Forecasting Sovereign External Debt Default via Mixed Panel Logit Simulation

Kenneth Rios* Shukrit Guha*

December 21, 2018

Abstract

We implement a panel data extension of a mixed logit model in order to forecast sovereign external debt defaults from 2010 to 2017. We fit a simulated version of the model on historical data using an assortment of lagged economic variables hypothesized to signal default and include year-fixed effects to control for year-specific influences. The marginal effects of the variables are allowed to be randomized across countries in order to reflect country-level heterogeneity in variable-specific influences. The model is simulated using Monte Carlo sampling over parameterized independent normal distributions. Maximum likelihood estimation is penalized with an L2 regularization term to improve out-of-sample performance. The optimal tuning parameter is estimated using 10-fold cross-validation. The fully calibrated model is then used to predict unconditional probabilities of external debt default using the test data. The performance of the model is assessed using the receiver operating characteristic.

Keywords: sovereign default, mixed panel logit model, Monte Carlo simulation, L2 regularization, receiver operating characteristic

JEL Codes: C530

*New York University, Department of Economics

1 Introduction

Sovereign defaults are one of the main sources of international country risk. A sovereign default is defined by various sources (such as Fitch's 2017 Sovereign Rating Criteria, Altman et. al. (2011), etc.) as an inability *or* an unwillingness of the chief finance minster or central bank of a sovereign to pay back installments or interest payments on foreign debt *in-full* and *on-time*. Information on potential future default events is a much sought after resource from foreign investors. To that extent, credit rating agencies have filled some of the void in the prediction market. Credit rating agencies, such as Moody's, Fitch, and S&P, have undertaken analyses that attempt to at least suggest the possibility of sovereign default. These agencies, at the behest of sovereigns, release annual credit ratings that are meant to measure a sovereign's "capacity to meet its financial commitment".

These agencies employ a rating scheme for sovereigns that generally ranges from a AAA rating (prime "investment grade"), which signifies sovereigns with the least default risk, to a D rating (default or default-imminent), which is reserved for the riskiest sovereigns. The models used to generate these ratings typically come from simple linear models, such as ordinary least squares regression. Fitted values from these regressions are assigned ratings based on ranges of fitted values. Foreign investors use the long-term foreign ratings, which are usually updated annually, as a way to categorize the potential for countries to default in the future.

There is much consternation over these ratings, however, both in the economic literature and in practice. The actual ratings themselves may not be so helpful in determining how much more likely one country is to default compared to another since the default ratings are not linear across the ratings. In other words, a one-notch jump from AA+ to AAA is **not** the same as a one-notch jump from BB+ to BBB-. There is often bias in the ratings as agencies typically employ a "qualitatively overlay" on the empirical results and can increase or decrease a rating by up to *three notche*s, essentially based off the sheer discretion of the analysts. There is also concern that ratings are in general "too optimistic", as agencies are incentivized to release higher ratings than predicted in order to keep those countries as customers. This hypothesis is supported by work done by Giacomino (2013). Lastly, it is unlikely that a simple linear model would be able to capture the intricate relationships between variables that lead a country to default.

Fitch's 2017 Methodology Report lists all the variables that their rating model uses. These will be discussed in some more detail in Section 2.1. We attempt to incorporate as many of the variables listed in their methodology that we were able to obtain with quality data. Cantor & Packer (1996), a seminal work in the sovereign default risk literature, run a linear regression which takes in several economic variables and uses them to predict sovereign default ratings for the following year. We also include versions of those variables used by Cantor & Packer. See Table 1 for a comprehensive list of the 21 variables which we were able to obtain for our analysis.

The "predictive" power of long-term sovereign ratings can be evaluated using a receiver operating characteristic (ROC) analysis. For an example of such an analysis, see Gaillard (2014). The long-term ratings in a given year are ranked from lowest (most risky) to highest (least risky), and the distribution is mapped to sequence of zeros and ones that correspond to defaults that occurred over a certain period of time after the ratings were released, such as 10 years. We employ a similar methodology to test our model's prediction, but instead of using ratings, we use predicted probabilities generated from our panel logit model for years 2010 to 2017.

Lastly, we make a distinction between generic sovereign debt default, which includes arrears on payments to both domestic and foreign investors, and defaults exclusively on foreign (external) debt. From the perspective of foreign investors, the latter defaults are the only ones of relevance that should be used in a model with predictive capacity. We believe our non-linear model will introduce valuable flexibility that will more accurately predict external debt default than traditional sovereign ratings.

2 Theory

2.1 Background and Justification

We model external debt defaults across countries as outcomes that are a function of the economic fundamentals of a given country as well as year-specific effects that may induce default. We proxy for countries' economic fundamentals using our most complete available assortment of economic variables that characterize a country's "structural features", "macroeconomic performance", "public finances", and "external finances". These four factors are considered by Moody's *Sovereign Rating Model*, which purports to classify countries according to their estimated ability to repay debt obligations in-full and on-time. We lag these predictors as is standard in much of the country risk literature (see Davis et. al. (2011), Papi et. al. (2015), etc.). Lagging variables allows us to implement fully forward-looking prediction, thus making the model of practical interest to forecasters. Furthermore, it is often the case that sovereign defaults are either underway or "officially begin" near the start of a calender year. Since we do not have data of finer level of granularity than annual (such as quarterly data), using lags is a safeguard against defaults occurring very early in a given year which would lead to end-of-year economic figures that are not of interest to the forecaster.

The year-specific effects we include may be interpreted as modeling contagion events within financial and foreign exchange markets that are internationally linked. Specifically, we expect year-fixed effects to be a reasonable approach to control for years of global recession that affected numerous countries, such as the global financial crisis of 2008.

There is an obvious distribution of the quality and characteristics of socioeconomic institutions across countries. Countries thus vary in their degree of default susceptibility on external debt according to the various predictors. For example, some countries have historically demonstrated to be particularly sensitive to changes in their inflation rates (e.g., Venezuela), whereas other countries possess the financial stability to handle increasing levels of public debt that countries like Venezuela may find unsustainable (e.g., United States). Thus, in order to address the likely heterogeneity of marginal effects of these variables across countries, we allow the vector of coefficients $\boldsymbol{\beta}_n$ to be randomized across countries, indexed by n. Allowing $\boldsymbol{\beta}$ to be randomized across countries not only accounts for this heterogeneity of marginal effects, but it also increases the flexibility of our model and allows us to simulate for predicted *unconditional* default probabilities. Our ultimate dependent variable of interest that we wish to predict is default status (=1 for a country defaulting in a given year, =0 otherwise).

Models that incorporate randomization in addition to fixed effects are known as *mixed models*. Since our ultimate dependent variable of interest is binary, we will be employing a *mixed logit model* whose distributional assumptions will be explained in Section 2.2. Mixed logit models represent a generalization of standard logit models, such as logistic regression. Unlike standard logit models, however, mixed logit models allow for randomization of marginal effects and the correlation of unobserved influences over time. In some literature, mixed logit models are considered either equivalent to or a type of *random coefficients* model, for obvious reasons.

Since we are interested in estimation over a panel dataset in order to model temporal influences, we introduce an extension to the mixed logit model that we call a *mixed panel logit model*, in the same vein as Cherchi & Cirillo (2008). This extension does not complicate the mathematics of the model substantially as we will assume that the noise terms present in the data-generating process are independent over time. Once closed-forms for the unconditional probabilities of default are derived via Monte Carlo simulation, we introduce L2 regularization that purports to improve the performance of our estimated model with the prediction (test) data. We decide on the optimal tuning parameter by implementing k-fold cross-validation. Finally, we fit the cross-validated model on the entirety of the training set and generate probability predictions using the test data. We wish to compare our model's prediction performance with that of credit rating agencies.

2.2 Mixed Panel Logit Model of External Debt Default

For a given country n in year t, the data-generating process decides whether the country ends up defaulting on its external debt obligations as $j = 1$ or not defaulting as $j = 0$ according to

$$D_{njt} = \boldsymbol{\beta}_n^{'} \boldsymbol{x}_{njt-1} + \boldsymbol{\alpha}^{'} \boldsymbol{\gamma}_t + \varepsilon_{njt} \tag{1}$$

$$\text{s.t. } d_{njt} = \begin{cases} 1 & if D_{n1t} > D_{n0t} \\ 0 & \text{otherwise,} \end{cases}$$

where d_{njt} is a dummy for default status, $\boldsymbol{\beta}_n \in \mathbb{R}^{k\times 1}$ is a column vector of k population coefficients randomized over countries, $\boldsymbol{x}_{njt-1} \in \mathbb{R}^{k\times 1}$ is a column vector of one-year lagged economic data corresponding to k economic variables, $\boldsymbol{\alpha} \in \mathbb{R}^{\max(T_n)\times 1}$ is a column vector of the time-fixed effects to be estimated, and $\boldsymbol{\gamma}_t = \mathbf{1}(\tau = t)^{\max(T_n)\times 1}$ for each $\tau \in \{1, \ldots, \max(T_n)\}$ is a dummy indicating year t, where $\mathbf{1}(\cdot)$ is the indicator function. Note that we are allowing the number of years T_n vary according to country n, thus allowing an unbalanced panel to be estimated. Indeed, the mixed panel logit model is successful in handling unbalanced panels. However, all years in the data must be accounted for and assigned fixed effects, hence the usage of $\max(T_n)$ in the dimensions of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}_t$.

D_{njt} is a latent variable unobserved by the econometrician that is characterized by (1). ε_{njt} is a random noise term that is also unobserved by the econometrician. We make the standard assumption of independent and identically distributed (i.i.d.) errors ε_{njt} over countries

n, years t, and default/non-default events j according to a Type I extreme value distribution. Given that the errors are i.i.d. extreme value, we can ascribe a closed form to the probability of country n experiencing a given sequence of default and non-default events conditional on observing $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_n$. In order to compress notation, we characterize a sequence of such events in a vector

$$\boldsymbol{d}_n = [d_{n_1}, \ldots, d_{n_{T_n}}]',$$

where the $\boldsymbol{d}_{n_t}$'th element of $\boldsymbol{d}_n$ is a realization of j for country n in year t. That is, $\boldsymbol{d}_n$ is a sequence of 1s (defaults) and 0s (non-defaults) for a given country n across all years for which data is available. Thus, the conditional probability of country n experiencing an arbitrary sequence of events $\boldsymbol{d}_n$ is given by

$$\begin{aligned} \pi_{n\boldsymbol{d}_n}(\boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t | \boldsymbol{\alpha}, \boldsymbol{\beta}) &= \prod_{t=1}^{T_n} \frac{e^{\boldsymbol{\beta}' \boldsymbol{x}_{nd_{n_t}t-1} + \boldsymbol{\alpha}' \boldsymbol{\gamma}_t}}{\sum_{j=0}^{1} e^{\boldsymbol{\beta}' \boldsymbol{x}_{njt-1} + \boldsymbol{\alpha}' \boldsymbol{\gamma}_t}} \\ &= \prod_{t=1}^{T_n} \frac{e^{\boldsymbol{\beta}' \boldsymbol{x}_{nd_{n_t}t-1} + \boldsymbol{\alpha}' \boldsymbol{\gamma}_t}}{e^{\boldsymbol{\beta}' \boldsymbol{x}_{n0t-1} + \boldsymbol{\alpha}' \boldsymbol{\gamma}_t} + e^{\boldsymbol{\beta}' \boldsymbol{x}_{n1t-1} + \boldsymbol{\alpha}' \boldsymbol{\gamma}_t}}. \end{aligned} \tag{2}$$

This formula is similar to the well-known logit formula for i.i.d. extreme value errors, but includes the product of logits across years for a given country. Recall that we are able to multiply the logits in such a manner due to our independence assumption over ε_{njt}'s. Note that since the conditional probability is indexed by n, we have suppressed the subscript on $\boldsymbol{\beta}_n$ in (2). From here on, we denote the vector of coefficients as simply $\boldsymbol{\beta}$.

The econometrician, however, does not know $\boldsymbol{\beta}$ either. Thus, in order to obtain a probability expression which can be estimated given the data, we must integrate the probabilities conditional on $\boldsymbol{\beta}$ over its joint density $f(\boldsymbol{\beta}) \in \mathcal{C}^1$, which will generate the unconditional default probability:

$$\pi_{n\boldsymbol{d}_n} = \int_{supp(f)} \pi_{n\boldsymbol{d}_n}(\boldsymbol{\alpha}, \boldsymbol{\beta}) f(\boldsymbol{\beta}) d\boldsymbol{\beta}.$$

The econometrician must also specify $f(\boldsymbol{\beta})$, which is known in the literature as the *mixing distribution*. For sake of computational simplicity, we take $\boldsymbol{\beta}$ to be normally distributed about a vector of means $\boldsymbol{\mu} \in \mathbb{R}^{k\times 1}$ with covariance $\boldsymbol{\Sigma} \in \mathbb{R}^{k\times k}$; that is, $\boldsymbol{\beta} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Thus the unconditional probability becomes

$$\pi_{n\boldsymbol{d}_n} = \int_{supp(\Phi)} \pi_{n\boldsymbol{d}_n}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \Phi(\boldsymbol{\beta} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) d\boldsymbol{\beta}, \tag{3}$$

which is now strictly a function of the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

2.3 Simulation

We simulate the unconditional probabilities generated by (3) using Monte Carlo sampling for *each* coefficient β_i in $\boldsymbol{\beta}$. In doing so, we assume that each β_i is independently and normally distributed with a density parameterized by a mean μ_i and standard deviation σ_i; that is,

$$\beta_i \sim \mathcal{N}(\mu_i, \sigma_i).$$

We ultimately wish to estimate these parameters for all coefficients $i = 1, \ldots, k$, which we now store in vectors $\boldsymbol{\mu} \in \mathbb{R}^{k\times 1}$ and $\boldsymbol{\sigma} \in \mathbb{R}^{k\times 1}$.

Using $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$, we take R draws from each of the k normal distributions and label each draw as $\boldsymbol{\beta}^r$. To simulate the unconditional probabilities we sum over the conditional probabilities $\pi_{n\boldsymbol{d}_n}(\boldsymbol{\beta}^r|\boldsymbol{\mu}, \boldsymbol{\sigma}; \boldsymbol{\alpha})$ calculated for each draw r using (2) and then normalize by the total number of draws R. The notation $\pi_{n\boldsymbol{d}_n}(\boldsymbol{\beta}^r|\boldsymbol{\mu}, \boldsymbol{\sigma}; \boldsymbol{\alpha})$ is meant to emphasize that each conditional probability expressed in (2) remains a function of the vector of year-fixed effects $\boldsymbol{\alpha}$ that has yet to be estimated as well as the vector of randomized coefficients $\boldsymbol{\beta}$, which has now been parameterized in terms of its mean $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$ that have also yet to be estimated.

The simulated unconditional probability of default for country n over $\boldsymbol{d}_n$ is thus

$$\hat{\pi}_{n\boldsymbol{d}_n} = \frac{1}{R}\sum_{r=1}^{R} \pi_{n\boldsymbol{d}_n}(\boldsymbol{\beta}^r|\boldsymbol{\mu}, \boldsymbol{\sigma}; \boldsymbol{\alpha}). \tag{4}$$

We can now proceed to estimate $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$ via maximum likelihood estimation. We interpret the sum of logged simulated unconditional probabilities for default events across countries as the log-likelihood of $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$ being the true population parameters *conditional* on the sample data the econometrician observes. Abstracting away from training and test set considerations, given a dataset with N countries where each country n has an observed event sequence $\hat{\boldsymbol{d}}_n$, the *net simulated log-likelihood function* is

$$\hat{\boldsymbol{\mathcal{L}}}(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma}|\boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t) = \sum_{n=1}^{N} \hat{\boldsymbol{d}}_n{}' \ln\left(\hat{\boldsymbol{\pi}}_{n\boldsymbol{d}_{n_t}}\right), \tag{5}$$

where we now decompress notation slightly for clarity and allow $\hat{\boldsymbol{\pi}}_{n\boldsymbol{d}_{n_t}}$ to represent a $T_n \times 1$-vector containing simulated unconditional probabilities of default $\hat{\pi}_{n\boldsymbol{d}_{n_t}}$ for country n, for *each* year $t \in \{1, \ldots, T_n\}$. (4) of course holds analogously with this time re-indexing, and each $\pi_{n\boldsymbol{d}_{n_t}}$ is simply equal to the expression in (2), but without the product symbol.

Before we proceed with optimization of the simulated log-likelihood function over the entire training dataset, we implement supervised regularization over the parameters of interest in order to improve the out-of-sample performance of our calibrated model. Employing regularization over the parameters which characterize the distributions of the randomized coefficients $\boldsymbol{\beta}$ will lead to a decrease in the dispersion of the estimated distributions and thus reduce the variance of the estimated model in exchange for an increase in bias. λ, which is the weight placed on the regularization penalty term, will be chosen via k-fold cross validation. λ^{CV} represents the optimal λ chosen from the cross-validation selection algorithm and indicates the point of optimal trade-off between variance reduction and bias increase as measured by a loss function.

We implement L2 regularization over the parameters $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$. L2 regularization is commonly referred to as ridge regression when implemented in the context of ordinary least squares estimation. It penalizes increases in the magnitudes of the parameters according to the sum of their squares. We augment the simulated log-likelihood function in (5) and define the *penalized simulated log-likelihood function* $\hat{\mathcal{L}}_{pen}$, which is now also function of λ:

$$\hat{\mathcal{L}}_{pen}(\lambda; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma} | \boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t) = \sum_{n=1}^{N} \hat{\boldsymbol{d}}_n{}' \ln\left(\hat{\boldsymbol{\pi}}_{n\boldsymbol{d}_{n_t}}\right) - \frac{\lambda}{2}\left(\boldsymbol{\alpha}'\boldsymbol{\alpha} + \boldsymbol{\mu}'\boldsymbol{\mu} + \boldsymbol{\sigma}'\boldsymbol{\sigma}\right). \tag{6}$$

An added benefit of employing this form of regularization to our parameters of interest is that the penalty term $\frac{1}{2}\left(\boldsymbol{\alpha}'\boldsymbol{\alpha} + \boldsymbol{\mu}'\boldsymbol{\mu} + \boldsymbol{\sigma}'\boldsymbol{\sigma}\right)$ can be interpreted as resulting from a standard normal distribution prior over the parameter matrix $\boldsymbol{\Omega}$ in a Bayesian *maximum a posteriori* (MAP) estimation framework. To see this, consider $\boldsymbol{\Omega}$ to be a matrix of random variables consisting of the elements in $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$ and then implement Bayes' theorem:

$$\Pr\left(\boldsymbol{\Omega} | \boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t\right) = \frac{\Pr\left(\boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t | \boldsymbol{\Omega}\right) \Pr\left(\boldsymbol{\Omega}\right)}{\Pr\left(\boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t\right)}. \tag{7}$$

If we take our prior over the parameters in $\boldsymbol{\Omega}$ to be independent standard normal distributions, then the probability of $\boldsymbol{\Omega}$ is

$$\Pr(\boldsymbol{\Omega}) = \frac{1}{\sqrt{2\pi}} \prod_{ij} e^{-\frac{\omega_{ij}^2}{2}} \qquad \forall ij \in \dim \boldsymbol{\Omega}.$$

Since $\Pr\left(\boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t\right)$ is fixed for the observed data and we have country panels that we must sum log-likelihoods over, taking logs of (7) gives us the net MAP log-likelihood

$$\begin{aligned}
\mathcal{L}_{MAP}\left(\boldsymbol{\Omega} | \boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t\right) &= \ln \Pr\left(\boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t | \boldsymbol{\Omega}\right) + \ln \Pr(\boldsymbol{\Omega}) \\
&= \sum_{n=1}^{N} \hat{\boldsymbol{d}}_n{}' \ln \boldsymbol{\pi}_{n\boldsymbol{d}_{n_t}}(\boldsymbol{\beta}^r | \boldsymbol{\Omega}) - \frac{1}{2}\sum_{ij} \omega_{ij}^2 + \ln\left(\frac{1}{\sqrt{2\pi}}\right) \\
&= \sum_{n=1}^{N} \hat{\boldsymbol{d}}_n{}' \ln \boldsymbol{\pi}_{n\boldsymbol{d}_{n_t}}(\boldsymbol{\beta}^r | \boldsymbol{\Omega}) - \frac{1}{2}\left(\boldsymbol{\alpha}'\boldsymbol{\alpha} + \boldsymbol{\mu}'\boldsymbol{\mu} + \boldsymbol{\sigma}'\boldsymbol{\sigma}\right) + \ln\left(\frac{1}{\sqrt{2\pi}}\right),
\end{aligned}$$

which is precisely the non-simulated version of (6) with $\lambda = 1$ and an additional immaterial constant term $\ln\left(\frac{1}{\sqrt{2\pi}}\right)$, re-parameterizing $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$ as $\boldsymbol{\Omega}$.

2.4 Cross-validation Selection of L2 Regularization Parameter

We now describe the cross-validation selection algorithm that chooses λ^{CV}. We begin by randomly partitioning our entire set of training data X into K disjoint folds of (roughly) equal size indexed by k such that $\cup_{k=1}^{K} X_k = X$ and $\cap_{k=1}^{K} X_k = \varnothing$. Then we decide on a range, or *grid*, of candidate λ's to iterate over. We thus pull each candidate $\lambda_i \in \{0, 1, 2, \ldots, L-1, L\}$, where $\lambda_i \in \mathbb{Z}_{\geq 0}$ and L is sufficiently large.

For *each* λ_i, we run the following procedure:

1. For each $k = 1, \ldots, K$, we estimate the parameters $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$ by maximizing the penalized simulated log-likelihood function in (6), but using only the training data in the remaining $K-1$ folds. The penalized simulated log-likelihood function is thus

$$\hat{\boldsymbol{\mathcal{L}}}_{pen}(\lambda_i; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma} | \boldsymbol{x}_{njt-1}, \gamma_t) = \sum_{n \in X \setminus X_k} \hat{\boldsymbol{d}}_n{}' \ln\left(\hat{\boldsymbol{\pi}}_{n\boldsymbol{d}_{n_t}}\right) - \frac{\lambda_i}{2}\left(\boldsymbol{\alpha}'\boldsymbol{\alpha} + \boldsymbol{\mu}'\boldsymbol{\mu} + \boldsymbol{\sigma}'\boldsymbol{\sigma}\right), \quad (8)$$

   where the summation index now acknowledges that the $K-1$ folds may contain only a proper subset of the N countries in the training data.

   Functions with the form of $\hat{\mathcal{L}}_{pen}$ are typically optimized numerically. We maximize $\hat{\mathcal{L}}_{pen}$ with respect to $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$ using the Nelder-Mead algorithm, with the assistance of Python's *pylogit* package. We defer a description of the Nelder-Mead algorithm to the Data Appendix. Using (8), the optimal $\hat{\boldsymbol{\alpha}}^k, \hat{\boldsymbol{\mu}}^k, \hat{\boldsymbol{\sigma}}^k$ for fold k are

$$\{\hat{\boldsymbol{\alpha}}^k, \hat{\boldsymbol{\mu}}^k, \hat{\boldsymbol{\sigma}}^k\} = \underset{\{\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma}\}}{\arg\max}\ \hat{\boldsymbol{\mathcal{L}}}_{pen}(\lambda_i; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma} | \boldsymbol{x}_{njt-1}, \gamma_t).$$

2. We take $\hat{\boldsymbol{\alpha}}^k, \hat{\boldsymbol{\mu}}^k, \hat{\boldsymbol{\sigma}}^k$, which were estimated using the data $X \setminus X_k$, and now back out the predicted probabilities using the data in fold k which was excluded in Step 1. In order to do so, we must re-simulate the unconditional probabilities of default as in (4), but this time by drawing values for the coefficients from normal distributions with means and standard deviations that have been estimated as $\hat{\boldsymbol{\mu}}^k$ and $\hat{\boldsymbol{\sigma}}^k$; that is,

$$\hat{\beta}_i^k \sim \mathcal{N}(\hat{\mu}_i^k, \hat{\sigma}_i^k),$$

   for $i = 1, \ldots, k$, where this k now refers to the total numbers of coefficients and **not** the index for the current fold.

   As in our initial simulation, we take R draws from each normal distribution and label the r'th draw of coefficients as $\hat{\boldsymbol{\beta}}^{k,r}$. The simulated unconditional probability is again the sum of conditional probabilities $\pi_{n\boldsymbol{d}_n}^k(\hat{\boldsymbol{\beta}}^{k,r} | \hat{\boldsymbol{\mu}}^k, \hat{\boldsymbol{\sigma}}^k; \hat{\boldsymbol{\alpha}}^k)$ evaluated for each draw r using data X_k, normalized by R:

$$\hat{\pi}_{n\boldsymbol{d}_n}^k = \frac{1}{R}\sum_{r=1}^{R} \pi_{n\boldsymbol{d}_n}^k(\hat{\boldsymbol{\beta}}^{k,r} | \hat{\boldsymbol{\mu}}^k, \hat{\boldsymbol{\sigma}}^k; \hat{\boldsymbol{\alpha}}^k). \quad (9)$$

   Re-indexing the probabilities in (9) by t and storing each unconditional probability of default for country n in year t present in X_k, we obtain the vector of probabilities $\hat{\boldsymbol{\pi}}_{n\boldsymbol{d}_{n_t}}^k$ that is present in the penalized simulated log-likelihood function.

3. We evaluate the fitted model's performance on the validation set X_k in order to estimate the cross-validation error for fold k. We specify the loss function to be the *negative* penalized likelihood, which is basically the negation of (8) evaluated at $\hat{\boldsymbol{\alpha}}^k, \hat{\boldsymbol{\mu}}^k, \hat{\boldsymbol{\sigma}}^k$, but instead only summed over the countries n that happen to be present in X_k:

$$NLL^k(\lambda_i) = -\left(\sum_{n \in X_k} \hat{\boldsymbol{d}}_n{}' \ln\left(\hat{\boldsymbol{\pi}}^k_{nd_{n_t}}\right) - \frac{\lambda_i}{2}\left(\hat{\boldsymbol{\alpha}}^{k'}\hat{\boldsymbol{\alpha}}^k + \hat{\boldsymbol{\mu}}^{k'}\hat{\boldsymbol{\mu}}^k + \hat{\boldsymbol{\sigma}}^{k'}\hat{\boldsymbol{\sigma}}^k\right)\right)$$

4. Steps 1, 2, and 3 are iterated over each fold k. In order to calculate the model's overall cross-validation error given λ_i, we average the cross-validation errors $NLL^k(\lambda_i)$ for each fold k, which is

$$NLL(\lambda_i) = \frac{1}{K}\sum_{k=1}^{K} NLL^k(\lambda_i).$$

The negative log-likelihoods $NLL(\lambda_i)$ generated by this procedure are stored. The chosen λ^{CV} is the λ_i with the **smallest** corresponding negative log-likelihood.

2.5 Prediction of Unconditional Default Probabilities

After cross-validation, we are now able to predict unconditional probabilities of default in our test dataset Z, which represents data that has occurred after the estimation procedure; that is, data spanning some number of years $t > \max(T_n)$. We begin by fitting (8) on the *entire* training dataset X using the optimal $\lambda_i = \lambda^{CV}$ and estimating $\boldsymbol{\alpha}, \boldsymbol{\mu}$, and $\boldsymbol{\sigma}$ using Nelder-Mead:

$$\{\boldsymbol{\alpha}^*, \boldsymbol{\mu}^*, \boldsymbol{\sigma}^*\} = \underset{\{\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\sigma}\}}{\arg\max}\ \hat{\boldsymbol{\mathcal{L}}}_{pen}(\lambda^{CV}; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\sigma} | \boldsymbol{x}_{njt-1}, \boldsymbol{\gamma}_t).$$

Predicted unconditional probabilities for country n are thus

$$\pi^*_{n\boldsymbol{d}_n} = \frac{1}{R}\sum_{r=1}^{R} \pi_{n\boldsymbol{d}_n}(\boldsymbol{\beta}^{*,r} | \boldsymbol{\mu}^*, \boldsymbol{\sigma}^*; \boldsymbol{\alpha}^*). \tag{10}$$

3 Empirical Analysis

3.1 Data

In order to estimate the model described in Section 2, we use data on external debt defaults from Reinhart & Rogoff's publicly available online dataset which supplements their 2009 book *This Time is Different: Eight Centuries of Financial Folly.* Since this dataset only covers external debt defaults until 2012, we augment this database with Table 1 of Laeven & Valencia (2018), which tabulates different types of economic crises, including sovereign defaults, across countries up until 2017.

The economic data for various countries were obtained from World Bank and the Economist Intelligence Unit. We could not find key data for some of the more industrialized countries and thus chose to omit them from our sample. The list of macroeconomic variables we used in our sample and their descriptions are displayed in Table 1. We combined data from both these sources to develop a larger and more complete sample of observations. For some of the variables such as inflation, unemployment, external balance on goods and services, real effective exchange rate, change in net exports, interest payments on government debt, and percentage change in bank lending to public and private sectors, data was unavailable for specific country-year pairs. These values were imputed from other sources that include Trading Economics, EconomyWatch, UNDP, OECD and countryecononomy.com. For a detailed explanation of these imputations, please see the Data Appendix.

Since we are interested in the applicability of our model to perform forward-looking prediction, we lag the predictors and use time as the cutoff criteria for dividing our entire dataset into a training set and a test set. Specifically, we bin pre-2010 observations into our training set and post-2009 observations into our test set. 2010 represents a natural temporal divide for our data as default histories are relatively quiet during the 2000s but jump in the early 2010s. We would like to measure how well our model captures this spike in default activity using historical data while controlling for year-specific influences. Partitioning the data in such a manner generates 707 observations in the training set, spanning 78 countries over 18 years. The corresponding test set contains 525 observations spanning 78 countries and 8 years.

Incorporating dummies for each year in the training set resulted in a model whose Nelder-Mead optimization did not converge, which was not surprising given the number of years and the complexity of the log-likelihood function. We thus decided to incorporate fixed effects only for years of US recession or international crisis, as those years were likely to have the greatest influence on default rates. Since our training data contains years from 1992 to 2009, we thus included dummy variables for the years 2001, 2007, 2008, and 2009. These are the years whose fixed effects are reported in Table 2.

3.2 Cross-validation Outcome

We refer to Section 2.4 for a discussion on how we performed k-fold cross-validation in order to estimate the optimal λ^{CV}. In our empirical analysis, we let $k = 10$ and performed 10-fold cross-validation while setting $R = 1000$, thus simulating 1000 draws from parameterized independent normal distributions. Our training set contains 525 observations, and thus the validation set contains $\frac{1}{10}(525) \approx 52$ observations and the bag of 9 folds contains $\frac{9}{10}(525) \approx$ 473 observations.

We use an iterative approach to find an appropriate range of λ given our data. In practice, we found that a range from $\lambda = 0$ to $\lambda = 100$ struck a balance between returning a stable number of local minima as well as being computationally feasible. On an i7 Core processor running at maximum clock speed, we found that the cross-validation procedure outlined in Section 2.4 took on-average 5 minutes to complete for values of λ near zero and over an hour to complete for values of λ near 100.

The optimal λ^{CV} was estimated to be **88** from 10-fold cross-validation. $\lambda = 88$ was thus used to fully calibrate the model and generate predictions using the test data.

Table 1: Description of Predictors

Regressor	Definition	Source
Net foreign assets (current LCU)	Net foreign assets are the sum of foreign assets held by monetary authorities and deposit money banks, less their foreign liabilities. Data are in current local currency	International Monetary Fund, International Financial Statistics
Inflation, consumer prices (annual %)	Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly. The Laspeyres formula is generally used	International Monetary Fund, International Financial Statistics
External balance on goods and services (% of GDP)	External balance on goods and services (formerly resource balance) equals exports of goods and services minus imports of goods and services (previously nonfactor services)	World Bank national accounts data, and OECD National Accounts data files
Current account balance (BoP, current US$)	Current account balance is the sum of net exports of goods and services, net primary income, and net secondary income. Data are in current U.S. dollars	International Monetary Fund, Balance of Payments Statistics Yearbook and data files
Net trade in goods and services (BoP, current US$)	Net trade in goods and services is derived by offsetting imports of goods and services against exports of goods and services. Exports and imports of goods and services comprise all transactions involving a change of ownership of goods and services between residents of one country and the rest of the world. Data are in current U.S. dollars	International Monetary Fund, Balance of Payments Statistics Yearbook and data files
Unemployment, total (% of total labor force)	Unemployment refers to the share of the labor force that is without work but available for and seeking employment	International Labour Organization, ILOSTAT database. Data retrieved in September 2018
XRRE	Trade-weighted basket of currencies converted to an index (1997=100) and adjusted for relative price movements	The Economist Intelligence Unit
DGDP	Percentage change in real GDP, over previous year	The Economist Intelligence Unit
RYPC	Percentage change in real gross domestic product per head	The Economist Intelligence Unit
CGEB	Change in net exports, as a percentage of real GDP in the previous period	The Economist Intelligence Unit
PSBR	Central government receipts minus central government outlays, as a percentage of GDP	The Economist Intelligence Unit
BINT	Interest payments on government debt (both domestic currency denominated and foreign currency debt), as a percentage of GDP	The Economist Intelligence Unit
PUDP	Total debt (both local and foreign currency) owed by government to domestic residents, foreign nationals and multilateral institutions such as the IMF, expressed as a percentage of GDP	The Economist Intelligence Unit
SODD	Percentage change in bank lending to public and private sectors, plus bank lending in domestic currency overseas	The Economist Intelligence Unit
CARA	Current-account balance as a percentage of nominal GDP	The Economist Intelligence Unit
IRTD	Total international reserves as a percentage of total external debt stock	The Economist Intelligence Unit
TDPX	Total external debt stock as a percentage of exports of goods, non-factor services, primary income, and workers remittances	The Economist Intelligence Unit
TDPY	Total external debt at end-period as a percentage of nominal GDP	The Economist Intelligence Unit
TSPY	Total external debt service paid as a percentage of nominal GDP	The Economist Intelligence Unit
INPS	Total interest payments made on total external debt as a percentage of total debt service paid	The Economist Intelligence Unit
INPY	Total interest payments made on total external debt as a percentage of nominal gross domestic product	The Economist Intelligence Unit

3.3 Mixed Panel Logit Estimation Output

Table 2 on the following page reports the estimated year-fixed effects for years 2001, 2007, 2008, and 2009; as well as the estimated means and standard deviations of the randomized coefficients for our lagged economic variables of interest.

Table 2: Mixed Logit Model Regression Output

	Dependent Variable:	default_RR	No. of observations:	18
	Model:	Mixed Logit Model	Df Residuals:	-28
	Method:	MLE	Df Model:	46
	Date:	Wed, 19th Dec 2018	Pseudo R-squared:	0.090
	Time:	22:08:20	Pseudo R-bar-squared:	-0.671
	AIC:	201.877	Log-Likelihood:	-54.939
	BIC:	242.834	LL-Null:	-60.391

Lagged Predictors	Coefficient	Std. Error	z	$P > \|z\|$
Year 2001	2.289e-06	2.804	8.16e-07	1.000
Year 2007	2.354e-07	17.639	1.33e-08	1.000
Year 2008	2.042e-05	11.919	1.71e-06	1.000
Year 2009	-8.875e-07	18.686	-4.75e-08	1.000
Net Foreign Assets (current LCU)	-3.859e-06***	6.54e-12	-5.9e+05	0.000
Inflation (annual % change in CPI)	-2.659e-07	0.084	-3.15e-06	1.000
External Trade Balance (% of GDP)	-5.124e-07	0.184	-2.78e-06	1.000
Current Account Balance (current USD)	4.531e-06***	3.11e-10	1.46e+04	0.000
Balance of Payments (current USD)	-4.975e-06***	3.37e-10	-1.48e+04	0.000
Unemployment (% of labor force)	1.186e-05	0.161	7.37e-05	1.000
Real GDP Growth (YOY % change)	7.266e-07	0.783	9.28e-07	1.000
Real GDP per capita growth (annual % change)	7.505e-07	0.757	9.91e-07	1.000
Change in Net Exports (% of GDP)	9.249e-07	0.367	2.52e-06	1.000
Central Government Balance (% of GDP)	7.166e-06	0.276	2.59e-05	1.000
Interest Payments on Government Debt (% of GDP)	-4.417e-06	0.044	-0.000	1.000
Total Debt owed by Government to Domestic and Foreign Creditors (% of GDP)	1.283e-05	0.043	0.000	1.000
Bank Lending to Public and Private Sectors (annual % change)	-1.55e-06	0.072	-2.15e-05	1.000
Current Account Balance (% of GDP)	-8.206e-07	0.264	-3.11e-06	1.000
Total International Reserves (% of external debt stock)	1.01e-06	0.041	2.47e-05	1.000
Total External Debt Stock (% of exports)	2.806e-06	0.015	0.000	1.000
Total External Debt (% of GDP)	2.535e-06	0.058	4.4e-05	1.000
Total External Debt Service Paid (% of GDP)	7.057e-06	0.722	9.78e-06	1.000
Total Interest Payments made on External Debt (% of total debt service paid)	-1.084e-07	0.093	-1.17e-06	1.000
Total Interest Payments made on External Debt (% of GDP)	1.392e-05	2.277	6.11e-06	1.000
Real Effective Exchange Rate (weighted by trade)	7.131e-06	0.032	0.000	1.000
Sigma Net Foreign Assets (current LCU)	-3.03e-06***	3.72e-12	-8.14e+05	0.000
Sigma Inflation (annual % change in CPI)	-6.995e-07	1.186	-5.9e-07	1.000
Sigma External Trade Balance (% of GDP)	-6.707e-06	1.455	-4.61e-06	1.000
Sigma Current Account Balance (current USD)	0.0001***	1.49e-10	9.87e+05	0.000
Sigma Balance of Payments (current USD)	-2.183e-05***	8.44e-10	-2.59e+04	0.000
Sigma Unemployment (% of labor force)	3.436e-06	2.436	1.41e-06	1.000
Sigma Real GDP Growth (YOY % change)	3.522e-06	3.848	9.15e-07	1.000
Sigma Real GDP per capita growth (annual % change)	-1.482e-06	5.738	-2.58e-07	1.000
Sigma Change in Net Exports (% of GDP)	8.241e-06	5.973	1.38e-06	1.000
Sigma Central Government Balance (% of GDP)	3.588e-06	6.230	5.76e-07	1.000
Sigma Interest Payments on Government Debt (% of GDP)	-9.455e-07	2.017	-4.69e-07	1.000
Sigma Total Debt owed by Government to Domestic and Foreign Creditors (% of GDP)	-2.845e-06	0.389	-7.32e-06	1.000
Sigma Bank Lending to Public and Private Sectors (annual % change)	1.028e-05	0.845	1.22e-05	1.000
Sigma Current Account Balance (% of GDP)	1.408e-05	3.113	4.52e-06	1.000
Sigma Total International Reserves (% of external debt stock)	8.047e-06	0.413	1.95e-05	1.000
Sigma Total External Debt Stock (% of exports)	-1.07e-06	0.067	-1.61e-05	1.000
Sigma Total External Debt (% of GDP)	-4.389e-06	0.277	-1.59e-05	1.000
Sigma Total External Debt Service Paid (% of GDP)	3.828e-06	2.661	1.44e-06	1.000
Sigma Total Interest Payments made on External Debt (% of total debt service paid)	3.461e-08	0.394	8.79e-08	1.000
Sigma Total Interest Payments made on External Debt (% of GDP)	3.751e-06	5.601	6.7e-07	1.000
Sigma Real Effective Exchange Rate (weighted by trade)	2.036e-06	0.197	1.03e-05	1.000

3.4 Predicted Probabilities and Out-of-Sample Performance

Table 3 lists the predicted probabilities of default for the top-30 likeliest default events in the test data, defined over country-year pairs. Three events of external debt default are present in the top-30 events out of 525 total. Note that we have six instances of default post-2009 in our test set, and only 21 overall. For comparison, there were a total of 131 external debt defaults that have occurred since 1980, with 9 occurring since 2010. Our very small subset of defaults in our training set is indicative of the great difficulty we had in finding quality data for those years. Despite these data limitations, our model performs well out-of-sample.

Table 3: Predicted Probabilities of Default for Top-30 Likeliest Default Events

	Country	Year	Default?	Predicted Probability
1	Mongolia	2017	No	0.0749
2	Sri Lanka	2017	No	0.0745
3	**Argentina**	**2013**	**Yes**	**0.0639**
4	Mongolia	2016	No	0.0629
5	Sri Lanka	2016	No	0.0628
6	Mongolia	2015	No	0.0623
7	Togo	2016	No	0.0585
8	Sri Lanka	2014	No	0.0582
9	Sri Lanka	2015	No	0.0503
10	New Zealand	2010	No	0.0478
11	Australia	2017	No	0.0469
12	New Zealand	2011	No	0.0437
13	El Salvador	2017	No	0.0413
14	Nicaragua	2010	No	0.0409
15	**Ukraine**	**2015**	**Yes**	**0.0402**
16	Jordan	2017	No	0.0398
17	Tunisia	2017	No	0.0397
18	Turkey	2017	No	0.0394
19	Australia	2015	No	0.0390
20	Australia	2014	No	0.0389
21	**Belize**	**2012**	**Yes**	**0.0387**
22	Australia	2016	No	0.0377
23	Panama	2017	No	0.0372
24	Malawi	2010	No	0.0367
25	Australia	2010	No	0.0366
26	Sudan	2016	No	0.0361
27	El Salvador	2014	No	0.0353
28	Jordan	2014	No	0.0352
29	Romania	2012	No	0.0351
30	Australia	2013	No	0.0350

We capture the overall predictive capability of our model using an ROC analysis. Figure 1 plots the true positive rates against the false positive rates. The kinks in the ROC curve correspond to actual default events observed in the test data.

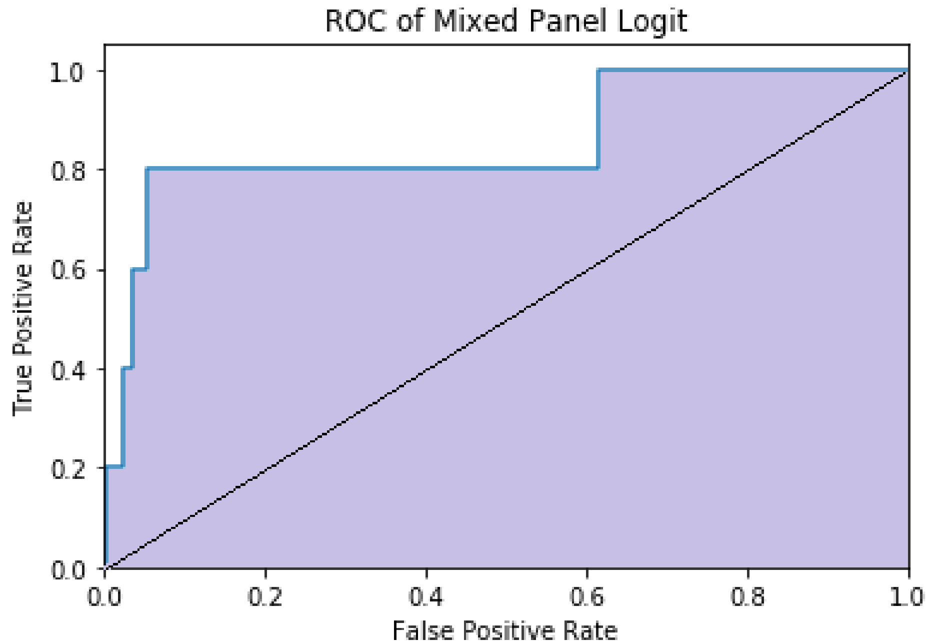

Figure 1: ROC Curve

We are able to calculate the area under the ROC curve, known as the $AUROC$ statistic, or "AUC" for short, using numerical methods implemented by the *sklearn* package. We calculate the AUC statistic of the model's ROC curve to be 0.85436. Using the AUC statistic, we can calculate the accuracy ratio AR, which is a measure of the fit of the predicted probabilities to the default histories in the test set. The AR is the proportion of the area above the 45-degree line, which corresponds a null model of 50% accuracy, that is consumed by the ROC curve. In other words,

$$\begin{aligned} AR &= \frac{\left(AUC - \frac{1}{2}\right)}{\frac{1}{2}} \\ &= 2\,(0.85346 - .5) \\ &\approx 0.707. \end{aligned}$$

Our model's predictions thus have an accuracy ratio of roughly 71%.

4 Concluding Remarks

Our results demonstrate that despite having data for only a fraction of all external debt defaults, our mixed logit model possessed sufficient flexibility to capture a sizable amount of the variation in default histories across countries. We believe that if we had data corresponding to all 131 default events since 1980, the calculated accuracy ratio would have improved significantly from 71%. We encountered that limitations on data availability was the major choke in improving model performance. Given the computationally-intensive nature of the simulation, however, we believe that generalizing our code to a model with an order of magnitude more variables would **not** be feasible given typical computing specifications.

The mixed panel logit regression output in Table 2 finds that a country's lagged net foreign assets, current account balance, and balance of payments are significant predictors for external debt default in the current year. When we observe the top-30 events listed in Table 3, we note that the model often reports countries that have traditionally been considered as developing or poor by international financial institutions, such as Mongolia, Sri Lanka, and Jordan.

We would have liked to compare our model's out-of-sample performance with that of long-term sovereign ratings issues by credit rating agencies in late 2009, using an ROC analysis similar to that used in section 3.4. We did not, however, have enough space to fit that analysis in this paper. Another avenue of further research we wanted to take was to use estimates of investment loss from default, known as "haircuts", and combine them with our predicted probabilities to generate an investment "risk score" for countries that would reflect the expected loss of investment.

Data Appendix

Code and data is available at `https://github.com/kenrios1993/mixed_panel_logit_default_risk`.

We used the Python package *pylogit* extensively for both Monte Carlo simulation and Nelder-Mead optimization. Source files and documentation for *pylogit* are available at `https://github.com/timothyb0912/pylogit`.

Data Inputs from Non-World Bank or EIU Sources

Regressor	**Countries**	**Alternative**	**Alternative Definition**	**Source**
Inflation, consumer prices (annual %)	Argentina, Venezuela	DCPN	Percentage change in consumer price index (end-period), over previous year	EIU
External balance on goods and services (% of GDP)	Trinidad & Tobago	-	External balance on goods and services (% of GDP)	Trading Economics
Unemployment, total (% of total labor force)	Seychelles, Dominica	-	Unemployment Rate (% of labor force)	Economy Watch
XRRE	Liberia*, Dominica**, Solomon Islands**	Real effective exchange rate index	*Real Effective Exchange Rate as Based on Consumer Price Index for Liberia, Index, Annual, Not Seasonally Adjusted; **Real effective exchange rate is the nominal effective exchange rate (a measure of the value of a currency against a weighted average of several foreign currencies) divided by a price deflator or index of costs	*FRED, Federal Reserve Bank of St. Louis; **International Monetary Fund, International Financial Statistics
CGEB	Iraq, Solomon Islands	Net trade in goods and services	Percentage change in Net trade in goods and services over previous year, **not as % of real GDP**. This is an imputation cost in estimation	International Monetary Fund, Balance of Payments Statistics Yearbook and data files
BINT	33 countries	Interest payments (% of revenue)	Interest payments include interest payments on government debt–including long-term bonds, long-term loans, and other debt instruments–to domestic and foreign residents	International Monetary Fund, Government Finance Statistics Yearbook and data files; Trading Economics (Cameroon)
BINT	Sudan*, Gabon**	*Interest payments on external debt (% of exports of goods, services and primary income); **Interest Payments (% of GDP)	*Total interest payments to exports of goods, services and primary income. Total interest payment is the sum of interest actually paid in currency, goods, or services on long-term debt, interest paid on short-term debt, and charges to the IMF; **Interest payments of Total Debt (as % of GDP)	*World Bank; **OECD Statistics
PUDP	Dominica*, Madagascar*, Solomon Islands*; Iraq**, Gambia**	Central government Debt (% of GDP)	Debt is the entire stock of direct government fixed-term contractual obligations to others outstanding on a particular date. It includes domestic and foreign liabilities such as currency and money deposits, securities other than shares, and loans. It is the gross amount of government liabilities reduced by the amount of equity and financial derivatives held by the government	*International Monetary Fund, Government Finance Statistics Yearbook and data files, and World Bank and OECD GDP estimates; **`countryeconomy.com`
SODD	Indonesia	Domestic credit to private sector by banks (% of GDP)	Domestic credit to private sector by banks refers to financial resources provided to the private sector by other depository corporations (deposit taking corporations except central banks), such as through loans, purchases of nonequity securities, and trade credits and other accounts receivable, that establish a claim for repayment. For some countries these claims include credit to public enterprises	International Monetary Fund, Government Finance Statistics Yearbook and data files

Nelder-Mead Algorithm

For a given function $f : \mathbb{R}^n \to \mathbb{R}$ that we wish to minimize, let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1}$ be the vertices of the convex hull of f. The Nelder-Mead algorithm iteratively produces a sequence of convex hulls used to estimate the minimum x of $f(x)$ according to the values of f evaluated at each of the vertices.

1. **Sorting**
   Evaluate f at each vertex such that
   $$f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \leq \ldots f(\mathbf{x}_n).$$

2. **Reflection**
   Compute
   $$\mathbf{x}_r = \bar{\mathbf{x}} + \alpha(\bar{\mathbf{x}} - \mathbf{x}_{n+1}).$$
   Evaluate $f_r = f(\mathbf{x}_r)$. If $f_1 \leq f_r < f_n$, replace $\mathbf{x}_{n+1}$ with $\mathbf{x}_r$.

3. **Expansion**
   If $f_r < f_1$, compute
   $$\mathbf{x}_e = \bar{\mathbf{x}} + \beta(\mathbf{x}_r - \bar{\mathbf{x}}).$$
   Evaluate $f_e = f(\mathbf{x}_e)$. If $f_e < f_r$, replace $\mathbf{x}_{n+1}$ with $\mathbf{x}_e$, otherwise replace with $\mathbf{x}_r$.

4. **Outside Contraction**
   If $f_n \leq f_r < f_{n+1}$, compute
   $$\mathbf{x}_{oc} = \bar{\mathbf{x}} + \gamma(\mathbf{x}_r - \bar{\mathbf{x}}).$$
   Evaluate $f_{oc} = f(\mathbf{x}_{oc})$. If $f_{oc} \leq f_r$, replace $\mathbf{x}_{n+1}$ with $\mathbf{x}_{oc}$, otherwise shrink according to Step 6.

5. **Inside Contraction**
   If $f_r > f_{n+1}$, compute
   $$\mathbf{x}_{ic} = \bar{\mathbf{x}} - \gamma(\mathbf{x}_r - \bar{\mathbf{x}}).$$

Evaluate $f_{ic} = f(\mathbf{x}_{ic})$. If $f_{ic} < f_{n+1}$, replace $\mathbf{x}_{n+1}$ with $\mathbf{x}_{ic}$, otherwise shrink according to Step 6.

6. **Shrinkage**
   While $2 \leq i \leq n+1$, set

$$\mathbf{x}_i = \mathbf{x}_1 + \delta(\mathbf{x}_i - \mathbf{x}_1).$$

Source: Gao & Han (2012)

References

Altman, I, Edward and Rijken, Herbert A. "Toward A Bottom-Up Approach to Assessing Sovereign Default Risk: An Update", *Journal of Banking & Finance*, 2004, vol. 28, issue 11, 2679-2714.

Brathwaite, Timothy, and Joan Walker. "Asymmetric, Closed-Form, Finite-Parameter Models of Multinomial Choice." arXiv preprint arXiv:1606.05900 (2016). http://arxiv.org/abs/1606.05900.

Cantor, Richard and Packer, Frank, (1996). "Determinants and impact of sovereign credit ratings", *Economic Policy Review*, issue Oct, p. 37-53.

Cherchi, Elisabetta and Cirillo, Cinzia. (2014). "Understanding variability, habit and the effect of long period activity plan in modal choices: a day to day, week to week analysis on panel data." *Transportation*, 41. 1245-1262. 10.1007/s11116-014-9549-y.

Davis, E, Karim, Dilruba and Liadze, Iana, (2011), "Should multivariate early warning systems for banking crises pool across regions?", *Review of World Economics* (Weltwirtschaftliches Archiv), 147, issue 4, p. 693-716.

Fitch 2017 Rating Criteria. *Fitch Ratings*. July 2017.

Gaillard, Norbert. "What is the Value of Sovereign Ratings?" Volume 15, Issue 1, Special Issue: *Special Issue on Government Debt in Democracies: Causes, Effects, and Limits.* February 2014, p. 208-224

Gao, F. and Han, L, "Implementing the Nelder-Mead simplex algorithm with adaptive parameters", *Comput Optim Appl* (2012) 51: 259. https://doi.org/10.1007/s10589-010-9329-3.

Giacomino, Paolo, (2013), "Are Sovereign Credit Ratings Pro-Cyclical? A Controversial Issue Revisited in Light of the Current Financial Crisis", *Rivista di Politica Economica*, issue 4, p. 79-111.

Laeven, Luc and Valencia, Fabian. (2018). "Systemic Banking Crises Revisited", *IMF Working Papers*. 18. 1. 10.5089/9781484376379.001.

Papi, Luca and Presbitero, Andrea Filippo and Zazzaro, Alberto, "IMF Lending and Banking Crises" (Jan 2015). *IMF Working Papers* No. 15/19.

Reinhart, Carmen M. and Kenneth S. Rogoff. "This Time is Different: Eight Centuries of Financial Folly." Figure P.1 (Princeton: Princeton University Press, 2009).

StataCorp. 2015. Stata 14 Base Reference Manual. College Station, TX: Stata Press.