COSTLY INFORMATION AND SOVEREIGN RISK *

April 14, 2021

Grace Weishi Gu† and Zachary R. Stangebye‡

The consequences of costly information acquisition for sovereign risk are explored in a quantitative sovereign default model. We identify information costs empirically using Bloomberg news-heat data. The calibrated model microfounds heteroskedasticity in the country risk spread: Crises are endogenously more volatile because more information is acquired and priced. Because effective risk tolerance depends on the information set, the model also suggests a novel layer of state-contingency in the risk premium. Welfare effects are small but non-monotone, as greater transparency can make the country mildly worse off by inducing more price volatility and default.

Keywords: costly information; sovereign default; heteroskedasticity; time-varying volatility; default risk inference; transparency

JEL code: F34, D83, G12

*The authors would like to thank Urban Jermann, Manuel Amador, Mark Aguiar, Sebnem Kalemli-Ozcan, Boragan Aruoba, Luminita Stevens, Felipe Saffie, John Shea, Juan Carlos Conesa, Marina Azzimonti, Ting Liu, Raoul Minetti, Charles Engel, Kamran Bilir, Kenneth West, Leonardo Martinez, Sandra Lizarazo, Ruediger Bachmann, Olivier Jeanne, Michal Szkup, Mark Wright, Marco Bassetto, Linda Tesar, Pablo Ottonello, Gita Gopinath and seminar participants at the University of Maryland, Michigan State University, Stony Brook University, the University of Pittsburgh, the University of Wisconsin, the International Monetary Fund, the Federal Reserve Bank of Chicago, the University of Michigan, and the NBER Summer Institute for useful comments and suggestions. All errors and mistakes are our own.

†E-mail: grace.gu@ucsc.edu, Phone: (831) 459-4791, Address: 463 Engineering 2, 1156 High Street, Santa Cruz, CA 95064

‡E-mail: zstangeb@nd.edu, Phone: (574) 631-1054, Fax: (574) 631-1054, Address: 3015 Nanovic Hall, Notre Dame, IN 46556 (Corresponding Author)

1. Introduction

Yields in sovereign bond markets in emerging economies largely reflect the risk that a domestic borrower may default on foreign creditors since it does not in principle have any reason to care about the well-being of these creditors. But a sovereign borrower's lack of welfare concern for its foreign lenders is not the only relevant friction that arises from the international nature of these markets: Information frictions also play a large role in cross-border financial transactions: Foreign investors are likely to be less informed about payoff-relevant shocks in emerging/developing countries (Hatchondo [2004], Van Nieuwerburgh and Veldkamp [2009], or Bacchetta and van Wincoop [2010]). For instance, Bloomberg publication and readership data suggest that only 30.5% of articles published pertain to emerging and developing economies, despite the fact that 47 out of 61 countries in their dataset fall under this category.[1]

This is true even for institutional investors, as few developing countries fully adopt International Public Sector Accounting Standards (IPSAS) or other international standards for government finance reporting. Further, their sovereign bonds are usually traded over-the-counter with limited information and thin trading volumes relative to assets traded in exchanges, such as stocks. Nevertheless, payoff-relevant information can often be acquired at a cost. For instance, institutions can spend resources to harmonize data across countries despite different accounting standards and can collect and analyze market trading data down to seconds, which then can be packaged as information products to sell to investors.

In this paper we seek to understand how costly information acquisition affects markets for sovereign bonds and default risk. To do so, we construct a model in which the sovereign's default and borrowing decisions as well as the lenders' receipt of payoff-relevant information are jointly endogenous: Lenders always observe some public states, such as output and debt levels, but cannot directly observe other potentially payoff-relevant states when making investment decisions, such as the recovery rate in the event of default or the severity of the

[1] The sample period for this database is 2008-2016.

recession implied by default in the borrowing country. Information regarding these sorts of shocks can only be acquired at a cost.

In any market-based model of costly information acquisition, one encounters the dilemma highlighted by Grossman (1976), which is that market prices tend to reveal too much information and thus kill incentives to acquire information in the first place. We circumnavigate this issue by separating the information acquisition decision from market participation, much like Angeletos and Werning (2006). An independent contractor/forecaster interested only in the integrity of its forecast such as the credit rating agencies in Manso (2013) or Holden et al. (2018) conducts the information acquisition. As market participants, lenders pay a small fee to gain access to the information that is obtained by the forecaster. The usefulness of the information depends on the acquisition efforts of the forecaster in that state.

Real world analogues of the forecaster might be financial analytics firms, such as Bloomberg or Reuters; alternatively they could be international entities such as the IMF or the World Bank, or credit rating agencies such as Moody's or S&P. Our calibration strategy reflects this market framework. In addition, we further extract signal games from the model by assuming that the sovereign does not have any informational advantage over lenders at the time of bond issuance and thus cannot signal its type through bond supply. All these practices allow us to quantitatively evaluate the impact of costly information acquisition in our model.

Costly information acquisition generates a nonlinear dependence of lender knowledge regarding unobserved states on those that are publicly observed. For instance, when debt levels are low and output is high, there is little to gain in terms of accurately inferring unobserved shocks since default risk is negligible for most of their realizations. However, for moderately high debt levels and low output, information is more valuable since unobserved shocks may tip the sovereign over to a default event. Intuitively, the forecaster will start acquiring more information at the beginning of a crisis, more carefully studying the borrower and its associated default risk. This will be reflected in more precise information being transmitted to lenders. Essentially, the forecaster will condition its information acquisition behavior regarding unobserved states on those that are publicly observed.

To quantify the impact of this mechanism, we calibrate the benchmark model to Russia and select Bloomberg as our empirical counterpart for the forecaster. Bloomberg not only provides investors with information about sovereign borrowers, but also provides a novel dataset that they call 'news-heat,' that we use as a proxy for information acquisition activity. News-heat tracks Bloomberg publications on sovereign borrowers as well as reading and searching activities on Bloomberg terminals, but is only available from 2008-2016. We use Russia because in this period it experienced substantial macroeconomic volatility but without the added complications of Eurozone or E.U. membership like turbulent Greece.

The model reveals several insights. First, it generates substantial heteroskedasticity, i.e., time-varying volatility, in the country-risk spread without assuming it in any of the exogenous fundamental processes and that its novel concomitant predictions are consistent with the data. Time-variation in macroeconomic volatility is a well-documented empirical fact (Justiniano and Primiceri [2008] or Bloom [2009]) but little has been done as of yet to understand its causes,[2] despite the fact that Fernández-Villaverde et al. (2011) show that second-moment fluctuations in the country-risk spread can have substantial first-moment effects on investment and output.

The intuition behind the heteroskedasticity in our model is as follows: During normal times when debt levels are low and output is high, the forecaster has little to gain from information acquisition regarding unobserved shocks since default risk is negligible for most of their realizations. Consequently, lenders receive little to no accurate information regarding these shocks. They assume them to be at their mean when pricing default risk, which implies that bond yields do not respond to their realizations and spread volatility is lower. However, during periods with high debt levels and low output, i.e., crises, the forecaster more carefully studies the borrower and its associated default risk since unobserved shocks may substantially affect default risk. Hence, lenders acquire additional information about these shocks, which

[2]Some notable recent exceptions are Seoane (2015) and Johri et al. (2015), but these papers explain time-varying volatility in the country risk spread by assuming exogenous time-varying volatility in fundamentals. Sedlacek (2016) provides an alternative solution, but it generates time-varying cross-sectional dispersion among firms. Chari and Kehoe (2003) also show that investor herding behavior can increase uncertainty during crises.

then become priced. This implies that bond yields *do respond* to realizations of unobserved shocks during crisis times, which increases spread volatility.

We propose a model-free metric of heteroskedasticity in the country risk spread that we call the *Crisis Volatility Ratio* (CVR) and use it to demonstrate the efficacy of our model against alternative models. The CVR has a simple construction: It is the sample average of the ratios of sample volatilities surrounding precisely defined 'crisis' events. It is a scalar measure of how much more volatile a country becomes on impact in the aftermath of a crisis event. If the underlying process features no heteroskedasticity, e.g., an AR(1), the CVR will approach unity as the sample size grows. We show that Russia featured a CVR of 2.46 for our sample period of 2008-2016, which is within the typical range of an emerging market in a longer time horizon (Table 2). Our model model brings this moment significantly closer to the data ($CVR = 2.13$) than an alternate calibration without state-contingent information flows[3] ($CVR = 1.73$) or even more reduced-form models such as an AR-ARCH-M ($CVR = 1.38$).[4]

Furthermore, we use Bloomberg news-heat data to empirically test our model's predictions on sovereign spread heteroskedasticity more broadly for 28 countries (2008-2016). Using a combined panel dataset of EMBI and news-heat, we conduct regression analysis to show that increased information flows are associated with greater spread levels and spread volatility in the sovereign bond market. The effect is state-contingent and is only statistically significant during crises, which accords with the predictions of our model.

Another insight revealed by our model regards the dynamics of risk premium. In standard default models, the spread on a bond can be decomposed into two components: Default, dilution, or recovery risk and a risk premium for that risk. The latter is affected not only by investors' degree of risk aversion, but also by the conditional distribution of payoff-relevant shocks. In our model, information acquisition reduces the variance of this conditional distribution from the perspective of the investor. Consequently, it exerts downward pressure

[3] de Ferra and Mallucci (2020) find that sovereign debt models generate heteroskedasticity naturally. We corroborate this finding, but also note that the additional state-contingent volatility generated by our model is necessary to boost the CVR into the empirically relevant range.

[4] Among the GARCH-class models, the data tend to favor the AR-ARCH-M, which is why we contrast our model to this.

on the risk premium. This novel channel alters the relative composition of default risk and the risk premia in the overall spread across publicly observed states.

The calibrated model suggests that around crisis times the risk premium constitutes about 33.5% of spreads,[5] whereas a standard model without costly information acquisition would put this figure as high as 35.5%. Importantly, this effect is state-contingent and plays a more significant role during crises. While this figure is fairly small, it is highly persistent, lasting at least 12 quarters. This has implications for the inference of default risk from spread data, which is a common practice in the literature (Bi and Traum [2012], Bocola [2016], Bocola and Dovis [2019], and Stangebye [2015]). In particular, it implies that, all else being equal, a standard sovereign default model without state-contingent information flows may *understate* default or recovery risk around crises by assuming a risk premium that is too large.

Finally, our model offers a natural framework to consider the impact of transparency policy. We interpret greater transparency here as a reduction in the information costs. We find that the welfare journey from costly information to full transparency is non-monotonic, as transparency works as a double-edged sword on sovereign utility. On the negative side, it increases price volatility as information is acquired and priced more often. This exposes the risk-averse sovereign to additional risk in its budget set and causes it to default more often, which works to hurt the country. At the same time, though, the sovereign benefits from transparency due to lower risk premia.

These different forces all at work in the benchmark model and the result is a welfare trajectory that is fairly flat but also features an inverted U-shaped as information costs fall. Initially, more transparency helps the country, but eventually it causes more default and price volatility than it is worth. This result suggests that some countries may be motivated to restrict or limit transparency policies to avoid these costs.

[5] We define a 'crisis' directly from spreads following Aguiar et al. (2016).

1.1. Contributions and Related Literature

Our paper contributes to the sovereign debt and asset pricing literature in three dimensions. First, our paper studies the quantitative impact of information friction on asset pricing and sovereign welfare dynamics. There has been work done in stylized theoretical environments but not much in standard quantitative models. We are able to introduce information friction in a tractable way by calibrating information costs using the Bloomberg investor-attention measurement and by shutting down signaling games among investors themselves as well as between investors and the borrower. Second, we find three important implications of costly information to asset pricing and beyond: time-varying state-contingent spread volatility, shifts in risk premium composition, and non-monotonic sovereign welfare. Thus, we also provide a micro-foundation for time-variation in volatility. Third, our paper empirically measures investors' attention to sovereigns and provides support to our model's predictions on asset pricing dynamics for a broad set of countries. In contrast, the existing works concentrate on investors' attention on firms and their stocks and are silent on sovereigns and their bonds.

Our model's focus on relations between a single sovereign borrower and its lenders over time and the implications for the ergodic price distribution distinguishes our analysis from the related work of Cole et al. (2016). These authors also explore a model of costly information acquisition in sovereign debt markets, but their focus is static. They highlight the potential for this channel to cause contagion effects across many countries and generate multiplicity. Angeletos and Werning (2006) and Carlson and Hale (2006) also explore how market-based information acquisition or rating agencies affect equilibrium multiplicity or uniqueness in variations on the canonical model of Morris and Shin (1998). Durdu et al. (2013) explore the impact of news shocks in a similar model. Recently, Bassetto and Galli (2019) also study the role of information on sovereign bond pricing in a two-period Bayesian trading game, focusing on implications for inflation risk.

Relatedly, it is worth distinguishing our paper from the vast literature that studies

sovereign defaults and information asymmetry. For instance, Cole and Kehoe (1998), Sandleris (2008), Catao et al. (2009), Wittenberg-Moerman (2010), Hellwig et al. (2014), Pouzo and Presno (2016), Croitorov (2016), Mihm (2016), and Perez (2017) among others have all shown that information asymmetries and uncertainties are key to explaining various features of sovereign bond markets, though none have considered the consequences of allowing information to be gathered at a cost. Another important distinction is that, although our model and these other papers all assume incomplete information for investors, our model mechanism does not require any information asymmetry between the borrower and the investors or any information asymmetry among investors. The model focuses on investors' costly information acquisition behavior regardless of information asymmetry.

de Ferra and Mallucci (2020) recently show that Arellano (2008) in its unaltered form generates endogenous heteroskedasticity in spreads that align with the data for its calibration. The mechanisms behind this, namely greater bond price elasticity at higher debt levels, is also at work in our model, as is evidenced by the model without information flows generating CVRs that indicate heteroskedasticity. However, we find that the additional inclusion of costly information acquisition is required to generate CVRs in line with the data. Further, our model has novel implications for risk premium composition and transparency policy that cannot be discussed in the standard Arellano (2008) framework.

This paper is also connected to a growing literature on corporate finance, stock returns, and investor information acquisition. Dang et al. (2015) and Babenko and Mao (2018) consider optimal security design problems in the financing of a risky investment with information acquisition. Pagano and Volpin (2012) show that issuers of asset-backed securities may choose to release coarse information to enhance primary market liquidity, even if it reduces secondary market liquidity. Barber and Odean (2008), Da et al. (2011), Ben-Rephael et al. (2017) show empirically that investor attention predicts stock returns and Cziraki et al. (2018) show that the geographic distribution of attention matters for stock returns. Most similar to our work are Vlastakis and Markellos (2012), Andrei and Hasler (2015), and Dimpfl and Jank (2016) who show that an increase in investor attention corresponds to an

increase in stock-return volatility.

Our empirical analysis is closely related to the recent work by Bi and Traum (2019). They use data about fiscal information from U.S. newspapers in the 1840s and find that state government bond prices were especially sensitive to fiscal news during crisis time, relative to normal time. Their results are consistent with the broad predictions of our model and empirical results.

The remainder of this paper is divided as follows: Section 2 describes the model; Section 3 discusses the data and quantitative implementation of the model; Section 4 demonstrates the model's key results; Section 5 tests the empirical predictions of the model in a broader dataset; and Section 6 concludes.

2. Model

We consider a small open economy model of endogenous sovereign default in the vein of Eaton and Gersovitz (1981). This is in part for tractability and in part to demonstrate our model's applicability and compare results to the recent, expanding quantitative literature, e.g., Aguiar and Gopinath (2006), Arellano (2008), Hatchondo and Martinez (2009), or Mendoza and Yue (2012). There is a sovereign borrower who issues long-term non-state-contingent debt to a unit mass of foreign lenders. This borrower lacks the ability to commit to repay this debt in subsequent periods and will default if it is optimal to do so ex post. The borrower also lacks the ability to commit to future borrowing behavior.

For clarity, we distinguish a random variable from its realization by placing a tilde over the former.

2.1. Shocks

There are two shocks in this model. The first is a transitory shock to the sovereign's endowment Y_t, whose log follows an AR process, $\log Y_t = \rho_y \log Y_{t-1} + \sigma_y \epsilon_t$, where ϵ_t is a standard normal. The endowment and its shocks are publicly observed by everyone. The publicly observed exogenous state follows a Markov process with transition density

$f_t(Y_{t+1}|Y_t)$.[6]

The second shock in the model is an i.i.d recovery rate on defaulted debt, m_t.[7] m_{t+1} is not publicly observed by anyone when it is realized in the middle of period t. However, we will assume that a professional forecaster may pay a cost to observe an imperfect correlate of it, x_t, at this time. This imperfect correlate (and its correlative structure with m_{t+1}) then becomes public knowledge at the time the debt is priced. m_t is assumed to be logit-normally distributed, which allows for it to be bounded on $[0, 1]$ while maintaining a jointly normal information structure with the correlate, x_t. We also assume that once the sovereign is in a default event, the value of m stays constant throughout the event until the default regime is over.

2.2. Timing

The timing of events is as follows: Period t begins with the realization of Y_t, following which the sovereign makes a default decision with knowledge of the m_t realization. Conditional on repayment, it then chooses a level of debt issuance B_{t+1} to maximize its expected utility prior to the realization of m_{t+1}.

Next, a professional forecaster, who observes the public states Y_t and B_{t+1}, chooses the accuracy with which he acquires information about m_{t+1} given some information cost. He designs a signal of the unobserved shock, x_t, and can pay a cost to increase its accuracy.

Following the information acquisition decision, m_{t+1} and x_t are jointly realized in the middle of period t. The market coordinates on the signal: Competitive lenders know both the signal and its accuracy, and then decide bond demand. The sovereign then determines an issuance price that clears the bond market and period t ends.

Notice that we assume that the sovereign cannot change its bond supply following the realization of m_{t+1}.[8] This allows us to focus on the role of information acquisition and avoid

[6] Note that our endowment structure features transitory shocks in the vein of Arellano (2008). This is not crucial to the novel mechanism; the intuition goes through with permanent shocks to income following Aguiar and Gopinath (2006) as well.

[7] Since default events are rare in both data and our model, most of the reset recovery rates each quarter do not show in the data nor matter to the sovereign borrow or the investors in the model. What matters are those few recovery rates that are actually used during the rare default events. Hence, it is difficult to estimate the recovery rate process for every quarter and assuming the rate being i.i.d keeps the model calibration simple.

[8] We do not specify whether sovereign gets to know m_{t+1} before or after the lenders do. It does not matter, since the

the complicated and, for our purposes, unnecessary signaling game that would ensue.

2.3. Sovereign Borrower

As is standard in the literature, we use a recursive, Markov-Perfect specification with limited commitment on the part of the sovereign. At the beginning of each period, the sovereign compares the value of repaying debt, $V_{R,t}$, with that of default, $V_{D,t}$, and chooses the option that provides a greater value:

$$V_t(Y_t, B_t, m_t) = \max\{V_{R,t}(Y_t, B_t), V_{D,t}(Y_t, B_t, m_t)\}$$

The sovereign has standard time-separable concave preferences over consumption and is a monopolist in his own debt market. Given the timing assumption, we can express the value of repayment at the beginning of period t as follows:

$$V_{R,t}(Y_t, B_t) = \max_{B_{t+1} \in \mathcal{B}_t} \; E_{\tilde{m}_{t+1}, \tilde{x}_t} \Big[u\left(C_t(\tilde{x}_t)\right) \; + \beta E_{\tilde{Y}_{t+1}|Y_t} V_{t+1}(\tilde{Y}_{t+1}, B_{t+1}, \tilde{m}_{t+1}) \Big] \tag{1}$$

$$\text{subject to} \;\; C_t(\tilde{x}_t) = Y_t - [\lambda + z(1-\lambda)]B_t + q_t(B_{t+1}|Y_t, \tilde{x}_t)[B_{t+1} - (1-\lambda)B_t]$$

The long-term bond matures with probability λ, following Chatterjee and Eyigungor (2012). For the bond that does not mature, each unit pays out a coupon amount of z. The determination of the issuance price schedule, $q_t(B_{t+1}|Y_t, x_t)$, will be discussed in the market clearing section below.[9]

When a default happens in period t, the sovereign ceases servicing the debt, is excluded from credit markets, and faces weakly positive, additive output losses $\psi(Y_t) = \max\{\theta_0 Y_t + \theta_1 Y_t^2, 0.0\}$. These costs could be interpreted as the usual consequences of tightening credit conditions, a disruption of trade credit, or a banking slump (Mendoza and Yue [2012] or Sosa-Padilla [2012]).

Our model also features recovery on defaulted bonds, as in Hatchondo et al. (2016). When a default happens, a debt swap is proposed with Poisson probability ϕ and the sovereign can choose to accept it to return to credit markets, or reject it to stay in financial autarky and wait for another proposal. The size of the recovery rate in these proposals is the source of

sovereign decides its bond supply before m_{t+1} realizes and cannot change its issuance decision in either case.

[9] If the budget set is ever empty, either before or after the realization of x_t, we follow the literature standard and assume that the sovereign defaults. In the calibrated model, we impose Inada conditions that ensure that the sovereign would never repay if there was a strictly positive probability that the budget set would be empty at the auction.

unobserved uncertainty in the period before default, i.e., m_t. After default, m_t is assumed to be constant until the sovereign accepts a proposal and returns to the credit markets. Under these assumptions, the value of default can be expressed recursively as follows:

$$V_{D,t}(Y_t, B_t, m_t) = u\big(C_t\big) \ + \beta E_{\tilde{Y}_{t+1}|Y_t}\Big[\phi V_{t+1}(\tilde{Y}_{t+1}, m_t B_t, m_t) + (1-\phi)V_{D,t+1}(\tilde{Y}_{t+1}, B_t, m_t)\Big]$$
$$\text{subject to} \ \ C_t = Y_t - \psi(Y_t) \tag{2}$$

Notice that the sovereign's choosing to accept or reject a debt swap proposal is embedded in the term $V_{t+1}(\tilde{Y}_{t+1}, m_t B_t, m_t)$. In particular, it may reject a proposal if its current endowment state makes it more valuable to remain in the default regime. Moreover, every time the sovereign rejects a proposal, the next proposal will further reduce the debt to be repaid by the same haircut rate. Hence, both the accumulated haircut by the time the sovereign accepts a proposal and the number of periods of continuous financial exclusion are endogenous and depend on how many times the sovereign rejects a proposal.

It is worth noting that the repayment value does not take m_t as an argument since the recovery rate m_t is irrelevant should the sovereign choose to repay because it is i.i.d. On the other hand, $V_{D,t}$ does take m_t as an argument because it determines the size of the haircut when it is eventually proposed. V_t is a function of both $V_{R,t}$ and $V_{D,t}$ and hence it takes m_t as an argument.

We define the sovereign's default decision with a binary operator:

$$d_t(Y_t, B_t, m_t) = \mathbf{1}\{V_{R,t}(Y_t, B_t) < V_{D,t}(Y_t, B_t, m_t)\}$$

2.4. Forecaster

There is an inherent difficulty associated with market-based information acquisition problems: The price tends to convey too much information. Perfect Bayesian investors can infer all relevant information from market prices, which gives them no incentive to acquire information in the first place (Grossman [1976], Kyle [1985], or Dow and Gorton [2006]). We circumvent this problem by designing the market to operate with complete rationality and transparency on a signal of the true hidden information.[10] No additional information regard-

[10]The other popular option is to add unseen noise to the aggregate supply, e.g., Grossman and Stiglitz (1980). In our model, however, supply is determined by the optimal behavior of the sovereign. Thus, there is little plausible room for such noise.

ing payoff-relevant states can be gleaned from the price besides what participants already know.

In particular, we separate the information acquisition decision from bond investment, much like Angeletos and Werning (2006). We assume that all lenders are fully rational, but that each acquires information by employing the same contractor, whom we call the forecaster. In a sense, the forecaster serves as a market maker by coordinating market information. Veldkamp (2011) argues that information is a commodity that is difficult to obtain but essentially costless to disseminate. Our market set-up reflects this feature of information production by allocating the task to a single specialist. Once the information is produced, it is transmitted to all market participants. Real-world analogues of the forecaster would be financial media and analytics firms such as Bloomberg or Reuters. Alternatively, they could be interpreted as financial news or media outlets such as the Wall Street Journal of the Financial Times, credit rating agencies such as Moody's or S&P, or supernational public institutions such as the International Monetary Fund.

The forecaster has a technology capable of gathering information regarding the unobserved shock, m_{t+1}, at a per-unit cost κ. He sells his services to lenders as an independent contractor for a small fee, $l_t \geq 0$. A subscription fee for such a service is not unrealistic. For instance, one pays a subscription fee to access IMF's International Financial Statistics data or a Bloomberg terminal. We also assume for simplicity that lenders must pay this subscription fee to have access to the sovereign bond market, such that all lenders in this market have access to a forecaster-provided signal of the unobserved shock.

Because the forecaster is interested in the integrity of its forecasts, much like the rating agencies in Holden et al. (2018) or Manso (2013), it actively weighs this objective against information acquisition costs. In each period it produces a signal, x_t, of the next period's unobserved shock, m_{t+1}. The signal and the logit of the unobserved true state are jointly normal, and the information contained in this signal is reflected in $\rho_{mx,t} = corr(x_t, m_{t+1}) \in [0, 1]$, which is the forecaster's choice.[11] That is, the forecaster can modify the signal to make

[11] Our restriction to signals with positive correlation is without loss of generality, since negatively correlated signals have the

it more or less informative about m_{t+1}: More informative signals will feature a larger $\rho_{mx,t}$. For this reason we will call $\rho_{mx,t}$ the *accuracy* or *precision* of the signal.[12] The forecaster uses this information to publish a forecast (distribution) over all future states, observed and unobserved: $\hat{f}_t(Y_{t+1}, m_{t+1}|Y_t, x_t) = f_t(Y_{t+1}|Y_t)g_{\rho_{mx,t}}(m_{t+1}|x_t)$. The goal is to minimize the mean-square-error of the default-risk forecast under this conditional distribution, that is, to minimize the conditional volatility of default risk, subject to information costs.

In our benchmark, we assume m_{t+1} and x_t to be orthogonal to observed states Y_t, but our framework is flexible enough to allow for some correlation with no change in the mechanism. The forecaster would simply acquire the residual information that is not conveyed through observed states.

The information required to obtain a signal is given by a time-invariant function, $I(\rho_{mx,t})$, which is increasing in signal accuracy. The per-unit cost of information is a constant κ. In the benchmark, we assume that $I(\cdot)$ is the reduction in entropy in m_{t+1} that comes from knowledge of x_t, but our results do not hinge on this functional form.[13] Any increasing function would work.

We formulate the forecaster's information acquisition problem as below, given Y_t and B_{t+1}:

$$\min_{\rho_{mx,t}\in[0,1]} \quad E_{\tilde{x}_t} E_{\tilde{m}_{t+1},\tilde{Y}_{t+1}|\tilde{x}_t,Y_t}\left[d_{t+1}(\tilde{Y}_{t+1}, B_{t+1}, \tilde{m}_{t+1}) - \bar{d}_{t+1}\right]^2 + \kappa I(\rho_{mx,t}) \tag{3}$$

$$\text{subject to} \quad \bar{d}_{t+1} = E_{\tilde{m}_{t+1},\tilde{Y}_{t+1}|Y_t}\left[d_{t+1}(\tilde{Y}_{t+1}, B_{t+1}, \tilde{m}_{t+1})\right]$$

where $d_{t+1}(\cdot)$ is the binary default identifier.[14] To see the benefit of information acquisition, notice that the variance of the interior expectation is decreasing in $\rho_{mx,t}$. Consequently, the variance in the forecaster's default forecast can be reduced if he is willing to undergo costly information acquisition. When Y_t and B_{t+1} indicate a greater risk of default, acquisition of more accurate information will be optimal, since the unobserved shock matters for the

same information content.

[12] This information structure is isomorphic to a Gaussian signal-noise model in which the forecaster chooses the variance in the noise. We prefer our specification to this, though, since it features a closed domain, i.e., an infinitely noisy signal would arise with some regularity along the equilibrium path.

[13] This notion of information was developed primarily by Shannon (1958) and applied to economics by Sims (2003, 2006).

[14] We prefer to think of the objective as a utility rather than a resource cost, so we do not incorporate their fees into the objective. Since the fee is cointegrated and we stationarize the model to solve it, such a model would be isomorphic to ours.

variance of the forecast. When these publicly known states indicate little to no default risk instead, the forecaster can save on information costs and provide imprecise or even orthogonal signals because m_{t+1} is adding little to no variance to the forecast.

In Appendix B, we consider a Rational Inattention (Sims [2003]) variation of the model in which the lenders themselves collect the information. In this specification, the lenders work not only to minimize uncertainty regarding default risk, but also of all other types of risk, i.e., dilution and recovery risk as well. The results are broadly similar to the benchmark model both qualitatively and quantitatively.

2.5. Foreign Lenders

There is a unit mass of risk-averse foreign lenders who invest in risky sovereign debt. These lenders act competitively, similarly to Lizarazo (2013) or Aguiar et al. (2016). Lenders arrive in overlapping generations and each lives for two periods. Each lender is endowed with wealth, w_t, and pays the contractor fee, l_t, to gain access to the forecaster's signal, x_t. Then, they solve a portfolio allocation problem to decide how much to invest in risky sovereign debt and how much to invest in a risk-free asset yielding a return, r.

Overlapping lenders can arrive in one of two states of the world: Sovereign inclusion in capital markets or sovereign default/exclusion. In states of inclusion, new lenders purchase debt both from the sovereign and the legacy lenders who held bonds that did not mature, i.e., both primary and secondary markets are active. In states of default/exclusion, legacy lenders alone supply debt to be purchased and thus only secondary markets are active.

2.5.1. Demand During Inclusion

Given the above setup, investor i takes the bond issuance price, q_t, as given and solves the following problem, knowing public states Y_t and B_{t+1} as well as the signal x_t and its

accuracy $\rho_{mx,t} = \rho_t(Y_t, B_{t+1})$:

$$\max_{b_{i,t+1}} \; E_{\tilde{Y}_{t+1},\tilde{m}_{t+1}|Y_t,x_t}\left[\frac{c_{i,t+1}^{1-\gamma_L}}{1-\gamma_L}\right] \tag{4}$$

$$\text{subject to} \;\; c_{i,t+1} = [w_t - l_t - b_{i,t+1}q_t](1+r)$$

$$+ [1 - d_{t+1}(\tilde{Y}_{t+1}, B_{t+1}, \tilde{m}_{t+1})]\{\lambda + (1-\lambda)[z + q_{t+1}(A_{t+1}(\tilde{Y}_{t+1}, B_{t+1})|\tilde{Y}_{t+1}, \tilde{x}_{t+1})]\}b_{i,t+1}$$

$$+ d_{t+1}(\tilde{Y}_{t+1}, B_{t+1}, \tilde{m}_{t+1})q_{D,t+1}(\tilde{Y}_{t+1}, B_{t+1}, \tilde{m}_{t+1})b_{i,t+1}$$

where A_{t+1} is the issuance policy in period $t+1$ and $q_{D,t+1}(\tilde{Y}_{t+1}, B_{t+1}, \tilde{m}_{t+1})$ is the price in the secondary market of a bond that has been defaulted on. Notice there is no Bayesian extraction problem to be undertaken as in Lucas (1972), Grossman (1976), or Bassetto and Galli (2019), since the forecaster has coordinated market expectations on its noisy signal.

We denote aggregate bond demand in non-defaulting periods by:

$$B_{D,t+1}(Y_t, x_t, B_{t+1}, q_t) \equiv \int_0^1 b^\star_{i,t+1}(Y_t, x_t, B_{t+1}, q_t)di.$$

Though we use the i-index to denote individual investors, there is no heterogeneity among them, so the indexing will be irrelevant to equilibrium dynamics.

2.5.2. Demand During Exclusion

A lender who purchases a bond during a period of sovereign exclusion also faces uncertainty. One of three things can happen: (1) Next period there is no proposed swap and the creditor sells all her holdings on the secondary market, i.e., $b_{i,t}q_{D,t+1}$; (2) Next period there is a proposed debt swap and the sovereign accepts, in which case the creditor receives payment on a fraction of the bond, i.e., $m_t b_{i,t}$; and (3) next period there is a proposed debt swap and the sovereign rejects, in which case the secondary bond supply shrinks by m_t and the creditor sells her reduced holdings at the secondary market price, i.e., $m_t b_{i,t} q_{D,t+1}$.

Given our recovery set-up spelled out above, secondary bond demand at a given price,

$q_{D,t}$, is determined as follows:

$$\max_{b_{i,t}} \; E_{\tilde{Y}_{t+1}|Y_t,m_t}\left[\frac{c_{i,t+1}^{1-\gamma_L}}{1-\gamma_L}\right] \tag{5}$$

$$\text{subject to} \;\; c_{i,t+1} = (w_t - l_t - b_{i,t}q_{D,t})(1+r) + (1-\phi)q_{D,t+1}(\tilde{Y}_{t+1}, B_t, m_t)b_{i,t}$$
$$+ \phi\left[1 - d_{t+1}(\tilde{Y}_{t+1}, m_tB_t, m_t)\right]\{\lambda + (1-\lambda)[z + q_{t+1}(A_{t+1}(\tilde{Y}_{t+1}, B_{t+1})|\tilde{Y}_{t+1}, \tilde{x}_{t+1})]\}m_tb_{i,t}$$
$$+ \phi d_{t+1}(\tilde{Y}_{t+1}, m_tB_t, m_t)q_{D,t+1}(\tilde{Y}_{t+1}, m_tB_t, m_t)m_tb_{i,t}$$

Notice that in this market m_t is constant and already known, with public states Y_t and B_{t+1}. Despite this, we assume that investors' subscription and access to the forecaster stand and thus still pay the contractor fee l_t.

2.6. Market Clearing

Once the information structure is chosen by the forecaster, the signal is realized and distributed to all lenders, who then enter a competitive market with a common information set. The sovereign issues its predetermined debt stock, B_{t+1} at the highest possible price. Market clearing requires that $B_{D,t+1} = B_{t+1}$ in states of inclusion. This yields a pricing schedule identical in structure to that in Aguiar et al. (2016) but with the inclusion of the signal realization as an additional state.

The forecaster plays no role in the exclusion states since they can mitigate no risk for buyers or sellers there. Again, lenders will be representative and the market price, $q_{D,t}$ will adjust until $B_{D,t} = B_t$, where $B_{D,t}$ is given by the solution to Equation 5.

2.7. Equilibrium Definition

Definition 1. *A* **Markov Perfect Equilibrium** *is a set of functions,* $\{V_t(Y_t, B_t, m_t), V_{R,t}(Y_t, B_t), A_t(Y_t, B_t), V_{D,t}(Y_t, B_t, m_t), \rho_t(Y_t, B_{t+1}), q_t(B_{t+1}|Y_t, x_t), q_{D,t}(Y_t, B_t, m_t)\}_{t=0}^{\infty}$ *such that* ***(1)*** $V_{R,t}(Y_t, B_t)$ *and* $V_{D,t}(Y_t, B_t, m_t)$ *solve Recursions 1 and 2 and imply the policy* $B_{t+1} = A_t(Y_t, B_t)$ *and* $V_t(Y_t, B_t, m_t) = \max\{V_{R,t}(Y_t, B_t), V_{D,t}(Y_t, B_t, m_t)\}$*;* ***(2)*** $\rho_t(Y_t, B_{t+1})$ *solves Problem 3;* ***(3)*** $q_t(B_{t+1}|Y_t, x_t)$ *ensures that the market clears in states of non-exclusion, where bond demand is derived from Problem 4; and* ***(4)*** $q_{D,t}(m_t, Y_t, B_t)$ *ensures that the secondary market clears in states of exclusion, where demand is derived from Problem 5.*

3. Calibration

To determine the impact that costly information acquisition has on the pricing of sovereign risk, we calibrate the model to match a set of empirical moments from Russian quarterly data between 2008 and 2016, which is the window in which the Bloomberg data is available. During this time, Russia both had access to credit markets and underwent a large, salient crisis in late 2014 and 2015. The intuition, however, is general enough that we could perform the exercise for other countries.

While the state space is not overly large, being at most 3-dimensional, the model requires the accurate computation of many multidimensional and highly computationally expensive expectations, especially in the forecaster's problem. Traditional methods prove insufficient for this task even with extensive parallelization and high-caliber computing resources. Thus, we solve the model using a specially tailored machine learning algorithm derived from the procedure described by Scheidigger and Bilionis (2019). Details regarding the solution of the model can be found in Appendix A, and the algorithm there contained is suitable not only for our particular model but for a wide range of sovereign debt models.

We draw data from three primary sources: First is the JP Morgan Emerging Market Bond Index (EMBI) database taken from Datastream; second is the World Bank International Debt Statistics database; and third is Bloomberg's news-heat database.

3.1. Information Cost Identification

First and foremost, to find a proper cost value per unit information, κ, we match the variability of information acquisition in the model and in the data. We find a value for κ that matches, in the model and the data, the fraction of periods in which intense attention is paid to the borrower country, where intense attention is defined as information flows greater than the midpoint (i.e., half of the maximum attention level that is ever paid to a country during the sample period). As information becomes infinitely costly, this fraction approaches zero; as it becomes free, this fraction approaches one. Further, this fraction increases as information costs decrease and thus identifies it.

The model has a direct measure of information acquisition, $I(\rho_{mx})$. The literature has proposed many different empirical measures of 'attention' or information acquisition ranging from news to trading volume to extreme returns to Google search trends.[15] Given our model set-up, we instead select an empirical entity, Bloomberg, to serve as the forecaster, and invoke their news-heat data as an empirical proxy for information acquisition. Since Bloomberg terminal users are mainly financial professionals who are likely to have both the incentives and financial resources to react to important news, this news-heat dataset can be used as a proxy for investor attention on a specific equity and its underlying company or sovereign.[16]

Bloomberg news-heat data come in two types: One measures the daily number of its news stories on companies and sovereigns; the other measures its users' news reading and news searching activities on Bloomberg terminals.[17] For instance, Ben-Rephael et al. (2017) use Bloomberg news-heat daily max readership data and find significant impact of institutional investors' attention on their trading behavior and stock price dynamics. To best mimic the structure of the model, we use Bloomberg daily total number of news story publications as our main indicator for investor attention, since it proxies for Bloomberg's information acquisition efforts.[18] It is similar to Bi and Traum (2019), who use the volume of newspaper articles related to fiscal policy as a measure of investor attention. This publication measure also has better data quality than the news-heat readership data in our case.

The sample period ranges from 3/3/2008 to 3/14/2016 for 61 countries.[19] We sum the daily publication data on a sovereign over a quarterly frequency and normalize each observation by the total number of publications on all available countries in the entire dataset for that quarter. Consequently, our measure of attention to a sovereign is a quarterly time-series

[15] Barber and Odean [2008], Gervais et al. [2001], Seasholes and Wu [2007], or Da et al. [2011].

[16] See Ben-Rephael et al. (2017) for a composition of Bloomberg clients' job titles and industries.

[17] Although Bloomberg does not track news-heat on bonds directly, according to Bloomberg one can still measure investors' attention on issuers by taking the news-heat data corresponding to the primary equity that is underlied by an issuer. We first use Bloomberg's World Countries Debt Monitor to identify equities that are associated with sovereigns and their debt, and then obtain news-heat data on these equities.

[18] In Section 5, we conduct formal econometric exercises on all emerging economies in a matched EMBI/news-heat dataset of 28 countries.

[19] Historical data are missing for the periods of 8/22/2008–10/19/2008 and from 3/15/2016 onward.

that can be described as a publication share. The fraction of periods with intense attention by this metric is 19% for Russia.

3.2. Other Parameters

To obtain the output process, we estimate via MLE an AR(1) process on HP-filtered log dollar-valued real Russian GDP from 2008-2016 at a quarterly frequency with the HP-smoothing parameter set to 1600. The MLE estimates that emerge from this procedure are $\rho_y = 0.7225, \sigma_y = 0.0343$. We endow the sovereign with power utility with a CRRA given by $\gamma = 2$, as is standard in the literature.

We capture empirically relevant recovery dynamics by ensuring that our logit-normal recovery process matches both the mean and volatility of bond recovery data taken from Cruces and Trebesch (2013). These authors find that the mean recovery rate is 0.63 and the volatility of recovery rates is 0.216. A logit-normal generates these two moments when the underlying Gaussian shock has a mean of 0.670 and a volatility of 1.136. The density of the estimated recovery shock distribution can be found in Figure 1.

We further assume that the risk-free rate is fixed at 1% quarterly; that the lenders exhibit constant relative risk-aversion preferences with CRRA γ_L; and that $\phi = 0.083$, which is the estimate used by Mendoza and Yue (2012) for an average duration of 6 years before returning to international bond markets.[20] For simplicity, we normalize lender wealth to be, $w_t = 1+l$, such that after paying fixed forecaster fees they have a unit mass of wealth to invest. The exact value of l is irrelevant given this normalization and all risk-aversion is driven by the choice of γ_L. The sovereign does not pay the forecaster since Bloomberg is our empirical analogue.

Following Arellano (2008), we take our debt moments from the World Bank International Debt Statistics database. The empirical debt analogue for our setup is Russian public and publicly guaranteed debt. The average debt/GDP in this data series is 43.0%. We calibrate the coupon on the bond to be $(1-\lambda)z = r$ following Aguiar et al. (2016), since this spec-

[20] We find, as do Hatchondo et al. (2016), that the sovereign takes the first deal proposed to exit autarky with probability near unity.

ification ensures that a risk-free bond has a price of one regardless of maturity. Given this coupon specification, we determine the maturity of the debt by matching debt-service/GDP, which delivers a maturity of $\lambda = 3.2\%$.

We calibrate the remaining five parameters, $\{\beta, \theta_0, \theta_1, \gamma_L, \kappa\}$, using simulated method of moments (SMM) to jointly match five moments: Average annual spread, annual spread volatility, average debt-to-GDP, the average risk premium, and the fraction of quarters with intense attention.[21] To compute the average risk premium, we follow Longstaff et al. (2011), who decompose the spread into default risk and a risk premium as follows: After estimating their model with a risk-averse pricing kernel, they price the exact same default risk using a risk-neutral pricing kernel and back out the implied spread, and find that about 1/3 of the spread is a risk premium. We repeat this exercise in our model, computing the average spread demanded by a hypothetical risk-neutral lender, and target the ratio of this default risk against our benchmark average spread, which is 2/3.

These parameters are given in Table 1. Each parameter is primarily identified by its corresponding target moment, though there are significant and non-negligible cross-partial effects. We present in Table 1 both a calibration for the benchmark model and a "no-information" calibration in which we set information costs to infinite and match all the targeted moments except the attention fraction. This alternate calibration will be useful in our discussion of the results.

4. Results

4.1. Typical Crises/Defaults

The acquisition of more information during crises has a number of results and implications. Before we exposit the properties that arise from this, it is worth exploring how typical crises and defaults occur in our model to provide context. From this point on, when we refer to the "standard" model, we are referring both to Arellano (2008) and its extension to long-

[21] The model is at a quarterly frequency; during calibration, model moments are annualized to match certain data moments at an annual frequency.

maturity debt, Chatterjee and Eyigungor (2012). In this section, we will discuss results from the no-information calibration of our model, since it is comparable with the standard model and still shares most of the crisis/default dynamics with the benchmark model without the added complication of information acquisition.

In the standard model, the sovereign borrows more during booms than busts. This is because bond prices are higher during booms, which tempts the sovereign to lever up (price effect) despite the natural consumption-smoothing desire to save when times are good.

Figure 2a reveals that in our model the sovereign instead borrows more during busts than booms. This is not driven by the new information or recovery structure, but rather by the calibrated maturity of the debt, which is nearly twice as long as Chatterjee and Eyigungor (2012). Because so little of the debt needs to be rolled over (only 3.2%), the consumption-smoothing motive to borrow more during recessions dominates the price effect. In keeping with this intuition, the net effect is attenuated at higher debt levels where prices become more responsive to debt issuance. Indeed, if we shorten the maturity, the sovereign begins to borrow more during booms as in the standard model.

In terms of pricing, bond demand as a function of issuance, which can be found in Figure 2b, is downward sloping and still resembles the standard model insofar as lower endowment realizations imply worse price schedules. This is because of debt dilution. As explained earlier, the sovereign borrows more during busts due to the long maturity. Since busts are persistent, low output today implies a high likelihood of high borrowing tomorrow, which dilutes prices for the vast majority of possible issuance levels. This puts pressure to raise bond yields, with low output and high debt levels where default risk is more prominent.[22]

In addition, calibrated default costs in our model exhibit reverse curvature relative to the standard model, i.e., default costs rise as output falls. This is entirely due to the fact that we calibrate default cost curvature to match spread volatility, which is a common practice in the literature. The standard model required lower default costs during downturns to

[22]However, this pressure can be partially offset by the reverse curvature in the default costs that we discuss in the next paragraph, as default costs rise with lower output.

match Argentina's high spread volatility of 4.4% during the 1990's. Flat default costs could not generate enough volatility, but curvature induces the sovereign to delever more slowly during periods of low output, tolerating higher bond yields in the process and so raising spread volatility.

But bond yield volatility has fallen substantially across most emerging markets since the 1990s (Born et al. [2020]). For our more recent Russian data the spread volatility is only 1.0%. Flat default costs for us generate too much volatility not because the model is fundamentally different than the standard model along this dimension, but because the target is much lower. Thus, we must tame the model's intrinsic spread volatility by reversing the curvature in the default costs rather than amplifying it. The reverse curvature induces the sovereign to delever more quickly during periods of low output, which works to eliminate prolonged periods of excessive spreads and thus decreases spread volatility.

Because of the long maturity and the reverse curvature, two forces (debt dilution and default costs) are at play as output and debt level change, and defaults happen differently in this model relative to the standard model. In the standard model, default typically happens following a surprise negative shock in the midst of a boom during which the sovereign levered up. In our model, default typically happens following a prolonged recession in which output steadily falls and debt levels steadily rise. Initially, the falling output lowers spreads as default costs go up, but as the recession goes on and debt levels rise, the latter eventually make borrowing more expensive. It is then at the beginning of a recovery when the default typically happens, since debt levels are still high due to the prior recession, but default costs are also getting lower. In fact, the default typically happens shortly after the sovereign is beginning the process of delevering but is met with an unexpectedly better output shock (near trend) that tempts them to default before they finish the delevering. These dynamics can be seen in Figure 3, which depicts the average default in a long simulation (half a million quarters).

It is worth noting that these deviations from the standard model are *not required* for endogenous information acquisition to generate the features we outline in this paper. If we

layered costly information acquisition into the standard model then we would see all of the same novel implications while maintaining the dynamics of the standard model. Rather, it is our choice of country and, more importantly, the time-period restriction set by our use of the News-Heat data that requires this particular calibration with these particular consequences. In Appendix C, we show that the benchmark model with default costs taken straight from Chatterjee and Eyigungor (2012) operates in much the same way, albeit generating spreads that are too volatile to match the data.

Our results from the benchmark calibration with endogenous information acquisition shares most of the above features and crisis dynamics (with a few exceptions of output and spread dynamics right before default events, which we highlight in the next section). This can be seen in Figures 4a and 4b, which give the policy and pricing functions, respectively, are very similar to the no-information case, featuring both more borrowing in busts than booms (with some attenuation at high debt levels) as well as monotonicity in the policy function with respect to output.

4.2. Information Acquisition Behavior

What is novel in this model is the state-contingent acquisition of information: More information is endogenously acquired and priced during crises. By 'crisis' here, we mean states of heightened spreads and default risk. It is in these states that the forecaster begins to pay more attention.

We can see this in the information acquisition policy function in Figure 4c, which provides the solution to the forecaster's problem in equilibrium. We can learn a lot from this one figure about how information acquisition works in this model.

First, the forecaster does not expend any information resources for the vast majority of debt levels. They only start paying attention for debt levels above 0.42 or so depending on the level of income. This is because there is no default risk at these lower debt levels and thus there is no uncertainty in the default risk forecast for the forecaster to reduce.

Second, the information acquisition policy is non-monotone in B_{t+1}. It increases from zero as debt levels rise into the upper reaches of the ergodic distribution, but it eventually turns

around and begins to fall in most cases. This is because the forecaster is only interested in the uncertainty associated with default risk, not the default itself. For very high debt levels, default is so certain that it is not worth learning more about. So the forecaster does not.

Third, the relation between information acquisition and output changes changes with debt level. When debt level is about 0.43 to 0.45, the upsloping portion of the information acquisition policy function is increasing in y_t, for each level of debt. This is a consequence of default being cheaper in booms than in busts, and thus more tempting for higher output levels; the forecaster needs to acquire more information about the heightened default risk when output is higher. When the debt level is above 0.45, the downsloping portion of the information acquisition policy function is decreasing in y_t, for each level of debt. This is because of the long maturity and debt dilution, lower output is associated with higher debt, higher spread, and higher default risk. Thus, the forecaster acquires more information about the heightened default risk when output is lower.

This state-contingent information acquisition has real consequences for prices. This can be seen in the price schedule viewed as a function of the signal, x_t, which is given in Figure 4d. For low debt levels, the forecaster does not acquire useful information about unobserved shocks; this implies that lenders do not pay attention to these uninformative signals and thus prices do not react. Hence, there is no difference across the three price schedules. As debt levels rise, though, the forecaster begins to acquire useful information, the lenders start to make use of the valuable signals, and the price schedules begin to diverge according to the different signal values. It's worth noting that because of bond's fixed income nature and because that default is a rare left-tail risk, there is an asymmetric response: A negative signal causes a substantial drop in the price when the information is useful, while a symmetric positive signal increases the price only slightly.

Figure 5 reveals how these new dynamics affect a typical default. The similar broad pattern in output, indebtedness, and spreads emerge that we saw in the no-information case Figure 3. However, it is worth noting that with information acquisition, average output tends to be slightly below trend at a default and lower than that in the no-information case.

This is more in accord with the findings of Tomz and Wright (2007) and Benjamin and Wright (2013) that about 60% of defaults take place when output is below trend.

Further, we see that information acquisition increases leading up to a default, climbing from a signal correlation of about 0.62 ten periods prior to the default to about 0.82 on the verge of the event. This is because while income is fluctuating prior to a typical default, indebtedness is increasing substantially. Thus, it is the increased borrowing leading up to a default that causes the forecaster to pay attention, not the recession.

It is also worth noting that spreads increase right before default in the benchmark, in contrast to the no-information case wherein they fell. This is because defaults are often associated with a bad debt-recovery shock (m) at high debt levels. The no-information case cannot price this and instead prices only the increase in output following a recession, which reduces yields as can be seen in Figures 4b and 5d. The benchmark model not only prices the output recovery, which reduces yields, but also the bad debt-recovery shock (m), which is partially observed through an accurate signal. This latter effect dominates the former effect and overall spreads tend to rise prior to the default.

4.3. Heteroskedasticity (Time-Varying Spread Volatility)

A key result following the information acquisition is that our model endogenously generates state-contingent variation in sovereign spread volatility (or spread heteroskedasticity). Spreads exhibit increased volatility during crises when lenders price the unobserved shocks more accurately rather than considering them to be at their mean, which they do during non-crisis times.

To assess our model's ability to generate state-contingency in the spread volatility, we propose a model-free metric of heteroskedasticity that we call the **Crisis Volatility Ratio** or CVR.[23] It is defined as follows: In a series of data, either simulated or empirical, let $\hat{T}_x$ denote the set of all periods in which the change in the spread from the prior period is above

[23]Typically, when measuring time-varying volatility, the literature imposes quite a bit of structure. For instance, Melino and Turnbull (1990) and Fernández-Villaverde et al. (2011) measure the impact of time-varying volatility in the context of a stochastic volatility model. Imposing this or another similar structure to measure the quantitative efficacy of the model would be inappropriate in our case, since we know that the data-generating process for the simulated data is not a stochastic volatility model.

the $(1-x)$ percentile of its distribution for some small, positive number x. We follow Aguiar et al. (2016) in calling such events "crises." By construction they are much more likely to happen than default events and are always observed in the data even if a default is not. With this notation, we define the CVR for a given crisis threshold x as

$$CVR_{x,w} = \frac{1}{|\hat{T}_x|} \sum_{t \in \hat{T}_x} \frac{\hat{\sigma}_{t:t+(w-1)}}{\hat{\sigma}_{t-w:t-1}}$$

where $\hat{\sigma}_{t_1:t_2}$ is the sample standard deviation calculated using the periods from t_1 to t_2. CVR compares the volatility in a window of w periods immediately prior to a crisis to the volatility in a window of w periods after. Neither window includes the crisis itself, i.e., from $t-1$ to t where crisis events occur at $t \in \hat{T}_x$. If the CVR is larger than one, then post-crisis periods tend to be more volatile than pre-crisis periods. Table 2 shows empirical CVRs for a panel of 23 emerging markets with a variety of frequencies; additional details can be found in Appendix D. The average across them at a quarterly frequency is 2.36 when $x = 2.5\%$ and $w = 5$. It suggests that a typical emerging market features crises that more than double the volatility on impact.

To give a sense of the metric, any AR(1) process would feature an average CVR equal to one regardless of x or w, since the volatility of the innovation is independent of any prior observation. However, a model designed to generate heteroskedasticity, such as an AR-ARCH or an AR-ARCH-M model, will feature CVRs greater than unity. To see this, we simulate both AR-ARCH and AR-ARCH-M models defined as follows:[24]

$$s_t = \kappa_0 + \rho_s s_{t-1} + \gamma\sigma_t + \sigma_t\epsilon_t, \quad and \quad \sigma_t = \omega + \alpha_1\epsilon_{t-1}^2$$

where s_t is the spread, ϵ_t is a sequence of iid errors following $N(0,1)$. Notice that α_1 governs the persistence of volatility in the above model and γ determines the impact of volatility on spreads. An AR-ARCH simply sets $\gamma = 0$ where an AR-ARCH-M allows it to be non-zero. We then calculate CVRs from simulating these two models for 1.5 million periods under different parameterizations of (α_1, γ).[25] Across different parameter pairs of (α_1, γ), we set

[24] An AR-GARCH-M or a similar variation could also be used to generate heteroskedasticity with similar results. We favor an ARCH to a GARCH specification for two reasons: First, the CVR behaves quantitatively very similarly across both; second, when we do estimate this model, the EMBI data appear to favor an ARCH to a GARCH, at least for the case of Russia. When we allow for the additional GARCH parameter, the ML procedure routinely sets it to zero, i.e., its lower bound.

[25] Note in order for the model to be stationary, α_1 is bounded above by unity. The same sequence of innovations is used

the AR parameter at $\rho_s = 0.6$ and adjust κ_0 and ω until the simulated series has both a mean of 2.5% and a volatility of 1.0%, which is consistent with the spread data for Russia. In Table 3, we report the CVR comparative statics and demonstrate that these models indeed can generate CVRs greater than unity, but in this case a large γ is required to bring the CVR into the empirically relevant range.

We first compute the CVR for the Russian data (the first column in Table 4) in the same time window. We can see that heteroskedasticity is a strong feature of the data by this metric: When $x = 2.5\%$ and $w = 5$, the CVR in the data is 2.46. Russia in this period is typical of emerging markets in general, since Table 2 tells us that this figure is within 10% of a standard deviation from the mean. A CVR being larger than unity comes as no surprise, as strong time-variation in the volatility of the country-risk spread is documented by Fernández-Villaverde et al. (2011).

Our benchmark model (the second column in Table 4) comes fairly close to matching this figure at 2.13. To get a sense of our model's relative performance, we consider three alternative common existing models. First, we re-calibrate our benchmark model assuming infinite information costs and not matching the attention fraction; it is effectively a variation of Aguiar et al. (2016) with recovery and default cost curvature. This alternate calibration can be found in Table 1 as well. It is known that calibrated default models can replicate some of the state-contingent second-order dynamics of bond spreads (de Ferra and Mallucci [2020]) as the bond price elasticity is higher in high debt regions. Our model corroborates this finding insofar as the no-information model offers a CVR at about 1.7, higher than an AR(1), but it also falls quite short of both the data and the benchmark model.

Next, we estimate via maximum likelihood an AR-ARCH and an AR-ARCH-M model on the Russian spread data. We report the CVRs in Table 4 as well.[26] For these latter models, the parameters are estimated to capture not only the first two spread moments, mean and variance, as our benchmark model does, but rather the entire spread distribution.

across all parameterizations.

[26]Parameter estimates and associated confidence bands can be found in Appendix E.

Nevertheless, both the AR-ARCH and the AR-ARCH-M yield CVRs less than 1.4, which is also short of both the data and the benchmark model. While these models, designed to generate conditional heteroskedasticity, *can* generate CVRs in the empirically-relevant range, they do not at the maximum likelihood parameters. In comparison, our benchmark still generates the most heteroskedasticity that is closest to the data as measured by CVR.

4.4. Comparative Statics

To understand further how information acquisition affects the model dynamics, we consider a comparative static across information costs. These comparative statics can be found in Figure 6. We see first that the spread volatility is monotonic in information costs, which is to be expected as more information is priced more often.

This monotonicity in the spread volatility induces a monotonicity in the default frequency, and thus average spread level, as well. The sovereign defaults more often as more price volatility reduces the value of repaying debt relative to defaulting, because sovereign does not have to endure this additional price volatility when defaulting. In contrast, the average debt level is relatively constant.

But there are many interesting non-monotonic effects. First, the CVR is hump-shaped in information costs. This is intuitive. Initially as information costs fall, attention increases first in periods of crises, which raises volatility in crises relative to non-crisis times. However, as information becomes very cheap, attention is allocated in non-crisis states as well, which works to shrink the CVR.

Second, the fraction of time periods with intense attention (i.e., attention fraction) is non-monotonic. This is because, in equilibrium, attention is affected by information costs directly as well as general equilibrium effects indirectly induced by information costs through default frequency. More specifically, default frequency changes the unconditional volatility of default risk, which alters the difficulty level of the forecaster's job, i.e., the minimization of conditional volatility of default risk.

For high levels of information costs, default events are rare and only prone to a small set of states (as seen in Figure 6). This causes the unconditional volatility of default risk itself

to be high, which makes forecasting default risk more difficult. Thus, the forecaster has to pay intense attention to get any signal to narrow down the default risk variance and obtain a decent conditional volatility of default risk, despite information costs. Of course when information costs are extremely high, the above effect is overpowered by the cost effect and thus no attention is ever paid.

As information costs go down, the heightened price volatility implies that default occurs more often in a broader set of states, hence the unconditional volatility of default risk shrinks, which makes forecasting default risk easier. The forecaster does not have to pay intense attention as often to minimize the conditional volatility of default risk. This latter effect dominates in the upper-middle-range of information costs and the time fraction with intense attention falls. As information costs continue to fall, though, the direct of effect of cheaper information kicks in inducing more attention and attention fraction begins to rise back up.

In the next two sections, we discuss the rest of Figure 6: the share of risk in spread (i.e., risk-neutral share) and sovereign welfare.

4.5. State-Contingent Risk Premia

The second result we highlight for this paper is that the composition of spreads during crisis times changes when we consider costly information acquisition. To understand this, first note that we can intuitively break the spread on sovereign debt into two categories: Risk (default, dilution, and recovery risk) and a risk premium for that risk.

When lenders pay the cost to observe normally unobserved states, they learn more about the realization of those shocks. In particular, the conditional volatility $\sigma_{m|x}$ shrinks. This pushes down the competitively priced risk premium. This occurs even in the case when signal $x_t = 0$ and merely supports the prior of the lenders, which is that $m_{t+1} = 0$.

What does this imply for the pricing behavior? While default risk is high, so too is the lenders' ability to learn and contain that risk since they are acquiring more precise signals about unobserved shocks. This implies that around a crisis their effective risk-aversion is lower and thus that the risk premium comprises a smaller share of the spread. Were an

econometrician not to take this effect into account, instead employing a more standard sovereign default model in which the investors' attention set is non-state-contingent, she would tend to *underestimate* underlying risk around crises by assuming the risk premium to be higher than it actually is.

Our model allows us to quantify this compositional shift in the risk spread. To do so, we construct an artificial, non-equilibrium price schedule in which $\kappa = \infty$, which implies that no information regarding m_{t+1} is ever transmitted to risk-averse lenders. To calculate risk premia for the benchmark case (BM) and the no-information case (NI), we compare them to the risk-neutral pricing of equilibrium risk, which is the same in both cases. The risk premium will be the difference between the spreads implied by these series. We compute the average share of the default, dilution, and recovery risk by dividing the risk-neutral spread by the total spread with risk-averse lenders for different levels of information costs. We find the share of risk (i.e., risk-neutral share) increases as the costs decrease (Figure 6), since the share of risk premium falls as obtaining information becomes cheaper.

We then consider the following event study: We define 'crisis' periods as in the CVR, continuing to use the 97.5%ile of the spread change distribution as the cutoff, and isolate crises given by the no-information (NI) counterfactual pricing schedule, which are not necessarily driven by learned information.[27] The only difference between the benchmark (BM) and no-information (NI) pricing schedules during such events is the risk premium: In the BM model it will be typically be lower than in the NI counterfactual, since more information is acquired and thus more risk premium is reduced in the benchmark model.

We isolate all such events in a simulation with length of half a million periods, and compute the median share of the default, dilution, and recovery risk by dividing the risk-neutral spread by the total spread with risk-averse lenders in the NI and BM case, respectively.[28] The stochastic impulse-response function can be found in Figure 7. We can see that the

[27] We could further condition on crises in which x_t is roughly zero, and thus the benchmark spread is not unusually high or low as a result of learning, since in these crises informed investors had their prior beliefs confirmed. However, we find that the implied price volatility in the benchmark from x_t averages out quickly and the results do not change.

[28] We use the median share of risk here as it is smoother than the average and their magnitudes are about the same.

share of risk in spread is always larger than the share of risk premium in spread, both for NI and BM cases.

In the normal course of events, the share of risk in spread is persistently smaller in the NI counterfactual than in the BM. That is, the share of risk premium is larger in the NI counterfactual, by about one half a percentage point. Because NI lenders never acquire information, always bear more uncertainty, and thus ask for more risk premium than BM lenders who acquire some moderate amount of information during normal times.

Something more interesting happens as the time approaches a crisis: The share of risk declines, i.e., the share of risk premium increases, in the NI model relative to the BM model. This is because a crisis is usually associated with a bad debt-recovery shock and the benchmark lenders learn about this shock and how dangerous it will be for the next period, but not the NI lenders. Then, one period after a crisis, the danger is made known, default risk becomes even more dominant than risk premium in the spread, the NI risk premium share falls relative to the BM model. After that, given the persistence of debt level and output process, the NI lenders realize they are still at the edge of another possible crisis, without knowing the debt-recovery shock that can tip the sovereign over to default at any time. NI lenders' risk premium share soon rebounds substantially larger than the benchmark for a considerable time.

This effect is fairly small, operating at an order of magnitude of about 2% of the spread, but it is particularly interesting for two reasons. One, it is fairly persistent. The relative jump in the NI risk premium typically takes at least 12 quarters to normalize to pre-crisis levels. Second, it follows from the fact that risk premium depends on *country-specific* states. Thus, risk premium cannot be controlled for only using global metrics, such as the CBOE VIX or the P/E ratio, as is often done (Aguiar et al. [2016] or Bocola and Dovis [2019]). Rather, our theory suggests that in order to accurately assess risk premium and the underlying default risk or dilution risk, some metric of forecaster/investor information acquisition must be controlled for.

It is worth noting that this effect is not so strong as to reverse the cyclicality of risk

premia level. The level of risk premia remain countercyclical in our model. After all, default risk is high during crises, and thus the expected repayment is low, which drives up the risk premia. This can also be seen in Figure 8, which gives the average of risk premium level, i.e., difference between the benchmark spread and the counterfactual risk-neutral spread, during a typical no-information crisis. We can see that it is lower before the crisis, i.e., when times are good and output is high, and jumps substantially during the crisis.

4.6. Transparency Policy

The last result we highlight regards the benefits and costs of transparency in light of our model mechanism. Costly information acquisition can be interpreted in many ways in the context of our model. One way the unit information cost κ can be understood as is the level of transparency that a sovereign has about its domestic affairs and finances. Interpreted this way, our model can also offer some intriguing insights for sovereign transparency.

Our model suggests that transparency is a double-edged sword. When there is zero transparency, i.e., information is infinitely costly, lenders always demand a risk premium for the unobserved shocks, especially during crisis times. This makes it more expensive for the sovereign to borrow and service debt. Having more transparency will benefit the sovereign by lowering the risk premium. On the other hand, when the sovereign is fully transparent, although risk premium for the unobserved shocks disappears, it is replaced by substantial spread volatility, since now what used to be unobserved shocks are always reflected in bond prices. This not only hurts the sovereign because the country is risk-averse, but it also works to increase the default frequency, as we saw in the previous section. This implies that there are generally more deadweight loss costs incurred in equilibrium

Therefore, transparency brings about two forces: There is the benefit of lower risk premia, but also the cost of higher price volatility and default. To illustrate this insight, we show the sovereign's welfare levels along different information costs in Figure 6d, which is evaluated at the steady state output and at zero debt.[29] We can see that as the information cost

[29]This is the welfare metric used by Chatterjee and Eyigungor (2012). The valley-shape pattern remains the same if we evaluate welfare at the ergodic mean of the debt-to-GDP ratio instead of zero debt.

decreases from the highest end, the sovereign's welfare first increases sharply in response to the lower risk premium. However, after about $\kappa = 0.028$ the welfare falls, as the increased volatility and frequency of default make the country worse off overall.

Because of these countervailing forces, the overall welfare effects are quite small, on the order of only 0.1% certainty-equivalent consumption. They are not, however, state-contingent, as the same pattern emerges for different levels of debt and output. Thus, the model suggests that countries may have a reason for increasing opacity or otherwise limiting the flow of relevant information to foreign investors.

5. Empirical Analysis

The calibrated model yields interesting implications and may leave one wondering whether these results are specific to our benchmark calibration or whether they hold more generally. Figure 9 shows a positive co-movement between investor attention and sovereign bond spreads for a broad range of countries.

To more formally test how sovereign bond spread moments relate to investor attention, we conduct a reduced-form empirical analysis for a abroad set of countries. We estimate the following equation using the daily EMBI spread and Bloomberg publication share data over the period of 2008/3/3–2016/3/14 for 28 emerging economies:

$$B_{it} = \beta_0 + \beta_1 IA_{it} + \beta_2 IA_{it} Crisis_{it} + \beta_3 z_i + \epsilon_{it} \tag{6}$$

The dependent variable B_{it} is the daily measure of sovereign bond spread moments for country i. Here we use two different measures to get a robust picture: logged daily EMBI and its (rolling window) standard deviation. We include the standard deviation as a dependent variable because our theoretical model also predicts a rise in spread volatility when investor attention increases. The standard deviation is calculated over the current date and the previous four days. For the standard deviation regressions, we also include logged daily EMBI level as a control variable to remove any alternate impact from level effects.

IA_{it} is the measure of investor attention: the share of Bloomberg news stories of country i

among those of all 61 countries on which Bloomberg reports news-heat data.[30] In Appendix G, we also use Bloomberg readership data as the investor attention measure in robustness checks.

$Crisis_{it}$ is a dummy variable that is equal to zero during normal times and one during crises. A crisis is defined the same way as before when CVR is calculated, that is, a period in which change in the EMBI from the prior period is above the 97.5 percentile of its distribution, follow Aguiar et al. (2016). We allow the publication variable to interact with the crisis dummy, which measures how the impact of the investor attention on bond spreads shifts during crises. z_i is a country fixed effect.

The first two columns in Table 5 show the daily-data regression results. We can see that the publication's relations with the EMBI and its standard deviation are statistically significant and positive during the sovereigns' crisis periods but not otherwise. This is precisely what the structural model predicts: During crises the correlation between investor attention and the bond spread becomes stronger; further, spread volatility increases with investor attention because more information is priced and the model suggests that this effect is stronger in crises when information flows contain more useful information.

These empirical results are consistent with the findings of Bi and Traum (2019), who use newspaper data from 1840s and find that, during crisis times, state fiscal information helped investors differentiate states with sound fiscal policy from insolvent ones and thus affect state government bond prices for states that were ill-prepared for downturns, whereas it was not the case during non-crisis times.

As robustness checks, we then extend the baseline regression with more country-specific control variables. The estimated equation becomes:

$$B_{it} = \beta_0 + \beta_1 IA_{it} + \beta_2 IA_{it} Crisis_{it} + \beta_3 z_i + \beta_4 X_{it} + \epsilon_{it} \tag{7}$$

where the new addition in the equation, X_{it}, denotes a vector of control variables and includes government balance as a share of GDP, debt-to-GDP ratio, RGDP growth, inflation, and

[30] There is no particular trend shown in the news heat data for bloomberg's publication behavior or for any country during our sample period. Nevertheless, we have run regressions with detrended data, the results are consistent with those below when we use non-detrended data.

unemployment rate. For the standard-deviation regression in the 4th column we also control for spread level. In order to run the extended regression, quarterly data are used instead. More specifically, we use the quarterly J.P. Morgan EMBI data and calculate its standard deviation by averaging the daily EMBI's standard deviation for each quarter. For publication shares, we first sum up the daily number of Bloomberg publications to the total publications each quarter and then use the quarterly publications to calculate the shares. The $Crisis_{it}$ dummy is calculated as before but using quarterly EMBI data.

The last two columns in Table 5 show the quarterly-data regression results. The publication's relations with the EMBI and its standard deviation remain statistically significant and positive during the sovereigns' crisis periods. More specifically, during crises, a one percentage point increase in a country's publication share is associated with a 38 percent increase in its EMBI and a 26 percent increase in its EMBI's standard deviation. The significant effect of additional information for investors during crisis is consistent with our earlier results.

6. Conclusion

Costly information acquisition plays an important role in the pricing of sovereign risk and we are able to study the quantitative impact of such an information friction. We constructed and calibrated a structural model of endogenous default and information acquisition using Bloomberg news-heat data as an attention metric to identify information costs. We demonstrated that the relationship between information flows and sovereign spread volatility is state-contingent and verified this prediction in the data. Moreover, all else being equal, without state-contingent information a standard sovereign default model may *understate* default or recovery risk around crises by assuming a risk premium that is too large. The inclusion of costly information acquisition in the model also allows us to study the welfare implication of transparency. Transparency is a double-edged sword, it not only can affect spread volatility but also has first-order effects on spread level and default risk. Countries may increase opacity or otherwise limit the flow of relevant information to foreign investors to maximize their welfare.

Tables and Figures

Table 1: Calibrated Parameters by Simulated Method of Moments

Benchmark	Value	Target Description	Data	Model
Discount factor	β=0.9693	Average Spread	2.5%	2.5%
Output cost (level)	θ_0=.0976	Average Debt/GDP	43%	43%
Output cost (curvature)	θ_1=-0.0634	Spread Volatility	1.0%	0.9%
Lender CRRA	$\gamma_L = 0.97$	(Avg. RN Spread)/(Avg. Spread)	67%	67%
Unit info cost	$\kappa = .0207$	Fraction of Quarters with $IA >$ Midpoint	19%	23%
$\kappa = \infty$				
Discount factor	β=0.97	Average Spread	2.5%	2.5%
Output cost (level)	θ_0=0.1578	Average Debt/GDP	43%	43%
Output cost (curvature)	θ_1=-0.1214	Spread Volatility	1.0%	1.0%
Lender CRRA	$\gamma_L = 0.83$	(Avg. RN Spread)/(Avg. Spread)	67%	67%

Notes: RN stands for risk neutral. The model is at a quarterly frequency; during calibration, model moments are annualized to match average annual spread and annual spread volatility.

Table 2: Empirical Crisis Volatility Ratios (1994-2018)

Country	**Q**	**M I**	**M II**	**W I**	**W II**	**D I**	**D II**
Argentina	4.23	3.87	0.98	1.88	0.75	1.96	0.92
Brazil	2.09	2.09	2.48	2.09	1.36	1.79	1.51
Bulgaria	1.91	1.70	2.97	1.96	1.20	1.62	1.41
Chile	2.10	2.28	2.39	2.48	3.11	1.76	2.04
China	1.82	3.44	1.12	2.38	1.31	1.85	1.75
Colombia	3.38	1.91	2.54	1.70	1.14	1.60	1.22
Hungary	3.42	2.29	2.37	1.70	2.30	1.99	2.29
India	2.56	2.50	2.81	1.48	4.10	1.27	2.30
Indonesia	1.59	1.99	2.86	3.26	1.91	2.32	2.84
Latvia	1.14	0.38	1.33	1.17	0.98	1.34	1.38
Lithuania	1.49	0.85	2.06	2.15	0.66	1.67	0.68
Malaysia	3.07	3.44	3.11	2.85	1.69	1.81	1.99
Mexico	1.96	4.89	2.23	2.01	0.63	1.84	2.19
Peru	1.49	1.62	2.21	1.65	2.03	1.83	1.90
Philippines	1.23	3.06	3.25	2.57	3.57	1.88	2.63
Poland	2.33	1.70	2.69	1.85	2.14	1.78	3.70
Romania	–	3.06	7.53	2.08	4.16	1.66	0.32
Russia	3.51	9.50	2.01	1.54	2.31	1.38	1.07
South Africa	1.54	1.82	1.63	1.90	1.48	2.07	1.11
Thailand	2.12	1.74	1.57	8.19	1.45	1.82	1.44
Turkey	1.10	1.75	1.32	1.34	2.30	1.68	1.25
Ukraine	2.51	2.00	4.29	1.90	1.65	2.38	2.45
Venezuela	1.38	1.19	1.64	1.56	1.64	1.90	1.24
Average	**2.36**	**2.57**	**2.50**	**2.25**	**1.91**	**1.79**	**1.72**
Stdev.	**1.04**	**1.82**	**1.35**	**1.38**	**1.01**	**0.27**	**0.77**

Notes: In the first column we use quarterly data and set the crisis threshold and window to $x = 2.5\%$ and $w = 5$, respectively, as in the benchmark calibration. Other columns use different frequencies of the same dataset. Q means quarterly, M monthly, W weekly, and D daily. A numeral I indicates the same CVR parameters ($x = 2.5\%$ and $w = 5$) at alternate frequencies, whereas a numeral II indicates those same frequencies with x and w adjusted to match that of the quarterly data, e.g., for the Monthly II, $x = 2.5 \times \frac{1}{3}\%$ and $w = 5 \times 3$.

In the event that a crisis event occurs close to the beginning or end of the sample, we compute the sample standard deviation using the minimum of the designated window size and the available data. Using this approach, we are only unable to compute the CVR for Romania in the first column because there is only one crisis in the sample and it occurs at the second observation. Consequently, we cannot compute a sample standard deviation prior to the crisis.

Table 3: Crisis Volatility Ratios: Comparative Statics with Simulated Data

Parameter	**AR-ARCH CVR**	**Parameter**	**AR-ARCH-M CVR**
$(\alpha_1, \gamma) = (0.05, 0.0)$	1.18	$(\alpha_1, \gamma) = (0.5, 1.0)$	1.28
$(\alpha_1, \gamma) = (0.25, 0.0)$	1.17	$(\alpha_1, \gamma) = (0.5, 5.0)$	2.18
$(\alpha_1, \gamma) = (0.5, 0.0)$	1.22	$(\alpha_1, \gamma) = (0.5, 10.0)$	2.61
$(\alpha_1, \gamma) = (0.75, 0.0)$	1.38	$(\alpha_1, \gamma) = (0.5, 20.0)$	2.76
$(\alpha_1, \gamma) = (0.95, 0.0)$	1.58	$(\alpha_1, \gamma) = (0.5, 50.0)$	2.80

Notes: For all cases, we set the AR parameter at $\rho_s = 0.6$ and adjust κ_0 and ω until the simulated series has both a mean of 2.9% and a volatility of 1.4% as in the Russian spread data.

Table 4: Crisis Volatility Ratios: Data, Benchmark, and Alternate Models with MLE

	Data	Benchmark	$\kappa = \infty$	AR-ARCH	AR-ARCH-M
$CVR_{x=2.5\%,w=5}$	2.46	2.13	1.73	1.37	1.38
$CVR_{x=5.0\%,w=5}$	2.67	1.98	1.72	1.38	1.39

Notes: AR-ARCH and AR-ARCH-M parameters are maximum likelihood estimates. Their details are in Appendix E. $\kappa = \infty$ is the case in which information costs are infinite but parameters are chosen via re-calibration to match the same moments rather than using the benchmark parameters.

Table 5: Bloomberg Publication Share and EMBI at Crisis Times

	Daily		**Quarterly**	
Expl. Variables	**log(EMBI)**	**log(Std. Dev.)**	**log(EMBI)**	**log(Std. Dev.)**
Pub. Share	0.00	0.19	−0.03	2.66
	(0.35)	(0.21)	(0.71)	(1.87)
Pub. Share× Crisis	2.23	5.83	37.74	26.38
	(1.11)	(0.67)	(13.63)	(7.10)
Other Controls	None	log(EMBI)	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	9682	9508	445	302
Adj. R-squared	0.008	0.165	0.192	0.292

Note: All regressions include fixed effects and robust standard errors. Other controls include government balance as a share of GDP, debt-to-GDP ratio, RGDP growth, inflation, and unemployment rate. For both spread standard deviation regressions, we also include $\log(EMBI)$ as a control variable. Data is from Bloomberg and CEIC. Standard errors in parentheses.

Figure 1:

Recovery Value Distribution

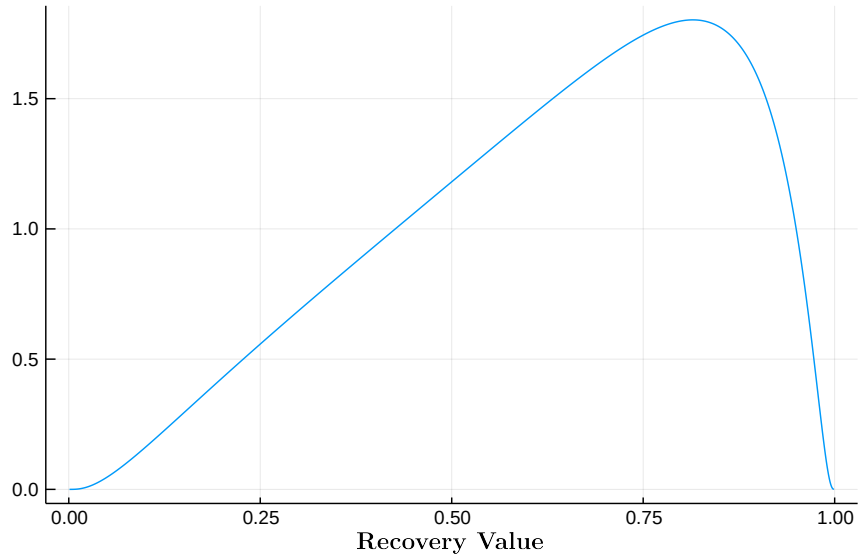
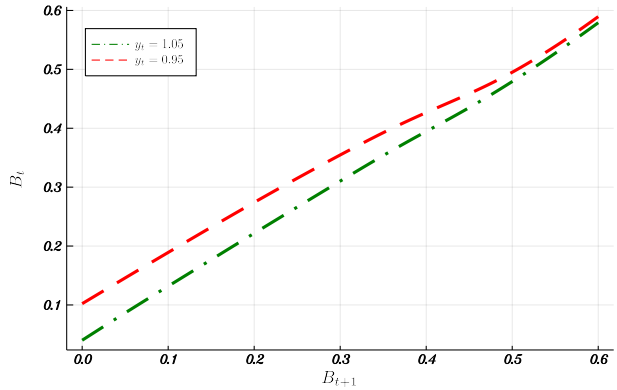


Figure 2: Model Objects: Typical Behavior for No-Info Calibration

(a)

Borrowing Policy Function:



(b)

Bond Pricing Function: $x_t = \bar{x}$

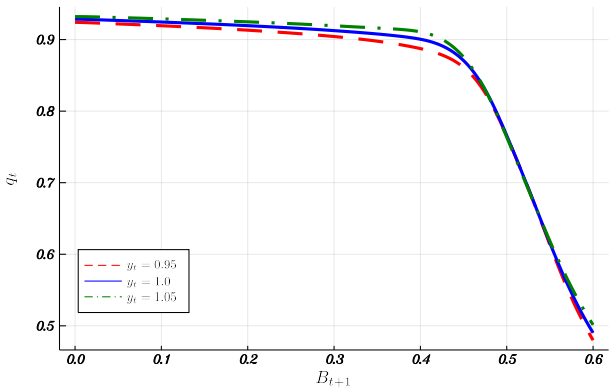


Figure 3: Typical Default for No-Info Calibration

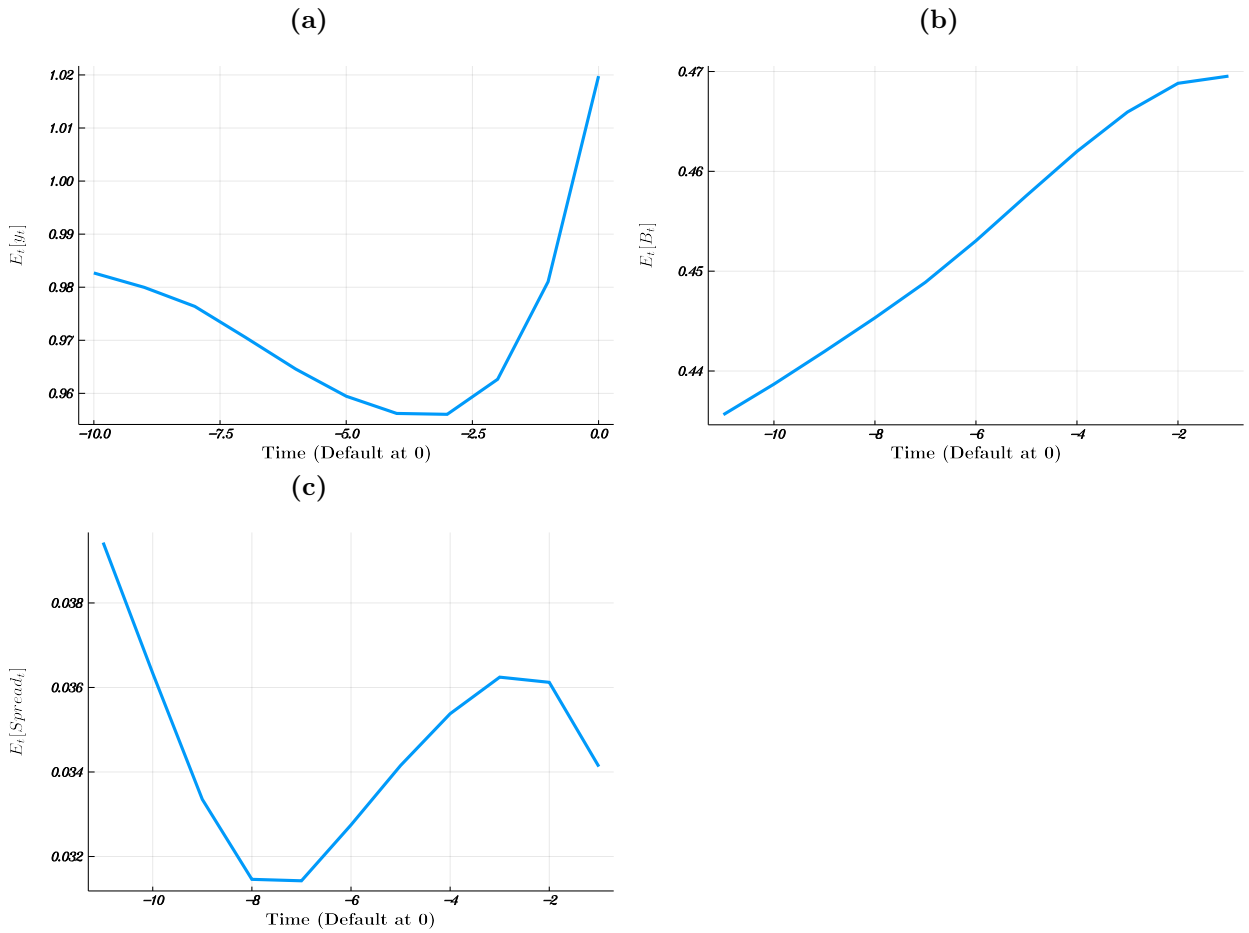


Figure 4: Information Acquisition: Model Objects

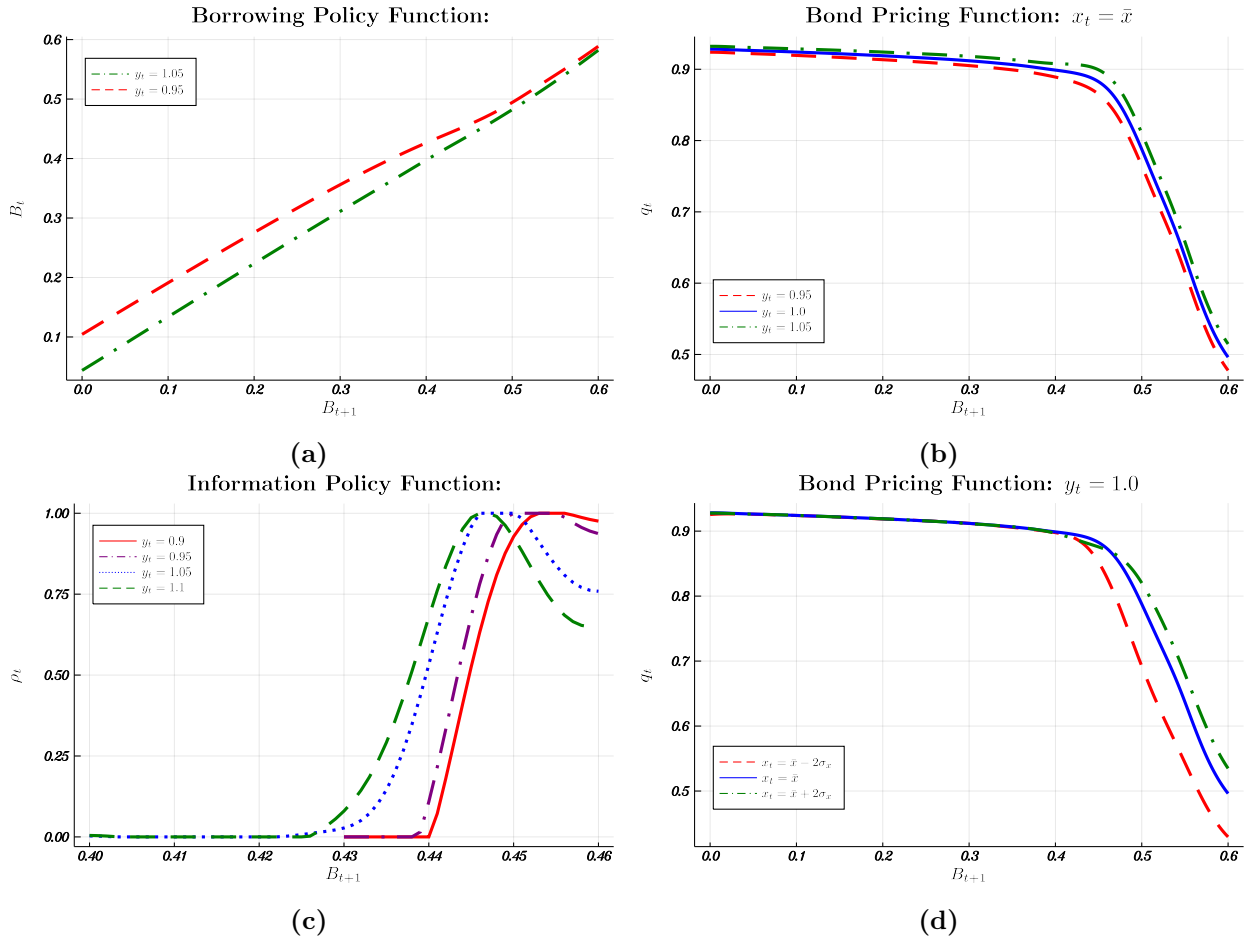


Figure 5: Typical Default for Benchmark Calibration

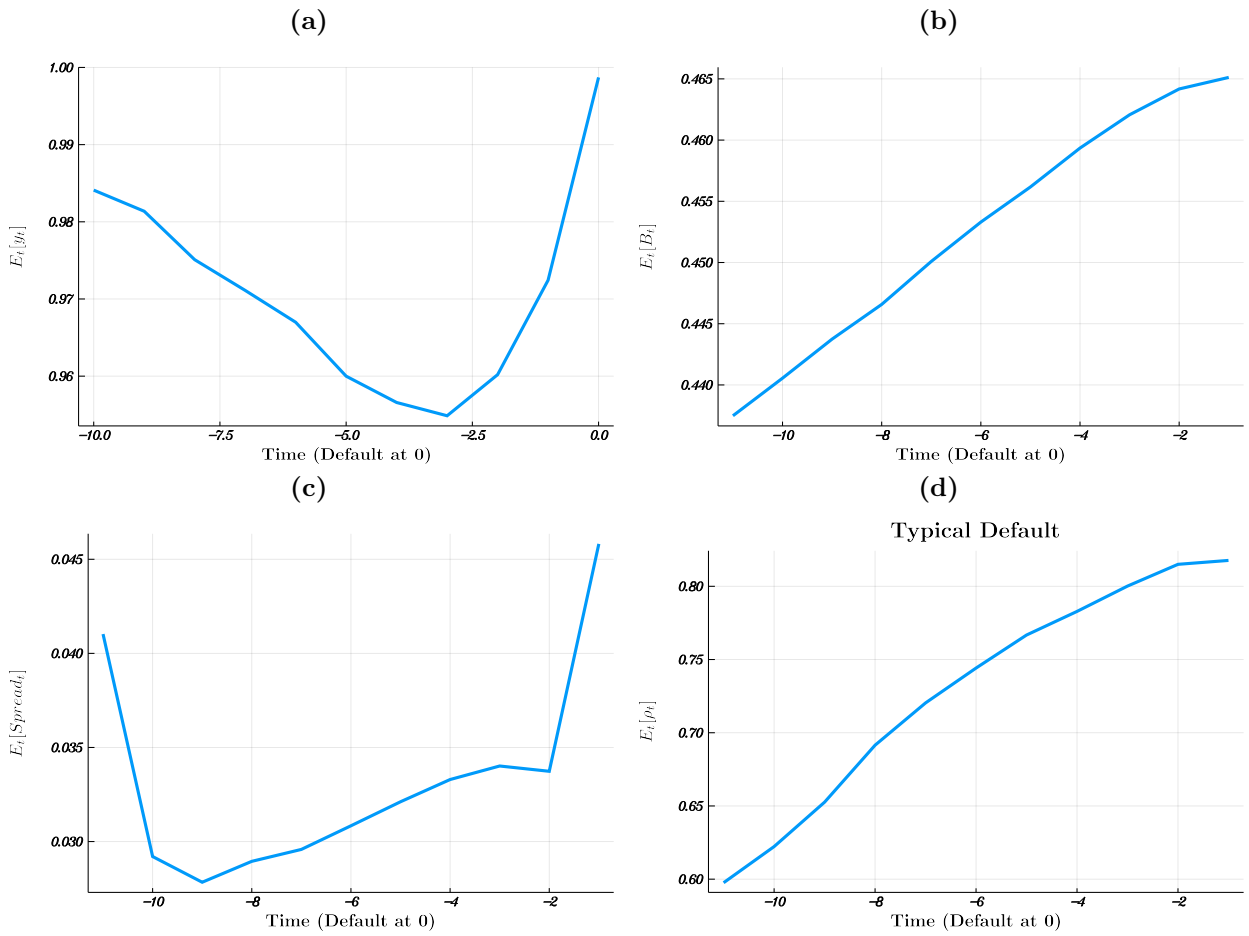
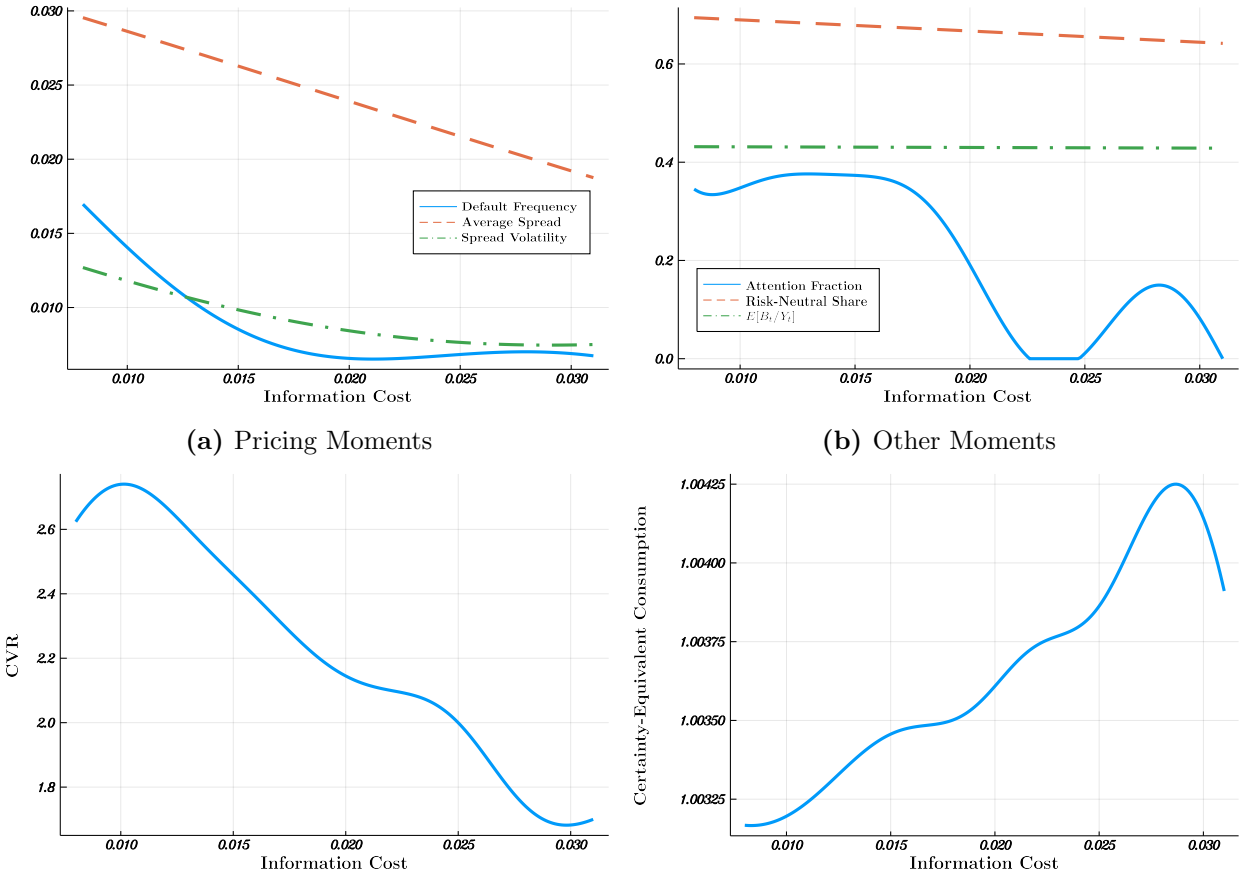


Figure 6: Comparative Static: Information Cost



(a) Pricing Moments

(b) Other Moments

(c) CVR

(d) Welfare

Figure 7: Spread Decomposition During Crises

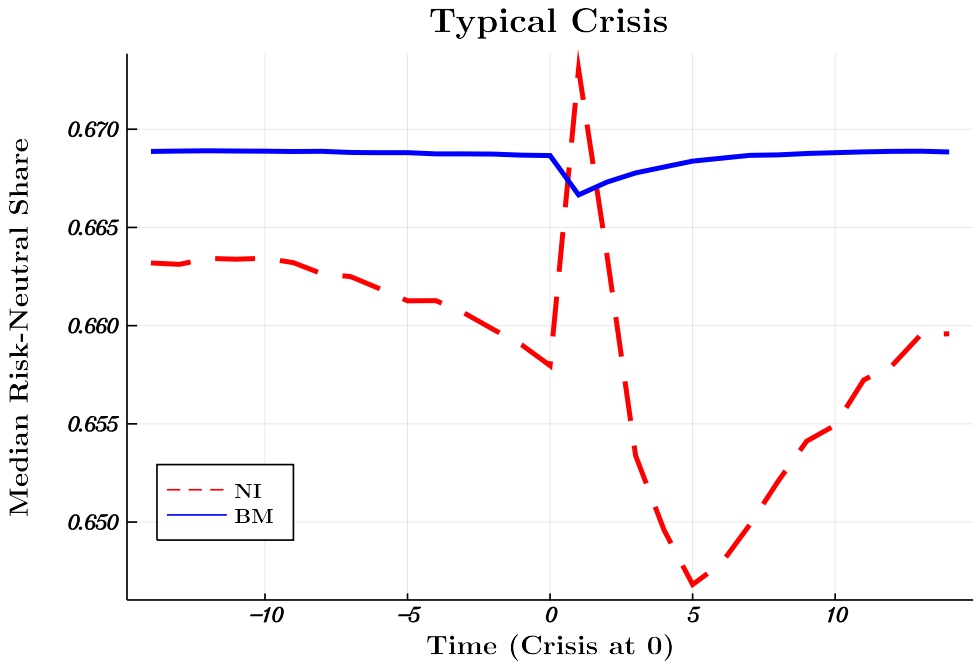


Figure 8: Risk Premium During NI Crisis: Benchmark Model

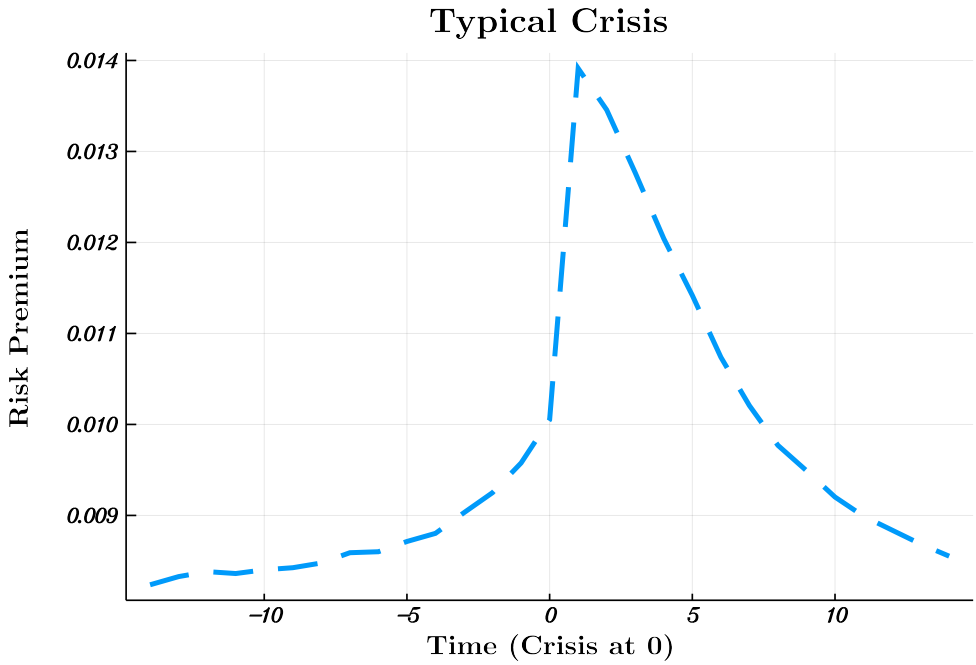
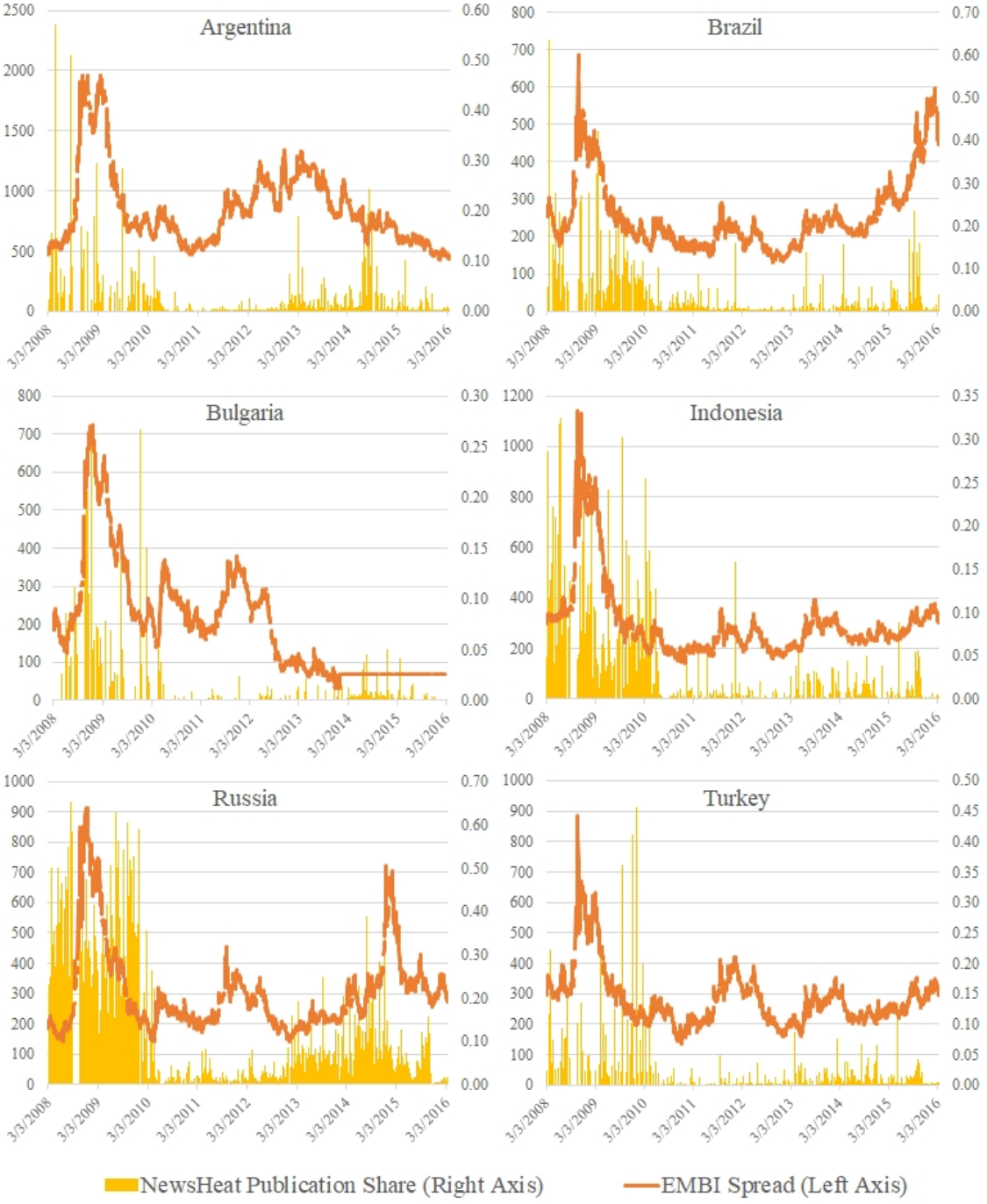


Figure 9: Bloomberg Daily Publication Share and EMBI, 2008/3/3-2016/3/14



Notes: News heat publication share is the number of Bloomberg news stories on a country divided by the total number of Bloomberg news stories on all the 61 countries that Bloomberg news heat data report, though in our empirical exercises we focus on only 28 emerging markets. There is a modest disconnect between 2010 and 2012 for many countries that seems largely to be driven by the Eurozone crisis. This crisis generated many news articles, which worked to push down the publication share of other emerging markets, but also generated contagion effects, which worked to push up their spreads.

Supplemental Appendices

Appendix A. Solution Algorithm

We solve the model using a variant of the Gaussian Process Dynamic Programming algorithm described by Scheidigger and Bilionis (2019), which is a machine learning algorithm. Every relevant equilibrium function is approximated by the mean of a Gaussian Process and we tighten their convergence criterion by a factor of one half to ensure accuracy. We modify the algorithm along only a handful of dimensions:

1. Rather than draw new training inputs (grid points) every iteration, we randomly select the training inputs from the .01% typical set (see Cover and Thomas [2006]) of the unconditional distribution over states, assuming all endogenous states to be distributed uniformly. These training inputs remain fixed across all iterations.
2. We use $40 \times D$ training inputs, where D is the dimensionality of the state space.[31]
3. We employ a Matern 5/2 Kernel rather than a Square Exponential Kernel.

All expectations are taken using a Gauss-Chebyshev quadrature with 20 nodes.

Approximating all value, price, and policy functions with the means of Gaussian processes, we solve for the stationary model as a limit of the finite-horizon game following Hatchondo and Martinez (2009). Since our solution method is continuous we find, as they do, no difficulty with convergence despite the inclusion of long-term debt.

Appendix B. Rationally Inattentive Lenders

It is worth considering how our benchmark model would behave if the investors themselves collected information as opposed to outsourcing this task to a third-party forecaster. One particular form of endogenous information acquisition that has gained considerable traction in recent years is Rational Inattention (RI), e.g., Sims (2003). In RI, decision-makers must pay information costs proportional to Shannon's entropy-based Mutual Information metric for correlating their actions with the underlying relevant states. In contrast to the benchmark model, which takes the structure of the signal as given, RI agents choose both the actions they take *and* the information structure that correlates their actions to the states. In other words, they design their own signal as well as choose how to act on it.

One big advantage of RI is that it is agnostic about whether these costs are a result of

[31] Scheidigger and Bilionis (2019) recommend $10 \times D$ for standard macro models, but we find that many more training inputs are required to capture the high degree of non-linearity in a model of endogenous default. Nevertheless, the critical feature that the number of grid points required expands only linearly in the dimensionality of the problem survives.

information acquisition or *information processing*. In the former, the agent can learn directly from public observables such as prices without paying the cost. In the latter the costs are always paid regardless of the source of the relevant information, that is, investors always need to pay costs to digest information, even for information coming from prices.

Interpreting it as information processing, which we will do here, offers the benefit of circumnavigating the Grossman puzzle laid out in the benchmark model. In particular, the interpretation is that the information is *always in the information set of the lenders*, but that they must pay a cost to process it in a way that can translate to their lending activity. Thus, it does not matter if information travels from equilibrium prices to the information set, since the bottleneck is not there but rather in the further translation of the information set to lending activity.

A general treatment of RI in the benchmark model will prove troublesome for reasons we'll describe momentarily. But we can easily explore an "RI stage game" to determine if it generates properties similar to the benchmark model. In particular, we assume that for a single period at time t, the model economy plays an RI stage game between the lenders and the sovereign in which lenders acquire all information themselves. From $t+1$ onward, then, the economy evolves according to the benchmark model. Thus, continuation values and future secondary market prices in the stage game are well-defined and plausible.

From a timing perspective, all that will change is that there will be no signal. The recovery shock in period $t+1$ will continue to realize in period t *after* the level of debt issuance is chosen but *before* the price is determined, just like the signal is in the benchmark model.[32] Since there is no signal, the market price will depend directly on m_{t+1} rather than a signal of it.

As a consequence, the only thing that will change on the sovereign's side is that the price schedule will now depend directly on the realization of recovery shock instead of the signal. Everything else remains the same: The sovereign continues to issue debt monopolistically to competitive foreigners at an equilibrium-determined pricing schedule.

Since there is no forecaster in the RI stage game, the lenders' problem is significantly different as they must process information regarding the recovery shock to translate it into actionable lending. We'll focus solely on the market for debt during inclusion in the stage game, since that is the market in which information acquisition takes place in the benchmark model as well.

In this market, an investor i takes aggregate states observed without friction (Y_t, B_{t+1}) as well as the bond issuance schedule across $q_t(m_{t+1})$. It then solves the following portfolio

[32] In the RI variant it actually does not matter whether the sovereign issues debt conditional on knowledge of m_{t+1} or not since borrowing cannot convey information anymore than the price can. But for the sake of similarity between this specification and the benchmark, we keep the timing the same.

choice problem with a rational inattention friction on the correlation of m_{t+1} and $b_{i,t+1}$. In particular, it designs its debt investment decision, $\tilde{b}_{i,t+1}$, to be a random variable from the space of all possible random variables $\mathcal{B}$, whose joint distribution with $\tilde{m}_{t+1}$ ensures that the prior marginal distribution over $\tilde{m}_{t+1}$ remains unchanged.

$$\max_{\tilde{b}_{i,t+1}\in\mathcal{B}} \; E_{\tilde{Y}_{t+1},\tilde{m}_{t+1},\tilde{b}_{i,t+1},\tilde{m}_{t+2}|Y_t}\left[\frac{\tilde{c}_i^{1-\gamma_L}}{1-\gamma_L}\right] - \kappa\mathcal{I}(\tilde{m}_{t+1};\tilde{b}_{i,t+1}) \tag{B.1}$$

$$\text{subject to} \;\; c_i = [e_t - b_{i,t+1}q_t(m_{t+1})](1+r) + b_{i,t+1}\times$$

$$\Big[[1-d_{t+1}(Y_{t+1},B_{t+1},m_{t+1})]\{\lambda+(1-\lambda)[z+q_{t+1}(A_{t+1}(Y_{t+1},B_{t+1})|Y_{t+1},m_{t+2})]\}$$

$$+\, d_{t+1}(Y_{t+1},B_{t+1},m_{t+1})q_{D,t+1}(Y_{t+1},B_{t+1},m_{t+1})\Big]$$

where e_t is lenders' endowment. $q_{D,t+1}(Y_{t+1},B_{t+1},m_{t+1})$ is the price in the secondary market of a bond that has been defaulted on. $\mathcal{I}$ is the mutual information between the recovery shock, $\tilde{m}_{t+1}$, and lender demand, $\tilde{b}_{i,t+1}$, i.e.,

$$\mathcal{I}(\tilde{m}_{t+1};\tilde{b}_{i,t+1}) = \sum_{m\in\mathbf{M}}\sum_{b\in\mathbf{B}} p(m,b)\log\left(\frac{p(m,b)}{p(b)\pi(m)}\right)$$

where $p(m,b)$ is the joint distribution of m and b, and $\pi(m)$ and $p(b)$ are the marginals of each. Notice that mutual information is always positive (see Cover and Thomas [2006]) and will only be zero if these two random variables are independent.

The solution to this problem will deliver the entire joint distribution of bond demand and the recovery shock. From this we can derive the conditional distribution of individual bond demand given some realized m_{t+1} and fixing aggregate, perfectly observed states: $p^*(b_{t+1}|m_{t+1};Y_t,B_{t+1})$. Given this, we denote aggregate bond demand from a unit mass of investors in non-defaulting periods by:

$$B_{D,t+1}(q_t|Y_t,B_{t+1},m_{t+1}) = \sum_{b\in\mathbf{B}} p^*(b|m_{t+1};Y_t,B_{t+1})b$$

Primary market clearing requires that $B_{D,t+1} = B_{t+1}$ in every possible state m_{t+1}, which implies that the market price in a state (Y_t,B_{t+1}) satisfies

$$B_{D,t+1}(q_t(\cdot)|Y_t,B_{t+1},m_{t+1}) = B_{t+1}, \quad \forall m_{t+1} \tag{B.2}$$

Note that even though lenders are representative ex-ante, they may demand different quantities ex-post. This is because RI typically implies stochastic solutions in choice variables. For instance, a typical lender may select $b_{i,t+1} = .75$ with marginal probability 50% and $b_{i,t+1} = .25$ with marginal probability 50%. Since this lender is representative, aggregate demand here would simply be $B_{D,t+1} = 0.5$.

Finally, for the sake of this exercise we will discretize both the recovery shock, m_{t+1}, into M grid points and the debt purchase decision, $b_{i,t+1}$, into B grid points. This allows us to solve the individual lenders' stochastic demand structure given a pricing vector using the

geometric method of Muller-Itten et al. (2020), which ensures global optimality. We consider a debt grid of 101 points and a recovery grid of 7 points.

We solve for the equilibrium price schedule as follows: Given a state (Y_t, B_{t+1}), we use a non-linear solver over the M-dimensional pricing vector to zero out the M-dimensional vector of excess demand (Equation B.2).

A valid concern is that there may be multiple such equilibrium pricing vectors given a state (Y_t, B_{t+1}). Indeed, we find this to be the case in many states. The endogeneity of prices in generating return risk that they themselves are intended to price in this model can lead to multiple valid equilibrium pricing vectors. We also have no theoretical reason to select one equilibrium over another, which is why we hesitate to construct the entire recursive Markov equilibrium from the RI stage game and instead invoke the forecaster in the benchmark model, which features determinate pricing conditional on a state (Y_t, B_{t+1}).

In the face of such multiplicity, we search for the equilibrium that most closely resembles our model by using as our initial guess in the non-linear solver the pricing vector derived from evaluating benchmark pricing schedule at the M discrete grid points. In the translation of prices from the benchmark model to guesses in the stage game, a given signal, x_t, implies a recovery rate, $m_t = \exp(x_t)/(1 + \exp(x_t))$ in our initial guess, i.e., $q_{init\ guess}(m_t) = q_t(Y_t, B_{t+1}, \log(x/(1-x)))$.

The equilibrium pricing function that zeros out excess demand starting from this guess and invoking a standard gradient-based non-linear solver is given in Figure Appendix B.1. In this figure, the value of the information cost, κ, is the same as the benchmark model. We can see that the same pattern as the benchmark model emerges. The prices are nearly identical to each other for debt levels less than about 0.45 and slope down gradually as in a standard default model. But as in the benchmark model, they diverge quickly once default risk becomes a real threat with information acquisition. High recovery rates imply less default risk and more recovery conditional on default, so prices rise modestly upon realization of a high recovery rate just as they did from a good signal in the benchmark model. Similarly, prices fall when the recovery rate is low, just like they did for a bad signal in the benchmark model.

The drop in prices upon news of low recovery is much more pronounced in the RI stage game than it is the benchmark model. Thus, it is possible that our benchmark model is understating the degree of state-contingency introduced into the pricing structure of debt by information acquisition. It suggests that, if anything, the results from the quantitative model would likely serve as a lower bound on what we could expect in a richer framework.

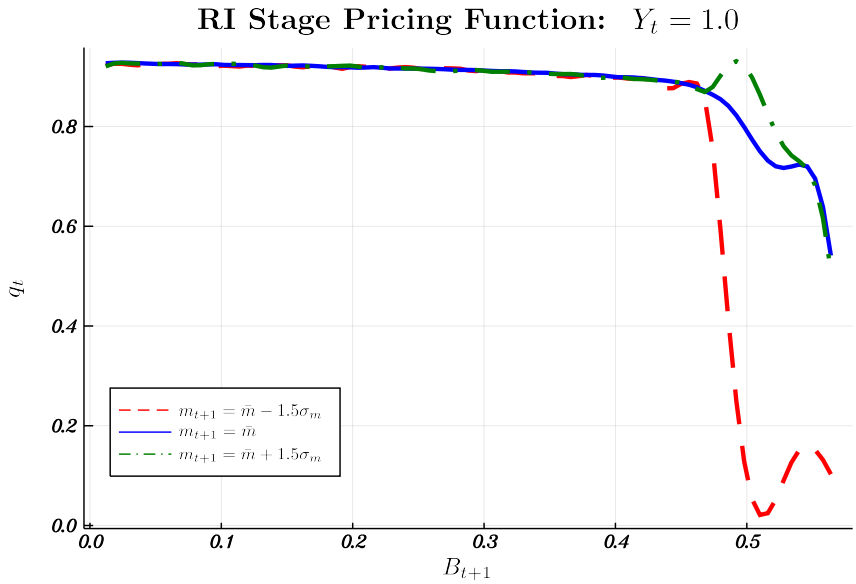


Figure Appendix B.1:

Appendix C. Typical Default Cost Curvature

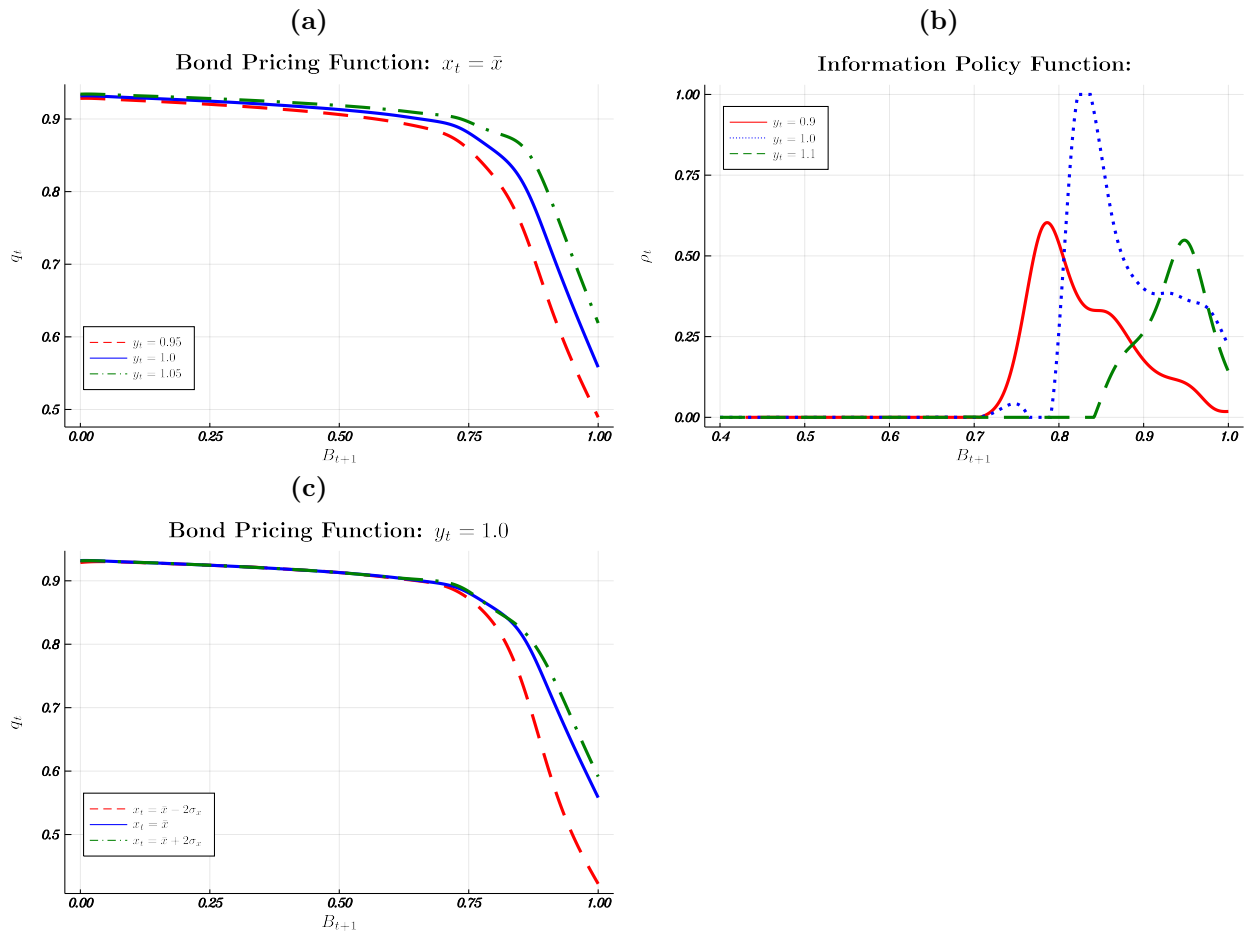
The benchmark model features reverse curvature in the default cost relative to more standard models, i.e., it is more expensive to default during busts than booms. This is entirely a consequence of the low spread volatility that we require the model to target rather than any feature of endogenous information acquisition.

Here, we explore how the benchmark model responds to a typical default cost curvature. In particular, we keep all benchmark parameters the same except the default cost parameters, which we supplant with those from Chatterjee and Eyigungor (2012), i.e., $\theta_0 = -.19$ and $\theta_1 = 0.25$. The pricing and information policy functions can be found in Figure Appendix C.1.

We observe that the dynamics of information acquisition are quite similar to the benchmark, but reversed as would be expected. First, the pricing schedules diverge more across income levels. This is because the country is tempted to default more at low income levels, not just borrow more, as was the case in the benchmark. Further, the information acquisition policy function reveals that attention is paid first in bad states of the world, as would be expected given that these are now more prone to default risk. Nevertheless, the implied pricing function across realizations of the signal looks very similar to the benchmark, with asymmetric responses across positive and negative realizations of the signal in high debt regions.

The only significant difference is that the information policy does not approach unity for some states of the world. In particular, low income and high incomes states max out their attention closer to 0.5, which stands in contrast to the benchmark wherein information

Figure Appendix C.1: Typical Default Cost Curvature Results



acquisition became something resembling binary. But this binary nature was a result, not an assumption, and it is not one that holds here. The steady state level features the most information acquisition at its peak debt level because it has the most uncertainty: It is equally likely to tip into recession as to a boom, which is not true in the more extreme cases. This difference from the benchmark can be dealt by choosing a proper information cost. A lower information cost than the benchmark would generate information policy functions resembling the benchmark along this dimension.

This all implies that the dynamics of interest that we consider here, e.g., heteroskedasticity, state-contingent risk premia, transparency welfare, etc., would go through in a model with more standard curvature.

Appendix D. CVR Cross-country Statistics

Here, we provide cross-country statistics for the CVR from the EMBI datset. Table 2 gives the CVR for the same panel of emerging economies considered in the empirical

work of Aguiar et al. (2016). Different columns indicate different parameters for the CVR parameters. In the event that a crisis is near the beginning or end of a sample and this window is insufficient, we use the largest window possible. The only cases in which this is impossible is if the crisis is the second or penultimate period in the sample. This happens only once in the sample, for Romania at a quarterly frequency.

In the first column we use quarterly data and set the crisis threshold and window to $x = 2.5\%$ and $w = 5$ respectively, as in the benchmark calibration. Across columns, Q stands for quarterly, M monthly, W weekly, and D daily. A numeral I indicates the same CVR parameters ($x = 2.5\%$ and $w = 5$) at alternate frequencies, whereas a numeral II indicates those same frequencies with x and w adjusted to match that of the quarterly data e.g. for the Monthly II, $x = 2.5 \times \frac{1}{3}\%$ and $w = 5 \times 3$.

Table 2 suggests that a typical emerging market features crises that are more than twice as volatile as non-crisis times, with the quarterly average for the dataset coming in at 2.36 and most other metrics coming in near or above 2 as well. Even the lowest metric, Daily II, yields an average CVR of 1.72, implying that a CVR greater than unity is a strong feature of the data. We also provide the correlations across these different metrics in Table Appendix D.1. There is broad positive correlation across these metrics, particularly with regard to the quarterly series, which is what we use in the benchmark model. The Standardized Generalized Standard Deviation (SenGupta [1987]) is 0.86, which is smaller than the volatilities of most of the individual series, which are given in the last row of Table 2.[33] This indicates broad comovement across the seven series. Consequently, the exact frequency or parameter choice is not pivotal in the computation of the CVR.

Table Appendix D.1: Empirical CVR (1994-2018): Correlation Table

	Q	**M I**	**M II**	**W I**	**W II**	**D I**	**D II**
Q	1.00						
M I	0.47	1.00					
M II	0.29	0.03	1.00				
W I	0.01	-0.08	-0.05	1.00			
W II	0.31	0.13	0.57	-0.07	1.00		
D I	0.03	-0.19	0.07	0.20	-0.21	1.00	
D II	0.16	-0.09	-0.01	0.04	0.18	0.33	1.00

Appendix E. Alternative Models: Parameter Estimates

We provide here in Table Appendix E.1 the maximum-likelihood parameter estimates for the Russian spread data used to compute the alternate-model-generated CVRs in Table 4.

[33] For reference, if there was zero correlation across a set of series, the Standardized Generalized Variance would simply be the geometric average of the variances of the series.

These results are robust to a wide range of initial parameters in the optimization, including all those from Table 3 designed to generate large CVRs. 95% confidence bands are estimated using the observed Fisher information matrix. Notice that in both cases, the conditional heteroskedasticity term, α_1, is fairly large and statistically significant. The mean term in the AR-ARCH-M model, while positive, is not statistically significant.

Table Appendix E.1: ML Estimates

Parameter	AR-ARCH	AR-ARCH-M
κ_0	.0239 [.0181,.0296]	.0209 [.0135,.0283]
ρ_s	.5529 [.3516,.7542]	.5100 [.3028,.7172]
$\sqrt{\omega}$	.0068 [.0046,.0089]	.0067 [.0046,.0088]
α_1	.7626 [.4422,1.0830]	.7534 [.4279,1.0789]
γ	N/A	.1835 [−.2151,.5821]

Appendix F. News-heat Data Summary Statistics

Table Appendix F.1 shows the summary statistics of the raw Bloomberg publication data at daily frequency during the period of 2008/3/3-2016/3/14 for 28 individual emerging economies and for all the 61 countries available, respectively. Most countries have about 4 articles on Bloomberg each day. Surprisingly, Dominican Republic has a large number of articles by Bloomberg, most of the publications actually were around June 2010–March 2011 during its election year. Among all countries, Japan had the most articles (24552) on one day, March 11, 2011, when a magnitude-9 earthquake and tsunami hit the country and were followed by a nuclear meltdown.

Table Appendix F.1: Summary Statistics of Bloomberg Daily Publication Number on Sovereigns (2008/3/3-2016/3/14)

Country	Obs.	Median	Max	Std. Dev.
Argentina	888	2	98	9
Brazil	768	4	126	14
Bulgaria	291	2	36	5
Chile	255	2	49	7
China	1695	44	649	46
Colombia	1428	2	56	4
Croatia	542	3	43	5
Dominican Republic	320	149	811	131
Ecuador	693	4	61	7
Hungary	761	3	58	7
India	1350	3	79	7
Indonesia	552	3	89	7
Kazakhstan	824	2	49	5
Lebanon	266	3	27	4
Lithuania	167	3	44	7
Mongolia	19	4	11	3
Nigeria	646	1	45	4
Pakistan	433	2	29	4
Panama	18	5	15	5
Philippines	273	3	57	8
Romania	348	3	47	6
Russia	1771	10	139	15
South Africa	303	2	49	7
Turkey	458	2	94	9
Ukraine	598	4	415	35
Uruguay	388	2	40	5
Venezuela	1182	2	123	9
Vietnam	273	1	63	6
All 61 Countries	34461	4	24552	417

Appendix G. Empirical Analysis: Robustness

Bloomberg readership data has more missing observations than the publication data, nevertheless, we use it to measure investor attention as robustness checks. It is defined as the maximum value of news heat, which reflects user activities (reading and searching news), on a specific country for the day. It is worth noting that news heat readership index does not report the absolute number of times an article being read, instead it measures readers' interest in an issuer relative to the previous 30 days, based on the number of times people call up stories with the issuer's equity ticker attached and the number of times they run the ticker looking for news. A score of 0 indicates readership is not widespread or is below the 30-day average. Scores of 1-4 indicate readership is unusually high, with 4 representing the top of the range. The daily and quarterly regression results are reported in Table Appendix G.1 and are consistent with those in the previous tables. Bloomberg users become much

more interested in an issuer during the issuers' crisis periods.

Table Appendix G.1: Bloomberg Readership and EMBI at Crisis Times

	Daily Data		Quarterly Data	
Expl. Variables	**EMBI**	**Std. Dev.**	**EMBI**	**Std. Dev.**
Readership	−2.49	0.06	1.17	0.01
	(4.21)	(0.24)	(1.38)	(0.04)
Readership× Crisis	36.01	6.78	7.00	0.10
	(12.56)	(2.28)	(1.06)	(0.04)
Other Controls	No	EMBI	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Observations	1269	1269	377	311
Adj. R-squared	0.020	0.319	0.445	0.300

Note: All regressions include fixed effects and robust standard errors. Other controls include government balance as a share of GDP, debt-to-GDP ratio, RGDP growth, inflation, and unemployment rate. For both spread standard deviation regressions, we also include $EMBI$ as a control variable. Data is from Bloomberg and CEIC. Standard errors in parentheses.

References

References

Aguiar, M., Chatterjee, S., Cole, H., and Stangebye, Z. R. (2016). Quantitative models of sovereign debt crises. *Handbook of Macroeconomics, Volume II*.

Aguiar, M. and Gopinath, G. (2006). Defaultable debt, interest rates, and the current account. *Journal of International Economics*, 69(1):64–83.

Andrei, D. and Hasler, M. (2015). Investor attention and stock market volatility. *Review of Financial Studies*, 28(1):33–72.

Angeletos, G.-M. and Werning, I. (2006). Crises and prices: Information aggregation, multiplicity, and volatility. *American Economic Review*, 96(5):1720–1736.

Arellano, C. (2008). Default risk and income fluctuations in emerging economies. *American Economic Review*, 98(3):690–712.

Babenko, I. and Mao, L. (2018). Information acquisition and corporate debt illiquidity. *SSRN Working Paper No. 2517974*.

Bacchetta, P. and van Wincoop, E. (2010). Infrequent portfolio decisions: A solution to the forward discount puzzle. *American Economic Review*, 100(3):870–904.

Barber, B. M. and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21(2):785–818.

Bassetto, M. and Galli, C. (2019). Is inflation default? the role of information in debt crises. *American Economic Review*, 109(10):3556–84.

Ben-Rephael, A., Da, Z., and Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *Review of Financial Studies*, 30(9):3009–3047.

Benjamin, D. and Wright, M. L. J. (2013). Recovery before redemption: A theory of delays in sovereign debt restructuring. *SSRN Working Paper No. 1392539*.

Bi, H. and Traum, N. (2012). Estimating sovereign default risk. *American Economic*

Review, 102(3):161–166.

Bi, H. and Traum, N. (2019). Sovereign risk and fiscal information: A look at the u.s. state default of the 1840s. *Mimeo*.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3):623–685.

Bocola, L. (2016). The pass-through of sovereign risk. *Journal of Political Economy*, 124(4):879–926.

Bocola, L. and Dovis, A. (2019). Self-fulfilling debt crises: A quantitative analysis. *American Economic Review*, 109(12):434–77.

Born, B., Muller, G. J., Pfeifer, J., and Wellman, S. (2020). Different no more: Country spreads in advanced and emerging economies. *CESif Working Paper No. 8083*.

Carlson, M. and Hale, G. B. (2006). Rating agencies and sovereign debt rollover. *Topics in Macroeconomics*, 6(2).

Catao, L. A., Fostel, A., and Kapur, S. (2009). Persistent gaps and default traps. *Journal of Development Economics*, 89(2):271–284.

Chari, V. and Kehoe, P. J. (2003). Hot money. *Journal of Political Economy*, 111(6):1262–1292.

Chatterjee, S. and Eyigungor, B. (2012). Maturity, indebtedness, and default risk. *American Economic Review*, 102(6):2674–2674.

Cole, H. L. and Kehoe, P. J. (1998). Models of sovereign debt: Partial versus general reputations. *International Economic Review*, 39(1):55–70.

Cole, H. L., Neuhann, D., and Ordonez, G. (2016). Debt crises: For whom the bell tolls. *NBER Working Paper No. 22330*.

Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. Wiley-Interscience, Hoboken, NJ.

Croitorov, O. (2016). Sovereign debt and asymmetric market information. *Mimeo*.

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

Cziraki, P., Mondria, J., and Wu, T. (2018). Asymmetric attention and stock returns.

mimeo.

Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *Journal of Finance*, 66(5):1461–1499.

Dang, T. V., Gorton, G., and Holmström, B. (2015). Ignorance, debt and financial crises. *Mimeo*.

de Ferra, S. and Mallucci, E. (2020). Sovereign risk matters: The effects of endogenous default risk on the time-varying volatility of interest rate spreads. *Federal Reserve Board International Finance Discussion Papers 1276*.

Dimpfl, T. and Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, 22(2):171–192.

Dow, J. and Gorton, G. (2006). Noise traders. *NBER Working Paper No. 12256*.

Durdu, C. B., Nunes, R., and Sapriza, H. (2013). News and sovereign default risk in small open economies. *Journal of International Economics*, 91(1):1–17.

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

Fernández-Villaverde, J., Guerron-Quintana, P. A., Rubio-Ramirez, J., and Uribe, M. (2011). Risk matters: The real effect of volatility shocks. *American Economic Review*, 101(6):2530–2561.

Gervais, S., Kaniel, R., and Mingelgrin, D. H. (2001). The high-volume return premium. *The Journal of Finance*, 56(3):877–919.

Grossman, S. (1976). On the efficiency of competitive stock markets where traders have diverse information. *Journal of Finance*, 31(2):573–585.

Grossman, S. and Stiglitz, J. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408.

Hatchondo, J. C. (2004). Asymmetric information and the lack of international portfolio diversification. *Working Paper University of Rochester*.

Hatchondo, J. C. and Martinez, L. (2009). Long-duration bonds and sovereign default. *Journal of International Economics*, 79(1):117–125.

Hatchondo, J. C., Martinez, L., and Sosa-Padilla, C. (2016). Debt dilution and sovereign default risk. *Journal of Political Economy*, 124(5):1383–1422.

Hellwig, C., Pancrazi, R., and de Soyres, C. (2014). Sovereign default and information frictions. *Mimeo*.

Holden, S., Natvik, G. J., and Vigier, A. (2018). Equilibrium rating and debt crises. *International Economic Review*, Forthcoming.

Johri, A., Kahn, S. K., and Sosa-Padilla, C. (2015). Interest rate uncertainty and sovereign default risk. *Mimeo*.

Justiniano, A. and Primiceri, G. E. (2008). The time-varying volatility of macroeconomic fluctuations. *American Economic Review*, 98(3):604–641.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335.

Lizarazo, S. V. (2013). Default risk and risk averse international investors. *Journal of International Economics*, 89:317–330.

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

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, 4(2):103–124.

Manso, G. (2013). Feedback effects of credit ratings. *Journal of Financial Economics*, 109(2):535–548.

Melino, A. and Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics*, 45(1-2):239–265.

Mendoza, E. G. and Yue, V. C. (2012). A general equilibrium model of sovereign default and business cycles. *Quarterly Journal of Economics*, 127(2):889–946.

Mihm, B. (2016). Mispricing of risk in sovereign bond markets with asymmetric information. *German Economic Review*, 17(4):491–511.

Morris, S. and Shin, H. S. (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, 88(3):587–597.

Muller-Itten, M., Armenter, R., and Stangebye, Z. (2020). Rational inattention via ignorance equivalence. *SSRN WP 3521379*.

Pagano, M. and Volpin, P. (2012). Securitization, transparency, and liquidity. *Review of Financial Studies*, 25(8):2417–2453.

Perez, D. J. (2017). Sovereign debt maturity structure under asymmetric information. *Journal of International Economics*, 108:243–259.

Pouzo, D. and Presno, I. (2016). Sovereign default risk and uncertainty premia. *American Economic Journal: Macroeconomics*, 8(3):230–66.

Sandleris, G. (2008). Sovereign defaults: Information, investment and credit. *Journal of International Economics*, 76(2):267–275.

Scheidigger, S. and Bilionis, I. (2019). Machine learning for high-dimensional dynamic stochastic economies. *Journal of Computational Science*, 33:68–82.

Seasholes, M. S. and Wu, G. (2007). Predictable behavior, profits, and attention. *Journal of Empirical Finance*, 14(5):590–610.

Sedlacek, P. (2016). Growth and uncertainty over the business cycle. *CEPR Working Paper No. 11296*.

SenGupta, A. (1987). Tests for standardized generalized variances of multivariate normal populations of possibly different dimensions. *Journal of Multivariate Analysis*, 23(2):209–219.

Seoane, H. D. (2015). Time-varying volatility, default and the sovereign risk premium. *Mimeo*.

Shannon, C. E. (1958). Channels with side information at the transmitter. *IBM Journal of Research and Development*, 2(4):289–293.

Sims, C. A. (2003). The implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

Sosa-Padilla, C. (2012). Sovereign defaults and banking crises. *MPRA Paper No. 41074*.

Stangebye, Z. R. (2015). Dynamic panics: Theory and application to the eurozone. *MPRA Working Paper No. 69967*.

Tomz, M. and Wright, M. L. J. (2007). Do countries default in "bad times"? *Journal of the European Economic Association*, 5(2/3):352–360.

Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, 64(3):1187–1215.

Veldkamp, L. (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.

Vlastakis, N. and Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking and Finance*, 36:1808–1821.

Wittenberg-Moerman, R. (2010). The impact of information asymmetry on debt pricing and maturity. *Mimeo*.