Model-Based Estimation of Sovereign Default Risk

Inci Gumus,† Munechika Katayama,‡ and Junko Koeda§*

January 2017

First version: October 2017

Abstract

We estimate a canonical sovereign debt crisis model from Arellano (2008) for Argentina via maximum simulated likelihood estimation. Despite its focus on idiosyncratic risk, the estimated model accounts for the overall default patterns of Argentina. The model-implied business cycle properties are consistent with Arellano’s findings, with some caveats. Our novel real-time default probability measure, which exploits model nonlinearity, performs better than a logit model in predicting the timing of default events.

JEL Classification: C13, E43, F34, O11, O19
Keywords: sovereign debt, default risk, maximum simulated likelihood estimation

† Assistant Professor, Faculty of Arts and Social Sciences, Sabanci University, Orhanli, Tuzla, Istanbul, 34956, Turkey. E-mail: incigumus@sabanciuniv.edu. Tel: +90-216-4839328. Fax: +90-216-4839250.
‡ Associate Professor, School of Political Science and Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169-8050, Japan. E-mail: mkatayama@aoni.waseda.jp. Tel: +81-3-5286-1224.
§ Corresponding author. Associate Professor, School of Political Science and Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169-8050, Japan. E-mail: jkoeda@waseda.jp. Tel: +81-3-3208-0752. Fax: +81-3-3203-9816.
* Acknowledgment. We thank Takashi Kano for his valuable and constructive inputs. We thank Junichi Fujimoto and seminar participants in the annual meeting of 2017 Japan Economic Association, the National Graduate Institute for Policy Studies, and the Center for Positive Political Economy’s macroeconomics-finance workshop at Waseda University for helpful comments. The research reported here uses the CPPE server at Waseda University and is supported by grants-in-aid from the Ministry of Education, Culture, Sports, Science, and Technology of the Japanese government (grant number 26870124).
1 Introduction

How informative are default risks estimated from a stochastic general equilibrium sovereign debt crisis model? There is an extensive theoretical literature on sovereign debt crises that builds on the endogenous sovereign default model of Eaton and Gersovitz (1981). However, there is little empirical work on how well existing sovereign debt crisis models explain actual default events.

To examine this question, we formally estimate the Arellano (2008) model for Argentina. We choose the Arellano model as our baseline model because it is the most basic stochastic general equilibrium sovereign debt crisis model. Although the model is parsimonious, Arellano (2008) is the first paper to quantitatively analyze the behavior of sovereign default and interest rates in relation to business cycles.

This model has a discrete default choice that faces two types of uncertainty: uncertainty with respect to the debtor country’s output and that with respect to the timing of regaining market access once the debtor country defaults. The nonlinearity of the policy function makes it impossible to derive an analytical (conditional) probability distribution of default events. We thus use a maximum simulated likelihood method to estimate the model. Our estimation uses only output and default data, allowing measurement (forecast) error in the observed default variable.

We find that the model-implied default decisions account for the overall default pattern in Argentina, especially the timing of default event occurrence in 1982 and 2001. Further, the model-implied default probability increases prior to the observed default events. This probability does not necessarily increase as output falls due to the non-monotonicity of default risk and output in the model. The difference between the model-implied default decision and default probability is the former is conditional on the current output whereas the latter is conditional on the previous period’s output. Despite use of only output and default data in our
estimation, the model-implied business cycle statistics with benchmark estimates are consistent with such Arellano (2008) findings as higher volatility in consumption relative to output, and a countercyclicality of interest rate spread.

We also provide a novel real-time default probability measure that can predict actual default events. We show that this measure can be derived via the likelihood function derivation. The measure better matches with the timing of default event occurrence than a logit-based measure. Further, it tends to be stable under the observed repayment years thanks to unbinding the endogenous debt ceiling below which the country chooses to repay.

The existing theoretical literature tends to focus on defaults of an idiosyncratic nature (Kaminsky and Vega-Garcia, 2016), occurring due to country-specific shocks. To address the possible role of systemic risk, we extend the Arellano (2008) model with a stochastic risk-free interest rate, assuming that the rate follows the AR(1) process. It turned out that the model-implied default profiles look similar to baseline with slight improvement in explaining the heightened default risk prior to observed default events.

There is a wide theoretical literature on sovereign default that extends Arellano (2008), and that addresses various aspects of sovereign default. Aguiar and Gopinath (2006) point out that the sovereign debt model of Arellano (2008) cannot match the countercyclicality of interest rates, the positive correlation of interest rates, and the trade balance without an asymmetric output cost for a country in default. Without such a cost, the probability of default, the volatilities of interest rate and trade balance, and the maximum spread that the model generates decrease considerably.1 Chatterjee and Eyigungor (2012) extend the Arellano (2008) model with long-term bonds. Using long-term debt significantly improves the model’s ability to match the average debt-to-output ratio observed in the data, while also matching the debt-service-to-output

---

1 Aguiar and Gopinath (2006) then show that using a productivity process characterized by a stochastic trend improves the model’s predictions in all of these dimensions. With shocks to trend, the model generates a countercyclical trade balance and interest rate, and matches the positive correlation between the two, albeit with smaller magnitudes compared to the data. While using a stochastic trend improves the predictions of the model compared to a case with shocks around a stable trend, the volatility of interest rates and the probability of default still fall short of the data.
ratio and generating a default frequency twice as high as Arellano (2008). The model’s performance also improves in terms of correlation of output with spreads and net exports, with no deterioration in other dimensions. Hatchondo and Martinez (2009) also analyze the effects of introducing long-term bonds to a sovereign default model. Without using an asymmetric output cost of default, these authors show that mean spread, spread volatility, and default frequency generated by the model with long-term debt are much higher than those obtained assuming one-quarter bonds as in the standard model. Yue (2010) incorporates debt renegotiation and endogenous debt recovery into a sovereign default model to study the connection between default, debt renegotiation, and interest rates. This author finds that debt recovery rates decrease with indebtedness, which in turn affects the country’s ex-ante incentive to default and the terms of borrowing. Interest rates increase with the level of debt, owing to the higher default probability and to the lower debt recovery rate. The quantitative results of this model are similar to models without debt renegotiation along many dimensions, generating slight improvement in some statistics.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 explains data, the estimation strategy, and the estimated results. Section 4 provides extension and robustness checks. Section 5 concludes.

2 Model

This section explains the key features of Arellano (2008), which we use as the baseline model in our model estimation. We first discuss the sequence of decisions in the model. We then explain how default decision and default risk are modeled. The details of this model are provided in Appendix A.

2.1 Sequence of decisions

In period $t$, a country faces debt obligation $-B_t$. It then observes output, $y_t$, the log of which follows an AR(1) process. If the country had repaid in the previous period, then it would be able to choose to repay or default in period $t$, denoted as $d_t = 0$ and $d_t = 1$ respectively. If the
country chooses to repay, then it also decides how much it borrows in that period 
\((-q(B_{t+1}, y_t)B_{t+1})\) where \(q\) is the price of asset \(B\). If it chooses to default, it can write off its 
debt obligations at the expense of losing a fraction of output and being excluded from world 
financial markets for a stochastic number of periods. The state variables \((d, B, y)\) are thus 
sequentially determined by \((d_{t-1}, B_t) \rightarrow y_t \rightarrow (d_t, B_{t+1}).\)

2.2 Default probability and default decision

The country decides whether to repay its debt or default by comparing the value function under 
default \((V^D)\) with the value function under repayment \((V^R)\). Thus, the default decision of the 
country is given by

\[ d_t = \begin{cases} 
1, & \text{if } V^D(y_t) > V^R(B_t, y_t) \\
0, & \text{otherwise}
\end{cases} \]

where \(B_t\) is pinned down by the savings policy function of \(B(B_{t-1},y_{t-1}).\)

The country’s choice of \(B_t\) in period \(t - 1\) implies a default probability for period \(t\) 
conditional on \((d_{t-1}, B_{t-1}, y_{t-1})\), i.e., before \(y_t\) is observed. The default probability is given by

\[ \Pr(d_t = 1|d_{t-1} = 0, B_{t-1}, y_{t-1}) = \delta(B_{t-1}, y_{t-1}), \]
\[ \Pr(d_t = 0|d_{t-1} = 0, B_{t-1}, y_{t-1}) = 1 - \delta(B_{t-1}, y_{t-1}), \]
\[ \Pr(d_t = 1|d_{t-1} = 1, B_{t-1}, y_{t-1}) = 1 - \lambda, \]
\[ \Pr(d_t = 0|d_{t-1} = 1, B_{t-1}, y_{t-1}) = \lambda, \]

where \(\lambda\) is the exogenous probability of regaining access to financial markets for a country that 
has previously defaulted, and \(\delta\) is defined by

\[ \delta(B_t, y_{t-1}) = \Pr(y_t \in I(B_t)). \]

\(I(B_t)\) is the set of \(y\)’s for which default is optimal for \(B_t\), defined as

\[ I(B_t) = \{y_t \in \mathcal{Y}: V^D(y_t) > V^R(B_t, y_t)\}.\]
The default decision is made after the current output, $y_t$ is realized. Thus, the model-implied default decision for $d_t$ is a nonlinear function of $(d_{t-1}, B_t, y_t)$,

$$d_t = d(d_{t-1}, B_t, y_t).$$

If the country had defaulted in period $t-1$, it would not be able to borrow in period $t$. The country can regain market access with a fixed probability $\lambda$ in period $t$.

3 Estimation

3.1 Data

We use annual data for the Argentine output and repayment regime for our estimation. For output ($y$), we use real GDP at constant national prices for Argentina from Penn World Table 9.0 (Feenstra, Inklaar and Timmer, 2015). We remove a stochastic trend from the log of the real GDP series by applying the Hodrick-Prescott (HP) filter (with the smoothing parameter equals to 100), and then use the detrended component as $\ln(y)$.

For the regime variable ($d$), we construct a dummy variable that takes the value 1 under default years and zero otherwise following the default years identified by Reinhart (2010). Specifically, we set the default years as 1951, 1956—1965, 1982—1993, and 2001—2005.

The black solid line in Figure 1 plots the series of output with default years in the shaded areas. Table 1 provides the summary statistics. The output series is available for the period 1950—2014, but we set the sample period as 1950—2010, dropping the last four years after HP filtering the data to address the end of sample problem and to be consistent with the Reinhart (2010) coverage of default years.

---

2 We have obtained it through FRED and its series ID is RGDPNAARA666NRUG.
3 We refer to Reinhart (2010) who identifies default years from 1950. Covering a more recent period after 1980, additional research has been conducted to estimate the years of exclusion from capital market allowing them to differ from the duration of default status (e.g., Gelos, Sahay, and Sandleris, 2011) and taking into account the role of haircuts (Cruches and Trebesch, 2013).
We do not use debt data for our estimation because we can compute a model-implied debt path given the paths of repayment/default regimes and output from the data. Further, publicly available aggregate debt stock data may lack accuracy as it is often difficult to take into account all publicly guaranteed debt outstanding, debt reductions, and reschedules.

### 3.2 Estimation strategy

We introduce i.i.d. measurement errors to the regime variable to allow a model-implied default path to deviate from the observed default events. Specifically, we assume that \( \Pr(d^o_t = 0 \mid d_t = i) = a_i \) for \( i = 1 \) or \( 0 \) where \( d^o_t \) denotes the observed default behavior in the data with \( d^o_t = 1 \) corresponding to default and \( d^o_t = 0 \) corresponding to repayment in year \( t \). The superscript “o” indicates that the corresponding variable is observed in the data. The state space representation is non-Gaussian and nonlinear as follows.

\[
\begin{align*}
\ln(y^o_t) &= \ln(y_t), \\
\begin{cases}
0 & \text{if } u_t < a_i \\
1 & \text{otherwise}
\end{cases}, \\
\ln(y_t) &= \rho \ln(y_{t-1}) + \varepsilon_t, \\
d_t &= d(d_{t-1}, B_t, y_t), \\
B_t &= B(B_{t-1}, y_{t-1}), \\
u_t &\sim i.i.d. \text{ uniform (0,1)}, \\
\varepsilon_t &\sim i.i.d. N(0, \eta),
\end{align*}
\]

where the first two equations are observation equations and the remaining three equations are the state equations. The functions \( f \) and \( g \) are highly nonlinear.

---

\(^4\) Specifically, the model implied debt path can be computed using the policy function and the lagged state variables as \( B_t = B(B_{t-1}, y_{t-1}) \times 1_{d_{t-1} = 0} \).
We apply a maximum simulated likelihood method to estimate the model.\(^5\) Let \(D^0 \equiv \{d^0_t\}, D \equiv \{d_t\} \text{ and } \mathcal{Y} \equiv \{y_t\}. \) The joint distribution of \(D^0\) and \(\mathcal{Y}\) implied by the model can be written as

\[
P(D^0, \mathcal{Y}; \theta) = P(D^0|\mathcal{Y})P(\mathcal{Y}),
\]

where \(\theta\) is the set of model parameters: \(\sigma\) (risk aversion), \(r\) (risk-free rate), \(\beta\) (discount factor), \(\lambda\) (reentry probability), \(\rho\) and \(\eta\) (coefficients in the output equation), \(\bar{y}\) (output cost), and \(B_0\) (initial asset level). The log likelihood function is

\[
\ln P(D^0, \mathcal{Y}; \theta) = \ln \sum_i [P(D^0|D_i)P(D_i|\mathcal{Y})] + \ln P(\mathcal{Y}).
\]

The difficulty is that there is no analytical representation of \(P(D_i|\mathcal{Y})\). However, we can simulate \(D_i\) from \(P(D_i|\mathcal{Y})\) from the model.\(^6\) Thanks to the parsimonious model feature that there are only eight parameters and many of them have specific ranges, we can carry out simulations for all possible parameter-value combinations with reasonably fine and widely-ranged grids.

In the benchmark estimation, for simplicity, we assume no measurement errors in the initial year (\(d_{1950} = d_{1950}^0\)). Further, we fix \(r = 0.025\) (the historical average of real interest

---


\(^6\) Specifically, the steps of calculation are as follows. Step 1: Given data \(\mathcal{Y}\) and the model, simulate \(D_i\) from distribution \(P(D_i|\mathcal{Y})\) many times; Step 2: Given data \(D^0\), calculate forecast error probability \(P(D^0|D_i)\) for each simulation \(D_i\); Step 3: Sum up \(P(D^0|D_i)P(D_i|\mathcal{Y})\) over simulations. We further explain our numerical method in Appendix D.
rate data series\(^7\) to make the benchmark results comparable to observed levels of \(r\). Lastly, to avoid overestimating the initial debt level, we fix the lower bound of \(-B_0\) at -0.1.\(^8\)

**3.3 Estimated results**

We find the unique parameter set that achieves the highest likelihood function value in our numerical maximization framework (see Appendix D for details). The estimated parameters are shown in the first column in Table 2 together with the asymptotic standard errors of the estimated parameters computed using the score vector for observations (see Proposition 7.9 in Hayashi (2000) for details).

However, there are two concerns regarding the estimated parameter values. First, it turns out that the estimated discount factor is very low (\(\beta = 0.53\)).\(^9\) Such a low \(\beta\) seems at odds with the Euler equation\(^10\) at the steady state,

\[
\beta = \left[ (1 + r) \left( \frac{\delta^*}{1 - \delta^*} \left( \frac{c^R}{c^D} \right)^{\sigma} + 1 \right) \right]^{-1},
\]

\[
\text{Table 2}
\]

\(^7\) We measure the risk-free interest rate as the nominal interest rate (three-year US Treasury securities) minus the average inflation rate (using the GDP deflator) over the current and subsequent two years in the US. This series is available from 1954.

\(^8\) We limit the size of initial debt because otherwise it can be overestimated to account for the 1951 default, which is observed right after the initial year. Given that the ratio of gross external central government debt to export is 0.2 based on Reinhart and Rogoff’s (2011) database, 0.1 seems to be a reasonable lower bound of initial debt level.

\(^9\) We still obtain a low \(\beta\) value even when we exclude the coup d’état era of 1950s and 60s from the sample period.

\(^10\) The Euler equation is given by,

\[
1 = \frac{\beta}{\sigma} E_t \left[ d_{t+1} \frac{u'(c^R_{t+1})}{u'(\bar{c}R)} + (1 - d_{t+1}) \frac{u'(c^D_{t+1})}{u'(\bar{c}D)} \right] = \beta (1 + r) E_t \left[ c^R_{t+1} \left( \frac{c^R}{c^R_{t+1}} \right)^{\sigma} + (1 - d_{t+1}) \left( \frac{c^R}{c^R_{t+1}} \right)^{\sigma} \right]
\]

where the second equality holds by the CRRA utility and bond pricing equation with risk neutral lenders. If \(c^R\) and \(c^D\) were constant at the steady state, the above equation is reduced to \(\beta = \left[ (1 + r) \left( \frac{\delta^*}{1 - \delta^*} \left( \frac{c^R}{c^D} \right)^{\sigma} + 1 \right) \right]^{-1}\), where the superscript * indicates the steady state values.
where superscript ∗ indicates the steady state values. Suppose \( c^R/R^* \) \( /c^D/D^* = 1.03 \) (roughly in line with Arellano’s (2008) calibrated value of \( \bar{y} \)), \( r = 0.025 \) (the historical average of the real interest rate data series mentioned above), \( \sigma = 2 \) (the commonly used calibrated value), \( \delta^* = 0.11 \) (the historical average of default years given that the country repays in the previous year). Then, the implied \( \beta \) is 0.86, notably higher than the estimated value of 0.53. Second, the estimated risk aversion is quite high (\( \sigma = 9.5 \)) compared to the value commonly used in the literature (\( \sigma = 2 \)). As a result of this high \( \sigma \), the average of simulated consumption volatility is lower than that of simulated output volatility (see Section 3.3.1 for a description of our model simulations).

Since such high \( \sigma \) and low \( \beta \) are difficult to justify, we fix \( \sigma = 2 \) and and \( \beta = 0.8 \) (the calibrated values used by Aguiar and Gopinath, 2006) in the benchmark estimation.\(^{11}\) We report the estimated parameters for the baseline model in the second column of Table 2. In the benchmark estimation, \( \lambda \) (probability of reentry) is 0.49, which is less than the value used by Arellano (2008). This value of \( \lambda \), however, is notably higher than the value implied by the historical average of default duration. The average duration of observed default years is 7 years\(^{12}\) in our sample. The value of \( \lambda \) that implies a 7-year default duration on average is 0.14. The value of \( \bar{y} \) (output cost, 0.99) is consistent with the calibrated values in the literature. The estimated coefficients for the output dynamics (\( \rho \) and \( \eta \)) are consistent with the simple AR(1) estimates. The relatively low value of \( \rho \) (0.55) reflects the low persistency of output gap at annual frequency obtained via HP filtering. The estimate of \( a_1 \) (the probability that the repayment is observed in the data, given the model implies default) is 0. Thus, we observe defaults whenever the country chooses to default in the model. The estimate of \( a_0 \) (the

\[^{11}\] A reasonable value of \( \sigma \) for a small open economy might be higher than 2 (Reinhart and Végh, 1995). We thus re-estimate the model now fixing \( \sigma=5 \) while keeping all other assumptions the same as the benchmark. It turned out that the implied default probabilities are very similar with those of the benchmark estimates and thus they are not shown here.

\[^{12}\] Uribe and Schmitt-Grohé (2017) review the existing estimates of years of exclusion from credit markets after default and find that on average countries regain full access to credit markets 8.4 years after emerging from default.
probability that the observed default variable is repayment given the model-implied default variable also indicates repayment) is 0.75, which is lower than the unrestricted estimate (0.88). These two measurement-error-related parameters are not needed in our simulations discussed below.

Why do the unrestricted estimates give such high σ and low β? The unrestricted estimate for λ (probability of reentry) is consistent with the observed duration of default years but is much lower than the benchmark estimate. The lower λ implies a higher penalty upon default; as a result, the country has less incentive to default. To offset this diminished incentive, the model parameters adjust and give a combination of high σ and low β. The higher the σ or the lower the β, the greater the incentive for the country to default.

3.3.1. Simulated default probabilities

Figure 2 compares the results from the model with the benchmark parameter estimates and the unrestricted parameter estimates. The probability that the default outcome after \( y_t \) is realized is shown by the red solid line for the benchmark estimates and by the blue dashed line for the unrestricted parameter estimates. We call this probability the ex-post default probability. In other words, it shows \( \Pr(d_t = 1 \mid y_t, B_t) \). Since the default decision is made after \( y_t \) has materialized, the default outcome becomes a certain event given the default decision. Thus, for a country that has access to the financial markets, the ex-post probability of default equals 1 if the country chooses to default, and zero if repayment is chosen. For a country in autarky, on the other hand, the probability of remaining in the default state or not depends on the exogenous probability of regaining access to markets, \( \lambda \). Therefore, the simulated ex-post default probability can fluctuate between zero and one after the decision to default due to the exogenous probability of reentry. The default probabilities are computed as averages of 10,000 simulations.

[Figure 2]

---

13 The unrestricted estimate is lower than 1 reflecting the model fails to predict the 1960 default.
With either set of estimates, the timing of the default decisions is quite similar as seen by the comovement of the two lines. While the benchmark estimation matches the default status of the country better than the unrestricted estimation during the repayment years, the unrestricted estimation performs better in the default years. This difference is due mainly to the value of $\lambda$ being much lower in the unrestricted estimation compared to the benchmark (0.14 versus 0.49). With a low $\lambda$ value, the probability of staying in autarky after the decision to default remains higher as seen by the smaller swings in the blue line. The low $\lambda$ value also leads to a slower decline in the probability of default once it increases; as a result, the default probability predicted by the unrestricted estimates in the years of repayment (the white areas) is higher than the benchmark estimates. Overall, the benchmark results explain the country’s decisions under the repayment years better at the expense of fitting the model to default years during which no decisions are modeled.

Figure 3 plots two default probabilities from the benchmark model. The red solid line shows the ex-post default probability, conditional on $y_t$ as in Figure 2. The blue dashed line shows the ex-ante default probability, which is conditional on $y_{t-1}$. These two default probabilities differ from each other based on whether they are conditional on the current or the previous period’s output.

[Figure 3]

The model matches the observed default events in 1982 and 2001 with the ex-post default probability, shown by the red line, equaling 1 in these years. The ex-post default probability continues to remain high in the years following the default events, identified as default years in the data, even though it fluctuates due to the exogenous reentry probability. The ex-post default probability falls close to zero in 1994 and 2006, when Argentina regained access to financial markets, matching the data perfectly. During the repayment years, it does not fluctuate with output and stays close to zero as long as the country remains below the model-implied endogenous debt ceiling. The model, however, predicts the 1956 default with a lag during the coup d’état era of the 1950s and 1960s.
The ex-ante default probability, which is the probability of default conditional on \( y_{t-1} \) and shown by the blue dashed line, moves similarly to the ex-post default probability in the default years, but usually follows the ex-post probability with a lag. This lagged pattern suggests that the decline in output observed at the time of default events is important for the model to predict a default.

### 3.3.2 Real-time default probability

One may be interested in directly computing default probability for the observed default events, because there is measurement error in the repayment regime variable. Formally, we may be interested in computing the following probability,

\[
Pr(d^o_t | D^o_{t-1}, Y_t),
\]

where \( D^o_{t-1} = \{d^o_{t-1}, \ldots, d^o_0\} \) and \( Y_t = \{y_t, \ldots, y_0\} \). We may call this probability real time default probability because it is the probability that the observed default variable in period \( t \) (\( d^o_t \)) takes a particular value given its past values (\( D^o_{t-1} \)) and the output data information up to that period (\( Y_t \)). In Appendix C, we show that this probability can be rewritten as the measurement-error and model-implied components using a part of the derivation used in the likelihood function derivation.

By definition, our default probability measure can be readily compared with the existing real-time indicators. For example, the dashed line in Figure 4 is EMBI Global for Argentina, a commonly used measure for individual country risk implied by the financial markets. Contrary to our real-time default probability measure, the EMBI Global indicator fluctuates under the observed repayment regime periods. In Section 4.2, we also provide a comparison with a logit model.

[Figure 4]
3.4 Other model variables

This section analyzes the business cycle statistics generated by the model. Table 3 compares the moments related to consumption, interest rate spread and net exports with those from the data and Arellano (2008). For each of these variables, we compute the average of 10,000 simulated paths given the benchmark parameter estimates and output data series.

[Table 3]  

The statistics from the model are broadly consistent with the data and Arellano (2008). Specifically, consumption is more volatile than output and interest spread is countercyclical. The model generates weakly procyclical net exports contrary to data. Trade balance is countercyclical during periods in which the country makes its repayment decisions. However, the correlation between trade balance and output becomes positive over default periods. It is difficult to generate unconditionally countercyclical trade balance as there are too many default periods in our data. Figure 5 shows the model-implied spread. It is highly volatile because of the very high spreads under the default regime as well as during the 1950s. The standard deviation of model-implied spread under the repayment regime after 1970 is only 8.1 percent.

[Figure 5]

4 Robustness and Extension

4.1 A stochastic risk-free rate

The baseline model focuses on idiosyncratic default risk. A natural way to consider a systemic risk is to introduce a stochastic risk-free interest rate to the model. This section extends Arellano (2008) by assuming that the risk-free interest rate \( r \) follows the AR(1) process. Thus, \( r \) becomes an additional state variable in the extended model. The sequence of decisions is a minor modification to that of the baseline,

\[
(d_{t-1}, B_t) \rightarrow y_t, r_t \rightarrow (d_t, B_{t+1}),
\]
where period $t$ risk-free real interest rate ($r_t$) does not affect debt repayment obligation in that period ($B_t$). We describe the extended model in Appendix A.

Does this additional model feature help to explain default episodes? We estimate the extended model by maximum simulated likelihood estimation, again fixing the values of $\beta$ and $\sigma$ as in the benchmark estimation. The corresponding likelihood function with stochastic $r$ is derived in Appendix B.2. For the risk-free interest rate ($r_t$), we use an ex-ante real interest rate because $r_t$ does not affect debt obligation while it does affect the default decision in period $t$. This is consistent with the above sequence of decisions in the model. In constructing this series, we follow the procedure outlined by Mishkin (1981) using data obtained from the FRED database. Specifically, we subtract the University of Michigan inflation expectation measure (MICH) from the 3-month treasury bill rate (TB3MS) and compute yearly averages. The former series is available from 1978. This ex-ante real interest rate series (the red dashed line is Figure 6) increases prior to observed default events.

[Figure 6]

Figure 7 shows the simulated default profiles in both baseline and extended models. The simulated default decision in either model explains the timing of the 1982 and 2001 default events quite well (Figure 7a). Both models have similar default probability profiles (Figure 7b) although the extended model explains the heightened default risk prior to default episodes slightly better as its simulated default probability increases (from 10 to 30 percent prior to the 1982 default, and from 5 to 17 percent prior to the 2001 default).

[Figure 7]

4.2 A comparison with a logit model

A popular reduced-form approach to estimate default probability is to estimate a logit model. For example, Kaminsky et al. (2016) estimate a logit model to examine idiosyncratic default risks—which have been emphasized in the theoretical sovereign debt models—for Argentina using data from 1820 to Great Depression.
Figure 8 plots the default probability estimated with the logit model that regresses $d_t^o$ on a constant, $d_{t-1}^o$ and $y_t$. The figure shows that our real-time default probability accounts for the timing of occurrence of default events better. Further, the logit-based default probability under the repayment regime is more volatile than our default probability measure. Thanks to the endogenous debt ceiling, the baseline model better accounts for repayment decisions under the repayment regime.

[Figure 8]

The logit and baseline models have similar fit to the data with the corresponding log likelihood value being -19 for the logit model and -21\textsuperscript{14} for the baseline model with unrestricted parameter estimate. The value implied by the baseline model with the benchmark parameter estimates (-32), however, is significantly lower, because of the poor performance of the baseline model during the observed default years and relatively low value of $a_0$ (0.75).

Predictions from the baseline model and the logit model are complements. In practice, our baseline model is more reliable for predicting the timing of default and we could use the logit model to forecast duration of default years, which is not well modeled in the model of Arellano (2008).

5 Conclusion

By formally estimating the Arellano (2008) model, we find that a sovereign debt crisis model is a useful indicator for Argentine default decisions. Despite using only output and default data in our estimation, the benchmark results account for overall default patterns of Argentina as well as business cycle properties consistent with the Arellano (2008) findings. Considering systemic risks by introducing a stochastic risk-free interest rate does not notably improve the model’s

\textsuperscript{14} This value corresponds to in $\mathcal{L}_A$ in Appendix B.
accountability for default events. These results may suggest that updating Argentine output information is a key to predicting its default events.

We also provide a novel real-time default probability measure that exploits the nonlinear nature of the model and allows a measurement error to the default variable. This real-time measure better agrees with the timing of observed default occurrence than the logit-based measure.

An important caveat on the model-implied business cycle properties is that if we use the unrestricted parameter estimates (that imply a high \( \sigma \)) for model simulation, consumption becomes less volatile than output. These estimates are affected by model fitting to default periods where no decisions are modeled. Going forward, following development in the theoretical literature in debt negotiation and restructuring, a more explicit modeling of default duration may help to improve model performance in accounting for business cycle properties.
References


Appendix A: The model

Arellano (2008)

This appendix summarizes the Arellano (2008) model. There are two regimes \( \langle d_t \rangle \): default regime \( \langle d_t = 1 \rangle \) and repayment regime \( \langle d_t = 0 \rangle \). The model is set up as a planner’s problem with the resource constraint given by

\[
\begin{align*}
  c_t &= y_t - q(B_{t+1}, y_t)B_{t+1} + B_t, \quad \text{under repayment}, \\
  c_t &= h(y_t), \quad \text{under default},
\end{align*}
\]

where \( y \) is output and \( h(y_t) = \overline{y} \) if \( y_t > \overline{y} \) and \( h(y_t) = y_t \) if \( y_t \leq \overline{y} \). \( c \) is consumption and \( q \) is the price of the asset. The log of output is assumed to follow the AR(1) process, i.e.,

\[
\ln(y_t) = \rho \ln(y_{t-1}) + \epsilon_t, \quad \epsilon_t \sim N(0, \eta). \tag{8}
\]

Denoting period \( t + 1 \) variables with prime and period \( t \) variables with no time subscript, the value functions are given by

\[
\begin{align*}
  V^D(y) &= u(h(y)) + \beta E[\lambda V^R(0, y') + (1 - \lambda)V^D(y')], \\
  V^R(B,y) &= \max_{B'} u(y - q(B', y)B' + B) + \beta E[\max \{V^D(y'), V^R(B', y')\}], \\
  &= u(y - q(B(B,y), y)B(B,y) + B) + \beta E[\max \{V^D(y'), V^R(B(B,y), y')\}],
\end{align*}
\]

where \( B(\ldots) \) is the savings policy function, \( \mu \) is the asset level under default, and \( \ln(y') = \rho \ln(y) + \epsilon' \). \( \lambda \) is the reentry probability.

With risk-neutral lenders, the bond price satisfies

\[
q(B(B,y), y) = \frac{1 - \delta(B(B,y), y)}{1 + r}, \tag{9}
\]

where \( \delta \) is endogenous default probability given by

\[
\delta(B(B,y), y) = \Pr(y' \in I(B')).
\]
with \( I(B) = \{ y \in \mathcal{Y} : V^D(y) > V^R(B, y) \} \).

**An extended model with a stochastic risk-free rate**

We extend the baseline model to include a stochastic process for the risk-free interest rate \( r \) as follows:

\[
r_t = \mu_r + \rho_r r_{t-1} + z_t, \quad z_t \sim N(0, \sigma_r).
\]

In this version of the model, the price of bonds is given by

\[
q(B', y, r) = \frac{1 - \delta(B', y, r)}{1 + r},
\]

where \( \delta(B', y, r) \) is the endogenous default probability, which in this case depends on the interest rate state \( r \) as well as \( B' \) and \( y \).

The resource constraint of the economy depending on the government’s default decision is given by

\[
c = \begin{cases} 
    y - q(B', y, r)B' + B, & \text{under repayment,} \\
    h(y), & \text{under default.}
\end{cases}
\]

The government observes the interest rate shock besides the income level, and chooses whether to repay or default given its existing debt, \( B \). The value of repayment, depending on the state \( s \), is given by

\[
V^R(B, y, r) = \max_{B'} u(y - q(B', y, r)B' + B) + \beta E[\max\{V^D(y', r'), V^R(B', y', r')\}].
\]

With a constant probability of reentry to financial markets \( \lambda \), the value function for default is given by

\[
V^D(y, r) = u(h(y)) + \beta E[\lambda V^R(0, y', r') + (1 - \lambda) V^D(y', r')].
\]

For a country that decides to repay its debt and chooses \( B' \) as the new debt level, the probability of default for the next period depending on the interest rate state is defined as
\[ \delta(B', y, r) = \Pr((y', r') \in I(B')) , \]

where \( I(B) \) is the set of \((y, r)\) pairs for which default is optimal for the debt level \( B \):\]

\[ I(B) = \{ (y, r) \in (X, R) : V_D(y, r) > V_R(B, y, r) \} . \]
Appendix B: The Likelihood Functions

This appendix derives the likelihood functions for the baseline and extended models. Data on output and default variables are used in the estimation, allowing measurement error on the observed default variables.

The likelihood function for Arellano (2008)

The likelihood function of the data is thus given by

\[ L = p(d_1^o, \ldots, d_T^o, y_1, \ldots, y_T|d_0^o, y_0), \]

where the superscript \( o \) indicates observed default variable. \( L \) can be rewritten as

\[ L = \frac{p(\bar{D}_T^o|d_0^o, Y_t) p(\bar{Y}_T|d_0^o, y_0)}{L_A} \]

where \( Y_t \equiv \{y_t, \ldots, y_0\}, \bar{Y}_t \equiv \{y_t, \ldots, y_1\}, \) and \( \bar{D}_t^o \equiv \{d_t^o, \ldots, d_1^o\} \).

\( L_A \) can be rewritten as

\[ p(\bar{Y}_T|d_0^o, y_0) = \prod_{t=1}^{T} f(y_t|d_0^o, Y_{t-1}), \text{(by seq.factorization)} \]

\[ = \prod_{t=1}^{T} f(y_t|y_{t-1}), \text{(the log of } y \text{ follows the AR(1)}) \]

\[ = \prod_{t=1}^{T} \phi \left( \frac{\ln y_t - \rho \ln y_{t-1}}{\eta} \right), \]

where \( \phi(.) \) is the pdf of the standard normal distribution.

\( L_A \) can be rewritten as
\[ p(D_T^0|d_0^0, Y_T) = \sum_{(d_T^0,d_0^0)} p(D_T^0, D_T|d_0^0, Y_T), \text{where } D_t = \{d_t, \ldots, d_0\}, \]
\[ = \sum_{(d_T^0,d_0^0)} p(d_T^0, D_T^{-1}|d_0^0, Y_T), \]
\[ = \sum_{(d_T^0,d_0^0)} p(d_T^0|D_T^{-1}, D_T, d_0^0, Y_T)p(D_T^{-1}, D_T|d_0^0, Y_T), \]
\[ = \sum_{(d_T^0,d_0^0)} p(d_T^0|D_T)p(D_T^{-1}, D_T, d_0^0, Y_T)p(D_T^{-2}, D_T|d_0^0, Y_T), \]
\[ = \sum_{(d_T^0,d_0^0)} \left[ \prod_{i=1}^{T} p(d_i^0|d_i) \right] p(D_T|d_0^0, Y_T). \quad (10) \]

By the model, \( p(D_T|d_0^0, Y_T) \) can be further rewritten as

\[ p(D_T|d_0^0, Y_T) = A \prod_{i=1}^{T} \Pr(d_i|d_{i-1}, d_0^0, Y_T), \]
\[ = A \prod_{i=1}^{T} \Pr(d_i|d_{i-1}, B_i, y_i; B_0), \]
\[ = A \prod_{i=1}^{T} \Pr(d_i|d_{i-1}, B(B_{i-1}, y_i), y_i; B_0). \quad (11) \]

where \( A = \Pr(d_0|d_0^0, Y_T). \) \( \Pr(d_i|d_{i-1}, B_i, y_i; B_0) \) in the second equality corresponds to the model-implied default decision rule which can be expressed as

\[
\Pr(d_t = d(d_{t-1}, B_t, y_t)|d_{t-1} = 0, B_t, y_t) = 1, \\
\Pr(d_t = 1|d_{t-1} = 1, B_t, y_t) = 1 - \lambda, \\
\Pr(d_t = 0|d_{t-1} = 1, B_t, y_t) = \lambda,
\]
where uncertainty arises only through the exogenous reentry probability $\lambda$ if the country defaulted in the previous period. The last equality holds by the saving policy function. The constant, $A$ in eq. (11) can be further rewritten as

$$\Pr(d_0^0|d_0^0, Y_t) = \frac{\Pr(d_0^0|d_0, Y_t)\Pr(d_0^0)}{\Pr(d_0^0)}, \text{ (by the Bayes rule)}$$

$$= \Pr(d_0^0|d_0 = 1)\Pr(d_0 = 1) + \Pr(d_0^0|d_0 = 0)\Pr(d_0 = 0).$$

By eqs. (8) and (10), the log likelihood function is given by

$$L = \ln \left[ \sum_{(d_T, \ldots, d_0)} \prod_{t=1}^{T} p(d_t^i|d_i) \prod_{i=1}^{T} \Pr(d_i|d_{i-1}, B(B_{i-1}, y_i), y_i; B_0) \right]$$

$$+ \sum_{t=1}^{T} \ln \left[ \phi \left( \frac{\ln y_t - \rho \ln y_{t-1}}{\eta} \right) \right],$$

where the parameter vector includes $\sigma, r, \beta, \lambda, \rho, \eta, \bar{y}, B_0, a^D, a^R$.

**The likelihood function with a stochastic risk-free rate**

The likelihood function of the data is thus given by

$$\mathcal{L} = p(d_1^0, \ldots, d_T^0, y_1^0, \ldots, y_T^0, r_1^0, \ldots, r_T^0|d_0^0, y_0^0, r_0^0).$$

In a similar manner as the baseline likelihood function derivation, we can show that the log likelihood function is given by

$$L = \ln \left[ \sum_{(d_T, \ldots, d_0)} \text{const} \left( \prod_{t=1}^{T} p(d_t^i|d_i) \right) p(D_T|d_0^0, Y_T, R_T) \right]$$

$$+ \sum\ln \left[ \phi \left( \frac{\ln y_t - \rho \ln y_{t-1}}{\eta} \right) \right] + \sum\ln \left[ \phi \left( \frac{\ln r_t - \rho_r \ln r_{t-1}}{\eta_r} \right) \right],$$

where the parameter vector includes $\sigma, r, \beta, \lambda, \rho, \eta, \bar{y}, B_0, \rho_r, \eta_r, a^D, a^R$. 

25
Appendix C: Real-Time Default Probability

This appendix shows the real-time default probability discussed in the text (eq. (7)) can be rewritten into the model-implied and measurement-error components.

\[
\Pr(d_t^o|D_{t-1}^o, Y_t) = \frac{\Pr(\tilde{D}_t^o | d_0^o, Y_t)}{\Pr(\tilde{D}_{t-1}^o | d_0^o, Y_t)} \\
\approx \frac{\sum_{(d_t^o-\cdots-d_0)} \prod_{i=1}^{t-1} p(d_i^o | d_i) p(D_t | d_0^o, Y_t)}{\sum_{(d_{t-1}^o-\cdots-d_0)} \prod_{i=1}^{t-1} p(d_i^o | d_i) p(D_{t-1} | d_0^o, Y_t)}.
\]

Appendix D: Numerical Maximization

The solution algorithm for the baseline model is as follows:

1. Start with an initial guess for the bond price function \( q(B', y) \) that corresponds to a default probability of zero for each point in the state space.

2. Using this initial price and initial guesses for \( V^R(B, y) \) and \( V^D(B, y) \), iterate on the Bellman equations to solve for the optimal value and policy functions.

3. Given the optimal default decision, update the price of bonds using equation (9). Repeat steps 2 and 3 until the bond price converges, i.e. until \( |q^{i+1} - q^i| < \varepsilon \), where \( i \) represents the iteration number and \( \varepsilon \) is a very small number.

There are only ten parameters \( (\sigma, r, \beta, \lambda, \rho, \eta, \tilde{y}, B_0, a_0, a_1) \) in the baseline model, and many of these parameter values have restrictions on their ranges. For example, the ranges of \( \beta, \lambda, \tilde{y}, a_0 \), and \( a_1 \) are between 0 and 1. The values of \( \rho \) and \( \eta \) should not be very different from the OLS estimates of the \( y \) equation alone. These restrictions enable us to compute likelihood values of all possible combinations of parameter values with reasonably fine grids.
Table 1. Summary statistics.

<table>
<thead>
<tr>
<th></th>
<th>y (output)</th>
<th>s (regime)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1.02</td>
<td>0</td>
</tr>
<tr>
<td>std. dev.</td>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td>min</td>
<td>1.11</td>
<td>0</td>
</tr>
<tr>
<td>max</td>
<td>0.96</td>
<td>0</td>
</tr>
<tr>
<td>mean</td>
<td>0.98</td>
<td>1</td>
</tr>
<tr>
<td>std. dev.</td>
<td>0.05</td>
<td>0</td>
</tr>
<tr>
<td>min</td>
<td>1.05</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>0.84</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2. Estimated parameters.

<table>
<thead>
<tr>
<th></th>
<th>Baseline model (unrestricted)</th>
<th>Baseline model (benchmark, $\sigma=2$)</th>
<th>Extended model</th>
<th>Arellano (2008) (annualized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$ (risk aversion)</td>
<td>9.5 (3.34)</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$\beta$ (discount factor)</td>
<td>0.53 (0.03)</td>
<td>0.80</td>
<td>0.80</td>
<td>0.82</td>
</tr>
<tr>
<td>$1+r$ (risk-free rate)</td>
<td>1.03</td>
<td>1.03</td>
<td>—</td>
<td>1.07</td>
</tr>
<tr>
<td>$B_0$ (initial asset level)</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.00</td>
<td>—</td>
</tr>
<tr>
<td>$\bar{y}$ (output cost)</td>
<td>0.95 (0.75)</td>
<td>0.99</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>$\lambda$ (reentry probability)</td>
<td>0.14 (0.05)</td>
<td>0.49</td>
<td>0.56</td>
<td>0.73</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.55 (0.11)</td>
<td>0.55</td>
<td>0.66</td>
<td>0.85</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.04 (0.003)</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>$a_0$</td>
<td>0.88 (0.03)</td>
<td>0.75</td>
<td>0.71</td>
<td>—</td>
</tr>
<tr>
<td>$\mu_r$</td>
<td>—</td>
<td>—</td>
<td>-0.06</td>
<td>—</td>
</tr>
<tr>
<td>$\rho_r$</td>
<td>—</td>
<td>—</td>
<td>0.89</td>
<td>—</td>
</tr>
<tr>
<td>$\eta_r$</td>
<td>—</td>
<td>—</td>
<td>0.01</td>
<td>—</td>
</tr>
</tbody>
</table>

In annualized values. The numbers in parentheses are standard errors. We prefix the coefficient values of risk aversion, risk-free rate and discount factor in our estimation for reasons discussed in Section 3.3. We do not report the standard error of $a_1$ and the initial asset level since they are estimated at the lower boundary of zero.
Table 3. Business cycle statistics.

<table>
<thead>
<tr>
<th></th>
<th>Data all periods</th>
<th>Model (benchmark, σ=2) on the lagged subsample of repayment regime</th>
<th>Arellano’s (2008) quarterly stat.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma(c)/\sigma(y) )</td>
<td>1.19</td>
<td>1.05</td>
<td>1.02</td>
</tr>
<tr>
<td>( \sigma(nx/y) )</td>
<td>2.58</td>
<td>0.86</td>
<td>0.94</td>
</tr>
<tr>
<td>( \sigma(spread) )</td>
<td>12.30</td>
<td>30.60</td>
<td>40.44</td>
</tr>
<tr>
<td>corr((c,y))</td>
<td>0.90</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>corr((nx/y,y))</td>
<td>-0.81</td>
<td>-0.13</td>
<td>0.03</td>
</tr>
<tr>
<td>corr((spread,y))</td>
<td>-0.81</td>
<td>-0.32</td>
<td>-0.35</td>
</tr>
</tbody>
</table>

Net exports are exports minus imports; the spread is in percentages. All series except net exports and the spread are in logs. All series have been HP filtered. Standard deviations are reported as percentages. All statistics are based on annual data. Sample periods are 1950-2010 for output and consumption, 1960-2010 for net exports, and 1983-2010 for the spread.
Figure 1. Output and default data series

The solid line plots a detrended output series for Argentina from 1950 to 2010. The shaded areas are default years.
This figure plots the averages of 10,000 simulated default decisions given the output data series and either the benchmark or unrestricted parameter estimates. The red solid line is the simulated default decisions with the benchmark estimates. The blue dashed line is that with unrestricted parameter estimates. In the figure, the initial default decision is set equal to the observed default variable. The shaded areas are default years.
Figure 3. Simulated ex-ante and ex-post default probabilities

This figure plots the averages of 10,000 simulated ex-ante (blue dashed) and ex-post (red solid) default probabilities given the output data series and the benchmark parameter estimates. We call default decisions the ex-post default probabilities. In the figure, the initial default decision is set equal to the observed default variable. The shaded areas are default years.
Figure 4. Estimated real time default probability

The solid line plots estimated real time default probability from 1951 to 2010 with the benchmark parameter estimates. The red dashed line is the end-of-year values of EMBI Global Argentina (stripped spread). The shaded areas are default years.
Figure 5. Simulated spread

In annualized rate in percent. The solid line is $1/q$ minus $(1+r)$. The solid line shows the simulated spread for the observed repayment years and the first years of default years. The red dashed line shows the spread data for Argentina. The spread data between 1983-1993 are taken from the dataset by Neumeyer and Perri (2005). The data for 1994-2010 is EMBI Global Argentina (stripped spread). The shaded areas are default years.
Figure 6. Output and the risk-free real interest rate

The red dashed line (right axis) is an ex-ante risk-free real interest rate series and the solid line is a detrended output series for Argentina for 1978-2010. The shaded areas are default years. These data series of output, default years, and real interest rate are used for the extended model estimation.
Figure 7. Simulated default profiles in the baseline vs. extended models

a) Simulated ex-post default probabilities

The red solid line plots the simulated results from the extended model and the blue dashed line plots those from the baseline model. The shaded areas are default years. The sample period for the extended model is 1978-2010 and the baseline model is 1950-2010.

b) Simulated ex-ante default probabilities
The red dashed line plots logit-based default probabilities obtained by regressing the observed default variable on a constant, the lagged observed default variable, and the current output. The black-solid and blue-dotted lines are real-time default probabilities with the unrestricted and benchmark parameter estimates respectively. The shaded areas are default years. The sample period is from 1950 to 2010.