Dissecting Mechanisms of Financial Crises: Intermediation and Sentiment

Arvind Krishnamurthy and Wenhao Li*

December 2020

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 principal finding is that both the frictional intermediation mechanism and fluctuations in beliefs are needed to match the crisis cycle patterns. A pure amplification mechanism quantitatively matches the crisis and aftermath period but fails to match the pre-crisis evidence. We consider two versions of sentiment fluctuations, a Bayesian belief updating process and one that overweights recent observations. Our second finding is that the data we consider does not allow one to clearly distinguish between these two mechanisms. Both fit the patterns well. Last, we 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 condition on the true belief process.

*Krishnamurthy: Stanford Graduate School of Business and NBER, a-krishnamurthy@stanford.edu; Li: University of Southern California, Marshall School of Business, liwenhao@marshall.usc.edu. We thank Xavier Gabaix, Paymon Khorrami, Christopher Hrdlicka, Yueran Ma, Guillermo Ordonez, Alp Simsek, Laura Veldkamp, Wei Xiong, and seminar participants at USC, Stanford, UW Foster, Virtual Finance Workshop, Barcelona Asset Prices, Finance and Macroeconomics Workshop, NY Fed, NBER Asset Pricing and Stanford SITE 2020 for comments.

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 (2009*a*), Schularick and Taylor (2012), Jordà, Schularick and Taylor (2011), Jordà, Schularick and Taylor (2013), Baron and Xiong (2017), and Krishnamurthy and Muir (2017). 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 beliefs, and harkens back to Kindelberger (1978). Agents pre-crisis see a string of good-news shocks that makes them optimistic about the path of the economy. Lending grows, risk spreads are low, and output grows. Bad-news events realize that lead agents to revise their views of the economy, creating the transition to the crisis. A slow-recovery follows as beliefs slowly revert back to a steady-state level. The key state variable in these models is agents' beliefs. There are two flavors of these models: one in which learning and belief updating is Bayesian (Moreira and Savov, 2017) and the other where updating is non-rational (Bordalo, Gennaioli and Shleifer, 2018). Bordalo, Gennaioli and Shleifer (2018) argues forcefully for a form of non-rational learning whereby beliefs over-react to current news. These authors argue that such over-reaction is essential to capture the crisis patterns.

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 crisis patterns, and to understand which elements of these mechanisms are essential. We build a model with a financial intermediary sector subject to capital constraints and financed in part by demandable debt. There are two sources of shocks in the model, 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 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 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 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).

This modeling of information about an illiquidity state captures the idea that a crisis is the result of an unusual confluence of events. Following Gorton and Ordonez (2014), the conditions for such a confluence remains opaque during normal periods. For example, pre-crisis, agents may not understand the extent of interconnectedness of the financial sector. While such information will be valuable, there are few signals produced in the normal operation of the economy that is relevant to interconnectedness. In our modeling, the only signal is the lack of a crisis, leading agents to think that the illiquidity state is unlikely. When an illiquidity state occurs, agents understand the stress points of the financial system and hence update their beliefs regarding the state. Thus, we aim to capture the Gorton and Ordonez (2014) opacity idea in a simple fashion.

We consider two flavors of this learning mechanism, a Bayesian rational updating process and a non-rational diagnostic updating process that overweighs current realizations. The Bayesian learning mechanism is fairly standard. Our modeling is closest to Moreira and Savov (2017). 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).

We report four principal results:

1. The model with financial frictions and a Bayesian belief process qualitatively 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 data is more murky. Both versions of the belief mechanism qualitatively generate the crisis cycle patterns, with the diagnostic updating process bringing the quantitative results closer to matching the pre-crisis data. The one pattern that the Bayesian model cannot match is the possibility that expected returns on risky assets may be negative in some states of the world pre-crisis. Baron and Xiong (2017) present evidence for this result, albeit weaker than that of other results presented in the paper and in the literature more broadly.

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 the drivers of belief fluctations 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 match 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. In the model, the frequency of the illiquidity shock follows a hidden two-state Markov process, and agents update their beliefs over the state based on history in a Bayesian fashion. Agents' beliefs of the shock probability is thus also a state variable driving the model's dynamics. If a crisis 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 spreads fall and credit grows. From this state, if an illiquidity shock arrives, beliefs jump towards the high likelihood state and banker wealth falls, leading 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 – they drive down spreads/risk premia and increase credit. They also take actions that 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 more risk tolerance 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), 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 negative shocks reduce banker wealth (at the same time, raising leverage). 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 consider an alternative diagnostic learning mechanism that leads beliefs to have a non-rational component, over-extrapolating from recent observations. Similarly to the Bayesian model, this model matches the crisis and aftermath evidence, but it additionally gets closer to quantitatively matching the pre-crisis froth evidence. The reason is that the sensitivity of bankers' leverage decisions to the true illiquidity state is higher under diagnostic beliefs than Bayesian beliefs. This "steepening" of the leverage decision is the key force that brings the model more in line with 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 data. However, our analysis also indicates that the qualitative patterns of the data do not clearly distinguish the two belief models.

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? Although our model is not suited for welfare analysis, it can shed light on the impact of a change in policy on economic outcomes. 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 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. Key to this similarity result is that both models are calibrated to common data but are not forced to having a common parameterization, and 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.

This paper's goals and approach is closest to that of two other recent papers, Greenwood, Hanson and Jin (2019) and Maxted (2019). Both of these papers 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) presents a macro-finance model that 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. Nevertheless, the paper considers only a subset of the crisis data that we aim to match in this paper. Like Greenwood, Hanson and Jin (2019), the paper does not allow an evaluation of a Bayesian belief process. Without diagnostic expectations, the model of the paper collapses to a pure intermediation model along the lines of He and Krishnamurthy (2019). One key point of difference relative to our model is that the diagnostic belief shifts the mean drift of the capital process, whereas in our model, beliefs are over the tail of the distribution (we are similar to Greenwood, Hanson and Jin (2019) in this regard). Although it is not entirely clear at this stage which approach (shifting mean versus shifting mass in tail) is the right way forward, one difference in these two approaches is that ours directly impacts risk premia, whereas the mean shift has no direct impact on risk premia. As noted earlier, our approach of modeling information is motivated by the opacity-during-normal-periods ideas in Gorton and Ordonez (2014).

Our paper also advances the recent continuous-time macro-finance literature (He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; Di Tella, 2017). The models in this literature feature non-linearities and are solved using global methods. Thus the advantage of these models is that they can characterize the non-linear dynamics in financial crises. However, a major disadvantage of these 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.

There are other recent papers that also touch on the issues of this paper. 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. Bordalo et al. (2019) introduce diagnostic beliefs into a relatively standard RBC model. Their model does not have an intermediation mechanism and thus helps to understand the role of diagnostic beliefs. Farboodi and Kondor (2020) present a model of time-varying sentiment that generates a credit cycle that is qualitatively in line with the facts. All agents in their model are rational, so that sentiment evolves in a Bayesian manner. 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. In our model, the illiquidity shock and the bank run are exogenous. In Gorton and Ordonez (2014, 2020), the debt crisis occurs when agents endogenously choose to acquire information and this information is bad-news, turning previously safe-debt risky.

Finally, this paper also contributes to a larger literature on beliefs and learning in macroeconomics models. 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. 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. 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.

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.

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 (2009*a*), 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 (2009*a*). 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. In light of the above evidence, we pick the medium value 4% as our target.

2. Baron and Xiong (2017) measure equity market crashes, defined as a fall in bank equity market prices in excess of 30%. They report that crashes occur with a frequency of 3.2% per quarter in a sample from 1920 to 2012. Note that not every equity crash corresponds to a real crisis, which is a point also emphasized by Greenwood, Hanson and Jin (2019).

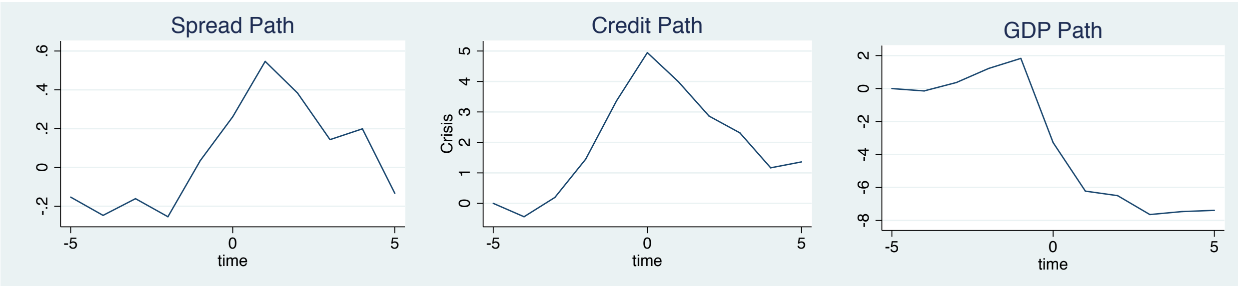


Figure 1: Mean path of credit spread, bank credit, and GDP across a sample of 41 financial crises identified in Jordà, Schularick and Taylor (2013). Units for spread path are 0.4 means that spreads are 0.4σs above their average for a given country. Units for credit path are that 5 indicates that credit/GDP 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 (2017)

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 (2017), 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 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, we see that:

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

4. GDP declines by 9.1%. Reinhart and Rogoff (2009*b*) 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 (2009*a*) 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 crisis, and the excess return on stocks rises during the crisis, indicated a generalized rise in asset market risk premia.

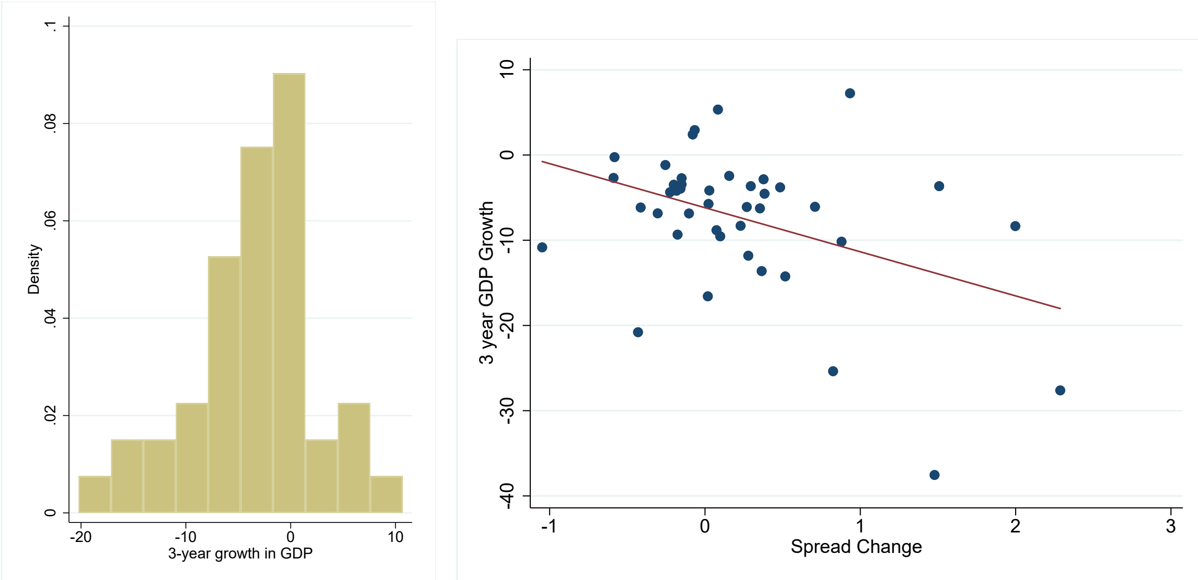
Aftermath and severity of crisis:

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

6. 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.

7. 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 (2017) 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,



(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..

8. Conditioning on a crisis at year t, and looking at the 5 years prior to the crisis, Krishnamurthy and Muir (2017) 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).

9. 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.

10. 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.

11. Conditioning on an episode where credit spreads are below their median value 5 years in a row, Krishnamurthy and Muir (2017) estimate that the conditional probability of a crisis rises by 1.76%.

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,

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

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.

We introduce two shocks that allow us to distinguish between financial crises and other fluctuations. The first is a Brownian shock dB_t that reflects every-day economic fluctuations. The second is a Poisson shock dN_t that we call a "financial illiquidity" shock. As will be clear, this shock triggers illiquidity and bank runs, and a possible financial crisis.

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.

Our key assumption is 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. Investment undertaken by an owner 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

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

The return on capital held by agent j =banker is

$$d\bar{R}_{j,t}^K = \frac{d(p_t k_{j,t})}{p_t k_{j,t}} + \frac{(\bar{A} - \phi(\mu_t^K))k_{j,t}}{p_t k_{j,t}} dt. \tag{6}$$

The return to capital held by a household, denoted by $d\underline{R}_{j,t}^K$, is the same except for the lower productivity $\underline{A}$.

The dynamics of capital price p_t is denoted as

$$\frac{dp_t}{p_{t-}} = \mu_t^p dt + \sigma_t^p dB_t - \kappa_{t-}^p dN_t, \tag{7}$$

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, Distress, and Bank Runs

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 rate r^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 a 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 if a bank raises F units of resources it pays a cost of $\alpha^0 F$. The cost can be thought of as a fire-sale liquidation cost when selling capital. Alternatively, the cost can be mapped into a premium on raising emergency financing from other banks or other households in the economy. In this latter case, we need to step outside the modeling and interpret the illiquidity event lasting longer than dt. Then, α^0 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. This assumption is not essential to the analysis.

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. Li (2019)'s objectives are normative, to study how policies forestall liquidity crises, whereas this study's objective is positive, to quantitatively understand mechanisms contributing to financial crises.

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 $\lambda_{L\to H}$. 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)**.**

$$d\lambda_t = \begin{pmatrix} (\lambda_L - \lambda_{t-})\lambda_{H\to L} + (\lambda_H - \lambda_{t-})\lambda_{L\to H} \\ -(\lambda_{t-} - \lambda_L)(\lambda_H - \lambda_{t-}) \end{pmatrix} dt + \frac{(\lambda_{t-} - \lambda_L)(\lambda_H - \lambda_{t-})}{\lambda_{t-}} dN_t \qquad (8)$$

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.

3.4 Diagnostic Expectations

Section 3.3 outlines our model when agents form expectations over $\tilde{\lambda}_t$ in a rational fashion, using Bayes rule. We also consider a version of our model where agents overweight recent observations. Specifically, we model the diagnostic beliefs of (Bordalo, Gennaioli and Shleifer, 2018). We adapt their model to our continuous 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

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

$$1 - \pi_t^\theta = (1-\pi_t) \cdot \left(\frac{1-\pi_t}{E_{t-t_0}[1-\pi_t]}\right)^\theta \frac{1}{Z_t} \tag{10}$$

where Z_t is a normalization to ensure that (9) and (10) 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

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

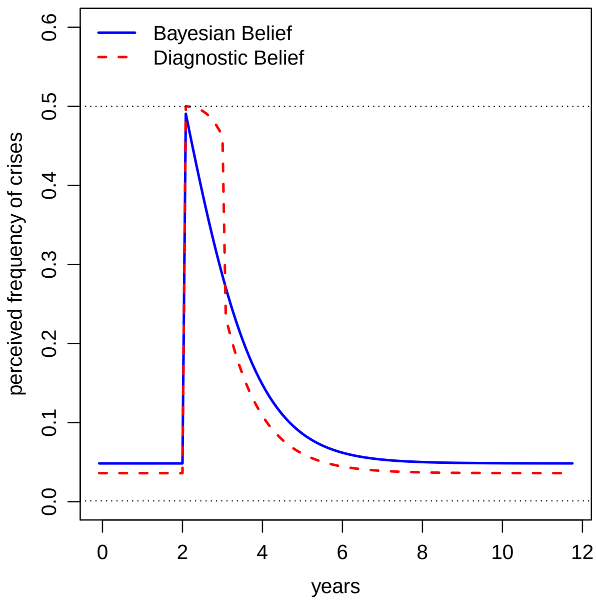
where E^θ is the expectation with respect to the probability distribution under the diagnostic belief. Then we have the following result:

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

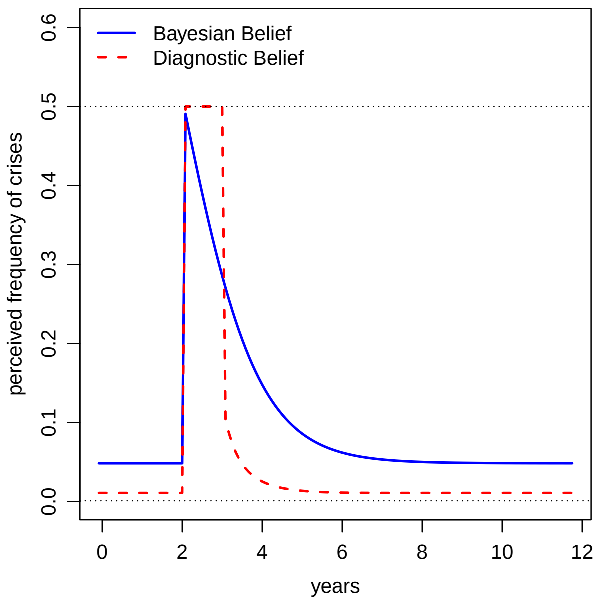
$$\lambda_t^\theta = \lambda_L + (\lambda_t - \lambda_L)\frac{(\lambda_H - \lambda_t) + (\lambda_t - \lambda_L)}{\left(\frac{\lambda_t^T - \lambda_L}{\lambda_H - \lambda_t^T} / \frac{\lambda_t - \lambda_L}{\lambda_H - \lambda_t}\right)^\theta (\lambda_H - \lambda_t) + (\lambda_t - \lambda_L)} \tag{11}$$

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 (11). We find that when θ is small, as shown in panel (a), the pre-illiquidity shock belief is slightly lower than the Bayesian belief, and then jumps to a higher level after a illiquidity shock. Initially, there is overreaction, but after one year, the perceived frequency of the illiquidity shock is below the Bayesian



(a) $\theta = 1$



(b) $\theta = 5$

Figure 3: Simulation of Beliefs with Different Values of the Diagnostic Parameter θ. The parameter $\theta \geq 0$ means the strength of the behavioral feature of the diagnostic belief. 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 (11).

belief. When θ is large, as shown in panel (b), the pre-illiquidity shock belief is much lower, and the post-illiquidity shock overreaction is stronger. One year after the illiquidity shock, the perceived frequency of illiquidity becomes much smaller.

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. 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.

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

$$w_t = \frac{W_t^b}{W_t^b + W_t^h}, \tag{12}$$

The second is the expected jump intensity λ. The final one is the total productive capital K_t. We construct an equilibrium whereby all relevant object scale linearly with capital.

This reduces the computational problem to solving a model with two state variables, w_t and λ.

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.[1] Bankers take this transition possibility into account in their optimization problems.

Bankers

Each banker can invest in productive capital and borrow from households or other banks via short-term debt at interest rate r_t^f. 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, and the interbank borrowing and lending as x_t^f with equilibrium rate r_t^f. Then the borrowing from household is $x_t^K + x_t^f - 1$. Total borrowing is $x_t^K + x_t^f - 1$. If $x_t^K + x_t^f - 1 > 0$ bankers lever up to own capital, while if $x_t^K + x_t^f - 1 < 0$, bankers save some of their wealth in riskless debt.

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

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

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.

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

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

$$V^h(w_t^h, w_t, \lambda_t) = \sup_{c_t^h \geq 0,\, y_t^K \geq 0} E[\int_t^\infty e^{-\rho(s-t)} \ln(c_s^h) ds \,\big|w_t^h, w_t], \tag{15}$$

subject to the solvency constraint

$$w_t^h \geq 0. \tag{16}$$

3.6 Equilibrium Definition

Denote the share of capital owned by bankers as

$$\psi_t = \frac{x_t^K W_t^b}{x_t^K W_t^b + y_t^K W_t^h}. \tag{17}$$

Then the aggregate production of consumption goods is

$$Y_t = (\psi_t \bar{A} + (1 - \psi_t)\underline{A})K_t. \tag{18}$$

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

$$\frac{dW_t^b}{W_t^b} = \frac{dw_t^b}{w_t^b} - \eta dt \tag{19}$$

$$\frac{dW_t^h}{W_t^h} = \frac{dw_t^h}{w_t^h} + \eta \frac{W_t^b}{W_t^h} dt, \tag{20}$$

where the second terms in both (19) and (20) 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. 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 $r(w_t, \lambda_t)$, household consumption wealth ratio $\hat{c}^h(w_t, \lambda_t)$ and lending $x^K(w_t, \lambda_t)$, banker consumption wealth ratio $\hat{c}^b(w_t)$ and lending $y^K(w_t, \lambda_t)$, such that*

- *Consumption, investment and portfolio choices are optimal.*
- *Capital good market clears*

$$W_t^b x_t^K + W_t^h y_t^K = p_t K_t. \tag{21}$$

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

$$W_t^b + W_t^h = p_t K_t. \tag{22}$$

- *Interbank market clears*

$$W_t^b x_t^f = 0 \tag{23}$$

- *Consumption goods market clears*

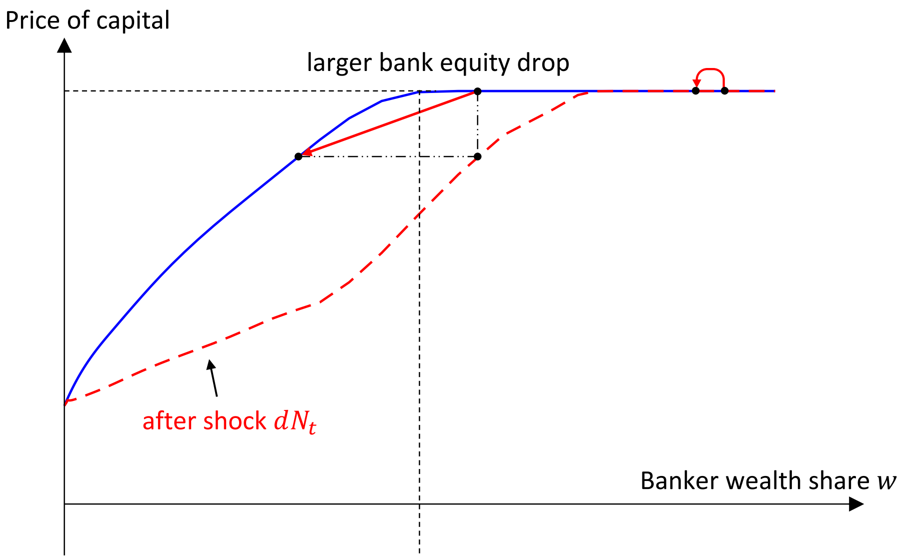
$$\hat{c}_t^b W_t^b + \hat{c}_t^h W_t^h = (\psi_t \bar{A} + (1-\psi_t)\underline{A})K_t - i_t K_t. \tag{24}$$

3.7 State-Dependence and Distress Dynamics

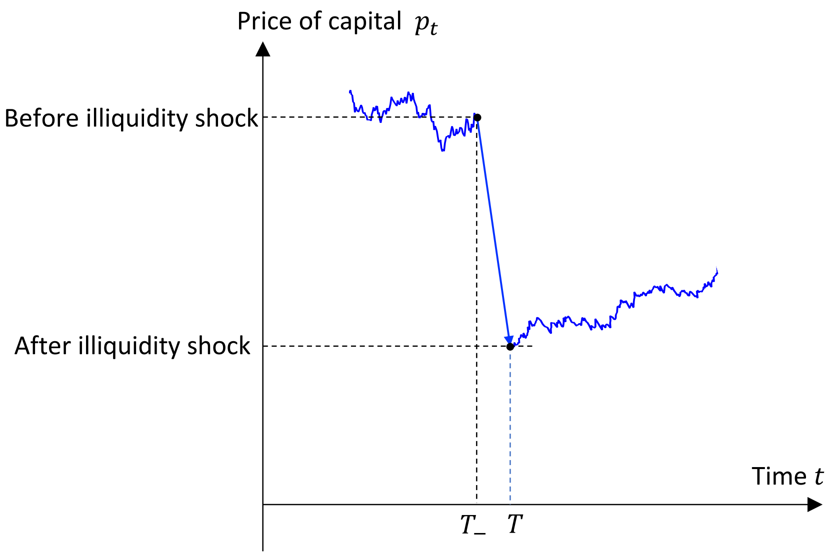
We solve the model and illustrate the nonlinear and state-dependent effects of a financial illiquidity event and the dynamics of the capital price around illiquidity shocks.

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



(a) Price of capital as a function of w_t, pre- and post- dN_t shock



(b) Path of the capital price around a bank run.

Figure 4: Illustration of Capital Price Dynamics

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.

Figure 4, Panel (b) illustrates the price path of capital in a case where one illiquidity shock occurs at time T and the wealth share is in the amplification region. We see that the pre-illiquidity shock price of capital follows a smooth path governed by the Brownian diffusion dZ_t. From $T-$ to T the price of capital jumps downwards. After T, the price of capital again follows a smooth path.

The rest of this section goes through this logic algebraically. For simplicity, we omit the t or $t-$ subscriptions in the following sections. For an individual bank, we can define the net funding withdrawal that has to be fulfilled by productive capital during a illiquidity as

$$\Delta x = (x^K + x^f - 1)^+ \tag{25}$$

To simplify the above expression, we prove that banks take leverage in equilibrium.

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

$$x^K \geq 1$$

Proof is provided in Appendix A.4. Because of Lemma 3, we have

$$x^d = \Delta x = x^K + x^f - 1 \tag{26}$$

From the banker optimization problem, we have the following first order conditions:

$$r^f - r^d = \lambda \frac{\alpha}{1 - x^K \kappa^p - \alpha \Delta x} \tag{27}$$

$$\mu^R + \frac{\bar{A}}{p} - r^f = (\sigma^K + \sigma^p)^2 x^K + \lambda \frac{\kappa^p}{1 - x^K \kappa^p - \alpha \Delta x} \tag{28}$$

where μ^R is the ex-dividend return of productive capital, with $\mu^R = \mu^p - \delta + \mu^K + \sigma^K \sigma^p - \phi(\mu^K)/p$. As clearly illustrated, if the total volatility $(\sigma^K + \sigma^p)$ increases, keeping the portfolio choice x^K the same, a banker requires a larger amount of risk compensation. Furthermore, if the expected intensity λ of the financial illiquidity shock rises, then the risk premium also rises. Finally, we observe that keeping everything else equal, a larger jump κ^p in the capital price leads to a higher risk premium.

Equation (28) also indicates how the belief λ affects bank leverage x^K. All else equal, we find that a higher λ results in a lower x^K. Further, in equilibrium, the lower x^K will result in less severe crisis (lower κ^p), which partly offsets the direct impact of λ on x^K.

We next derive the excess expected return on capital. We rewrite the banker budget dynamics as

$$\frac{dw^b}{w^b} = \left(r^f + x^K(\mu^R + \frac{\bar{A}}{p} - \lambda\kappa^p - r^f) - x^d(r^d - r^f) - \lambda\alpha x^d \right) dt - \hat{c}dt$$

$$+ x^K(\sigma^K + \sigma^p)dB_t - \kappa^b(dN_t - \lambda dt)$$

where the last component is the compensated Poisson process $dN_t - \lambda dt$, which is a Martingale. It is clear that the excess return of capital above the risk-free rate of bankers is $\mu^R + \bar{A}/p - \lambda\kappa^p - r^f$. Using (28), we can express this capital risk premium as

$$\mu^R + \frac{\bar{A}}{p} - \lambda\kappa^p - r^f = (\sigma^K + \sigma^p)^2 x^K + \lambda\kappa^p \frac{x^K\kappa^p + \alpha\Delta x}{1 - x^K\kappa^p - \alpha\Delta x} \tag{29}$$

which takes into account the downward impact of asset returns due to the realizations of illiquidity shocks. From this equation, we find that this premium is strictly positive in a Bayesian model.

3.8 Credit Spreads

We define a credit spread in this section that is needed in mapping the model to credit spread data. It is important to state at the outset that the defaultable bonds we price are in zero net supply. They are not issued by banks or households and do not affect the general equilibrium. 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. Define τ as the expected maturity of the bond. We assume that the bond matures based on the realizations of the 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.,

$$\pi\lambda_t + \lambda_t^\tau = 1/\tau \tag{30}$$

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 $m \cdot \kappa^p_{t-} + \hat{\kappa}_0$. The first term varies with economic conditions. It contains capital price drop κ^p_{t-}, 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

$$\hat{\kappa}_t = (m \cdot \kappa^p_{t-} + \hat{\kappa}_0) dN_t \tag{31}$$

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

$$\frac{\pi \lambda_t}{\pi \lambda_t + \lambda^\tau_t} = \tau \pi \lambda_t \tag{32}$$

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 as $v_t = v(w_t, \lambda_t)$, and the market value of the safe bond as $\bar{v}_t$. Then we define the credit spread as

$$\mathcal{S}_t(p_{t_0}) = \frac{1}{\tau} \log(1/v_t) - \frac{1}{\tau} \log(1/\bar{v}_t) \tag{33}$$

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. Details are provided in Appendix A.7.

4 Model Solution and Calibration

In this section, 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 (i.e., $\lambda_L < \lambda_H$).
2. Diagnostic (non-rational) Model: Agents form beliefs over the illiquidity state via diagnostic expectations, and belief varies over time (i.e., $\lambda_L < \lambda_H$).
3. Static belief Model: Agents' belief are constant at $\bar{\lambda}$.

Under diagnostic beliefs, we assume that while agents' beliefs are diagnostic, they think that their and all other agents' beliefs are Bayesian. In other words, the policy functions 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 states are different due to the underlying difference between diagnostic belief and the truth. 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 $\theta > 0$.

There are four parameters governing beliefs in the Bayesian model: λ_H, λ_L, $\lambda_{L \to H}$, and $\lambda_{H \to L}$. However, as we set λ_L near zero, the model has three parameters. 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)). We explain how these parameters are calibrated below.

We also present results from a version of the model where we turn off the belief mechanism ("Static Belief Model"). This model has only one parameter $\bar{\lambda}$ governing the crisis frequency process which is constant over time. As will be clear, this static belief model fails to match important features of the data. We present this static belief model primarily to explain the mechanisms of the belief models.

4.1 Solution Methodology

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 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.

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 model, and 8 parameters for the diagnostic model) 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

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

- 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.

The algorithm is time-consuming. We parallel the process and solve it using high-performance clusters.

4.2 Model Simulation

We simulate the model at a monthly frequency but analyze simulations at a yearly frequency to be consistent with the data. The procedure of simulation is as follows for each version of the model.

- From initial values $w_t = 0.1$ and $\lambda_t = \bar{\lambda}$, we draw shocks.
- 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.

- 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.

In order to map model outputs to data, we define the following events:

- A financial distress: in the year, there is at least one illiquidity shock $dN_t = 1$.
- A financial crisis: bank credit/GDP in a given year falls into the lowest 4% quantile of the distribution of bank credit/GDP (Fact 1).

In our model, large output declines in a year coincide with the financial illiquidity events. Therefore, financial crises under the above definition are a subset of financial illiquidity events.

4.3 Parameter Calibration and Estimation

Our calibration strategy is to identify each model parameter with a corresponding moment. 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 a given parameter.

A list of the calibrated parameters for the core model (not including the credit spread) 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^0), 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 $\alpha^0 \cdot \beta$ in Li (2019). Alternatively, we can interpret this liquidation loss as a funding premium. The value of $\alpha^0 = 0.05$ translates to a 10% premium for a illiquidity event that lasts 6 months.

Table 1: Calibrated Parameters for the Core Model

	Parameters	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^0	Distress illiquidity costs	0.05	Data

Table 2: 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

For the credit spread, we have the following calibration (summary in Table 2)

- 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 (2017).
- 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 (2017) 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.

Then we proceed to estimate other parameters, including λ_H, λ_L, $\lambda_{H\to L}$, $\lambda_{L\to H}$, $\bar{A}$, $\underline{A}$, σ^K, η, and θ. 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. After experimentation with the model, we find that the following moments to be particularly informative for each parameter:

1. Yearly frequency of bank equity crashes (fact 2): This moment maps to the frequency of financial illiquidity shocks and helps discipline λ_H. From Baron and Xiong (2017), the probability of a equity return below -30% is 3.2% at the quarterly frequency, which implies an annual frequency of about 12%, i.e. $1 - (1 - 3.2\%)^4$.

2. Credit spread changes during a crisis (fact 3). 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.

3. Half-life of credit spread recovery (fact 5). According to Krishnamurthy and Muir (2017), 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.

4. 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 $\underline{A}$.

5. Average output decline during a crisis (fact 4): We target -9.1% as explained in Section 2. This moment is most directly related to the productivity differential $\bar{A} - \underline{A}$.

6. 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.

7. 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. More details are provided in Appendix B. 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.

8. The diagnostic parameter θ is disciplined by fact 8. Conditioning on a crisis at year t, and looking at the 5 years before the crisis, Krishnamurthy and Muir (2017) 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).

The models have different sets of estimated parameters, as represented in Table 3. For each model, we only use moments that are related to the economics of that model. For the Bayesian model, we use moments 1–7. For the diagnostic model, we use moments 1–8. In this way, both models are exactly identified. For the static belief model, we use moments 1–7 (overidentified) to give it the best chance to match the data as the Bayesian model.

Table 3: Comparison of Model Parameters to be Estimated

This table lays out the set of estimated parameters in different models. "–" denotes not having the parameter, while "✓" denotes the opposite.

Parameters	Static Belief Model	Bayesian Belief Model	Diagnostic Belief Model
λ_H	✓	✓	✓
$\lambda_{L\to H}$	–	✓	✓
$\lambda_{H\to L}$	–	✓	✓
$\bar{A}+\underline{A}$	✓	✓	✓
$\bar{A}-\underline{A}$	✓	✓	✓
σ^K	✓	✓	✓
η	✓	✓	✓
θ	–	–	✓

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 4 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_0, but we opt to keep α_0 constant across all of the models to better illustrate the mechanisms underlying the models.

[TABLE 4 HERE]

Figure 6 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 60% in the model (that is, 0.6 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 0.8σs 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 6 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 beliefs model, spreads are slightly higher than normal, 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 5, 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. The jump to the λ process leads to a jump in beliefs, which leads to a fall in asset prices and hence a jump in w to the left. Additionally, on such a realization, λ_t is temporarily high so that more jumps are realized. This creates an increased mass at low w states.

In the diagnostic model (panels c and d), the realization of a jump leads to a large adjustment in w, relative to the Bayesian model, because agents shift from over-optimistic to over-pessimistic. As a result, more mass is shifted to low-w states.

[FIGURE 5 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 5 summarizes the model's performance in matching non-targeted moments.

Panel A reports that all of the models are able to fit 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. 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 are in the ballpark of the data. Note that this is true even in the static belief model, despite the fact that this model misses on the mean spike in credit spreads.

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. Note that the diagnostic model explicitly targets the mean pre-crisis spread in the calibration. 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. The second row of Panel C asks whether within crises, worse crises are preceded by even lower spreads. This is true in the data, as seen in the data column. The Bayesian belief model is unable to match this fact, while the diagnostic model gets closer, but still fails. We will also explain this failure later in the paper. 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, with the diagnostic model coming nearer to quantitatively matching the data moment.

[TABLE 5 HERE]

5.4 Mechanism 1: Frictional Intermediation and Leverage

Figure 7 graphs the histogram of 3-year GDP growth in crises. Focus first on the black dashed line corresponding to the static belief model. In a model with no financial amplification and only diffusion shocks to AK_t, output growth would be normally distributed. Thus, we conclude that the left-skewed output growth distribution in line with the data can be generated by the financial amplification mechanism.

[FIGURE 7 HERE]

Adding beliefs to the mix, as in the diagnostic and Bayesian models, produces more left-skew and places more mass in the left tail. The reason is that a illiquidity event not only triggers a jump down in w and a fall in GDP, but also an increase in λ. The amplification mechanism plus the change in belief drive an output distribution that is somewhat more left-skewed than the data.

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 5 evaluates the relationship between the jump in credit spreads in this model and the fall in GDP, which is in line with the data. 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).

We note that all of the variants of the model, even the static belief model, produces comovements between credit spreads, bank leverage, and output growth, in line with the data. We have also seen in Panel B of Table 5 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_0 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.

5.5 Mechanism 2: Beliefs and Leverage

We report in Table 5 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.[2] To understand the economics here, consider Figure 8. 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 risk/return tradeoff that drives their 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. When λ is low and leverage is high, if a 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. This endogenous relationship between illiquidity risk and vulnerability generates low credit spread before crises.

The diagnostic model generates a magnitude in line with the data, but note that the pre-crisis spread is the moment we have explicitly targeted to pin down the belief distortion parameter θ. The Bayesian model gets a spread that is below normal, and about half-way towards matching the data fact that spreads are about 0.34σs lower than normal in the pre-crisis period. Moreover, in the data, more severe crises are preceded by even lower spreads. The Bayesian model fails to capture this fact (see Panel C in Table 5). The reason can be understood by examining Figure 9. Here we plot the relationship between spreads right before an illiquidity shock triggers a crisis, and the GDP decline in the crisis. The blue line corresponds to the Bayesian belief model. Note that the relationship is non-monotonic, and

[2]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.

for low spreads, the slope is negative. There are two forces driving spreads pre-crisis. One is the belief mechanism, as reflected in Figure 8. The other is the w mechanism of the static belief model. This latter mechanism drives a negative relation between spreads and output growth; i.e., spreads are lower when w is higher, and hence vulnerability to an illiquidity shock is lower. These two forces compete against each other. In our calibrated model, the w mechanism dominates at low values of spread (the reason is that λ is as low as possible, weakening the belief mechanism).

[FIGURE 8 HERE]

[FIGURE 9 HERE]

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 8. 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. Returning to Figure 9, we see that by strengthening the leverage/lambda relationship, the diagnostic model generates a stronger relationship between the extent of low spreads and crisis output declines. As a result, the diagnostic model better fits the spread/output relation as reflected in Panel C of Table 5.

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 8. 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 also deliver this steepening. This is an important observation because it indicates that while the data patterns allow us to easily rule out the static belief models, the diagnostic vs. Bayesian belief models are not convincingly identified in the data.

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

Next, we consider the evidence that high bank credit and low spreads predict crises and not just precede crises. To see the difference, note that the former conditions in the event of a crisis. Table 5 Panel D presents the crisis prediction result. To replicate the predictability regressions in Krishnamurthy and Muir (2017), we define "high froth" as the past 5-year average of a dummy that indicates whether the credit spread is below its median value. Similarly, we define "high credit" as the past 5-year average of a dummy that indicates whether bank credit/GDP is above its median value.

In the dynamic belief models, we find the variables have the right signs, although the models are low in terms of magnitudes. The static belief model fails again, generating a sign that is the opposite of the data. See Table 5 Panel D.

To understand what drives the mechanism in the dynamic belief models, consider Figure 10. 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 panels (b) and (c). In these cases, we progressively condition on higher levels of credit and hence lower effective banker risk 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 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 5, 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]

Figure 11 plots the distribution of GDP growth over the next year conditional on different levels of credit in the diagnostic 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 top to bottom panel in the figure, the mass in the left tail rises. The improvement of the diagnostic model is again due to steepening the leverage/lambda relationship.

[FIGURE 11 HERE]

Figure 12 examines the predictive relation in a different way. In the figure, we plot the banker's wealth return conditional on different 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 (top panel) driven by a significant mass spread between -0.2 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 bottom 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 surprising result that in the Bayesian model, even if illiquidity events are less likely (low λ), crises are more likely.

[FIGURE 12 HERE]

Once one understands these mechanisms, it becomes clear that the result is more general than our model's specific mechanisms. High credit growth occurs when bankers are effectively less risk-averse. This leads to the relation between high credit and low expected returns/risk premia. Additionally, the less risk-averse banker is more willing to take risks and as a result, GDP outcomes have mass pushed out to the tails. In our model, time variation in beliefs regarding illiquidity risk generates variation in banker's willingness to take the risk.

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 Table 5 Panel D. As we will explain, this relation holds for the belief models in the parameterization we study, but need not hold generally. Figure 13 draws density plots of next-year GDP growth for the 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 Bayesian model, on the other hand, succeeds in this dimension.

[FIGURE 13 HERE]

The logic here 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: (1) higher λ means more illiquidity events and hence higher spreads; (2) worse crises mean lower loss-given default (via κ_t^p) and hence higher spreads. If we imagine shutting down effect (2), then we can understand the froth relation easily. When λ is low, the banker is effectively less risk-averse, and hence the economy is more subject to large GDP downturns. This force pushes more mass into crisis states but does not increase credit spreads ex-ante. Hence we arrive at the positive relation between froth and crises. Now, if we add back effect (2), the froth relation is weakened and can potentially flip the sign to resemble the static belief model. The reason is that more crises imply larger losses given default and hence higher ex-ante spreads. The sign of the froth relation depends quantitatively on the exact cyclicality of recoveries in default. We have calibrated our model to the history of recoveries on BAA bonds in the U.S., as reported by Moodys.

Table 6 illustrates this point for the extreme case where we set $\kappa_0 = 0$, and hence recovery has no fixed component. We now see that the froth relation has a negative sign and no longer matches the data. The high credit relation continues to match the data.

[TABLE 6 HERE]

Figure 14 draws the GDP distributions conditional on credit spreads for the diagnostic model. 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. As a result, the diagnostic model gets closer to matching the data in Table 5 Panel D.

[FIGURE 14 HERE]

5.8 Conditional Risk Premium

A negative result of our analysis is that the qualitative patterns of the crisis cycle do not allow one to distinguish between diagnostic and Bayesian belief models. We find 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 in the right direction, with the diagnostic model getting the quantitative results closer to the data. The key to this improvement is that the diagnostic model steepens the leverage/lambda relationship.

Are there other data that help identify diagnostic beliefs? Bordalo, Gennaioli and Shleifer (2018) argues for the importance of survey expectations to measure agent beliefs. Rationality requires that the frequency of financial crises be consistent with agents' beliefs, and measuring such beliefs may allow one to discipline a non-Bayesian component of beliefs. The difficulty with this approach is that financial crises are infrequent, and survey measures do not cover the breadth of history and countries necessary to investigate this possibility rigorously.

An approach that is more in line with that of this paper in terms of matching data from historical crises is from Baron and Xiong (2017). They observe that the expected returns on bank equity as well as the market can be negative conditional on bank-credit growth in the highest quartile of its distribution (see Table V and Figure III of their paper). The statistical strength of this result is weak relative to other results in the paper. It only holds at longer horizons and only for bank equity and not for broad equity returns. However, let us take this as a fact since such evidence is inconsistent with any model of rationality.[3] Figure 15 plots the associated figures from our model (the equivalent of their Figure III). Our model matches the general pattern of a negative slope, but our calibration does not generate returns that fall below zero. To be clear, this is not an impossibility result: parameters do exist such that diagnostic expectations generate negative expected returns (see Greenwood, Hanson and Jin (2019)). The result illustrated by the figure is in the context of our specific model and calibration that matches parameters, as indicated in Table 3.

[FIGURE 15 HERE]

6 Policy Impact under Diagnostic and Bayesian Beliefs

Does it matter for policy purposes whether we live in a world with diagnostic beliefs or one with Bayesian beliefs? Although our model is not suited for welfare analysis, it can shed light on the impact of a policy change on economic outcomes. In this section, 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.

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

[3]Note that we are discussing equity returns which have a positive β with respect to systematic risk factors. For other assets, such as fixed income or housing, it is possible that a collateralizability premium or convenience yield can generate a negative expected return.

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 16. 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 between 1% and 2% over the next two years. The bottom left panel plots the credit spread response. The recap leads 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 15%.

The key message from Figure 16 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 rational model (i.e. neither over-optimistic or pessimistic).

In our second experiment, we probe whether knowledge of the exact divergence from rational 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 17. We again find that the impulse response gaps are quite similar across these two cases. The key to this result is that the initial condition in terms of observables is the same in this experiment. Finally, note that the plots in Figure 17 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 18 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 than in Figure 17. 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 rational 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 our hands by forcing the diagnostic model's parameters and states to match the *same observables*. In Figure 19, 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 7 HERE

A comparison of the simulation experiments is shown in Table 7.

7 Conclusion

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.

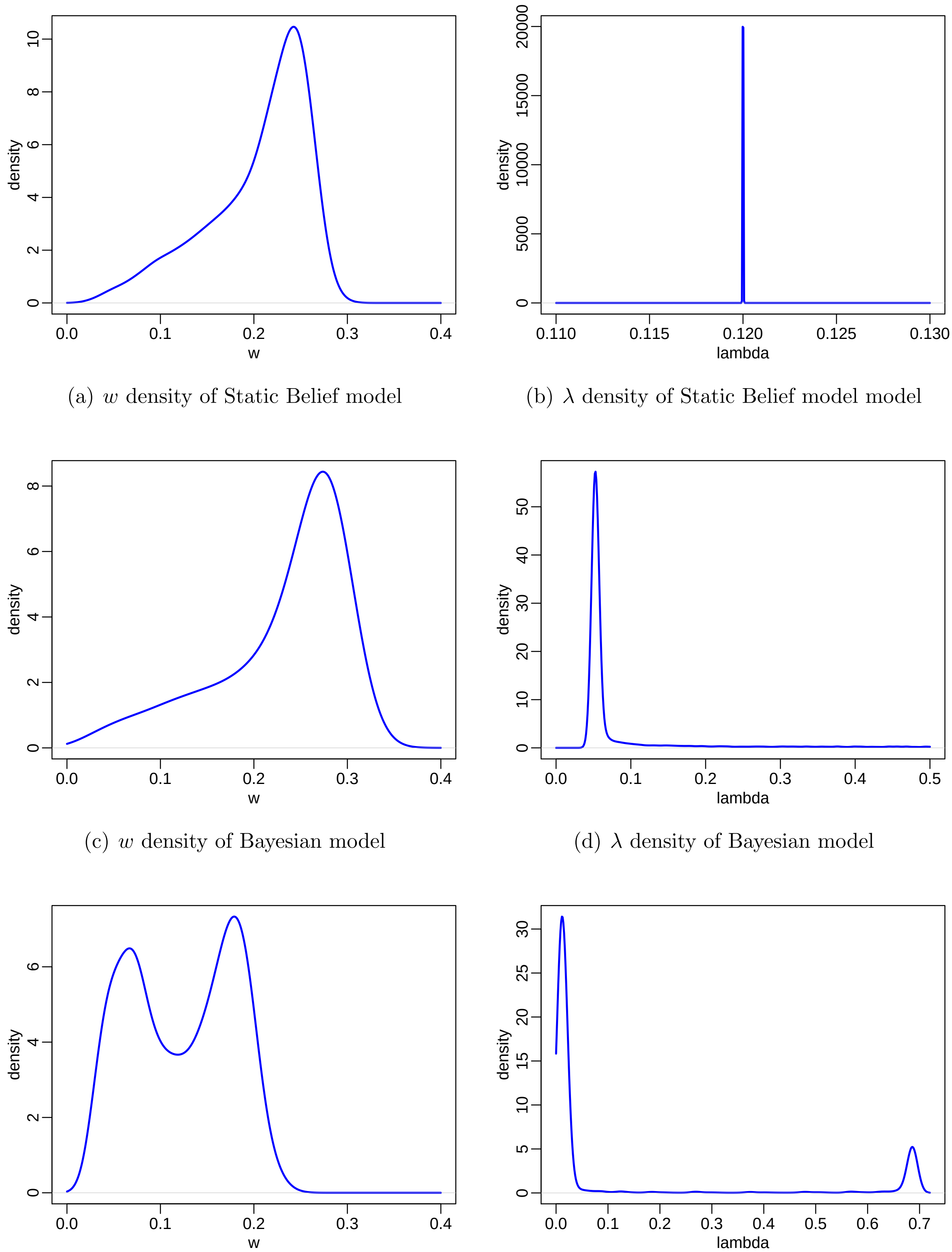
We offer three main conclusions that may guide future research:

1. On the positive side, a minimal model with two state variables, one that governs financial frictions and one that governs beliefs, can match the crisis cycle facts. Both the Bayesian and diagnostic models are driven by two state variables, w, and λ. The dynamics of these variables drive the model's fit on the dimensions on which we evaluate. Our analysis shows that these variables can have the "right" dynamics under both Bayesian and diagnostic belief updating.

2. On the negative side, based on our analysis, we are not compelled that the data identifies a non-rational component of beliefs. Apparently, even the Bayesian model gets the broad patterns correct; the non-rational component just adds a kick to get these patterns closer to the data. Considering that our baseline model is quite minimalist, it seems premature to put too much weight on the success from the extra kick. Hence our conclusion is that the data does pose an identification challenge for sorting among models of belief formation.

3. Despite this identification problem, our analysis suggests that for many policy experiments it may not matter whether beliefs are Bayesian or diagnostic. The response of the economy to a given shock, such as a fall in bank capital, 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, impulse response of the models we study are quite similar.

Figures and Tables



(a) w density of Static Belief model

(b) λ density of Static Belief model 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 5: Stationary Distribution of State Variables in Each Model

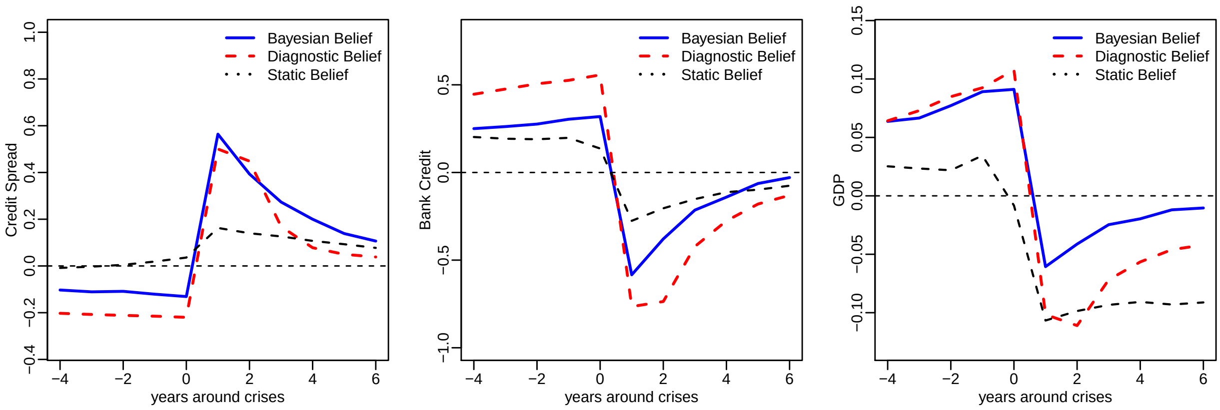


Figure 6: Dynamics of Different Models Around Crises. Credit spread and bank credit are measured as standard-deviations from the mean value. For example, credit spread raising to 0.2 means that it is larger than the value at year 0 by 0.2σs. GDP is measured in terms of percentage deviation from the long-run mean value.

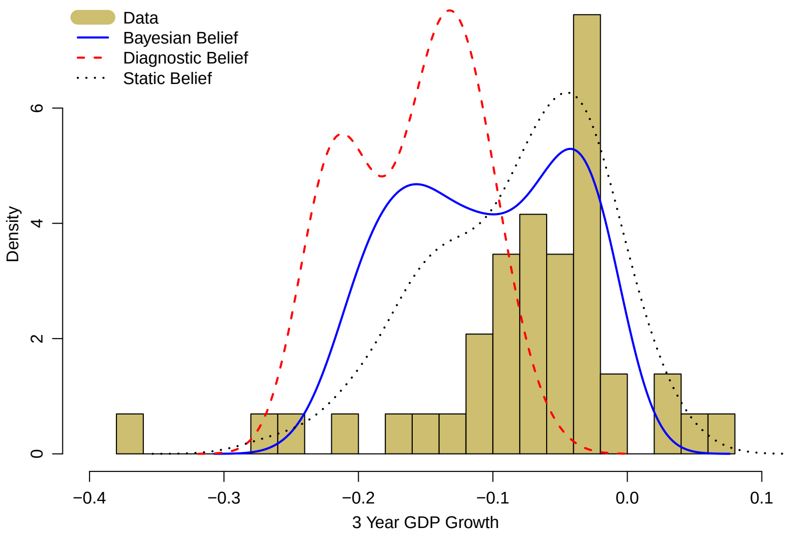


Figure 7: 3-Year GDP Growth: Model versus Data

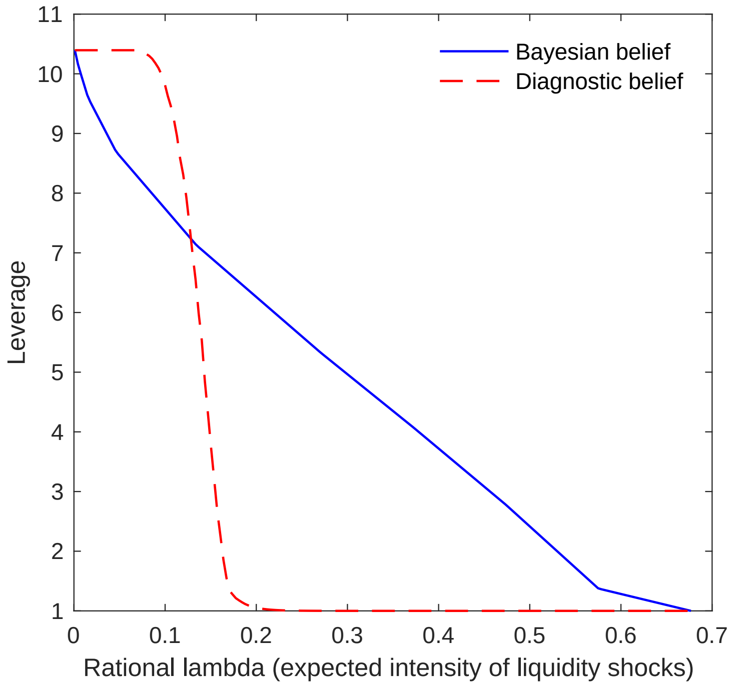


Figure 8: Expected Distress Frequency and Bank Leverage. This figure plots the leverage of banks as a function of the rational belief λ, given the same state variable w. We simulate the diagnostic model to derive the model-implied relationship between rational λ and the diagnostic belief λ^{θ}, and show the corresponding leverage choice.

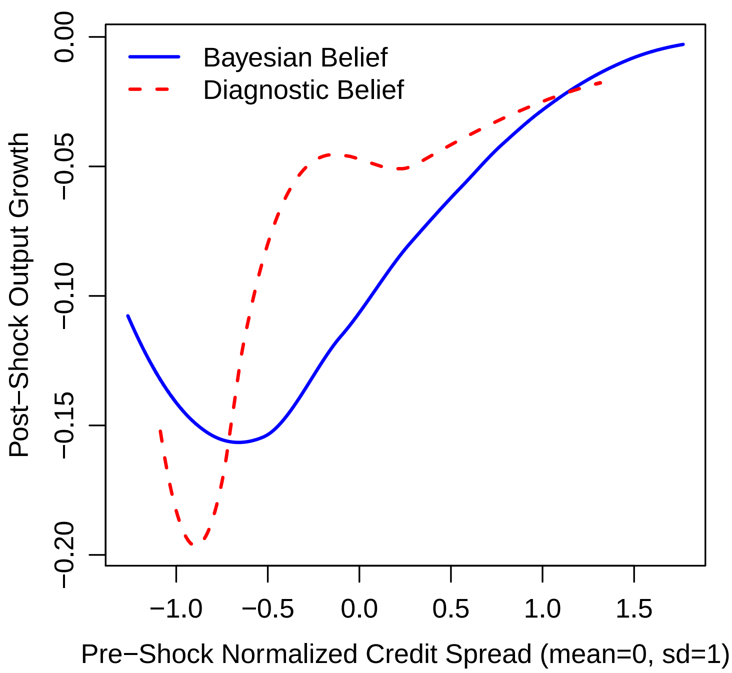
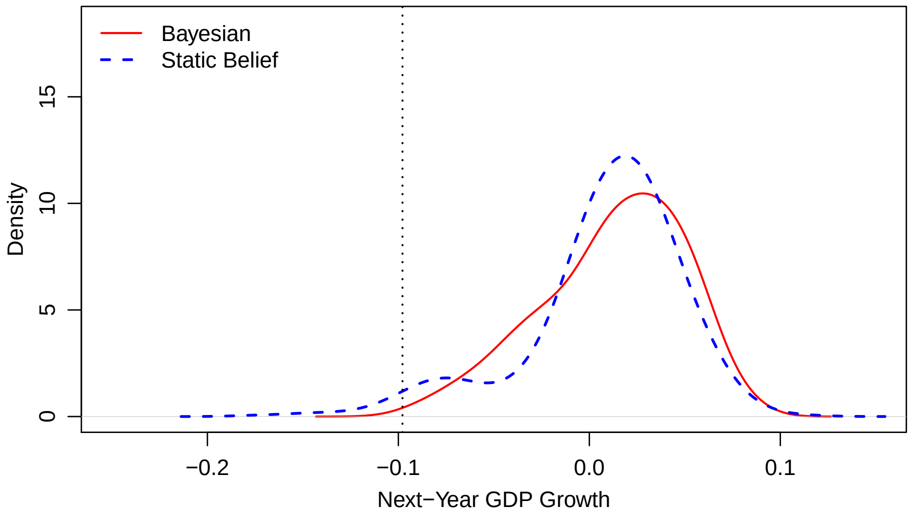
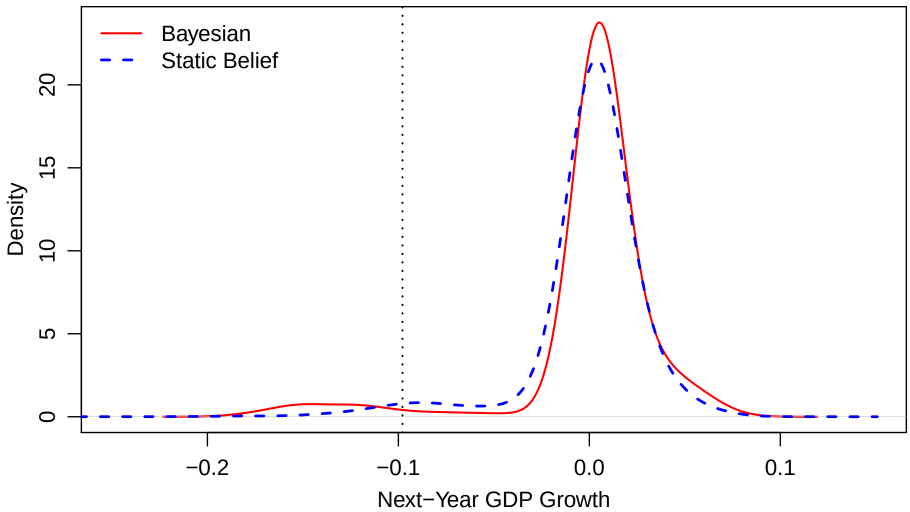


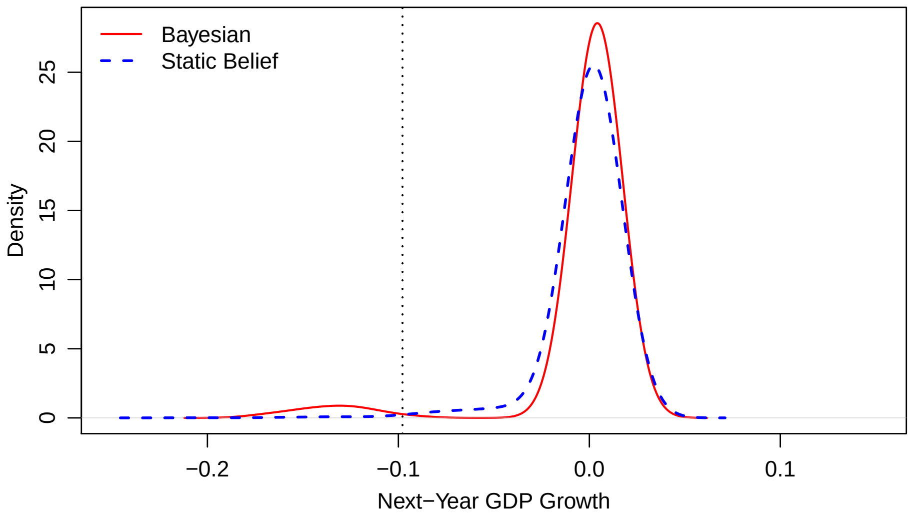
Figure 9: Average Output Growth in Year after a dN_t Shock vs Pre-Shock Credit Spread



(a) Conditioning on Low Bank Credit/GDP

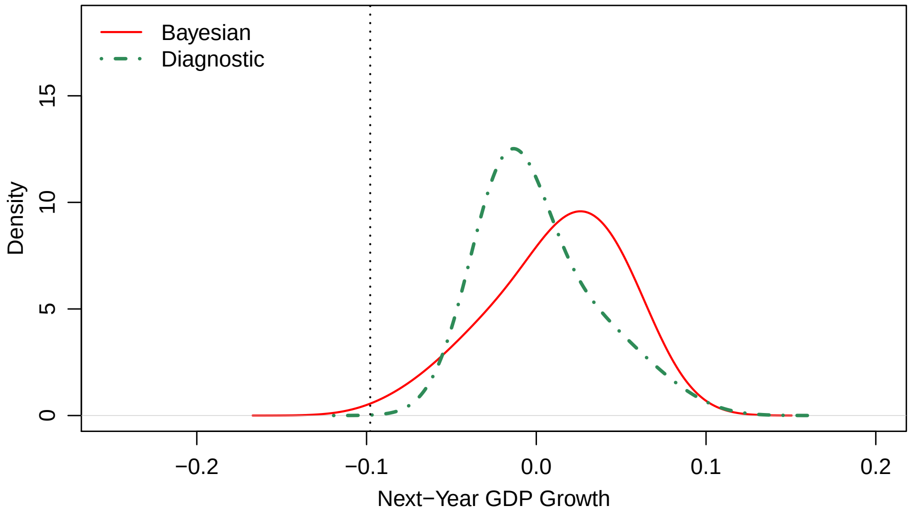


(b) Conditioning on Medium Bank Credit/GDP

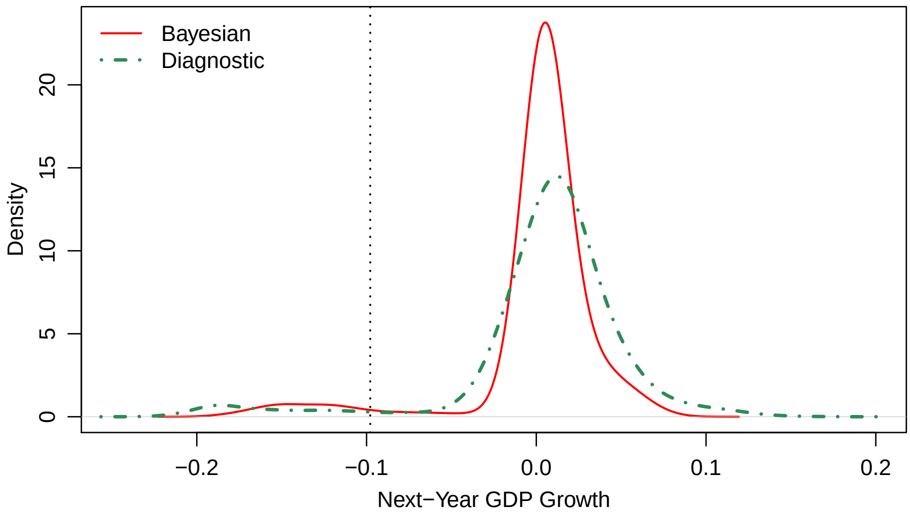


(c) Conditioning on High Bank Credit/GDP

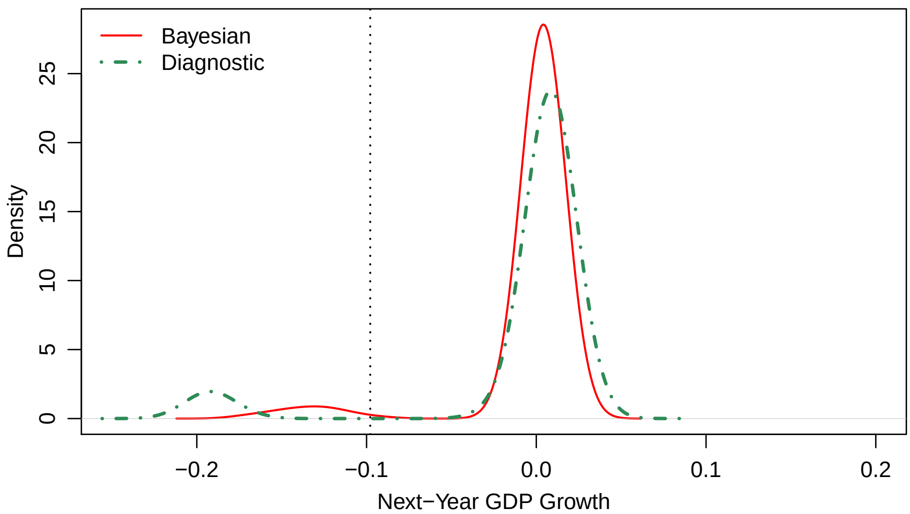
Figure 10: Density of Next-Year GDP Growth in Static Belief and Bayesian Models Conditional on Bank Credit/GDP. Cutoffs are 30% quantile and 90% quantile of bank credit/GDP.



(a) Conditioning on Low Bank Credit/GDP

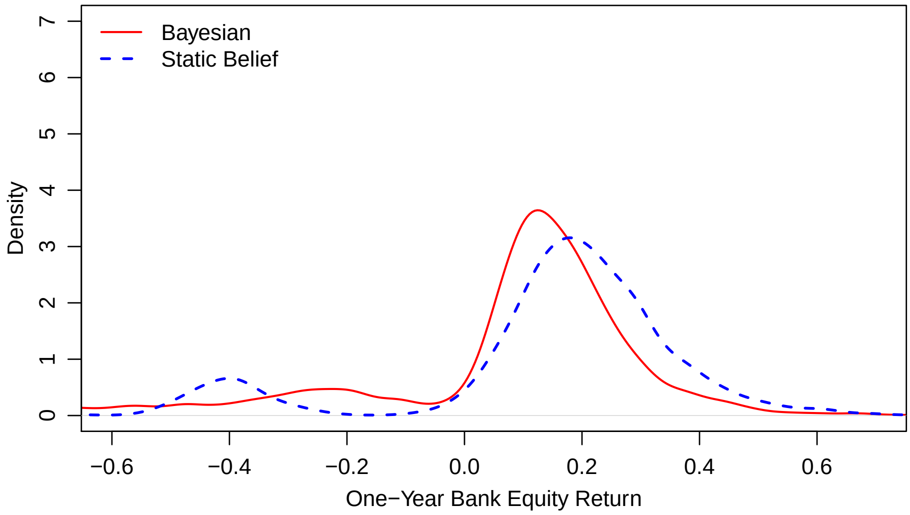


(b) Conditioning on Medium Bank Credit/GDP

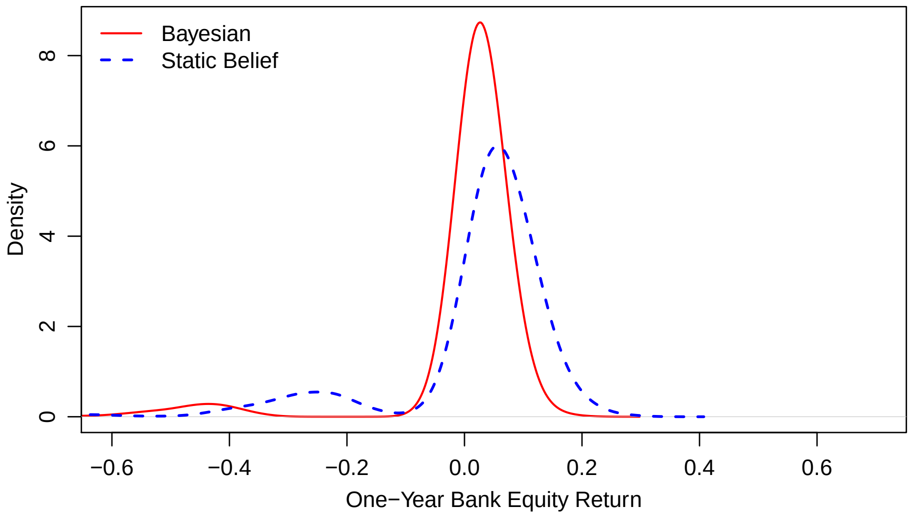


(c) Conditioning on High Bank Credit/GDP

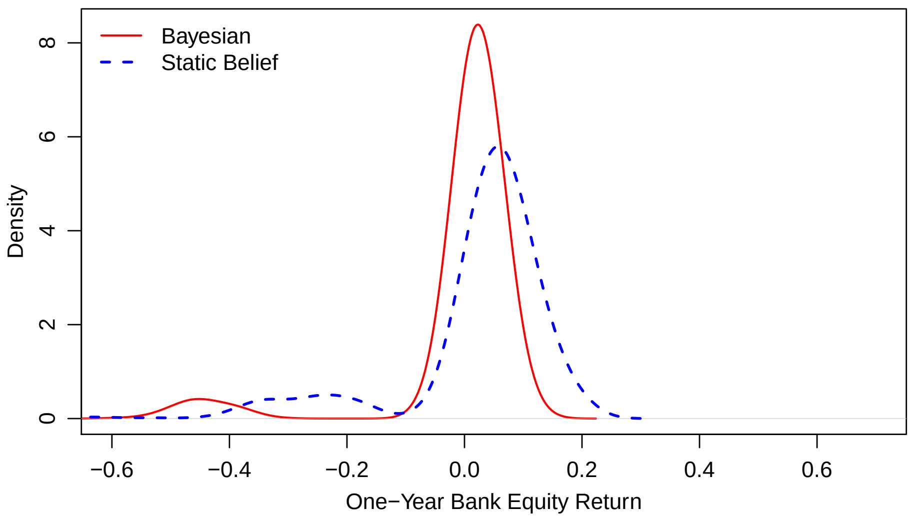
Figure 11: Density of Next-Year GDP Growth in Bayesian and Diagnostic Models Conditional on Bank Credit/GDP. Cutoffs are 30% quantile and 90% quantile of bank credit/GDP.



(a) Conditioning on Low Bank Credit

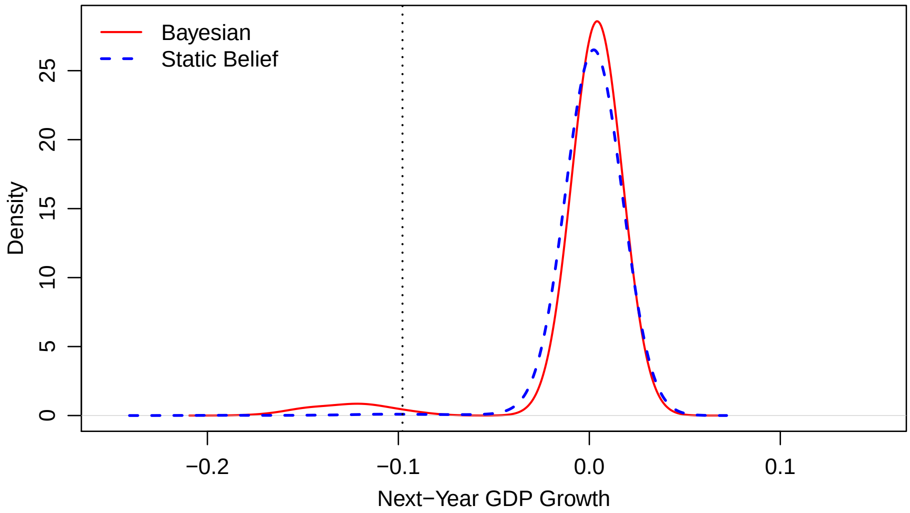


(b) Conditioning on Medium Bank Credit

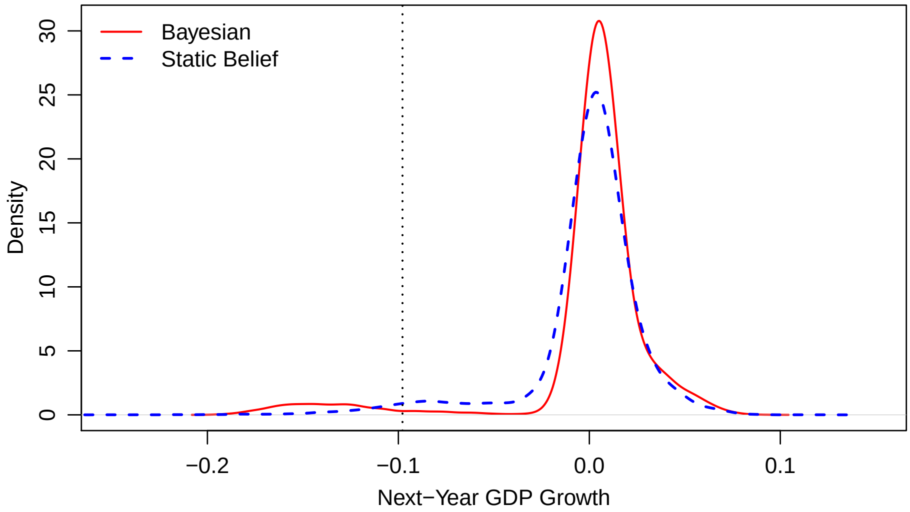


(c) Conditioning on High Bank Credit

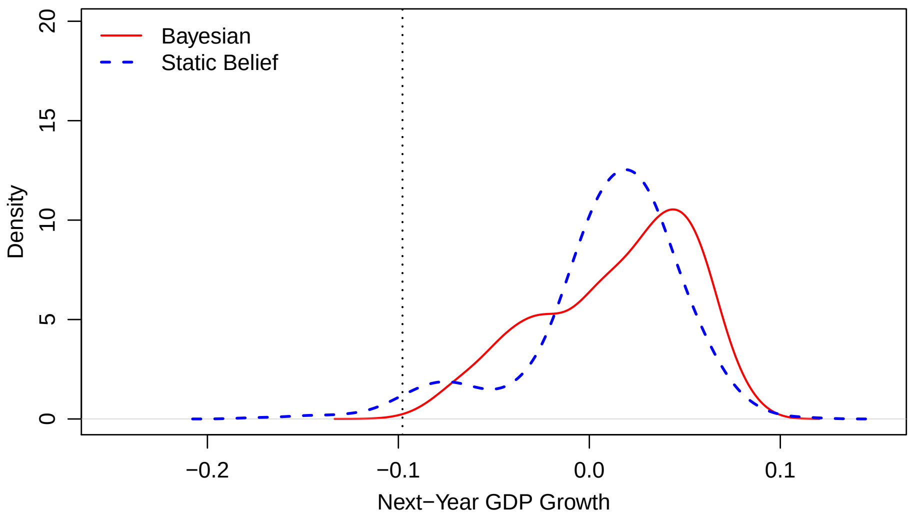
Figure 12: Density of Bank Equity Return in Three Models Conditional on Bank Credit/GDP. Cutoffs are 30% quantile and 90% quantile of bank credit/GDP.



(a) Conditioning on Low Credit Spread

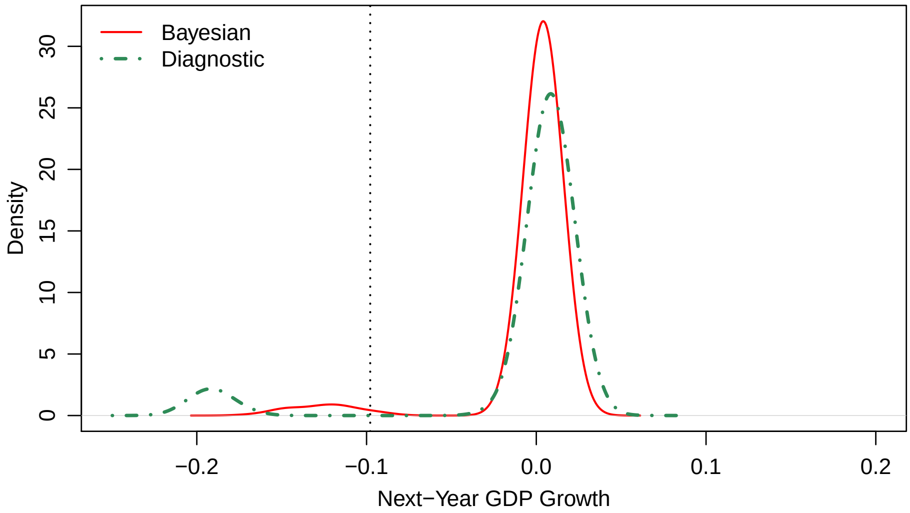


(b) Conditioning on Medium Credit Spread

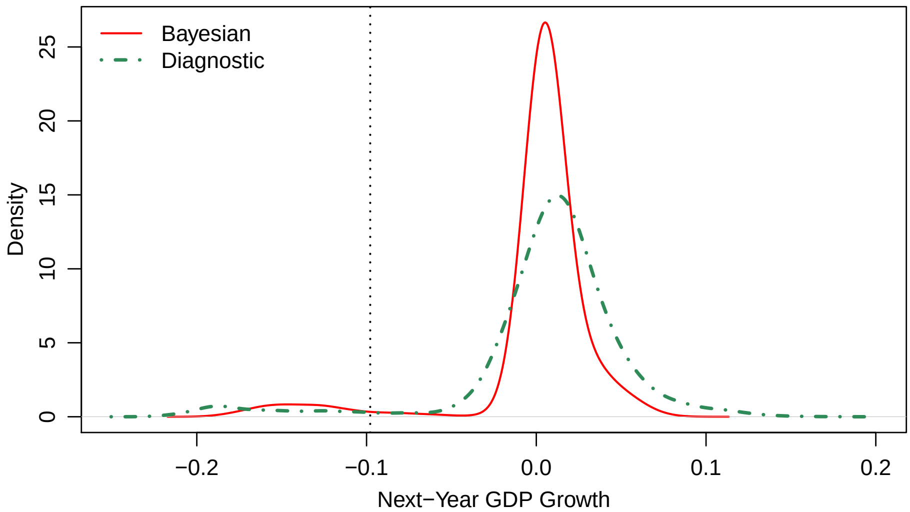


(c) Conditioning on High Credit Spread

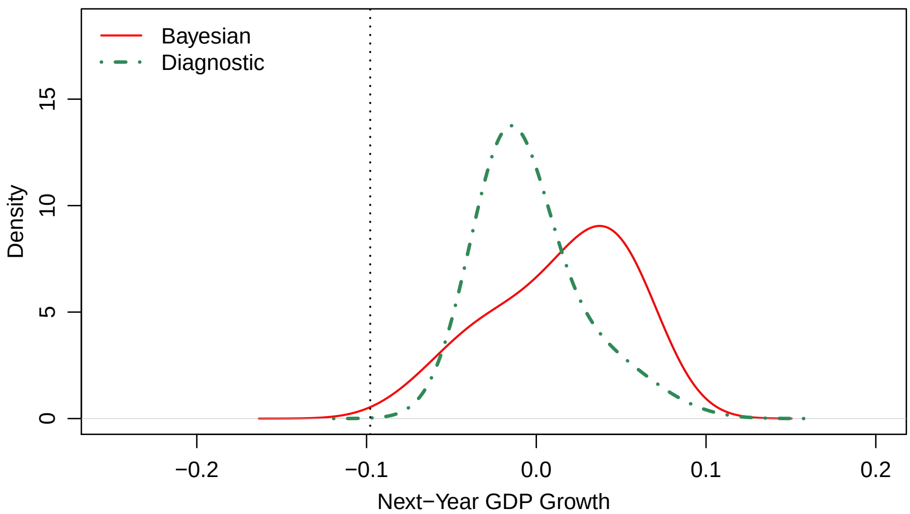
Figure 13: Density of Next-Year GDP Growth in Three Models Conditional on Credit Spread. Cutoffs are 30% quantile and 90% quantile of credit spread.



(a) Conditioning on Low Credit Spread

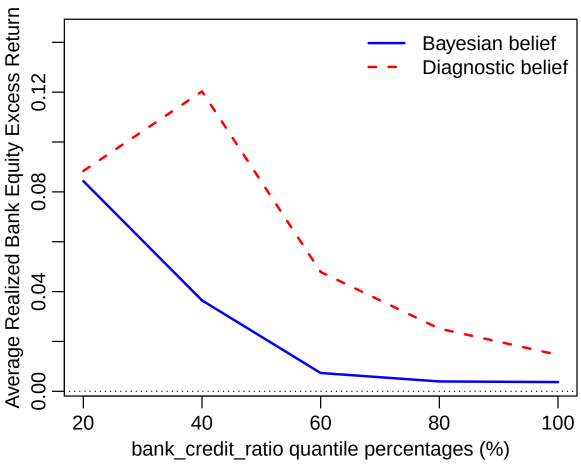


(b) Conditioning on Medium Credit Spread

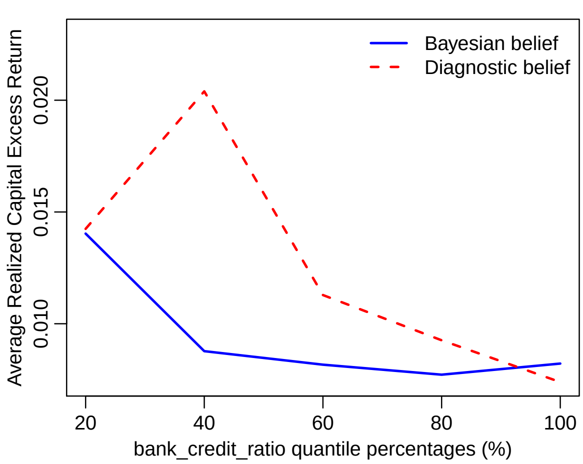


(c) Conditioning on High Credit Spread

Figure 14: 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.



(a) Bank Credit and Bank Equity Excess Return



(b) Bank Credit and Capital Excess Return

Figure 15: Bank Credit and Risk Premium

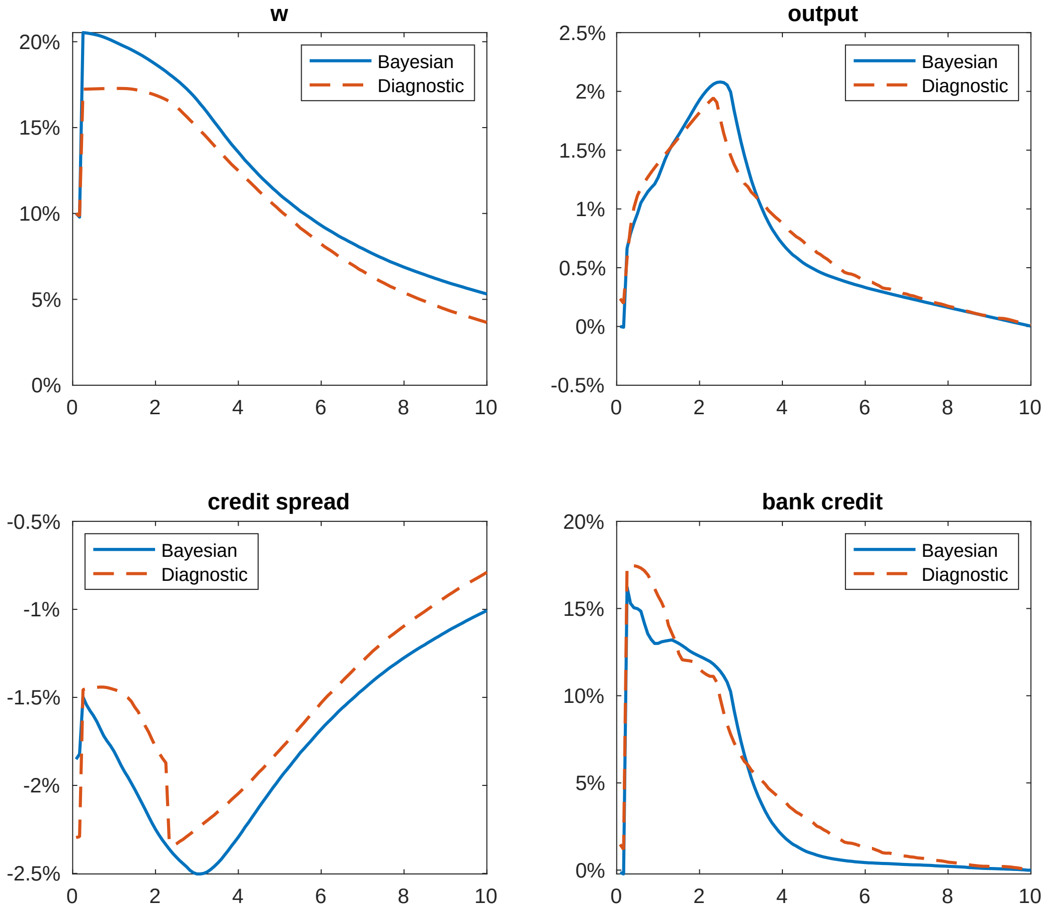


Figure 16: Impulse Responses of Experiment 1. In this figure, we show the impulse responses to a recapitalization policy at $t = 0$ that increases banker wealth share w by 10%, in order to "lean against the wind" and avoid future losses in a liquidity distress. The starting state is a "boom state", solved 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 4. 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.

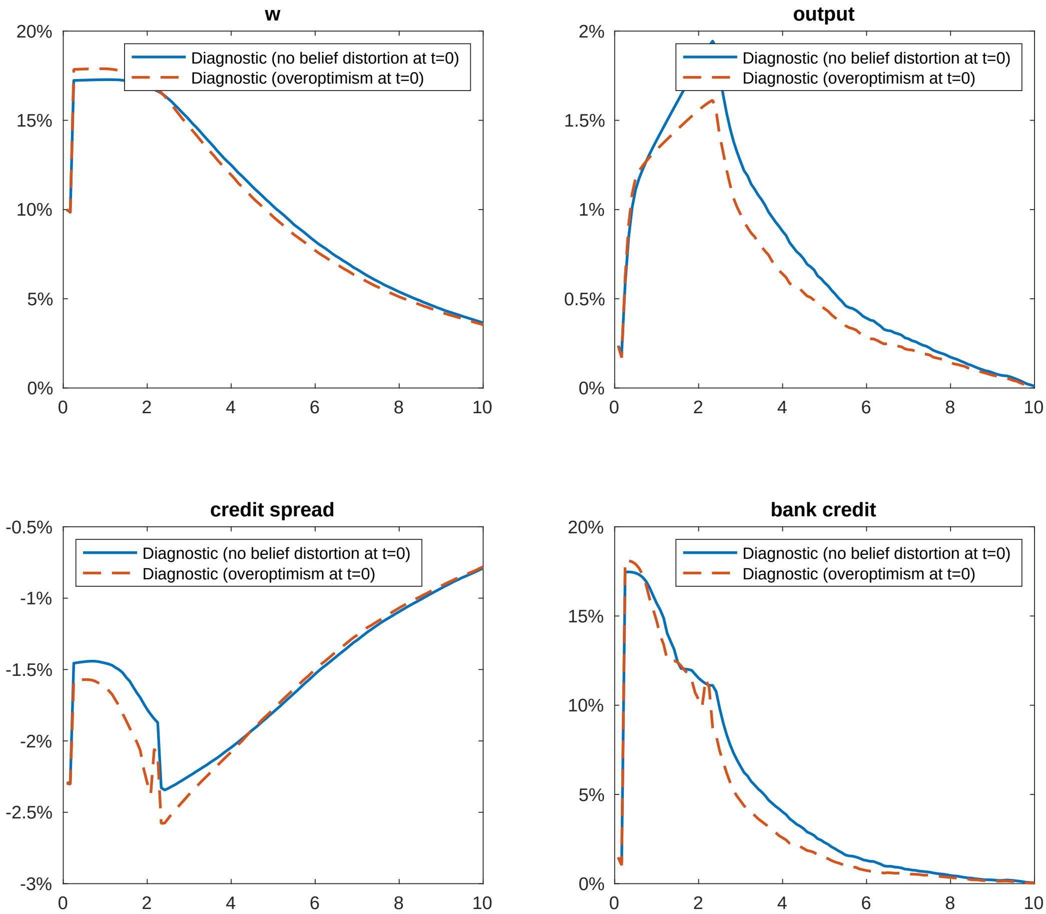


Figure 17: Impulse Responses of Experiment 2. Experiment 2 is identical to experiment 1 except for $\lambda_0^\theta > \lambda_0$ (overoptimism) in simulating the diagnostic model. Specifically, both Bayesian and diagnostic models have the same credit spread and bank leverage at $t = 0$, but the true frequency of distress in the diagnostic model is higher than the believed frequency. More descriptions are provided in Table 7 and footnotes of Figure 16.

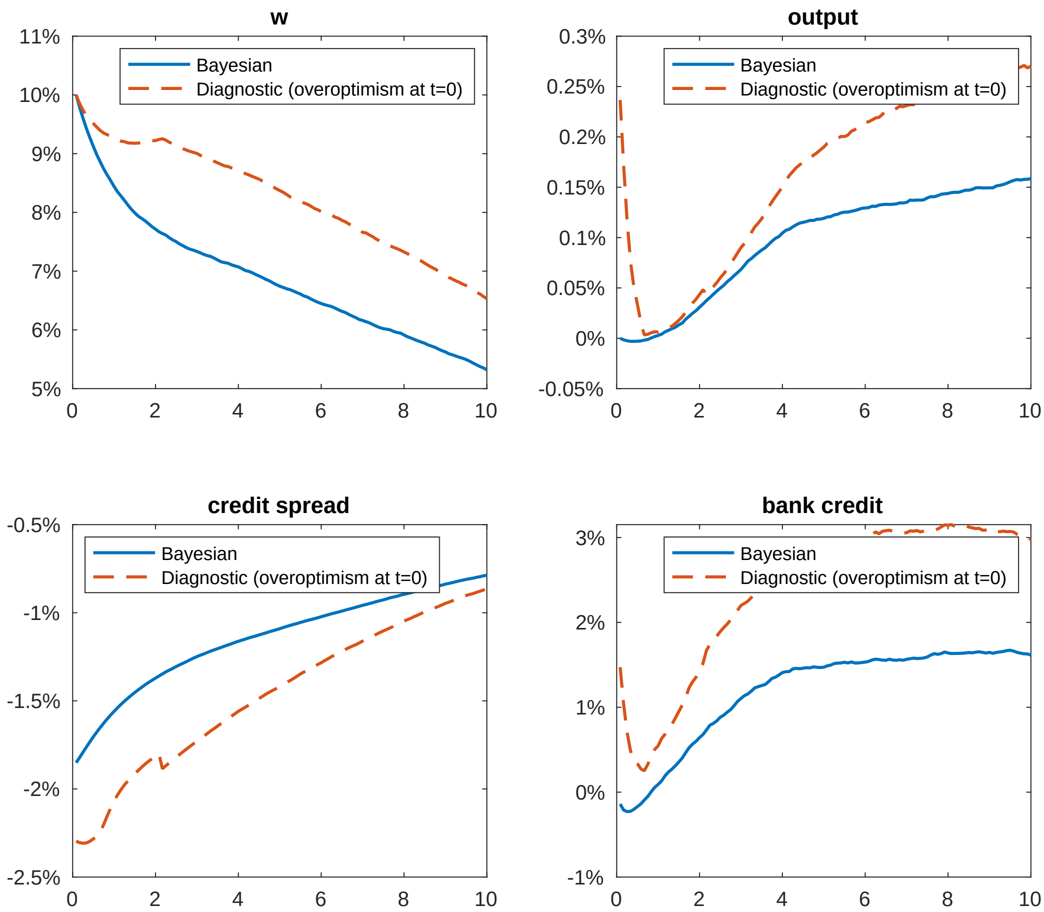


Figure 18: Expected Impulse Responses of Experiment 2. In this figure, we illustrate the expected impact of the recapitalization policy in experiment 2, by simulating dN_t according to the underlying process instead of setting $dN_t = 1$ at $t = 1/12$. More descriptions of experiment 2 are provided in Table 7 and footnotes of Figure 16 and 17.

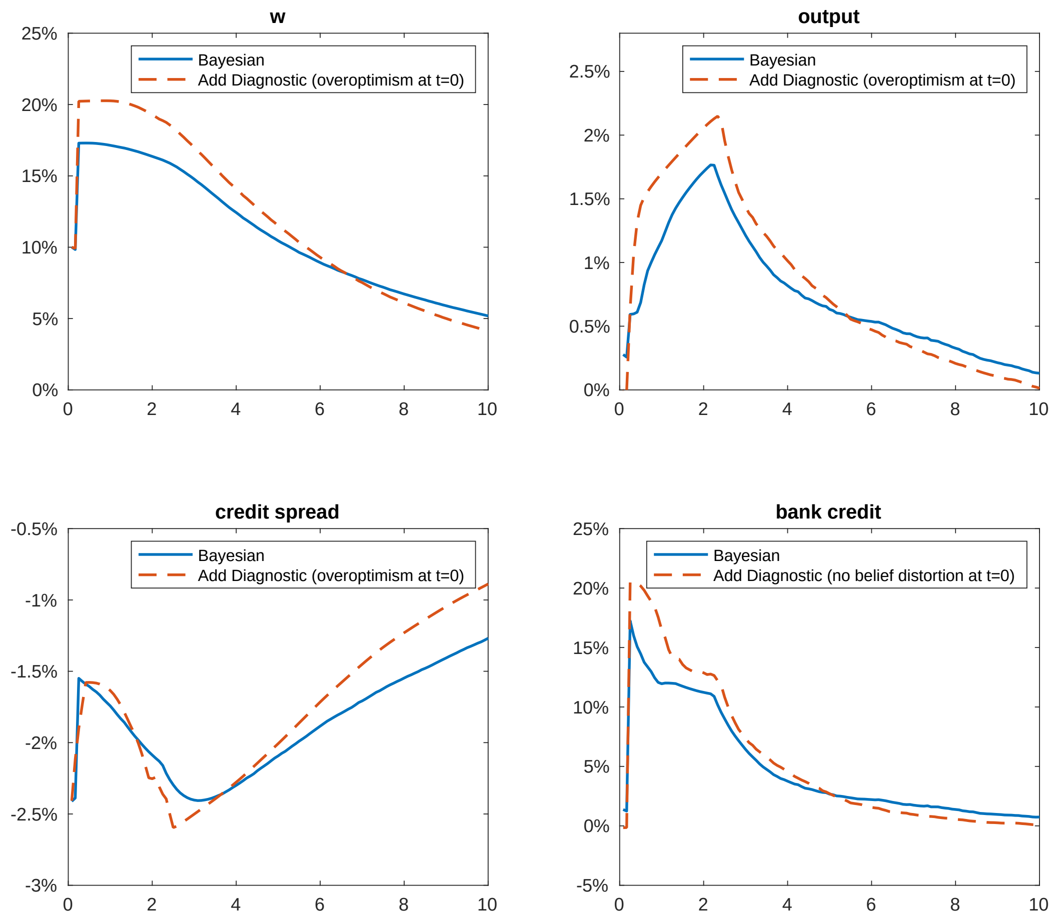


Figure 19: Expected Impulse Responses of Experiment 3. Experiment 3 shows the typical evaluation of the diagnostic belief in the literature, by keeping the Bayesian component fixed (same parameters and same starting states) but introducing the diagnostic element. In experiment 3, unlike experiment 1 and 2, the $t = 0$ observables including credit spread and bank leverage could be different across the Bayesian and diagnostic model (refer to Table 7). 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 7). 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.

Table 4: Comparison of Calibrated Model Moments

	Data	Static Belief	Bayesian	Diagnostic
1. Frequency of financial distress	13%	13%	12%	13%
2. Avg credit spread change in crises	70%	15%	63%	67%
3. Half-life of credit spread recovery (years)	2.5	2.3	3.2	2.6
4. Investment/capital ratio	14%	14%	18%	14%
5. Avg 3-year output drop in crises	-9%	-8%	-8%	-11%
6. Output growth volatility	4%	3%	4%	5%
7. Average bank leverage	5.0	5.2	4.8	5.2
8. Pre-crisis credit spread	-34%			-32%

Table 5: Model Simulation and Data: Non-targeted Moments

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$_t$	−6.19		−4.07		−3.94		−7.46 (0.16)	
$(\frac{\text{bank credit}}{\text{GDP}})_t$*crisis$_t$		−1.40		−2.61		−3.72		−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
$(\frac{\text{bank credit}}{\text{GDP}})_t$	−0.02	−0.01	−0.01	−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 Table 3 of Baron and Xiong (2017). 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)	(5)	(6)	(7)
pre-crisis indicator	0.22	−0.14		−0.32		−0.34 (0.15)	
mild crisis			−0.16		−0.31		−0.27 (0.15)
severe crisis			−0.09		−0.33		−0.45 (0.18)
Observations						634	634

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

Panel D: Predicting Crises

	Dependent variable: crisis$_{t+1 \text{ to } t+5}$							
	Static Belief		Bayesian		Diagnostic		Data	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HighFroth $_t$	−0.76		0.06		0.38		1.76 (0.91)	
HighCredit$_t$		−0.90		0.09		0.38		0.55 (0.46)
Observations							528	549

Note: HighFroth measures if spreads have been abnormally low in the last 5 years. HighCredit measures if credit growth has been abnormally high in the last 5 years.

Table 6: Predicting Crises when $\hat{\kappa_0} = 0$ in the Bayesian Model

	Dependent variable: $crisis_{t+1\ to\ t+5}$			
	Model simulation		Data	
	(1)	(2)	(3)	(4)
Froth_t	-0.14		1.76	
			(0.91)	
High Credit_t		0.14		0.55
				(0.46)
Observations			528	549

Note: HighFroth measures if spreads have been abnormally low in the last 5 years. HighCredit measures if credit growth has been abnormally high in the last 5 years.

Table 7: Policy Experiments

	Experiment 1	**Experiment 2**	**Experiment 3**
	Same Observables	Same Parameters, Overoptimism	Same Parameters and States
Initial States of the Bayesian Model ($\theta = 0$)			
Bank Leverage_0	4.2	same as experiment 1	same as experiment 1
Credit Spread_0	0.95		
w_0	0.239		
λ_0	0.112		
λ_0^θ	0.112		
Initial States of the Diagnostic Model ($\theta = 1.38$)			
Bank Leverage_0	4.2	4.2	4.2
Credit Spread_0	0.95	0.95	0.85
w_0	0.234	0.234	0.239
λ_0	0.117	**0.168**	0.112
λ_0^θ	0.117	0.117	**0.034**

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.

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. 2019. "Real Credit Cycles." Harvard University.

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.

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.

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.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos. 2017. "Deposit competition and financial fragility: Evidence from the us banking sector." *American Economic Review*, 107(1): 169–216.

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."

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.

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.

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.

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 Tyler Muir. 2017. "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.

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

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

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. 2009*a*. "The Aftermath of Financial Crises." *American Economic Review*, 99(2): 466–72.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009*b*. *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

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

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

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

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

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

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

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

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

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

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

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

which results in

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

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 $\lambda_{L\to H}\Delta t$. 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

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

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

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

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

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

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

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

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

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

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

Therefore, we have

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

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

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

which can be simplified as

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

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

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

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

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

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 (8). Our task now is to prove that the discrete-time diagnostic belief process converges to a continuous-time process as in (11). By definition, the diagnostic belief at period k is

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

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

with

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

which implies

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

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

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

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

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 Wealth Dynamics

To solve the model, we start with deriving the wealth dynamics of households and bankers. In order to simplify notations, we omit the subscripts t and $t-$.

First, from (24) and (22), we get the following equation that links consumption, production and investment:

$$w\dot{c}^b + (1-w)\dot{c}^h = \frac{\psi A^H + (1-\psi)A^L - i}{p}. \tag{41}$$

Second, from (21) and (22), we get the following portfolio equation on capital

$$x^K w + y^K(1-w) = 1. \tag{42}$$

Third, we can rewrite (17) as a function of state variables and portfolio choices, i.e.

$$\psi = \frac{x^K w}{x^K w + y^K(1-w)} = x^K w, \tag{43}$$

where the first equality is by definition and the second equality is by (42).

To proceed, we need to express the evolution dynamics of state variable w. Define

$$\mu^R = \mu^p - \delta + \mu^K + \sigma^K\sigma^p - \frac{\phi(\mu^K)}{p} \tag{44}$$

Return on banker wealth is

$$\begin{aligned}\frac{dw^b}{w^b_{t-}} &\triangleq \mu^b dt + \sigma^b dB - \kappa^b_{t-}dN \\ &= \left(r^d + x^K(\mu^R + \frac{A^H}{p} - r^d) + x^f(r^f - r^d) - \rho\right)dt + x^K(\sigma^K + \sigma^p)dB - \kappa^b_{t-}dN\end{aligned} \tag{45}$$

The jump component is

$$x^K\kappa^p + \alpha\Delta x \tag{46}$$

where

$$\Delta x = (x^K + x^f - 1)^+ \tag{47}$$

The return on household wealth is

$$\begin{aligned}\frac{dw^h}{w^h_{t-}} &\triangleq \mu^h dt + \sigma^h dB - \kappa^h_{t-}dN \\ &= \left(r^d + y^K(\mu^R + \frac{A^L}{p} - r^d) - \rho\right)dt + y^K(\sigma^K + \sigma^p)dB - \kappa^h_{t-}dN\end{aligned} \tag{48}$$

where

$$\kappa^h = y^K\kappa^p - \kappa^{fs}$$

A.4 Proof of Lemma 3

According to the equilibrium market clearing condition for the interbank market, we have

$$x^f = 0$$

in equilibrium. As a result, in equilibrium,

$$\Delta x = (x^K - 1)^+$$

Suppose that in equilibrium, $x^K < 1$. This implies that $\Delta x = 0$. Then we can easily derive the first order condition for households and bankers holding capital as

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

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

which together imply that

$$\frac{\bar{A} - \underline{A}}{p} = \left((\sigma^K + \sigma^p)^2 + \frac{\lambda(\kappa^p)^2}{(1 - x^K\kappa^p)(1 - y^K\kappa^p)} \right) (x^K - y^K) \tag{49}$$

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, due to the budget constraint

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

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

$$y^K > x^K$$

which implies that the right-hand side of (49) should be negative. This is a contradiction since the left-hand side of (49) 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.

A.5 Equilibrium Solutions

Optimal investment rate is

$$\mu^{K^*} = \frac{p-1}{\chi} + \delta \tag{50}$$

The resulting optimal investment is

$$i(p) = \phi(\mu^{K^*}) = \frac{(p-1)^2}{2\chi} + \frac{p-1}{\chi} + \delta \tag{51}$$

Then we can apply Ito's lemma on the definition of wealth share in (12) and get the dynamics of w as

$$\begin{aligned}\frac{dw}{w} &\triangleq \mu^w dt + \sigma^w dB - \kappa^w_{t-} dN \\ &= (1-w)\left(\mu^b - \mu^h + (\sigma^h)^2 - \sigma^b\sigma^h - w(\sigma^b - \sigma^h)^2 - \eta\right) dt \\ &\quad + (1-w)(\sigma^b - \sigma^h)dB - (1-w_{t-})\frac{1 - \frac{1-\kappa^b_{t-}}{1-\kappa^h_{t-}}}{1 + w_{t-}(\frac{1-\kappa^b_{t-}}{1-\kappa^h_{t-}} - 1)} dN.\end{aligned} \tag{52}$$

With dynamics of the state variable w, we apply Ito's lemma on price function $p(w)$ to get

$$\begin{cases} \mu^p = p_w w \mu^w + \frac{1}{2} p_{ww}(w\sigma^w)^2 + p_\lambda \mu^\lambda(\lambda) \\ \sigma^p = p_w w(1-w)(\sigma^b - \sigma^h) \\ \kappa^p_{t-} = 1 - p(w_{t-}\frac{1-\kappa^b_{t-}}{1-\kappa^h_{t-} - w_{t-}(\kappa^b_{t-} - \kappa^h_{t-})}, \lambda_t))/p(w_{t-}, \lambda_{t-}). \end{cases} \tag{53}$$

To fully characterize the economy, we also need to know the dynamics of aggregate capital quantity K, although it is not a state variable since everything else is scalable with respect to K. Denote the Ito process for K as

$$\frac{dK}{K_{t-}} = \mu^{K*} dt - \delta dt + \sigma^K dB, \tag{54}$$

With (46), (45), (48), and (53), we get a system of equations for other jumps:

$$\begin{cases} \kappa^b = x^K \kappa^p + \alpha \Delta x \\ \kappa^h = y^K \kappa^p - \kappa^{fs} \\ \kappa^{fs} = \alpha \Delta x \frac{w}{1-w} \\ \kappa^p = 1 - p(w\frac{1-\kappa^b}{1-\kappa^h - w(\kappa^b - \kappa^h)}, \lambda + \kappa^\lambda(\lambda))/p(w, \lambda) \end{cases} \tag{55}$$

From (45), (48), and (53), we get the relation between capital price volatility and volatility of the banker's return on wealth as follows:

$$\begin{cases} \sigma^p = p_w w(1-w)(\sigma^b - \sigma^h) \\ \sigma^h = y^K(\sigma^K + \sigma^p) \\ \sigma^b = x^K(\sigma^K + \sigma^p). \end{cases} \tag{56}$$

Denote fire sale benefits for each unit of household wealth as κ^{fs}, which is net wealth transfer from bankers to households due to the temporary market pressure. By market clearing, we have

$$\underbrace{(1-w)}_{\text{total household wealth}} \cdot \kappa^{\text{fs}} = \underbrace{I^B}_{\text{bankruptcy indicator}} \cdot \underbrace{w}_{\text{total banker wealth}} \cdot \underbrace{\frac{\Delta x}{(1-\alpha^0)p_t}}_{\text{fire sale quantity for each unit of banker wealth}} \cdot \underbrace{\alpha^0 p_t}_{\text{wealth transfer for each unit sale}}$$

$$\Rightarrow \kappa^{\text{fs}} = \alpha \Delta x \frac{w}{1-w} \tag{57}$$

Then we have the following household first order condition:

$$\mu^R + \frac{\underline{A}}{p} - r^d \leq (\sigma^K + \sigma^p)^2 y^K + \lambda \frac{\kappa^p}{1 - y^K \kappa^p + \kappa^{\text{fs}}}, \text{ equality if } y^K > 0 \tag{58}$$

In equation (58), the left hand side is the excess return on productive capital over bank debt, while the right hand side includes the cost of the additional risks from productive capital compared to bank debt. When the excess return is lower than the cost, households do not hold productive capital and set $y^K = 0$.

On the other hand, the first order condition on bank productive capital holding is

$$\mu^R + \frac{\bar{A}}{p} - r^d = (\sigma^K + \sigma^p)^2 x^K + \lambda \frac{\kappa^p + \alpha}{1 - x^K \kappa^p - \alpha \Delta x} \tag{59}$$

since banks always hold a positive amount of productive capital.[4] The excess return of productive capital over debt consists of three components: volatility, endogenous price decline and fire sale losses in case of financial distress shocks.

Combining (58) and (59), we have

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

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

[4]Suppose not, then banks are not subject to financial distress shocks by increasing its productive capital holding from 0 to a small positive number, but increases profit strictly. Thus we easily arrive at a contradiction.

Next, the bank first order condition over inter-bank lending is

$$r^f - r^d = \lambda \frac{\alpha}{1 - x^K \kappa^p - \alpha \Delta x} \tag{60}$$

which implies that the bank debt interest rate is lower than bank risk-free rate, because bank debt funding is runnable.

Combining (59) and (60), we arrive at the excess return expression for the productive capital in normal time as:

$$\mu^R + \frac{\bar{A}}{p} - r^f = (\sigma^K + \sigma^p)^2 x^K + \lambda \frac{\kappa^p}{1 - x^K \kappa^p - \alpha \Delta x} \tag{61}$$

We note that productive capital is also subject to the losses of $\kappa^p x^K$ during a distress, which arrives at intensity λ. As a result, the full excess return expression should be

$$\mu^R + \frac{\bar{A}}{p} - \lambda \kappa^p - r^f = (\sigma^K + \sigma^p)^2 x^K + \lambda \kappa^p \frac{x^K \kappa^p + \alpha \Delta x}{1 - x^K \kappa^p - \alpha \Delta x} \tag{62}$$

Intuitively, equation (62) implies that the excess return of productive capital above the risk-free rate is compensating the volatility of the productive capital, as well as the potential price drop. Note that we did not attribute the $\alpha \Delta x$ component to x^K, since it is related to the amount of shot-term debt funding. Another way to look at the above problem is to rewrite the bank welath dynamics in terms of x^K and x^d, which leads to

$$\frac{dw^b}{w^b} = \left(r^f + x^K (\mu^R + \frac{\bar{A}}{p} - \lambda \kappa^p - r^f) - x^d (r^d - r^f) - \lambda \alpha x^d \right) dt - \hat{c} dt$$

$$+ x^K (\sigma^K + \sigma^p) dB_t - \kappa^b (dN_t - \lambda dt)$$

Diagnostic Beliefs

When we solve the equilibrium jumps upon dN_t with diagnostic beliefs, additional adjustments are needed to accomodate the distortions induced by the Diagnostic beliefs. As we have assumed, households believe that they have Bayesian beliefs and make decisions with the Bayesian policies. However, the realizations during a crisis will be different from their expectations, which may cause additional disruptions. 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 Other Measures of Risk Premium

In this subsection, we discuss other measures of the risk premium. In the main text, we use the credit spreads of long-term bonds as a measure of risk premium, mainly because of the data availability of the same measure. In this part, we show several other measures of risk premium. The first one is the bank equity excess return. we show that although bank equity is not tradable in the model, the excess returns are still positive under Bayesian expectations. The second one is Sharpe ratio. We show how to adjust the definition to accommodate jumps in the model.

Bank Equity Excess Returns

We note that bank equity return is not simply dw^b/w^b, since this wealth growth term also incorporates banker consumption, which should be interpreted as a dividend payment. Formally, the expected return of bank equity is

$$r^e = r^f + x^K(\mu^R + \frac{\bar{A}}{p} - r^f - \lambda\kappa^p) - x^d(r^d - r^f) - \lambda\alpha x^d$$

Expressing the right-hand terms with (27) and (28), we have

$$r^e - r^f = \underbrace{x^K\left((\sigma^K + \sigma^p)^2 x^K + \lambda\kappa^p \frac{x^K\kappa^p + \alpha\Delta x}{1 - x^K\kappa^p - \alpha\Delta x}\right)}_{\text{total risk compensation for holding productive capital}} + \underbrace{\lambda\alpha x^d \frac{x^K\kappa^p + \alpha\Delta x}{1 - x^K\kappa^p - \alpha\Delta x}}_{\text{total risk compensation for taking debt}}$$

where two terms of risk compensations appear. The first term is the total risk compensation for holding productive capital, including the risk premium of volatility and the decline in wealth due to financial distress shocks. The second term is the total risk compensation for raising short-term debt. Alternatively, we can write the risk compensation as the following

$$r^e - r^f = \underbrace{x^K}_{\text{exposure to } dB_t \text{ shock}} \cdot \underbrace{(\sigma^K + \sigma^p)^2 x^K}_{\text{compensation to } dB_t \text{ shock}} + \underbrace{(x^K\kappa^p + \alpha x^d)}_{\text{exposure to } dN_t \text{ shock}} \cdot \underbrace{\frac{\lambda(x^K\kappa^p + \alpha\Delta x)}{1 - x^K\kappa^p - \alpha\Delta x}}_{\text{compensation to } dN_t \text{ shock}}$$

As a result, in this Bayesian model, the bank equity excess return should always be above zero.

For the Diagnostic model with diagnostic expectations, there is a "surprise" element in the realized excess returns during a crisis, due to the additional jumps in the price of capital. We just need to take into account the impact on the excess return by the additional jumps.

Sharpe Ratio

Another measure of the risk premium in the economy is the Sharpe ratio. However, since we have Poisson jumps in a continuous time economy, it is not enough to only incorporate the Brownian terms to measure risk. For any bank-held assets with process

$$dR_t = \mu dt + \sigma dB_t - \kappa dN_t$$

we denote the Sharpe ratio as

$$SR = \frac{E[R_{t+\Delta t} - R_t] - r^f \Delta t}{var(R_{t+\Delta t} - R_t)} \approx \frac{\mu - \lambda\kappa - r^f}{(\sigma)^2 + \lambda|\kappa|} \tag{63}$$

where we have taken the perspective of Δt being small but positive. For productive capital, the modified Sharpe ratio is

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

According to (29), the numerator is positive. Therefore, the model implied Sharpe ratio for productive capital is always positive.

A.7 Credit Spread

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

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

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

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, 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,

$$\begin{aligned}\frac{\partial v}{\partial w}w\mu^w + \frac{1}{2}\frac{\partial^2 v}{\partial w^2}w^2(\sigma^w)^2 + \frac{\partial v}{\partial\lambda}\mu^\lambda - r^f v = x^K(\sigma^K + \sigma^p)\frac{\partial v}{\partial w}w\sigma^w \\ + \lambda\pi\frac{1}{1-\kappa^b}\left(v - (1-\kappa^p-\hat{\kappa}^0)\right) + \lambda(1-\pi)\frac{1}{1-\kappa^b}\kappa^v v + \lambda^\tau(v-1)\end{aligned} \tag{64}$$

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 (64), we can get the HJB equation for $\tilde{v}$ as

$$\frac{\partial \tilde{v}}{\partial t} = \lambda\pi\frac{1}{1-\kappa^b}\left(v-(1-\kappa^p-\hat{\kappa}^0)\right)+\lambda(1-\pi)\frac{1}{1-\kappa^b}\kappa^v v+\lambda^\tau(v-1)$$

$$+x^K(\sigma^K+\sigma^p)\frac{\partial v}{\partial w}w\sigma^w+r^f\tilde{v}-\left(\frac{\partial \tilde{v}}{\partial w}w\mu^w+\frac{1}{2}\frac{\partial^2\tilde{v}}{\partial w^2}w^2(\sigma^w)^2+\frac{\partial \tilde{v}}{\partial \lambda}\mu^\lambda\right)$$

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 (64) into

$$\begin{aligned}\frac{\partial v(w,\lambda)}{\partial\lambda}\mu^\lambda(\lambda)-r^f(w,\lambda)v(w,\lambda)&=\lambda\pi\frac{1}{1-\kappa^b(w,\lambda)}\left(v(w,\lambda)-(1-\hat{\kappa}^0)\right)\\+\lambda(1-\pi)\frac{1}{1-\kappa^b(w,\lambda)}\kappa^v(w,\lambda)v(w,\lambda)&+\lambda^\tau(\lambda)(v(w,\lambda)-1),\quad w\in\{0,1\}\end{aligned} \tag{65}$$

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 (65) 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 $\varepsilon > 0$.

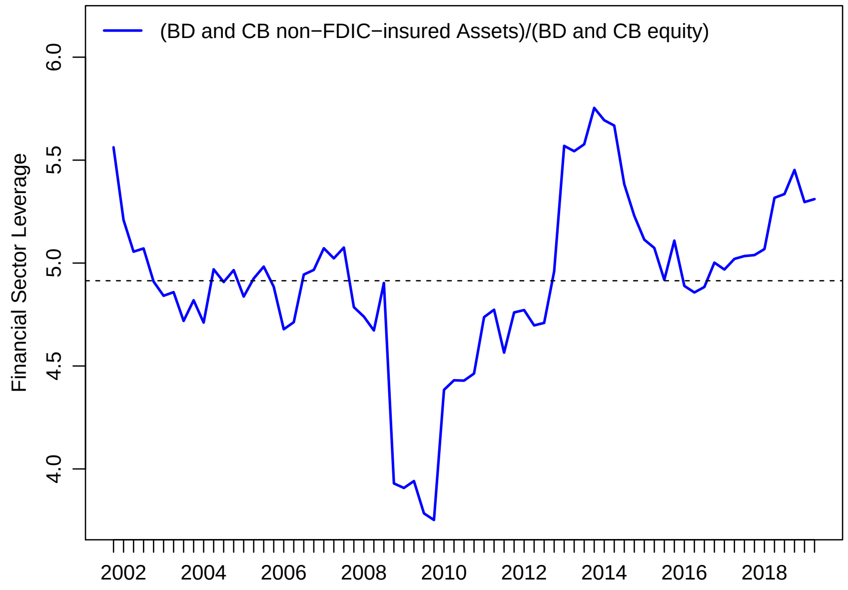
Finally, we notice that once the $\lambda = \lambda^*$, it will not go up or down unless there is a dN_t shock. Once we know the jump component, we can solve $v(w,\lambda^*)$ along the w dimension as an ODE. The ODE is

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

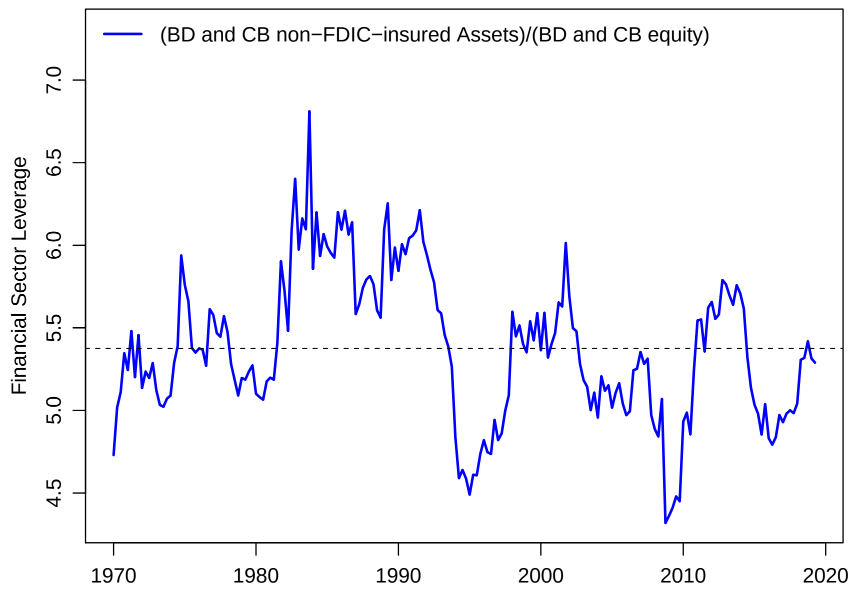
for $w \neq 0, 1$.

B Leverage Target

Here we provide more details about the calculations of moment target for bank leverage. We use the Flow of Funds data to calculate average bank leverage from runnable (non-insured) deposits. As shown in Figure 20, bank leverage based on non-FDIC insured bank deposits is slightly below 5. We use Egan, Hortaçsu and Matvos (2017) estimate that the fraction of FDIC-insured deposits is 0.5.



(a) Bank Leverage (using data on FDIC-insured deposits)



(b) Bank Leverage (FDIC-insured deposits are inferred)

Figure 20: Bank Leverage in the Flow of Funds Data. This figure illustrates the time series of bank leverage in the U.S. using flow of funds data. We map banks in the model to "private depository institutions" and "security broker dealers" in flow of funds. Bank equity is measured as total bank assets minus bank liabilities. Since the model only captures banks' runnable liabilities liabilities, we measure effective bank liability as total liability minus FDIC-insured deposits, and calculate bank leverage as (effective bank liability + bank equity)/bank equity. The data on FDIC-insured deposits are only available after 2002, so the data sample starts from 2002 in panel (a). An alternative way is to assume a constant fraction of FDIC-insured deposits over total deposits (here we assume a half). In panel (b), we show the results under this alternative method.