Asymmetric Information: Persuasion and Underpricing of Heightened Sovereign Default Risk

Zachary R. Stangebye*

September 4, 2024

Abstract

A model of sovereign debt with asymmetric information is explored quantitatively with machine learning methods. Sovereigns come in two hidden persistent types with different propensities to borrow and default. The 'bad' type defaults significantly more than the 'good' type. Asymmetric information both increases default risk overall and results in systemic underpricing of this heightened default risk in that average risk-neutral spreads are significantly lower than average default frequencies. This happens because the bad type often persuades lenders that his default risk is lower than it is, especially during debt crises.

Keywords: Sovereign debt, asymmetric information, underpricing

JEL: E44, F34, G01, H63

*Department of Economics, University of Notre Dame. 3015 Nanovic Hall, Notre Dame, IN 46556. United States. E-mail: zstangeb@nd.edu. Phone: (574) 631-1054. Fax: (574) 631-4783

1 Introduction

One of the perils of foreign investments is that investors typically have less access to payoff-relevant information about the foreign country than they have about their own country. The resulting asymmetric information problem has significant consequences for market dynamics that the literature has explored in some depth (Cole and Kehoe 1998, Sandleris 2008, D'Erasmo 2008, Phan 2017, Amador and Phelan 2021, Stangebye and Wright 2024, or Fourakis 2023 among others).

In this paper, I document two novel quantitative consequences of asymmetric information in a sovereign debt model. First, default frequencies are increased relative to the case of full information. And second, this heightened sovereign default risk is systemically underpriced. By this, I mean that in canonical quantitative environments, asymmetric information can introduce a substantial divergence between risk-neutral bond spreads and average default frequencies, which are closely linked in standard models (Arellano 2008 or Aguiar and Gopinath 2006). For standard parameter calibrations, average risk-neutral spreads can be far *lower* than the associated unconditional default risk that they are pricing.

These phenomena arise in a simple extension of the sovereign default model of Arellano (2008) with asymmetric information regarding the discount factor of the government. There are two optimizing types: A bad type with a low discount factor and a good type with a higher discount factor. There is some exogenous switching between the types, but they are persistent. An additional transitory, unobserved liquidity shock works to obscure perfect revelation and introduces a plausible signal-noise inference framework into the investors' problem: A particular fiscal position or default event may be driven by the persistent type *or* by the realization of the liquidity shock.

Asymmetric information raises default probabilities for both types, and thus raises unconditional default frequencies overall. The good type defaults more often because he is not rewarded with a better budget set for being the good type; instead, his budget set is typically contracted toward that of the full-information bad type. This reduces the good type's value of repayment and causes him to default more often. The bad type defaults more often because his budget set is typically expanded mildly by virtue of the fact that investors sometimes perceive him to be the good type, i.e., be less of a credit risk than he actually is. He takes advantage of this and borrows into higher debt regions where default frequencies are higher.

Equally fascinating is the fact that this additional risk is systematically underpriced. To understand why requires a brief discussion of the dynamics of debt crises in the model. In most states of the world, risk spreads are at or near zero regardless of reputation. Thus, in most states of the world, reputation has virtually no impact on aggregate dynamics, which are driven largely by price fluctuations (see Aguiar et al., 2016).

During a debt crisis, however, when output is unexpectedly low, debt levels are high, and the sovereign is close to his borrowing capacity, then the contribution of reputation can have a substantial impact on borrowing costs, sometimes nearly doubling spreads even for small price differences. Because the bad type tends to borrow more than the good type, it is almost always the bad type who finds himself in such crises. In these situations the bad type *deliberately* tries to improve his reputation by deleverageing, which lowers spreads below actual default probabilities.[1]

The success of this persuasion-strategy taken by the bad type during debt crises is reminiscent of the findings of Kamenica and Gentzkow (2011), who note that in asymmetric information environments with fully rational agents the sender of a signal can strictly benefit by manipulating the information set of the receiver, even though the receiver is fully aware that this is taking place. These authors note that Bayesian inference restricts the expectation of posterior beliefs but places no other constraints on the distribution. Consequently, as long as investor's investment decisions are not linear in their beliefs, which they are not, the sovereign can (potentially) strictly benefit from this manipulation. In this case, he does so by currying lower spreads during debt crises.

As a consequence, nearly all defaults on the equilibrium path are initiated by bad types who recently have gone out of their way to persuade investors that they are not as risky as they truly are. As most strictly positive risk-spreads occur during periods like this, average risk-spreads are uniformly lower than they should be relative to the true default risk they price. For a reasonable parameterization, default risk can be more than $4.5\times$ the average spread.

While this divergence is quite large in plausible cases, not all parameterization foster these dynamics. For instance, if fundamental observable uncertainty is too large, then both types will default with more similar frequencies. This mutes or even eliminates these underpricing dynamics, which rely on a divergence

[1]Such behavior is not fully revealing because of the presence of the liquidity shock. By deleverageing here, the bad type is effectively trying to communicate to lenders that he got a better liquidity shock than he actually received.

in default probabilities across types. This is perhaps why this phenomenon has not been documented in other related quantitative work, e.g., D'Erasmo (2008), Fourakis (2023), or Chatterjee et al. (2023).[2]

2 Model Description

The benchmark model is an extension of Arellano (2008) to allow for asymmetric information with respect to the sovereign's discount factor. It follows the basic structure of Stangebye and Wright (2024) with the inclusion of observable endowment shocks and re-entry following default.[3]

2.1 Shocks and Information Structure

Publicly observable output is characterized by a Markov process following Arellano (2008). In particular, $z_t = \log(Y_t)$ and

$$z_t = \rho_z z_{t-1} + \sigma_z \epsilon_t$$

where ϵ_t is an iid standard normal. I succinctly write $s_t = [z_t, Y_t]$ as the vector of observable states.

Asymmetric information enters in two ways. First, the sovereign's discount factor, β, is unknown to the investors. It will take one of two values, β^L or β^H, with the former being smaller. I call this the sovereign's 'type' and allow for some infrequent and stochastic switching across types.

In particular, there is a symmetric and persistent transition probability across these types given by

$$Pr(\beta^H|\beta^L) = Pr(\beta^L|\beta^H) = p_\beta < 0.5$$

Second, there additional fully transitory but unobserved shock to the sovereign's marginal utility from consumption that I denote m_t.[4]

The transitory shock has any potentially type-specific distribution, $m_t \sim F_{\beta_t}$. I assume that it is

[2]Models of asymmetric information are notorious for engendering equilibrium multiplicity. Because the solution technique here is fundamentally different than these works, it is also possible that the underpricing phenomenon may manifest in equilibria besides their benchmark equilibrium.

[3]Indeed, the model described here can be found as an extension from Appendix B of Stangebye and Wright (2024), though there is no exploration of heightened default risk or underpricing of that default risk, as the function of the model in that work is to highlight the attenuation of default behavior across types and budget sets across beliefs.

[4]The transitory shock is iid and can be interpreted either as an endowment shock or a preference shock to the level of subsistence consumption, which directly affects the sovereign's incentives to borrow. See Chatterjee and Eyigungor (2012) for a proof of this result.

continuous and iid.

This dual shock structure serves a few purposes. First, it is empirically plausible. When lenders see a sudden surge in indebtedness or default risk, they are often left trying to tease apart how much is due to underlying institutional change in the borrowing country and how much is due to random, uncontrollable liquidity shocks that can happen even to a well-behaved government. Second, it increases numerical stability (see Chatterjee and Eyigungor (2012). Third, and most importantly, it gives the sovereign access to a mechanism he can use at any time to improve his reputation in any state regardless of his underlying type. Because of this shock is unobserved, by delevering the sovereign can effectively persuade investors that he got a better liquidity shock than he actually received, which causes lenders to assign more weight to the probability that he is the good type.

I bundle these unobserved states together into a vector $u_t = [\beta_t, m_t]$. The debt stock is observed by all parties, but for expositional clarity I do not bundle it with the other observed states.

The lenders will attempt to infer the sovereign's persistent type from his borrowing and default behavior. Their beliefs at any time t can be summarized by the scalar

$$\rho_t = Pr(\beta_t = \beta^L | x^t)$$

Every period, the lenders update their beliefs using Bayes' rule according to x^t, which is the lenders' entire information set at time t, which consists of the history of all endowment realizations, borrowing and default decisions, and prices.

2.2 Timing

Investors will be given at most two opportunities to update beliefs in a given period. A sample period t takes place as follows:

1. Period t begins with the realization of observed (s_t) and unobserved (u_t) states together with some inherited debt, B_t, and current beliefs, ρ_t

2. Given $[s_t, u_t, B_t, \rho_t]$, the sovereign decides whether or not to repay B_t.

   (a) If the sovereign defaults, beliefs change to ρ_t^D via an equilibrium updating rule. The sovereign receives a default value given by $[s_t, u_t, \rho_t^D]$. Following default, the sovereign will enter financial autarky so there are no other actions he can take to signal information in period t.

(b) If the sovereign repays, beliefs change to ρ_t^R via an equilibrium updating rule. The sovereign's period t debt issuance problem thus assumes the state $[s_t, u_t, B_t, \rho_t^R]$. The choice of debt issuance further reveals information, changing ρ_t^R to ρ_t^{R+A} via an equilibrium updating rule.

3. End-of-period lender beliefs update to ρ_{t+1} from either ρ_t^D or ρ_t^{R+A} via the expected trajectory of the Markov process for β_t.

2.3 Sovereign

The sovereign borrower suffers from limited commitment. In period t, he cannot commit to either borrowing or default behavior in period $t+1$. He has a time-separable utility function, $u(\cdot)$, and is a monopolist in his own debt market. Debt is potentially long-term and matures with probability λ while issuing a fraction κ of coupons regardless of maturity.

The sovereign begins each period by making a default-repayment decision. His value at period t can be expressed as follows

$$
\begin{aligned}
V_t(s_t, u_t, \rho_t, B_t) &= \max\{V_{R,t}(s_t, u_t, \rho_t^R, B_t), V_{D,t}(s_t, u_t, \rho_t^D)\} \qquad (1)\\
\text{s.t.} \quad \rho_t^R &= G_{R,t}(s_t, \rho_t, B_t)\\
\rho_t^D &= G_{D,t}(s_t, \rho_t, B_t)
\end{aligned}
$$

where $G_{R,t}$ and $G_{D,t}$ are belief updating rules that will be described momentarily. Notice that beliefs can only be updated using publicly observed information, s_t and B_t. Reputation matters for the default value because there is the possibility of re-entry at some point, as will be described shortly.

Debt is short-term and pays a coupon, κ in the next period, implying that the Bellman equation conditional on repayment of the current debt stock, B_t is given by

$$V_{R,t}(s_t, u_t, \rho_t^R, B_t) = \max_{B_{t+1}\in\mathcal{B}} u\left(C_t - m_t\right) + \beta_t E_t \left[V_{t+1}(\tilde{s}_{t+1}, \tilde{u}_{t+1}, \rho_{t+1}, B_{t+1})\right]$$
$$\text{s.t. } C_t = Y_t - (1+\kappa)B_t + q_t(B_{t+1}|s_t, \rho_t^R, B_t)B_{t+1} \tag{2}$$
$$\rho_{t+1} = p_\beta + (1-2p_\beta)G_{A,t}(s_t, \rho_t^R, B_t, B_{t+1})$$

$G_{A,t}$ is a belief updating rule that takes into account information revealed at issuance. Notice that belief updating also accounts for the expected trajectory of the Markov process between periods t and $t+1$.

If the sovereign defaults he is excluded from credit markets temporarily, receiving a value $V_{D,t}$, which entails no liquidity shocks (m_t) but a state-contingent, weakly positive default cost $\psi(Y_t)$. There will be exogenous and stochastic re-entry with recovery following Arellano (2008). In particular, during the period of exclusion the sovereign faces a Poisson probability of market re-access. Upon re-entry, debt levels are set to zero.

The Bellman for default is thus given by

$$V_{D,t}(s_t, u_t, \rho_t^D) = u(Y_t - \psi(Y_t)) + \tag{3}$$
$$\beta_t E_t \left[(1-\phi)V_{D,t+1}(\tilde{s}_{t+1}, \tilde{u}_{t+1}, \rho_{t+1}) + \phi V_{t+1}(\tilde{s}_{t+1}, \tilde{u}_{t+1}, \rho_{t+1}, 0)\right]$$
$$\rho_{t+1} = p + (1-2p)\rho_t^D$$

I follow Chatterjee and Eyigungor (2012) and assume that default costs must be always weakly positive and that $\psi(Y_t) = \max\{0, \psi_1 Y_t + \psi_2 Y_t^2\}$. Notice that, once in default, the type and belief process evolve exogenously with no chance for sovereign actions to influence beliefs.

I finally assume that the sovereign defaults if it is not possible for him to raise enough revenue to satisfy liquidity needs.

2.4 Foreign Lenders

Foreign lenders are competitive and risk-neutral. They arrive in overlapping generations with a wealth w_t and have access to a gross risk-free return of R against which they price default risk.

Upon arriving in an observable state, (s_t, B_t, ρ_t^R), they observe aggregate issuance, B_{t+1}, and then solve the following problem.

$$\max_{b_{t+1}} \; E_t\left[\tilde{c}_{L,t+1}\right] \tag{4}$$

$$\text{where} \;\; c_{L,t+1} = (w_t - q_t b_{t+1})R + b_{t+1}\left[\rho_{t+1}R^L(s_{t+1}, m_{t+1}, \rho_{t+1}, B_{t+1}) + (1-\rho_{t+1})R^H(s_{t+1}, m_{t+1}, \rho_{t+1}, B_{t+1})\right]$$

$$\text{and} \;\; R^i(s_{t+1}, m_{t+1}, \rho_{t+1}, B_{t+1}) = \big(1 - D_{t+1}(s_{t+1}, (\beta^i, m_{t+1}), \rho_{t+1}, B_{t+1})\big) \times [1+\kappa]$$

$$\rho_{t+1} = p + (1-2p)G_{A,t}(s_t, \rho_t^R, B_t, B_{t+1})$$

R^i is defined to be the gross return function given a hidden type $i \in \{L, H\}$.

2.5 Belief Updating

In equilibrium, lenders update their beliefs regarding β every period in response to sovereign borrowing and default behavior. They do so optimally, i.e., according to Bayes' rule.

I begin with default behavior. Denote the default policy by $D_t(s_t, u_t, \rho_t, B_t)$. I define a subset of the domain of the m-shock by $M_D^i(s_t, \rho_t, B_t) = \{m | D_t(s_t, (\beta_i, m), \rho_t, B_t) = 1\}$. $G_{D,t}(s_t, \rho_t, B_t)$ describes the immediate window in which investors update their beliefs following a default: The first argument is the current beliefs were the sovereign to repay and the second is the current level of debt.

It is necessary to specify off-equilibrium beliefs, which cannot by nature be subject to Bayes' rule. For plausibility, I assume that all off-equilibrium default results in an immediate assignment of the type to bad in the eyes of the lender. All on-equilibrium default is governed by Bayes' rule. This leads to the following belief-updating rule.

$$G_{D,t}(s_t, \rho_t, B_t) = \tag{5}$$
$$\begin{cases} 1, \quad M_D^i(s_t, \rho_t, B_t) \text{ empty for } i \in \{L, H\} \\ \dfrac{\rho_t \int_{m \in M_D^L(s_t, \rho_t, B_t)} f_{\beta_L}(m)dm}{\rho_t \int_{m \in M_D^L(s_t, \rho_t, B_t)} f_{\beta_L}(m)dm + (1-\rho_t)\int_{m \in M_D^H(s_t, \rho_t, B_t)} f_{\beta_H}(m)dm}, \qquad o/w \end{cases}$$

I define a symmetric function for the case when the sovereign repays, noting that the repayment set over m will be the complement of the default set, i.e., $M_D^{i,C}(s_t, \rho_t, B_t)$. The only difference here will be off-equilibrium beliefs. If the sovereign ever repays when both types prescribe default for all possible unobserved states, then lenders instantly assume him to be the good type.

$$G_{R,t}(s_t, \rho_t, B_t) = \tag{6}$$
$$\begin{cases} 0, \quad M_D^{i,C}(s_t, \rho_t, B_t) \text{ empty for } i \in \{L, H\} \\ \frac{\rho_t \int_{m \in M_D^{L,C}(s_t,\rho_t,B_t)} f_{\beta_L}(m)dm}{\rho_t \int_{m \in M_D^{L,C}(s_t,\rho_t,B_t)} f_{\beta_L}(m)dm + (1-\rho_t)\int_{m \in M_D^{H,C}(s_t,\rho_t,B_t)} f_{\beta_H}(m)dm}, \qquad o/w \end{cases}$$

Now for the Bayesian response to borrowing. To define this, let $A_t(s_t, u_t, \rho_t^R, B_t)$ be the sovereign's borrowing policy function in equilibrium. For a given (s_t, ρ_t^R, B_t), define $M^i(s_t, \rho_t^R, B_t, B_{t+1}) = \{m | A_t(s_t, (\beta_i, m), \rho_t^R, B_t) = B_{t+1}\}$. In states of repayment, one can define the following function:

$$G_{A,t}(s_t, \rho_t^R, B_t, B_{t+1}) = \tag{7}$$
$$\begin{cases} 1, \quad B_{t+1} > A_t(s_t, (\beta_i, m), \rho_t^R, B_t) \ \forall m, \ i \in \{L, H\} \\ 0, \quad B_{t+1} < A_t(s_t, (\beta_i, m), \rho_t^R, B_t) \ \forall m, \ i \in \{L, H\} \\ \frac{\rho_t^R \int_{m \in M^L(s_t,\rho_t^R,B_t,B_{t+1})} f_{\beta_L}(m)dm}{\rho_t^R \int_{m \in M^L(s_t,\rho_t^R,B_t,B_{t+1})} f_{\beta_L}(m)dm + (1-\rho_t^R)\int_{m \in M^H(s_t,\rho_t^R,B_t,B_{t+1})} f_{\beta_H}(m)dm}, \qquad o/w \end{cases}$$

Belief responsiveness to off-equilibrium behavior is intuitive: Off-equilibrium overborrowing is instantly perceived as the bad type while off-equilibrium underborrowing is instantly perceived as the good type. Given the relative propensities of the two types to borrow, this will lead to continuity in $G_{A,t}$.

2.6 Market Clearing

The debt market is the only active market here, so the market clearing condition during periods of repayment is simply that

$$\underbrace{B_{t+1}}_{Sovereign\ Issuance} = \underbrace{B_{t+1}^D(s_t, \rho_t, B_t, B_{t+1})}_{New\ Lender\ Demand} \tag{8}$$

The price of bonds, q_t, adjusts to ensure this equation holds.

2.7 Equilibrium Definition

A **Markov Perfect Bayesian Equilibrium (MPBE)** is a set of value, policy, belief updating, and price functions such that Equations (1)-(8) are satisfied.

3 The Pricing of Default Risk

Before analyzing the quantitative results, it will be worthwhile to discuss how exactly default risk is priced in this model. We consider first the full-information case and then how asymmetric information changes things. In the full-information case, the risk-neutral lender's problem (Equation 4 with no hidden types) implies the following no-arbitrage condition, as it does in all such models:

$$q_t = \frac{1+\kappa}{R} \times E_t\left[1 - \tilde{D}_{t+1}\right] \tag{9}$$

where $\tilde{D}_{t+1}$ is binary default variable. We can define the spread in this environment as follows

$$s_t = 1 - \frac{q_t R}{1+\kappa} = E_t\left[\tilde{D}_{t+1}\right] \tag{10}$$

Notice that the conditional spread in any state at time t must equal the expected default risk in $t+1$.

Using the law of iterated expectations, we can take a further average of each side conditioning only on the fact that the country is issuing. I denote such an expectation by E_0. When there is *no* asymmetric information, this operation is straightforward and we arrive at the usual equality between average spreads and unconditional default frequencies.

$$\text{Full Information:} \quad E_0\left[\tilde{s}_t\right] = E_0\left[\tilde{D}_{t+1}\right] \tag{11}$$

Things are more complicated in the asymmetric information case. This is because the operation E_0 averages over the *lenders' known states* according to the *stationary distribution.* This is not a well-defined transformation of the general default function (as it is in the full-information case), as default is also type-specific. Thus, the analogous expression for Equation 11 is instead

$$\text{Asymmetric Information:} \quad E_0\left[\tilde{s}_t\right] = E_0\left[E_t\left[\tilde{D}_{t+1}|\beta_L\right]\tilde{\rho}_{t+1} + E_t\left[\tilde{D}_{t+1}|\beta_H\right](1-\tilde{\rho}_{t+1})\right] \tag{12}$$

One can unpack this a bit more to see concretely how default risk is now priced, paying careful attention to the potential interaction between reputation and the other states in the stationary distribution.[5]

$$E_0\left[\tilde{s}_t\right] = \tag{13}$$

$$E_0\left[\tilde{D}_{t+1}|\beta_H\right] + \left[\left(E_0\left[\tilde{D}_{t+1}|\beta_L\right] - E_0\left[\tilde{D}_{t+1}|\beta_H\right]\right)\right]E_0[\rho_{t+1}] + Cov_0\left(E_t\left[\tilde{D}_{t+1}|\beta_L\right] - E_t\left[\tilde{D}_{t+1}|\beta_H\right], \tilde{\rho}_{t+1}\right)$$

Compare this to the *true* default frequency. To see this, we denote the average over observed *and unobserved* states conditional on issuance according to their stationary distribution by $E_{true,0}$.

$$E_{true,0}[\tilde{D}_{t+1}] = E_{true,0}\left[E_{\beta_L,0}\left[\tilde{D}_{t+1}|\beta_L\right]\tilde{\mathbf{1}}\{\beta_L\} + E_{\beta_H,0}\left[\tilde{D}_{t+1}|\beta_H\right]\left(1-\tilde{\mathbf{1}}\{\beta_L\}\right)\right] \tag{14}$$

Importantly, the average over the default behavior conditional on type uses a different expectation than it does in Equation 13. It is the average of default behavior over the stationary distribution *conditional on a particular type* as opposed to *unconditionally* across both types. As such, I employ a different notation: $E_{\beta_L,0}$ for the low type and $E_{\beta_H,0}$ for the high type.

This difference between $E_{\beta_L,0}$ and E_0, for instance, is significant because the bad type tends to borrow more than the good type. This implies that $E_{\beta_L,0}\left[\tilde{D}_{t+1}|\beta_L\right] \geq E_0\left[\tilde{D}_{t+1}|\beta_L\right]$; not because the default policy is any different, but because default frequencies increase with average indebtedness.[6] This is the fundamental source of what appears to be underpricing of default risk: Investors must price default risk according to the unconditional stationary distribution for both types rather than conditional distribution for either type. This reduces average expected default for the bad type significantly.

This discrepancy in expectations could be undone by the covariance term in Equation 13 (indeed that is why it shows up). If it is the case that bad reputations always accompany heightened default risk, then

[5] It is worth observing that generally $E_0[\tilde{\rho}_{t+1}] \neq 0.5$. While completely unconditionally this equality must hold, when we condition on issuance we are conditioning on the sovereign having access to credit markets. As β_L will typically default more than β_H, β_H is more likely to be the one with access to credit markets, i.e., $E_0[\tilde{\rho}_{t+1}] \leq 0.5$.

[6] The reverse will also be true of the good type, but the difference will typically be much smaller (often truncated zero) as the good type rarely defaults to begin with.

average spreads would rise back toward the true default frequency. However, if the bad type engages in additional deleverageing during debt crises and deliberately tries to improve his reputation to curry better prices (which he will do), then that covariance term will be attenuated. Kamenica and Gentzkow (2011) note that such persuasion is possible even when fully rational receivers of signals (here, the lenders) are aware that it is happening.

This technical description of the differences between $E_0\left[\tilde{s}_t\right]$ and $E_{true,0}\left[\tilde{D}_{t+1}\right]$ can be understood in a fairly intuitive way. The distribution of spreads and default risk is highly skewed: Most of the time they are both at or near zero. In those rare instances when they are not near zero, it will typically be the bad type causing the trouble. However, the bad type will employ a costly deleverageing strategy to convince lenders that he is not as risky as he is and thus persuade them to offer better prices. This implies that in those few instances that drive both the right-tail risk and the average, spreads will be less than the implied default risk.

Figure 1 depicts how investors update their beliefs following the issuance of debt. We observe that it is increasing in debt issuance, which allows the sovereign the capacity to delever to improve his reputation in almost any state and regardless of type. It also transitions smoothly into the prescribed off-equilibrium beliefs.

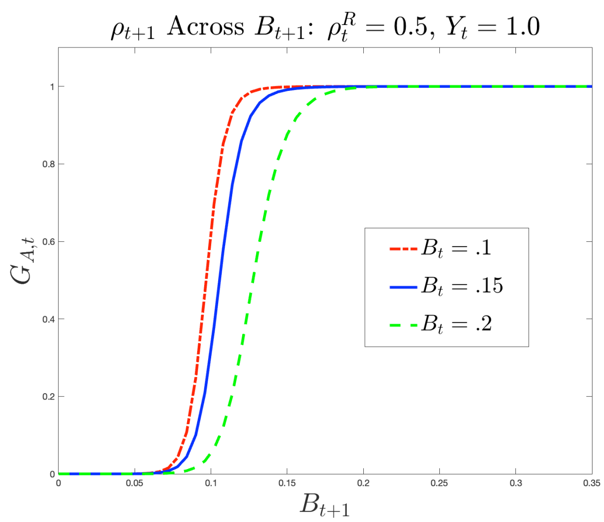

Figure 1: Beliefs Following Borrowing

In the next section, we will explore numerically a simple example in which this sort of 'underpricing' can arise.

4 Quantitative Implementation

When solving for an equilibrium, I always compute the MPBE that is the limit of the finite-horizon game as in Hatchondo and Martinez (2009). While multiple MPBEs may exist, this mechanism will allow us to select consistently from among them.

4.1 Solution and Calibration

I solve the model using machine learning techniques. In particular, I employ the Gaussian Process Dynamic Programming algorithm described in Gu and Stangebye (2023), which builds on work by Scheidigger and Bilionis (2019).[7] I iterate on the finite-horizon game until stationarity in the value functions is achieved via convergence on the order of $1.0e-7$.

The goal is not to match a particular empirical regularity but rather to seek out a plausible calibration in which 'underpricing' occurs. As such, the calibration will be guided by proximity to the extant literature rather than determined by a moment-matching exercise. The model is calibrated at a quarterly frequency with the parameters found in Table 1.

Table 1: Parameter Calibration

Parameter	Value	Source
Sovereign $u(\cdot)$	$u(C) = C^{1-\gamma}/(1-\gamma)$ with $\gamma = 2$	Arellano (2008)
R	1.017	Arellano (2008)
(ψ_1, ψ_2)	$(-.956, .998)$	Arellano (2008)
κ	$R - 1$	Aguiar et al. (2016)
ϕ	.083	Mendoza and Yue (2012)
(β_L, β_H)	$(.93, .975)$	Proximity to Arellano (2008)
(ρ_z, σ_z)	$(.9212, .0226)$	Gu and Stangebye (2023)
σ_m	.005	Unobservable shocks $\approx 1/4$ size of observable shocks
p_β	.025	Types change $\approx$ 10 years

Only a few parameters are worth some discussion. Arellano (2008) employs a default cost function wherein defaulting with output below a certain threshold is costless and defaulting above this threshold reduces output to the level of the threshold. The chosen parameters (ψ_1, ψ_2) engender very nearly the same

[7] The only relevant deviation is the following: Gu and Stangebye (2023) assume a constant lengthscale across all states, while here I allow for the reputation state (ρ_t) to have a potentially different lengthscale than the other states. The algorithm could thus find reputation to be less, more, or as important as other states. Endogenously, I find as a result that it is less important.

dynamic in the smoother cost function of Chatterjee and Eyigungor (2012). In particular, they imply that default is costless when $Y_t \leq 0.96$ and that output is nearly flat at .96[8] for all $Y_t > 0.96$.[9] We also select the discount factors to be modestly below and above that of Arellano (2008), in which the constant discount factor is $\beta = 0.953$.

I take the duration of default from Mendoza and Yue (2012) rather than Arellano (2008), who sets it to be shorter. A longer default duration makes for more costly default, which will allow for slightly more borrowing. I find that this helps generate countercyclical net exports and spreads, which are empirically relevant. I follow Aguiar et al. (2016) and set the coupon to be equal to the lenders' discount factor, which implies that a risk-free bond will always have a unity price.

The output process is taken from Gu and Stangebye (2023), who employ a bandpass filter on Russian output data at business cycle frequencies. The resulting process very closely resembles that used by Arellano (2008) or Chatterjee and Eyigungor (2012), who calibrate to Argentine output data from the 1990s.

The introduction of the two new parameters in the model, i.e., unobservable liquidity shock volatility and the probability of switching types, are disciplined as follows. σ_m is set such that it is non-trivial in size but also significantly smaller than observable output shocks (σ_z). This jointly ensures that the investor's inference problem is non-trivial and that the model's aggregate dynamics closely resemble its predecessors in the literature. p_β is set such that types change once every ten years on average.

4.2 Key Numerical Results

The headline results can be observed in the ergodic moments implied by the model. These can be found in the Column 1 of Table 2.[10] We first observe that average spreads, default frequency, and debt-to-GDP are in the rough empirical range for a wide variety of emerging markets (see Aguiar et al., 2016). We also observe that spreads and the trade share are countercyclical, which is another prominent feature of emerging markets generated by this class of models.

[8] Arellano (2008) sets this value at 0.969. The model does not converge to the required tolerance at exactly this value, but converges for the nearby value here of 0.96. This difference is immaterial quantitatively.

[9] In particular, it ensures that output three unconditional standard deviations above the steady state, which is roughly 1.1903, is 0.23 lower in default relative to repayment.

[10] Of the many columns in Table 2, only two are proper equilibria: Columns 1 and 4 (both labeled "Benchmark"), which correspond to the asymmetric information and full information cases.

Table 2: Simulated Moments (Conditional on Issuance)

Moment	Asymmetric Information			Full Information		
	Benchmark	L-Only	H-Only	Benchmark	L-Only	H-Only
Average Debt/GDP	13.6%	18.8%	9.3%	8.6%	12.1%	5.4%
Annual Default Rate	2.8%	7.0%	0.2%	0.3%	.8%	0.0%
Average Annualized Spread	0.6%	1.3%	0.4%	0.2%	.4%	0.0%
Annualized Spread Volatility	2.0%	2.9%	0.3%	0.9%	1.4%	0.0%
Spread Cyclicality	−.05	−.11	−.03	−.06	−.10	.00
Trade Share Cyclicality	−.19	−.17	−.23	−.18	−.21	−.15

Table 2 reveals two significant consequences of asymmetric information relative to full-information.

4.2.1 Heightened Default Risk

First, there is a significant increase in default behavior for both types. This can be seen by comparing default frequencies in Columns 2 and 3 to those in Columns 5 and 6, all of which give counterfactual default frequencies by assuming the type never changes in a long-form simulation.

There are two forces at work driving this. First, the good type is significantly worse off under asymmetric information relative to full information. The reason is because, as described in Stangebye and Wright (2024), his budget set contracts toward that of the bad type. He is not 'rewarded' for being the good type with better prices and thus chooses to default more often. This can be seen in certainty-equivalent-consumption figures, which are almost always worse for the good type relative to the case of full-information (see Stangebye and Wright, 2024). Because the good type is almost always worse off, he defaults more. In this case, he actually goes from a situation of never defaulting at all to one in which he defaults with strictly positive probability on the equilibrium path.

The second force driving this result is that the bad type is given more freedom to borrow. Lenders rarely believe fully that he is the bad type, and so his budget set is often slightly larger than it is under full information. The bad type often takes advantage of this, borrowing into high-debt regions of the state-space where default is more likely and defaulting more often as a consequence. This latter effect is quite strong quantitatively, causing default frequencies to increase by more than an order of magnitude relative to full-information (see Columns 2 and 5 of Table 2).

4.2.2 'Underpricing' of Default Risk

The second major consequence of asymmetric information is that default risk is underpriced in the sense that unconditional default frequencies are significantly larger than average spreads. We discussed in Section 3 why these two are generally tied together under full information and why they might deviate in under asymmetric information. Columns 1 and 4 in Table 2 clearly show this to be the case. Under full information, the default rate and its associated spread are roughly equal to each other, while under asymmetric information the former is nearly 5× the latter.

We can use the decomposition in Equation 13 to unpack why this is the case for this particular calibration. Table 2 gives annualized moments, but since the model is quarterly we construct the decomposition at that frequency as follows:

$$\underbrace{E_0\left[\tilde{s}_t\right]}_{.0014} = \tag{15}$$
$$\underbrace{E_0\left[\tilde{D}_{t+1}|\beta_H\right]}_{.0009} + \underbrace{\left[\left(E_0\left[\tilde{D}_{t+1}|\beta_L\right] - E_0\left[\tilde{D}_{t+1}|\beta_H\right]\right)\right]}_{.0016-.0009}\underbrace{E_0[\rho_{t+1}]}_{.34} + \underbrace{Cov_0\left(E_t\left[\tilde{D}_{t+1}|\beta_L\right] - E_t\left[\tilde{D}_{t+1}|\beta_H\right], \tilde{\rho}_{t+1}\right)}_{.0003}$$

whereas for the true default frequency, we get the following decomposition of Equation 14.

$$\underbrace{E_{true,0}[\tilde{D}_{t+1}]}_{.0087} = \underbrace{E_{\beta_H,0}\left[\tilde{D}_{t+1}|\beta_H\right]}_{.0005} + \underbrace{\left[E_{\beta_L,0}\left[\tilde{D}_{t+1}|\beta_L\right] - E_{\beta_H,0}\left[\tilde{D}_{t+1}|\beta_H\right]\right]}_{.0180-.0005} \times \underbrace{E_{true,0}\left[\mathbf{1}\{\beta_L = \beta\}\right]}_{0.47} \tag{16}$$

We derive the conditional default frequencies for Equation 16 from Columns 2 and 3 of Table 2, which counterfactually simulates the model under the assumption that the type never switches.[11]

We observe immediately that at both the quarterly and annual frequency the unconditional default frequency is more than 5× the annual spread. This is driven by two forces. First, the difference between the average default frequency of the bad type *in the unconditional stationary distribution*, i.e., as given in Equation 15, and the default frequency of the bad type *in the conditional stationary distribution implied by his own behavior*, i.e., as given in Equation 16. This difference is driven by the fact that debt levels are

[11] Annualizing this default rate gives a default rate slightly over 3%. This is just mildly higher than the benchmark simulation of a default rate just below 3%. This discrepancy is due to small sampling error in the translation of the ergodic to the stationary distribution across multiple counterfactual models for a very rare event such as default. The average spread, which corresponds to an object realized in every period in a single long-form simulation, features less sampling error and is annualized to 0.6% in Equation 15 and in Column 1 of Table 2.

much higher in the latter, and thus there is much more default (a .018 quarterly rate instead of .0016).

Second, the bad type successfully hides the fact that he is the bad type, especially during a debt crisis. This manifests itself most immediately in the fact that reputations tend to be better than the reality on average, i.e., $E_0[\rho_{t+1}] < E_{true,0}\left[\mathbf{1}\{\beta_L = \beta\}\right]$. But it also manifests itself in the covariance term in Equation 15, is nearly zero. If it were positive, this term would more closely connect bad default behavior with high spreads by souring reputations when risk is high. The fact that it is near zero reveals that reputations do not go hand-in-hand with default risk, especially during debt crises. This implies that the bad type deliberately (and successfully) tries to hide his true identity during high-risk states to curry better prices.

The underpricing dynamics make for interesting debt crises. Figure 2 depicts the most typical (highest likelihood) default in the simulation. In the early stages, the bad type has obscured his type so well with reduced borrowing that investors are convinced that he is the good type. This engenders better prices until very nearly the moment of default, when he eventually gives his type away.

4.3 Robustness

It is worth noting that the underpricing phenomenon is reliant on (1) the differences between the default frequencies of the two types, (2) the differences in the borrowing behavior of the two types, and (3) the dynamics of persuasion. These three do not always hold in conjunction for all plausible parameterizations of this model. An example would be the exact calibration of Arellano (2008). The differences between this and the benchmark model are not substantial but they do affect the underpricing result. They are as follows.

- The output process is slightly more persistent and volatile: $\rho_z = 0.945$ and $\sigma_z = 0.025$
- The coupon is zero, i.e., $\kappa = 0$.
- The expected duration of exclusion is shorter, i.e., $\phi = 0.282$.

We maintain the benchmark values for the discount factor, which are symmetrically placed around those in Arellano (2008). The results can be found in Table 3.

With this calibration, we find an annual default rate of 4.0% and an average annualized spread that is much nearer (though still less) of 3.4%. Thus, the underpricing dynamics are muted. The primary

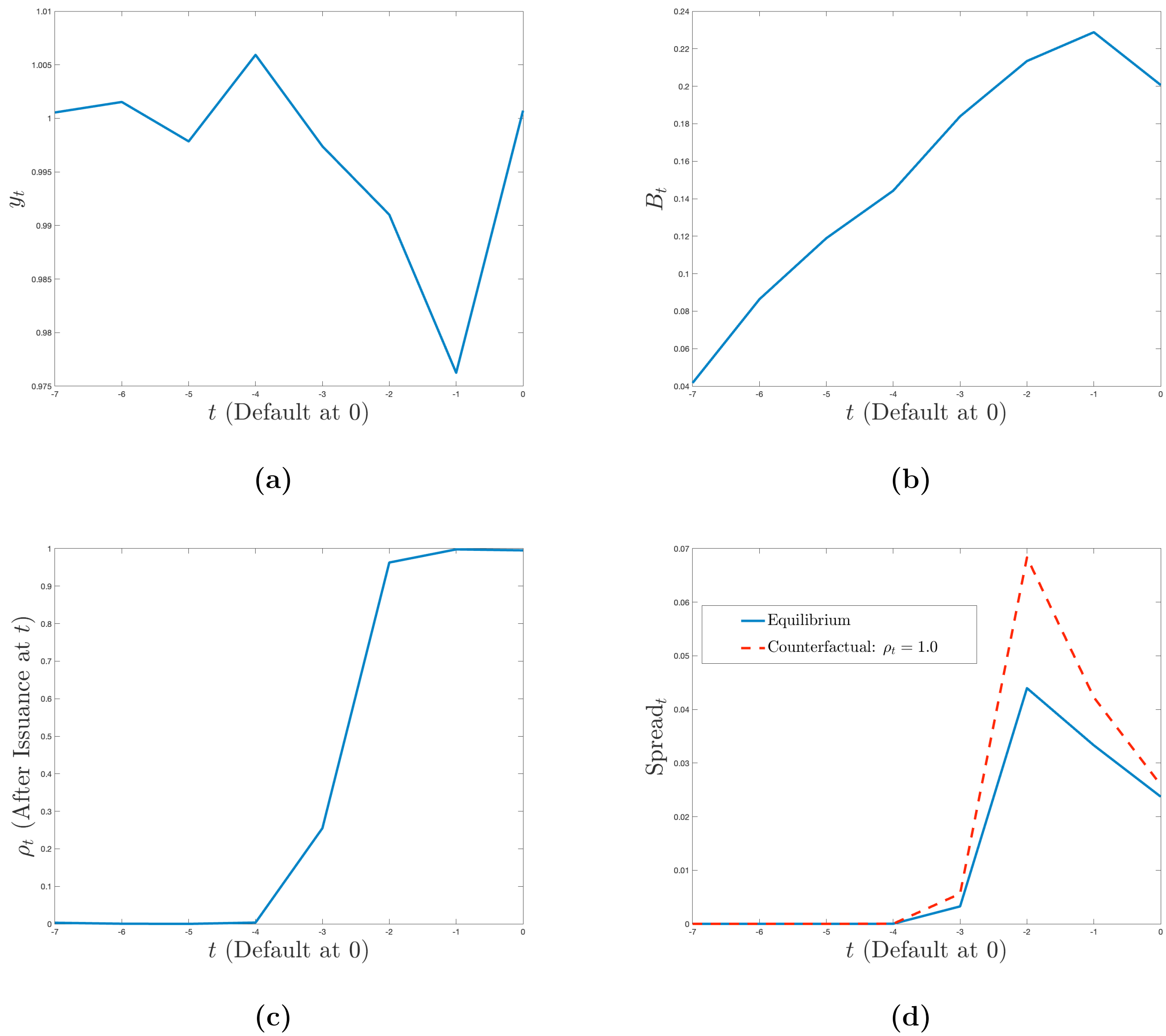

Figure 2: Dynamics of a Typical Default

reason for this is the difference in the output process. Relative to the benchmark calibration, output is more persistent and more volatile. This implies that the sovereign has larger booms and busts relative to the benchmark *and* the surprise output shocks are typically larger. This induces more default risk, which works to both reduce overall indebtedness (lenders are not willing to lend) and increase default frequencies for both types. The tighter budget set for both types prevents the bad type from borrowing significantly more than the good type, a key feature in driving underpricing for the benchmark calibration. The greater

Table 3: Simulated Moments (Conditional on Issuance) for Robustness

	Asymmetric Information			Full Information		
Moment	Benchmark	L-Only	H-Only	Benchmark	L-Only	H-Only
Average Debt/GDP	3.9%	4.2%	3.7%	1.8%	2.2%	1.4%
Annual Default Rate	4.0%	6.1%	2.1%	0.1%	0.1%	0.0%
Average Annualized Spread	3.4%	4.9%	2.0%	0.0%	0.0%	0.0%
Annualized Spread Volatility	23.1%	31.3%	12.5%	0.3%	0.4%	0.0%
Spread Cyclicality	.10	.10	.13	.07	.09	.03
Trade Share Cyclicality	−.05	−.06	−.05	−.06	−.06	−.06

default frequency for the good type also works against the underpricing, as it highlights the fact that default is driven largely by exogenous shocks rather than the sovereign's location in the (type-specific) ergodic distribution.

It is worth noting that even in this robustness exercise, we still find heightened default risk relative to the full-information model.[12] In the full-information model, the annualized default rate is a much lower 0.1% unconditionally. Thus, while the underpricing dynamic is muted, the heightened default risk dynamic is certainly not.

References

[1] **Aguiar, Mark and Gita Gopinath**, "Defaultable Debt, Interest Rates, and the Current Account," *Journal of International Economics*, 2006, *69* (1), 64–83.

[2] **___ , Satyajit Chatterjee, Harold Cole, and Zachary R. Stangebye**, "Quantitative Models of Sovereign Debt Crises," *Handbook of Macroeconomics, Volume II (Forthcoming)*, 2016.

[3] **Amador, Manuel and Christopher Phelan**, "Reputation and Sovereign Default," *Econometrica*, 2021, *89* (4), 1979–2010.

[4] **Arellano, Cristina**, "Default Risk and Income Fluctuations in Emerging Economies," *American Economic Review*, 2008, *98* (3), 690–712.

[5] **Chatterjee, Satyajit and Burcu Eyigungor**, "Maturity, Indebtedness, and Default Risk," *American Economic Review*, 2012, *102* (6), 2674–2699.

[12]Neither the average spreads nor the default frequencies in any of Columns 4-6 of Table 3 are true zeros, but just rounded to the nearest tenth of a percent annually.

[6] ——, **Dean Corbae, Kyle Dempsey, and Jose-Victor Rios-Rull**, "A quantitative theory of the credit score," *Econometrica*, 2023, *91* (5), 1803–1840.

[7] **Cole, Harold L. and Patrick J. Kehoe**, "Models of Sovereign Debt: Partial Versus General Reputations," *International Economic Review*, 1998, *39* (1), 55–70.

[8] **D'Erasmo, Pablo**, "Government Reputation and Debt Repayment," *Manuscript, University of Texas Austin*, 2008.

[9] **Fourakis, Stelios**, "Sovereign Default and Government Reputation," *Mimeo*, 2023.

[10] **Gu, Grace Weishi and Zachary R. Stangebye**, "Costly information and sovereign risk," *International Economic Review*, 2023.

[11] **Hatchondo, Juan Carlos and Leonardo Martinez**, "Long-Duration Bonds and Sovereign Default," *Journal of International Economics*, 2009, *79* (1), 117–125.

[12] **Kamenica, Emir and Matthew Gentzkow**, "Bayesian Persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615.

[13] **Mendoza, Enrique G. and Vivian C. Yue**, "A General Equilibrium Model of Sovereign Default and Business Cycles," *Quarterly Journal of Economics*, 2012, *127* (2), 889–946.

[14] **Phan, Toan**, "Sovereign Debt Signals," *Journal of International Economics*, 2017, *104*, 157–165.

[15] **Sandleris, Guido**, "Sovereign Defaults: Information, Investment and Credit," *Journal of International Economics*, 2008, *76* (2), 267–275.

[16] **Scheidigger, Simon and Ilias Bilionis**, "Machine Learning for High-Dimensional Dynamic Stochastic Economies," *Journal of Computational Science*, 2019, *33*, 68–82.

[17] **Stangebye, Zachary R. and Mark L.J. Wright**, "Living in Infamy: Bad Reputations in Emerging Markets," *Mimeo*, 2024.