, namely AIRB and FIRB. In the FIRB methodology it is sufficient to estimate PD for exposures, and apply a set of inputs for other parameters predefined by supervisors. For instance, an exposure without collateral LGD is set to either 45% for senior (higher priority) claims or 75% for subordinated (lower priority) exposures. Contrarily,

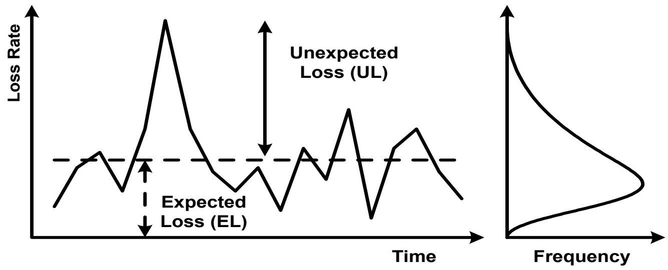

Figure 6: Realized losses over time and credit risk loss distribution (Basel Committee on Banking Supervision 2005, p. 2)

the AIRB option permits banks to implement models to estimate each input risk parameter (Trueck and Rachev 2009, pp. 41-42).

2.2.3 IRB Formulas and Economic Capital

In previous sections we have defined capital requirements as a measure of regulatory capital. However, one needs to note that regulatory capital and EC differ slightly in their respective definitions. While regulatory capital implies the minimum capital requirements set by regulators, EC represents a bank's internal capital estimate, denoted in currency amount. EC is commonly used to manage risks across the entire portfolio of assets as well as making strategic decisions (Elizalde and Repullo 2006, p. 1). Mausser and Rosen (2007, pp. 681) focus more detailedly on *economic credit capital*, which purpose is to absorb large *unexpected losses* (UL) specifically linked to credit events of obligors, including defaults, credit migrations (up- or downgrades) and credit spread changes. Nonetheless, capital calculations for regulatory capital and EC are performed equivalently under the IRB framework, and therefore we will not discriminate between the two terminologies.

For sovereign exposures, there are a set of IRB formulas for which asset correlation R, maturity adjustment b, capital requirement K, and risk-weighted assets RWA are calculated. The derivation of the formulas is provided in Trueck and Rachev (2009, pp. 43-50) for the curious reader.

$$R = 0.12 \cdot \frac{1 - \exp(-50 \cdot PD)}{1 - \exp(-50)} + 0.24 \cdot \left(1 - \frac{1 - \exp(-50 \cdot PD)}{1 - \exp(-50)}\right) \tag{2.3}$$

$$b = (0.11852 - 0.05478 \cdot \ln(PD))^2 \tag{2.4}$$

$$K = LGD \cdot \left(\Phi \left(\frac{\Phi^{-1}(PD) + \sqrt{R} \cdot \Phi^{-1}(0.999)}{\sqrt{1-R}} \right) - PD \right) \cdot \frac{1 + (M - 2.5) \cdot b}{1 - 1.5 \cdot b} \tag{2.5}$$

$$RWA = K \cdot 12.5 \cdot EAD \tag{2.6}$$

With regard to the formula for capital requirement K, Φ is the standard normal distribution, $\Phi^{-1}(PD)$ represents the inverse of the standard normal distribution applied to PD in order to derive the default threshold, $\Phi^{-1}(0.999)$ is the inverse of the standard normal distribution applied to the 99.9% confidence level, and the term outside the brackets is the maturity adjustment factor (Basel Committee on Banking Supervision 2006, pp. 63-64). The confidence level of 99.9% is set by supervisory authorities in order to protect against model uncertainties and estimation errors of risk factors such as PD, LGD and EAD (Munniksma 2006, p. 29).

Intuitively, the asset correlation R declines with increasing PD. In other words, as PD of a sovereign increases, the sovereign becomes less dependent on aggregate factors, and consequently more affected by country-specifics (Trueck and Rachev 2009, p. 46).

Regarding maturity, it is considered as an explicit risk component under the IRB framework, in order to reflect potential credit quality deterioration of credits with longer maturities. Empirical evidence also indicate that long-term credit exposures hold more risk than short-term exposures (Munniksma 2006, p. 28).

RWA is determined by the capital required times 12.5, resulting from the fact that the required capital is 8% of RWA.

If we revise the part within the brackets from formula for capital requirement, equation 2.5, and extract LGD, we obtain a formula for UL:

$$UL = \underbrace{LGD \cdot \Phi \left(\frac{\Phi^{-1}(PD) + \sqrt{R} \cdot \Phi^{-1}(0.999)}{\sqrt{1-R}} \right)}_{\text{EL+UL}} - \underbrace{LGD \cdot PD}_{\text{EL}} \tag{2.7}$$

Conclusively, the required capital is based exclusively on UL, and not the sum of EL and UL Trueck and Rachev (2009, p. 50). Figure 7 illustrates how the credit risk loss distribution relates to UL and EL.

In terms of actually utilizing the IRB formulas, table 3 illustrates the resulting risk weights and required capital from a set of PD levels. Since the required capital is denoted in % of EAD, EC can be calculated by multiplying the required

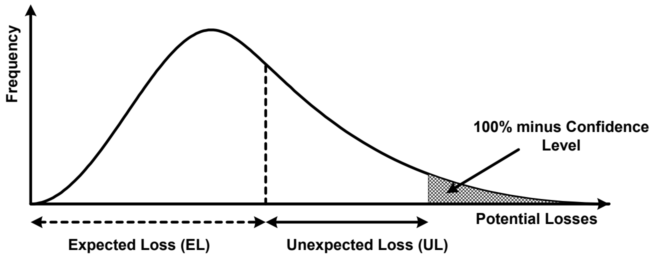

Figure 7: Credit risk loss distribution (Basel Committee on Banking Supervision 2005, p. 3)

capital by EAD. In addition, note that LGD and M assume standard FIRB input values of 45% and 2.5 years respectively (Basel Committee on Banking Supervision 2006). EAD is set to $1,000,000,000$ EUR.

Probability of Default (%)	Risk Weight (%)	Required Capital (%)	Economic Capital (EUR)
0.01	7.53	0.6	6000000
0.02	11.32	0.91	9000000
0.03	14.44	1.16	11600000
0.05	19.65	1.57	15700000
0.1	29.65	2.37	23700000
0.25	49.47	3.96	39600000
0.5	69.61	5.57	55700000
1	92.32	7.39	73900000
2	114.86	9.19	91900000
3	128.44	10.28	102800000
4	139.58	11.17	111700000
5	149.86	11.99	119900000
10	193.09	15.45	154500000
15	221.54	17.72	177200000
20	238.23	19.06	190600000

Table 3: Illustrative IRB risk weights and EC for sovereign exposures. $LGD = 45\%$, $EAD = 1,000,000,000$ EUR, $M = 2.5$ years.

2.2.4 Minimum Requirements under IRB Approach

When developing a rating model within the IRB framework, a set of minimum requirements determined by supervisors must be met. The most significant requirements for the thesis' purpose are listed in paragraphs $416 - 417$ and $461 - 463$ in Basel Committee on Banking Supervision (2006, pp. 93-102), and hold for sovereign, corporate and bank exposures:

- When estimating PD for each rating grade, IRB models must account for the long-run experience, and also be estimated using one or a combination of techniques based on internal default experience, mapping to external data, or statistical default models.

- Irrespective of which data source is used, the length of the underlying historical data must correspond to at least one business cycle, i.e. at least five years.
- The input data must form an intuitive set of predictors, and must be assessed on accuracy, completeness and appropriateness. Furthermore, the dataset used to build the model must also be representative for the bank's sovereign exposures.
- Banks must adopt a conservative bias when limited data is available, for instance in the case of LDP.

2.3 Modeling Sovereign Credit Risk

In the following section, various approaches to modeling sovereign credit risk will be reviewed. Particular attention will be given to the shadow rating approach (SRA), which represents the adopted modeling methodology when estimating PD for sovereign exposures.

2.3.1 Overview of Credit Risk Approaches

One of the most frequently used approaches within academic literature is contingent claim analysis models, which are typically found under the category structural models in figure 8. Contingent claim analysis makes use of the Merton model and Black and Scholes option pricing theory pioneered in the 1970s, and was primarily developed for corporation credit valuation (Brandorf and Holmberg 2010, p. 11). Several researchers however have applied structural

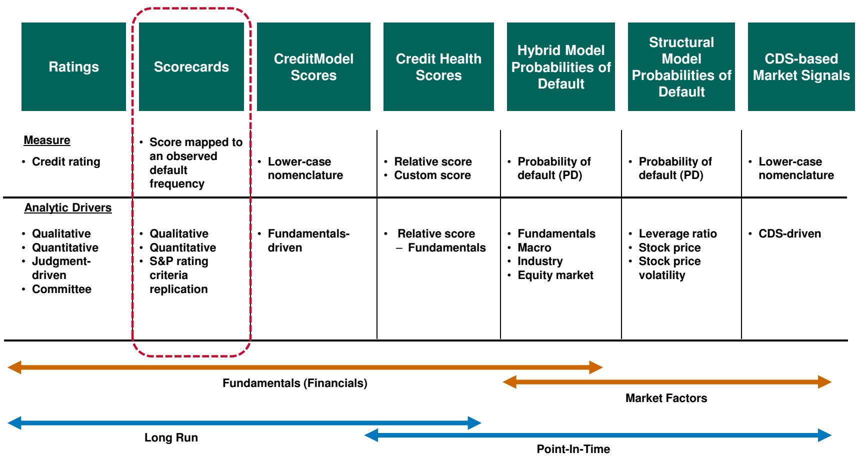

Figure 8: Spectrum of credit risk measures (Guglielmo 2013, p. 4)

models to estimate sovereign credit risk. Amongst many others, the most notable studies include Gray and Jobst (2011) and Gray et al. (2007). Duyvesteyn and Martens (2012, p. 1) applied the model to emerging markets sovereigns, which rendered an underestimation of sovereign credit spreads as well as assigning near-zero PDs. A possible explanation might be that structural models are suitable for PD estimation when balance sheet data is the most reliable and preferred source of data. Additionally, structural models often lack flexibility when fitting to a certain spread term structure, making pricing of credit derivatives difficult (Brandorf and Holmberg 2010, p. 13). Due to its drawbacks when applying to sovereigns, the model will not be taken into further consideration.

An alternative to structural models are CDS-driven reduced-form models, which in recent years have been given attention. Contrary to structural models, which have an explicit link between capital structure and PD, reduced-form models treat defaults as exogenous stochastic events that occur at an unknown time of default. In other words, default is not dependent on items of the balance sheet, and instead is a stochastic variable (Bomfim 2005, p. 183). Although structural models may have a clearer economic intuition, reduced-form models fit data better (Bomfim 2005, p. 199). In addition, two well-cited studies on reduced-form models are Ang and Longstaff (2011) and Pan and Singleton (2008).

Furthermore, Text Medic (2013b) highlights several modeling approaches for sovereign default risk, consisting of both qualitative and quantitative approaches. Scoring models such as Euromoney's weighted factor scoring model, are typical qualitative approaches, while discriminant analysis, principal component analysis, and logit or probit approaches constitute popular quantitative scoring approaches (Rosenberg and Singenellore 2013, p. 22). Both types of approaches makes use of quantitative inputs, such as economic, fiscal, monetary and market factors, as well as qualitative components such as political risk scores. For instance, sovereign methodologies of CRAs rely on a mix of multiple quantitative and qualitative inputs (Text Medic 2013b).

Continuing to focus on scorecards, Izzi et al. (2011) compile a brief overview of three major banking practice methodologies used when developing a PD model. When analyzing small- and medium-sized enterprises or large corporations, a good/bad analysis or pure expert ranking method is applied respectively. However, these two methodologies rely on a significant amount of historical default data, which makes any use of the models inapplicable to LDP segments such as sovereigns or banks. In contrast, one of the banking practice methodologies suitable for LDPs is the SRA, which will constitute the main focus throughout the rest of the section.

2.3.2 Shadow Rating Approach

The SRA is a commonly used banking practice methodology, applied when there is an insufficient amount of default data to build a prediction model for internal EC calculation. The main requirement of the approach is the availability of external ratings from CRAs for the substantial part of a portfolio. Erlenmaier (2011, p. 43) recommends long-term, local currency ratings, due to Basel regulations and measuring transfer risk separately from an obligor's credit rating. Additionally, SRA typically utilizes macroeconomic ratios and indicators as well as qualitative factors (Erlenmaier 2011, pp. 39-40).

In short, one must firstly identify risk factors that represent significant predictors of future default. SRA may then be applied with the objective to weight the risk factors in order to mimic external ratings using a statistical model. In order to use SRA for PD estimation, external ratings have to be calibrated to PDs, i.e. a PD is attached to each external rating grade. In other words, the objective is to rank sovereigns' creditworthiness by replicating external ratings, more specifically external rating PDs, using a selection of quantitative and qualitative risk factors (Erlenmaier 2011, pp. 39-40).

In order to obtain estimates of external PDs which the SRA is based upon, Erlenmaier (2011, p. 45) suggests the use of annually published historical default rates from CRAs. Default rates are defined as the annual number of defaults in relation to the total number of sovereigns within each rating category. However, these default rates do not provide sufficiently accurate estimates, especially for highly rated sovereigns which have a zero default rate. This issue is addressed in detail in section 2.4.

It is also important to note that the observations SRA utilizes are not split into default and non-defaults, which normally is the case for corporate default analysis. Instead, SRA follows the external rating scale which consists of multiple groups, each group corresponding to a rating grade. These grades are ordered according to creditworthiness, thus providing a finer granularity in comparison to corporate analysis. Therefore, determining factors with high explanatory power might be easier. Nevertheless, since SRA is based on external ratings and not observed defaults, information on the quality of creditworthiness is not as significant. Moreover, SRA is strongly dependent on the quality of external ratings, which implies certain limitations (Zaalberg 2013, p. 7). One limitation mentioned in section 2.1.4 stems from the fact that CRAs prefer rating stability rather than frequently altering credit ratings.

Next, one may inquire why banks do not utilize external ratings directly, and instead apply the SRA by attaching a PD to each rating grade. There are several incentives for banks to adopt the latter approach:

- In general, CRAs do not provide sufficiently detailed rating methodologies. Internal models therefore provide additional insight to how ratings are determined.
- Banks are able to incorporate their own opinion when rating counterparties, since IRB models allow for rating modifications and expert input.
- Internal models are able to provide ratings for externally unrated exposures.
- The Basel framework encourages the use of IRB approaches, and less reliance on external ratings (Zaalberg 2013, p. 7).

The SRA methodology will be thoroughly presented in chapter 4.

2.4 PD Calibration of Low-Default Portfolios

Due to the lack of defaults in historical data of sovereigns and in order to facilitate the use of SRA, the first step is to calibrate a PD scale for external ratings. The following section will therefore give a brief overview to a number of PD calibration approaches, as well as a thorough outline of the preferred Empirical Bayes (EB) estimator methodology.

2.4.1 Overview of PD Calibration Approaches

LDP PD estimation has been given reasonable attention in literature since the introduction of the Basel II framework. The main concern for supervisors is the possibility of credit risk being underestimated due to the lack of historical default data, which effects model development, estimation of parameters and model validation (Dzidzevičiūtė 2012, p. 132). This is highlighted by the Basel committee in the IRB minimum requirements stated in section 2.2.4.

The issue of LDP is present when using long-run averages on one-year cumulative default rates (DR) in order to estimate external rating PDs. In reply, a range of statistical techniques have been applied to enforce sufficiently conservative PD estimates for regulatory capital allocation, such as confidence level based approaches (Pluto and Tasche (2005)), CAP curve calibration (van der Burgt (2007)), and marginal of conservatism methodologies (Benjamin et al. (2006)). Nonetheless, there is no consensus on which technique represents best practice (Dzidzevičiūtė 2012, p. 134).

One of the latest contributions to LDP modeling is the Bayesian approach, which will be the focus of our forthcoming analysis. Bayesian parameter estimation is useful in cases of small samples, providing flexibility in choosing the prior probability distribution, the possibility of incorporating expert opinion as well as not producing as conservative estimates as other LDP approaches. The final argument may sound peculiar, since the aim of LDP modeling is to produce conservative estimates. However, for instance the confidence level based approach is known to produce estimates which stakeholders of banks find too conservative, and are therefore difficult to buy-in (Clifford et al. 2013, p. 17).

2.4.2 Empirical Bayes Estimator

Since data on sovereign default are scarce, PD estimates based purely on sovereign DR are insufficient for capital allocation according to the Basel framework. In order to make DR applicable, the EB approach makes use of a data-driven estimation procedure for the prior probability distribution, which provides additional information to PD estimations (Orth 2011, p. 5). Orth (2011) applies the EB to sovereign PD estimation by using corporate data to estimate the prior probability distribution. In the following paragraphs, the EB estimator will be formally defined.

Suppose that all obligors have rating r at time t, $t = 1, \dots, T$ form a cohort. Further, let $N^r_{t,1}$ be the number of obligors that constitute the cohort at period t, and similarly let $N^r_{t,s}$ be the number of obligors that still has not defaulted at period $t + s$ or is not censored in the first $s - 1$ periods. Examples of censored data are withdrawal of external ratings or gaps in dataset. Moreover, denote $D^r_{t,s}$ as the number of defaults in period $t + s$, and $L^r_{t,s}$ be the number of censored in period $t + s$. Define λ^r_s as the marginal default rate, i.e. conditional on surviving the first $s - 1$ periods λ^r_s represents the probability of an r-rated obligor defaulting s periods later (Orth 2011, p. 3).

Suppose there exists $G \geq 2$ different groups or rating categories, and that the marginal DR for each group g, $g = 1, \dots, G$, are a priori beta distributed and have the same prior parameters,

$$\lambda^{r,g}_s \sim beta(\alpha^r_s, \beta^r_s). \tag{2.8}$$

Additionally, the conditional distribution of the number of defaults in period s is binomially distributed,

$$D^{r,g}_s | \lambda^{r,g}_s \sim Bin(\tilde{N}^{r,g}_s, \lambda^{r,g}_s). \tag{2.9}$$

Note that the notations above in equations 2.8 and 2.9 are somewhat simplified, where $D_s^{r,g} = \sum_{t=1}^{T} D_{t,s}^{r,g}$ and $\tilde{N}_s^{r,g} = \sum_{t=1}^{T}(N_{t,s}^{r,g} - L_{t,s}^{r,g}/2)$. $\tilde{N}_s^{r,g}$ is the adjusted number of non-defaulted obligors, where $L_{t,s}^{r,g}/2$ represents the assumption that censored obligors have survived on average half of the period (Orth 2011, pp. 4-5).

The framework above is often referred to as the beta-binomial model, and the beta distribution produces parameters bounded within $[0, 1]$, which is suitable for PD estimation. One should however note the binomial assumption of conditional independence of default events. Nonetheless, Orth (2011) does provide evidence that the estimator performs well for data that involves dependencies through common shocks (Orth 2011, p. 6).

Next, we define the prior mean of $\lambda_s^{r,g}$ as $\mu_s^r = \frac{\alpha_s^r}{(\alpha_s^r+\beta_s^r)}$, and the prior precision as $\tau_s^r = \frac{1}{(1+\alpha_s^r+\beta_s^r)}$. These two prior parameters are estimated by using Method of Moments in formulas 2.10 and 2.11. The derivation is found in Kelinman (1973).

$$\hat{\mu}_s^r = \sum_{g=1}^{G} w_s^{r,g} \frac{D_s^{r,g}}{\tilde{N}_s^{r,g}} = \sum_{g=1}^{G} w_s^{r,g} \hat{\lambda}_s^{r,g} \tag{2.10}$$

$$\hat{\tau}_s^r = \frac{\frac{G-1}{G}\sum_{g=1}^{G} w_s^{r,g}(\hat{\lambda}_s^{r,g} - \hat{\mu}_s^{r,g})^2 - \hat{\mu}_s^r(1-\hat{\mu}_s^r)\left(\sum_{g=1}^{G} w_s^{r,g}(1-w_s^{r,g})/\tilde{N}_s^{r,g}\right)}{\hat{\mu}_s^r(1-\hat{\mu}_s^r)\left(\sum_{g=1}^{G}(1-1/\tilde{N}_s^{r,g})w_s^{r,g}(1-w_s^{r,g})\right)} \tag{2.11}$$

The weights can be determined in a number of ways. Most simplistically, one may initially set weights equally $w_s^{r,g} = 1/G$ or according to the number of observations for each group $w_s^{r,g} = \tilde{N}_s^{r,g}/\sum_{g=1}^{G}\tilde{N}_s^{r,g}$. However, Kelinman (1973) proves that inclusion of $\hat{\tau}_s^r$ provides the optimal weights, and proposes the use of one iteration to re-estimate the prior estimates by setting the weights to

$$w_s^{r,g} = \frac{\tilde{N}_s^{r,g}}{1+\hat{\tau}_s^r(\tilde{N}_s^{r,g}-1)} \Big/ \sum_{j=1}^{G} \frac{\tilde{N}_s^{r,j}}{1+\hat{\tau}_s^r(\tilde{N}_s^{r,j}-1)}. \tag{2.12}$$

Note also that it is not guaranteed that $\hat{\tau}_s^r$ will be within $[0, 1]$, which is a necessity. Therefore, $\hat{\tau}_s^r$ is truncated at zero and one (Orth 2011, p. 6).

Having now estimated prior parameters $\hat{\mu}_s^r$ and $\hat{\tau}_s^r$, the aim is to reach the posterior distribution, which is done by applying the Bayesian theorem. The posterior mean, i.e. the EB estimator for $\lambda_s^{r,g}$, is estimated according to equation

2.13.

$$\hat{\lambda}_{s,EB}^{r,g} = \frac{1-\hat{\tau}_s^r}{1+\hat{\tau}_s^r(\tilde{N}_s^{r,g}-1)}\hat{\mu}_s^r + \frac{\hat{\tau}_s^r \tilde{N}_s^{r,g}}{1+\hat{\tau}_s^r(\tilde{N}_s^{r,g}-1)}\hat{\lambda}_s^{r,g} \tag{2.13}$$

As seen in equation 2.13, the EB estimator is a weighted average of the prior mean $\hat{\mu}_s^r$, and the marginal default rate estimate $\hat{\lambda}_s^{r,g}$ for group g. Furthermore, the EB estimator brings the marginal default rate estimates closer towards the prior means, which are equal for all groups g. Additionally, if the number of observations for each group $\tilde{N}_s^{r,g}$ increases, the less affect the prior means will have on the EB estimator, which is intuitively appealing (Orth 2011, p. 7).

Ultimately, PD estimates are produced by applying equation 2.14 to $\hat{\lambda}_{j,EB}^{r,g}$ (Orth 2011, p. 7).

$$P\hat{D}_{s,EB}^{r,g} = 1 - \prod_{j=1}^{s}(1-\hat{\lambda}_{j,EB}^{r,g}) \tag{2.14}$$

3 Benchmark Methodologies

The following two sections will briefly review two separate sovereign methodologies currently in use. Firstly, SRA aims to replicate external ratings, which naturally propels an assessment of CRA S&P's sovereign government rating methodology. Secondly, the purely quantitative based Bloomberg SRSK model will examined.

3.1 S&P's Sovereign Government Rating Methodology

Standard and Poor's (2013c, pp. 3-40) summarize their sovereign government rating methodology, based upon a foundation of five areas indicating a sovereign's willingness and ability to pay. Each area, depicted in figure 9, is determined by both quantitative and qualitative factors as well as subjective judgment. In addition to the information listed below, the curious reader is encouraged to view the URL link found in references for further details within each area.

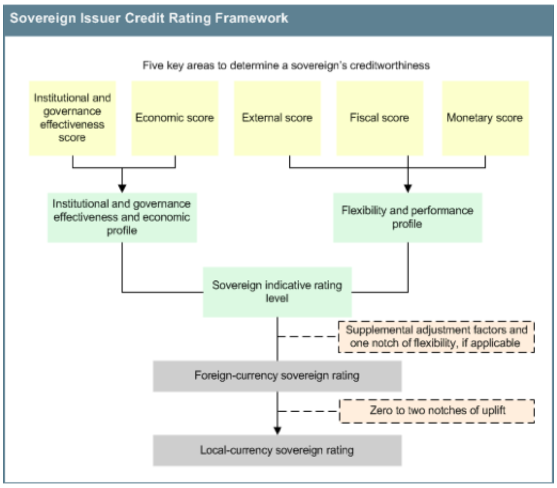

Figure 9: S&P's sovereign issuer credit rating framework (Standard and Poor's 2013c, p. 5)

- **Institutional and Governance Effectiveness Score**: Primarily dependent on the effectiveness, stability, and predictability of a sovereign's policy-making and political institutions. A sovereign's credit rating may be adjusted by its perceived debt payment culture and external security risks.
- **Economic Score**: Key drivers of the economic structure and growth prospects are income levels, growth prospects, economic diversity, and economic volatility of a sovereign.

- **External Score**: Three factors determine a sovereign's external liquidity and internal investment position, namely the status of the currency in international transactions, external liquidity and external indebtedness of the sovereign.
- **Fiscal Score**: Reflects the fiscal flexibility and performance of a sovereign, combined with its debt burden and sustainability of deficits. Qualitative factors do exist in the assessment of revenue and expenditure flexibility, vulnerabilities, and long-term trends.
- **Monetary Score**: A sovereign's monetary flexibility is analyzed by its ability to coordinate monetary policy and exchange rate, inflation trends over an economic cycle, and the development of financial markets as well as debt markets. Sovereigns part of monetary unions naturally have less flexibility in relation to sovereigns with an own central bank, which weakens the monetary score for sovereigns part of monetary unions.

In summary, the key quantitative indicators of S&P's sovereign rating methodology are listed in table 4.

Score	Key Indicators
Economic **Monetary**	GDP per capita Real GDP per capita (% change) Consumer price index (% change) Depository corporation claims (% change) Monetary base
External	Current account receipts (CAR) Official reserves Usable reserves Gross external financing needs (% of CAR plus usable reserves) Narrow net external debt/CAR (%) Current account balance/CAR (%) Net foreign direct investment (FDI)/GDP (%) Net external liabilities/CAR (%) Terms of trade
Fiscal	General government Change in general government debt as a percentage of GDP Net general government debt/GDP (%) General government liquid financial assets Gross general government debt/GDP (%) General government interest/general government revenues (%)

Table 4: Key quantitative indicators of S&P's sovereign rating methodology (Standard and Poor's 2013c, pp. 39-40)

In addition, Standard and Poor's (2013a, pp. 8-14) state that external imbalances associated with public or private sector excesses and political issues, often are the principal default indicators when rating sovereigns. In terms of the eurozone, S&P's believe that external imbalances represent the foundation of the recent crisis. What is more, common traits among defaulting sovereigns are external indebtedness and a weakening currency. Notably, the study indicates that GDP per capita and real economy indicators show mixed features for defaulting sovereigns. S&P's study also shows that changes in government debt represent a more accurate indicator of a decline in creditworthiness than general government deficits.

Nevertheless, S&P is of the opinion that there simply does not exist one single measure that consistently acts as adequate indicator for sovereign default. What makes sovereign default assessment even more challenging is that economic indicators, in particular fiscal and current account deficits, improve prior to default. One possible explanation might be an inflation rise boosting revenues before expenditures, thus initially easing the fiscal deficit. Additionally, nominal GDP rises with higher inflation, which improves deficit and debt burden ratios, but the improvements may not be sustainable (Standard and Poor's 2013a, p. 14).

One should note that the former Moody's senior director Joffe (2012, p. 350) states that credit rating agencies do not apply contemporary and academically developed techniques, such as logit or probit regression models, to model sovereign exposures.

3.2 Bloomberg's SRSK Model

Focusing on Bloomberg's proprietary sovereign risk model (SRSK), it is based exhaustively on financial, economic and political risk factors, with no subjective judgment as in the case of CRA models. The model firstly divides sovereigns into two types, reserve currency (RC) and non-reserve currency (NRC) sovereigns. RC sovereigns, such as USA, Japan and EMU members, have debt denominated in their own currency. On the contrary, NRC sovereigns which do not have debt denominated in their own currency, are instead required to hold significant quantities of RC sovereigns' currencies as foreign exchange reserves (Text Medic 2013b).

In comparison to S&P's methodology, Bloomberg's SRSK model only utilizes a few drivers of sovereign risk, listed in table 5. The main reflection Bloomberg makes is that RC sovereigns must continually be aware of not only revenues and debt, but also expenditures. RC sovereigns depend solely on surplus or deficit, i.e. the quantity available for interest payments. On the other hand, the focus for NRC countries is a measure of solvency, more specifically the ratio of currency reserves to external debt. In its denotation of debt, Bloomberg includes both short-term and a fraction of long-term debt. Additionally, debt

	Reserve-Currency	**Non-Reserve Currency**
Financial	Revenues Expenditures Debt	Reserves Debt
Economic	GDP Growth Banking Sector Health	GDP Growth Banking Sector Health
Political/Social	Political Risk Score	Political Risk Score

Table 5: Sovereign risk drivers in Bloomberg's SRSK model (Rosenberg and Singenellore 2013, p. 25)

and banking sector health in the SRSK model can be seen as a measure of the ability to pay back debt, while the political risk score measures the willingness to pay back debt (Text Medic 2013b). The issue of quantifying political risk may be solved by using qualitative estimates from either the Economist Intelligence Unit or the PRS Group, which both publish a global political risk index (Text Medic 2013a).

In comparison to the list of sovereign default risk drivers from Rosenberg and Singenellore (2013, p. 11) in table 1, Bloomberg's SRSK model incorporates significantly fewer factors. Bloomberg motivates this methodology by stating that factors such as trade deficits, foreign exchange trends and capital flows are strongly related to the factors in the SRSK model. In a statistical sense, the model would therefore not benefit from additional factors (Text Medic 2013b).

The SRSK model quantifies a one-year PD range by utilizing a ten-year period for which all input data is presented. Notably, the model is independent of CRA and credit market data, and instead of calibrating a CDS model based on CDS market prices, Bloomberg uses the factors that drive sovereign default risk to also predict an intrinsic estimate of 5-year CDS contract (Text Medic 2013b).

4 Rating Model Development

The following section presents each step to developing a rating model, specifically for the purpose of designing a SRA model for sovereign PD estimation. The main phases of the rating model development are found in figure 10. All steps will be described in detail apart from the final phase, *extensions*, which includes possible additions to the model. Nonetheless, before turning specifically towards the SRA, the methodology of calibrating a PD scale from external ratings will be reviewed.

Figure 10: Rating model development steps (Kuhn 2012, p. 17)

4.1 Calibration of External Ratings to PD

The reason for calibrating to external ratings specifically and no other external data source such as CDS spreads, stems from the fact that we wish to map to an external default frequency. Typically, such approaches are through-the-cycle (TTC) oriented, i.e. long-run oriented, rather than point-in-time (PIT). CRAs adopt a TTC methodology when constructing credit ratings, which makes them the most suitable option to apply in our case. Nevertheless, having compared the respective characteristics of the external market data sources in section 2.1.4, it is clear that there is no leading proxy for sovereign credit risk.

As previously reviewed in section 2.4, the works of Orth (2011) on the EB estimator will be utilized. In order to produce EB estimates, sovereign marginal DR are obtained from S&P's long-term credit ratings and default histories on 130 rated sovereigns between the period January 1975 and April 2011. The fact that the set only contains 15 default events represents a typical case of LDP. In order to estimate the prior in the EB methodology, data from S&P Capital IQ on S&P's default histories and ratings of North American public firms between January 1981 and April 2011 will be used. Just as for the sovereign credit ratings, the public firms' ratings are long-term issuer credit ratings. The corporate dataset is substantially larger containing 5355 rated firms, and 755 defaults over the time period.

However, in our case the data used to compute the EB estimator are limited to pooled rating grades, measured on an ordinal rating scale of seven rating grades.

This seven-point ordinal rating scale is not preferable to apply when calculating EC, since a downgrade would imply a very large jump in capital requirement. Therefore, a more fine granular scale is required, which is done by adding a (+) or (-) to ratings 'AA' to 'CCC' of the rating scale in appendix section A, to showcase relative standing within each rating category. Additionally, one of the most important requirements of PD estimates is the property of monotonicity, i.e. a better rating grade *must* have a lower PD estimate than a worse rating grade. When using DR to estimate PDs, the resulting estimates may not always possess this trait.

In order to obtain monotonicity as well as expanding the rating scale for (+) and (-) grades, we will proceed from a method outlined by Erlenmaier (2011, pp. 47-48). Erlenmaier (2011) firstly regresses the logarithm of calibrated PDs (EB estimates in our case) to the ordinal rating scale, and secondly interpolates in place of non-monotonic PDs, thus generating a monotonic PD curve. The logarithmic function is applied since $ln(PD)$ are approximately linear against the ordinal rating scale. However, when re-transforming the logarithmic PDs using the exponential function, resulting PDs may theoretically exceed 1. This would imply that the PD exceeds 100%, which is clearly not feasible. In other words, it is necessary to bound PDs within the interval $(0,1)$. This is obtained by applying the logit function to PDs as a replacement of the logarithmic transformation, since the logistic distribution function maps the regression to the interval $(0,1)$.

The logit transform is defined as

$$logit\,(x) = ln\left(\frac{x}{1-x}\right), \tag{4.1}$$

where $\frac{x}{1-x}$ is called the *odds ratio*, i.e. the relation between probability of default and probability of survival (Trueck and Rachev 2009, p. 22). The regression performed can therefore be defined as

$$logit\,(EB) = \beta_0 + \beta_1 ORS, \tag{4.2}$$

where EB represents the EB estimates, ORS the ordinal rating scale, β_0 the intercept and β_1 the coefficient for the ordinal rating scale. Next, we interpolate between non-zero EB estimates, and if necessary extrapolate for zero EB estimates, typically including ratings above 'A'. The extrapolation is justified since logit-transformed EB estimates are approximately linear with the ordinal rating scale. Ultimately, the inverse logit function is applied

$$logit^{-1}(x) = \frac{1}{1 + exp(-x)} \tag{4.3}$$

to obtain the calibrated PD scale. In coming sections, let the resulting PDs corresponding to each rating category be denoted as ratings-implied probability of default (RIPD).

4.2 Modular Design

Having calibrated external ratings to PDs, the rating model development begins with designing the database. Firstly, four modular approaches will be developed when designing the database. In order to maximize statistical and economic relevance, each approach will consist of a homogeneous sub-segment. The necessary selection criterion to include a sovereign in each development sample is the existence and availability of external ratings. Since the primary interest is to evaluate EU sovereigns and sovereigns in close geographical proximity, the development samples will consist exclusively of such sovereigns.

The first approach is determined by World Bank's country classification (World Bank 2013), based on each sovereign's gross national income per capita. The development and validation samples are divided into sets of high income OECD sovereigns and non-OECD sovereigns, which are depicted in the appendices B, table 16. GIIIPS sovereigns as well as Cyprus and Slovenia are not included in the sample.

In the second approach, detailed in appendices table 17, sovereign are divided according to Bloomberg's methodology of reserve currency and non-reserve currency. The classification of reserve currencies is based on International Monetary Fund's currency composition of official foreign exchange reserves (COFER) classification (International Monetary Fund 2013). GIIIPS sovereigns as well as Cyprus and Slovenia are not included in the sample.

The third approach consists of sovereigns which have been hit hardest by the European sovereign-debt crisis, in terms of both growing government debt levels as well as severely diminishing creditworthiness. The approach does not include a validation sample, as it is specifically developed towards GIIIPS, Cyprus, and Slovenia exposures. Table 18 in appendix section B provides an overview of the approach.

The fourth approach combines the prior three approaches, i.e. the development sample consists of sovereigns included in the development samples of three first approaches. Similarly, the validation sample consists of sovereigns included in the two first approaches' development samples. Table 19 in appendix section B overviews the fourth approach.

Briefly commenting on the modular approaches, if GIIIPS, Cyprus and Slovenia were to be included in either the World Bank's country classification and IMF's COFER classification approaches, the resulting PD estimates for GIIPS, Cyprus and Slovenia would be largely underrated in relation to RIPDs. Therefore, these have been chosen as a separate sub-segment.

Moreover, transcontinental sovereigns such as Turkey are part of development samples, because of the ever increasing relationship between the EU and Turkey. Meanwhile, for instance Russia is chosen as part of the validation samples, as the prospects of Russia embracing a stronger link than currently with the EU are slim.

4.3 Factor Long-List and Data Generation

When constructing the factor long-list of potential factors indicating a sovereign's creditworthiness, both quantitative and qualitative categories must be taken into account. Since SRA is applied, the aim is to replicate the TTC oriented external ratings. S&P's methodology reviewed in chapter 3 is based exclusively on country-specific factors, implying that country-specific factors are of primary interest as factor inputs. Moreover, S&P's and Bloomberg's methodology are both based on economic, external, fiscal, monetary, and political indicators. In addition to S&P's and Bloomberg's categories, Izzi et al. (2011) have compiled a list of quantitative and qualitative categories, detailed in table 6.

Quantitative	**Qualitative**
Banking System	Debt Servicing Record
Current Account	Economic Conditions
Debt	Foreign Relations
Government Finance	Stability of the Financial System
Growth	Social and Political Conditions
Liquidity	
Monetary Policy	

Table 6: Categories for quantitative and qualitative factor long-list (Izzi et al. 2011).

The data on quantitative and qualitative factors have primarily been collected from IMF and World Bank, in order to obtain a cohesive dataset. Using several data sources may cause deviations in data. In addition, the data generation includes the collection of external ratings for each included sovereign. Izzi et al.

(2011) indicate the minimum data requirements of the SRA, for which the development and validation sample should include at least 100 and 50 observations respectively. In order to cope with the requirements of Izzi et al. (2011) as well the IRB minimum requirements listed in section 2.2.4, annual data between years $2000 - 2011$ will be included in the samples, i.e. the datasets contain one observation per year for a given sovereign. The available 2012 data will be used out-of-sample for EC calculations. The restriction of annual data is mainly dependent on the fact that several indicators are only published on an annual basis. Conscious of the minimum requirements stated by Izzi et al. (2011), the GIIIPS, Cyprus and Slovenia dataset does fulfill the requirements of at least 100 observations. Nonetheless, placing the sovereigns as part of a separate sub-segment makes both economical and statistical sense, than having them as part of the first and second approach.

It is also important to note that the time interval between factors and external ratings are shifted according to the PD time horizon. Intuitively, since the aim is to estimate PD one year from the evaluation of a sovereign, quantitative and qualitative factors relative to year t will correspond to RIPD relative to year $t + 1$. Therefore, one must account for the one-year time lag, and construct the dataset accordingly (Izzi et al. 2011).

The most demanding part of the rating model development is in fact data generation and identification of economically meaningful factors. The main challenges lie in the fact that certain sovereigns may have different macroeconomic indicators that are not fully comparable with other sovereigns' indicators. Furthermore, it is not feasible to estimate a separate model for each sovereign, due to the data restrictions previously mentioned. Therefore, it is of significant importance to find indicators that are comparable across a sub-segment of sovereigns. Potential testing of variations in factors, such as the use of variance or absolute value of factors may be relevant for the single-factor analysis (SFA), detailed in section 4.4.

Moreover, some degree of data cleansing is required to obtain reliable observations. For instance, observations with detected outliers should be removed, and missing factor values may be filled by interpolation or result in a deletion of the corresponding observation.

The long-list of factors is not provided in detail, since the vast majority are filtered out in the SFA.

4.4 Single-Factor Analysis

In short, the aim of SFA is to attain a factor shortlist explaining the creditworthiness of a sovereign. This is accomplished by initially performing a factor transformation as well as using a measure of discriminatory power, namely the shadow accuracy ratio (SAR). Additionally, factors should be checked for representativeness as well as the existence of highly correlated factors.

4.4.1 Factor Transformation

The first part of SFA consists of transforming factors in order to attain more insightful results in the multi-factor analysis (MFA), detailed in section 4.5. Kuhn (2012), Erlenmaier (2011) and Zaalberg (2013) all suggest the use of the previously introduced inverse logit function in section 4.1, which maps a factor score on the real axis to $(0, 1)$. The fact that all factors are given the same range as well as diminishing effects of outliers, represent the two major reasons to apply the transformation. When performing the inverse logit transformation, factor distributions inversely related to creditworthiness are adjusted in order to obtain an intuitive final model. In other words, factors are adjusted resulting in higher factor scores being positively correlated with creditworthiness, and therefore correspond to lower RIPD. In addition, the transformation is monotone increasing, which preserves the ordering of the factor scores (Zaalberg 2013, p. 15).

In detail, equation 4.4 depicts the applied transformation, where parameters a and b represent the horizontal translation and steepness of the transform respectively.

$$f(x) = \frac{1}{1 + exp(a + bx)} \tag{4.4}$$

The parameters are determined separately for each factor, providing the best possible fit according to the empirical distribution function of the factor scores. Factors inversely related to creditworthiness receive parameters $-a$ and $-b$ to adjust for positive relation with creditworthiness (Zaalberg 2013, p. 16).

Shadow Accuracy Ratio

Moving to the second part of SFA, Izzi et al. (2011) define SAR as a measure of single-factor rank ordering power, i.e. a factor's predictive power on a stand-alone basis. In general, SAR compares single-factor scores to the ideal model, which orders sovereigns in the best possible way according to RIPD. The closer a factor is to the ideal model, the more useful the factor is for replicating external ratings, and subsequently receives a larger SAR value. A SAR near one implies high discriminatory power, while a value near zero represents a factor with no discriminatory power.

In order to compute SAR, the ranking power (RP) of a factor needs to be calculated. This is accomplished by determining areas defined in figure 11 by the random model curve, power curve, crystal ball curve, and perfect discriminatory model curve (Izzi et al. 2011). For a single-factor, the *power curve* is obtained by first ordering factor scores ascendingly. Next, for a certain factor score, the sum of RIPDs of observations corresponding to lower or equal factor scores is divided by the sum of RIPDs corresponding to all factor scores. The proportion determined is then graphically plotted against the proportion of all observations with lower or equal factor score. This procedure is done for each factor score, thus constructing the power curve (Zaalberg 2013, p. 13).

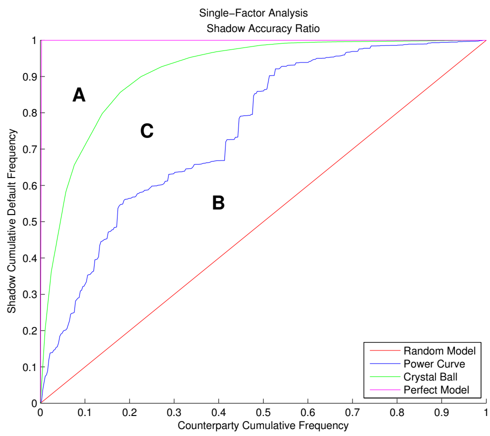

Figure 11: SAR of a single factor, recreated from Izzi et al. (2011).

The methodology behind the construction of the power curve may be formalized by determining the shadow cumulative default frequency (SCDF), which represents the y-axis in figure 11 as

$$SCDF_1 = \frac{RIPD_1}{\sum_{j=1}^{n} RIPD_j} \tag{4.5}$$

$$SCDF_i = SCDF_{i-1} + \frac{RIPD_i}{\sum_{j=1}^{n} RIPD_j},\ \forall\, i = 2, \ldots, n \tag{4.6}$$

where n is the number of observations (Izzi et al. 2011).

Furthermore, the *crystal ball* represents the ideal model. That is, if factor scores have full discriminatory power, it would correspond to having a crystal ball and predicting future creditworthiness of sovereigns flawlessly. In a scenario where factor scores follow the crystal ball model, factor scores will be ordered inversely to the ordering of RIPDs, since higher factor scores imply lower RIPD as a result of the factor transformation (Zaalberg 2013, p. 14).

If factor scores have no discriminatory power, the *random model* will be imminent. In such cases, the factor score of an observation will not determine any sort of creditworthiness at all, therefore yielding a straight line with constant slope when plotting SAR (Zaalberg 2013, p. 14).

The *perfect discriminatory model* curve is determined by calculating the shadow default rate (SDR) as the average RIPD across the sub-segment:

$$SDR = \frac{\sum_{j=1}^{n} RIPD_j}{n}. \tag{4.7}$$

It is now possible to determine RP_{factor} and $RP_{\text{crystal ball}}$ as a ratio of several areas:

$$RP_{\text{factor}} = \frac{Area(A)}{Area(A) + Area(B)} \tag{4.8}$$

$$RP_{\text{crystal ball}} = \frac{Area(C)}{Area(C) + Area(B)} \tag{4.9}$$

Ultimately, SAR_{factor} is computed as the ratio between RP_{factor} and $RP_{\text{crystal ball}}$:

$$SAR_{\text{factor}} = \frac{RP_{\text{factor}}}{RP_{\text{crystal ball}}}. \tag{4.10}$$

A maximum SAR value of one would imply that the power curve is identical to the crystal ball curve (Izzi et al. 2011).

4.4.2 Factor Rank-Correlation Analysis and Representativeness

Having selected a number of factors with high discriminatory power, one has to make sure that the factors are not highly correlated. Therefore, the next step is to perform a factor rank-correlation analysis. Each factor is evaluated on its pairwise correlation with the other factors. The correlation level for which two variables should be excluded is not rigid. Instead, it is of higher importance to distinguish if highly correlated factors belong to the same information category or are different measures in an economical sense (Izzi et al. 2011). In general, one should be alarmed if two factors have a correlation of more than 75% with one another (Jo 2008, p. 22).

Furthermore, factor representativeness should be checked. Representativeness indicates whether a factor represents a predictor of creditworthiness for an entire homogeneous sub-segment. This analysis may for instance be performed by comparing risk factor distributions of the development and validation sample (Erlenmaier 2011, p. 62).

4.5 Multi-Factor Analysis

Having completed SFA and ultimately selected a factor shortlist according to discriminatory power, factor-rank correlation, and representativeness, the next step of the rating model development is the multi-factor analysis. The major part of the MFA consists of regressing logit-transformed RIPDs to a subset of the factor shortlist best replicating RIPDs, in order to obtain a multi-factor model with factor-specific weights. In other words, the aim is to construct a model for the discrete dependent variable $logit(RIPD)$ based on the explanatory subset of shortlist factors, while accounting for their interdependencies. Among the strongest models, the economically most meaningful will be chosen (Kuhn 2012, p. 17). In addition, practitioners are often interested in determining each factor's influence on the external ratings within a specific sub-segment of sovereigns. Consequently, a measure of influence will be constructed based on the factor-specific weights.

4.5.1 Multiple Linear OLS Regression

Essentially, attention must first be given to model selection, which includes choice of model type as well as determining the subset of factors included. With regard to the choice of model type, the most frequently applied MFA technique is multiple linear regression with logit-transformed external PDs as dependent variable, and a set of explanatory factors. The logit transform fulfills the compatibility requirement of linear relationship assumed by the regression model, as well as maintaining calibrated PDs within the interval $(0, 1)$. Additionally, a stepwise method is commonly used to determine the combination of factors included in the model (Izzi et al. 2011).

Focusing on the stepwise method, the factors selected for the final model are firstly chosen from the SFA produced shortlist of k factors. As a result, the final model may consist of 2^k possible factor combinations. If k is large, it is recommended to utilize a stepwise method instead of using brute force to check all possible combinations. A stepwise forward selection builds the model step by step, initially adding the factor yielding highest stand-alone predictive power. The second factor is chosen based on the combined predictive power of the first and second factor. The iteration continues until additional factors do not yield a significant improvement of the model within a certain threshold. The significance threshold in our analysis is set to 5%. Additionally, factors are only added if its sign is intuitive. Since all factors have previously been transformed, a higher factor score should correspond to a lower RIPD, i.e. factors are only added they have a negative sign (Zaalberg 2013, pp. 17-18).

Focusing on the model type, the linear regression for SRA rating systems is typically denoted:

$$logit(RIPD_i) = \beta_0 + \beta_1 x_{i,1} + \ldots + \beta_m x_{i,m} + \epsilon_i, \ \forall\, i = 1, \ldots, n \tag{4.11}$$

where $x_{i,j}$ is the score of factor j corresponding to observation i, ϵ_i the i:th observation residual, and $\beta_0, \ldots, \beta_m$ the regression coefficients, i.e. factor-specific weights. The regression coefficients are estimated by means of ordinary least squares (OLS), i.e. minimizing $\sum_{i=1}^{n} \epsilon_i^2$. For the derivation of the OLS estimator, we refer to section C in the appendices. Furthermore, three essential stochastic assumptions of the residuals need to be fulfilled. Residuals must namely be normally distributed, independent of each other as well as homoskedastic, i.e. all residuals having the same standard deviation (Erlenmaier 2011, pp. 66-68).

When considering normality, residual distribution plots and statistical tests are primarily used. In the event of statistical tests rejecting the hypothesis of normality, it is worth noting that estimators still are BLUE (best linear unbiased estimator). In addition, if sample size is sufficiently large, convergence of estimates' confidence intervals and related statistical tests is achieved, implying that residuals are approximately normally distributed. However, violating the assumption of homoskedasticity or independence is more severe. Both assumptions are checked by statistical tests on the covariance matrix of the residuals. The derivation of the covariance matrix is found in appendix section C. Homoskedasticity implies identical values of each entry of the diagonal, and independence is present if non-diagonal entries are zeros. If either heteroskedasticity, i.e. differing standard deviations of residuals, or autocorrelation, i.e. serially correlated residuals, the structure of the covariance requires adjustment in order to obtain consistent estimates (Erlenmaier 2011, pp. 68-69). The potential adjustment is conducted through Matlab's built-in function `hac`, which is a heteroscedasticity and autocorrelation consistent covariance estimator for OLS coefficient estimates.

4.5.2 Measuring Factor Influence

From the estimated factor-specific weights, i.e. regression coefficients, each factor's influence on the external ratings may be determined. Erlenmaier (2011, pp. 70-71) details a commonly used method for measuring factor influence. Firstly, an adjusted factor coefficient is constructed by multiplying a factor's weight by the factor's standard deviation. Each factor's influence is then determined by mapping the adjusted factor coefficients to the interval $[0, 1]$. As a result, the sum of the absolute value of all mapped coefficients should add up to one. Interpreting the obtained coefficients, a coefficient x_j depicts to which degree it affects logit PDs predicted by the regression model, given all other mapped factor coefficients $x_{k,k \neq j}$ are held constant. Formally, weight w_j for measuring influence of factor x_j with regression coefficient β_j is computed

$$w_j = \frac{w_j^*}{|w_1^*| + \ldots + |w_m^*|}, \tag{4.12}$$

where the adjusted factor coefficient w_j^* is calculated by the use of the standard deviation operator denoted σ,

$$w_j^* = \beta_j \sigma(x_j). \tag{4.13}$$

4.6 Manual Adjustments and Model Calibration

Once the MFA has provided factor-specific weights from the regression as well as each factor's influence on the external ratings, the measure of influence may be used for manual adjustments in order to maximize the fit to RIPDs, and subsequently calibrating PDs from the derived model.

The reason to manually adjust estimates from the statistical model may stem from inadequacies such as unsatisfactory representativeness of the development sample, lacking empirical basis for the specific sub-segments or potential expert judgement that differs from estimates. However, when performing a manual adjustment, the model must not experience a significant decrease in discriminatory power. This is ensured by applying the validation methods described in section 4.7 (Erlenmaier 2011, p. 72).

In detail, the manual adjustment is performed by first constructing an aggregated score S_i from factor scores $x_{i,m}$, and its respective factor weight w_m for measuring influence:

$$S_i = w_1 x_{i,1} + \ldots + w_m x_{i,m}, \ \forall \, i = 1, \ldots, n \tag{4.14}$$

The aggregated score is then regressed against logit-transformed RIPDs:

$$logit(RIPD_i) = c_0 + c_1 S_i + \epsilon_i, \; \forall\, i = 1, \ldots, n. \tag{4.15}$$

Thus, the regression provides estimates for coefficients c_0, which corresponds to the average predicted logit PD of the sub-segment sample, and c_1, which measures the rate of which predicted logit PDs vary across the sample. If the estimates for c_0 and c_1 are close to β_0 and $\beta_{\sum} = \beta_1 + \ldots + \beta_m$ in equation 4.11, the general properties of the model will not have been vastly altered. Ultimately, calibrated PDs are computed by applying the inverse logit transform to each logit PD observation i predicted by the regression in equation 4.15 (Erlenmaier 2011, p. 72).

4.7 Model Validation

Turning our attention towards model validation, the SRA rating model development does not allow for a significantly extensive validation process (Erlenmaier 2011, p. 74). Therefore, only the most essential validation measures are presented and utilized. Briefly itemized, the following validation tools are applied:

Outlier and Residual Analysis: Verification of stochastic assumptions on the residuals, in addition to potential removal of observation outliers, which may distort regression estimates.

Discriminatory Power: Computing SAR for calibrated development and validation sample PDs.

PD Bucketing: Verification that calibrated PDs are within one or two rating notches of the rating grades corresponding to RIPDs.

Formal Statistical Tests: Spiegelhalter test measuring the quality of the rating model's PD calibration.

The outlier and residual analysis is exhaustively based on multiple linear regression residual and diagnostics plots. These tools are not specific to the SRA development process, and will therefore not be reviewed in detail. Furthermore, the discriminatory power measure SAR is reviewed in full detail in section 4.4. Nevertheless, PD bucketing as well as the Spiegelhalter test require some attention.

4.7.1 PD Bucketing

When banks develop rating models, the final step usually involves mapping calibrated PDs to an internal master scale for internal use. The mapping is done by creating buckets for the calibrated PDs, i.e. a certain interval for the calibrated PDs correspond to a certain bucket (Zaalberg 2013, p. 20). Following intuition, a simple mapping strategy based on intermediate values of RIPDs as bucket thresholds is applied. For instance, if $RIPD('AAA') = 1\%$ and $RIPD('AA+') = 2\%$, then the threshold dividing the two rating buckets is simply 1.5%. For validation purposes, it possible to verify that the calibrated PD buckets are within one or two rating notches of rating grades corresponding to RIPDs (Izzi et al. 2011). As for EC calculations, the bucket PD estimates for the 2012 data sample is used as PD input parameter in the IRB formulas.

4.7.2 Spiegelhalter Test

Ultimately, the Spiegelhalter test is briefly reviewed. As a statistical test reflecting the quality of a rating model, the Spiegelhalter test is based around the mean square error

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(RIPD_i - PD_i)^2, \tag{4.16}$$

where N is the number of observations, $RIPD_i$ is the RIPD corresponding to observation i, and PD_i is the calibrated PD for observation i. In general, a low MSE is an indication of a well-performing rating model. Spiegelhalter tests whether the observed MSE differs from its expected value. Formally, the following hypotheses are tested:

$$H_0 : MSE = \mathbb{E}[MSE] \tag{4.17}$$

$$H_1 : MSE \neq \mathbb{E}[MSE] \tag{4.18}$$

Under H_0, the expected value and variance of MSE is:

$$\mathbb{E}[MSE] = \frac{1}{N}\sum_{i=1}^{N} PD_i \cdot (1 - PD_i) \tag{4.19}$$

$$\mathbb{V}[MSE] = \frac{1}{N^2}\sum_{i=1}^{N}(1 - 2PD_i)^2 \cdot PD_i \cdot (1 - PD_i). \tag{4.20}$$

Under H_0, the distribution of the standardized MSE, Z_s, is approximately standard normally distributed according to the central limit theorem.

$$\begin{aligned} Z_S &= \frac{MSE - \mathbb{E}[MSE]}{\mathbb{V}[MSE]^{0.5}} = \\ &= \frac{\frac{1}{N}\sum\limits_{i=1}^{N}(RIPD_i - PD_i)^2 - \frac{1}{N}\sum\limits_{i=1}^{N} PD_i \cdot (1 - PD_i)}{\sqrt{\frac{1}{N^2}\sum\limits_{i=1}^{N}(1 - 2PD_i)^2 \cdot PD_i \cdot (1 - PD_i)}} \end{aligned} \tag{4.21}$$

Therefore, a joint test can be applied at suitable significance level in order to see if MSE is sufficiently small to suggest that the calibrated PDs equal true PDs (Erlenmaier 2011, pp. 319-320).

5 Empirical Findings and Model Analysis

The following section will detail the empirical findings, and review every aspect of the model output. For intuition, the section will in the main follow the outline of the rating model development.

5.1 Empirical Bayes PD Estimates

Firstly, EB estimates and the corresponding calibrated RIPDs will be displayed. In table 7, default rates for sovereigns and corporates are listed together with the resulting EB estimator. Regarding sovereign DR, one may initially note the

	Sovereign DR (%)	**Corporate DR (%)**	**Sovereign EB Estimator (%)**
AAA	0.00	0.00	0.00
AA	0.00	0.00	0.00
A	0.00	0.06	0.06
BBB	0.00	0.20	0.20
BB	0.56	0.77	0.76
B	2.60	4.47	3.88
CCC-C	32.27	24.20	24.38

Table 7: One-year DR for sovereigns and corporates, as well as sovereign EB estimates (%)

several zero default rates, which is an unsatisfactory feature considering that CDS spreads are traded for highly-rated sovereigns, i.e. there exists credit risk for exposures towards such sovereigns. In order to enforce more conservative estimates, corporate DR are used as the prior to attain EB estimates. Due to the sovereign sample size, which is small relative to the corporate sample, one may clearly see that the corporate DR dominates the EB estimates. This follows from the nature of Bayesian analysis. However, EB estimates for 'B'-rated sovereigns are not as close to corporate DR as other rating grades, since we have reasonable amount of information for 'B'-rated sovereigns. There exists a few defaults in the sample of 'B'-rated sovereigns, in addition to not having too few 'B'-rated sovereigns overall in the sample. In general, the EB estimator enforces a reasonable degree of conservatism to highly rated obligors.

With regard to the final RIPD estimates, the fine granular RIPD scale as well as logit-transformed RIPDs are displayed in figure 12. Also, the resulting RIPD scale is more detailedly defined in table 8. A monotonic PD scale is obtained as well as extrapolated estimates for zero EB estimates. The estimates from the extrapolation are a result of the logit-transformed RIPDs approximative linear relationship with the ordinal rating scale, shown by the best fitted least-squares.

	Calibrated RIPD (%)
AAA	0.0044
AA+	0.0074
AA	0.0125
AA-	0.0211
A+	0.0356
A	0.0600
A-	0.0896
BBB+	0.1339
BBB	0.2000
BBB-	0.3123
BB+	0.4875
BB	0.7600
BB-	1.3152
B+	2.2669
B	3.8800
B-	6.3548
CCC+	10.2400
CCC	16.0920
CCC-	24.3800
CC	31.8030

Table 8: Calibrated RIPD scale (%)

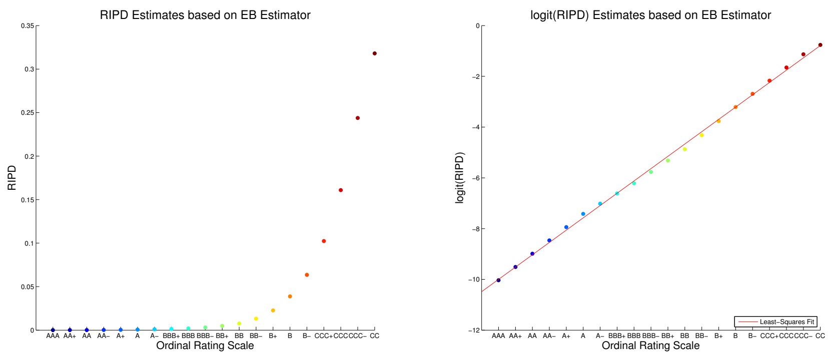

Figure 12: Calibrated RIPD scale and logit-transformed RIPDs.

5.2 Significant MFA Factors

Next, the significant factors included in the MFA of each modular approach as well as results on each factor's discriminatory power will be presented. In table 9, having excluded observations outliers, the factors included in the MFA of each modular approach are listed. Notably, GDP per capita is the only indicator part of all approaches. There exist two factors which have been adjusted in respect to their raw data format, namely absolute deviation inflation and 3Y GDP growth volatility. Regarding the first factor, the possibility of a sovereign having negative inflation exists. In order to facilitate that negative inflation does not imply a higher creditworthiness, while maintaining that high inflation implies low creditworthiness, one may set an optimal level of inflation rate.

The factor absolute deviation inflation may then be calculated by applying the absolute value to the difference between the optimal level and a sovereign's inflation rate. Furthermore, 3Y GDP growth volatility is calculated by taking the standard deviation of the previous three years of GDP growth. The factor has been found to have substantially larger discriminatory power than using the raw yearly GDP growth data.

Modular Approach 1: World Bank's Country Classification

OECD	**Non-OECD**
Absolute Deviation Inflation (%)	Absolute Deviation Inflation (%)
Constant GDP per Capita (USD)	Constant GDP per Capita (USD)
EIU Political Risk Score	Non-Performing Loans (% Total Gross Loans)

Modular Approach 2: IMF's COFER Classification

RC	**NRC**
Absolute Deviation Inflation (%)	Absolute Deviation Inflation (%)
Constant GDP per Capita (USD)	Constant GDP per Capita (USD)
EIU Political Risk Score	Non-Performing Loans (% Total Gross Loans)
	EIU Political Risk Score

Modular Approach 3: GIIIPS, Cyprus & Slovenia

Constant GDP per Capita (USD)
Gross National Savings (% GDP)
Non-Performing Loans (% Total Gross Loans)

Modular Approach 4: Full Dataset

Absolute Deviation Inflation (%)
Constant GDP per Capita (USD)
3Y GDP Growth Volatility (%)
Non-Performing Loans (% Total Gross Loans)
EIU Political Risk Score

Table 9: List of significant indicators in MFA

Note that certain manual adjustments have been made to the choice of factors by the automatic stepwise method. For instance, the factor gross government debt has relatively high discriminatory power and is included by the stepwise method for OECD and RC approaches. However, in spite of improving the model sample, the factor distorts the validation sample. As a result, the factor lacks sufficient representativeness for the sub-segment, and is therefore excluded from the MFA. In addition, the stepwise method includes the EIU political risk score instead of GDP per capita for the non-OECD sample, since EIU provides better fit for the model sample. Replacing the EIU political risk score with GDP per capita, the model sample performance is not altered much. However, the validation sample sees a significant improvement, which justifies the switch between the factors. Note that it is not possible to include both factors for the non-OECD model sample, since one of the factors will not be significant.

Furthermore, in table 10 each factor's ranking power and SAR is listed. In addition, the ranking power of the crystal ball approach is also listed, since it is used to compute SAR. In the main, SAR values for significant factors lie in the range $0.6 - 1$, which provides sufficient discrimination towards RIPDs. Nevertheless, having a number of factors with high discriminatory power does not automatically render sufficiently good results in the MFA. The interactions between factors are vital, which strongly impacts factor-specific weights.

Approach 1 - OECD	RP_{factor}	$RP_{\text{crystal ball}}$	SAR_{factor}
Absolute Deviation Inflation	0.4612	0.6019	0.7663
Constant GDP per Capita	0.5928	0.6019	0.9850
EIU Political Risk Score	0.5895	0.6019	0.9794
Approach 1 - Non-OECD	RP_{factor}	$RP_{\text{crystal ball}}$	SAR_{factor}
Absolute Deviation Inflation	0.4013	0.6190	0.6483
Constant GDP per Capita	0.5271	0.6190	0.8516
Non-Performing Loans	0.3999	0.6190	0.6460
Approach 2 - RC	RP_{factor}	$RP_{\text{crystal ball}}$	SAR_{factor}
Absolute Deviation Inflation	0.4884	0.6061	0.8057
Constant GDP per Capita	0.5947	0.6061	0.9811
EIU Political Risk Score	0.5830	0.6061	0.9618
Approach 2 - NRC	RP_{factor}	$RP_{\text{crystal ball}}$	SAR_{factor}
Absolute Deviation Inflation	0.5703	0.7209	0.7910
Constant GDP per Capita	0.6676	0.7209	0.9260
Non-Performing Loans	0.5577	0.7209	0.7737
EIU Political Risk Score	0.6819	0.7209	0.9459
Approach 3 - GIIIPS	RP_{factor}	$RP_{\text{crystal ball}}$	SAR_{factor}
Constant GDP per Capita	0.9351	0.9755	0.9586
Gross National Savings	0.9719	0.9755	0.9963
Non-Performing Loans	0.9715	0.9755	0.9959
Approach 4 - Full Dataset	RP_{factor}	$RP_{\text{crystal ball}}$	SAR_{factor}
Absolute Deviation Inflation	0.7340	0.8258	0.8888
Constant GDP per Capita	0.8053	0.8258	0.9752
3Y GDP Growth Volatility	0.6749	0.8258	0.8172
Non-Performing Loans	0.7426	0.8258	0.8992
EIU Political Risk Score	0.8108	0.8258	0.9819

Table 10: SFA results: RP and SAR of significant MFA factors.

Additionally, the factors listed in table 9 and 10 all have sufficient representativeness for each sub-segment, and have passed the factor-rank correlation analysis.

	Approach 1		**Approach 2**		**Approach 3**	**Approach 4**
	OECD	Non-OECD	RC	NRC	GIIIPS	Full Dataset
(Intercept)	-8.1391***	-6.1912***	-8.2142***	-4.8019***	-7.1616***	-6.7713***
	(0.2246)	(0.5655)	(0.2714)	(0.6794)	(0.6530)	(0.5078)
Constant GDP per Capita	-1.4234***	-1.9153***	-2.2831***	-1.8836*	-1.3045*	-0.9611*
	(0.3275)	(0.4897)	(0.3861)	(0.7872)	(0.6061)	(0.4239)
Absolute Deviation Inflation	-0.7343***	-1.7920***	-0.9019***	-1.2579*		-1.3650**
	(0.2246)	(0.4183)	(0.2187)	(0.5203)		(0.4227)
EIU Political Risk Score	-1.7787***		-0.9939***	-3.8930***		-4.1823***
	(0.2882)		(0.2881)	(1.0321)		(0.6571)
Non-Performing Loans		-1.2990***		-1.0889**	-3.4192***	-1.2629***
		(0.3696)		(0.3528)	(0.6513)	(0.3385)
Gross National Savings					-3.4069***	
					(0.6377)	
3Y GDP Growth Volatility						-0.4921*
						(0.2173)
R^2	0.820	0.568	0.810	0.864	0.587	0.852
N	166	125	120	174	76	368

Standard errors corrected for heteroskedasticity and autocorrelation in brackets.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: MFA results: Heteroskedasticity and autocorrelation corrected MFA regression output

5.3 MFA Regression, Measure of Influence and Calibration

In the following section, results obtained from the MFA are detailed. More specifically, table 11 displays the multiple linear OLS regression output, including intercepts, regression coefficients, factor significance levels p, goodness-of-fit R^2 as well as number of observations N of each modular approach. Most notably, the GIIIPS and non-OECD model samples provide low R^2. The GIIIPS model sample includes the observation of Greece's default, which severely diminishes the overall regression fit. In the end, the inclusion of the observation is justified since it provides more accurate factor-specific weights when computing PD for 2012 data. The non-OECD model sample consists of mainly low-rated sovereigns, and the low R^2 indicates that factors indicative of default seem to vary between these sovereigns. This is evident in table 10, since non-OECD SAR values are significantly lower in comparison to other approaches. The segment may benefit if divided into sub-segments, such as upper income non-OECD, middle income non-OECD and lower income non-OECD sovereigns. However, in such a case the issue of number of observations is imminent, as well as the fact that several non-OECD sovereigns have only recently began receiving external ratings. Similarly, the GIIIPS sample only contains 76 observations, which is not optimal. Nonetheless, including these particular sovereigns as part of OECD, non-OECD, RC and NRC would cause distortions in terms of determining which explanatory factors to include, and their respective weight in the MFA.

Furthermore, table 12 overviews factor regression coefficients x_j as well as each factor's influence w_j. Having accounted for factors' standard deviation, one may see that the influence of factors differ from the regression coefficients. For instance, in the non-OECD and NRC approach, absolute deviation inflation and non-performing loans nearly have the same influence, but the latter is given a notably smaller regression coefficient. This is due to non-performing loans having a larger standard deviation across the set of observations than absolute deviation inflation, therefore implying a larger factor weight of influence. Also notable in the GIIIPS approach, is that gross national savings receives a larger influence weight than non-performing loans, in contrast to the regression coefficients which are reversely related.

Using the factor influence weights, the regressions which calibrate PD estimates are listed in table 13. One may discern only slight deviations to R^2 in comparison to table 11, which is promising. In general, the intercepts are only slightly altered, indicating that the intercept of the MFA regression is nearly

Approach 1 - OECD	x_j	w_j
Absolute Deviation Inflation	-0.7343	0.1702
Constant GDP per Capita	-1.4234	0.3618
EIU Political Risk Score	-1.7787	0.4680
Approach 1 - Non-OECD	x_j	w_j
Absolute Deviation Inflation	-1.7920	0.3270
Constant GDP per Capita	-1.9153	0.3467
Non-Performing Loans	-1.5044	0.3263
Approach 2 - RC	x_j	w_j
Absolute Deviation Inflation	-0.9019	0.2082
Constant GDP per Capita	-2.2831	0.5434
EIU Political Risk Score	-0.9939	0.2483
Approach 2 - NRC	x_j	w_j
Absolute Deviation Inflation	-1.2579	0.1354
Constant GDP per Capita	-1.8836	0.2230
Non-Performing Loans	-1.0889	0.1373
EIU Political Risk Score	-3.8930	0.5043
Approach 3 - GIIIPS	x_j	w_j
Constant GDP per Capita	-1.3045	0.1546
Gross National Savings	-3.4069	0.4266
Non-Performing Loans	-3.4192	0.4188
Approach 4 - Full Dataset	x_j	w_j
Absolute Deviation Inflation	-1.3650	0.1391
Constant GDP per Capita	-0.9611	0.1254
3Y GDP Growth Volatility	-0.4921	0.0578
Non-Performing Loans	-1.2629	0.1470
EIU Political Risk Score	-4.1823	0.5307

Table 12: MFA results: Comparison of regression coefficient weights x_j and factor influence weights w_j of significant MFA factors

optimally estimated according to the fit to RIPDs. In terms of the OECD and RC approach, the intercept signals a low average predicted PD for sovereigns in relation to for instance non-OECD and NRC, which receive an intercept closer to zero indicating a higher average predicted PD. This is in line with expectations, since OECD and RC consist of a larger proportion of highly rated sovereigns in comparison to the sets of non-OECD and NRC. The full dataset has the largest deviation in terms of intercept, which suggests that the approach required further fitting to RIPDs. Moreover, the full dataset exhibits a score indicating that the set contains sovereigns rated across the entire rating scale. Reversely, the OECD score is less than half of the full datasets score, which relates to the fact that it exhaustively contains highly rated sovereigns. Similarly, at some point between $2000 - 2011$ the NRC and GIIIPS set contains both highly rated sovereigns in addition to lower rated, thus indicating a score similar to the full dataset.

	Approach 1		Approach 2		Approach 3	Approach 4
	OECD	Non-OECD	RC	NRC	GIIIPS	Full Dataset
(Intercept)	-8.089***	-6.2941***	-8.1874***	-4.6928***	-7.1514***	-6.5175***
	(0.0588)	(0.0954)	(0.0687)	(0.0805)	(0.1804)	(0.0508)
Score	-3.8929***	-5.1748***	-4.1623***	-8.0234***	-8.0848***	-8.1107***
	(0.1423)	(0.40935)	(0.1855)	(0.2429)	(0.7879)	(0.1774)
R^2	0.820	0.565	0.810	0.864	0.587	0.851
N	166	125	120	174	76	368

Standard errors in brackets.
*** $p < 0.001$

Table 13: MFA results: Manual adjustment regression output

5.4 Model Validation

Having analyzed the MFA output, the next step is to cast an eye on the model validation results, which essentially measure the performance of the different modular approaches. As a first measure, the discriminatory power provides an indication of model performance, displayed in table 14. However, one must note that the performance of calibrated PDs are tested in a discriminatory manner, and not in absolute terms of deviations from RIPD. Upon first sight, OECD, RC, Full Dataset and GIIIPS perform very well according to the SAR values. Notably, non-OECD and NRC perform similarly well in-sample, but worse out-of-sample. This possibly indicates that the set of indicators are not suitable for certain sovereigns part of the sub-segment or estimated factor-specific weights require adjustment. One should note that the validation samples contains sovereigns not part of EU region. Applying a discriminatory measure to 2012 samples would not yield sufficiently indicative results, due to the limited number of observations.

Approach	SAR_{model}	$SAR_{\text{validation}}$
OECD	0.9842	0.9533
Non-OECD	0.8836	0.6629
RC	0.9844	0.9093
NRC	0.9495	0.7076
GIIIPS	0.9996	
Full Dataset	0.9824	0.9160

Table 14: Validation results: Discriminatory power of model and validation samples

A second validation measure is the verification of rating class deviations of PD buckets, which illuminates the performance of calibrated PDs in absolute terms. The results in table 15 indicate that OECD and RC approaches perform very well, both in and out-of-sample as well as for computing PD based on 2012 data. Furthermore, the measure confirms that non-OECD and NRC approaches perform adequately in-sample, but lacks out-of-sample performance. In terms of the full dataset approach, the performance is not as strong as the OECD or RC approaches, but delivers better results overall in comparison to non-OECD and NRC. One may be crude, and conclude that the approach delivers a weighted performance of OECD/RC and non-OECD/NRC. Interestingly, even though the discriminatory power of the GIIIPS approach was very high, calibrated PDs evidently do not perform as well in absolute terms. In particular the results from 2012 sample are notable, in which PD estimates for Greece, Cyprus and Spain lie outside two rating notches. Since the GIIIPS sample only contains eight sovereigns, this affects the results in table 15 substantially.

	Model Sample		**Validation Sample**		**2012 Sample**	
	Within One	Within Two	Within One	Within Two	Within One	Within Two
OECD	90.96	100	87.50	98.86	86.36	100
Non-OECD	59.20	88.80	38.36	58.90	40.91	68.18
RC	92.50	100	90.91	97.73	87.50	93.75
NRC	69.54	82.76	47.01	64.10	46.67	70
GIIIPS	59.21	81.58			50	62.50
Full Dataset	71.20	86.96	56.52	77.02	51.85	66.67

Table 15: Validation results: Verification of PD bucket sample within one or two rating class deviations (%)

In terms of the Spiegelhalter test, all samples except the GIIIPS model sample pass the test when applying a significance level of 5 %. This is due to the inclusion of the observation corresponding to the default of Greece in the GIIPS model sample, which receives a significantly lower calibrated PD than the corresponding RIPD. Again, the inclusion of the observation is justified, since removing it from the sample decreases accuracy when computing PDs for the 2012 sample.

6 Concluding Remarks

This thesis aimed to estimate economic capital (EC) for sovereign exposures within a found internal ratings-based (FIRB) model framework, by estimating a sovereign's probability of default (PD) based on both quantitative and qualitative factors according to the shadow rating approach (SRA). Standard and Poor's (S&P's) current sovereign rating methodology as well as Bloomberg's SRSK model represented benchmark methodologies. The developed PD model was tested on four different modular approaches consisting of six model samples and five validation samples. The samples were divided according to World Bank's country classification, IMF's COFER classification, GIIIPS and closely related sovereigns, as well as the entire dataset at hand. The model was evaluated by employing validation procedures, namely the SAR measuring discriminatory power, verification through PD buckets, and the statistical Spiegelhalter test. The results indicate that OECD and RC datasets perform very well overall, while non-OECD and NRC perform adequately within the model sample, but not as well for the validation sample. The segments may benefit from a more fine granular division into sub-segments, mindful of obtaining a substantially large set of observations. Furthermore, GIIIPS in general provide reliable results, but special attention should be given to the performance of PD estimates in absolute terms. The full dataset performs adequately overall, but grouping sub-segments such as OECD sovereigns provide more accurate estimates. In the main, economic, monetary and political factors as well as banking sector health are found to best replicate external ratings.

It may be possible to obtain performance enhancements by classifying sovereigns according to the rating scale, i.e. sovereigns currently rated 'AAA'-'A' should be grouped together. However, in such approaches there is no economical relevance other than the creditworthiness opinionated by the CRA.

In regards to alternative sovereign credit risk practices, the reason for calibrating to external ratings specifically was chosen primarily due to its suitability for EC calculations, and the aim of mapping to an external default frequency due to the scarcity of sovereign defaults historically. Other sovereign credit risk practices, such as the widespread CDS-based models might represent more accurate short-term PD estimates. However, implied PD estimates from CDS spreads are severely more volatile than credit ratings, which would imply large fluctuations in capital requirements from year to year. For the purpose of calculating EC, banks do not find volatile PD estimates a pleasant feature.

In terms of model uncertainty, the calibrated RIPD scale and the methodology behind it represents one such area. LDP PDs in general hold a certain level of uncertainty, since the objective is to estimate PDs including a degree of conservatism. Similarly, one may contest the degree of conservatism of the EB estimator for sovereign PD estimation. In addition, obtaining RIPD estimates for (+) and (-) ratings using interpolation involves a certain degree of uncertainity.

Nonetheless, a major issue discussed in section 2.1.3, is the fact that empirical studies have provided conflicting conclusions on sovereign default determinants. There simply does not exist a set of common and consistent determinants for default, but rather a complex range of conditional circumstances that may cause default. Hence, the thesis' conclusions regarding the indicators for each subsegment may not be seen as general facts in terms of determinants of sovereign default. Nevertheless, the thesis' results clarifies which indicators are most important to take into account when replicating sovereign external ratings.

Further studies towards sharpening the tools for SRA rating development may closer investigate panel type correlations models as a substitute for the multiple linear OLS regression in order to incorporate correlations with respect to time of observations. In addition, alternative measures for factor influence, and manual adjustments may be desirable to develop. PIT adjustments to external TTC ratings could also be of interest to explore further. Nevertheless, it is the author's hope that the reader has gained insight into the topic of sovereign SRA model development, and that further attention will be given to studying the field of sovereign credit risk.

References

Altman, E. I. and Rijken, H. A. (2006). A point-in-time perspective on through-the-cycle ratings. *Financial Analysts Journal.*. 62, 54–70. URL http://www.jstor.org/stable/4480730.

Ang, A. and Longstaff, F. A. (2011). Systematic sovereign credit risk: Lessons from the U.S. and Europe. Working Paper Series 16982. National Bureau of Economic Research. URL http://www.nber.org/papers/w16982.

Arezki, R., Candelon, B., and Sy, A. N. (2010). Bad news spreads. *Finance and Development*. 47, 36–37. URL http://www.imf.org/external/pubs/ft/fandd/2010/12/pdf/arezki.pdf.

Bai, J. and Wei, S.-J. (2012). When is there a strong transfer risk from the sovereigns to the corporates? Property rights gaps and CDS spreads. Staff Report 579. Federal Reserve Bank of New York. URL http://www.newyorkfed.org/research/staff_reports/sr579.pdf.

Bandiera, L., Cuaresma, J. C., and Vincelette, G. A. (2010). Unpleasant surprises: Sovereign default determinants and prospects. World Bank Policy Research Working Paper 5401. The World Bank. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1660302.

Basel Committee on Banking Supervision (2005). An explanatory note on the Basel II IRB risk weight functions. Technical report. Bank for International Settlements. URL http://www.bis.org/bcbs/irbriskweight.pdf.

Basel Committee on Banking Supervision (2006). International convergence of capital measurement and capital standards. BIS papers. Bank for International Settlements. URL http://www.bis.org/publ/bcbs128.pdf.

Benjamin, N., Cathcart, A., and Ryan, K. (2006). Low default portfolios: A proposal for conservative estimation of default probabilities. Technical report. Financial Services Authority. URL http://www.fsa.gov.uk/pubs/international/default_probabilities.pdf.

Benzin, A., Trueck, S., and Rachev, S. T. (2003). Approaches to credit risk in the new Basel Capital Accord. Technical report. Universität Karlsruhe. URL http://statistik.econ.kit.edu/download/doc_secure1/benzin_trueck.pdf.

Bomfim, A. N. (2005). *Understanding Credit Derivatives and Related Instruments*. Academic Press Advanced Finance.

Brandorf, C. and Holmberg, J. (2010). *Determinants of Sovereign Credit Default Swap Spreads for PIIGS - A Macroeconomic Approach*. Master's thesis. Lund University, Department of Economics. URL http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=1608010&fileOId=1608014.

Clifford, T., Marianski, A., and Sebestyen, K. (2013). Low default portfolio (LDP) modelling. Technical report. Deloitte. URL http://www.business-school.ed.ac.uk/waf/crc_archive/2013/23.pdf.

Duyvesteyn, J. and Martens, M. (2012). Forecasting sovereign default risk with Merton's model. Working paper series. Robeco Asset Management. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1839470.

Dzidzevičiūtė, L. (2012). Estimation of default probability for low default portfolios. *Ekonomika*. 91, 132–156. URL http://www.leidykla.eu/fileadmin/Ekonomika/91_1/132-156.pdf.

Elizalde, A. and Repullo, R. (2006). Economic and regulatory capital in banking: What is the difference? Technical report. CEMFI. URL http://www.abelelizalde.com/pdf/economic%20regulatory%20actual.pdf.

Erlenmaier, U. (2011). *The Basel II Risk Parameters*. chapter 4. Springer Berlin Heidelberg. URL http://www.hkfrm.org/resources/Risk_Parameters.pdf.

Flannery, M. J., Houston, J. F., and Partnoy, F. (2010). Credit default swap spreads as viable substitutes for credit ratings. *University of Pennsylvania Law Review*. 158, 2085–2123. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1666350.

Gray, D. F. (2009). Modeling financial crises and sovereign risks. *The Annual Review of Financial Economics*. 1, 117–144.

Gray, D. F. and Jobst, A. A. (2011). Modelling systemic financial sector and sovereign risk. Sveriges Riksbank Economic Review 2. Sveriges Riksbank. URL http://www.tcmb.gov.tr/yeni/konferans/Gray_2.pdf.

Gray, D. F., Merton, R. C., and Bodie, Z. (2007). Contingent claims approach to measuring and managing sovereign credit risk. *Journal of Investment Management*. 5, 5–28. URL http://www.people.hbs.edu/rmerton/ContingentClaimsApproachJOIM.pdf.

Guglielmo, C. (2013). Best practices for credit risk assessment of low default portfolios. Technical report. Standard and Poor's Capital IQ. URL http://www.bvdinfo.com/BvD/media/

Case-studies/Frankfurt/BvD%20Corporate%20Risk%20Day/S-P_Best-Practices-for-Credit-Risk-Assessment-of-low-default-Portfolios.pdf.

Hannoun, H. (2011). Sovereign risk in bank regulation and supervision: Where do we stand? Bis papers. Bank for International Settlements. URL https://www.bis.org/publ/bppdf/bispap72y.pdf.

Huberdeau, P. (2013). How to provide insurance against sovereign risk? *Insurance Economics*. URL https://www.genevaassociation.org/media/178587/GA2013-IE67-Huberdeau.pdf.

International Monetary Fund (2013). Currency Composition of Official Foreign Exchange Reserves (COFER). URL http://www.imf.org/External/np/sta/cofer/eng/index.htm. Last visited 2013-12-19.

Izzi, L., Oricchio, G., and Vitale, L. (2011). *Basel III Credit Rating Systems: An Applied Guide to Quantitative and Qualitative Models*. Palgrave Macmillan.

Jedidi, O. (2013). Predicting sovereign debt crises: A panel data approach using composite indices. Technical report. CREM-CNRS, University of Rennes. URL http://gdre2013.conference.univ-poitiers.fr/sites/gdre2013.conference.univ-poitiers.fr/IMG/pdf/Jedidi.pdf.

Jo, I. (2008). Shadow bond approach through large corp scorecard (LCS). Presentation. Standard Chartered.

Joffe, M. D. (2012). Rating government bonds: Can we raise our grade? *Econ Journal Watch*. 9, 350–365. URL http://ideas.repec.org/a/ejw/journl/v9y2012i3p350-365.html.

Joy, M. (2012). Sovereign default and macroeconomic tipping points. Research Technical Papers 10/RT/12. Central Bank of Ireland. URL http://www.centralbank.ie/publications/documents/10rt12.pdf.

Kelinman, J. C. (1973). Proportions with extraneous variance: Single and independent sample. *Journal of the American Statistical Association*. 68, 46–54.

Kindleberger, C. (1978). *Manias, Panics, and Crashes: A History of Financial Crisis*. Basic Books.

Kraay, A. and Nehru, V. (2006). When is external debt sustainable? *World Bank Economic Review*. 20, 341–365.

Kuhn, G. (2012). Introduction to credit risk rating models. Lecture: Application of Probability Theory in Risk Management, Technische Universität München WS 2012/13.

Longstaff, F. A., Pan, J., Pedersen, L. H., and Singleton, K. J. (2011). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*. 3, 75–103. URL http://pages.stern.nyu.edu/~lpederse/papers/SovereignLPPS.pdf.

Manasse, P., Roubini, N., and Schimmelpfennig, A. (2003). Predicting sovereign debt crises. IMF Working Paper WP/03/221. International Monetary Fund. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=880911.

Mausser, H. and Rosen, D. (2007). Chapter 16 economic credit capital allocation and risk contributions. *Handbooks in Operations Research and Management Science*. 15, 681–726.

Menéndez, L. (2012). The spread of the european sovereign debt crisis. Technical report. The University of Iowa Center for International Finance and Development. URL http://ebook.law.uiowa.edu/ebook/sites/default/files/Spread%20of%20the%20European%20Sovereign%20Debt%20Crisis.pdf.

Munniksma, K. (2006). Credit risk measurement under Basel II. BMI paper. Vrije Universiteit.

O'Kane, D. (2012). The link between Eurozone sovereign debt and CDS prices. Working paper. EDHEC-Risk Institute. URL http://docs.edhec-risk.com/mrk/000000/Press/EDHEC-Risk_Working_Paper_Link_Sovereign_Debt_CDS.pdf.

Orth, W. (2011). Default probability estimation in small samples - with an application to sovereign bonds. Technical report. University of Cologne. URL http://mpra.ub.uni-muenchen.de/33778/1/MPRA_paper_33778.pdf.

Pan, J. and Singleton, K. J. (2008). Default and recovery implicit in the term structure of sovereign CDS spreads. *The Journal of Finance*. 63. URL http://www.mit.edu/~junpan/Pan_Singleton.pdf.

Pluto, K. and Tasche, D. (2005). Estimating probabilities of default for low default portfolios. Technical report. Deutsche Bundesbank. URL http://arxiv.org/pdf/cond-mat/0411699.pdf.

Remolona, E. M., Scatigna, M., and Wu, E. (2007). Interpreting sovereign spreads. *BIS Quarterly Review*. URL http://www.bis.org/repofficepubl/arpresearch_fs_200703.02.pdf.

Rosenberg, M. and Singenellore, R. (2013). Quantitative approaches to modeling sovereign risk. Technical report. Global Association of Risk Professionals.

Santis, R. A. D. (2012). The Euro area sovereign debt crisis: Safe haven, credit rating agencies and the spread of fever from Greece, Ireland and Portugal. Working Paper Series 1419. European Central Bank. URL http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1419.pdf.

Standard and Poor's (2012). Standard and Poor's ratings definitions. Technical report. Standard and Poor's Ratings Services. URL http://www.standardandpoors.com/spf/general/RatingsDirect_Commentary_979212_06_22_2012_12_42_54.pdf.

Standard and Poor's (2013a). Common characteristics of rated sovereigns prior to default. *CariCRIS RatingMonitor*. URL http://www.cfsc.com.bb/photos/Rating%20Monitor%20July%202013%20(2)1.pdf.

Standard and Poor's (2013b). Default study: Sovereign defaults and rating transition data, 2012 update. Technical report. Standard and Poor's Ratings Services. URL http://www.standardandpoors.com/spf/upload/Ratings_US/Sovereign_DefaultandRatingTransition_Data2012Update.pdf.

Standard and Poor's (2013c). Sovereign government rating methodology and assumptions. RatingsDirect. Standard and Poor's.

Text Medic (2013a). Modeling sovereign risk. part 1: Emerging markets. URL http://textmedic.ca/modeling-sovereign-risk-part-1-emerging-markets/. Last visited 2013-09-30.

Text Medic (2013b). Modeling sovereign risk. part 2: Quantification. URL http://textmedic.ca/modeling-sovereign-risk-part-2-quantification/. Last visited 2013-09-30.

Tomz, M. and Wright, M. L. J. (2007). Do countries default in bad times? *Journal of the European Economic Association*. 5, 352–360. URL http://www.stanford.edu/~tomz/pubs/TW2007.pdf.

Tomz, M. and Wright, M. L. J. (2012). Empirical research on sovereign debt and default. *Annual Reviews of Economics*. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2182179.

Trueck, S. and Rachev, S. T. (2009). *Rating Based Modeling of Credit Risk: Theory and Application of Migration Matrices*. 1st edition. Academic Press Advanced Finance.

van der Burgt, M. J. (2007). Calibrating low-default portfolios, using the cumulative accuracy profile. *Journal of Risk Model Validation*. 2, 17–33. URL http://www.risk.net/digital_assets/5043/jrm_v1n4a2.pdf.

World Bank (2013). How we classify countries. URL http://data.worldbank.org/about/country-classifications. Last visited 2013-12-19.

Zaalberg, R. (2013). *Quantifying uncertainty in credit rating model development*. Master's thesis. University of Twente. URL http://essay.utwente.nl/63919/1/Thesis_RebeccaZaalberg.pdf.

Appendices

A S&P's Long-Term Issuer Credit Ratings

Long-Term Issuer Credit Ratings*

Category	Definition
AAA	An obligor rated 'AAA' has extremely strong capacity to meet its financial commitments. 'AAA' is the highest issuer credit rating assigned by Standard & Poor's.
AA	An obligor rated 'AA' has very strong capacity to meet its financial commitments. It differs from the highest-rated obligors only to a small degree.
A	An obligor rated 'A' has strong capacity to meet its financial commitments but is somewhat more susceptible to the adverse effects of changes in circumstances and economic conditions than obligors in higher-rated categories.
BBB	An obligor rated 'BBB' has adequate capacity to meet its financial commitments. However, adverse economic conditions or changing circumstances are more likely to lead to a weakened capacity of the obligor to meet its financial commitments.
BB; B; CCC; and CC	Obligors rated 'BB', 'B', 'CCC', and 'CC' are regarded as having significant speculative characteristics. 'BB' indicates the least degree of speculation and 'CC' the highest. While such obligors will likely have some quality and protective characteristics, these may be outweighed by large uncertainties or major exposures to adverse conditions.
BB	An obligor rated 'BB' is less vulnerable in the near term than other lower-rated obligors. However, it faces major ongoing uncertainties and exposure to adverse business, financial, or economic conditions which could lead to the obligor's 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. Adverse business, financial, or economic conditions will likely impair the obligor's capacity or willingness to meet its financial commitments.
CCC	An obligor rated 'CCC' is currently vulnerable, and is dependent upon favorable business, financial, and economic conditions to meet its financial commitments.
CC	An obligor rated 'CC' is currently highly vulnerable.
R	An obligor rated 'R' is under regulatory supervision owing to its financial condition. During the pendency of the regulatory supervision the regulators may have the power to favor one class of obligations over others or pay some obligations and not others. Please see Standard & Poor's issue credit ratings for a more detailed description of the effects of regulatory supervision on specific issues or classes of obligations.
SD and D	An obligor rated 'SD' (selective default) or 'D' is in payment default on one or more of its financial obligations (rated or unrated) unless Standard & Poor's believes that such payments will be made within five business days, irrespective of any grace period. The 'D' rating also will be used upon the filing of a bankruptcy petition or the taking of similar action if payments on a financial obligation are jeopardized. A 'D' rating is assigned when Standard & Poor's 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' rating is assigned when Standard & Poor's believes that the obligor has selectively defaulted on a specific issue or class of obligations, but it will continue to meet its payment obligations on other issues or classes of obligations in a timely manner. A selective default includes the completion of a distressed exchange offer, whereby one or more financial obligation is either repurchased for an amount of cash or replaced by other instruments having a total value that is less than par.
NR	An issuer designated 'NR' is not rated.

*The ratings from 'AA' to 'CCC' may be modified by the addition of a plus (+) or minus (-) sign to show relative standing within the major rating categories.

Figure 13: S&P's long-term issuer credit ratings (Standard and Poor's 2012, p. 7).

B Modular Sample Approaches

Modular Approach 1

High Income OECD		Non-OECD	
Development Sample	**Validation Sample**	**Development Sample**	**Validation Sample**
Austria	Australia	Albania	Brazil
Belgium	Canada	Belarus	Egypt
Czech Republic	Chile	Bosnia & Herzegovina	India
Denmark	Israel	Bulgaria	Kazakhstan
Estonia	Japan	Croatia	Peru
Finland	New Zealand	Hungary	Russia
France	South Korea	Latvia	Uruguay
Germany	USA	Lithuania	
Luxembourg		Macedonia	
Netherlands		Malta	
Norway		Montenegro	
Poland		Romania	
Slovakia		Serbia	
Sweden		Turkey	
Switzerland		Ukraine	
United Kingdom			
166 Observations	88 Observations	125 Observations	73 Observations

Table 16: Approach 1: World Bank's country classification (World Bank 2013)

Modular Approach 2

Reserve Currency		Non-Reserve Currency	
Development Sample	**Validation Sample**	**Development Sample**	**Validation Sample**
Austria	Australia	Albania	Brazil
Belgium	Canada	Belarus	Chile
Estonia	Japan	Bosnia & Herzegovina	Egypt
Finland	USA	Bulgaria	India
France		Croatia	Israel
Germany		Czech Republic	Kazakhstan
Luxembourg		Denmark	New Zealand
Malta		Hungary	Peru
Netherlands		Latvia	Russia
Slovakia		Lithuania	South Korea
Switzerland		Macedonia	Uruguay
United Kingdom		Montenegro	
		Norway	
		Poland	
		Romania	
		Serbia	
		Sweden	
		Turkey	
		Ukraine	
120 Observations	44 Observations	174 Observations	117 Observations

Table 17: Approach 2: IMF's COFER classification (International Monetary Fund 2013)

Modular Approach 3	
Development Sample	
Cyprus	Italy
Greece	Portugal
Iceland	Slovenia
Ireland	Spain
84 Observations	

Table 18: Approach 3: GIIIPS, Cyprus and Slovenia

Modular Approach 4		
Development Sample		**Validation Sample**
Albania	Lithuania	Australia
Austria	Luxembourg	Brazil
Belarus	Macedonia	Canada
Belgium	Malta	Chile
Bosnia & Herzegovina	Montenegro	Egypt
Bulgaria	Netherlands	India
Croatia	Norway	Israel
Cyprus	Poland	Japan
Czech Republic	Portugal	Kazakhstan
Denmark	Romania	New Zealand
Estonia	Serbia	Peru
Finland	Slovakia	Russia
France	Slovenia	South Korea
Germany	Spain	USA
Greece	Sweden	
Hungary	Switzerland	
Iceland	Turkey	
Ireland	Ukraine	
Italy	United Kingdom	
Latvia		
368 Observations		161 Observations

Table 19: Approach 4: Combination of approach 1, 2 and 3

C Derivation of OLS Estimator and Covariance Matrix

Let N be the number of observations, k the number of explanatory factors, and $\hat{\beta}$ the OLS estimator minimizing the sum of all ϵ^2. The linear regression model may then be denoted in the following matrix notation:

$$Y = X\beta + \epsilon, \tag{C.1}$$

where Y, β and ϵ are vectors of size $N \times 1$, $k \times 1$, $N \times 1$ respectively, while X is a $N \times k$ matrix. The OLS estimator minimizes $\epsilon^T \epsilon$:

$$\epsilon^T \epsilon = (Y - X\beta)^T (Y - X\beta) = Y^T Y - 2\beta^T X^T Y + \beta^T X^T X\beta. \tag{C.2}$$

From C.2, the derivative of $\epsilon^T \epsilon$ with respect to β^T is then given by:

$$\frac{\partial \epsilon^T \epsilon}{\partial \beta^T} = 2X^T Y + X^T X\beta. \tag{C.3}$$

Setting $\partial \epsilon^T \epsilon / \partial \beta^T = 0$, provides the OLS estimator $\hat{\beta}$:

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T Y \tag{C.4}$$

(Zaalberg 2013, p. 74).

Assuming $Var(\epsilon) = \sigma^2 I$, i.e. ϵ are uncorrelated and homoskedastic, the OLS estimator is unbiased:

$$E(\hat{\beta}) = E\left(\left(X^T X\right)^{-1} X^T Y\right) = \left(X^T X\right)^{-1} X^T E(Y) = \left(X^T X\right)^{-1} X^T X\beta = \beta \tag{C.5}$$

It is possible to retrieve the covariance matrix $Var(\hat{\beta})$ through the following equations:

$$\hat{\beta} - \beta = \left(X^T X\right)^{-1} X^T (X\beta + \epsilon) - \beta = \left(X^T X\right)^{-1} X^T \epsilon \tag{C.6}$$

$$\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)^T = \left(X^T X\right)^{-1} X^T \epsilon\epsilon^T X \left(X^T X\right)^{-1} \tag{C.7}$$

The covariance matrix $Var(\hat{\beta})$ can now be calculated as follows:

$$\begin{aligned} Var(\hat{\beta}) &= \left(X^T X\right)^{-1} X^T E\left(\epsilon\epsilon^T\right) X \left(X^T X\right)^{-1} = \\ &= \sigma^2 \left(X^T X\right)^{-1} X^T X \left(X^T X\right)^{-1} = \sigma^2 \left(X^T X\right)^{-1} \end{aligned} \tag{C.8}$$

(Zaalberg 2013, p. 26).