Determinants of Sovereign Defaults

Yaseen Ghulam[a,b] and Julian Derber

a. University of Portsmouth, Portsmouth Business School,

Economics and Finance Subject Group

Richmond Building, Portland Street, PO1 3DE, UK

Email: **yaseen.ghulam@port.ac.uk**

Phone: +442392844127

b. Al Yamamah University 7010 King Fahd Road,

Al Qirawan, Riyadh 13541, Saudi Arabia

Determinants of Sovereign Defaults

Highlights

- Political uncertainty/worsening international monetary conditions promote default

- Export (import) growth reduces (increases) the probability of default

- Higher debt/GDP ratio, and inflation is linked to higher probability of default

- A previous banking crisis is related to higher chances of sovereign defaults

- Higher US treasury rates would initiate sovereign defaults and hinders to come out

Determinants of Sovereign Defaults

Abstract

We study major sovereign defaults from 1970 to 2010 using an advanced duration analysis method. Descriptive analysis of the data indicates a cyclical nature of these defaults over a longer period. Regression results highlight the importance of the international monetary conditions as the volatilities of US treasury bills rates and USD-denominated LIBOR exert significant impacts on defaults. Political uncertainty increases the probability of default. Export (import) growth reduces (increases) the probability of default. Similarly, a 1% increase in inflation would increase the probability of defaults by 7%. Higher debt/GDP ratio is also linked to higher probability of default. A 1% increase in external debt would lead to a five to 7% increase in the probability of default. Higher GDP per capita reduces the probability of default. A previous banking crisis is linked to higher chances of sovereign defaults. Further analysis of entry into (out of) sovereign defaults indicates that higher US treasury rates would initiate sovereign defaults and would make it difficult for countries to come out of default. The same is true for central government debt/GDP, higher current account deficit and exchange rate volatility.

1. Introduction

Reinhart & Rogoff (2009) outlined how throughout history "this time is different" became a credo for private and public investors extensively engaged in sovereign lending, just to eventually face "yet another crisis." In this regard, the recent European sovereign debt crisis is no exception. It started in 2007 with the U.S. subprime financial crisis, which came as a shock to most economies, thus cooling down growth internationally. This forced the newly elected Greek government to announce that the 2009 budget deficit would largely exceed expectations. In the subsequent (2010) Greek sovereign debt crisis, the government struggled not only with high public debt and low competitiveness compared to the rest of the euro zone, but also from a credibility deficit view point too (Tsoulakis, 2012). Before this incident, economists and capital markets seemed to have shared the belief that sovereign default would be an issue only experienced by developing countries. A defaulting developed country, and a member of the European Monetary Union (EMU), was not perceived as a possibility by the market. This allowed Greece (and other EMU members) to borrow at exceptionally low spreads over German bunds, creating a strong incentive to over borrow and accumulate remarkably high amounts of sovereign debt (Tsoulakis, 2012).

Compared to late 1980s and early 1990s when financial markets cooled down and sovereign defaults were few leading to less academic research on foreign borrowing crisis, the number of studies since 2014 on sovereign debt crisis have been regularly forthcoming in particular of European countries. Some of these studies include Ucler & Kirmizioglu (2015), Tamborini (2015), Broto & Perez-Quiros (2015), Popov & Van Horen (2015), Smeets (2016), Moisescu & Giurescu (2016), Stamatopoulos et al. (2016), Gómez-Puig & Sosvilla-Rivero (2016), Afonso & Silva (2017), Cencini (2017), Reusens & Croux (2017), and Ehrmann & Fratzscher (2017). A significant number of these studies have used time series models to evaluate credit default swaps (CDS) and other similar instruments to assess the riskiness of a country due to the ever rising burden of sovereign debt. This paper aims to develop an empirical model that helps to identify risk factors influencing the probability of sovereign default. The study builds a deep understanding of related economic theory and the findings of other

researchers that tried to identify "early warning signs" (such as Manasse & Roubini, 2009). We use a panel dataset distributed over 70 countries (developing as well as developed) from 1970 to 2010.

The inclusion of a number of important variables in our study to model external debt defaults comprising political, regional, macroeconomic, financial and external trade and debt is useful and significantly advanced compared to existing empirical literature. Compared to a number of other studies mentioned above, our study uses more indicators to approximate in particular political conditions of a country. Our study's use of advanced statistical models such as duration (survival time) analysis is a significant addition and advancement to existing literature and a more up-to-date dataset comprising of both developed and developing countries makes this analysis a significant piece of research and a contribution to the ever developing empirical literature. The paper develops an understanding that the use of time-to-event methodology can have a significant advantage over other techniques such as simple logit or probit used so far for the analysis of sovereign debt analysis.

2. Modelling the Determinates of Sovereign Defaults – A Review of Literature

An interesting question is how creditors measure sovereign risk. While the calculation of country risk is more an art than a scientific econometric analysis (especially because of only qualitatively assessable political factors), sovereign risk as a part of the country risk is rated by well-known international agencies. These agencies only give hints on which variables they include in their analysis, but do not disclose details of their procedures (Fight, 2004). Fight (2004) asserts that bank managers often faithfully refer to these ratings. This is consistent with Larraín et al. (1997), who present econometric evidence for the period 1987 to 1996, finding a highly significant announcement effect on financial markets for negative changes in credit ratings. They also state that sovereign risk ratings can be primarily explained with publicly available macroeconomic data (see also Cantor & Packer, 1996; Manasse & Roubini, 2009; Eichler & Maltritz, 2012) and dollar-bond-spreads.

In one of the most-cited recent empirical studies investigating determinants of sovereign default, Manasse and Roubini (2009, p. 3) complain that many "policymakers and analysts continue to use simple rules of thumb to

judge risks and to assess fiscal sustainability (IMF, 2003), as well as the soundness of macroeconomic policies." According to Manasse and Roubine (2009), economists and practitioners struggled to properly understand macroeconomic and structural weaknesses leading to sovereign defaults. They also notice little comparative empirical work on the sovereign debt crises of the 1990s and early 2000s. The recent European sovereign debt crisis, however, seems to have drawn new attention to the issue. Villemont (2011) explains that even though there is a large body of literature trying to empirically identify determinates of sovereign defaults, most results can hardly be interpreted as representing causalities. In particular, endogeneity issues limit these efforts to identify risk factors associated with sovereign defaults. Popular approaches focus on central economic indicators, risk ratios or market estimates of sovereign risk (Das et al., 2012). The most popular explanatory variables are external or public debt to GDP as well as the public debt or debt service to government revenues ratios.

Most studies use probit or logit regressions or signal models (Manasse & Roubini, 2009) to model sovereign defaults. Referring to the existing literature in this regard, Manasse and Roubini (2009) suggest that an empirical model attempting to predict sovereign default should include measures of solvency and liquidity as well as political, institutional or other variables that could represent a country's willingness to pay. Many authors also emphasize the inclusion of variables on the macroeconomic state and volatility (Hilscher & Nosbusch, 2010; Villemont, 2011; Das et al., 2012), because they capture not only the ability, but also the willingness to pay (Manasse & Roubini, 2009). Das et al. (2012) describe research by Sturzenegger and Zettelmeyer (2006, p. 6) which analyses default and restructuring clusters during the last 200 years and finds that debt crises were preceded by either a worsening in terms of trade, a recession in capital providing countries, a surge in the international costs of capital, or a crisis in a contagion-causing country.

With a range of 50 predictor variables, Manasse and Roubini (2009) use a classification and regression tree (CART) methodology on a panel dataset including 47 market economies from 1970 to 2002. They claim to reach excellent predictive capacities in-sample, while the out-of-sample prediction includes less correct

predictions, but also less false alarms than the "Early Warning Signal (EWS)" literature. Moreover, they could identify ten variables as sufficient for the classification or prediction of a sovereign debt crisis. These include total external debt/GDP ratio; short-term debt reserves ratio; real GDP growth; public external debt/fiscal revenue ratio; CPI inflation; number of years to the next presidential election; U.S. treasury bills rate; external financial requirements (current account balance plus short-term debt as a ratio of foreign reserves); exchange rate overvaluation; and exchange rate volatility.

Manasse and Roubini (2009) specify characteristics of a relatively "safe" country: total external debt over GDP ratios below 49.7%; short-term debt over reserves under 130%; public external debt as% of fiscal revenue not being higher than 214; and a maximum overvaluation of the exchange rate of 48%. They also outline three risk types: 1) unsustainability risk can be identified by external debt to GDP ratios over 49.7% in combination with monetary or fiscal imbalances and large needs of refinancing; 2) liquidity risk types show moderate debt levels, with short-term debt over reserves exceeding 130% and political uncertainty (no upcoming elections) as well as tight international capital markets; 3) macro-exchange rate risk types are significant in terms of low growth (-5.5%) combined with relatively fixed exchange rates.

Our study contributes to the existing theoretical and empirical literature reviewed above. We model the sovereign default by using a number of variables discussed above and employing an advanced econometric method: survival (duration) analysis framework. The empirical analysis of our study is rich and would provide a good deal of information for creditors as well as debtor nations to manage future sovereign defaults and related costs. In the following, we discuss the duration (survival) analysis methodology for modelling sovereign defaults alongside justification for the chosen method.

3. Methodology

As mentioned before, the majority of the literature in this field has used logit or probit regression to identify risk factors associated with sovereign default. With the use of CART, Manasse and Roubini (2009) applied a

new methodology. Reinhart et al., (2003 as cited in Das et al., 2012), call past defaults a main predictor of defaults and debt restructurings. The large number and strong influence of time-dependent variables used in the existing empirical literature to explain external sovereign defaults in combination with the strong role of past events may make survival (duration) analysis an interesting option and a more appropriate method to define determinants or warning signs predictor of sovereign default.

Although, to the author's knowledge, there has been very few applications of duration (survival) analysis to model sovereign default (one of such is Cruces and Trebesch, 2013), it is repeatedly applied to measure the credit risk of bank loans or other financial instruments (see, for example, Li, 1999). For credit scoring or the prediction of default on credit cards, the use of statistical methods such as logistic regression has developed into a standard (Thomas et al., 2002). Survival analysis can improve this process by providing "forecasts as a function of time" (Banasik et al., 1999). It "allows us to model not just if a borrower will default, but when" (Bellotti & Crook, 2009, p. 1). Bellotti & Crook (2009) formulate another advantage of survival-analysis: compared to standard regression or logistic regression models, it allows for very natural incorporation of time-varying covariates such as macroeconomic time-series data.

Most data in social sciences can be considered continuous time data, as the behavioral processes that the data represent occur continuously. For example a person could lose job any time despite signing a long term contract. Nevertheless, it is usually grouped into time-intervals, forming discrete survival time data. Where the ratio of grouping-interval length to typical spell length is small, Jenkins (2005) recommends the use of continuous time specifications, as in a complementary log-log (cloglog) model. Although, the "cloglog" model seems attractive because of its approximation to the continuous-time piecewise constant exponential (PCE) model, the logit model is more commonly used in social sciences (Steele, 2011). It is also called "proportional odds model" (Jenkins, 2005) or "discrete-time proportional hazard model" (Steele, 2011). Sueyoshi (1995) demonstrated that a logistic model with interval-specific intercepts can be representative for an underlying continuous-time process, although by default it considers the underlying process to be a discrete-time one. Steele (2011) argues

that in practice both models produce very similar predictions as long as probabilities are small. Furthermore, some studies on sovereign default like Reinhart and Rogoff (2009) argue a period of up to three years around the actual default date could be considered sovereign debt crisis. When the actual date of an event within an interval is less important, "we can happily treat the process as truly discrete" (Steele, 2012, p. 3).

This paper will, therefore, focus on using a logit model on country-period panel data (duration data) in order to analyze factors influencing the probability that a sovereign debtor defaults on its external debt obligations. Another point to consider is that most of the applications of survival analysis in financial contexts, however, only need to consider one default per "patient" (country in our case), because the lifetime of a debt contract usually ends with repayment or default. Throughout history, most sovereigns – especially now developed Western nations – defaulted multiple times (Reinhart & Rogoff, 2009). The occurrence of multiple spells per subject is usually discussed under the label "multiple failure-time" (Cleves, 2009). It is, however, important to distinguish between the analyses of "competing risks" where subjects are at risk of different exclusive events and "recurrent events."

With regards to employment transitions, Steele (2011, p. 2) describes recurrent events data as "a type of two-level hierarchical structure with episodes of being at risk of an event nested within individuals, and individuals (subject) may themselves be nested within geographical areas or institutions (e.g. employers)…" This definition seems applicable to the data used in this study too: each country is continuously at risk of a transition either from default and "non-default" periods or vice versa. Cleves (2009) points out that this clustering of failure times within subjects is violating the assumption made in traditional survival analysis that failure times would be independent. The easiest way to avoid this violation is examining the time to first event (default) only. This, however, often leaves a large amount of data unused, wasting information that may have been relevant (Cleves, 2009). Nevertheless, because of its simplicity, this study will initially follow this approach, excluding all country-related observations after a country defaults for the first time, using a "proportional odds model."

Analyzing the time to first default will only be beneficial for developing insights and patterns in the underlying data. For the analysis of the whole dataset, Steele (2011) proposes an extension to the simple proportional hazard model into what the author calls a random-effect or two-state model. Cleves (2009) describes two basic model classes used to analyze multiple event data: Frailty models and variance-corrected models. The model outlined by Steele (2011) would fall under Cleve's (2009) definition of a frailty model, where "the association between failure times is explicitly modelled as a random-effect term, called the frailty."

Since individual country heterogeneity would remain unobserved in a single-level model, such as the simple logit model, subject (country)-specific effects will bias the duration dependence of the population hazard (Vauper et al., 1979). The inclusion of random-effects terms or "frailties" makes it possible to interpret coefficients as subject-specific or conditional, instead of population-averaged as is the case for single-level coefficients (Neuhaus et al., 1991). Additionally, it is important to respect the residual standard deviation $\hat{\sigma}_u$ when interpreting the estimates in a random effects model. With larger inter-individual variance, the population-averaged coefficients will increase in magnitude compared to estimates from the single-level model (Snijders & Bosker, 1994, p. 227-8).

The data used in the following empirical analysis is continuous time data with interval-censored spell lengths. It will be analyzed using discrete-time methods. This means the continuous time axis can be split into a number of non-overlapping intervals $[0 = a_0, a_1), (a_1, a_2], (a_2, a_3], \dots, (a_{k-1}, a_k = \infty]$. The interval boundaries $a_0 = 0, a_1, a_2, \dots, a_k$ are then represented by dates on the time axis. Nevertheless, applying the proportional odds model implicitly assumes an intrinsically discrete distribution of time, handling survival time T as a discrete random variable with the probability density function $f(j) \equiv f_j = \Pr(T = j)$ where $j \in \{1, 2, 3, \dots\}$ representing consecutive cycles instead of equally distributed intervals. This assumption is not visible in the dataset, because both sets of survival times would follow the same notation based on positive integers j. However, it influences the interpretation of survivor function and discrete time hazard functions (Jenkins, 2005).

The survivor function $S(j)$ shows the probability of "survival" or, put more generally, of not experiencing an event (default in our case) within j cycles. The whole function is given by:

$$S(j) = \Pr(T \geq j) + \sum_{k=j}^{\infty} f_k.$$

(1)

Correspondingly, the function $h(j)$ is the conditional probability of experiencing an event during cycle j. This is called the discrete time hazard defined as:

$$h(j) = \Pr(T = j | T \geq j) = \frac{f(j)}{S(j-1)}.$$

(2)

The cumulative distribution function of T, also called failure function, is denoted by $F(j)$. The discrete time failure function is:

$$F_j = F(j) = 1 - S(j) = 1 - \prod_{k=1}^{j}(1 - h_k).$$

(3)

Based on these assumptions, the "proportional odds model" can then be written as:

$$\frac{h(j,X)}{1 - h(j,X)} = \left[\frac{h_0(j)}{1 - h_0(j)}\right]\exp(\beta'X)$$

$$\Rightarrow \log it\,[h(j,X)] = \log\left[\frac{h(j,X)}{1 - h(j,X)}\right] = a_j + \beta'X.$$

(4)

This models the discrete time hazard $h(j,X)$ as the relative probability of an event in cycle j based on survival up to this point in time. This probability is influenced by two factors: relative odds that take the same value for the whole population and a scaling-factor that is specific to each individual country. While X is a matrix containing the observations of explanatory variables and β represents the corresponding coefficients, $a_j = $ logit $[h_0\,(j)]$ with $h_0\,(j)$ being the so called "baseline hazard" for $X = 0$. Thus,

$$h(j, X) = \frac{1}{1 + \exp(-a_j - \beta'X)},$$

(5)

This is the "logistic hazard model" (as outlined in Jenkins, 2005), which can be estimated using maximum likelihood methods. In Stata its estimation is possible with the **logit** command.

As Jenkins (2005, p. 44) puts it: "The choice of shape of hazard function in these models is up to the investigator…" One possibility is to assume a baseline hazard, a time-dependent probability for an event that is underlying each individual's specific risk, based on economic theory. One interesting option could be to assume the duration dependent baseline hazard to be represented by $a_j = r \log(j)$. This would be a discrete-time equivalent to the continuous-time Weibull model. It also allows for some flexibility, because the coefficient r would be estimated alongside the intercept and the β-values. Where $r > 0$ ($r < 0$), the baseline hazard will be monotonically increasing (decreasing). A constant baseline hazard will be represented by $r = 0$.

Another option to derive the baseline hazard, especially for discrete-time data is the use of life table estimators that produce estimates for the survival and hazard functions, using an "average" estimate of n_j, the number of individual countries that have left the sample at the midpoint of each interval of time I_j with $j = 1, ..., J$: $I_j : [t_j, t_{j+1})$. The adjustment is needed to account for the fact that continuous data is grouped in discrete intervals; where the underlying process is assumed to be intrinsically discrete, it becomes unnecessary (Jenkins, 2005).

Jenkins (2005) specifies that for each interval I_j, the number of countries at risk of default is N_j (the so-called "risk sample"), the number of observed defaults is d_j, and the number of censored spell endings is m_j. Then $n_j = N_j - \frac{d_j}{2}$ and thus the estimate for the survival function is:

$$\hat{S}(j) = \prod_{k=1}^{j}(1 - \frac{d_k}{n_k}).$$

(6)

With $S(t) = 1 - F(t)$ this gives an estimate for the density function

$$\hat{f}(j) = \frac{\hat{F}(j+1) - \hat{F}(j)}{t_{j+1} - t_j} = \frac{\hat{S}(j) - \hat{S}(j+1)}{t_{j+1} - t_j}$$

(7)

The hazard rate can then be estimated using $\tilde{S}(k) = (\hat{S}(k) + \hat{S}(k+1))/2$ as:

$$\hat{\theta}(j) = \frac{[\hat{f}(j)]}{\tilde{S}(j)}.$$

(8)

The command ltable in STATA offers life table estimates of the survival function. With ltable, the corresponding hazard rates can also be accessed.

In order to estimate the probability of a transition either into default or out of default, a re-specification of the above outlined logit model is necessary. In the following, a two-state model like that developed by Steele (2011) will be developed. The single-state logistic model is looking at the conditional probability that an individual country j experiences an event (default) at time t given that it did not have an event, yet. This holds the expectation that each individual country can only experience one event over the observation period. In order to account for the possibility that more than one event may occur and that there is more than one type of event with corresponding states ($s = 1, 2$), it is necessary to distinguish the time of the study t into the time spent in each episode i.

An event such as the exit from state 1 automatically implies entrance into the other state (2). At this point it would be necessary to start counting the periods spent in this state until another transition back into state 1 or the end of the study. Where entry into the study or exit from it, as well transitions between states are considered

13

an event, each individual country's time in the study can be split into one or more episodes defined as all consecutive periods between two events. The state individual country j inherits at the beginning of a cycle interval (or cycle) t in episode i is denoted by S_{tij}. Implementing these considerations, Steele (2011, p. 9) describes the "multilevel event history model for transitions between the two states" as:

$$\text{logit}(p_{stij}) = \boldsymbol{a}_s' \boldsymbol{z}_{stij} + \boldsymbol{\beta_s}' \boldsymbol{X}_{stij} + u_{sj}, s = 1, 2.$$

(9)

Here, $a_s' z_{stij}$ specifies the baseline logit-hazard (defined by α_j in the single-state logit model). With y_{tij} being the binary indicator for state transition during interval t, the probability that individual country j leaves state s in interval (or cycle) t conditional on no previous state-transition in that episode is:

$$p_{stij} = (y_{tij} = 1 | y_{t'ij} = 0 \text{ for } t' < t, S_{tij} = s).$$

(10)

One feature of this model is that the baseline logit-hazard, the explanatory variables with their coefficients, and the frailty can vary between the two states. This model can easily be estimated in Stata using the xtmelogit or xtlogit command with y_{tij} as the dependent variable and interacting z_{stij} and X_{stij} with dummy variables for each state as explanatory variables. The state-identifying dummies are also included as explanatory variables and replace the intercept.

4. Data

The data used in this study can be grouped in different broad areas: data identifying financial crises, variables describing a country's economic situation, political situation, or the quality of a government's financial and monetary policy. Apart from these explanatory variables, the longitudinal panel dataset contains structural variables (e.g., different time scales and indicators of the beginnings and ends of default episodes) that are necessary to conduct the analyses described in the previous section.

In all the empirical analysis conducted in this study, the dependent variable is based on Reinhart and Rogoff's (2010a) dates for external sovereign debt default. They define external debt crisis as involving "outright default

14

on payment of debt obligations incurred under foreign legal jurisdiction, including non-payment, repudiation, or the restructuring of debt into terms less favorable to the lender than in the original contract" (Reinhart & Rogoff, 2010a, p. 1679). This limits the dataset to the 70 countries analyzed by Reinhart and Rogoff (2010b) between 1800 and 2010. The option to extend this indicator by constructing a default variable based on Standard and Poor's studies and the occurrence of "large" IMF loans, as suggested by Roubini & Setser (2003), Tomz and Wright (2007), or Manasse and Roubini (2009), was considered but discarded in favor of consistency in the dependent variable.

The publicly available sources for (macro-) economic data, for example, the International Monetary Fund (2013a, 2013c), the World Bank (2013a, 2013b), or the OECD (2009), limit the scope of the constructed dataset to yearly observations from 1960 to 2010 for each country. This time span was reduced to 1970 to 2010 due to gaps in the data, resulting in a total of 2870 observations. Possible explanatory variables like ratings of sovereign bonds, and mainly debt-related variables like short-term debt over reserves and volumes of debt services were excluded because of their limited range and large gaps.

The number of observations in the resulting full dataset is large compared to other studies: Manasse et al (2003) and Masse and Roubini (2009) use a panel dataset for 76 countries from 1970 to 2002, but with limited range (1995 to 2002) for about 30 transition economies and further gaps in the data. Hilscher and Nosbusch's (2010) data ranges from 1970 to 2007 for 31 emerging market countries and Villemont's (2012) works with 93 countries between 1970 and 2004. Most other studies that tried to identify risk indicators on sovereign default focus on one or a few chosen explanatory aspects (Das et al., 2012). As will be explained in the following section, the analysis in this study is conducted in consecutive steps. In a first step, only data until the first default will be considered. Therefore, all country-specific periods after the first year in default were deleted. For countries that have been in default in the first observation periods, these observations were excluded, because the aim of the analysis is to identify determinants influencing a transition into default not characteristics of default periods. In subsequent estimations, the whole dataset is used, but observations are distinguished by

whether or not the country is in default. Table 1 lists all 70 countries in the dataset with their default (if any) periods, assigned "id," the years of the observations considered for the first-default-only analysis, whether they defaulted (value equal to "1" otherwise "0") and default episodes in the full dataset. For each explanatory variable used, Table 2 includes a brief description and its origin.

5. Estimation and Explanation

As outlined in the previous section, this study follows a stepwise extension of the used methodology in order to make use of the whole scope of the collected data. In a first step, a logit or "proportional odds model" will be used to estimate the probability of external sovereign default. Life table estimates of the survivor (not going into default) and hazard function (risk of going into default) should help to specify an appropriate baseline logit-hazard. A number of explanatory variables are used subsequently to model first default probability. In a next step, the same explanatory variables will be used with the two-stage model on the full dataset. This model should give an impression on the effect of these variables not only on the occurrence, but also the end of a default period. The main purpose of this is to show how multiple-failure methodology can be used to model external sovereign default compared to single-event methodology.

Figure 1 shows the survivor function plotted from life table estimates on the first-default-only dataset. The shape of the survivor function shows a steadily decreasing, but still high probability (> 90%) of survival for the first ten years in the risk set. Within the next five years, a sudden drop reduces the probability of survival to an average level around 60%. From then on, the survival rates reduce fairly slowly and in a linear pattern to a mean 47% after 40 years. Correspondingly, the estimated hazard rates (Figures 2) are almost constant (around 1.5%) until year 12 of the observation period with the interval from year ten to eleven being an exception with a hazard rate of 4.65%. The hazard rates spike in the following three intervals with average values of ten, nine and 17%. From there onwards, the hazard rates quickly decline to 2.7% and stay almost constant for all intervals between periods 22 and 40, where the observation period ends with a hazard rate of 3.0%.

The spikes in the intervals between the 12th and 15th year may be explained historically. The observation period starts in 1970. This means the observed surge in defaults in the years 12 to 15 is caused by the Latin American sovereign debt crisis of the 1980s[1]. This raises the following question: how does one handle and treat these hazard rates in the estimation of the logistic model? A look at the 95%-confidence intervals of the estimated hazard rates (represented in Figure 2 and Figure 3 by the vertical lines) reveals that assuming an almost linear hazard function would not be too implausible. This could mean that the distribution is actually biased by the actual historical events. Another possibility is that the pattern represents certain cyclical effects regarding external sovereign defaults. It could be interesting to investigate this further using different samples and maybe with multiple-default data. Figure 3 shows the hazard rates when grouped into more or less than 15 years at risk of default. While the function for the intervals up to year 15 resembles the upward reaching half of a parable, the pattern in the remaining periods is almost constant. Thus, we consider two different logit models: One "general model" assuming a baseline hazard of $a_j = r \log(j)$, as described in the previous section, allows for some flexibility as it estimates the actual shape along with other parameters; and a more "specific model" with the observations after year 15 being represented by a dummy variable.

The regression results can be seen in Table 3. For each version of the model, we estimated three specifications to observe the sensitivity of estimates. These include full model (coefficients and odd ratios presented in column 1 & 2), excluding political and regional variables (column 3) and dropping variables representing volatility (column 4). Both versions (general and specified) produce significant estimates for the baseline hazard. However, the "specific model" produces overall more significant coefficients and has a slightly higher Pseudo R-squared (0.76 compared to 0.74 in the general model). An interesting effect is that the tendency in some already very large (or very small) odds ratio takes values suspiciously larger (smaller) in the specific model.

---

[1] Excluding Latin American countries from the sample produced significantly lower hazard rates (around 5% to 6% for these time periods).

The variables "regime" and "ht_region2" are categorical variables that represent a country's political regimes and geographical regions. In order to interpret their effects appropriately, the different categories would need to be included separately in the form of dummy variables. This, however, was not possible since Stata logit routines omitted multiple categories for "perfectly predicting" sovereign defaults in the first-default-only sample. While the estimates for regime are not significant, the geographical region appears to have a significant influence (5%-significance level).

Hughes & MacDonald (2002) and Fight (2004) stress the influence of political, economic, and social factors for sovereign risk. In particular, the use of political factors makes the analysis more complicated, since there is very limited agreement on how to quantify them. A proper theory is missing. However, approaches like Borensztein & Panizza (2009), who refer to the decreased probability of re-election for defaulting governments as a cost of default, demonstrate the importance of political factors. The variables "preselec," no_preselec," "t_pres," "legelec," and "t_leg" have been included in light of these theoretical considerations. For presidential and parliamentary (/legislative) election, a dummy (preselec, legelec) indicates election years. The dummies no_preselec and no_legelec have been constructed to indicate periods where there was no prospect for coming elections (1 = election times; 0 = no elections). The variables t_pres and t_leg are the times since the last election interacted with no_legelec and no_preselec in order to avoid missing values in no-election years.

The only significant estimate is on the 10%-level for preselect in the specified model. According to the odds ratio, the risk of a sovereign default during a presidential election year would be 27 times that of any other year, which seems to be a rather extreme value. One possible explanation could be that the newly-elected government is at minimum risk of not being re-elected, because the next election date will be further away. Another reason, however, comes from Manasse et al. (2003), who use presidential election years to account for political uncertainty that results in economic uncertainty, which increases the risk of default.

"Van_index" and "fh_status" are other political indicators. While for the van_index higher values stand for more democratic institutions, the Freedom House status takes values of one, two or three for relatively free, partially free, or not free countries. While the insignificant estimates for the Van-Index would suggest that more democratic institutions are less likely to default, the highly (1%-level) significant Freedom House status coefficients indicate much less default in less free countries. These interpretations would make sense if one assumes that democratic countries are perceived as less likely to default in the market, and can therefore more easily amass higher debt levels, which increase the risk of a sovereign default. In an effort to test this further, we dropped fh_status variable and run the regression again. This produced a statistically significant and positive coefficient of van_index (10% significance level). This further confirms the above point. The correlation coefficient between van_index and fh_status is -0.8383 which perhaps explains the above result in the sense that fh_status captures the main effect in the presence of both variables been used in the regression. Manasse and Roubini (2009) also argue with proportional representation that the most frequent electoral system would also be the most frequent to default. A more detailed inspection of the data reveals that majority of the countries with fh_status values equal to 3 are those one where decision to default is made predominantly not based on market feedback and ground realities but rather by one or few individuals at the helm of affairs[2].

The variables "g7_us," "g7_us_vol" and "libor_6us_vol" are included to represent monetary conditions in important capital-lending countries. Other G7-state yields have been considered, but were discarded for being highly collinear. Manasse and Roubini (2009) describe how contractions in the G7 may result in reduced capital flows to other markets. While they do not have a significant influence in the general model, in the specific model the approximated volatilities of US treasury bill and USD-denominated LIBOR are significant at the 5%-level. An interesting result is that both volatilities seem to have a strong impact (large coefficients/odds ratios) which is negative for the US treasury bill, but positive for the LIBOR. The negative coefficient for the volatility of the treasury bills should probably be interpreted together with the coefficients for its absolute value.

---

[2] Some of these countries include Zimbabwe, Tunisia, Thailand, Russia, Myanmar, Egypt, Chili, Algeria and China.

The inclusion and choice of macroeconomic variables and their approximated volatilities are motivated mainly from the insights of Catao and Sutton (2002), Manasse et al. (2003), and Manasse and Roubini (2009). These are the inflation rate ("infl;" "infl_vol3"), exchange rates ("xrat_avg;" "xrat_avg_vol3"), and the terms of trade ("tot"). Other variables ("exp_gr" = export growth in %; "exp_imp" = exports to imports ratio; "trade_gdp" = volume of exports plus imports as a ratio of the GDP) have been added to investigate effects of direction and volume of international trade flows. Sturzenegger and Zettlemeyer (2006) identify a worsening in terms of trade as a source of default risk (Das et al., 2012). Theoretically, default reduces trade volumes as it hurts a country's credibility and reduces its access to trade credits (Yeyati & Panizza, 2011). Less favorable terms of trade reduce incentives to conduct trade, which reduces the expected costs of a default and thereby the incentive for repayment. Furthermore, Hilscher and Nobusch (2010) found that countries strong in exports are exposed to external shocks more severely. According to them, the volatility of the terms of trade serves as a statistically and economically significant predictor for sovereign default spreads.

While neither of the models finds the terms of trade or trade as a ratio of GDP to be significant, each model reports highly significant negative estimates of the export growth and significantly positive coefficients for the exports to imports ratio. In different previous versions of the model, the growth of import volumes was included and seemed to positively affect the default probability. Here, an increase in the export volume growth of one% reduces the annual hazard rate by ten%. The most obvious explanation is that export streams generate revenues to the country which positively influence its solvency, whereas imports are associated with an outflow of capital and therefore negatively affect the trade balance.[3] It is then interesting to see that the export over import ratio has such a strong positive influence for the default probability (note that exp_imp is not denoted in % but in absolute terms, which relativizes the size of this effect). The key to this lies in the negative relationship between overall trade volumes and the risk of default (see, for example the negative coefficient for

---

[3] The significance of the trade balance may explain the insignificance of the current account balance ("ca_net") that could only contribute to the analysis through influential changes in the factor income or cash transfers.

trade_gdp or the above explanations regarding the terms of trade). This could be an indication that it is more likely that the export/import-ratio increases by a reduction in the import volumes than by an increase in the volume of exports. This overall reduction in the trade volumes would explain the positive coefficient.

Both models identify the inflation and exchange rate variables as highly significant. Manasse et al. (2003) use highly volatile inflation to identify political uncertainty and high inflation as a measure of monetary mismanagement. Low inflation, for example, would indicate macroeconomic stability which appeals to potential investors. Both the general and specific models show similar effects: while a one% increase in the absolute inflation rate leads to a seven% higher per annum risk of default (or hazard), this risk decreases by the same amount for each point increase in the approximated 5-year inflation volatility. This finding is consistent with the classical economic theory, particularly Eaton and Gersovitz (1981), where sovereign borrowing serves as insurance against output contractions and defaulters are excluded from future lending. More specifically, the authors argue that higher volatilities act as a deterrent to sovereign defaults. All those countries with volatile macroeconomic conditions mostly depend on foreign borrowing to smooth their future consumption. Defaulting on foreign debt would invite harsh penalties for 'bad borrowers' in particular of exclusion from future borrowing[4]. In this regard, higher macroeconomic volatility would make a country more dependent on its foreign investors and therefore less likely to default (for further discussion in this regard, see also Catao & Sutton, 2002).

Similarly, direct effects can be observed for the exchange rate and its approximated five-year volatility. The (highly) significant coefficients of the debt-to-GDP ratios support the existing literature. Not surprisingly, they have a positive influence on the risk of default. The influence of a one% increase in the total amount of external debt is five% (seven% for the specified model) higher hazard rates, which is one (three)% larger than the

---

4 A number of authors found evidence for an exclusion form international credit markets following a default (Aguiar and Gopinath, 2006; Amador, 2009; Arellano, 2008; Asonuma, 2010; Mendoza and Yue, 2008; Tomz and Wright, 2007; and Yue, 2010; as cited in Das et al., 2012)

influence of a one% increase in the total central debt over GDP ratio. Also not surprising is the very significant negative relationship between current GDP per capita or its growth rates and the occurrence of a default. Potential endogeneity-issues have not been addressed; therefore, the estimated coefficient can only be interpreted as a correlation not causation.

One more explanatory variable, which is significant at the 1%-level, is a dummy indicating banking crises ("cr_bk_lag1"). It has been lagged by one period to account for the controversial causality between banking and currency crises. It appears that in the general model, the risk of entering into default is 35 times higher in a period following a period with a banking crisis (27 times in the specified model). Both models have a significantly negative intercept and baseline hazard. The baseline hazard in the specified model is distinctly more negative: While the coefficient r is -2.87 for all periods in the general model, periods one to 14 in the specified model follow a baseline logit-hazard with a coefficient of -14.74, the consecutive periods have a constant baseline hazard which reduces the intercept by -11.54. The lower baseline hazards in the specified model probably make up for the larger odds ratios for the explanatory variables.

In a next step, the two-state model with the same explanatory variables should give an impression of whether or not the findings from the first-default-only dataset can be applied to the full default dataset. One issue that may influence the transferability is that the exclusion of each country after its first default makes the analysis highly vulnerable to historical occurrences such as international financial crises. Countries are connected via international trade, financial- and capital markets. Ongoing globalization makes them more vulnerable to systemic risk and contagion (Reinhart & Rogoff, 2009). Reinhart & Rogoff (2009) claim that an analysis of sovereign default with data limited to only a few recent decades would be insufficient to identify and account for the very long cycles of sovereign defaults. The two-way model could in this regard be a first step to explore the usability of duration (survival) methodology for the analysis of these long cycles. The separation of observations into two or more different states may also give an impression of how certain indicators change from entrance to exit of a spell like sovereign default crisis.

Table 4 contains the estimates of the two-way model as specified in the methodology section. Similar to first default analysis, we estimated three different specifications for both out of and into default (column 1, 3 & 4). For subsequent discussion, we mainly focus on our main specification estimates and related odds ratios (column 1&2). Overall, there are fewer significant coefficients. Furthermore, many estimates became very insignificant and would probably be excluded from a more thorough analysis. For the purpose of this study, it is, however, interesting to see how the estimates change with the use of multiple-spell data and a frailty-model that addresses individual heterogeneity. The standard deviation of the random effects terms $\hat{\sigma}_u$ of 0.001 seems to be fairly low indicating that the magnitudes of the coefficients would differ too much than the ones in the single-level model (Snijders & Bosker, 1994, p. 227-8). For the transition out of default, the constant is significantly positive and the significant coefficient for the baseline logit-hazard is distinctly more negative than for the transition into default.

One important point to notice is that the baseline hazard for the transition into default, which can be thought of as the probability of defaulting ceteris paribus (given none of the explanatory variables in the model change), seems not to be appropriately represented by $r \log(j)$ anymore. This may be due to the fact that the multiple-default dataset analyses the whole default and non-default episodes whereas in the first-default-only dataset it was unknown when the spell at the beginning of the observation period originally started. Focusing only on significant coefficients, the two-state model allows stating that the interest rate of the US treasury bill is associated negatively with the end of a sovereign debt crisis and positively with its beginning. Manasse et al. (2003), Borri and Verdelhan (2011) and Longstaff *et al.* (2011) find the same effect and argue that monetary conditions in the G7 influence emerging markets' abilities to refinance, which may drive them to a liquidity crisis. The total central debt over GDP ratio seems to be positively associated with defaulting and negative with the end of a default spell. Interesting findings include the positive coefficient of the current account balance for transition into default and the change in direction for the effect of the approximated exchange rate volatility on default, which is now positive. It is, however, important to note that the model was not tested for misspecification which is probably an issue with so many insignificant coefficients.

6. Conclusion

Sovereign debt crises are frequently reoccurring, often contagious international events. History has seen multiple episodes of amassed sovereign defaults, mainly in developing countries. Reinhart and Rogoff (2009) call their occurrence cyclical. Even though all involved economic agents should be aware of the threat of sovereign defaults, it appears that the international capital markets repeatedly underestimate this risk. Related decision-making is often based on perception rather than a thorough technical analysis (Hughes & MacDonald, 2002). There simply is no such thing as a reliable prediction of sovereign default as the analysis is more art than science (Fight, 2004).

The aim of this paper is to estimate determinants of sovereign default using a comprehensively constructed panel-dataset with yearly observations for 70 countries between 1970 and 2010. The gaps in the data reduce the number of observations in the final model to 1953. Survival (duration) analysis was used for this purpose as this methodology addresses some distinct features of sovereign debt data: as most data in the social sciences is continuous time data that is grouped into discrete intervals; most explanatory variables are time-varying; and it can accommodate country-specific heterogeneity when estimating population-related coefficients. The analysis followed two steps: in a first step, a logit or proportional odds model was used to analyze the times from 1970 until the first observed default. This type of model is commonly used in social sciences for the estimation of discrete-time data. It allows for easy inclusion of time-varying covariates and can be approximately representative for underlying continuous time processes.

Several explanatory variables have been excluded based on economic theory and experiences of earlier approaches to model the risk of sovereign default. A basic problem all estimations face is that identified effects can hardly be interpreted as causalities due to the involvement of complex macroeconomic processes. For example, the dependent variable can often cause trends in one or more of the independent variables (Villemont, 2011). A good example is the simultaneous occurrence of different financial crises, for example, banking and sovereign debt crises. A sovereign default can cause domestic contagion and thus motivate banking crises.

24

However, banking crises also possess the power to cool down an economy, forcing a sovereign into default. Zettelmeyer (2012) speaks of a standard reverse causality problem. The estimated logit model finds significant negative effects on the risk of survival (not defaulting) from macroeconomic volatilities in the interest rates of the US treasury bill and LIBOR, as well as inflation. The seminal model by Eaton and Gersovitz (1981) can help to explain this: it assumes sovereign borrowing to be an insurance against the occurrence of output contractions. Creditors punish with exclusion from future lending when a sovereign debtor defaults. Macroeconomic volatility may increase the perceived need for insurance against future "bad times." This finding is consistent with Catao and Sutton (2002), Manasse et al. (2003), Manasse and Roubini (2009).

Manasse and Roubini (2009) explain that in a presidential election year, the rising political risk increases the economic risk and thus the probability for a default to happen. This study supports this finding with a highly significant positive effect of presidential election years. While the terms of trade are not significant, growth of the export volumes is representing rising revenues to the country, and thus significantly indicates reduced default probabilities. However, rises in the export over import ratio cause a higher likelihood of default, which can be explained assuming that it is more likely to be caused by reduced imports than increased exports. The overall reduced trade levels positively affect the default risk. Low exchange rates indicate political and monetary stability, which is attractive to investors. The coefficient, thus, has a negative sign. This study can also replicate the widely-used results that debt-over-GDP ratios have positive effects and GDP levels as well as growth rates have negative effects on the risk of default. A one-period lagged indicator of banking crises positively indicates sharp rises in the default probability.

References

Afonso, A., & Silva, J. (2017). Debt crisis and 10-year sovereign yields in Ireland and in Portugal. Applied Economics Letters, 1-6.

Banasik, J., Crook, J. N., & Thomas, L. C. (1999). Not if but when will borrowers default. Journal of the Operational Research Society, 1185-1190.

Bellotti, T., & Crook, J. (2009). Credit scoring with macroeconomic variables using survival analysis. Journal of the Operational Research Society, 60(12), 1699-1707.

Borensztein, E., & Panizza, U. (2009). The costs of sovereign default, IMF Staff Papers, 56(4), 683-741.

Bormann, N.-C. & Golder, M. (2013). Democratic electoral systems around the world, 1946-2011 [data file]. Retrieved from https://files.nyu.edu/mrg217/public/elections.html

Borri, N., & Verdelhan, A. (2011). Sovereign risk premia.

Broto, C., & Perez-Quiros, G. (2015). Disentangling contagion among sovereign CDS spreads during the European debt crisis. Journal of Empirical Finance, 32, 165-179.

Cantor, R., & Packer, F. (1996). Determinants and impact of sovereign credit ratings. Economic Policy Review, 2(2), 37-54.

Catao, L., & Sutton, B. (2002). Sovereign defaults: the role of volatility (IMF Working Paper WP/02/149).

Cencini, A. (2017). The sovereign debt crisis: The case of Spain. Cuadernos de Economía, 40(112), 1-13.

Cleves, M. (2009). Analysis of multiple failure-time survival data. Retrieved from www.stata.com/support/faqs/statistics/multiple-failure-time-data/

Consortium for Elections and Political Process Strengthening. (2013). Election Guide. Retrieved from http://www.electionguide.org

Cruces, J.J. & Trebesch, C., (2013). Sovereign defaults: The price of haircuts. American Economic Journal: Macroeconomics, 5(3), pp.85-117.

Das, U. S., Papaioannou, M. G., & Trebesch, C. (2012). Sovereign debt restructurings 1950-2010: Literature survey, data, and stylized facts. International Monetary Fund.

Eaton, J. & Gersovitz, M. (1981). Debt with potential repudiation: Theoretical and empirical analysis. The Review of Economic Studies, 48(2), 289-309.

Ehrmann, M., & Fratzscher, M. (2017). Euro area government bonds–Fragmentation and contagion during the sovereign debt crisis. Journal of International Money and Finance, 70, 26-44.

Eichler, S., & Maltritz, D. (2012). The term structure of sovereign default risk in EMU member countries and its determinants. Journal of Banking & Finance, 36.

Feenstra, R. C., Inklaar, R., & Timmer, M. P. (2013). Penn World Table (Version 8) [data file and codebook]. Retrieved from www.ggdc.net/pwt

Fight, A. (2004). Understanding International Bank Risk. Chichester: John Wiley & Sons Ltd.

Golder, M. (2005). Democratic electoral systems around the world, 1946-2000 [data file]. Retrieved from https://files.nyu.edu/mrg217/public/elections.html

Gómez-Puig, M., & Sosvilla-Rivero, S. (2016). Causes and hazards of the euro area sovereign debt crisis: Pure and fundamentals-based contagion. Economic Modelling, 56, 133-147.

Heston A., Summers R., & Aten, B. (2012). Penn World Table (Version 7.1) [data file and codebook]. Retrieved from https://pwt.sas.upenn.edu/php_site/pwt_index.php

Hilscher, J., & Nosbusch, Y. (2010). Determinants of sovereign risk: macroeconomic fundamentals and the pricing of sovereign debt. Review of Finance, 14, 235-262.

Hughes, J. E., & MacDonald, S.B. (2002). International Banking: Text & Cases. New York: Person Education.

International Monetary Fund. (2013a). International Financial Statistics (2013) [database]. Retrieved from http://elibrary-data.imf.org/DataExplorer.aspx

International Monetary Fund. (2013b). International Financial Statistics (Edition: 2013) [database]. Retrieved from http://dx.doi.org/10.5257/imf/ifs/2013-08

International Monetary Fund. (2013c). World Economic Outlook (Edition: April 2013) [data file and codebook]. Retrieved from http://www.imf.org/external/pubs/ft/weo/2013/01/weodata/index.aspx

Jenkins, S. P. (2005). Survival analysis. Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Keefer, P. E. (2013). Database of Political Institutions (Edition: Jan. 2013) [data file and codebook]. Retrieved from http://go.worldbank.org/2EAGGLRZ40

Lane, P. R., & Milesi-Ferretti, G. M. (2007). Updated and Extended "External Wealth of Nations" Dataset, 1970-2011[data file]. Retrieved from http://www.philiplane.org/EWN.html

Larraín, G., Reisen, H., & von Maltzan, J. (1997). Emerging market risk and sovereign credit ratings (Working Paper No. 124). OECD Development Centre.

Longstaff, F. A., Pan, J., Pedersen, L. H., & Singleton, K. J. (2011). How sovereign is sovereign credit risk? American Economic Journal: Macroeconomics, 3(2), 75-103.

Manasse, P., & Roubini, N. (2009). "Rules of thumb" for sovereign debt crises. Journal of International Economics, 78, 192-205.

Manasse, P., Roubini, N., & Schimmelpfennig, A. (2003). Predicting sovereign debt crises (IMF Working Paper, WP/03/221).

Moisescu, E. R., & Giurescu, A. (2016). Sovereign debt crisis. From challenges to solutions. Theoretical and Applied Economics, 23(1 (606), Spring), 195-202.

Neuhaus, J. M., Kalbfleisch, J. D., & Hauck, W. W. (1991). A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. International Statistical Review/Revue Internationale de Statistique, 25-35.

Organisation for Economic Development and Cooperation. (2009). National Accounts (Edition: July 2009). Retrieved from DOI: 10.5257/oecd/na/2009-07

Popov, A., & Van Horen, N. (2015). Exporting sovereign stress: Evidence from syndicated bank lending during the euro area sovereign debt crisis. Review of Finance, 19(5), 1825-1866.

Reinhart, Carmen, Kenneth Rogoff, & Miguel Savastano. 2003. Debt intolerance. Brookings Papers on Economic Activity, 2003(1): 1–70.

Reinhart, C. M., & Rogoff, K. S. (2009). This time it's different: eight centuries of financial folly. Princeton: Princeton University Press.

Reinhart, C. M., & Rogoff, K. S. (2010a). From financial crash to debt crisis. American Economic Review, 101(5), 1676-1706.

Reinhart, C. M., & Rogoff, K. S. (2010b). From financial crash to debt crisis [data files including explanations]. Retrieved from http://www.reinhartandrogoff.com/data

Reusens, P., & Croux, C. (2017). Sovereign credit rating determinants: A comparison before and after the European debt crisis. Journal of Banking & Finance, 77, 108-121.

Roubini, N., & Setser, B. (2003, March). Improving the sovereign debt restructuring process: Problems in restructuring, proposed solutions, and a roadmap for reform. In Prepared for the Conference on Improving the Sovereign Debt Restructuring Process. Institute for International Economics and Institut Français des Relations Internationales, Paris, March (Vol. 9).

Smeets, D. (2016). Financial Contagion during the European sovereign debt crisis. Journal of Economic and Financial Studies, 4(02), 46-59.

Snijders, T.A.B. & Bosker, R.J. (1994). Modeled variance in two-level models. Sociological Methods and Research, 22, 342-363.

Stamatopoulos, T. V., Arvanitis, S. E., & Terzakis, D. M. (2016). The risk of the sovereign debt default: the Eurozone crisis 2008–2013. Applied Economics, 1-15.

Steele, F. (2011). Multilevel discrete-time history models with applications to the analysis of recurrent employment transitions. Australian & New Zealand Journal of Statistics, 53(1), 1-26.

Sturzenegger, F., & Zettelmeyer, J. (2006). Debt defaults and lessons from a decade of crises. MIT press.

Sueyoshi, G. T. (1995). A Class of binary response models for grouped duration data. Journal of Applied Econometrics, 10(4), 411-431.

Tamborini, R. (2015). Heterogeneous market beliefs, fundamentals and the sovereign debt crisis in the Eurozone. Economica, 82(s1), 1153-1176.

Teorell, J., Charron, N., Dahlberg, S., Holmberg, S., Rothstein, B., Sundin, P., & Svensson, R. (2013, August). The Quality of Government Dataset (Edition: 15May13) [data file and codebook]. Retrieved from http://www.qog.pol.gu.se

The Conference Board. (2013). Total Economy Database (Edition: January 2013) [data files]. Retrieved from http://www.conference-board.org/data/economydatabase/

Thomas, L. C., Edelman, D. B., & Crook, J. N. (2002). Credit scoring and its applications. Society for industrial and Applied Mathematics.

Tomz, M., & Wright, M. L. J. (2007). Do countries default in "bad rimes"? Journal of the European Economic Association, 5(2-3), 352-360.

Ucler, G., & Kirmizioglu, H. (2015). The reasons of Eurozone sovereign debt crisis and an empirical analysis over permanency of the crisis. International Journal of Economics and Financial Issues, 5(1), 86.

World Bank (2013a) International Debt Statistics (Edition: 2013) [database]. Retrieved from http://data.worldbank.org/data-catalog/international-debt-statistics

World Bank (2013b) World Development Indicators (Edition: July 2013) [database]. Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators

Yeyati, E. L., & Panizza, U. (2011). The elusive costs of sovereign defaults. Journal of Development Economics, 94, 95-105.

Zettelmeyer, J. (2012). How to do a sovereign debt restructuring in the Euro area: Lessons from emerging-market debt crisis. In W. R. Cline, & G. B. Wolff (Eds.). Resolving the European Debt Crisis (pp. 165-186). Washington, DC: Peterson Institute for International Economics, Bruegel.

Table 1 Sample of Countries and Default Periods

Country	id	First Default Only Dataset		Multiple Default Dataset	Country	id	First Default Only Dataset		Multiple Default Dataset
		Years	Default	Default Periods			Years	Default	Default Periods
					Korea (South)	36	1970 - 2010	0	
Algeria	1	1970 - 1991	1	1991 – 1996	Malaysia	37	1970 - 2010	0	
Angola	2	1970 - 1985	1	1985 – 2003	Mauritius	38	1970 - 2010	0	
Argentina	3	1970 - 1982	1	1982 – 1993 2001 - 2005	Mexico	39	1970 - 1982	1	1982–1990
Australia	4	1970 - 2010	0		Morocco	40	1970 - 1983	1	1983 1986–1990
Austria	5	1970 - 2010	0		Myanmar	41	1970 - 2002	1	2002–2010
Belgium	6	1970 - 2010	0		Netherlands	42	1970 - 2010	0	
Bolivia	7	1970 - 1980	1	1980 – 1984 1986 – 1997	New Zealand	43	1970 - 2010	0	
Brazil	8	1970 - 1983	1	1983 – 1994 2002	Nicaragua	44	1970 - 1979	1	1979–2010
Canada	9	1970 - 2010	0		Nigeria	45	1970 - 1982	1	1982–1992 2001 2004–
Central African Republic	10	1970 - 1981	1	1981 1983 - 2010	Norway	46	1970 - 2010	0	
Chile	11	1970 - 1983	1	1972 1974–1975 1983–	Panama	47	1970 - 1983	1	1983 – 1996
China	12	1970 - 2010	0		Paraguay	48	1970 - 1986	1	1986 – 1992 2003–2004
Colombia	13	1970 - 2010	0		Peru	49	1970 - 1976	1	1976 1978 1980 1984–1997
Costa Rica	14	1970 - 1981	1	1981 1983 – 1990	Philippines	50	1970 - 1981	1	1981–1992
Cote d'Ivoire	15	1970 - 1983	1	1983 – 1998 2000–2010	Poland	51	1970 - 1981	1	1981–1994
Denmark	16	1970 - 2010	0		Portugal	52	1970 - 2010	0	
Dominican Republic	17	1970 - 1982	1	1982 – 1994 2005	Romania	53	1970 - 1981	0	1981 – 1983 1986
Ecuador	18	1970 - 1982	1	1982 – 1995 1999–2000	Russia	54	1987 - 1991	1	1970 – 1986 1991–2000
Egypt	19	1970 - 1985	1	1984	Singapore	55	1970 - 2010	0	
El Salvador	20	1970 - 2010	0		South Africa	56	1970 - 1985	1	1985 – 1987 1989 1993
Finland	21	1970 - 2010	0		Spain	57	1970 - 2010	0	
France	22	1970 - 2010	0		Sri Lanka	58	1970 - 1979	1	1979 1981–1983
Germany	23	1970 - 2010	0		Sweden	59	1970 - 2010	0	
Ghana	24	1971 - 1974	1	1970 1974 1987	Switzerland	60	1970 - 2010	0	
Greece	25	1970 - 2010	0		Taiwan	61	1970 - 2010	0	
Guatemala	26	1970 - 1986	1	1986 1989	Thailand	62	1970 - 2010	0	
Honduras	27	1970 - 1981	1	1981 – 2010	Tunisia	63	1970 - 1979	1	1979– 982
Hungary	28	1970 - 2010	0		Turkey	64	1970 - 1978	1	1978 – 1979 1982 2001
Iceland	29	1970 - 2010	0		UK	65	1970 - 2010	0	
India	30	1970 - 1972	1	1972 – 1976	US	66	1970 - 2010	0	
Indonesia	31	1971 - 1998	1	1970 1998–2000 2002	Uruguay	67	1970 - 1983	1	1983–85 1987 1990–91
Ireland	32	1970 - 2010	0		Venezuela	68	1970 - 1983	1	1983–88 1990 1995–97
Italy	33	1970 - 2010	0		Zambia	69	1970 - 1983	1	1983 – 1994
Japan	34	1970 - 2010	0		Zimbabwe	70	1975 - 2000	1	1970–1974 2000–2009
Kenya	35	1970 - 1994	1	1994 – 2003					

Note: This table contains information on the overage of countries and defaults episode periods. First default analysis data is comprised of data rows for each country until the first incident of default (denoted by default =1 and 0 otherwise). For multiple defaults and transition into and out of default, the dataset comprises of each country name alongside the periods when a country experienced default. The dummy variable =1 is recorded for all those periods when a country remained unable to service debt and 0 otherwise.

Table 2 Description and Source of Dependent and Independent Variables

Variable name	Description	Source[5]
cr_esd	Dummy indicating external sovereign default (=1)	RR
y	Dummy, indicating transition into and out of default	Constructed using cr_esd
y_def	Dummy indicating a transition into default (=1)	Constructed using cr_esd
y_nodef	Dummy indicating a transition out of default (=1)	Constructed using cr_esd
s_def	Dummy indicating the country is in default (=1)	Constructed using cr_esd
s_nodef	Dummy indicating the country is not in default (=1)	Constructed using cr_esd
t	Counting the years from start (1970) to end (2010) of the observation period with 1970=1; 2010=41. In the first-default-only dataset, t has been adjusted for the four countries with delayed entry, taking "0" in the entry-year.	
defaultct	Counting a country's transitions into default. (1970=0)	Constructed using cr_esd
cr_bk_lag1	One year lagged dummy indicating the country suffers from a banking crisis	cr_bk
ht_region2	Categorical variable for the geographical region: (1) Eastern Europe and post Soviet Union (including Central Asia) (2) Latin America (including Cuba, Haiti & the Dominican Republic) (3) North Africa & the Middle East (including Israel, Turkey & Cyprus) (4) Sub-Saharan Africa (5) Western Europe and North America (including Australia &New Zeeland) (6) East Asia (including Japan & Mongolia) (7) South-East Asia (8) South Asia	Hadenius, Teorell & Wahman (2007) as cited in QoGst
regime	Categorical variable for the form of government: (1) Parliamentary democracy; (2) Semi-presidential democracy; (3) Presidential democracy; (4) Civilian dictatorship; (5) Military dictatorship; (6) Royal	Constructed using data from Cheibub et al. (2010), Marshall & Jaggers (2011), and as cited in QoGst; Bormann & Golder (2013); DPI
preselect	Dummy indicating presidential election year (=1)	Created using data from Golder (2005); Bormann, N.-C. & Golder, M. (2013); Consortium for Elections and Political Process Strengthening. (2013)
legelec	Dummy indicating legislative/parliamentary election year (=1)	Created using data from Golder (2005) Bormann, N.-C. & Golder, M. (2013); Consortium for Elections and Political Process
no_preselec	0= no presidential elections conducted or observed since establishment of regime or 1946; 1=known last election date and regime conducts elections	Created based on legelec; DPI
t_leg	leg_last*no_legelec	
t_pres	pres_last*no_preselec	
Van_index	"This index combines two basic dimensions of democracy – competition and participation –measured as the %age of votes not cast for the largest party (Competition) times the %age of the population who actually voted in the election (Participation). This product is divided by 100 to form an index that in principle could vary from 0 (no democracy) to 100 (full democracy). (Empirically, however, the largest value is 49.)"	Vanhanen (2011) as cited in QoGst

---

[5] Abbreviations stand for:
DPI   Keefer (2013)
IFS   International Monetary Fund (2013a)
OECD   Organisation for Economic Development and Cooperation (2009)
PWT7   Heston et al. (2012)
PWT8   Feenstra et al. (2013)
QoGst   Teorell et al. (2013)
RR   Reinhart & Rogoff (2010b)
TED   The Conference Board (2013)
WDI   World Bank (2013b)
WEO   International Monetary Fund (2013c)

fh_status	Freedom House status; (1) free (2) partial free (3) not free	Freedom House (2013) as cited in QoGst; where single values were missing the value of the previous year was used
wpr_3spot	World oil-price index (Index number: 00176AADZF)	International Monetary Funds (2013b)
wpr_3spot_growth	Logarithmic returns of wpr_3spot (in %)	
gdp_pc_curr	GDP per capita (current US$)	WDI
gdp_pc_growth	GDP per capita growth (annual %)	WDI
g7_us	US T-Bill	IFS
g7_us_vol	US T-Bill volatility (approximated with the square root of the 5-year variance)	IFS
libor_6us_vol	libor_6_us volatility (approximated with the square root of the 5-year variance)	Calculated using data from DataStream, Series: USY60LDE (Retrieved August 2013)
infl	Inflation in% per annum	RR
Infl_vol3	Volatility of infl (approximately, using the square-root of infl_variance).	
xrat_avg	National Currency per USD, period average	IFS; for the Euro Zone members the rate for 'Euro Area' was used; for Ecuador 1960-1979: WDI ("Official exchange rate, period average"); for Indonesia 1960-1966: PWT7 ("XRAT"); for Russia 1970-1992: United nations 2013 ("unna_er") as cited in QoGst; Taiwan 1970-2010: PWT7("XRAT")
xrat_avg_vol3	xrat_avg volatility (approximated with square root of xrat_avg_variance)	
exp_gr	Exports of goods and services (annual% growth)	WDI
exp_imp	exp_usd divided by imp_usd	
tot	"Terms of trade adjustment (constant LCU) Net barter terms of trade index is calculated as the%age ratio of the export unit value indexes to the import unit value indexes, measured relative to the base year 2000"	WDI
trade_gdp	"…the sum of exports and imports of goods and services measured as a share of gross domestic product." (% of GDP)	WDI
ca_net	Current Account, Net (excluding exceptional financing)	IFS; for Taiwan & Algeria (1991-2004), Belgium (1980-2001), Central Africa (1995 onwards), China 1980-1 & Hungary 1980-1 WEO; El Salvador(up to '75),Egypt (up to '76), Ecuador ('70-'75), Denmark, ('70-'74), Cote D'Ivoire, ('70-'74), Costa Rica ('70-'76), Chile ('70-'74), Brazil ('70-'74), Algeria prior to 1977, Angola prior to 1985, Argentina prior to 1976, Bolivia prior to 1976 from Lane, P. R., & Milesi-Ferretti, G. M. (2007)
debt_tot_ctr_gdp	Total (domestic plus external) gross central government debt/GDP	Calculated using data from RR; WDI; WEO; OECD
debt_tot_ext_gdp	Total external (public plus private) gross external Debt/GDP	Calculated using data from RR; WDI; WEO; OECD

Note: This table contains information on the variable names (dependent and independent) alongside description and detail of each variable and sources.

Table 3 Determinants of Sovereign Defaults (proportional odds model – first default only)

|	General Logit Model				Specified Logit Model			
	Coef.	Odds Ratio	Coef.	Coef.	Coef.	Odds Ratio	Coef.	Coef.
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
regime	0.6140	1.8479		0.8863	1.2375	3.4468		0.9978*
ht_region2	0.9296**	2.5335		0.6471*	0.9501**	2.5860		0.5947*
preselect	2.1489	8.5756		2.7407**	3.3011*	27.1434		2.5070**
no_preselec	3.5363	34.3413		1.8808	2.5156	12.3742		1.4823
t_pres	0.1114	1.1179		0.1877**	0.1327	1.1419		0.1738*
legelec	-0.9910	0.3712		-1.2538	-1.7373	0.1760		-1.0520
t_leg	-0.3066	0.7360		-0.2395	-0.2975	0.7426		-0.1982
van_index	-0.0408	0.9600		-0.0095	-0.0535	0.9479		-0.0113
fh_status=2	-5.4559***	0.0043		-4.2457***	-6.6527***	0.0013		-4.4019***
fh_status=3	-6.4132***	0.0016		-4.8308***	-7.2892***	0.0007		-5.0669***
g7_us	0.0818	1.0852	0.1386	0.0063	0.5291	1.6975	0.2288	0.1449
g7_us_vol	-2.3358	0.0967	-1.8334		-6.2885**	0.0019	-2.3581	
libor_6us_vol	2.1200	8.3310	1.7297		5.4688**	237.1661	2.2109	
ca_net	2.26e-11	1.0000	6.89e-12	7.01e-12	2.26e-11	1.0000	9.81e-12	5.82e-12
infl	0.0575***	1.0591	0.0285***	0.0135***	0.0566***	1.0582	0.0281***	0.0122***
infl_vol3	-0.0617**	0.9402	-0.0298**		-0.0605**	0.9413	-0.0292**	
wpr_3spot	0.0787	1.0819	0.0436	0.0782**	0.0966	1.1015	0.0325	0.0675**
wpr_3spot_growth	-0.0146	0.9855	-0.0036	0.0045	-0.0072	0.9928	-0.0011	0.0012
xrat_avg	-0.0197**	0.9805	-0.0103*	-0.0001	-0.0204**	0.9798	-0.0101**	-0.0003
xrat_avg_vol3	0.0329**	1.0335	0.0175**		0.0334**	1.0340	0.0170**	
exp_gr	-0.1009**	0.9040	-0.0769**	-0.0689**	-0.1200**	0.8869	-0.0768**	-0.0658**
tot	-8.96e-14	1.0000	1.60e-14	1.11e-14	-8.83e-14	1.0000	1.56e-14	1.05e-14
exp_imp	4.5991*	99.3923	4.1333**	4.2083**	5.4681*	236.9991	4.0618**	4.2866**
trade_gdp	-0.0263	0.9740	-0.0233	-0.0263	-0.0395	0.9613	-0.0249	-0.0345
debt_tot_ext_gdp	0.0520**	1.0534	0.0344**	0.0289	0.0717***	1.0743	0.0360**	0.0368*
debt_tot_ctr_gdp	0.0394*	1.0402	0.0203	0.0377**	0.0434**	1.0443	0.0191	0.0340**
gdp_pc_curr	-0.0021***	0.9979	-0.0012***	-0.0017***	-0.0021***	0.9980	-0.0012***	-0.0016***
gdp_pc_growth	-0.3452**	0.7081	-0.1522**	-0.2822***	-0.3928***	0.6752	-0.1496*	-0.2605**
cr_bk_lag1	3.5673***	35.4200	1.9367***	3.6698***	3.2828***	26.6516	1.8599**	3.1232***
logt	-2.8656*	0.0570	-0.9982	-2.6993**				
logt14					-7.0198**	0.0009	-0.9234	-3.0743
e15					-14.7395**	0.0000	-2.3725	-8.1062*
constant	-12.3770**	4.21E-06	-8.9720***	-9.9068**	-11.5388**	9.74E-06	-9.6127***	-9.6616**
Log likelihood	-32.2681		-46.765	-38.8176	-30.1528		-46.9330	-39.8057
Pseude R_square	0.7419		0.6313	0.6902	0.7589		0.6300	0.6823
Prob > chi^2	0.0000		0.000	0.0000	0.0000		0.0000	0.0000
Number of observations	1186		1270	1198	1186		1270	1198

Note: This table contains regression estimates of duration (survival) analysis models. Two different models are specified and estimated. One without taking into account of significant increase in hazard up to 15 years and second which does. In both cases, table also presents odd ratio of first specification. For the sake of sensitivity analysis, table also contains coefficient estimates of two other specifications for each type of model. More specifically, in second specification (column 3), we drop political and regional variables while in third specification (column 4), we drop volatility variables. *, **, ***significant at 10%, 5% and 1% level of significance. Dependent variable in each model specification is whether a country defaulted or not (1, 0 dummy variable (represented by cr_esd as explained in table 2)). The number of observations used in each model estimation differ due to missing values of certain independent variables.

Table 4 Determinants of Transition Out of and into Default (two-state model)

|	Transition out of Default				Transition into Default			
	Coef.	Odds Ratio	Coef.	Coef.	Coef.	Odds Ratio	Coef.	Coef.
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
regime	0.4053	1.4997		0.2710	0.3905	1.4778		0.3248
ht_region2	-0.1199	0.8870		-0.1292	-0.1166	0.8899		-0.1658
preselec	0.0577	1.0594		-0.0360	0.2133	1.2378		0.2613
no_preselec	-1.6719	0.1879		-1.6406	0.9848	2.6772		0.7761
t_pres	-0.0889	0.9149		-0.1169	-0.0137	0.9864		0.0105
legelec	-0.3211	0.7254		-0.1879	-0.0410	0.9599		-0.1188
t_leg	-0.0165	0.9836		0.0084	0.1184	1.1256		0.1057
van_index	0.0384	1.0392		0.0307	-0.0172	0.9829		-0.0137
fh_status=2	-0.1395	0.8698		-0.1666	-0.8460	0.4291		-0.7250
fh_status=3	-0.5682	0.5665		-0.3627	-2.9979***	0.0499		-2.7327***
g7_us	-0.1865**	0.8299	-0.1586**	-0.1758**	0.2604***	1.2975	0.2569***	0.2379***
g7_us_vol	-0.1101	0.8957	0.2564		-0.4702	0.6249	0.2031	
libor_6us_vol	-0.1016	0.9034	-0.4202		0.2966	1.3453	-0.2950	
ca_net	5.55e-11	1.0000	8.53e-11*	3.88e-11	1.63e-11 ***	1.0000	9.97e-12*	1.51e-11**
infl	0.0005	1.0005	0.0006	-0.0001	0.0184**	1.0186	0.0162**	0.0094***
infl_vol3	-0.0012	0.9988	-0.0012		-0.0134	0.9867	-0.0158	
wpr_3spot	-0.0032	0.9968	-0.0061	-0.0013	0.0159**	1.0160	0.0132**	0.0101
wpr_3spot_growth	1.6469	5.1909	1.5369	1.9461	-0.5892	0.5548	-0.6209	-0.8780
xrat_avg	-0.0000	1.0000	-0.0000	-0.0000	-0.0002	0.9998	-0.0003	0.0003**
xrat_avg_vol3	-	1.0000	-		0.0029*	1.0029	0.0030**	
exp_growth	0.0064	1.0064	0.0112	0.0008	0.0025	1.0025	-0.0057	0.0024
tot	-6.23e-14	1.0000	-6.52e-14	-4.83e-14	-9.03e-16	1.0000	4.71e-15	2.38e-14
exp_imp	-1.5325	0.2160	-1.9535**	-0.9920	2.0400**	7.6908	1.3966*	2.2200***
trade_gdp	-0.0019	0.9981	-0.0004	-0.0005	-0.0082	0.9919	-0.0140*	-0.0085
debt_tot_ext_gdp	0.0076	1.0077	0.0046	0.0080*	0.0018	1.0018	0.0055*	0.0030
debt_tot_ctr_gdp	-0.0235***	0.9768	-0.0211**	-0.0218***	0.0264***	1.0267	0.0176***	0.0238***
gdp_pc_curr	0.0003	1.0003	0.0004**	0.0002	-0.0001	0.9999	-0.0001***	-0.0001**
gdp_pc_growth	0.0047	1.0047	0.0122	0.0312	-0.1556***	0.8559	-0.1681***	-0.1530***
cr_bk_lag1	-0.8138	0.4432	-0.7659	-0.8429*	0.6407	1.8977	0.6309	0.6454
logt	-0.5472**	0.5786	-0.6275**	-0.5392**	-0.0925	0.9117	-0.3818*	-0.0132
constant	3.6014*	36.6501	3.4820**	2.9775	-9.5209***	0.0001	-6.0924***	-8.9457***
/lnsig2u	-13.3696		-13.3276	-10.9666	-13.3696		-13.3276	-10.9666
sigma_u	0.0012		0.0013	0.0042	0.0012		0.0013	0.0042
e(sigma_u)^2	1.562e-06		1.629e-06	1.727e-05	1.562e-06		1.629e-06	1.727e-05
rho	4.75e-07		4.95e-07	5.25e-06	4.75e-07		4.95e-07	5.25e-06
LR-test(rho=0)	0.4980		0.4980	0.4970	0.4980		0.4980	0.4970
Log likelihood	-219.5420		-243.0418	-227.7908	-219.5420		-243.0418	-227.7908
Number of observations	1953		2017	1968	1953		2017	1968

Note: This table contains regression estimates of duration (survival) analysis models. For the sake of sensitivity analysis, three different specifications of the model are estimated. Table contains coefficient estimates of two other specifications alongside main specification results (column 1). More specifically, in second specification (column 3), we drop political and regional variables while in third specification, we drop volatility variables. Table also presents odd ratio of first specification. *, **, ***significant at 10%, 5% and 1% level of significance. Dependent variable in each model specification is whether a country is out or into default (represented by y as explained in table 2)). The number of observations used in each model estimation differ due to missing values of certain independent variables.

Figure 1 Survivor Function (first default only)
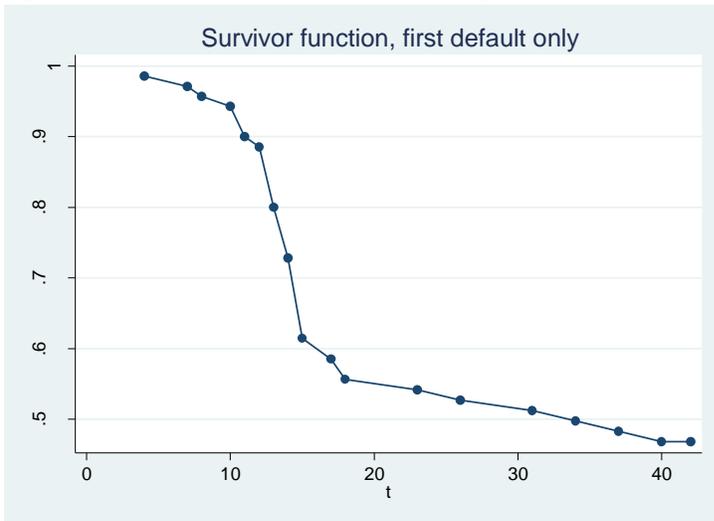
Note. Survivor function estimates (proportion of countries not defaulting on sovereign loan payment at each time period).

Figure 2 Hazard Function (first default only)
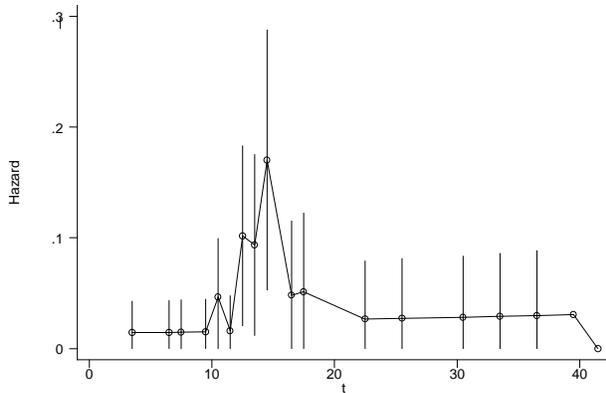
Note. Hazard function estimates. Vertical lines show the 95% confidence interval.

Figure 3 Differences in Hazard Rates Before/After t=15 (e15=1 up to 15 years and e15=1 beyond 15 years)