Dissecting Mechanisms of Financial Crises: Intermediation and Sentiment

Arvind Krishnamurthy and Wenhao Li[*]

October 2021

Abstract

We develop a model of financial crises with both a financial amplification mechanism, via frictional intermediation, and a role for sentiment, via time-varying beliefs about an illiquidity state. We confront the model with data on credit spreads, equity prices, credit, and output across the financial crisis cycle. In particular, we ask the model to match data on the frothy pre-crisis behavior of asset markets and credit, the sharp transition to a crisis where asset values fall, disintermediation occurs and output falls, and the post-crisis period characterized by a slow recovery in output. Our model with the frictional intermediation mechanism and fluctuations in beliefs provides a parsimonious account of the entire crisis cycle. The model with only the frictional intermediation mechanism misses the frothy pre-crisis behavior; fluctuations in beliefs resolve this problem. On the other hand, modeling the belief variation via *either* a Bayesian or diagnostic model match the broad patterns, with each missing some targets to different extents. We also show that a lean-against-the-wind policy has a quantitatively similar impact in both versions of the belief model, indicating that policy need not "get into the minds" of investors and condition on the true belief process.

1 Introduction

Financial crises have a common character. There is a pre-crisis period that is marked by a runup in credit, leverage, low risk spreads, and an expansion in output. Credit and asset valuations appear frothy before a crisis. The transition to the crisis is sharp. There are losses to the financial sector, defaults and bank-runs, a jump in risk spreads, and contraction in credit and output. The aftermath of the crisis is a gradual recovery in credit, output, and fall in risk spreads. These patterns emerge from a large and growing body of research examining financial crises episodes across countries and time, dating back to the 19th century. See Bordo et al. (2001), Borio and Lowe (2002), Claessens, Kose and Terrones (2010), Reinhart and Rogoff (2009a), Schularick and Taylor (2012), Jordà, Schularick and Taylor (2011), Laeven and Valencia (2013), Jordà, Schularick and Taylor (2013), Baron and Xiong (2017), Krishnamurthy and Muir (2020), and Baron, Verner and Xiong (2021). This empirical research describes and quantifies these common patterns.

Theoretical research on crises has fallen into two categories. The first emphasizes frictions in financial intermediation that drive an amplification mechanism. The key idea is that the fragility of the financial sector, measured typically as high leverage or low levels of equity capital-to-assets, is an endogenous state variable. An unexpected large-loss event hitting the economy in a state where the financial sector is fragile sets in motion mechanisms whereby the shock is amplified, there is disintermediation, a rise in risk spreads and contraction in output. Recovery takes time, tracking a gradual re-intermediation. The amplification model speaks directly to the transition to crisis and the aftermath of the crisis. See work by Gertler and Kiyotaki (2010), He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014), He and Krishnamurthy (2019), and Li (2019).

The second line of research emphasizes the role of information and beliefs, and harkens back to Kindelberger (1978). There are two key ideas in this research. First, agents experience a period of prosperity and come to believe that risks are low. Second, the crisis is a informational event – a "Minsky (1992) moment" – where risk is re-assessed leading to swings in asset prices, credit, and macroeconomic outcomes. In the work of Gorton and Ordonez (2014) and Dang, Gorton and Holmström (2020), the shift in beliefs occurs because financial sector information is hidden, by design, during prosperous periods, and a crisis is the event when negative information comes to light and agents reassess risks. The shift from no-information to information is at the heart of their narrative of crises. The work of Bordalo, Gennaioli and Shleifer (2018) has instead argued that a sharp shift in beliefs in a crisis reflects a change from over-optimistic to over-pessimistic beliefs. Extrapolative expectations are at the heart of their narrative of the belief shift in a crisis.

This paper builds a model that integrates both of these elements, frictional financial intermediation and time-variation in beliefs, into a quantitative macro-finance model. Our

objective is to understand the extent to which these mechanisms can account qualitatively and quantitatively for the macro crisis patterns, and to clarify which elements of these mechanisms are essential. Our model has a financial intermediary sector subject to capital constraints and financed in part by demandable debt. There are two sources of shocks, a Brownian shock to the return on capital and an illiquidity shock where the market for capital assets temporarily freezes up, and debtors refuse to roll over their debts, as in a bank run. In this latter state, sales of capital assets by banks incur a liquidation cost, or alternatively, loans against capital are charged an illiquidity premium. The economy transits through booms and busts driven by the Brownian shock and its impact on the dynamics of real capital and the equity capital of the financial sector. Crises are events where both the financial sector equity capital is low and the illiquidity shock occurs. In this case, there are runs on banks leading to disintermediation, declines in asset values, and a reduction in output. The illiquidity shock captures a financial panic, such as occurred in both fall 2008 and spring 2020, with differences in macroeconomic outcomes driven in part by differences in financial sector fragility. We also note that our illiquidity shock impacts the economy indirectly via a financial amplification mechanism and not directly via its impact on productivity and output. This approach to modeling leads to endogenous crises in which the financial sector is the key factor. The modeling is motivated by our objective to shed light on financial crises such as the 2008 global financial crisis and not on rare consumption disasters such as the 2020 COVID recession. The financial frictions model of our paper is a variant of Li (2019). It draws on ingredients from the recent macro-finance literature on financial crises and intermediation frictions, and particularly He and Krishnamurthy (2013); Brunnermeier and Sannikov (2014); Gertler and Kiyotaki (2015).

Agents in the economy make decisions based on their beliefs about the likelihood of the illiquidity shock. The illiquidity shock is a Poisson event, the intensity of which follows a hidden two-state Markov process. Agents infer the state and hence the likelihood of the illiquidity shock based on history. A string of no-shock realizations leads them to believe that shocks are unlikely (i.e., the true state is the low-intensity state). A shock occurrence leads them to think that shocks are more likely (i.e., the true state is the high-intensity state). After an extended period with no shocks, banks downplay liquidity risk and increase leverage. The shock triggers a "Minsky moment:" agents' beliefs regarding liquidity risk rises and is then amplified and propagated to the macroeconomy depending on the leverage of the financial sector.

We consider two flavors of the learning mechanism, a Bayesian updating process closest to Moreira and Savov (2017) and non-rational updating process, along the lines of Bordalo, Gennaioli and Shleifer (2018), where beliefs over-react to current news.[1]

---

[1]The diagnostic updating process is motivated by the work of Bordalo, Gennaioli and Shleifer (2018), and is also related to the models of Greenwood, Hanson and Jin (2019) and Maxted (2019). Bordalo, Gennaioli and Shleifer (2018), Bordalo et al. (2019b), and Bordalo et al. (2020) examine data on survey

We report four principal results:

1. The model with financial frictions and a time-varying belief process matches the main features of the pre-crisis, crisis, and aftermath.

2. A model with only financial frictions generates the amplification needed to match crisis and post-crisis patterns but fails to match the pre-crisis froth evidence. That is, both a financial frictions mechanism and a mechanism involving fluctuations in beliefs are needed to match the crisis cycle evidence.

3. While belief fluctuations are essential, whether one needs a Bayesian belief process or the non-rational diagnostic process to fit the crisis patterns is more murky. The diagnostic belief model, calibrated to the evidence from Bordalo, Gennaioli and Shleifer (2018), matches the crisis patterns qualitatively. But so does the Bayesian belief model. These two learning variants each fit different dimensions of the data better (and worse), with the diagnostic model's principal success over the Bayesian model being that the model's pre-crisis froth is quantitatively closer to the data.

4. The impulse responses of both the Bayesian and diagnostic model, conditional on a state chosen to match the same measured credit spread and bank leverage, are quantitatively similar. That is, for many policy experiments, distinguishing between agents' true learning mechanism is not necessary.

The model has two key state variables: one governing the wealth-share of bankers and the other describing agents' beliefs over the intensity of the illiquidity shock. The wealth-share variable, coupled with financial frictions, governs a financial amplification mechanism studied in prior work. We show that this amplification mechanism helps the model match data on the crisis and its aftermath. In particular, the financial amplification mechanism of the model generates a sharp drop in asset prices, credit, and output. The mean drop in our model is in line with the data, but more telling, the skewness of these variables and their comovement also aligh with data counterparts. That is, a key feature of financial crises is non-linearity, reflected in a skewed distribution of output declines. The model's amplification mechanism generates skew in line with that of the data. The model also generates a slow recovery due to the persistence mechanism of financial frictions models.

While the financial frictions wealth-share mechanism is the key to understanding the model's match of the crisis and aftermath, the belief state variable is needed for the model to match the pre-crisis patterns. If an illiquidity shock has not occurred for some time,

---

agent beliefs drift towards the low likelihood state. Bankers choose to increase leverage as they are less concerned about liquidity risk. Risk and credit spreads fall and credit grows. From this state, if an illiquidity shock arrives, beliefs jump towards the high likelihood state and banker wealth falls and they endogenously choose less leverage given the heightened liquidity risk. Both forces lead to financial amplification of the shock and persistence as in a crisis. The belief mechanism helps explain why spreads are low and credit is high *before* the crisis. More surprisingly, low spreads and high credit help *predict* a crisis. The reason is that bankers act more risk-tolerant in the pre-crisis period when liquidity risk is low – they drive down spreads/risk premia and increase credit. They also increase leverage, taking liquidity risk, and effectively shift GDP outcomes into tail states. It may be surprising that we find that there are times when crises are more likely and yet risk prices are low and bankers take more leverage. Our model ties these observations together by generating bankers' willingness to take on illiquidity risk in the pre-crisis period, driven by the beliefs state variable.

We probe this model in two dimensions. First, we find that if the belief intensity is held constant (i.e., no learning mechanism, but illiquidity shocks still occur), the model fails to match the pre-crisis patterns. In such a model, only the banker wealth-share is a state variable. The fragility of the economy to a crisis is measured by the banker wealth-share state variable. When this is low, a negative shock triggers a crisis. Thus a crisis is more likely when banker wealth is low (and leverage is high). However, this means that forward-looking asset prices will account for the increased fragility as the wealth share state variable falls. As a result, the model implies that credit spreads will rise, and bank credit will fall in the period before a crisis, contrary to the data. On the other hand, we find that this static-belief model is able to match the data for the crisis and its aftermath, clarifying that the financial amplification mechanism drives these patterns. We also show that this model generates a negative relation between bank credit and equity market excess returns (risk premia), as documented by Baron and Xiong (2017). This occurs in our model because variation in the wealth-share drives variation in bankers' risk tolerance that generates the required comovement between credit and risk premia. It is worth emphasizing that this result arises in a model with no variation in beliefs.

Second, we compare the Bayesian and diagnostic learning mechanisms. There are differences in the magnitude of the crisis and post-crisis match: the diagnostic model generates too much comovement, relative to data, between bank credit and crisis, for example. But the principal difference is in the amount of pre-crisis froth of these models. The key force in matching the pre-crisis evidence is the sensitivity of the bankers' endogenous leverage decision to the true (not perceived) illiquidity state. This sensitivity has the required sign (negative) in both learning mechanisms, but the sensitivity is higher under diagnostic beliefs. In our calibrated models, the Bayesian mechanism gets about half-way to matching the data quantitatively, while the diagnostic mechanism brings the froth in the model gets

4

even closer to the data.

Putting this together, our analysis indicates that a financial amplification mechanism plus a belief mechanism provides a parsimonious account of the main crisis facts. The static belief model fails to match the pre-crisis froth in the data. However, our analysis also indicates that the qualitative patterns of the data do not clearly distinguish the two belief models we consider. Within the bounds of how much one can push the quantitative fit of our parsimonious equilibrium model, both models fit the moments we consider.

Lastly, we ask the question, does it matter for policy purposes whether we are living in a world with diagnostic beliefs or one with Bayesian beliefs? It is well understood that in models with financial frictions, leverage restrictions can improve welfare by alleviating a pecuniary externality. See Bianchi (2011) for example. Our model fits within this framework. When agents have distorted beliefs, leverage policies can also improved welfare under a paternalistic welfare criterion. See Dávila and Walther (2020). Thus in our model, it is interesting to study a leverage restriction and ask how the nature of beliefs affects policy assessment.

We consider an unexpected policy that transfers wealth from households to bankers so that the banker wealth share increases by 10% during a pre-crisis boom period. The policy captures the impact of increasing bank equity (or equivalently, reducing bank leverage) to lean-against-the-wind, along the lines of Gertler, Kiyotaki and Prestipino (2020). Under each version of the model, we pick an initial condition in terms of credit spreads and bank leverage and map these into the state variables in each model (they map to different values of the state variable across the models). We then simulate the path of the economy with and without the recapitalization policy. We calculate the difference in quantities and prices between the with- and without-recapitalization and repeat this across both models. Our main finding is that these impulse response differences are quite similar across both models. The policy raises the mean path of output and credit, and the conditional response of these variables to an illiquidity shock, but these responses are quantitatively similar across both models. The two keys to this similarity result are that (1) both models are calibrated to common data but are not forced to having a common parameterization, and (2) both models are tied to the same initial condition in terms of observables. In particular, the diagnostic model does not just take the Bayesian model parameters and add a new diagnostic parameter. In this case, which is not economically meaningful, the impulse responses are no longer similar.

The main contribution of this paper is to bridge the recent theoretical work on non-linear macro-finance models (He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; Di Tella, 2017) with the empirical literature on financial crises cited earlier. The models in this theoretical literature feature non-linearities and are solved using global methods. Thus these models are well-suited to characterize the non-linear dynamics in financial crises. But the work thus far is either purely theoretical or aims to match a single crisis-event (e.g.,

the 2007-2009 financial crisis in He and Krishnamurthy (2019) and Gertler, Kiyotaki and Prestipino (2020)). The empirical crisis literature on the other hand has largely documented systematic patterns in the data rather than assess this data from the standpoint of models. Our paper bridges this gap.[2]

This paper's objective of matching the boom-bust of the crisis cycle is closest to that of a few other papers that precede ours.[3] Boissay, Collard and Smets (2016) develop a dynamic model of banking crises that generates the pattern in line with the data that credit booms precede credit market collapse and crises. The key idea in the model is that banks' absorption capacity is reduced during a boom, with the economy potentially hitting a cliff where the credit market collapses. Thus, the probability of a financial crisis rises in the boom. Relative to their analysis, we aim to match the asset market fact that risk premia and credit spreads are low during the boom, which we reconcile with the learning mechanism of our model. Greenwood, Hanson and Jin (2019) and Maxted (2019) construct models of the boom-bust crisis cycle with a role for beliefs. Greenwood, Hanson and Jin (2019) present a model where lenders extend credit based on beliefs over the default probabilities of borrowers. There is a feedback between realized default and beliefs regarding default probabilities, similar to the model of this paper, that creates a persistence and amplification mechanism. Like us, their paper aims to match facts on credit growth, credit spreads, and risk premia. But their model is not a full macroeconomic model, and thus does not speak to other macroeconomic data such as output and the conditional distribution of output growth. Their model also does not have an intermediary sector, so it cannot assess the role of intermediary frictions relative to beliefs. Finally, lenders are risk-neutral in their model, so that without diagnostic expectations, risk premia are zero. As a result, their model does not give the Bayesian belief process a chance of explaining the data. Maxted (2019)'s macrofinance model is closer to ours. There is an intermediation sector that is central to crisis dynamics. The paper also considers a full macroeconomic setting, and can thus speak to more macro data. One key point of difference relative to our model is that Maxted (2019)'s diagnostic belief modeling extrapolates the mean growth of capital productivity, whereas in our model, beliefs over a tail illiquidity shock are distorted.[4] Thus in Maxted (2019),

---

[2]Another contribution of our paper to the non-linear macro finance work is the model. A major disadvantage of the current models is that they are computationally challenging, and current models restrict attention to one or two-state variables following a Brownian diffusion process. In this paper, we present and solve a model with two state variables and endogenous jumps. Our methodology helps broaden the scope of the literature to encompass richer dynamics with sudden and large disruptions, which are plausibly central to financial crises.

[3]There are other recent macro-finance papers, not explicitly about the boom-bust cycle, but that aim to match crisis facts. Gertler, Kiyotaki and Prestipino (2020) introduces bank runs into a macro-intermediation model. Beliefs, modeled via a sunspot, play a role in driving crisis dynamics. The objective of their paper is to study the 2007-2009 financial crisis rather than disentangling mechanisms underlying the crisis cycle facts. Camous and Van der Ghote (2021) builds on Maxted (2019) and considers diagnostic expectations and financial frictions in a multi-sector model. The model can generate a build-up of instability and a safety trap with low growth. Gopalakrishna (2020) introduces state-dependent bank exit into a quantitative continuous-time macro-finance model and generates a slow recovery in line with empirical evidence.

[4]Ma, Paligorova and Peydro (2021) presents survey evidence that assessments by banks over the downside

optimism occurs after a period of growth in productivity which can sow the seeds of the crisis, while in our mechanism optimism occurs after a quiet period of no illiquidity shocks, leading endogenously to increased risk-taking. There is evidence that productivity booms are good booms that do not lead to crises (Gorton and Ordonez, 2020), and certainly the period before the 2008 financial crisis was marked by a slow-down in productivity growth (Fernald, 2015).[5]

Finally, this paper also contributes to a larger literature on beliefs and learning in macroeconomics models. Closest to our paper is Kaplan, Mitman and Violante (2020) who dissect the U.S. housing boom-bust cycle around the 2008 crisis to evaluate the role of beliefs and financial constraints in driving the cycle. They conclude that a shift in beliefs during the boom are essential to matching the cycle. Note that we consider banking crises and not housing crises, and broaden our scope to include patterns across many crisis episodes. Van Nieuwerburgh and Veldkamp (2006) show that asymmetry in learning about productivity can generate asymmetries in business cycles. Simsek (2013) explores the interaction of beliefs and credit, building a model where beliefs over upside versus downside payoffs have an asymmetric impact on asset valuations, total credit and fragility of the economy. Simsek (2013) studies the role of belief heterogeneity, which is absent in our model with homogeneous beliefs. Motivated by the slow recovery from the 2008 recession, there is research tying learning to slow recoveries. In Fajgelbaum, Schaal and Taschereau-Dumouchel (2017), information flows slowly in times of low activity and uncertainty remains high, discouraging investment. Liu, Wang and Yang (2020) show that the uncertainty and learning about banks' peers can lead to a slow recovery. In Kozlowski, Veldkamp and Venkateswaran (2020), agents learn about the parameters of the economic shock process, and a large negative shock realization as in a deep recession alters agents' estimates of these parameters, leading to a persistent impact of the shock on economic growth. Bordalo et al. (2019a) introduce diagnostic beliefs into a relatively standard real business cycle model. Their model helps to understand the role of diagnostic beliefs in driving business cycles.

The rest of this paper is as follows. In Section 2, we review general patterns of the crisis cycle in the data. In Section 3, we set up a model that combines financial intermediation frictions and beliefs regarding an illiquidity shock. In Section 4, we solve and explain how we calibrate the the model(s). In Section 5, we evaluate the model, explaining its fit and the role of beliefs. In Section 6, we consider how the Bayesian and diagnostic models may inform policy. We then conclude in Section 7. An appendix follows.

---

tail, more than the mean of economic forecasts, explain bank lending decisions.

[5]Farboodi and Kondor (2020) present a model of time-varying sentiment that generates a credit cycle that is qualitatively in line with the facts. Sentiment evolves in a Bayesian manner in their model. Thus, like us, they show that the basic facts of the credit cycle can be generated within a Bayesian model. The objective of the paper is different than ours, as their model is not suited to a quantification exercise and does not have an intermediary sector.

2 The Crisis Cycle

This section reviews broad patterns of the crisis cycle, drawn from the empirical literature on crises. Along the way, we list (numbered below) specific quantitative estimates from the literature which guide our modeling exercise.

What is a financial crisis? Jordà, Schularick and Taylor (2011) state:

> In line with the previous studies, we define financial crises as events during which a country's banking sector experiences bank runs, sharp increases in default rates accompanied by large losses of capital that result in public intervention, bankruptcy, or forced merger of financial institutions.

We focus on events, as per the quotation, as financial crises. These events are banking crises and do not necessarily include currency crises or sovereign debt crises, which are other crises of interest, unless such events coincide with a banking crisis. Jordà, Schularick and Taylor (2011)'s dating of banking crises is closely related to the approach of Bordo et al. (2001), Reinhart and Rogoff (2009a), and Laeven and Valencia (2013). Bordo and Meissner (2016) discuss the approaches that researchers have taken to crisis-dating as well the drawbacks of different approaches.

1. We target an unconditional frequency of financial crises of 4%. In an article written for the Annual Review of Economics, Taylor (2015) reports the historical frequency of financial crises to be 6%. This data point is obtained from a sample of countries in both developing and advanced stages, and covers the period after 1860. The Handbook of Macroeconomics chapter by Bordo and Meissner (2016) reports numbers in the range of 2 to 4% across the studies by Bordo et al. (2001) and Reinhart and Rogoff (2009a). Another evidence comes from Jordà, Schularick and Taylor (2013), which shows that the average frequency of crises is 3.6% using data from multiple countries.

Figure 1 plots the mean path of credit spread, credit, and GDP across a sample of 41 international financial crises identified by Jordà, Schularick and Taylor (2013). The figure is drawn from Krishnamurthy and Muir (2020), which includes data on credit spreads relative to other studies of crises. Date 0 on the figure corresponds to the date of a financial crisis. The top-left panel plots the path of the mean across-country credit spread, relative to the mean spread for country-i, from 5-years before the crisis to 5-years after the crisis. The units here are that 0.4 means that spreads are 0.4σs larger than the country's time-series average spread, while -0.2 means that spreads are 0.2σs below the country's time-series average. The data is annual from 14 countries spanning a period from 1879 to 2013.

We see that spreads run below their average value in the years before the crisis. They rise in the crisis, going as high as 0.4σs over their mean value in the year after the crisis
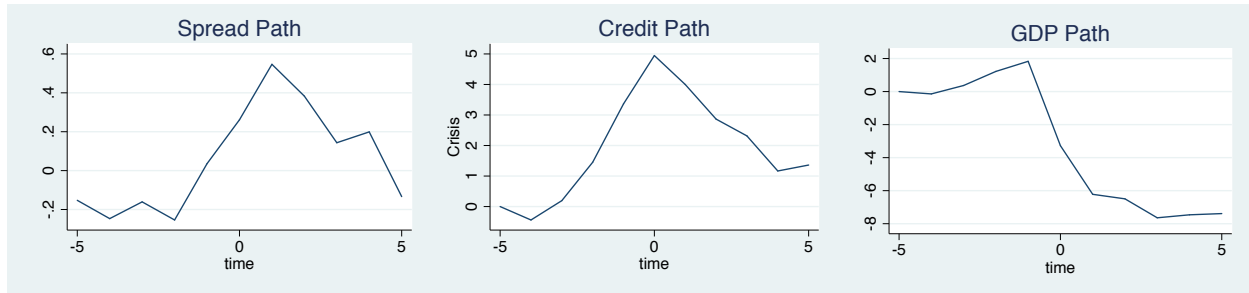
Figure 1: Mean path of credit spread, bank credit, and GDP across a sample of 41 financial crises identified in Jordà, Schularick and Taylor (2013). Units for spread path are 0.4 means that spreads are 0.4σs above their average for a given country. Units for credit path are that 5 indicates that credit/GDP is 5% above the trend for a given country. Units for GDP path are that -8 means that GDP is 8% below trend for a given country. Source: Krishnamurthy and Muir (2020)

date, before returning over the next 5 years to the mean value. The half-life of the credit spread recovery is 2.5 years in this figure.

The top-right panel plots the path of the quantity of bank credit divided by GDP. The credit variable is expressed as the average across-country percentage change in the quantity of credit/GDP from 5-years before the crisis to a given year, after demeaning by the sample growth rate in credit for country-i. The value of 5 for time 0 means that credit/GDP is 5% above the country trend. We see that credit grows faster than average in the years leading up to the crisis at time zero. After this point, credit reverses so that by time $+5$ the variable is back near the country average.

The bottom-left panel plots GDP, again as an average percentage change from 5-years before the crisis, after demeaning by the sample growth rate in GDP for country-i. GDP grows slightly faster than average in the years preceding the crisis. GDP falls below trend in the crisis and remains low up to 5 years after the crisis.

Transition to crisis: A crisis is characterized by a sharp jump in credit spreads, a reversal in the quantity of credit and a decline in GDP. From the data underlying Figure 1:

2. Credit spreads rise by 0.7σs of their mean value at the crisis.

3. GDP declines by 9.1%. Reinhart and Rogoff (2009b) report a peak-to-trough decline in GDP across a larger sample of crises of 9.3%. Jordà, Schularick and Taylor (2013) report a 5-year decline in GDP from the date of crisis of around 8%. Cerra and Saxena (2008) report output losses from banking crises of 7.5% with these losses persisting out to 10 years. We will use the 9.1% number in our quantitative exercise.

The rise in credit spreads in the year of the crisis is mirrored in other asset prices. Reinhart and Rogoff (2009a) report that equity prices decline by an average of 55.9% during banking crises. Muir (2017) shows that the price-dividend ratio on the stock market falls in
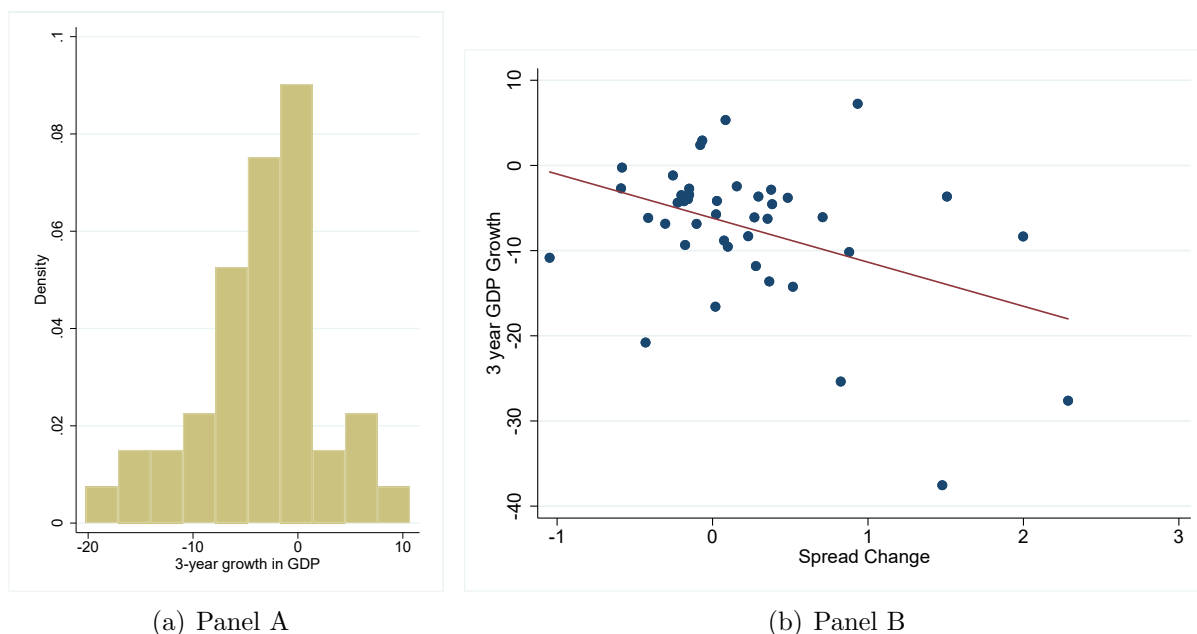
(a) Panel A	(b) Panel B

Figure 2: Panel A presents a histogram of 3-year GDP growth from the start of a crisis, as dated by Jordà, Schularick and Taylor (2013). Panel B presents a scatter plot of the spike in spreads in the year of the crisis against 3-year GDP growth after the crisis..

a crisis, and the excess return on stocks rises during the crisis, indicated a generalized rise in asset market risk premia.

Aftermath and severity of crisis:

4. The half-life of the recovery of the credit spread to its mean value is 2.5 years.

5. There is variation in the severity of the crisis. Figure 2, Panel A presents data on the variation in the severity of the crisis, as measured by 3-year GDP growth following a crisis. The figure reflects significant variation in crisis severity.

6. The variation in the severity of the crisis is correlated with the increase in spreads measured at the transition into the crisis, as illustrated in Figure 2, Panel B. Krishnamurthy and Muir (2020) report a coefficient of -7.46 (*s.e.* 1.46) from a regression of 3-year GDP growth following a crisis on the increase in credit spreads from the year before the crisis to the year of a crisis.

Pre-crisis period: In the pre-crisis period, credit markets appear frothy, reflecting low credit spreads and high credit growth. In particular,

7. Conditioning on a crisis at year t, and looking at the 5 years prior to the crisis, Krishnamurthy and Muir (2020) show that credit spreads are 0.34σs below their country mean (where this country mean is defined to exclude the crisis and 5 years after the crisis).

8. Conditioning on a crisis at year t, credit/GDP in the 5 years before the crisis is 5% above country mean. The relation between a lending boom and subsequent crisis is well documented in the literature. See Gourinchas et al. (2001), Schularick and Taylor (2012), and Baron and Xiong (2017).

Predicting Crises: There is also evidence that periods of frothy conditions predict and not just precede crises. There are two quantitative estimates that we will aim to match.[6]

9. Schularick and Taylor (2012) find that a one-standard deviation increase in credit growth over the preceding 5 years (= 0.07 in their sample) translates to an increased probability of a financial crisis of 2.8% over the next year.

10. Conditioning on an episode where credit spreads are below their median value 5 years in a row, Krishnamurthy and Muir (2020) estimate that the conditional probability of a crisis rises by 16% over the next 5 years.

11. Baron and Xiong (2017) find that a one-standard deviation increase in credit growth over the preceding 3-years increases the probability of bank equity crash (defined as decline in bank equity by over 30%) by 5.4%.

3   A Model of Financial Crises with Amplification and Sentiment

In this section, we present a model of financial crises that incorporates both a financial amplification mechanism and a role for sentiment. We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and assume all stochastic processes are adapted to this space and satisfy the usual conditions. The economy evolves in continuous time. It is populated by a continuum of a unit mass of two classes of agents, households, and bankers. For clarity, aggregate variables are in capital letters, and individual variables are in lower case letters. The basic setup is a variant of Li (2019), which is drawn from Brunnermeier and Sannikov (2014) and Kiyotaki and Moore (1997).

3.1   Agents and Assets

Households maximize expected value of the discounted log utility,

$$\int_0^\infty e^{-\rho t} \log(c_t^h) dt \tag{1}$$

---

[6] Greenwood et al. (2020) present further evidence in line with froth predicting crises. In post-war cross-country data, they document that periods of high credit growth coupled with periods of high returns in the stock market substantially increase the likelihood of a financial crisis.

and bankers optimize expected value of the same form of discounted log utility,

$$\int_0^\infty e^{-\rho t} \log(c_t^b) dt \tag{2}$$

The expectation could be either Bayesian or diagnostic, as we will specify later.

Output is produced by capital. We will simplify by assuming that the capital is held directly by either banks or households. In a richer and more realistic model, the capital will be held and operated by firms that receive loans from banks or households, along the lines of Holmstrom and Tirole (1997). We simplify by collapsing firms into banks, and assuming the banks own the capital.

We assume that credit flowing through banks allows the economy to achieve higher output and returns to capital. Intermediation is a socially valuable service, and for example, disintermediation in a crisis reduces output. We capture this feature by assuming that banker-operated capital has productivity $\bar{A}$, which is higher than the household-operated capital productivity of $\underline{A}$.

The dynamic evolution of productive capital owned by agent $j \in \{\text{banker}, \text{household}\}$ is

$$\frac{dk_{j,t}}{k_{j,t}} = \mu_t^K dt - \delta dt + \sigma^K dB_t \tag{3}$$

where the rate of new capital installation μ_t^K is endogenously determined through investment, δ is the exogenous depreciation rate, and σ^K is exogenous capital growth volatility.

Denote the price of productive capital as p_t (i.e., "q" in the standard Q-theory). Investment undertaken by an owner, either banker or household, of productive capital is chosen to solve:

$$\max_{\mu_t^K} \ p_t \mu_t^K - \phi(\mu_t^K),$$

where $\phi(\cdot)$ is an investment adjustment cost:

$$\phi(\mu^K) = \mu^K + \frac{\chi}{2}(\mu^K - \delta)^2. \tag{4}$$

That is, we assume quadratic costs to investment, leading to the q-theory of investment

$$p_t = \phi'(\mu_t^K) \quad \Rightarrow \quad \mu_t^K = \delta + \frac{p_t - 1}{\chi}. \tag{5}$$

The dynamics of capital price p_t is denoted as

$$\frac{dp_t}{p_{t-}} = \mu_t^p dt + \sigma_t^p dB_t - \kappa_{t-}^p dN_t, \tag{6}$$

where μ_t^p, σ_t^p, and κ_{t-}^p are all endogenously determined. The "minus" notation (i.e. p_{t-})

reflects a pre-jump asset price, as will be made clear.

3.2 Financing, Liquidity Risk, and Bank Runs

The Brownian shock dB_t in equation (3) reflects business cycle fluctuations in the effective productivity of capital. We introduce a second shock that we call a "financial illiquidity" shock. We model this as a Poisson shock dN_t that triggers illiquidity and bank runs, and a possible financial crisis if the endogenously chosen bank leverage is sufficiently high.

Since banker held capital is more productive than household held capital, there is room for an intermediation relationship whereby households provide some funds to bankers to invest in capital. We assume that the only form of financing is short-term (instantaneous) debt at the [endogenous] interest rate r_t^d. Bankers cannot raise equity, long-term debt, or other forms of financing. When we refer to bank equity, we mean the net-worth of bankers, w_t^b. That is, the financing side of the model is one of inside equity and outside short-term debt. These model simplifications do sweep aside important issues, but we nevertheless go down this path because we aim to build a simple quantitative amplification mechanism and see how well it matches data, rather than explore the micro-foundations of intermediary models.

We assume that in the event of an illiquidity shock, all short-term debt holders run to their own bank and withdraw financing in a coordinated fashion. Raising resources to cover this withdrawal is temporarily costly. That is, asset markets are temporarily illiquid in the illiquidity event. We assume that a cost of α is incurred when capital is liquidated to meet the funding withdrawal during the illiquidity shock. We can think of this cost as liquidation cost or, alternatively, the cost can be mapped into a premium on raising emergency financing from other banks or other households in the economy against the capital. In this latter case, we need to step outside the modeling and interpret the illiquidity event lasting longer than dt. Then, α is proportional to the spread over the riskless rate that the bank pays to obtain funds over the illiquidity episode (if the event lasts dt then a financing spread maps into a cost of order dt). Finally, we assume that the cost is not dissipated but is paid to households proportional to their wealth. This assumption is not essential to the analysis but ensures that the illiquidity shock has no direct impact on output.

The illiquidity shock captures a financial panic, such as occurred in both fall 2008 and spring 2020, with differences in macroeconomic outcomes driven by differences in financial sector fragility. We also note that our illiquidity shock impacts the economy indirectly via a financial amplification mechanism and not directly via its impact on productivity and output as would arise in a rare consumption disasters model. Our approach to modeling leads to endogenous crises in which the financial sector is the key factor.

Note that we do not model a Diamond and Dybvig (1983) bank-run game. We simply assume that the shock leads all debtors to pull their funding. It is possible to model the game in detail following Li (2019) whose model is the basis for this paper. However, we learn from that study that the model's positive implications are almost the same with and without the deeper model of the bank-run game.

3.3 Beliefs and Crises

The intensity of the illiquidity shock process dN_t follows a two state continuous-time Markov process, $\tilde{\lambda}_t \in \{\lambda_L, \lambda_H\}$. This intensity changes from λ_L to λ_H at rate $\lambda_{L \to H}$, and changes from λ_H to λ_L at rate $\lambda_{H \to L}$. Agents, neither bankers nor households, observe $\tilde{\lambda}_t$. Instead agents infer $\tilde{\lambda}_t$ from observing the history of N_t, i.e., via realizations of the shock process.

We denote the Bayesian expectation as $\lambda_t = E_t[\tilde{\lambda}_t]$. Using Bayes rule,

Lemma 1 (Bayesian Belief Process).

$$d\lambda_t = \left(\begin{array}{c} (\lambda_L - \lambda_{t-})\lambda_{H \to L} + (\lambda_H - \lambda_{t-})\lambda_{L \to H} \\ -(\lambda_{t-} - \lambda_L)(\lambda_H - \lambda_{t-}) \end{array} \right) dt + \frac{(\lambda_{t-} - \lambda_L)(\lambda_H - \lambda_{t-})}{\lambda_{t-}} dN_t \quad (7)$$

Therefore, if illiquidity occurs, the expected intensity λ_t jumps up. As time goes by, without further illiquidity shocks, the expected intensity λ_t gradually falls.[7]

3.4 Diagnostic Expectations

We also consider a version of our model where agents overweight recent observations motivated by the diagnostic belief model of (Bordalo, Gennaioli and Shleifer, 2018). We adapt their model to our continuous time dynamic equilibrium environment.

Denote the Bayesian belief for the probability of $\tilde{\lambda}_t = \lambda_H$ as π_t, and the diagnostic belief for the probability of $\tilde{\lambda}_t = \lambda_H$ as π_t^θ. Then we define the diagnostic beliefs as

$$\pi_t^\theta = \pi_t \cdot \left(\frac{\pi_t}{E_{t-t_0}[\pi_t]} \right)^\theta \frac{1}{Z_t} \quad (8)$$

$$1 - \pi_t^\theta = (1 - \pi_t) \cdot \left(\frac{1 - \pi_t}{E_{t-t_0}[1 - \pi_t]} \right)^\theta \frac{1}{Z_t} \quad (9)$$

---

[7]In theory, when $\lambda_t \to \lambda_L$, the drift of $d\lambda_t$ can be positive. The reason is that the underlying intensity process $\tilde{\lambda}_t$ switches between λ_L and λ_H and the average is between the two values, so when λ_t is close to λ_L the dynamics of $\tilde{\lambda}_t$ dominates the information effect and $d\lambda_t$ is positive. However, states with positive $d\lambda_t$ are transient, i.e., λ_t never get back to those states once it drifts outside. In the long run, those states are reached with zero probability and do not matter quantitatively.

where Z_t is a normalization to ensure that (8) and (9) add up to 1. We call the lag t_0 as the "look-back period," which is one in the discrete time model of Bordalo, Gennaioli and Shleifer (2018). In our case, the diagnostic beliefs of the process are simply distorted Bayesian beliefs with the benchmark from t_0 time ago. The process π_t^θ features both overreaction and underreaction, depending on the gap between current π_t and past π_{t-t_0}.

Denote the diagnostic belief for the expected intensity of illiquidity shocks as

$$\lambda_t^\theta = E_t^\theta[\tilde{\lambda}_t] :\stackrel{\Delta}{=} \pi_t^\theta \lambda_H + (1 - \pi_t^\theta)\lambda_L$$

where E^θ is the expectation with respect to the probability distribution under the diagnostic belief. Then we have the following result:

Lemma 2 (Diagnostic Belief Process). *The diagnostic belief* $\lambda_t^\theta = E_t^\theta[\tilde{\lambda}_t]$ *is*

$$\lambda_t^\theta = \lambda_L + (\lambda_t - \lambda_L)\frac{(\lambda_H - \lambda_t) + (\lambda_t - \lambda_L)}{(\frac{\lambda_t^T - \lambda_L}{\lambda_H - \lambda_t^T} / \frac{\lambda_t - \lambda_L}{\lambda_H - \lambda_t})^\theta (\lambda_H - \lambda_t) + (\lambda_t - \lambda_L)} \tag{10}$$

where $\lambda_t^T = E_{t-T}[\tilde{\lambda}_t]$ *is the expected value of* $\tilde{\lambda}_t$ *under the Bayesian expectation.*

In Figure 3, we plot the evolution dynamics of the Bayesian and diagnostic belief processes, where the diagnostic belief process is described by (10). We note that when $\theta = 0$, $\lambda_t^\theta = \lambda_t$ so that the diagnostic belief is the same as the Bayesian belief. When θ is above 0, the pre-illiquidity shock belief is lower than the Bayesian belief, and then jumps to a higher level after an illiquidity shock. Right after an illiquidity shock, there is over pessimism. However, after one year, the perceived frequency of the illiquidity shock is below the Bayesian belief so that diagnostic agents are overly optimistic.

Under the diagnostic belief, we assume that all agents are unaware of their belief bias (i.e., they think λ_t^θ as if it is λ_t) and apply rational decision rules.[8] As a result, although we need to keep track of both λ^θ and λ for simulating the model dynamics, we only need λ^θ for a "snapshot" of the economy. For this reason, in what follows, we only discuss the model solutions under the Bayesian belief. The diagnostic model easily follows through by replacing λ with λ^θ in the policy functions.

---

[8]One could consider a model where a single diagnostic agent thinks they are unbiased, but other agents are biased. This is in the spirit of the heterogenous belief models (Simsek, 2013). We conjecture that in such a model the equilibrium impact of belief distortions will be weakened (e.g., bankers will take less leverage if they observe that other bankers are over-levered).
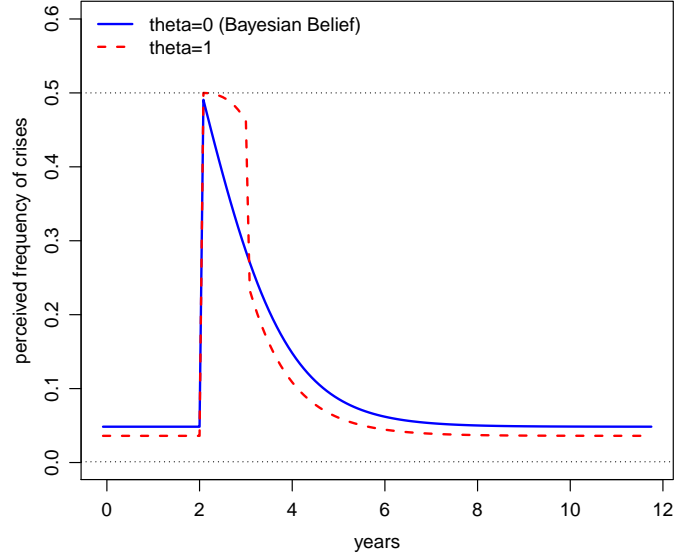
Figure 3: Simulation of Diagnostic Belief under Different Diagnostic Parameter θ. The parameter $\theta \geq 0$ means the strength of the behavioral feature of the diagnostic belief, and it becomes the Bayesian belief when $\theta = 0$. Other parameters are set as $\lambda_L = 0.001$, $\lambda_H = 0.5$, $\lambda_{H \to L} = 0.5$, $\lambda_{L \to H} = 0.1$. These parameters imply that a financial illiquidity shock happens once about each 12 years. The diagnostic belief process is fully described by (10). In the figure, as θ increases from 0 to 1, the believed frequency of illiquidity shocks in a pre-crisis boom decreases by 30%.

3.5   State Variables and Decisions

We define the total wealth of banks as W_t^b and the total wealth of households as W_t^h. Then we have three state variables. One is the wealth share of bankers, denoted by

$$w_t = \frac{W_t^b}{W_t^b + W_t^h}, \tag{11}$$

The second is the expected jump intensity λ_t. The final one is the total productive capital K_t. We construct an equilibrium where all relevant object scale linearly with K_t. This reduces the computational problem to solving a model with two state variables, w_t and λ_t.

Denote w_t^b as the wealth of a representative banker. Similarly, denote w_t^h as the wealth of a representative household. Let the associated value function be $V^b(w_t^b, w_t, \lambda_t)$ and $V^h(w_t^h, w_t, \lambda_t)$, respectively, at time t. To guarantee a non-degenerate wealth distribution, we assume bankers randomly transit to becoming households at rate η.[9] Bankers take this transition possibility into account in their optimization problems.

---

[9]Without this assumption, the banker, who earns a higher return on capital, will come to own almost all of the wealth of the economy.

Bankers

Each banker can invest in productive capital and borrow/lend from households or other banks via short-term debt at interest rate r_t^d. Note that short-term debt is riskless even though the price of capital will jump in equilibrium. This is because a forward-looking banker with log utility will never make a portfolio choice that leaves him with negative wealth in any state.

Denote the banker's portfolio choice (as a fraction of the banker's wealth w_t^b) in productive capital as x_t^K. Then the borrowing from household is

$$x_t^d = x_t^K - 1 \tag{12}$$

We will later show that banks always borrow from households and take leverage so we always have $x_t^d > 0$.

Starting from time t, the time that banker will switch to becoming a household is denoted as T, which is exponentially distributed with rate η. A banker with wealth w_t^b solves the problem

$$V^b(w_t^b, w_t, \lambda_t) = \sup_{c_t^b \geq 0, \, x_{t-}^K \geq 0} E\Big[\int_t^T e^{-\rho(s-t)} \log(c_s^b) ds + e^{-\rho T} V^h(w_T^b, w_T) \,\big|\, w_t^b, w_t\Big], \tag{13}$$

subject to the solvency constraint

$$w_t^b \geq 0. \tag{14}$$

The second part of the objective function is the transition to a household, which changes the continuation value from V^b to V^h. The dynamic bank budget constraint is:

$$\frac{dw_t^b}{w_{t-}^b} = \underbrace{x_{t-}^K(\mu_{t-}^R + \frac{\bar{A}}{p_{t-}})dt}_{\text{return from capital}} - \underbrace{x_{t-}^d r_{t-}^d dt}_{\text{deposit funding}} - \underbrace{\frac{c_{t-}^b}{w_{t-}}dt}_{\text{consumption}} + \underbrace{x_{t-}^K(\sigma^K + \sigma_{t-}^p)dB_t}_{\text{volatility of capital}} - \underbrace{(x_{t-}^K \kappa_{t-}^p + \alpha(x_{t-}^d)^+)dN_t}_{\text{losses in illiquidity shock}},$$

where $(x^d)^+ = \max\{x^d, 0\}$ measures the net borrowing from households, and the "non-dividend" component of capital return is:

$$\mu_t^R = \underbrace{\mu_t^p}_{\text{price appreciation}} - \underbrace{\delta}_{\text{depreciation}} + \underbrace{\sigma^K \sigma_t^p}_{\text{Ito term}} + \underbrace{\mu_t^K - \frac{\phi(\mu_t^K)}{p_t}}_{\text{net investment return}}, \tag{15}$$

In equation (15), the bank obtains returns from capital investment and pays the funding costs to depositors and dividends (i.e., banker consumption) to bank shareholders, subject to the Brownian risks of capital volatility, and losses caused by the liquidity shocks. The return from capital can be classified into a dividend component denoted by $\bar{A}/p_t$, and a

non-dividend component denoted by μ_t^R, which as shown in equation (15) consists of capital price appreciation, capital depreciation, the Ito term on capital volume and price volatility, and finally the net investment returns. During a liquidity shock, the banker suffers from both an exogenous liquidation cost (α) and a valuation drop on their capital holdings (κ_{t-}^p). Note that the net funding withdrawal that has to be fulfilled during an illiquidity episode by selling productive capital is x_{t-}^d.

Households

Each household chooses the consumption rate c_t^h and capital holding y_t^K as a fraction of household wealth for the following objective

$$V^h(w_t^h, w_t, \lambda_t) = \sup_{c_t^h \geq 0, \, y_t^K \geq 0} E[\int_t^\infty e^{-\rho(s-t)} \ln(c_s^h) ds \, | w_t^h, w_t], \tag{16}$$

subject to the solvency constraint

$$w_t^h \geq 0, \tag{17}$$

and the budget constraint

$$\frac{dw_t^h}{w_{t-}^h} = \underbrace{y_{t-}^K(\mu_{t-}^R + \frac{A}{p_{t-}})dt}_{\text{return from capital}} + \underbrace{y_{t-}^d r_{t-}^d \, dt}_{\text{deposit interest}} - \underbrace{\frac{c_{t-}^h}{w_{t-}}dt}_{\text{consumption}} + \underbrace{y_{t-}^K(\sigma^K + \sigma_{t-}^p)dB_t}_{\text{volatility of capital}} - \underbrace{\kappa_{t-}^h dN_t}_{\text{liquidity exposure}} \tag{18}$$

where in the liquidity shock, they also suffer losses on their holdings of capital, but receive a transfer (the exogenous liquidation cost paid by the banker):

$$\kappa_{t-}^h = \underbrace{y_{t-}^K \kappa_{t-}^p}_{\text{valuation drop}} - \underbrace{\alpha(x_{t-}^d)^+ \frac{w_{t-}}{1 - w_{t-}}}_{\text{transfer}}. \tag{19}$$

Relative to the bank budget constraint in (15), the household budget constraint (18) differs mainly in two ways: First, households earn a lower dividend return compared to bankers, $\underline{A}/p_t < \bar{A}/p_t$; Second, during the liquidity shock, households provide emergency funding to banks and earn a profit, while bankers lose net worth due to the financing costs. In practice such a profit is likely intermediated by the central bank, which we have omitted in our modeling. Our modeling implies that there is no destruction of wealth in a liquidation shock, so the household financing benefits (per unit of wealth) multiplied by household total wealth $1 - w_t$ equals to the banker financing costs (per unit of wealth) multiplied by banker total wealth w_t. We could alternatively model the liquidation cost as a deadweight loss. t not to, primarily to ensure that the liquidity shock is purely financial and has no direct impact on aggregate output.

3.6 Equilibrium Definition

Denote the share of capital owned by bankers as

$$\psi_t = \frac{x_t^K W_t^b}{x_t^K W_t^b + y_t^K W_t^h}. \tag{20}$$

Then the aggregate production of consumption goods is

$$Y_t = (\psi_t \bar{A} + (1 - \psi_t)\underline{A})K_t. \tag{21}$$

Because $\bar{A} > \underline{A}$, output is increasing in ψ_t.

Given that there is no heterogeneity within bankers and within households, we can express the dynamics of aggregate wealth as

$$\frac{dW_t^b}{W_{t-}^b} = \frac{dw_t^b}{w_{t-}^b} - \eta dt \tag{22}$$

$$\frac{dW_t^h}{W_{t-}^h} = \frac{dw_t^h}{w_{t-}^h} + \eta \frac{W_{t-}^b}{W_{t-}^h}dt, \tag{23}$$

where the second terms in both (22) and (23) are due to the transition of bankers to households.

We derive a Markov equilibrium where all choices only depend on the state variables w_t and λ_t.[10] Let $\hat{c}^b = c^b/w^b$ be the consumption of a representative banker as a fraction of the banker's wealth, and $\hat{c}^h = c^h/w^h$ similarly. The following formalizes the equilibrium definition.

Definition 1 (Equilibrium). *An equilibrium is a set of functions, including the price of capital $p(w_t, \lambda_t)$, bank debt yield $r^d(w_t, \lambda_t)$, household consumption wealth ratio $\hat{c}^h(w_t, \lambda_t)$ and capital holdings $y^K(w_t, \lambda_t)$, banker consumption wealth ratio $\hat{c}^b(w_t, \lambda_t)$ and capital holdings $x^K(w_t, \lambda_t)$, such that*

- *Consumption, investment and portfolio choices are optimal.*

- *Capital good market clears*

$$W_t^b x_t^K + W_t^h y_t^K = p_t K_t. \tag{24}$$

---

[10] Under diagnostic beliefs, while agents' beliefs are diagnostic they think that theirs and all other agents' beliefs are Bayesian. In other words, the policy functions and state variables are the same as those under Bayesian beliefs. However, because these policy functions are evaluated under diagnostic beliefs, the equilibrium outcomes are different. Furthermore, the dynamics of the state variables are different due to the underlying difference between diagnostic belief and the true process. The solution strategy for the diagnostic belief model is to solve the Bayesian decision rules under Bayesian belief $\theta = 0$, and then apply the same policy functions and simulate the diagnostic model with the diagnostic belief of $\theta > 0$.
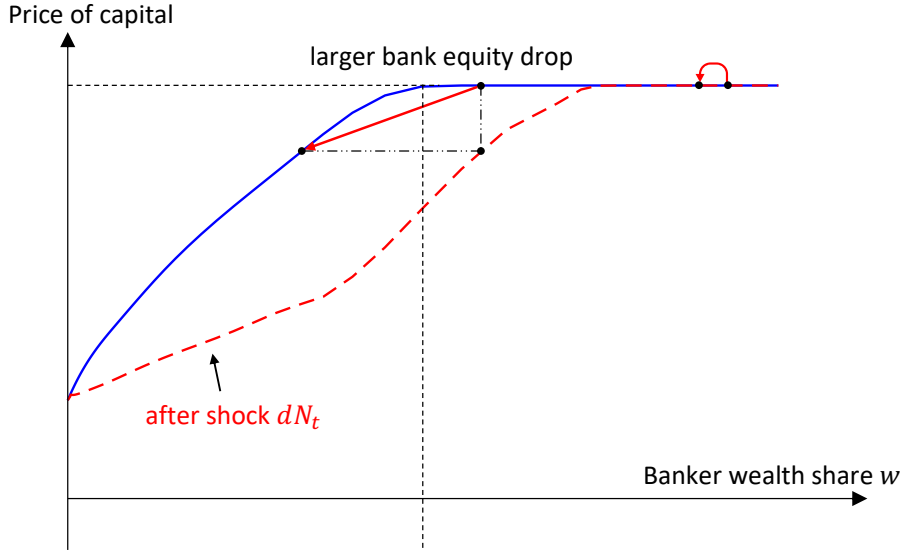
Figure 4: Price of capital as a function of w_t, pre- and post- dN_t shock.

- *The aggregate non-financial wealth of households and banks equal to total value of capital*

$$W_t^b + W_t^h = p_t K_t. \tag{25}$$

- *Consumption goods market clears*

$$\hat{c}_t^b W_t^b + \hat{c}_t^h W_t^h = (\psi_t \bar{A} + (1 - \psi_t)\underline{A})K_t - i_t K_t. \tag{26}$$

3.7   State-Dependence and Distress Dynamics

We solve the model and illustrate the nonlinear and state-dependent effects of a financial illiquidity event and the dynamics of the capital price around illiquidity shocks.

Figure 4, Panel (a) graphs the price of capital in blue as a function the banker's wealth share, w_t, which is one of the state variables in the equilibrium (λ_t is the other state variable). We note that the price of capital is increasing in w_t up to a point and then is flat thereafter. In the increasing portion, both bankers and households own capital. As the wealth share increases, more of the capital is in the bankers' hands, and hence more of the capital produces a higher dividend of $\bar{A}$. This force leads to a positive relationship between the price of capital and the wealth share. To the right of the dashed line, all of the capital is in the bankers' hands. Now, it will be the case that as the wealth share of bankers rises to the right of the dashed line, the risk premium required by bankers to absorb capital risk falls, which by itself would raise capital prices. However, because of log utility, the interest rate rises to offset the fall in the risk premium, and the net effect on the discount rate is to keep the price of capital constant to the right of the dashed-line.

There are two cases of interest. If the illiquidity shock occurs when banker wealth share is high – on the right side of the dashed line in panel (a) – bankers suffer the exogenous liquidation loss, which means that the post-shock wealth share jumps to the left, as indicated by the red arrow. But since at this new wealth share, the price of capital is the same as at the old wealth share, there is no endogenous fall in the price of capital. On the other hand, on the left side of the dashed line, the exogenous loss leads to a fall in banker wealth share, which leads to an endogenous fall in the price of capital, which implies further losses to bankers, and so on. The post-shock capital price traces along the red dashed line, reflecting a downward jump in the capital price and the banker wealth share state variable. The exogenous loss is amplified in this case. Our model thereby captures an amplification mechanism, where the degree is state-dependent.

3.8 Leverage, Risk and Liquidity Premia

For an individual bank, the net funding withdrawal that has to be fulfilled during an illiquidity episode by selling productive capital is $(x_t^d)^+ = (x_t^K - 1)^+$. In Appendix A.3, we prove that:

Lemma 3. *In equilibrium, banks always borrow from households and take leverage, i.e.,*

$$x_t^K \geq 1.$$

The statement is true because banks earn higher returns on holding productive capital than households. Thus, we have

$$(x_t^d)^+ = x_t^d \tag{27}$$

With the results in Lemma 3 and the properties of log utility, we write the equivalent banker's optimization problem as:

$$\max_{c_t^b, x_t^d, x_t^K} \left\{ \log(c_t^b) + \frac{1}{\rho} \left(E_{t-}[\frac{dw_t^b}{w_{t-}^b}]/dt - \frac{1}{2} \left(\frac{dw_t^b}{w_{t-}^b} \right)^2 /dt \right) \right\} \tag{28}$$

subject to the bank budget constraint, (15), rewritten as,

$$\frac{dw_t^b}{w_{t-}^b} = -\frac{c_{t-}^b}{w_{t-}^b}dt + r_{t-}^d dt + x_{t-}^K \cdot \underbrace{(\mu_{t-}^R + \frac{\bar{A}}{p_{t-}} - r_{t-}^d)}_{\text{capital excess return}} dt + \underbrace{x_{t-}^K(\sigma^K + \sigma_{t-}^p)dB_t}_{\text{Brownian risks}} - \underbrace{(\alpha x_{t-}^d + x_{t-}^K \kappa_{t-}^p)dN_t}_{\text{losses in illiquidity shocks}} \tag{29}$$

The objective has the familiar mean-variance form over the evolution of wealth that comes from log-utility. We note the key quantities that enter into this mean-variance tradeoff: (1) purchasing capital funded by deposits earns the capital excess return; (2) but this return is at the cost of capital price risk (Brownian risk) and the costs in a bank run (losses in an

illiquidity shock). Note that the net funding withdrawal that has to be fulfilled during an illiquidity episode by selling productive capital is x^d_{t-}.

Denote the return to a bank on holding capital as dR^b_t. Then we have the following first order condition for the excess return earned by the banker in purchasing capital funded by deposits:

$$E_{t-}[dR^b_t] - r^d_{t-} = \underbrace{(\sigma^K + \sigma^p_{t-})^2 x^K_{t-}}_{\text{Brownian risk premium}} + \underbrace{\lambda_{t-}(\alpha + \kappa^p_{t-})\frac{x^K_{t-}\kappa^p_{t-} + \alpha x^d_{t-}}{1 - x^K_{t-}\kappa^p_{t-} - \alpha x^d_{t-}}}_{\text{liquidity risk premium}} \qquad (30)$$

where the first term is the required compensation for taking on Brownian risk. The two sources of Brownian risk are the exogenous capital shock, σ^K, and the endogenous capital price risk, σ^p_t. The second term is a liquidity risk premium. Deposits are subject to run risk, in which case the bank has to sell capital, suffering the exogenous loss of α and the endogenous fire sale loss of κ^p_{t-}. This possible loss requires a compensation, which is the liquidity risk premium.[11]

Equation (30) can also be used to understand the leverage decision of a banker, which is x^K_t. In particular, consider how news that leads the banker to revise upwards his estimate of λ_t will affect leverage. Since purchasing capital funded by runnable deposits exposes the banker to liquidity risk, this higher liquidity risk will lead the banker to take on less leverage. Figure 9 illustrates this negative relationship in our model. A useful intuition to help understand our model's results is:

$$\text{Prob of crisis} \propto \underbrace{\text{Leverage}}_{\uparrow \text{ as } \lambda \downarrow} \times \underbrace{\lambda}_{\text{Prob of liquidity shock}} \qquad (31)$$

We return to this relation in Section 5.

3.9 Spreads and Bank Pricing of Liquidity and Credit

In this section, we define spreads that enable us to align the model with data. First, we define the spread on a hypothetical instantaneous loan with interest rate r^C_t and no capital price risk. While there is no price risk when making this loan, we assume it is subject to illiquidity costs of α in the event of a bank run. It is straightforward to show[12] that the

---

[11] Note that the expected loss under the physical probability is $\lambda_{t-}(\alpha + \kappa^p_{t-})$. The term for the liquidity risk premium in (30) reflects the risk compensation for being exposed to these losses.

[12] The derivation detail is in Appendix A.6.

spread on this loan relative to the deposit rate is:

$$r_{t-}^{C} - r_{t-}^{d} = \underbrace{\frac{\lambda_{t-}}{1 - x_{t-}^{K}\kappa_{t-}^{p} - \alpha x_{t-}^{d}}}_{\text{risk-adjusted probability}} \cdot \underbrace{\alpha}_{\text{liquidity loss}} \tag{32}$$

This object is a pure <u>liquidity</u> spread and reflects banks concern over liquidity risk. We use this spread to help calibrate the unconditional mean intensity of the liquidity shock.

Second, we aim to match the crisis-cycle pattern of bank's credit pricing, which reflect pre-crisis froth, a sharp tightening in the crisis, and a gradual post-crisis recovery. A natural model object that will reflect bank's credit pricing is $E_{t-}[dR_t^b] - r_{t-}^d$ which is banks' required return on holding capital (i.e., loans) over its funding cost. Loosing speaking, $E_{t-}[dR_t^b] - r_{t-}^d$ is a bank's required loan spread. However, the exact historical data we match over the crisis cycle is not loan spreads but credit spreads (see Section 2). There is considerable empirical support for the association between credit spreads and bank lending standards. See Gilchrist and Zakrajsek (2012). We next define a credit spread that is needed to map the model to the credit spread data.

We define a zero net-supply defaultable bond, matching the characteristics of the credit spreads in the data presented in Section 2. These defaultable bonds are priced by the banker's pricing kernel. This last point is worth stressing, as the model-defined credit spread will thus pick-up endogenous variation in bankers' attitude towards risky lending. We define the credit spread as the yield differential between a risky zero-coupon bond and a zero-coupon safe bond with the same [expected] maturity. We model the default intensity of the bond as related to the intensity of the illiquidity shock, λ_{t-}. In default, the losses to bond holders are affine in the capital price decline κ_{t-}^p. Details on this specification, the bond pricing solution, and the calibration are provided in Appendix A.7.

Figure 5 plots the credit spread in the calibrated model against the liquidity premium, $r_{t-}^C - r_{t-}^d$, and the loan spread $E_{t-}[dR_t^b] - r_{t-}^d$. The variation in the spreads is generated by model's variation in the state variable λ_t. The figure plots this relation for two different values of w, one at the median, and one at a higher value of w. The upshot from this figure is that all of these spreads move together. We use the liquidity premium and the credit spread as targets for calibration because they have measured data counterparts.

3.10   Solution Methodology and Simulation

The challenge of solving this model comes from both multiple state variables and the endogenous jumps in the state variables. To ensure stability, we use a functional iteration method that begins with an initial guess of the capital price function $p^{(0)}(w, \lambda)$, and then iterates over the equilibrium equation system to get an updated price $p^{(1)}$. This updating

(a) Liquidity Premium (Equation (32)) (b) Bank Loan Spread (Equation (30))
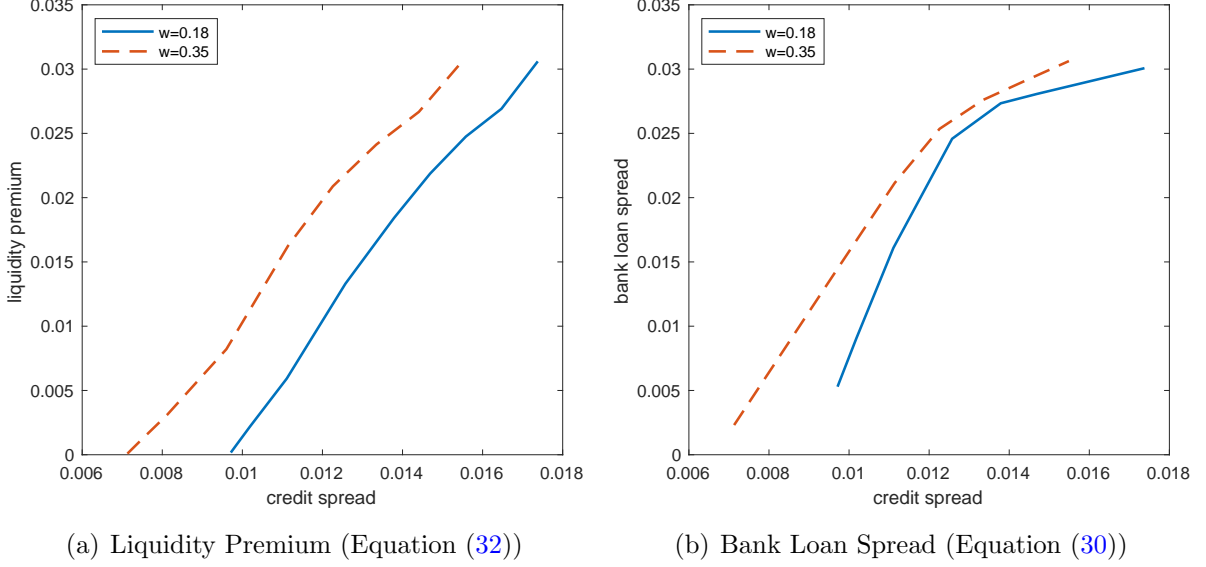
Figure 5: Credit Spread against Bank Loan Spread and Liquidity Premium in the calibrated Bayesian Model. In this figure, we show the relationship between credit spread, liquidity premium, and the bank loan spread. We fix the state w and then trace the relationship among different spreads along the λ dimension.

step involves solving a fixed-point problem at each state (w, λ). Then we iterate until at step k, we have

$$\int_0^1 \int_{\lambda_L}^{\lambda_H} |p^{(k+1)}(w, \lambda) - p^{(k)}(w, \lambda)| d\lambda dw < \varepsilon$$

for a small positive number ε.

We simulate the model at a monthly frequency but analyze simulations at a yearly frequency to be consistent with the data.

- We set the simulation interval as $dt = 1/12$ (a month), and generate the independent Brownian shocks $dB_t \sim \mathcal{N}(0, \sqrt{dt})$, as well as an independent frequency of illiquidity shock process $\tilde{\lambda}_t$. Based on the illiquidity shock process $\tilde{\lambda}_t$, we generate illiquidity shocks dN_t that hits with probability $\tilde{\lambda}_t dt$ for the time interval dt.

- Once shocks are generated, we solve for the dynamics of state variables, including w_t, λ_t, and K_t. For the static belief model, $\lambda_t = \bar{\lambda}$. For the diagnostic belief model, we need to generate λ_t^θ based on λ_t.

- With state variables determined, we generate all other quantities and prices of the model.

- We discard the first one thousand data points of each simulation path collected in this manner. As a result, the initial values do not affect our computed moments. The simulation approximates picking initial conditions from the ergodic distribution of the state variables.

24

- Finally, we average all of the monthly quantities for a given year to get an annual data set. For prices, we use the first observation of every year.

4 Calibration

In order to map model outputs to data, we need to define a financial crisis. Crises are the events where the growth in bank credit/GDP in a given month falls into the lowest 4% quantile of the distribution of monthly bank credit/GDP growth rates, and there has not been another such event in the previous three years. This latter criterion is to ensure that a longer crisis is still dated as a single crisis, as is done in the empirical literature. This crisis corresponds to a disintermediation event, and in the simulation almost always involves an illiquidity shock and bank run, although as crises are endogenous, not all illiquidity events are crises. We target the 4% number based on fact 1 of Section 2. We also consider a crisis definition based on bank equity crashes, as in Baron and Xiong (2017), in Section 5.8.

We solve and calibrate three variants of the model:

1. Bayesian (rational) Model: Agents form beliefs over the illiquidity state following Bayes rule, and this belief varies over time.

2. Diagnostic (non-rational) Model: Agents form beliefs over the illiquidity state via diagnostic expectations, and belief varies over time.

3. Static-belief Model: Agents' beliefs are constant.

We apply a combination of calibration and estimation for model parameters. Specifically, we directly set parameter values for those with standard values in the literature. Then we estimate the rest of parameters based on moments chosen to best reflect the economics of those parameters.

Table 1: Calibrated Parameter Values

Choice	Moment		
δ	Depreciation rate	10%	Depreciation rate in the literature
ρ	Time discount rate	4%	Discount rate in the literature
χ	Investment adjustment cost	3	Adjustment cost in literature
α	Distress illiquidity costs	0.05	Data
θ	Diagnostic belief weight	0.90	Literature

A list of the calibrated parameters for the model (not including the credit spread which is given in Appendix A.7) are shown in Table 1. We follow the macroeconomics literature to set annual depreciation rate $\delta = 0.1$ (Gertler and Kiyotaki, 2010), annual time discount rate $\rho = 4\%$ (Gertler and Kiyotaki, 2010), and investment adjustment cost $\chi = 3$ (He and Krishnamurthy, 2019). For the emergency liquidity costs (α), we do not have good data for the historical financial crises to pin these down. From data of the 2008 crisis, the effective liquidation loss is about 0.05, which is the value of $\alpha \cdot \beta$ in Li (2019). Alternatively, we can interpret this liquidation loss as a funding premium. The value of $\alpha = 0.05$ translates to a 10% premium for a illiquidity event that lasts 6 months. Last, in our investigation of beliefs in the model, we choose the diagnostic parameter θ based on the research by Bordalo, Gennaioli and Shleifer (2018), Bordalo et al. (2019b), and Bordalo et al. (2020). These authors estimate θ based on the dynamics of forecasts for financial and economic variables. We set θ equal to 0.9, which is the value used by Bordalo, Gennaioli and Shleifer (2018) and Bordalo et al. (2019b).[13]

Table 2: Moments and Model Estimates

Panel A. Moments

	Data	Static	Bayesian	Diagnostic
Average liquidity premium	0.90%	0.90%	0.93%	0.75%
Avg credit spread change in crises	70%	7%	51%	55%
Half-life of credit spread recovery (years)	2.5	2.7	2.6	2.3
Output/capital ratio	14%	16%	17%	15%
Avg 3-year output drop in crises	-9.1%	-7.9%	-8.8%	-9.4%
Output growth volatility	3.8%	3.5%	2.8%	3.6%
Average bank leverage	5.0	5.1	5.0	5.0

Panel B. Estimated Parameter Values

	Parameter	Static	Bayesian	Diagnostic
Avg frequency of liquidity shock	$\bar{\lambda}$	0.08	–	–
High intensity of liquidity shock	λ_H	–	0.51	0.58
Low to high transition	$\lambda_{L \to H}$	–	0.11	0.11
High to low transition	$\lambda_{H \to L}$	–	0.52	0.48
Household productivity	A_L	0.12	0.17	0.13
Bank lending advantage	$A_H - A_L$	0.055	0.03	0.02
Volatility of capital growth	σ^K	0.055	0.03	0.03
Banker-household transition rate	η	0.135	0.06	0.04

---

[13]Bordalo et al. (2020) report an estimate of θ of 0.5.

Then we proceed to estimate other parameters, reported in Table 2 Panel B. The Static Belief model has only one parameter $\bar{\lambda}$ governing the crisis frequency process which is constant over time. There are four parameters governing beliefs in the Bayesian model: λ_H, λ_L, $\lambda_{L \to H}$, and $\lambda_{H \to L}$. We note that as long as λ_L is close to zero, the impact of its value is negligible. Therefore, we pick $\lambda_L = 0.001$ directly. The diagnostic model adds θ as one more degree of freedom (the 'look-back period" parameter t_0 is set to 1, the implicit value from discrete-time diagnostic belief process such as Bordalo, Gennaioli and Shleifer (2018)). Finally, we note that the Bayesian Belief and Diagnostic Belief models are exactly identified, while the Static Belief model has two more moments than parameters. After experimentation with the model, we find that the following moments to be particularly informative for the belief parameters:

- The average liquidity premium will reflect banks' assessment of liquidity risk, and thus the average value of λ. See equation (32). The spread between P2 rated 3-month commercial paper and 3-month T-bills in data from 1974 to 2018 is 94 basis points. We target a liquidity premium of 90 basis points. Krishnamurthy and Vissing-Jorgensen (2015) estimate the average liquidity premium on long-term Treasury bonds relative to AAA corporate bonds to be 75 basis points. We focus on a short-term bond in our exercise and thus target a higher spread. Our estimate reported in Panel B implies an average λ of 0.08, which translates to a liquidity event once over 12.5 years. In the high illiquidity state, the λ_H for the Bayesian and Diagnostic models are a little over 0.5 implying a liquidity event roughly every two years.

- Credit spread changes during a crisis (fact 2). The spike in the credit spread is 0.7σs. This moment helps determine $\lambda_{L \to H}$, which affects the degree of surprise in beliefs due to the realizations of illiquidity shocks.

- Half-life of credit spread recovery (fact 4). According to Krishnamurthy and Muir (2020), the half-life is 2.5 years. This moment primarily determines $\lambda_{H \to L}$, since the speed of recovery of beliefs after a illiquidity shock is directly affected by the underlying transition probability.

The parameters $\bar{A}$, $\underline{A}$, σ^K, and η govern the output process both in and out of crises. The following targets inform these parameter choices:

- Average output decline during a crisis (fact 3): We target -9.1% as explained in Section 2. This moment is most directly related to the productivity differential $\bar{A} - \underline{A}$.

- Investment to capital ratio: We use the same target as He and Krishnamurthy (2019). This moment mainly affects the average of productivity parameters, $\bar{A}$ and $\underline{A}$.

- Average output growth volatility: According to Bohn's historical data, the volatility of real GDP growth from 1791 to 2012 for the U.S. is 4%. This moment mainly affects the capital volatility σ^K.

- We map banks in the model to depository institutions and broker dealers in the flow of funds. Bank equity is defined as total bank assets minus total bank liabilities. Since our model only captures runnable liabilities, we define effective bank liabilities as total liabilities minus insured deposits. Then we calculate bank leverage as (bank equity + effective bank liabilities)/bank equity. Using all data available, we find that bank leverage is approximately 5. This moment disciplines η, the transition rate from bankers to households, which affects the stationary distribution of leverage in the model. For example, setting η very low leads to a stationary distribution where almost all of the wealth is in bankers' hands and average leverage in equilibrium is very low.

To search for parameter values that best match moments, we need to repeatedly solve the model for a large combination of parameter values. A simple discretization of the parameter space (5 parameters for the benchmark, 7 parameters for the Bayesian and diagnostic models) renders the task computationally infeasible. To resolve this difficulty, we apply the Smolyak grid method (Judd et al., 2014) to generate a discretized state space. For each version of the model, we follow the estimation procedure:

- Discretize the state space of parameters around their initial values. We pick a discretization level of 3 in the Smolyak discretization. This results in 177 combinations for the static belief model, 241 combinations for the Bayesian model, and 389 combinations for the diagnostic model. Simulate all of these models and collect their moment values.

- Denote the moments in the data as $m_1, \cdots, m_J$, and the moments from the model as $\hat{m}_1, \cdots, \hat{m}_J$. From all of the parameter combinations, pick the one that minimizes the objective

$$\sum_{j=1}^{J} weight_j \frac{|\hat{m}_j - m_j|}{m_j}.$$

  Here $weight_j$ reflects the importance of a given target in the estimation. We set the weight for the liquidity premium to be three and the rest of the weights to be one. The average liquidity premium determines the frequency of illiquidity shocks which is a particularly important parameter in the model.

- Once we have picked a set of parameters, we search in a smaller region around this set of parameters and find a new best set of parameters in the smaller region. We iterate the above process until the difference between the optimized objective value between two iterations is below a threshold.

5 Model Evaluation

This section evaluates the models we consider and explains the mechanisms that help the models match the crisis data patterns.

5.1 Targeted Moments

We target means across crises in choosing parameters. Table 2 presents the model's fit in hitting the targets. We re-calibrate the model parameters to best match moments for each version of the model, thus giving each model the best chance to represent the data. Although each version of the model is at least exactly identified (static belief model is overidentified), because the state-space is restricted, we do not fit all of the moments accurately. The static belief model, in particular, misses the spread change in the crisis by a wide margin. It is possible to fit this moment if we increase the exogenous liquidation cost α, but we opt to keep α constant across all of the models to better illustrate the mechanisms underlying the models.

[TABLE 2 HERE]

Figure 7 plots the path of the model-generated credit spread, bank credit/GDP and GDP around a crisis at $t = 0$. The credit spread and bank credit variables are plotted in units of standard-deviations from their mean value over the sample. The figure should be compared to the data in Figure 1. We see that the model is able to generate the jump in spreads, contraction in credit, and drop in GDP. For both the Bayesian and Diagnostic model, the magnitudes of the spread spike and GDP decline are also in line with the data. During a crisis, spreads jump about 50% in the model (that is, 0.5 σs) and 70% in the data. As noted above, the magnitude of the spread spike in the static belief model is too small relative to the data. The magnitude of the credit contraction of around 0.55σs is larger than the data counterpart of 0.33σs. This is likely because in our model all credit is extended via banks, while in the data, there are other intermediaries involved in the credit process. Note that we have not explicitly targeted the credit contraction in the calibration.

[FIGURE 7 HERE]

All of the models match the sharp transition in the crisis, driven by the model's amplification mechanism, and output that is below trend for a sustained period post-crisis. The figures also reveal how the pre-crisis patterns vary across the models. In the years before the crisis, bank credit and GDP are rising while credit spreads are below normal in both dynamic belief models. In the static belief model, spreads are slightly higher than normal,

29

while credit is falling. This contrast points to the need for time-variation in beliefs to fit the data.

5.2 Ergodic Distributions

In Figure 6, we graph the ergodic distributions of the state (w_t, λ_t) for the three models. Underlying movements in w are driven by three forces: the exogenous diffusion shocks to capital shift wealth, creating paths from the center of the distribution to both right and left; paths that go to the left are pushed back to the middle because in low w states, risk premia are high and bankers expected wealth growth is high; the transition rate of bankers into households, η, result in a drift in w of $-\eta w$, which pushes all paths to the left. The result of these forces is a mean-reverting w process and the single-peaked distribution. In the diagnostic and Bayesian models, the realization of a jump leads to a larger adjustment in w relative to the static beliefs model, because agents belief shift from the low illiquidity to the high illiquidity state. As a result, more mass is shifted to low-w states. Broadly, all three of the models generate a similar left-skewed output distribution.

[FIGURE 6 HERE]

5.3 Non-targeted Moments

The success in matching the mean patterns of crises verifies that our model's mechanisms can speak to the data. However, as we have noted, our calibration explicitly targets the means. We next describe the model's fit in the **cross-section of crises**, which are non-targeted moments. Within the sample of crises, there are smaller and larger crises. The moments we report measure variation within these crises. We discuss the model's fit of these non-targeted moments in this section, and delve further into the fit in the next sections. Table 3 summarizes the model's performance in matching non-targeted moments.

Panel A reports that all the models' fit on the data patterns on the crisis and its aftermath. First, in the data, episodes where credit spreads increase more are followed by larger output contractions. The first row of Panel A reports these moments from the models and data. All the model get the signs right, but the static belief model comes closest to matching the data. Credit spreads are likely too informative in our model relative to the data, because in our model credit spreads only reflect aggregate downside risk, when in the data idiosyncratic risk likely also drives spreads. Second, crises that are preceded by a run up in bank credit are also more severe crises. The second row of Panel A reports the models' fit with data on this dimension. In general, all three of the models get the signs right and are in the ballpark of the data, but again the diagnostic and Bayesian models are over-sensitive.

We note that these crisis facts are well modeled even in the static belief model which only has a financial intermediation mechanism.

Panel B reports that all of the models are able to fit the negative data relationship between bank credit and risk premia. The panel is not explicitly about financial crises, but more generally about the relationship between movements in credit and risk premia. In the data, credit growth is negatively correlated with excess equity returns (Baron and Xiong, 2017). Periods of high credit growth are followed by low returns, and periods of low credit growth are followed by high returns. We verify that all of the models we consider deliver this relation. They do so via time variation in the supply of risk-bearing capacity. The state variables of the model, such as w, capture variation in the effective risk aversion of the banking sector. When effective risk aversion is low, banks lend more and credit grows, while risk premia are low; the opposite pattern holds when risk aversion is high. This mechanism thus delivers the relation between bank credit and risk premia. The fact that this relationship holds even in the model with static beliefs bears stressing: a sentiment/belief mechanism is not necessary to replicate the credit/risk premia relationship.

Panels C and D consider the pre-crisis patterns where we see divergence across the models. In Panel C, we examine whether the model can reproduce the fact that spreads are below normal before crises. The first row considers the mean pre-crisis spread. Both the diagnostic model and the Bayesian model deliver the below normal spread, while the static belief model delivers an above normal spread. We explain this failure in further detail below. Panel D considers the predictive relationship between measures of credit market excess and subsequent crises. We again see that the static belief model fails to generate a sign in keeping with the data. Both of the belief models succeed in this dimension, although each model does better (and worse) in some dimensions.

[TABLE 3 HERE]

5.4   Mechanism 1: Frictional Intermediation and Leverage

Figure 8 graphs the histogram of 3-year GDP growth in crises for all three models. In a model with no financial amplification and only diffusion shocks to AK_t, output growth would be normally distributed. All three models, and particularly the static model with only an amplification mechanism, generates the skew in line with the data. Thus, we conclude that the left-skewed output growth distribution in line with the data can be generated by a pure financial amplification mechanism.

[FIGURE 8 HERE]

In the data, the skewness in output growth matches the skewness of the jump in credit

spreads in the crisis (fact 7). Panel A in Table 3 evaluates the relationship between the jump in credit spreads in this model and the fall in GDP. The bottom row of Panel A evaluates the relation between the run-up in bank credit at the start of the crisis and the subsequent severity of the crisis. This is a relation reported by several empirical studies (Jordà, Schularick and Taylor, 2013). As noted all of the variants of the model get the signs right, while the credit variables in the diagnostic and Bayesian model are over-sensitive to crisis information. We have also seen in Panel B of Table 3 that all of the models generate the negative relation between bank credit and excess equity returns (risk premia). The models capture this moment because all of the models embed variation in the supply of risk-bearing capacity that drives both leverage and risk premia.

These observations indicate that the frictional intermediation mechanism, which is the only mechanism present in the static belief model, can capture the patterns of the economy in a crisis and its aftermath. Again, it is possible to improve the quantitative fit of the static belief model for the crisis and if its aftermath if we allow α to vary across models and be determined via the estimation. We choose not to go down this path because, as we explain next, this static belief model fails to fit the pre-crisis facts even qualitatively. It is likely also possible to reduce the over-sensitivity of spreads and credit in the belief models if we include an idiosyncratic component of risk in the determination of credit spreads.

5.5 Mechanism 2: Beliefs and Leverage

We report in Table 3 Panel C that the static belief model generates a spread that is higher than normal in the pre-crisis period, contrary to the data. The failure can be understood as follows. The amplification mechanism of the model, which is what drives the response of the economy to the illiquidity shock, is governed by the single state-variable w. If w is low (and leverage is high), a negative shock triggers a large fall in GDP and a crisis. However, since the credit spread is forward-looking, variation in the spread is also driven by w. The economy is more vulnerable when w is low, and hence credit spreads are higher when w is low. As a result, the static belief model generates an above normal spread before a crisis, contrary to the data.

The belief models are able to generate a spread with the right sign of the data.[14] To understand the economics here, consider Figure 9. We graph the policy function of bankers, for both Bayesian and diagnostic models, in choosing leverage as a function of the true state λ (denoted "rational" in the figure). Bankers in our model lever up to gain high returns on capital, but at the cost of the illiquidity event where they suffer bankruptcy costs from liquidating capital. Thus there is a bankruptcy risk/return tradeoff that drives their

---

[14] We report the results of a regression of spreads on a dummy that takes the value of one for the 5 years before a crisis. This regression also includes a control for the 5 years after the crisis so that the pre-crisis dummy indicates the level of spreads relative to non-crises periods.

leverage decision. When λ is low, the illiquidity event is less likely, and the banker chooses high leverage; hence, the negative slope in the curves in the figure. A useful relation to keep in mind is:

$$\text{Severity of crisis} \propto \text{Leverage} \quad and, \quad (\text{Leverage} \uparrow, \text{Spreads} \downarrow) \text{ as } \lambda \downarrow$$

When λ is low and leverage is high, if an illiquidity shock dN_t occurs, then its impact on GDP will be severe and more likely to result in the large GDP decline of a crisis. Finally, when λ is low, spreads are low, as is evident from Figure 5. This endogenous relationship between illiquidity risk and vulnerability generates the low credit spread before crises.

The diagnostic model with the calibration of $\theta = 0.9$ generates a magnitude in line with the data. The Bayesian model gets a spread that is below normal, but not as low as the data fact that spreads are about 0.34σs lower than normal in the pre-crisis period. The diagnostic model strengthens the belief mechanism and helps bring the model closer to matching the pre-crisis froth patterns. Consider the red dashed curve in Figure 9. We plot the banker's leverage decision as a function of the true lambda – not the agent's perceived diagnostic lambda. Clearly, at lambda of zero, the true and diagnostic lambda are the same. But as lambda becomes larger than zero, the diagnostic agent chooses higher leverage than the Bayesian agent. This is because the banker is overoptimistic and thinks lambda is lower than it actually is. When the true lambda is larger than a threshold, the banker is on average over-pessimistic and thinks lambda is higher than it actually is, thus choosing lower leverage. As a result, the leverage/lambda curve steepens under the diagnostic model and the diagnostic model better fits the spread/output relation as reflected in Panel C of Table 3.

[FIGURE 9 HERE]

This analysis indicates a "recipe": to strengthen the pre-crisis relationship, a model needs to steepen the leverage/lambda curve, even beyond that of the curves in Figure 9. Increasing the belief distortion helps in this regard. But it worth stressing that other specifications of the banker's problem – altering the corporate financing frictions, for example – can likely also deliver this steepening.

5.6 Pre-crisis: Predicting a Crisis with High Bank Credit

Next, we consider the evidence that high bank credit predict crises and not just precede crises. To see the difference, note that the former conditions on the event of a crisis. Table 3 Panel D, second row, presents the crisis prediction result. For the bank credit regression, we aim to match the result in Schularick and Taylor (2012) that a one-sigma increase in bank credit/GDP increases the probability of a crisis over the next year by 2.8%.

In the dynamic belief models, we find the variables have the right signs, although the models are somewhat high in terms of magnitudes for spreads, and low for the credit quantity variable. The static belief model fails again, generating a sign that is the opposite of the data. See Table 3 Panel D.

To understand what drives the mechanism in the dynamic belief models, consider the following intuition:

$$\text{Prob of crisis} \propto \underbrace{\text{Leverage}}_{\uparrow \text{ as } \lambda \downarrow} \times \underbrace{\lambda}_{\text{Prob of liquidity shock}} \tag{33}$$

There are two competing forces at work. As λ falls, endogenous leverage rises, but the probability of the illiquidity shock falls. If the leverage force is stronger, as it is in both versions of the calibrated belief models, we match the data relationship between high leverage and higher probabilities of a crisis.

Figure 10 illustrates this further. We plot the density of GDP growth over the next year conditional on the level of credit/GDP today. The red lines correspond to the Bayesian model and the dashed-blue lines correspond to the static-belief case. In panel (a) of the figure, we condition on low bank credit/GDP, which is typically the outcome when w is low and/or λ is high. This is a case where the banker faces higher illiquidity risk and endogenously chooses lower leverage. As a result, the economy is faced with moderate volatility of GDP but this volatility is confined to the center of the distribution and there is little mass at the left tail. Next, consider panel (b) where condition on high levels of credit and hence lower effective banker illiquidity aversion. The dotted black vertical line on the figure indicates the cutoff we have used to define a financial crisis. Mass is now pushed from the center of the distribution towards the left-tail crisis states. Effectively, the more risk-tolerant banker is willing to take on more liquidity risk when making decisions. There is less risk at the center of the distribution, but more mass in the tail. As a result, high credit states forecast more left-tail events.

The static belief model has only w as the state variable to drive effective risk aversion. With only this state variable driving decisions, the banker chooses leverage in a manner that crises are avoided when w and credit are higher. As shown in Panel C and D of Table 3, the signs on the credit-crisis relationship are the opposite of that in the data. This result reinforces a lesson of our analysis that we do need a model with two state variables to explain the entire crisis cycle.

[FIGURE 10 HERE]

Panels (c) and (d) of Figure 10 plots the distribution of GDP growth over the next year conditional on different levels of credit in the diagnostic model relative to the Bayesian

34

model. We plot the diagnostic's model distribution in green dashed lines and the Bayesian model in red. We can see that the forces that work to generate the relation between high credit and crises are similar but stronger in the diagnostic model compared to the Bayesian model. As we go from left to right panel in the figure, the mass in the left tail rises. The improvement of the diagnostic model is again due to steepening of the leverage/lambda relationship.

Figure 11 examines the predictive relation in a different way. In the figure, we plot the banker's wealth return conditional on high and low values of credit. Recall that our banker has log utility, so the mean and variance of this distribution are the key statistics driving banker utility and the leverage decision. The banker's wealth volatility is highest in the low credit case (left panel) driven by a significant mass spread between -0.1 and 0.4 at the center of the distribution. Distress and bankruptcy costs are salient to the banker, and thus he chooses low leverage. In the right panel high credit case, the output distribution is tight so that over most of the distribution, there is little distress for the banker. While there is a tail of wealth losses in crisis states, the banker's decision to take high leverage is largely driven by the tight central peak of the distribution. The banker understands that the typical negative shock will have small effects on his wealth, and is willing to gamble on avoiding the large tail shock. Note also that the banker's wealth process is different from the economy-wide GDP process, as should be expected in a model where banks drive systemic risk. Banker wealth is more sensitive than GDP to small shocks, and since such shocks are more likely, they are the drivers of the banker's leverage decision. As a result, the model produces the result that in the Bayesian model, even if illiquidity events are less likely (low λ), crises are more likely.

[FIGURE 11 HERE]

5.7 Pre-crisis: Predicting a Crisis with Low Credit Spread

We next turn to the relation driving froth (low credit spreads) and crises as reflected in the first row Table 3 Panel D. To replicate the spread predictability regressions in Krishnamurthy and Muir (2020), we define "high froth" as a dummy that indicates whether the credit spread is below its median value at time t. In Krishnamurthy and Muir (2020), the froth definition is based on credit spreads being below median over a 5 year period, which is necessary because a crisis in the data is not sharp 0-1 phenomena as in our model (spreads typically rise before the historian-dated crisis). We predict the likelihood of a crisis over the next 5 years in the model, in line with the data moment.

As we will explain, the froth relation holds for the belief models in the parameterization we study, but need not hold generally. Figure 12 draws density plots of next-year GDP

growth for the diagnostic, Bayesian and static belief model conditional on different levels of the credit spread. We can see that the static belief model gets the sign of the mass shift wrong. The diagnostic and Bayesian models, on the other hand, succeed in this dimension. We see again that the relative to the Bayesian model, the diagnostic model shifts more mass to the left tail when spreads are low, and leverage is endogenously high. We also see that the larger shift of the diagnostic model brings the coefficient more in line with the data, albeit still too small. See Table 3 Panel D.

[FIGURE 12 HERE]

The logic behind froth is more nuanced than for the high credit relation of the last section. There are two forces driving variation in the credit spread that are salient for understanding the mechanisms: (i) lower λ means less illiquidity events and hence lower spreads; (ii) worse crises mean higher loss-given default (via κ_t^p) and hence higher spreads. If we imagine shutting down effect (ii), then we can understand the froth relation easily from equation (33). Now, if we add back effect (ii), the froth relation is weakened. The reason is that more crises imply larger losses given default and hence higher ex-ante spreads. If we consider the extreme case where $\kappa = 0$, and hence recovery has no fixed component, the regression coefficient on froth falls to 0.96, thus reduced further relative to the data counterpart. The sign of the froth relation depends quantitatively on the exact cyclicality of recoveries in default and thus the relation between λ and spreads. We have calibrated our model to the history of recoveries on BAA bonds in the U.S., as reported by Moodys.

5.8 Bank equity crises

Baron and Xiong (2017) and Baron, Verner and Xiong (2021) define financial crises in terms of large (<30%) declines in bank equity values. They note that many of the crisis patterns documented in the narrative crisis dating literature (e.g., Laeven and Valencia (2013), Jordà, Schularick and Taylor (2011)) hold for this quantitative definition of financial crisis. In this section we define an equity-crash crash as an event where the return on bank equity in a given quarter is below $-X\%$, where X is chosen to yield a frequency of bank equity crashes of 4%. Because crashes can cluster in our simulation, we define the crash-crisis as the first crash that occurs after at least 3 years of no crash-crises. Thus we are effectively defining a crash as a single crisis. In our simulations, $X = 42\%$ for the Bayesian model and 41% in the diagnostic model.

Table 4 Panel A reports the declines in GDP in the 3 years subsequent to the crash-crisis. Baron, Verner and Xiong (2021) report that a crash-crisis is followed by a GDP decline of around 4.5%. Our numbers are larger than theirs. They also consider a definition of a crisis which involves a crash and a banking panic. In this event, they show the GDP declines are

about 6%. This latter definition is more in line with our model, as a crash almost always occurs with a liquidity shock. We also report the interaction regression, describing how bank credit pre-crisis worsens GDP outcomes in an equity crash. Analogous to our earlier results, bank credit is a vulnerability indicator for GDP declines in a crisis. Note that this is a regression we do in the model, but is not presented in Baron, Verner and Xiong (2021).

Table 4 Panel B presents predictive regressions, analogous to Table III of Baron and Xiong (2017), of bank credit/GDP on the likelihood of an equity crash. Baron and Xiong (2017) report that the the marginal probability of an equity crash rises by around 5.4% (column 7, top row of their table) in response to a one-sigma increase in bank credit/GDP growth. We run this regression in our simulated data and evaluate the change in the probability of an equity crash for a one-sigma increase in bank credit/GDP, evaluated at the mean value of bank credit/GDP. Our model regressions are considerably lower than the data regression. A part of this is that in Baron and Xiong (2017) the one-year probability of a crash using the -30% cutoff is 8% rather than 4%. However, this factor accounts for only half of the discrepancy. The main reason for the discrepancy is that in the model, equity crashes can occur when w is low and bank credit is low – essentially a situation where much of the capital is held outside the banking sector. In this case, the model can produce an equity crash from a state of low bank credit/GDP. When using contractions in bank credit/GDP to define the crisis, as in our the main crisis definition we have used, this situation does not arise.

[TABLE 4 HERE]

Finally, note that we have not reported these regressions for the static belief model. That model implies the wrong sign relative to the data. We can see this in Figure 13, which is a plot of the return on bank equity from month t to $t+1$ if an illiquidity shock occurs against the time t value of bank credit/GDP. Over the entire range, we see that the relation is positive rather than negative .

6   Policy Impact under Diagnostic and Bayesian Beliefs

We have shown that a financial friction mechanism plus a belief mechanism can capture the main features of the crisis cycle. We also learn that both the Bayesian and diagnostic belief models work, with the diagnostic model generating over-sensitivity relative to the data for the crisis and post-crisis patterns, while the Bayesian model is under-sensitive in its fit of the pre-crisis froth. Our take on these results, generated from a fairly parsimonious model, is that nailing down the source of belief fluctuations is not essential to a researcher aiming to match the main patterns of the crisis cycle.

We next turn to evaluating our model's implications for a policy experiment. We ask, does it matter for policy purposes whether we live in a world with diagnostic beliefs or one with Bayesian beliefs? We run the following experiment: we start the economy in a boom state (time $t = 0$) with low credit spread and high leverage, and then consider an unexpected policy that transfers wealth from households to bankers so that the banker wealth share w_0 increases by 10%. The idea is to recapitalize banks to "lean against the wind" so that the severity of a crisis is reduced if an illiquidity shock dN_t hits. This exercise is similar to Gertler, Kiyotaki and Prestipino (2020). Under each version of the model, we calculate the "derivative" of the nonlinear impulse response of quantities and asset prices with respect to the recapitalization policy. We focus on aggregate variables to assess the impact of policies.[15,16]

In our first experiment, we require that both models are calibrated to common data and we simulate the models from an initial condition in terms of observables that is the same across both models. In particular, we study a recap policy conditional on given initial bank leverage and credit spread, which are both observables in the data. These observables pin down the underlying states (w_0, λ_0) in the Bayesian model. In the diagnostic model, we also need to know the reference belief λ_0^T at $t = -T$. Since, on average, the diagnostic belief is equal to the rational belief, to reflect the average scenario, we assume $\lambda_0^T = \lambda_0$. We simulate the model at interval $dt = 1/12$ (one month), and introduce $dN_t = 1$ at the first month (but zero otherwise). To reflect the dynamics of the other shocks, we randomly generate dB_t in the simulation and simulate each model 10,000 times. For each model, we compute the average impulse responses across simulation runs with and without the recap policy and plot the difference between these responses in Figure 14. That is, we are plotting the difference of the impulse response to the recap, across both types of models. In the top-left panel, we plot the path-difference in w. At $t = 0$, due to the recapitalization policy, the response is +10% in both cases as expected. At $t = dt$ (monthly simulation so that $dt = 1/12$), the illiquidity shock dN_t hits and the nonlinear amplification mechanism turns on so that the response becomes larger than the initial 10% difference. The output recoveries (top right panel) after the illiquidity-shock are similar across Bayesian and diagnostic models. Since we start the economy in a boom state that features high bank credit, the additional 10% of bank equity has little impact on output initially. Upon the illiquidity shock, in both models, the output is higher in the recap relative to no-recap, by around 1.5% over the next two years. The bottom left panel plots the credit spread response. The recap leads

---

[15]In models with distorted beliefs, there is an open question of what should be the appropriate welfare criterion. By focusing on aggregates, we are implicitly adopted a paternalistic criterion. Dávila and Walther (2020) studies optimal policy in a credit market setting with heterogenous agents with distorted beliefs. Under a paternalistic criterion, he shows that leverage restrictions depend on the source of the belief distortion.

[16]The literature has observed that in models with financial frictions, there is a pecuniary externality that can motivate restrictions on bank leverage during booms (Bianchi, 2011). This motive carries over to our analysis.

to a smaller rise in the credit spread. Finally, the bottom right panel illustrates the bank credit response. The initial response-gaps are close to zero for both Bayesian and diagnostic models, subsequently rising to about 20%.

The key message from Figure 14 is that the derivative of the impulse responses to a recapitalization policy in the Bayesian and diagnostic models are quantitatively similar. The result arises because the initial state is the same in terms of observables, the models are calibrated to common data, and the initial condition for the diagnostic model is near the Bayesian model (i.e. neither over-optimistic or pessimistic).

In our second experiment, we probe whether knowledge of the exact divergence from Bayesian beliefs in the diagnostic model matters for outcomes. The question is whether "getting into the minds of agents" is important for the impulse responses. The answer is not obvious because we are simulating both models based on the same observable initial conditions, e.g., credit spreads, which reflect agents' beliefs. We calculate the diagnostic model's impulse responses in a case of overoptimism, where the initial state has the same bank leverage and credit spread, but with $\lambda_0^\theta < \lambda_0$. Then we compare the results with the no-belief distortion case of the diagnostic model in Figure 15. We again find that the impulse response gaps are quite similar across these two cases. We thus learn that the key element to the similarity-result is that the initial condition in terms of observables is the same. The dynamics of the economy, conditional on a state defined by the same observables, although different underlying state variables, are quite similar.

Finally, note that the plots in Figure 15 are conditional on an illiquidity shock. It is also interesting to examine the unconditional response. In the overoptimism case the true expected path of the economy will differ from the agent's beliefs over this path. Figure 16 plots these average impulse response gaps. Now we see that the recap policy has a more beneficial effect in the overoptimism case. However, the y-axis scale in these plots is far smaller in magnitude than in Figure 15. That is, these average gaps will be hard to discern in data.

In our last experiment, we do something closer to a pure comparative static exercise. We set the parameters and the initial conditions of the Bayesian model. Then we use the *same* parameters and initial conditions in terms of state variables and simulate the diagnostic model. We think this exercise is the least economically relevant, but sheds light on why we find similar responses in our earlier exercises. Our previous experiments tie the model's hands by forcing the diagnostic model's parameters and states to match the *same observables*. In Figure 17, we present the results. Both the Bayesian model and the diagnostic model have the same parameters (other than the diagnostic parameter) and the same initial state ($w_0 = \bar{w}, \lambda_0 = 0.9\bar{\lambda}$). We also set the initial diagnostic belief to feature overoptimism at $t = 0$. We see that the recapitalization policies driver larger impulse response differences. Output is higher by about 0.3% in the diagnostic model. We conclude

that with the freedom to choose the degree of overoptimism, we can generate a larger impact of policy. However, as noted above, this experiment is likely the least relevant in terms of informing policy.

[TABLE 5 HERE]

A comparison of the simulation experiments is shown in Table 5.

7 Conclusion

Financial crises have clear regularities. The 2008 global financial crisis was not a unique event. Over the last two decades, researchers have documented a number of common empirical patterns of financial crises. The main contribution of our paper is to apply a model in the class of the recent non-linear macro-finance models (He and Krishnamurthy (2013); Brunnermeier and Sannikov (2014); Di Tella (2017); Gertler, Kiyotaki and Prestipino (2020)) plus a learning mechanism (Moreira and Savov (2017); Bordalo, Gennaioli and Shleifer (2018)) to matching these patterns. We have shown that our model with a financial amplification mechanism plus belief dynamics, either driven by Bayesian or extrapolative expectations, is able to generate patterns on the crisis cycle consistent with the empirical literature on financial crises. The model matches the pre-crisis froth and leverage build-up. It matches the sharp transition to a crisis, the left-skewed distribution of output declines and asset price declines, and the slow post-crisis recovery. The quantitative fit of the model does leave room for improvement: the model generates froth pre-crisis, but not as much as the data, and while the post-crisis credit market comovement with output is positive, it is too strong relative to the data.

Our research also helps to clarify the role of beliefs and learning in matching the crisis cycle. In our model, the crisis is triggered by a "Minsky moment;" a shock that sharply shifts agents' beliefs regarding liquidity risk and is then amplified and propagated to the macroeconomy depending on the leverage of the financial sector. The work of Gorton and Ordonez (2014) and Dang, Gorton and Holmström (2020) argues that such a shift in beliefs occurs because financial sector information is hidden, by design, during normal periods, and a crisis is the event when negative information comes to light. The shift from no-information to information is at the heart of their narrative of crises. The work of Bordalo, Gennaioli and Shleifer (2018) has instead argued that a sharp shift in beliefs in a crisis reflects a change from over-optimistic to over-pessimistic beliefs. Extrapolative expectations are at the heart of their narrative of the belief shift in a crisis. In both of these narratives, the pre-cursor to a crisis is a period where agents' perceive risk to be low, either because risk is hidden or because it is misperceived. Our work suggests that either of these narratives fit the variation

in beliefs over the crisis cycle as needed to match the crisis cycle facts. Indeed, it is likely that other models of belief fluctuation such as Kozlowski, Veldkamp and Venkateswaran (2020) where agents update their models of tail risk based on the realization of tail risk can likely also be used to address the macro crisis-cycle facts.

A reader may ask, does our analysis favor a rational or non-rational account of the crisis cycle. Our answer is that both are good models to describe the crisis cycle facts. As with any moment-matching exercise, one can include further moment targets to help discriminate among these alternative models of belief fluctuations. We show that the elements we have included in the model match the "cake" of macro crisis-cycle facts. Discriminating further among these alternative models amounts to choosing flavors of "icing." For example, Baron and Xiong (2017) show that if credit growth is high, at the tails of the credit growth distribution, returns on equities going forward is negative. This evidence is hard to square with any model of rational belief formation, thus favoring models of over-extrapolation such as Bordalo, Gennaioli and Shleifer (2018). On the other hand, Dang, Gorton and Holmström (2020) points to the importance of debt as a factor in financial crises, and argues that debt is the financial contract that creates opacity. In their work, all agents are Bayesian. Note that there is nothing in our research that rules out that both of these mechanisms are at play, possibly for different agents. Our research just shows that the belief fluctuations generated by these models are consistent with the macro crisis-cycle facts we have presented.

Finally, we have also shown that for a leaning-against-the-wind macro policy experiment it does not matter whether beliefs are Bayesian or diagnostic. The response of the economy to this experiment depends on observables such as the credit spread and leverage. Different models map these observables to different values of the state variables, but given the observables, the impulse response of the models we study are quite similar. There are further policy experiments that our model can be used to address. For example, in our model, the regulation is introduced via an "MIT" shock. It will be interesting to consider a leverage tax or a state-contingent capital regulation, in a manner that agents anticipate, and recompute the equilibrium of the model. The model can be used to see how the stochastic properties of output are altered across these policies, as well as across belief models. Our preliminary investigation suggests that it should be possible to do these exercises.
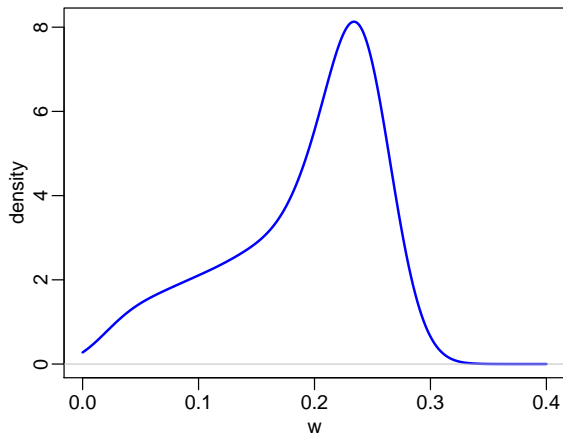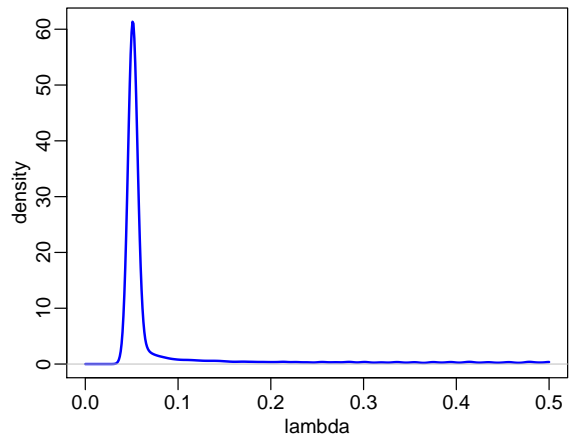
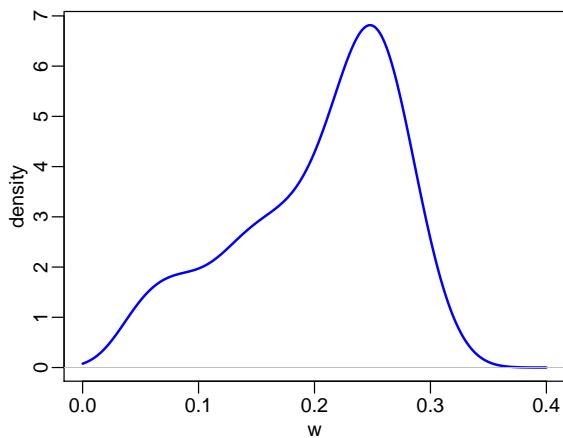# Figures and Tables
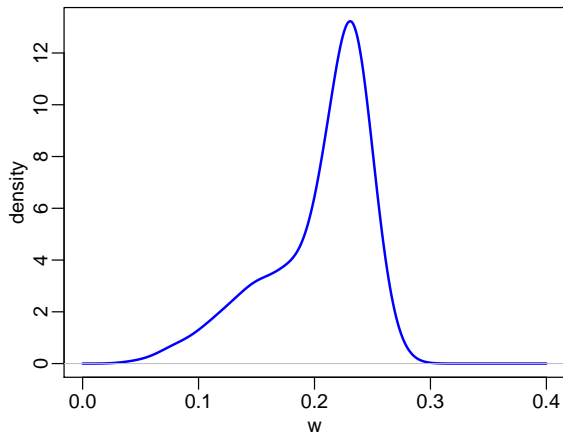
(a) w density of Static Belief model
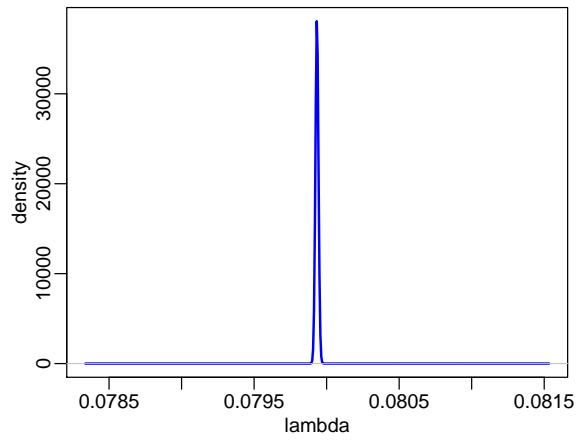
(b) λ density of Static Belief model
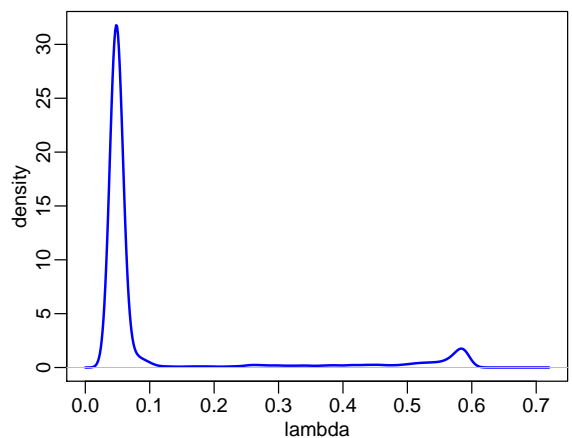
(c) w density of Bayesian model

(d) λ density of Bayesian model

(e) w density of diagnostic model

(f) λ density of diagnostic model

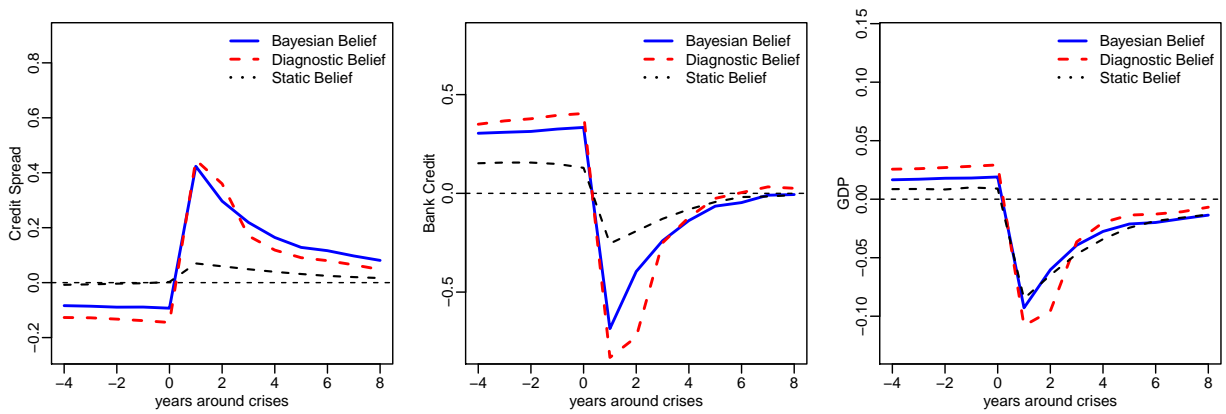Figure 6: Stationary Distribution of State Variables in Each Model

Figure 7: Dynamics of Different Models Around Crises. Credit spread and bank credit are measured as standard-deviations from the mean value. For example, credit spread raising to 0.2 means that it is larger than the value at year 0 by 0.2σs. GDP is measured in terms of percentage deviation from the long-run mean value.
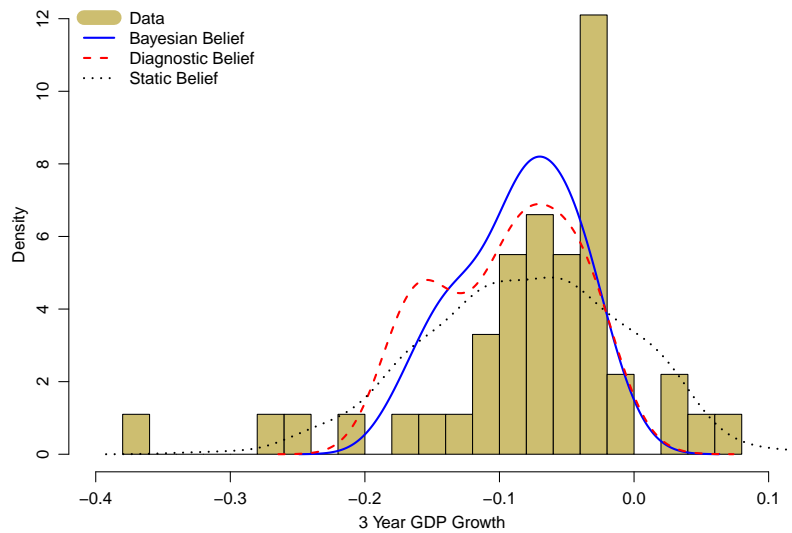

Figure 8: 3-Year GDP Growth: Model versus Data
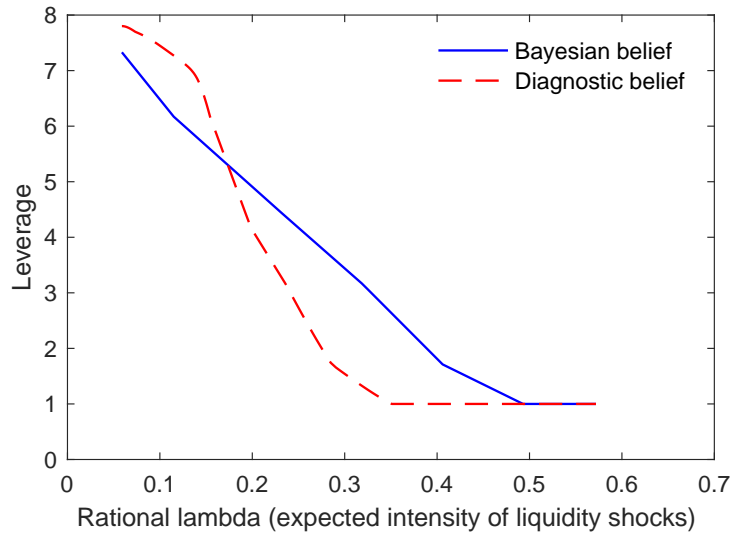
Figure 9: Expected Distress Frequency and Bank Leverage. This figure plots the leverage of banks as a function of the rational belief λ, given the same state variable w. We simulate the diagnostic model to derive the model-implied relationship between rational λ and the diagnostic belief λ^θ, and show the corresponding leverage choice.
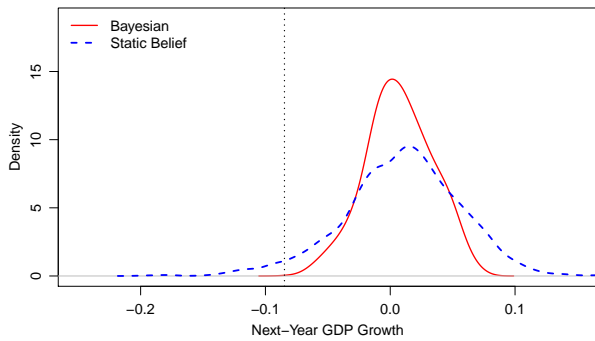
(a) Static Belief v.s. Bayesian: Low Bank Credit
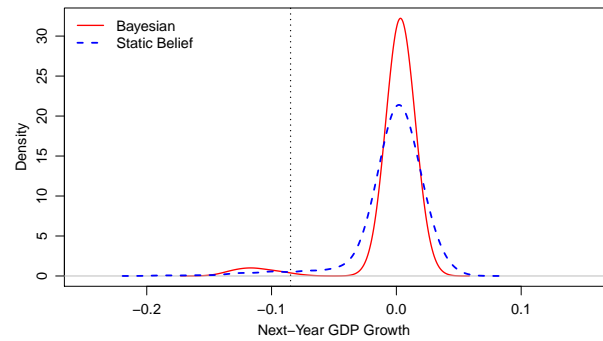
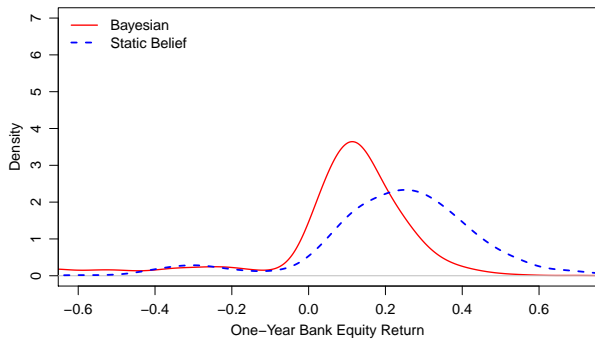(b) Static Belief v.s. Bayesian: High Bank Credit
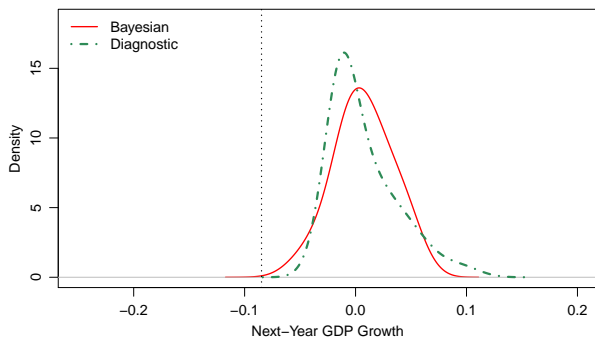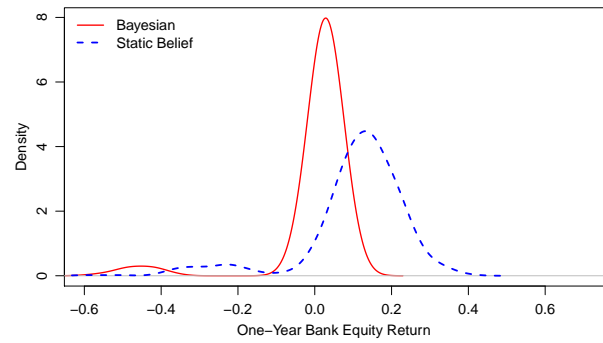
(c) Bayesian v.s. Diagnostic: Low Bank Credit
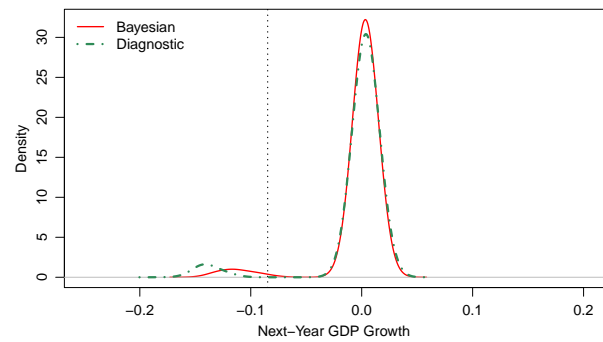
(d) Bayesian v.s. Diagnostic: High Bank Credit

Figure 10: Density of Next-Year GDP Growth Conditional on Bank Credit/GDP. Cutoffs are 30% quantile and 90% quantile of bank credit/GDP.
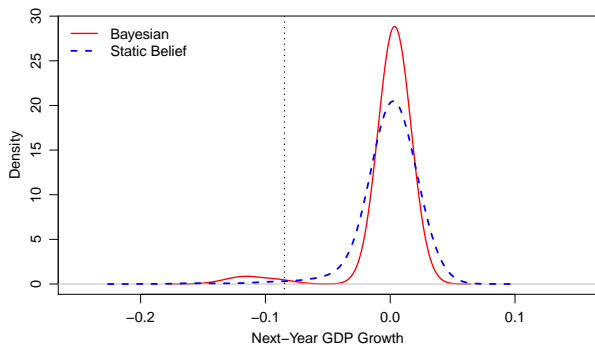
44

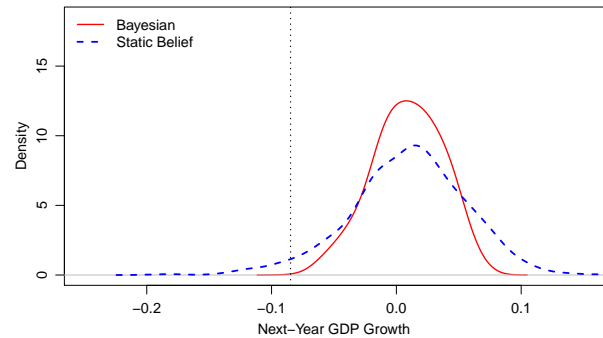(a) Static Belief v.s. Bayesian: Low Bank Credit

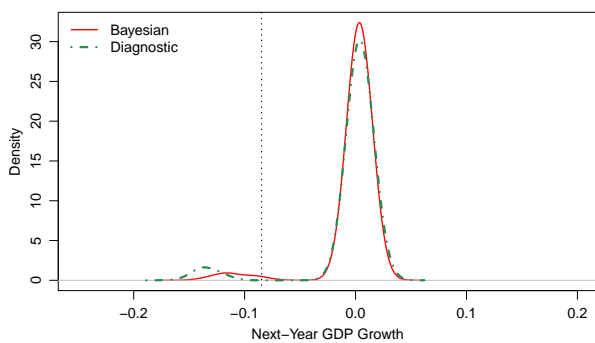(b) Static Belief v.s. Bayesian: High Bank Credit

Figure 11: Density of Bank Equity Returns Conditional on Bank Credit/GDP. Cutoffs are 30% quantile and 90% quantile of bank credit/GDP.

(a) Static Belief v.s. Bayesian: Low Credit Spread

(b) Static Belief v.s. Bayesian: High Credit Spread

(c) Bayesian v.s. Diagnostic: Low Credit Spread
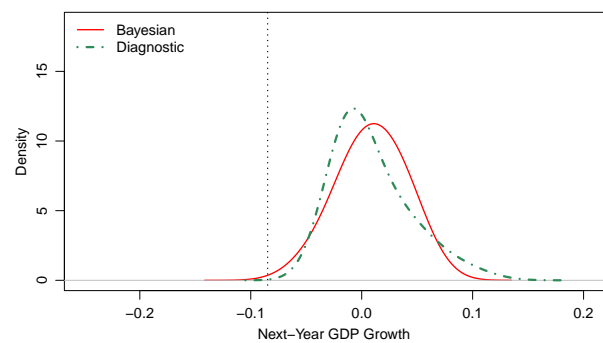
(d) Bayesian v.s. Diagnostic: High Credit Spread

Figure 12: Density of Next-Year GDP Growth in Bayesian and Diagnostic Models Conditional on Credit Spread. Cutoffs are 30% quantile and 90% quantile of bank credit/GDP.
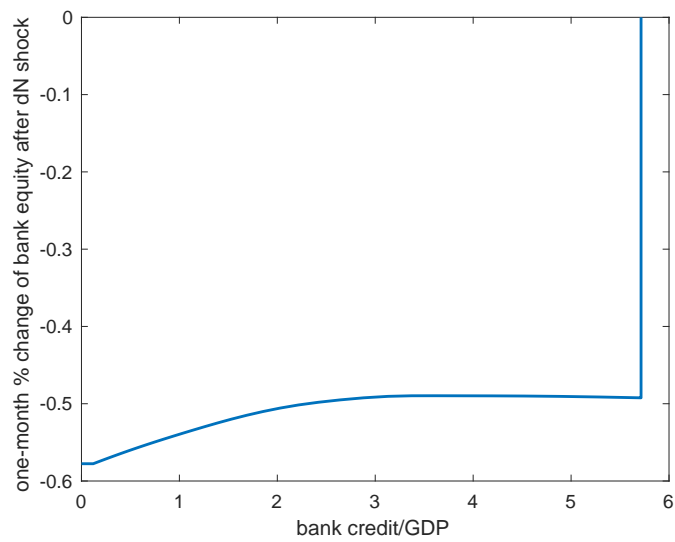
Figure 13: Bank Credit/GDP and Next-Month Bank Equity Returns in the Static-Belief
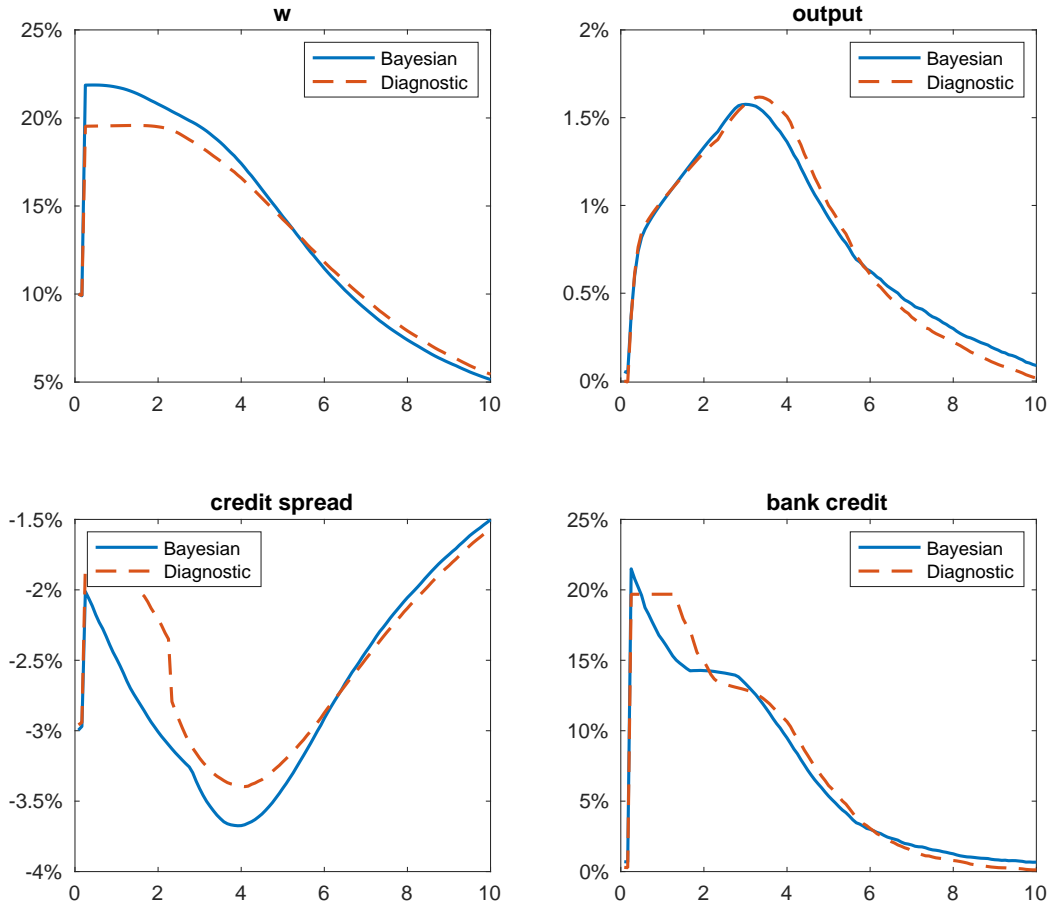Model

Figure 14: Impulse Responses of Experiment 1. In this figure, we show the impulse responses to a recapitalization policy at $t = 0$ that increases banker wealth share w by 10%, in order to "lean against the wind" and avoid future losses in a liquidity distress. The starting state is a "boom state", solved by matching a normal bank leverage but a credit spread 5% below its average. In the diagnostic model, $\lambda_0^\theta = \lambda_0$ so that the diagnostic belief is correct at the beginning. Both the Bayesian and the diagnostic models are the calibrated versions as in Table 2. The impulse responses are percentage deviations between with and without the recapitalization policy. In both cases, we introduce a $dN_t = 1$ shock at the first month $(t = 1/12)$, but set $dN_t = 0$ otherwise. The Brownian shocks dB_t are randomly generated. We simulate the model by 10000 times and show the average impulse responses in the graph.
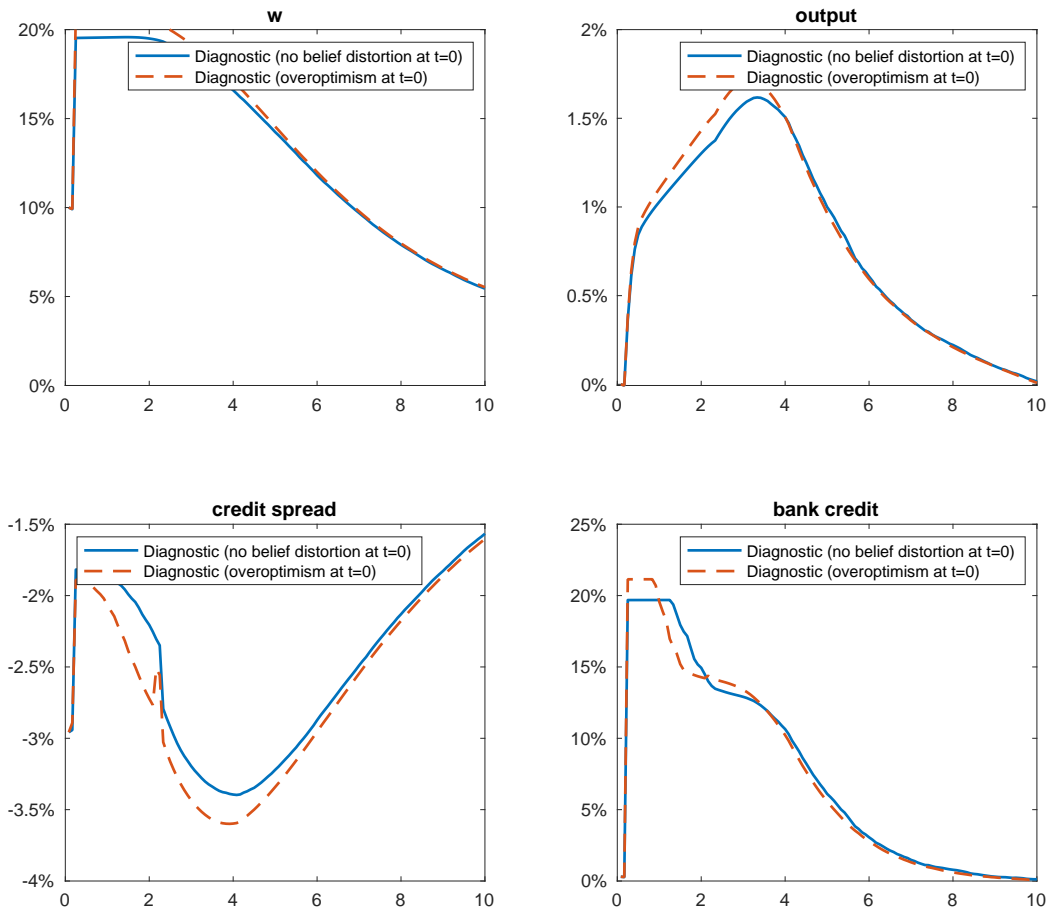
Figure 15: Impulse Responses of Experiment 2. Experiment 2 is identical to experiment 1 except for $\lambda_0^\theta > \lambda_0$ (overoptimism) in simulating the diagnostic model. Specifically, both Bayesian and diagnostic models have the same credit spread and bank leverage at $t = 0$, but the true frequency of distress in the diagnostic model is higher than the believed frequency. More descriptions are provided in Table 5 and footnotes of Figure 14.
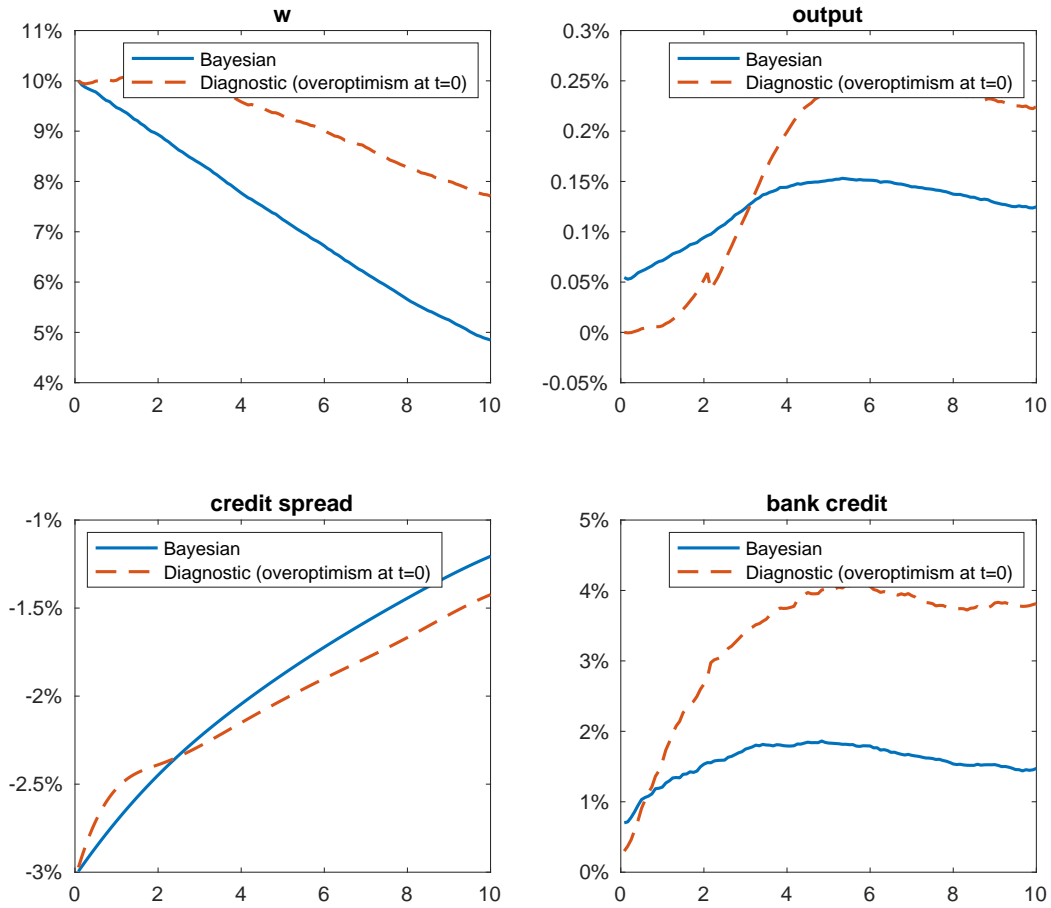
Figure 16: Expected Impulse Responses of Experiment 2. In this figure, we illustrate the expected impact of the recapitalization policy in experiment 2, by simulating dN_t according to the underlying process instead of setting $dN_t = 1$ at $t = 1/12$. More descriptions of experiment 2 are provided in Table 5 and footnotes of Figure 14 and 15.
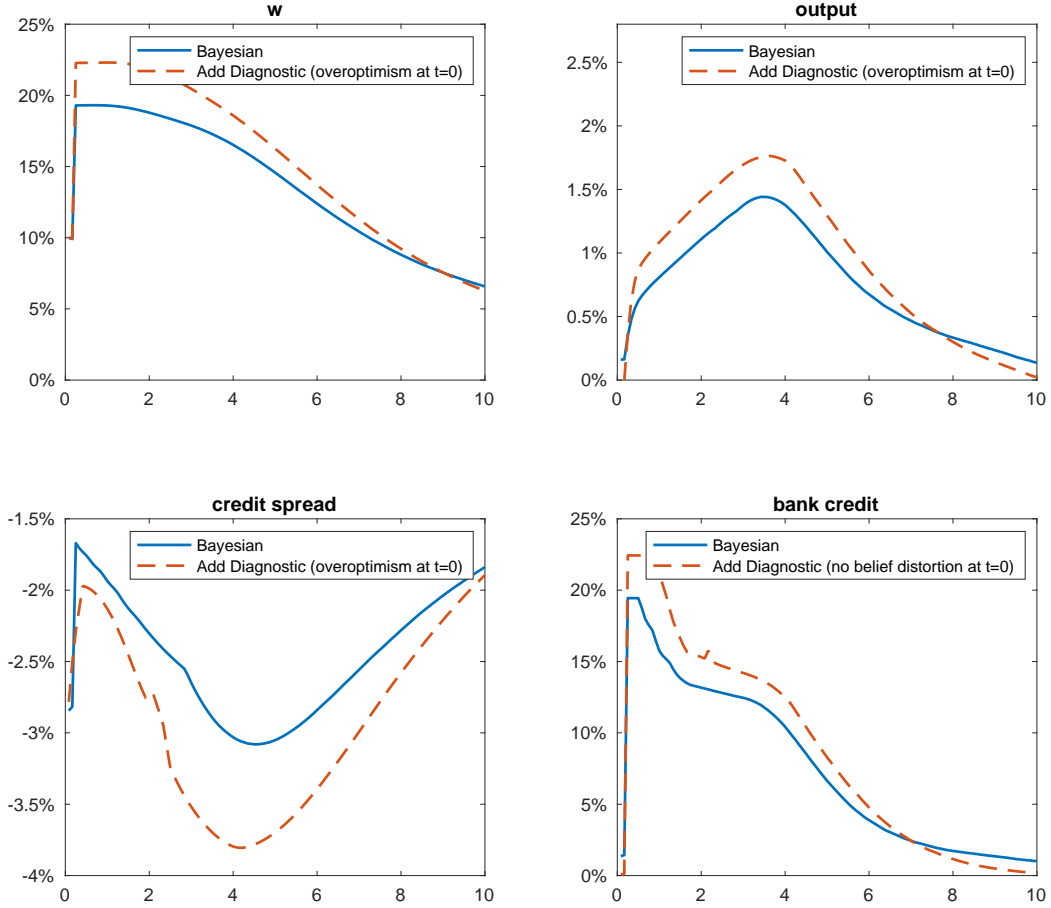
Figure 17: Expected Impulse Responses of Experiment 3. Experiment 3 shows the typical evaluation of the diagnostic belief in the literature, by keeping the Bayesian component fixed (same parameters and same starting states) but introducing the diagnostic element. In experiment 3, unlike experiment 1 and 2, the $t = 0$ observables including credit spread and bank leverage could be different across the Bayesian and diagnostic model (refer to Table 5). All impulse responses are with respect to a recapitalization policy at $t = 0$ that increases banker wealth share w by 10%, in order to "lean against the wind" and avoid future losses in a liquidity distress. The starting state is a "boom state", with (w_0, λ_0) matched to the same values as the rational model in experiment 1 (refer to Table 5). The diagnostic belief features overoptimism so that $\lambda_0^\theta < \lambda_0$. The impulse responses are percentage deviations between with and without the recapitalization policy. In both cases, we introduce a $dN_t = 1$ shock at the first month ($t = 1/12$), but set $dN_t = 0$ otherwise. The Brownian shocks dB_t are randomly generated. We simulate the model by 10000 times and show the average impulse responses in the graph.

Panel A: Credit Spread, Bank Credit, and Crisis Severity

	Dependent variable: GDP Growth from t to $t+3$
	Static Belief		Bayesian		Diagnostic		Data
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δcredit spread$_t$*crisis$_t$	-4.99		-3.20		-3.44		-7.46 (0.16)	
$(\frac{\text{bank credit}}{\text{GDP}})_t$*crisis$_t$		-0.97		-2.42		-3.23		-0.95 (0.30)
Observations							641	641

Note: Model and data regressions are normalized so that the coefficients reflect the impact of one sigma change in spreads, and bank credit/GDP.

Panel B: Bank Credit and Risk Premia

	Dependent variable: Average realized excess return$_{t+1}$
	Static Belief	Bayesian	Diagnostic	Data
$(\frac{\text{bank credit}}{\text{GDP}})_t$	-0.01	-0.01	-0.03	-0.02 (0.01)
Observations				867

Note: Model excess return is defined as the return to capital minus the risk-free rate. Data excess return is from Online Appendix of Baron and Xiong (2017) (Table 3, column 1 of Panel B). To ensure comparability, the model return to capital has been normalized to equal the standard deviation of returns reported by Baron and Xiong (2017).

Panel C: Credit Spread Before Crises

	Dependent variable: credit spread$_t$
	Static Belief	Bayesian	Diagnostic	Data
	(1)	(2)	(3)	(4)
pre-crisis	0.31	-0.15	-0.28	-0.34
Observations				634

Note: regression is: $s_t = \alpha + \beta \cdot 1\{t \text{ is before a crisis}\} + controls$. For the model, "pre-crisis" is defined as within 1 year before the next crisis. For the data, "pre-crisis" is defined as within 5 years before the next crisis. For both model and data, controls include an indicator of within 5 years after the last crisis. The data regression has more controls such as country fixed effect.

Panel D: Predicting Crises

	Dependent variable: crisis$_{t+1,t+5}$
	Static Belief		Bayesian		Diagnostic		Data
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Froth$_t$	-10.25		3.33		7.95		18.0	
Bank Credit/GDP$_t$		-6.79		1.47		2.44		2.80
Observations							528	1272

Note: Froth in the model measures if the credit spread is below the median at date t. In the data regression, froth measures if credit spread is below the median over $t-5$ to t (see Krishnamurthy and Muir (2020)). In both model and data we run a Logit regression of crisis occurring over the next 5 years on the froth measure and report the probability. Bank credit/GDP is the current ratio of bank credit over GDP. The data regression of crisis over the next year on bank credit/GDP is from Schularick and Taylor (2012), and we report the probability of the crisis.

Table 4: Using Bank Equity Crash to Define a Crisis

Panel A: Crisis, Bank Credit and Severity

	Dependent variable: GDP Growth from t to $t+3$				
	Bayesian		Diagnostic		Data
	(1)	(2)	(3)	(4)	(5)
crisis$_t$	-8.59		-9.26		-4.5
$\left(\frac{\text{bank credit}}{\text{GDP}}\right)_t*\text{crisis}_t$		-2.64		-3.72	
Observations					2548

Note: Model and data regressions are normalized so that the coefficients reflect the impact of one sigma change in spreads, and bank credit/GDP. The coefficient in column (5) is from Table I (column 4) of Baron, Verner and Xiong (2021).

Panel B: Predicting Equity Crashes

	Dependent variable: equity crash from $t+1$ to $t+3$		
	Bayesian	Diagnostic	Data
	(1)	(2)	(3)
$\left(\frac{\text{bank credit}}{\text{GDP}}\right)_t$	0.37	0.51	5.4
Observations			316

Note: The coefficient on Bank Credit/GDP is the sensitivity of crisis probability (%) to a one standard deviation increase in bank credit/GDP. The data regression is from Table III (column 7) of Baron and Xiong (2017).

Table 5: Policy Experiments

	Experiment 1 Same Observables	Experiment 2 Same Parameters, Overoptimism	Experiment 3 Same Parameters and States
Initial States of the Bayesian Model ($\theta = 0$)			
Bank Leverage$_0$	5		
Credit Spread$_0$	0.9	same as experiment 1	same as experiment 1
w_0	0.196		
λ_0	0.041		
λ_0^θ	0.041		
Initial States of the Diagnostic Model ($\theta = 0.9$)			
Bank Leverage$_0$	5	5	5
Credit Spread$_0$	0.9	0.9	0.9
w_0	0.200	0.200	0.196
λ_0	0.037	**0.104**	0.041
λ_0^θ	0.037	0.037	**0.005**

Note: This table compares the initial states of the three simulation experiments. In experiment one and two, the Bayesian model and diagnostic model are both calibrated to the same set of moments, and they have the same bank leverage and credit spread at the beginning of the simulation. In experiment one, the diagnostic belief is correct at $t = 0$, but in experiment two, the diagnostic belief features overoptimism as the underlying $\lambda_0 > \lambda_0^\theta$. In experiment 3, both the Bayesian and the diagnostic model have the same parameters as the calibrated Bayesian model, and same starting states (w_0, λ_0). However, the behavioral belief λ_0^θ is below λ_0 and there is overoptimism.

References

Baron, Matthew, and Wei Xiong. 2017. "Credit expansion and neglected crash risk." The Quarterly Journal of Economics, 132(2): 713–764.

Baron, Matthew, Emil Verner, and Wei Xiong. 2021. "Banking crises without panics." The Quarterly Journal of Economics, 136(1): 51–113.

Bianchi, Javier. 2011. "Overborrowing and systemic externalities in the business cycle." American Economic Review, 101(7): 3400–3426.

Boissay, Frédéric, Fabrice Collard, and Frank Smets. 2016. "Booms and banking crises." Journal of Political Economy, 124(2): 489–538.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2018. "Diagnostic expectations and credit cycles." The Journal of Finance, 73(1): 199–227.

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and S Terry. 2019a. "Real Credit Cycles." Harvard University.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2019b. "Diagnostic expectations and stock returns." The Journal of Finance, 74(6): 2839–2874.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in macroeconomic expectations." American Economic Review, 110(9): 2748–82.

Bordo, Michael, Barry Eichengreen, Daniela Klingebiel, and Maria Soledad Martinez-Peria. 2001. "Is the crisis problem growing more severe?" Economic policy, 16(32): 52–82.

Bordo, Michael D, and Christopher M Meissner. 2016. "Fiscal and financial crises." In Handbook of macroeconomics. Vol. 2, 355–412. Elsevier.

Borio, Claudio EV, and Philip William Lowe. 2002. "Asset prices, financial and monetary stability: exploring the nexus."

Brunnermeier, Markus K, and Yuliy Sannikov. 2014. "A macroeconomic model with a financial sector." American Economic Review, 104(2): 379–421.

Camous, Antoine, and Alejandro Van der Ghote. 2021. "Financial Cycles under Diagnostic Beliefs." Available at SSRN 3786680.

Cerra, Valerie, and Sweta Chaman Saxena. 2008. "Growth Dynamics: The Myth of Economic Recovery." The American Economic Review, 98(1): 439–457.

Chen, Long, Pierre Collin-Dufresne, and Robert S Goldstein. 2008. "On the relation between the credit spread puzzle and the equity premium puzzle." The Review of Financial Studies, 22(9): 3367–3409.

Claessens, Stijn, M Ayhan Kose, and Marco E Terrones. 2010. "The global financial crisis: How similar? How different? How costly?" Journal of Asian Economics, 21(3): 247–264.

Dang, Tri Vi, Gary Gorton, and Bengt Holmström. 2020. "The information view of financial crises." Annual Review of Financial Economics, 12: 39–65.

Dávila, Eduardo, and Ansgar Walther. 2020. "Prudential Policy with Distorted Beliefs." Available at SSRN 3694722.

Diamond, Douglas W, and Philip H Dybvig. 1983. "Bank runs, deposit insurance, and liquidity." Journal of political economy, 91(3): 401–419.

Di Tella, Sebastian. 2017. "Uncertainty shocks and balance sheet recessions." Journal of Political Economy, 125(6): 2038–2081.

Fajgelbaum, Pablo D, Edouard Schaal, and Mathieu Taschereau-Dumouchel. 2017. "Uncertainty traps." The Quarterly Journal of Economics, 132(4): 1641–1692.

Farboodi, Maryam, and Péter Kondor. 2020. "Rational Sentiments and Economic Cycles."

Fernald, John G. 2015. "Productivity and Potential Output before, during, and after the Great Recession." NBER macroeconomics annual, 29(1): 1–51.

Gertler, Mark, and Nobuhiro Kiyotaki. 2010. "Financial intermediation and credit policy in business cycle analysis." In Handbook of monetary economics. Vol. 3, 547–599. Elsevier.

Gertler, Mark, and Nobuhiro Kiyotaki. 2015. "Banking, liquidity, and bank runs in an infinite horizon economy." American Economic Review, 105(7): 2011–43.

Gertler, Mark, Nobuhiro Kiyotaki, and Andrea Prestipino. 2020. "A macroeconomic model with financial panics." The Review of Economic Studies, 87(1): 240–288.

Gilchrist, Simon, and Econ Zakrajsek. 2012. "Credit Spreads and Business Cycle Fluctuations." American Economic Review.

Gopalakrishna, Goutham. 2020. "A Macro-Finance model with Realistic Crisis Dynamics." SSRN Electronic Journal.

Gorton, Gary, and Guillermo Ordonez. 2014. "Collateral crises." American Economic Review, 104(2): 343–78.

Gorton, Gary, and Guillermo Ordonez. 2020. "Good booms, bad booms." Journal of the European Economic Association, 18(2): 618–665.

Gourinchas, Pierre-Olivier, Rodrigo Valdes, Oscar Landerretche, et al. 2001. "Lending Booms: Latin America and the World." Economía Journal, 1(Spring 2001): 47–100.

Greenwood, Robin, Samuel G Hanson, and Lawrence J Jin. 2019. "Reflexivity in credit markets." National Bureau of Economic Research.

Greenwood, Robin, Samuel G Hanson, Andrei Shleifer, and Jakob Ahm Sørensen. 2020. "Predictable financial crises." National Bureau of Economic Research.

He, Zhiguo, and Arvind Krishnamurthy. 2013. "Intermediary asset pricing." American Economic Review, 103(2): 732–70.

He, Zhiguo, and Arvind Krishnamurthy. 2019. "A macroeconomic framework for quantifying systemic risk." American Economic Journal: Macroeconomics, 11(4): 1–37.

Holmstrom, Bengt, and Jean Tirole. 1997. "Financial intermediation, loanable funds, and the real sector." the Quarterly Journal of economics, 112(3): 663–691.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor. 2011. "Financial crises, credit booms, and external imbalances: 140 years of lessons." IMF Economic Review, 59(2): 340–378.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor. 2013. "When credit bites back." Journal of Money, Credit and Banking, 45(s2): 3–28.

Judd, Kenneth L, Lilia Maliar, Serguei Maliar, and Rafael Valero. 2014. "Smolyak method for solving dynamic economic models: Lagrange interpolation, anisotropic grid and adaptive domain." Journal of Economic Dynamics and Control, 44: 92–123.

Kaplan, Greg, Kurt Mitman, and Giovanni L Violante. 2020. "The housing boom and bust: Model meets evidence." Journal of Political Economy, 128(9): 3285–3345.

Kindelberger, Charles P. 1978. Manias, Panics, and Crashes: A History of Financial Crises.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." Journal of Political Economy, 105(2): pp. 211–248.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2020. "The tail that wags the economy: Beliefs and persistent stagnation." Journal of Political Economy, 128(8): 000–000.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2015. "The impact of Treasury supply on financial sector lending and stability." Journal of Financial Economics, 118(3): 571–600.

Krishnamurthy, Arvind, and Tyler Muir. 2020. "How credit cycles across a financial crisis."

Laeven, Luc, and Fabian Valencia. 2013. "Systemic banking crises database." IMF Economic Review, 61(2): 225–270.

Liptser, Robert, and Albert N Shiryaev. 2013. Statistics of Random Processes: II. Applications. Vol. 5, Springer Science & Business Media.

Liu, Xuewen, Pengfei Wang, and Zhongchao Yang. 2020. "Delayed Crises and Slow Recoveries." Available at SSRN 3558492.

Li, Wenhao. 2019. "Public Liquidity and Financial Crises." Available at SSRN 3175101.

Maxted, Peter. 2019. "A Macro-Finance Model with Sentiment." Tech. rep.

Ma, Yueran, Teodora Paligorova, and José-Luis Peydro. 2021. "Expectations and bank lending." Work. Pap., Chicago Booth Sch. Bus., Chicago Google Scholar Article Location.

Minsky, Hyman P. 1992. "The financial instability hypothesis." The Jerome Levy Economics Institute Working Paper, , (74).

Moreira, Alan, and Alexi Savov. 2017. "The macroeconomics of shadow banking." The Journal of Finance, 72(6): 2381–2432.

Muir, Tyler. 2017. "Financial crises and risk premia." The Quarterly Journal of Economics, 132(2): 765–809.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009a. "The Aftermath of Financial Crises." American Economic Review, 99(2): 466–72.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009b. This time is different: Eight centuries of financial folly. Princeton, NJ:Princeton University Press.

Schularick, Moritz, and Alan M Taylor. 2012. "Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008." American Economic Review, 102(2): 1029–61.

Simsek, Alp. 2013. "Belief disagreements and collateral constraints." Econometrica, 81(1): 1–53.

Taylor, Alan M. 2015. "Credit, financial stability, and the macroeconomy." Annu. Rev. Econ., 7(1): 309–339.

Van Nieuwerburgh, Stijn, and Laura Veldkamp. 2006. "Learning asymmetries in real business cycles." Journal of monetary Economics, 53(4): 753–772.

A   Model Solutions

A.1   Proof of Lemma 1

We will derive the Bayesian belief process λ_t in two different ways. The first method is by applying the theorem in Liptser and Shiryaev (2013). The second one is by taking the continuous-time limit of a discrete-time process. The reason that we show the second method is because we will use the connection between discrete-time and continuous-time processes to prove the results for the diagnostic belief in Lemma 2.

Method 1

We can represent the Poisson process of bank-run as

$$N_t = \int_0^t \mathbf{1}_{\tilde{\lambda}_s = \lambda_L} dN_t^L + \int_0^t \mathbf{1}_{\tilde{\lambda}_s = \lambda_H} dN_t^H = A_t + M_t$$

where N_t^H and N_t^L are two independent Poisson processes, M_t is a martingale, and A_t is a previsible process

$$A_t = \int_0^t (\mathbf{1}_{\tilde{\lambda}_s = \lambda_L} \lambda_L + \mathbf{1}_{\tilde{\lambda}_s = \lambda_H} \lambda_H) dt$$

Denote $\mathcal{F}_t^N = \sigma\{N_s, 0 \le s \le t\}$, $\tilde{\theta} = \mathbf{1}_{\tilde{\lambda}_t = \lambda_H}$, and

$$\theta_t = E[\tilde{\theta}_t | \mathcal{F}_t^N] = P(\tilde{\lambda}_t = \lambda^H | \mathcal{F}_t^N)$$

Then according to Theorem 18.3 of Liptser and Shiryaev (2013), the compensator of N_t that is measurable with respect to $\mathcal{F}_t^N$ is

$$\bar{A}_t = \int_0^t E[(\mathbf{1}_{\tilde{\lambda}_s = \lambda^L} \lambda_L + \mathbf{1}_{\tilde{\lambda}_s = \lambda^H} \lambda_H) | \mathcal{F}_{s-}^N] ds = \int_0^t ((1 - \theta_{s-}) \lambda_L + \theta_{s-} \lambda_H) ds$$

Moroever, the compensator of θ_t is

$$\int_0^t \left(\mathbf{1}_{\tilde{\lambda}_s = \lambda_H} (-\lambda_{H \to L}) + \mathbf{1}_{\tilde{\lambda}_s = \lambda_L} \lambda_{L \to H} \right) ds$$

and the $\mathcal{F}_{t-}^N$ measurable version is

$$\int_0^t \left(\theta_{s-} (-\lambda_{H \to L}) + (1 - \theta_{s-}) \lambda_{L \to H} \right) ds$$

Finally, the martingale component of $\tilde{\theta}_t$ is independent from the jumps in N_t. Thus we can

apply Theorem 19.6 of Liptser and Shiryaev (2013) to get

$$d\theta_t = (\theta_{t-}(-\lambda_{H\to L}) + (1-\theta_{t-})\lambda_{L\to H})\, dt + E[\tilde{\lambda}_t(\frac{dA_t}{d\bar{A}_t} - 1)|\mathcal{F}_{t-}^N]d(N_t - \bar{A}_t)$$

$$= (\theta_{t-}(-\lambda_{H\to L}) + (1-\theta_{t-})\lambda_{L\to H})\, dt + E[\mathbf{1}_{\tilde{\lambda}_t=\lambda^H}(\frac{\mathbf{1}_{\tilde{\lambda}_t=\lambda^L}\lambda_L + \mathbf{1}_{\tilde{\lambda}_t=\lambda^H}\lambda_H}{(1-\theta_{t-})\lambda_L + \theta_{t-}\lambda_H} - 1)|\mathcal{F}_{t-}^N](dN_t - ((1-\theta_{t-})\lambda_L + \theta_{t-}\lambda_H$$

$$= (\theta_{t-}(-\lambda_{H\to L}) + (1-\theta_{t-})\lambda_{L\to H})\, dt + \frac{\theta_{t-}(1-\theta_{t-})(\lambda_H - \lambda_L)}{(1-\theta_{t-})\lambda_L + \theta_{t-}\lambda_H}(dN_t - ((1-\theta_{t-})\lambda_L + \theta_{t-}\lambda_H)dt)$$

$$= (\theta_{t-}(-\lambda_{H\to L}) + (1-\theta_{t-})\lambda_{L\to H} - \theta_{t-}(1-\theta_{t-})(\lambda_H - \lambda_L))\, dt + \frac{\theta_{t-}(1-\theta_{t-})(\lambda_H - \lambda_L)}{(1-\theta_{t-})\lambda_L + \theta_{t-}\lambda_H}dN_t$$

Denote $\lambda_t = E[\tilde{\lambda}_t|\mathcal{F}_t^N]$. We can get the motion of λ_t from

$$\lambda_t = E[\mathbf{1}_{\tilde{\lambda}_t=\lambda_H}|\mathcal{F}_t^N]\lambda_H + E[\mathbf{1}_{\tilde{\lambda}_t=\lambda_L}|\mathcal{F}_t^N]\lambda_L$$

$$\Rightarrow \theta_t = \frac{\lambda_t - \lambda_L}{\lambda_H - \lambda_L}$$

which results in

$$d\lambda_t = \left(\begin{array}{c} (\lambda_L - \lambda_{t-})\lambda_{H\to L} + (\lambda_H - \lambda_{t-})\lambda_{L\to H} \\ -(\lambda_{t-} - \lambda_L)(\lambda_H - \lambda_{t-}) \end{array} \right) dt + \frac{(\lambda_{t-} - \lambda_L)(\lambda_H - \lambda_{t-})}{\lambda_{t-}}dN_t$$

Method 2

Consider a discrete-time Markov process $\tilde{\lambda}_k$ with two states λ_H and λ_L. We define $\Delta t * \tilde{\lambda}_k$ as the probability of a financial distress shock within a single period. The transition probability from high to low is $\lambda_{H\to L}\Delta t$, and the transition probability from low to high is $\lambda_{L\to H}\Delta t$. We note that as $\Delta t \to 0$, this discrete-time Markov chain converges to the continuous-time Markov chain in our main model.

Agents observe the realizations of financial distress shocks, and update their beliefs. Denote the crash realization process as $N_k \in \{0, 1\}$, and the filtration as $\mathcal{F}_k = \sigma\{N_1, N_2, \cdots, N_k\}$. Denote the updated belief at period k as $\lambda_k = \mathbb{E}[\tilde{\lambda}_k|\mathcal{F}_k]$, with $\tilde{\lambda}_k$ the state of the hidden Markov process. In each period, the financial distress shock first realizes, and then the agent updates belief for that period.

Suppose that the belief on the probability at high state λ_H is π_k at period k. Then the relationship between π_k and λ_k is as follows:

$$\lambda_k = \pi_k\lambda_H + (1-\pi_k)\lambda_L$$

Observing $N_{k+1} = n_k \in \{0, 1\}$, the belief π_{k+1} is

$$\pi_{k+1} = P(\tilde{\lambda}_{k+1} = \lambda_H | N_{k+1} = n_{k+1}, \pi_k)$$

$$= \frac{P(N_{k+1} = n_{k+1}|\tilde{\lambda}_{k+1} = \lambda_H, \pi_k)P(\tilde{\lambda}_{k+1} = \lambda_H|\pi_k)}{P(N_{k+1} = n_{k+1}|\tilde{\lambda}_{k+1} = \lambda_H, \pi_k)P(\tilde{\lambda}_{k+1} = \lambda_H|\pi_k) + P(N_{k+1} = n_{k+1}|\tilde{\lambda}_{k+1} = \lambda_L, \pi_k)P(\tilde{\lambda}_{k+1} = \lambda_L|\pi_k)}$$

Note that the probabilities $P(\tilde{\lambda}_{k+1} = \lambda_H|\pi_k)$ and $P(\tilde{\lambda}_{k+1} = \lambda_L|\pi_k)$ can be calculated from the Markov one-step transition

$$\begin{pmatrix} \pi_k \\ 1 - \pi_k \end{pmatrix}^T \begin{pmatrix} 1 - \lambda_{H \to L}\Delta t & \lambda_{H \to L}\Delta t \\ \lambda_{L \to H}\Delta t & 1 - \lambda_{L \to H}\Delta t \end{pmatrix} = \begin{pmatrix} \pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t \\ \pi_k\lambda_{H \to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L \to H}\Delta t) \end{pmatrix}^T$$

which results in

$$P(\tilde{\lambda}_{k+1} = \lambda_H|\pi_k) = \pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t$$

and

$$P(\tilde{\lambda}_{k+1} = \lambda_L|\pi_k) = \pi_k\lambda_{H \to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L \to H}\Delta t)$$

Therefore, the belief π_{k+1} is

$$\pi_{k+1} = \frac{((n_{k+1}\lambda_H\Delta t + (1 - n_{k+1})(1 - \lambda_H\Delta t))(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t))}{\begin{pmatrix} (n_{k+1}\lambda_H\Delta t + (1 - n_{k+1})(1 - \lambda_H\Delta t))(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t) \\ + (n_{k+1}\lambda_L\Delta t + (1 - n_{k+1})(1 - \lambda_L\Delta t))(\pi_k\lambda_{H \to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L \to H}\Delta t)) \end{pmatrix}}$$

Now it is easier to separately discuss $n_{k+1} = 0$ and $n_{k+1} = 1$. Suppose that no financial distress shock happens ($n_{k+1} = 0$), then we have

$$\pi_{k+1} = \frac{(1 - \lambda_H\Delta t)(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t)}{\begin{pmatrix} (1 - \lambda_H\Delta t)(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t) \\ + (1 - \lambda_L\Delta t)(\pi_k\lambda_{H \to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L \to H}\Delta t)) \end{pmatrix}}$$

Suppose that a financial distress shock happens ($n_{k+1} = 1$), then we have

$$\pi_{k+1} = \frac{\lambda_H\Delta t(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t)}{\begin{pmatrix} \lambda_H\Delta t(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t) \\ + \lambda_L\Delta t(\pi_k\lambda_{H \to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L \to H}\Delta t)) \end{pmatrix}}$$

$$= \frac{\lambda_H(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t)}{\begin{pmatrix} \lambda_H(\pi_k(1 - \lambda_{H \to L}\Delta t) + (1 - \pi_k)\lambda_{L \to H}\Delta t) \\ + \lambda_L(\pi_k\lambda_{H \to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L \to H}\Delta t)) \end{pmatrix}}$$

Note that taking $\Delta t \to 0$ will result in $\pi_{k+1} = \pi_k$ when $n_{k+1} = 0$. This is reasonable, because this is like calculating $\mu_t dt$ for the λ_t process in continuous time, which is a small

order term. An appropriate way to derive the time limit is to calculate

$$\lim_{\Delta t \to 0} \frac{\pi_{k+1} - \pi_k}{\Delta t}\Big|_{n_{k+1}=0,\mathcal{F}_k}$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \begin{pmatrix} (1 - \lambda_H \Delta t)\left(\pi_k(1 - \lambda_{H\to L}\Delta t) + (1 - \pi_k)\lambda_{L\to H}\Delta t\right) \\ -\pi_k(1 - \lambda_H \Delta t)(\pi_k(1 - \lambda_{H\to L}\Delta t) + (1 - \pi_k)\lambda_{L\to H}\Delta t) \\ -\pi_k(1 - \lambda_L \Delta t)(\pi_k \lambda_{H\to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L\to H}\Delta t)) \end{pmatrix}$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \begin{pmatrix} (1 - \pi_k)(1 - \lambda_H \Delta t)\left(\pi_k(1 - \lambda_{H\to L}\Delta t) + (1 - \pi_k)\lambda_{L\to H}\Delta t\right) \\ -\pi_k(1 - \lambda_L \Delta t)(\pi_k \lambda_{H\to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L\to H}\Delta t)) \end{pmatrix}$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \begin{pmatrix} (1 - \pi_k)\left(\pi_k - \pi_k \lambda_{H\to L}\Delta t + (1 - \pi_k)\lambda_{L\to H}\Delta t - \lambda_H \pi_k \Delta t\right) \\ -\pi_k\left(\pi_k \lambda_{H\to L}\Delta t + (1 - \pi_k)(1 - \lambda_{L\to H}\Delta t) - \lambda_L(1 - \pi_k)\Delta t\right) \end{pmatrix} \quad \text{(removing } \Delta t^2 \text{ terms)}$$

$$= -\pi_k \lambda_{H\to L} + (1 - \pi_k)\lambda_{L\to H} - (\lambda_H - \lambda_L)\pi_k(1 - \pi_k)$$

Therefore, we have

$$\lim_{\Delta t \to 0} \frac{\pi_{k+1} - \pi_k}{\Delta t}\Big|_{n_{k+1}=0,\mathcal{F}_k} = -\pi_k \lambda_{H\to L} + (1 - \pi_k)\lambda_{L\to H} - (\lambda_H - \lambda_L)\pi_k(1 - \pi_k) \quad (34)$$

To build an exact connection to λ_k, we can write λ_k in terms of π_k as

$$\pi_k = \frac{\lambda_k - \lambda_L}{\lambda_H - \lambda_L} \quad (35)$$

Then the limit of $\Delta t \to 0$ expressed with λ_k is

$$\frac{1}{\lambda_H - \lambda_L} \frac{\lambda_{k+1} - \lambda_k}{\Delta t}\Big|_{n_{k+1}=0,\mathcal{F}_k} = -\frac{\lambda_k - \lambda_L}{\lambda_H - \lambda_L}\lambda_{H\to L} + \frac{\lambda_H - \lambda_k}{\lambda_H - \lambda_L}\lambda_{L\to H} - (\lambda_H - \lambda_L)\frac{\lambda_k - \lambda_L}{\lambda_H - \lambda_L}\frac{\lambda_H - \lambda_k}{\lambda_H - \lambda_L}$$

which can be simplified as

$$\lim_{\Delta t \to 0} \frac{\lambda_{k+1} - \lambda_k}{\Delta t}\Big|_{n_{k+1}=0,\mathcal{F}_k} = (\lambda_L - \lambda_k)\lambda_{H\to L} + (\lambda_H - \lambda_k)\lambda_{L\to H} - (\lambda_k - \lambda_L)(\lambda_H - \lambda_k) \quad (36)$$

Suppose that a financial distress shock happens ($n_{k+1} = 1$). By taking $\Delta t \to 0$, the updating is

$$\pi_{k+1}\big|_{n_{k+1}=1,\mathcal{F}_k} = \frac{\lambda_H \pi_k}{\lambda_H \pi_k + \lambda_L(1 - \pi_k)}$$

Using (35), the updating is

$$\frac{1}{\pi_{k+1}} = 1 + \frac{\lambda_L}{\lambda_H}\frac{1 - \pi_k}{\pi_k}$$

$$\lambda_{k+1} = \frac{\lambda_H(\lambda_k - \lambda_L)}{\lambda_k} + \lambda_L = \frac{(\lambda_H + \lambda_L)\lambda_k - \lambda_H \lambda_L}{\lambda_k}$$

which implies

$$\lambda_{k+1} - \lambda_k|_{n_{k+1}=1,\mathcal{F}_k} = \frac{(\lambda_H + \lambda_L)\lambda_k - \lambda_H\lambda_L}{\lambda_k} - \lambda_k = \frac{(\lambda_H - \lambda_k)(\lambda_k - \lambda_L)}{\lambda_k}$$

Finally, we express the above with the continuous-time notation dN_t and dt to get

$$d\lambda_t = \left(\begin{array}{c} (\lambda_L - \lambda_{t-})\lambda_{H\to L} + (\lambda_H - \lambda_{t-})\lambda_{L\to H} \\ -(\lambda_{t-} - \lambda_L)(\lambda_H - \lambda_{t-}) \end{array} \right) dt + \frac{(\lambda_H - \lambda_{t-})(\lambda_{t-} - \lambda_L)}{\lambda_{t-}} dN_t$$

which is the same as method 1.

A.2   Proof of Lemma 2

To prove Lemma 2, we start with discrete time process and then take the continuous-time limit. The discrete-time distress frequency process $\tilde{\lambda}_t$ is the same as Section A.1. Specifically, the process has two states λ_H and λ_L, with transition probability from high to low as $\lambda_{H\to L}\Delta t$, and the transition probability from low to high as $\lambda_{L\to H}\Delta t$. Agents observe the realizations of financial distress shocks, and update their beliefs. Denote the crash realization process as $N_k \in \{0, 1\}$, and the filtration as $\mathcal{F}_k = \sigma\{N_1, N_2, \cdots, N_k\}$. Denote the updated belief at period k as $\lambda_k = \mathbb{E}[\tilde{\lambda}_k|\mathcal{F}_k]$, with $\tilde{\lambda}_k$ the state of the hidden Markov process. Also denote the probability $\pi_k = P(\tilde{\lambda}_k = \lambda_H)$, which implies

$$\lambda_k = \pi_k\lambda_H + (1 - \pi_k)\lambda_L$$

We choose the period length Δt so that $T(\Delta t) = t_0/\Delta t$ is an integer, where t_0 is the "look-back period" for the diagnostic belief. Then we denote the reference probability for the diagnostic belief at period k as

$$\pi_k^T = P(\tilde{\lambda}_k = \lambda_H|\pi_{k-T(\Delta t)})$$

We already know from method 2 of Section A.1 that when $\Delta t \to 0$, the continuous-time limit of the Bayesian belief process results in (7). Our task now is to prove that the discrete-time diagnostic belief process converges to a continuous-time process as in (10). By definition, the diagnostic belief at period k is

$$\pi_k^\theta = \pi_k \cdot \left(\frac{\pi_k}{\pi_k^T}\right)^\theta \frac{1}{Z_k}$$

$$1 - \pi_k^\theta = (1 - \pi_k) \cdot \left(\frac{1 - \pi_k}{1 - \pi_k^T}\right)^\theta \frac{1}{Z_k}$$

with

$$Z_k = \frac{1}{\pi_k \cdot \left(\frac{\pi_k}{\pi_k^T}\right)^\theta + (1 - \pi_k) \cdot \left(\frac{1-\pi_k}{1-\pi_k^T}\right)^\theta}$$

which implies

$$\pi_k^\theta = \pi_k \left(\frac{\pi_k}{\pi_k^T}\right)^\theta \frac{1}{\pi_k \left(\frac{\pi_k}{\pi_k^T}\right)^\theta + (1 - \pi_k)\left(\frac{1-\pi_k}{1-\pi_k^T}\right)^\theta}$$

$$= \pi_k \frac{1}{\pi_k + (1 - \pi_k)\left(\frac{\pi_k^T}{1-\pi_k^T} \Big/ \frac{\pi_k}{1-\pi_k}\right)^\theta}$$

Therefore, if $\pi_k^T < \pi_k$, then $\pi_k^\theta > \pi_k$, leading to an overreaction. Now we can replace the probability with λ_t. Define the expected $\tilde{\lambda}_k$ under the diagnostic belief as λ_k^θ. Then we have

$$\lambda_k^\theta - \lambda_L = (\lambda_k - \lambda_L)\frac{(\lambda_H - \lambda_k) + (\lambda_k - \lambda_L)}{\left(\frac{\lambda_k^T-\lambda_L}{\lambda_H-\lambda_k^T} \Big/ \frac{\lambda_k-\lambda_L}{\lambda_H-\lambda_k}\right)^\theta (\lambda_H - \lambda_k) + (\lambda_k - \lambda_L)}$$

where

$$\lambda_k^T = \pi_k^T \lambda_H + (1 - \pi_k^T)\lambda_L$$

The key is to derive π_k^T and λ_k^T under the limit of $\Delta t \to 0$ while keeping $t = k\Delta t$ constant. Using the probability transition matrix, we get

$$\begin{pmatrix} P(\lambda_k = \lambda_H | \pi_k^T) \\ P(\lambda_k = \lambda_L | \pi_k^T) \end{pmatrix}' = \begin{pmatrix} \pi_{k-T} \\ 1 - \pi_{k-T} \end{pmatrix}' \begin{pmatrix} 1 - \lambda_{H\to L}\Delta t & \lambda_{H\to L}\Delta t \\ \lambda_{L\to H}\Delta t & 1 - \lambda_{L\to H}\Delta t \end{pmatrix}^T$$

where the $'$ notation denotes transpose of a matrix. The limit of the above expression with $\Delta t \to 0$ is effectively the transition of a continuous time Markov chain, with rate matrix

$$Q = \begin{pmatrix} -\lambda_{H\to L} & \lambda_{H\to L} \\ \lambda_{L\to H} & -\lambda_{L\to H} \end{pmatrix}$$

A decomposition reveals that the two eigenvalues of this matrix are 0 and $-(a + b)$, where $a = \lambda_{H\to L}$ and $b = \lambda_{L\to H}$. The associated eigenvector formed matrix is

$$\bar{Q} = \begin{pmatrix} 1 & -a \\ 1 & b \end{pmatrix}$$

with the inverse

$$\bar{Q}^{-1} = \frac{1}{a + b}\begin{pmatrix} b & a \\ -1 & 1 \end{pmatrix}$$

Then we can decompose

$$Q = \bar{Q}\begin{pmatrix} 0 & \\ & -(a + b) \end{pmatrix}\bar{Q}^{-1}$$

Then the transition for t units of time is

$$\bar{Q}\begin{pmatrix} 1 & \\ & e^{-(a+b)t} \end{pmatrix}\bar{Q}^{-1} = \frac{1}{a+b}\begin{pmatrix} b + ae^{-(a+b)t} & a - be^{-(a+b)t} \\ b - be^{-(a+b)t} & a + be^{-(a+b)t} \end{pmatrix}$$

Using the t notation ($t = k * \Delta t$), and taking the limit $\Delta t \to 0$ while keeping t unchanged, we have

$$\lim_{\Delta t \to 0}\begin{pmatrix} P(\lambda_k = \lambda_H | \pi_k^T) \\ P(\lambda_k = \lambda_L | \pi_k^T) \end{pmatrix}^T = \begin{pmatrix} P(\lambda_t = \lambda_H | \pi_{t-t_0}) \\ P(\lambda_t = \lambda_L | \pi_{t-t_0}) \end{pmatrix}^T$$

$$= \begin{pmatrix} \pi_{t-t_0} \\ 1 - \pi_{t-t_0} \end{pmatrix}^T \frac{1}{a+b}\begin{pmatrix} b + ae^{-(a+b)t_0} & a - be^{-(a+b)t_0} \\ b - be^{-(a+b)t_0} & a + be^{-(a+b)t_0} \end{pmatrix}$$

$$\triangleq \begin{pmatrix} a_H \pi_{t-t_0} + a_L(1 - \pi_{t-t_0}) \\ b_H \pi_{t-t_0} + b_L(1 - \pi_{t-t_0}) \end{pmatrix}^T$$

where

$$\begin{pmatrix} a_H & b_H \\ a_L & b_L \end{pmatrix} = \frac{1}{a+b}\begin{pmatrix} b + ae^{-(a+b)t_0} & a - ae^{-(a+b)t_0} \\ b - be^{-(a+b)t_0} & a + be^{-(a+b)t_0} \end{pmatrix} \tag{37}$$

Therefore, the intensity process follows

$$\lambda_t^\theta - \lambda_L = (\lambda_t - \lambda_L)\frac{(\lambda_H - \lambda_t) + (\lambda_t - \lambda_L)}{\left(\frac{\lambda_t^T - \lambda_L}{\lambda_H - \lambda_t^T} / \frac{\lambda_t - \lambda_L}{\lambda_H - \lambda_t}\right)^\theta (\lambda_H - \lambda_t) + (\lambda_t - \lambda_L)} \tag{38}$$

where

$$\lambda_t^T - \lambda_L = a_H(\lambda_{t-t_0} - \lambda_L) + a_L(\lambda_H - \lambda_{t-t_0}) \tag{39}$$

$$\lambda_H - \lambda_t^T = b_H(\lambda_{t-t_0} - \lambda_L) + b_L(\lambda_H - \lambda_{t-t_0}) \tag{40}$$

When the total transition rates $a + b$ are low, we have $a_H \approx 1$, $a_L \approx 0$, $b_H \approx 0$, and $b_H \approx 1$. Then we have $\lambda_t^T \approx \lambda_{t-t_0}$. When $\lambda_t^T > \lambda_t$, i.e., the likelihood of a crisis is decreasing, then the subjective probability is even lower, with $\lambda_t^\theta < \lambda_t$. When $\lambda_t^T < \lambda_t$, i.e., the likelihood of a crisis is increasing, then the subjective probability is even higher, with $\lambda_t^\theta > \lambda_t$. These predictions are perfectly consistent with the spirit of the diagnostic expectations. The extent of such extrapolation is larger as θ becomes larger, and we have $\lambda_t^\theta = \lambda_t$ when $\theta = 0$.

A.3 Proof of Lemma 3

To save on notation, we omit the subscripts t and $t-$.

Suppose that in equilibrium, $x^K < 1$. This implies that $(x^d)^+ = (x^K - 1)^+ = 0$, which

leads to the following first order conditions for households and bankers:

$$\mu^R + \frac{\bar{A}}{p} - r^d = (\sigma^K + \sigma^p)^2 x^K + \lambda\kappa^p\frac{1}{1 - x^K\kappa^p}$$

$$\mu^R + \frac{\underline{A}}{p} - r^d = (\sigma^K + \sigma^p)^2 y^K + \lambda\kappa^p\frac{1}{1 - y^K\kappa^p}$$

Subtracting the above two equations, we obtain

$$\frac{\bar{A} - \underline{A}}{p} = \left((\sigma^K + \sigma^p)^2 + \frac{\lambda(\kappa^p)^2}{(1 - x^K\kappa^p)(1 - y^K\kappa^p)} \right) (x^K - y^K) \tag{41}$$

The first bracket on the right hand side is always positive, since the nonnegative wealth constraint implies $x^K\kappa^p < 1$ and $y^K\kappa^p < 1$. However, from market-clearing conditions (24) and (25),

$$wx^K + (1 - w)y^K = 1$$

Under the assumption of $x^K < 1$, we must have

$$y^K > x^K$$

which implies that the right-hand side of (41) should be negative. This is a contradiction since the left-hand side of (41) is positive.

Importantly, all of the above derivations go through regardless of whether we use the Bayesian Bayesian belief or the diagnostic belief, as long as bankers and households have the same belief.

In summary, we have $x^K \geq 1$ in equilibrium. In other words, bankers borrow from households in the debt market.

A.4 First-Order Conditions

In this section, we derive bank and household first-order conditions. To save on notation, we omit the subscripts t and $t-$.

From equation (27) and a bank's optimization problem in (28) and (29), we obtain the first-order condition over x^K,

$$\mu^R + \frac{\bar{A}}{p} - r^d = (\sigma^K + \sigma^p)^2 x^K + \lambda\frac{\kappa^p + \alpha}{1 - x^K\kappa^p - \alpha\Delta x} \tag{42}$$

As a result, the excess return on capital is

$$E[dR^b] = \mu^R + \frac{\bar{A}}{p} - r^d - \lambda(\kappa^p + \alpha)$$

$$= (\sigma^K + \sigma^p)^2 x^K + \lambda(\alpha + \kappa^p)\frac{x^K \kappa^p + \alpha x^d}{1 - x^K \kappa^p - \alpha x^d}$$

(43)

The household objective function can be equivalently written as

$$\max_{c^h, y^d, y^K} \left\{ \log(c^h) + \frac{1}{\rho} \left(E[\frac{dw^h}{w^h}]/dt - \frac{1}{2} \left(\frac{dw^h}{w^h} \right)^2 /dt \right) \right\}$$

(44)

Combined with household budget dynamics in (18), we obtain the first-order condition over y^K as

$$\mu^R + \frac{A}{p} - r^d \leq (\sigma^K + \sigma^p)^2 y^K + \lambda\frac{\kappa^p}{1 - \kappa^h}, \text{with equality if } y^K > 0$$

(45)

In equation (45), the left hand side is the yield spread on productive capital over bank debt, while the right hand side includes the risk-adjusted losses of productive capital in liquidity shocks. When the yield spread is lower than the cost, households do not hold productive capital and set $y^K = 0$.

Combining (45) and (42), we have

$$\frac{\bar{A} - \underline{A}}{p} \geq (\sigma^K + \sigma^p)^2 (x^K - y^K) + \lambda\frac{\kappa^p + \alpha}{1 - x^K \kappa^p - \alpha \Delta x} - \lambda\frac{\kappa^p}{1 - \kappa^h}$$

where the equality holds when $y^K > 0$.

A.5 Equilibrium Solutions

With log utility, the optimal consumption rule is $\dot{c}^b = \dot{c}^h = \rho$. Then we simplify the equilibrium conditions into the following equations:

$$\rho = \frac{\psi A^H + (1 - \psi)A^L - i}{p}.$$

(46)

$$x^K w + y^K(1 - w) = 1.$$

(47)

$$\psi = \frac{x^K w}{x^K w + y^K(1 - w)} = x^K w,$$

(48)

Next, b we derive the dynamics of state variables. We apply Ito's lemma on the definition

of wealth share in (11) and get the dynamics of w as

$$
\begin{aligned}
\frac{dw}{w} &\overset{\Delta}{=} \mu^w dt + \sigma^w dB - \kappa^w dN \\
&= (1-w)\left(\mu^b - \mu^h + (\sigma^h)^2 - \sigma^b \sigma^h - w(\sigma^b - \sigma^h)^2 - \eta\right) dt \\
&\quad + (1-w)(\sigma^b - \sigma^h)dB - (1-w)\frac{1 - \frac{1-\kappa^b}{1-\kappa^h}}{1 + w(\frac{1-\kappa^b}{1-\kappa^h} - 1)}dN.
\end{aligned}
\tag{49}
$$

where all variables in the right hand side should have subscripts $t-$ which we omit. Then we can apply Ito's lemma on price function $p(w)$ to get

$$
\begin{cases}
\mu^p = p_w w \mu^w + \frac{1}{2}p_{ww}(w\sigma^w)^2 + p_\lambda \mu^\lambda(\lambda) \\
\sigma^p = p_w w(1-w)(\sigma^b - \sigma^h) \\
\kappa^p = 1 - p(w\frac{1-\kappa^b}{1-\kappa^h - w(\kappa^b - \kappa^h)}, \lambda))/p(w, \lambda).
\end{cases}
\tag{50}
$$

To fully characterize the economy, we also need to know the dynamics of aggregate capital quantity K (although all policy functions are scalable with respect to K). Denote the Ito process for K as

$$
\frac{dK}{K} = \mu^K dt - \delta dt + \sigma^K dB,
\tag{51}
$$

We collect the system of equations for jumps from (6), (15), and (19) as follows:

$$
\begin{cases}
\kappa^b = x^K \kappa^p + \alpha x^d \\
\kappa^h = y^K \kappa^p - \alpha x^d \frac{w}{1-w} \\
\kappa^p = 1 - p(w\frac{1-\kappa^b}{1-\kappa^h - w(\kappa^b - \kappa^h)}, \lambda + \kappa^\lambda(\lambda))/p(w, \lambda)
\end{cases}
\tag{52}
$$

From (15), (18), and (50), we collect the exposure to Brownian shocks as

$$
\begin{cases}
\sigma^p = p_w w(1-w)(\sigma^b - \sigma^h) \\
\sigma^h = y^K(\sigma^K + \sigma^p) \\
\sigma^b = x^K(\sigma^K + \sigma^p).
\end{cases}
\tag{53}
$$

Diagnostic Beliefs

We solve the model with diagnostic beliefs as follows. As households act as if their beliefs are the true ones, their policy functions are the same as the model with Bayesian beliefs. However, the true (physical) frequency of jumps with differ from that of the agents' beliefs. There are two steps to clear the market during a jump with diagnostic belief:

- First, the agents interpret λ_t^θ as the Bayesian belief. After a crisis shock dN_t, the market price of capital switches to the level under this "Bayesian belief".

- The realization of belief, however, is different from the Bayesian expectation, because the diagnostic belief formation. Now additional price adjustment is needed to clear the market under the diagnostic belief.

A.6 Loan Spread

To price the loan spread, denote the bank holding of this risk-free but illiquid loan as x_t^C. Then the equivalent bank optimization problem in (28) will have two additional terms involving x_t^C:

$$... + x_{t-}^C (r_{t-}^C - r_{t-}^d) dt - x_{t-}^C \alpha dN_t$$

which implies the FOC on x_{t-}^C as

$$r_{t-}^C - r_{t-}^d = \frac{\lambda_{t-}}{1 - x_{t-}^K \kappa_{t-}^p - \alpha x_{t-}^d - \alpha x_{t-}^C} \alpha$$

Since the illiquid loan does not take up the balance sheet (i.e., we are using the existing bank SDF to price the loan), we have $x_{t-}^C = 0$ and

$$r_{t-}^C - r_{t-}^d = \frac{\lambda_{t-}}{1 - x_{t-}^K \kappa_{t-}^p - \alpha x_{t-}^d} \alpha$$

A.7 Credit Spread

In this section, we define the credit spread used in the calibration, derive the jump differential equation for the credit spread and provide the solution methodology.

Define τ as the expected maturity of the bond. We assume that the bond matures based on the realizations of a Poisson event with intensity $1/\tau$. This modeling allows for a simple recursive formulation for bond pricing. Moreover, we suppose that a fraction of the maturity events result in default, while another fraction result in full repayment. In particular, we assume that a bond matures in two cases: (1) conditional on the financial illiquidity dN_t shock, the bond matures with probability π; (2) conditional on another independent Poisson process dN_t^τ (with intensity λ_t^τ), the bond matures with probability 1. The two intensities sum up to a fixed number, i.e.,

$$\pi \lambda_t + \lambda_t^\tau = 1/\tau \tag{54}$$

where τ can be interpreted as the maturity of the bond. We can see that

$$1/\tau \geq \pi \lambda^H$$

and therefore,

$$\tau \leq \frac{1}{\pi \lambda^H}$$

which is the maximum maturity of bonds that we can define with this method.

Each risky bond has a face value of 1. One unit value of a risky asset is continuously posted to back this risky bond, i.e., the bond is fully collateralized if the bond matures as long as there is no jump in the value of the risky asset. If dN_t hits when the bond matures, the underlying risky asset's value jumps downwards by $m \cdot \kappa_{t-}^p + \hat{\kappa}_0$. The first term varies with economic conditions. It contains capital price drop κ_{t-}^p, and a multiplier m that measures the exposure of the collateral to capital price decline. The second term here a constant "baseline" loss given default. If maturity occurs with no illiquidity event, we assume that the bond pays back in full. Thus, the loss function upon maturity for the risky bond is

$$\hat{\kappa}_t = (m \cdot \kappa_{t-}^p + \hat{\kappa}_0)dN_t \tag{55}$$

This structure gives a time-varying default probability. Specifically, when a bond matures, the probability of default is

$$\frac{\pi \lambda_t}{\pi \lambda_t + \lambda_t^\tau} = \tau \pi \lambda_t \tag{56}$$

Therefore, the unconditional probability of default is $\tau \pi \bar{\lambda}$, where $\bar{\lambda}$ is the unconditional average of the expected illiquidity frequency.

Denote the current market value of this risky bond, priced using the banker's pricing kernel, as $v_t = v(w_t, \lambda_t)$, and the value of the safe bond as $\bar{v}_t$. Then we define the credit spread as

$$\mathcal{S}_t(p_{t_0}) = \frac{1}{\tau} \log(1/v_t) - \frac{1}{\tau} \log(1/\bar{v}_t) \tag{57}$$

We expect $\mathcal{S}_t \geq 0$, given that risky bonds may default, and default occurs in high marginal utility states. Solving for this credit spread involves solving an endogenous jump equation with second-order derivatives.

HJB Equations

From Ito's lemma, we have

$$dv(w, \lambda) = \frac{\partial v(w, \lambda)}{\partial w}(w\mu^w dt + w\sigma^w dB_t) + \frac{1}{2}\frac{\partial^2 v(w, \lambda)}{\partial w^2}w^2(\sigma^w)^2 dt$$

$$+ \frac{\partial v(w, \lambda)}{\partial \lambda}\mu^\lambda(\lambda)dt + (v(w + \Delta w, \lambda + \Delta \lambda) - v(w, \lambda))dN_t$$

Denote

$$\frac{dv(w, \lambda)}{v(w, \lambda)} = \mu^v dt + \sigma^v dB_t - \kappa^v dN_t$$

Matching the coefficients, we have

$$v(w, \lambda)\mu^v = \frac{\partial v(w, \lambda)}{\partial w}w\mu^w + \frac{1}{2}\frac{\partial^2 v(w, \lambda)}{\partial w^2}w^2(\sigma^w)^2 + \frac{\partial v(w, \lambda)}{\partial \lambda}\mu^\lambda(\lambda)$$

$$v(w, \lambda)\sigma^v = \frac{\partial v(w, \lambda)}{\partial w}w\sigma^w$$

$$v(w, \lambda)\kappa^v = v(w, \lambda) - v(w + \Delta w, \lambda + \Delta\lambda)$$

From banker's perspective, the optimization problem is

$$\frac{dw_t^b}{w_t^b} = \ldots + x_{t-}^v \left(\frac{dv_t}{v_{t-}} - \frac{v_{t-} - (1 - \hat{\kappa}_t)}{v_{t-}}\xi_t dN_t - \kappa_{t-}^v(1 - \xi_t)dN_t + \frac{v_{t-} - (1 - \hat{\kappa}_t)}{v_{t-}}dN_t^\tau \right)$$

with $\lambda_t^\tau = 1/\tau - \pi\lambda_t$, $\xi_t \in \{0, 1\}$, $P(\xi_t = 1) = \pi$, and $\{\xi_t\}$ is an i.i.d. process that is independent from everything else. The jump κ_{t-}^v is the amount of decline of bond price upon the distress shock if the bond does not mature during the financial distress shock.

Rewriting the above and omitting the time subscripts, we have

$$\frac{dw^b}{w^b} = \left(r^f + x^K(\mu^R + \frac{A^H}{p} - r^f) + x^d(r^f - r^d) + x^v(\mu^v - r^f) - \rho \right)dt$$

$$+ \left(x^K(\sigma^K + \sigma^p) + x^v\sigma^v \right)dB_t - (x^K\kappa^p + \alpha x^d + x^v\xi\frac{v - (1 - \kappa^p - \hat{\kappa}_0)}{v} + x^v(1-\xi)\kappa^v)dN_t - x^v\frac{v-1}{v}dN_t^\tau$$

where I have omitted the subscripts t and $t-$ for simplicity. To solve the price of the safe bond $\bar{v}$, we can simply replace the notation v with $\bar{v}$, and set the term κ^p and $\hat{\kappa}^0$ both to zero.

The first order condition over x^v is

$$\mu^v - r^f - \lambda\pi\frac{\frac{v - (1 - \kappa^p - \hat{\kappa}^0)}{v}}{1 - (x^K\kappa^p + \alpha x^d + x^v\frac{v-(1-\kappa^p-\hat{\kappa}^0)}{v})} - \lambda(1-\pi)\frac{\kappa^v}{1 - (x^K\kappa^p + \alpha x^d + x^v\kappa^v)} - \lambda^\tau\frac{\frac{v-1}{v}}{1 + x^v\frac{v-1}{v}}$$

$$- \underbrace{(\sigma^v)^2 x^v}_{\text{compensation for change in risk - bearing capacity}} - \underbrace{x^K\sigma^v(\sigma^K + \sigma^p)}_{\text{compesnation for covariance}} = 0$$

Given that in equilibrium $x^v = 0$, we have

$$\mu^v - r^f = \lambda\pi\frac{1}{1 - \kappa^b}\frac{v - (1 - \kappa^p - \hat{\kappa}^0)}{v} + \lambda(1 - \pi)\frac{1}{1 - \kappa^b}\kappa^v + \lambda^\tau\frac{v - 1}{v} + x^K\sigma^v(\sigma^K + \sigma^p)$$

with

$$\lambda^\tau = \frac{1}{\tau} - \pi\lambda$$

Therefore, the excess return has three components: (1) the compensation for losses during a distress shock, (2) the compensation for losses (negative losses mean positive benefits) in a maturity event without distress shock, and (3) the compensation for exposure to the

volatility risk dB_t, where the price of risk is $x^K(\sigma^K + \sigma^p)$. This equation together with the matched coefficients form an HJB equation for the value of bonds,

$$
\begin{aligned}
&\frac{\partial v}{\partial w} w \mu^w + \frac{1}{2} \frac{\partial^2 v}{\partial w^2} w^2 (\sigma^w)^2 + \frac{\partial v}{\partial \lambda} \mu^\lambda - r^f v = x^K(\sigma^K + \sigma^p) \frac{\partial v}{\partial w} w \sigma^w \\
&+ \lambda \pi \frac{1}{1 - \kappa^b} \left(v - (1 - \kappa^p - \hat{\kappa}^0) \right) + \lambda(1 - \pi) \frac{1}{1 - \kappa^b} \kappa^v v + \lambda^\tau (v - 1)
\end{aligned}
\tag{58}
$$

Solution Methods

We will use the "false time derivative" method, by introducing a time dependence of v. Define such a function as $\tilde{v}(w, \lambda, t)$. Following a similar derivation as (58), we can get the HJB equation for $\tilde{v}$ as

$$
\begin{aligned}
\frac{\partial \tilde{v}}{\partial t} &= \lambda \pi \frac{1}{1 - \kappa^b} \left(v - (1 - \kappa^p - \hat{\kappa}^0) \right) + \lambda(1 - \pi) \frac{1}{1 - \kappa^b} \kappa^v v + \lambda^\tau (v - 1) \\
&+ x^K(\sigma^K + \sigma^p) \frac{\partial v}{\partial w} w \sigma^w + r^f \tilde{v} - \left(\frac{\partial \tilde{v}}{\partial w} w \mu^w + \frac{1}{2} \frac{\partial^2 \tilde{v}}{\partial w^2} w^2 (\sigma^w)^2 + \frac{\partial \tilde{v}}{\partial \lambda} \mu^\lambda \right)
\end{aligned}
$$

We can start with a function $\tilde{v}$ that satisfies $\tilde{v}(0, \lambda, T) = v(0, \lambda)$, and $\tilde{v}(1, \lambda, T) = v(1, \lambda)$, and has linear interpolation in other regions. By taking T large enough, we are going to have convergence before t reaches 0, i.e., two iterations have close to zero differences. Denote the converged solution as $\tilde{v}(w, \lambda, 0)$. From the property of convergence, we must have $\partial \tilde{v}(w, \lambda, t)/\partial t|_{t=0} = 0$. As a result, $\tilde{v}(w, \lambda, 0)$ satisfies the original PDE of $v(w, \lambda)$, which implies that $v(w, \lambda) = \tilde{v}(w, \lambda, 0)$.

Next, we show how to solve the boundary conditions at $w = 0$ and $w = 1$.

Boundary Conditions

We note that $w = 0$ and $w = 1$ are two absorbing boundaries. At both $w = 0$ and $w = 1$, we have $p = \underline{p}$ or $\bar{p}$ forever, and $\mu^w = \sigma^w = \kappa^p = 0$. Thus, we can simplify the HJB equation (58) into

$$
\begin{aligned}
&\frac{\partial v(w, \lambda)}{\partial \lambda} \mu^\lambda(\lambda) - r^f(w, \lambda) v(w, \lambda) = \lambda \pi \frac{1}{1 - \kappa^b(w, \lambda)} \left(v(w, \lambda) - (1 - \hat{\kappa}^0) \right) \\
&+ \lambda(1 - \pi) \frac{1}{1 - \kappa^b(w, \lambda)} \kappa^v(w, \lambda) v(w, \lambda) + \lambda^\tau(\lambda)(v(w, \lambda) - 1), \quad w \in \{0, 1\}
\end{aligned}
\tag{59}
$$

Suppose that $\kappa^v = 0$ when $\lambda = \lambda^*$ (defined as $\mu^\lambda(\lambda^*) = 0$). Then we get

$$
v^{(0)}(w, \lambda^*) = \frac{\lambda^* \pi \frac{1}{1 - \kappa^b(w, \lambda^*)} (1 - \hat{\kappa}^0) + \lambda^\tau(\lambda^*)}{\lambda^* \pi \frac{1}{1 - \kappa^b(w, \lambda^*)} + r^f(w, \lambda^*) + \lambda^\tau(\lambda^*)}, \quad w \in \{0, 1\}
$$

Denote the value function at iteration k as $v^{(k)}(w, \lambda)$. Then for $w = 1$ or $w = 0$, the

algorithm works as follows:

- Step k: Solve for the jump $\kappa^v v = v(w, \lambda) - v(w + \delta w, \lambda + \delta \lambda)$ using $v = v^{(k)}$. Denote this value as $\Delta v^{(k)}$. With such jump solved, we translate the jump equation (59) into an ODE of $v(w, \lambda), w \in \{0, 1\}$ as a function of λ. The ODE solution starts with the initial value $v(w, \lambda^*) = v^{(k)}(w, \lambda^*)$, $w \in \{0, 1\}$. Solve this ODE and denote the solution as $v^{(k+1)}$.

- Stop if

$$\int_{\lambda_L}^{\lambda_H} |v^{(k+1)}(w, \lambda) - v^{(k)}(w, \lambda)| \, d\lambda < \varepsilon, \quad w \in \{0, 1\}$$

  for a small $\varepsilon > 0$.

Finally, we notice that once the $\lambda = \lambda^*$, it will not go up or down unless there is a dN_t shock. Once we know the jump component, we can solve $v(w, \lambda^*)$ along the w dimension as an ODE. The ODE is

$$\frac{\partial^2 v}{\partial w^2} = \frac{\left(\begin{array}{c} \lambda^* \pi \frac{1}{1-\kappa^b} \left(v - (1 - \kappa^p - \hat{\kappa}^0) \right) + \lambda(1 - \pi) \frac{1}{1-\kappa^b} \kappa^v v \\ + \lambda^\tau (v - 1) + x^K (\sigma^K + \sigma^p) \frac{\partial v}{\partial w} w \sigma^w + r^f v - \frac{\partial v}{\partial w} w \mu^w \end{array} \right)}{\frac{1}{2} w^2 (\sigma^w)^2}$$

for $w \neq 0, 1$.

Credit Spread Calibration

Table 6 summarizes the credit spread calibration.

Table 6: Calibrated Parameters for the Credit Spread Construction

Parameters		Choice	Moment
τ	Risky bond maturity	7 Years	Maturity of 7 years.
π	Maturing probability in illiquidity	0.31	Average default intensity of 0.04
$mE_{\text{crises}}[\kappa_t^p] - mE_{\text{non-crises}}[\kappa_t^p]$	Additional loss in crises	0.1	Additional loss of 10% in crises
$mE_{\text{crises}}[\kappa_t^p] + \hat{\kappa}_0$	Baseline default loss	0.55	Average loss rate of 0.55

- In our baseline calibration, we target the an average maturity of $\tau = 7$ years, which is the average maturity of bonds used in Krishnamurthy and Muir (2020).

- According to Chen, Collin-Dufresne and Goldstein (2008), the 10-year BAA (AAA) default rate is 4.89% (0.63%). The difference in their default rates is 4.26%. We use 4% as

our target. In the model, the default rate is

$$\pi \bar{\lambda} = 0.04$$

where $\bar{\lambda}$ is the average frequency of financial illiquidity, which is 12.8% according to our calibration. Therefore, we have $\pi = 0.31$.

- The total loss given default is $m \cdot \kappa_t^p + \hat{\kappa}^0$ if a illiquidity shock dN_t hits, where κ_t^p is the percentage decline of capital price p_t during a crisis shock. The price jump component κ_t^p is large during crises but close to zero otherwise. We calibrate the loss given default to that of BAA bonds, which from Moodys data has been 55% on average over the last three decades and rose by 10% during the 2008 crisis. As a result, we set m so that $m \cdot \kappa_t^p$ during crises is 10% larger than other defaults. Then we set the average of losses during default to 55% to get $\hat{\kappa}_0$.

Finally, we should note that we define our spread measures in units of standard-deviation differences relative to the unconditional mean value of the credit spread. This is what Krishnamurthy and Muir (2020) do in their empirical work. As a result of this normalization, the results are relatively insensitive to the exact values of the credit-spread calibration.