A Macro-Finance Model with Sentiment *

PRELIMINARY AND INCOMPLETE

Peter Maxted

Harvard University

June 10, 2019

Abstract

This paper studies financial market and macroeconomic dynamics in a continuous-time macroeconomic model with both financial frictions and diagnostic (non-rational) expectations. Diagnostic expectations overweight future states that are representative of recent news. When recent news has been positive, market sentiment will be elevated and market prices, which reflect subjective expectations, will be dislocated from fundamentals. The model quantifies the causal role of elevated sentiment in amplifying financial market fragility. The model also shows how diagnostic expectations and nonlinear financial frictions interact to generate boom-bust cycles in investment and output.

*I am grateful to Francesca Bastianello, Pedro Bordalo, Daniele d'Arienzo, Emmanuel Farhi, Paul Fontanier, Xavier Gabaix, Nicola Gennaioli, Thomas Graeber, David Laibson, Matteo Maggiori, Andrei Shleifer, and Jeremy Stein for helpful comments. The computations for this paper were run on the Research Computing Environment supported by the Institute for Quantitative Social Science in the Faculty of Arts and Sciences at Harvard University.

1 Introduction

The 2007-2008 Financial Crisis and the Great Recession that followed are a powerful reminder of the debilitating effects that financial market turmoil can have on the broader economy. Since the crisis, macro-finance theory has turned to frictions in financial intermediation to account for financial crises and their consequences for the real economy. Financial frictions prevent risks from being shared and funds from flowing to productive investments, especially in times when frictions are particularly acute. Crises are modeled as nonlinearities where occasionally-binding constraints on financial intermediaries generate sharp increases in risk premia and contractions in output (e.g., He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014)).

While models of financial frictions can capture nonlinearities in risk premia and the transmission of financial sector vulnerabilities into aggregate downturns, a growing body of empirical evidence questions the ability for models with rational expectations to capture the initial triggers of financial fragility. The 2007-2008 Financial Crisis is a particularly salient example. By the summer of 2007 it was clear that the housing market was deflating, yet financial market indicators such as credit spreads and the VIX showed little sign of the systemic risks that quickly revealed themselves following the bankruptcy of Lehman in September 2008 (Gennaioli and Shleifer, 2018). This evidence of neglected risk preceding crises is not unique to 2007-2008 episode. Examining an international panel of financial crises, Krishnamurthy and Muir (2017) find that spreads are regularly "too low" before financial crises, and that frothy financial markets predict future crises. Krishnamurthy and Muir (2017) directly conclude that this pre-crisis evidence is difficult to square with rational models of financial fragility.

Patterns of excessive optimism preceding financial market and macroeconomic downturns appear consistently. Greenwood and Hanson (2013) show that the credit quality of corporate

debt issuers deteriorates during credit booms, and that deteriorating issuer quality predicts a tightening of credit markets, a widening of spreads, and low realized returns. Moreover, issuer quality is poor following periods of low defaults, consistent with investors extrapolating past default rates. Taking this result to the macroeconomy, López-Salido et al. (2017) estimate that periods of loose credit and narrow spreads predict a decline in future economic activity once the sentiment-driven financial cycle predictably reverses.

Direct expectations data from professional forecasters supports this evidence of cyclical overreaction to recent economic trends.[1] Using professional forecasts of the Baa-Treasury spread from Blue Chip Financial Forecasts, Bordalo et al. (2018a) find that forecast errors are predictable: when the credit spread is narrow the expected future spread is too narrow, and when the credit spread is wide the expected future spread is too wide. Bordalo et al. (2018c) document a similar pattern of overreaction in professional forecasters' expectations across a variety of macroeconomic outcomes, such as real GDP and consumption growth.

Given the adverse effects that financial fragility can have on the real economy, it is imperative to develop a macro-finance model that is consistent with this evidence of non-rational expectations and can characterize how sentiment interacts with financial frictions to jointly drive financial market and business cycle dynamics. To do so, this paper builds a quantitative macroeconomic model which combines financial frictions with extrapolative expectations. The macroeconomic model is based on He and Krishnamurthy (2019). He and Krishnamurthy (2019) present a continuous-time RBC model augmented with a financial intermediary sector in order to generate financial frictions. The financial intermediary sector faces an occasionally-binding constraint on its ability to raise equity funding from households. The intermediary sector's funding capacity becomes a key state variable for

[1]Data on the expectations of financial market and macroeconomic professionals is most consistent with the model. As will be detailed later, restrictions on asset market participation imply that it is sophisticated financial intermediaries who are responsible for pricing risky assets (see equation (22)). Similar extrapolative behavior is exhibited by households, as discussed in Fuster et al. (2010) and Greenwood and Shleifer (2014).

the economy, capturing the severity of financial frictions. In non-distress periods the model behaves similarly to a frictionless RBC model and is calibrated to quantitatively represent the U.S. economy. However, a sequence of poor returns leads to financial distress as the intermediary sector moves closer to its funding constraint. In periods of financial distress the nonlinearities arising from financial frictions become quantitatively important, thereby adding a systemic risk element to the model. The model is solved globally in order to fully characterize the nonlinear effects of financial frictions on the economy's dynamics.

I depart from rational expectations by introducing behavioral frictions to this model of financial frictions. This paper develops a method for extending rational models with a continuous-time variant of diagnostic expectations (Bordalo et al., 2018a). Diagnostic expectations are a forward-looking model of extrapolative expectations in which agents overweight future states that are *representative* of new information (Kahneman and Tversky, 1972). When recent shocks have tended to be positive, sentiment is elevated and agents are over-optimistic about future economic growth. The reverse holds when recent shocks have tended to be negative. Relative to Bordalo et al. (2018a), the key innovation of this paper is to model diagnostic expectations such that they can be applied to endogenous processes. This implies not only that recent economic performance affects expectations (a standard feature of extrapolative expectations), but also that expectations can then feed back into the dynamical system to alter the future evolution of the processes on which expectations are formed. This important methodological advancement allows this paper to derive a number of important consequences resulting from the interaction of behavioral and financial frictions.

First, the model quantifies the causal role of over-optimism in triggering the amplification of financial distress into a full-blown financial crisis. Elevated sentiment erodes the balance sheets of financial intermediaries by lowering returns relative to expectations. This sentiment-driven build-up of systemic risk is neglected by agents in the model because diagnostic expectations drive a wedge between the true and perceived probability of a future

financial crisis. One important implication of diagnostic expectations is that, unlike rational expectations, endogenous market prices will not always reflect the true risks building in the economy.

The model's second prediction is that financial crises are less likely to occur when expectations are diagnostic than when they are rational. This is in direct opposition to the typical narrative that extrapolative beliefs create financial instability (e.g., the Financial Instability Hypothesis of Minsky (1977)). This *stability from beliefs* result arises because sentiment tracks recent economic shocks, and therefore sentiment is typically depressed as the economy nears a financial crisis. Such excessive pessimism leads the financial sector to earn higher returns than expected, thus allowing the financial sector to quickly rebuild its balance sheet. While it is still the case that the interaction of financial distress plus elevated sentiment is highly predictive of financial crises, this dangerous combination occurs rarely. It is much more likely for sentiment to be depressed as the financial sector nears its funding constraint, generating a hedging effect which lowers the unconditional probability of financial crises under diagnostic expectations.

Third, diagnostic expectations interact with nonlinear financial frictions to generate boom-bust cycles in investment and output. When sentiment is elevated, investment in capital will rise because capital is valued highly. This creates a short-run investment boom. However, elevated sentiment simultaneously leads to a deterioration of the financial sector's balance sheet. This increases financial frictions and impairs investment. The sentiment effect is immediate while the balance sheet effect materializes slowly, such that the interaction generates a sentiment-driven boom in the short run followed by a financial-frictions-driven bust in the long run.

In addition to developing a model that quantifies the effect of sentiment on macroeconomic and financial market dynamics, this paper makes two methodological contributions to the behavioral finance and macroeconomic literatures. First, the results of this paper high-

light the importance of using (continuous-time) global solution methods for characterizing the effect of beliefs on economic dynamics.[2] Indeed, the narrative on behavioral macroeconomic dynamics is one of sentiment-driven expansions and slumps (e.g., Keynes (1936)). This is fundamentally an analysis of the cyclical effects of expectations away from the steady state, and global solution methods allow for the complete characterization of the nonlinear dynamical system.

Second, an often-noted negative consequence of introducing behavioral elements into an economic model is that the model must typically be simplified elsewhere in order to keep the solution tractable. This is not the case in this paper. I provide a microfounded model of non-rational expectations that is used to generalize a cutting-edge dynamic stochastic general equilibrium macroeconomic model with financial frictions. This allows for the study of non-rational expectations without compromising the rich equilibrium dynamics on which sentiment can interact.

Related Literature The macroeconomic model in this paper follows from a large literature on financial frictions. Seminal work includes Kiyotaki and Moore (1997) and Bernanke et al. (1999). Recent research has tended to use continuous-time methods in order to study global dynamics in models with nonlinearities. Examples include Adrian and Boyarchenko (2012), He and Krishnamurthy (2013, 2019) Brunnermeier and Sannikov (2014), Maggiori (2017), and Moreira and Savov (2017). The contribution of this paper is the introduction of non-rational diagnostic expectations.

Diagnostic expectations align this set of models more closely with the empirical findings on the behavioral triggers of financial distress. The literature mentioned above provides a partial summary of this evidence. In the banking sector, Baron and Xiong (2017) find

[2]The macro-finance literature (e.g., He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014)) utilizes continuous-time global solution methods to characterize the nonlinearities arising from occasionally-binding constraints. These nonlinear financial friction dynamics are also important in this paper.

that rapid bank credit expansion predicts increased equity crash risk and lower bank equity returns. In the cross-section of U.S. banks, Fahlenbrach et al. (2017) show that rapid loan growth predicts equity underperformance and elevated crash risk over the next three years. Adrian and Brunnermeier (2016) find that systemic risk measurements suffer from a "procyclicality pitfall" in which tail risks build up in the background of low-volatility environments. Mian et al. (2017) observe that increases in the household debt-to-GDP ratio predict lower future GDP growth, and also that economic forecasters are over-optimistic at the end of household debt expansions. Schularick and Taylor (2012) conclude that financial crises are often "credit booms gone wrong."

There is also a growing theoretical literature focusing on behavioral credit cycles and the behavioral triggers of crises. Closest to the current paper, Bordalo et al. (2018a) develop the original model of discrete-time diagnostic expectations and apply these expectations to a simple frictionless macroeconomic model. This work is extended in Bordalo et al. (2019b), who study the business cycle implications of diagnostic expectations in a quantitative heterogeneous firm model.[3] Gennaioli and Shleifer (2018) summarize the research on diagnostic expectations and present a belief-driven narrative of the 2007-2008 Financial Crisis. Gertler et al. (2017) examine how waves of optimism about future capital returns can generate an ex-ante buildup of leverage which increases the fragility of the banking sector once expectations disappoint. Greenwood et al. (2019) develop a model of firm financing and default in which extrapolative expectations about past default rates generate boom-bust credit market cycles. Jin (2015) presents a model in which crash risk is extrapolated from recent experiences. Gennaioli et al. (2012) model neglected risk and its effect on financial fragility. Barberis (2018) provides a summary of research on the psychology of extrapolative expectations and their application in macro-finance.

[3]See Bordalo et al. (2019a) and Bordalo et al. (2018b) for discrete-time applications of diagnostic expectations to financial markets.

This paper also contributes more broadly to a budding literature on general equilibrium macroeconomic theory augmented with behavioral features. A partial list of recent examples includes Sims (2003), Fuster et al. (2012), Woodford (2013), Gabaix (2016a,b), Farhi and Werning (2017), and García-Schmidt and Woodford (2019). These papers focus on various forms of cognitive limits and typically imply sluggish expectations.[4] Here, the model of diagnostic expectations generates overreaction to new information.

The organization of this paper is as follows. Section 2 describes the macroeconomic model. In section 2, only the reduced-form beliefs process is specified. Section 3 outlines the solution strategy and the baseline calibration. The global solution is characterized in section 4. The model's financial market results are presented in section 5. Section 6 discusses the macroeconomic implications of diagnostic expectations. Section 7 provides a microfoundation for diagnostic expectations.

2 Macroeconomic Model

The macroeconomic model builds on He and Krishnamurthy (2019, henceforth HK). The HK model is chosen because HK is one of the first quantitative papers in the continuous-time macro-finance literature, which has instead focused mainly on qualitative results to date. The HK model is successful at quantitatively matching the downside macroeconomic risks precipitated by financial instability. However, the authors acknowledge that a failure of their model is its inability to decouple the market price of risk from the probability of a future financial crisis, as was observed in financial markets preceding the collapse of Lehman.

I extend HK in two ways.[5] Most importantly, I generalize the model to allow for diagnostic expectations. Second, I introduce a simple labor income margin in order to improve

[4] One exception is Fuster et al. (2012), which generates overreaction to recent data.

[5] Both of these extensions can be shut down under certain parameter assumptions, and therefore the model in this paper nests He and Krishnamurthy (2019).

the model's quantitative fit. Whenever possible, I adopt the same notation as HK.

2.1 Model Setup

Time is continuous with t denoting the current period. The economy has two sectors: households and financial intermediaries. The economy has two types of capital: productive capital K_t and housing capital H. The housing supply is fixed and normalized to $H \equiv 1$. The price of a unit of capital is denoted q_t and the price of housing is denoted P_t. These prices are endogenous and will be determined in equilibrium.

Only financial intermediaries possess the requisite skills to operate capital, and therefore the intermediary sector directly holds K_t and H. Intermediaries fund these purchases by issuing debt and equity to households. The key financial friction in this model is that each intermediary faces an *equity capital constraint* which restricts the intermediary's ability to raise equity funding. When binding, the intermediary sector must replace its equity funding with additional debt funding.

The economy features an "AK" production technology, with flow output Y_t given by:

$$Y_t = AK_t, \tag{1}$$

where A is a positive constant determining the productivity of capital. Productive capital evolves according to:

$$\frac{dK_t}{K_t} = (i_t - \delta)dt + \sigma dZ_t, \tag{2}$$

where i_t is the endogenous rate of capital installation at time t and δ is the exogenous depreciation rate. The term σdZ_t is a capital quality shock.[6] $\{Z_t\}$ is a standard Brownian

[6]In the RBC literature it is more common to have shocks hit productivity. Having shocks directly alter capital improves tractability by reducing the requisite number of state variables. Capital quality shocks are

motion. Capital quality shocks σdZ_t are the only source of uncertainty in the model.

Investment in capital is subject to quadratic adjustment costs. For a gross capital installation of $i_t K_t$, the cost of this installation is:

$$\Phi(i_t, K_t) = i_t K_t + \frac{\xi}{2}(i_t - \delta)^2 K_t.$$

2.2 Diagnostic Expectations

Overview I depart from rationality by assuming that all agents have extrapolative expectations about the capital stock. Specifically, the model of expectations in this paper is an extension of the model of diagnostic expectations developed in Bordalo et al. (2018a, henceforth BGS).

Diagnostic expectations are based on Kahneman and Tversky's *representativeness heuristic*, defined as follows:

> "an attribute is representative of a class if it is very diagnostic; that is, the relative frequency of this attribute is much higher in that class than in the relevant reference class." (Tversky and Kahneman, 1983, p. 296)

Following BGS, the reference class should reflect the absence of new information.[7] In the realm of expectations, this implies that future states become more representative, and are therefore overweighted, when they become more likely to occur in light of incoming data. Put differently, agents overweight future states that are diagnostic of new information.

The model of diagnostic expectations developed in this paper features two innovations relative to BGS. First, diagnostic expectations are cast in continuous time. Continuous-time

also used by Gertler and Kiyotaki (2010) and Brunnermeier and Sannikov (2014).

[7]For example, after seeing a patient test positive for a rare disease, doctors tend to overweight the probability that the patient is truly sick (Casscells et al., 1978). Here, the positive test is new information. The state of the world where the patient is truly sick becomes representative when evaluated relative to the reference class in which the patient never takes the test (Bordalo et al., 2018a).

methods are particularly applicable for the study of non-rational expectations in dynamic equilibrium models. Over short (instantaneous) horizons, Itô's lemma provides a straightforward method for calculating forecast errors on endogenous equilibrium objects (e.g., asset returns). At longer horizons, the Kolmogorov forward equation can be used to characterize the distribution over future states that agents perceive, as well as the true equilibrium distribution over future states. Additionally, as was discussed in the introduction, global solutions enable the dynamics of non-rational expectations to be evaluated over the entire state space.

Second, the methodology developed here allows for diagnostic expectations to be applied to endogenous Itô processes. This is a crucial methodological advancement because it introduces a channel for expectations to feed back into the equilibrium and alter the endogenous future path of the processes over which expectations are formed. This is my main improvement on BGS, who can only apply diagnostic expectations to exogenous AR(N) processes.

A goal for this paper's formulation of diagnostic expectations is to constitute a tractable and portable methodology that enables rational models to be augmented with diagnostic expectations using a single additional state variable. In this section I specify the reduced-form diagnostic expectations process as applied to the macroeconomic model. The microfoundation is provided in section 7 .

Expectations of Capital All agents have diagnostic expectations about the log capital stock.[8] Capital is the fundamental in this economy — capital alone determines output ($Y_t = AK_t$), and capital quality shocks are the sole source of uncertainty. Note that the evolution of capital is endogenous since it depends on the endogenous investment rate i_t (see equation (2)). This endogeneity introduces expectations-driven macroeconomic dynamics, where non-rational expectations alter investment which in turn affects the growth rate of

[8]Using Itô's lemma, the log capital stock evolves according to $dk_t = (i_t - \delta - \frac{\sigma^2}{2})dt + \sigma dZ_t$.

capital and output.

The psychology of diagnostic expectations is that the agent has in the back of her mind the true model, yet overweights future states of the world that are representative of new information. Here, new information corresponds to recent capital quality shocks. This notion is formalized by the following measure of new information at time t:

$$\mathcal{I}_t \equiv \int_0^t e^{-\kappa(t-s)} \sigma dZ_s. \tag{3}$$

$\mathcal{I}_t$ is a weighted integral of past shocks to capital, where the weight decays exponentially at rate κ as shocks occur further in the past.[9] Thus, $\mathcal{I}_t > 0$ when recent capital quality shocks have tended to be positive, and $\mathcal{I}_t < 0$ when recent capital quality shocks have tended to be negative. $\mathcal{I}_t$ can be expressed in differential form as:

$$d\mathcal{I}_t = -\kappa \mathcal{I}_t dt + \sigma dZ_t. \tag{4}$$

Equation (4) illustrates that $\mathcal{I}_t$ drifts back to 0 at rate κ, subject to the new shocks that hit capital. Formally, $\mathcal{I}_t$ is an Ornstein-Uhlenbeck process. Intuitively, it is the continuous-time analogue of an AR(1) process.

Throughout, I will use hat-notation to denote the perceptions of diagnostic agents. Under diagnostic expectations, agents believe that the instantaneous evolution of capital is:

$$\frac{\widehat{dK_t}}{K_t} = (i_t - \delta)dt + \theta \mathcal{I}_t dt + \sigma dZ_t. \tag{5}$$

Parameter $\theta \geq 0$ governs the extent to which agents "judge by representativeness." $\theta = 0$ recovers rationality. When $\theta > 0$, agents overweight future states that are diagnostic of

[9] In discrete time, any individual shock can itself represent new information. In continuous time, I integrate over the past sequence of shocks because any individual shock σdZ has only an infinitesimal effect on the capital stock. This specification is similar to the definition of "sentiment" in Barberis et al. (2015).

recent information. This overreaction to recent capital quality shocks results in a biased perception of the drift of capital.

Equation (5) specifies that diagnostic expectations of the instantaneous evolution of capital from period t to period $t + dt$ are biased by $\theta \mathcal{I}_t dt$. At longer horizons, the agent forms expectations iteratively by simulating the economy forward state-by-state. When doing so, the agent forms expectations about the future evolution of capital as if $\mathcal{I}_t$ decays deterministically over time according to:

$$\widehat{d\mathcal{I}_t} = -\kappa \mathcal{I}_t dt. \tag{6}$$

Equation (6) implies that that diagnostic agents have incorrect expectations about their own future expectations because they do not know that they are diagnostic. In particular, a comparison of equations (4) and (6) illustrates that diagnostic agents do not understand that future capital quality shocks will alter the bias of their future expectations.

In summary, agents make two mistakes when $\theta > 0$. First, they hold incorrect beliefs about the drift of capital. Second, they have incorrect expectations about their own future expectations because they do not understand that they are diagnostic.

Figure 1 provides an illustrative example of diagnostic expectations applied to an arithmetic Brownian motion (ABM).[10] The blue line plots the ABM process. The vertical black line marks time t when a prediction is made about the future evolution of the ABM. The solid red line plots the diagnostic prediction and the dashed black line plots the rational prediction. Because recent shocks to the ABM have been positive at the time of prediction, diagnostic expectations are over-optimistic about the future evolution of the ABM process.

[10]Note that log capital k_t would follow arithmetic Brownian motion if investment rate i_t were constant.

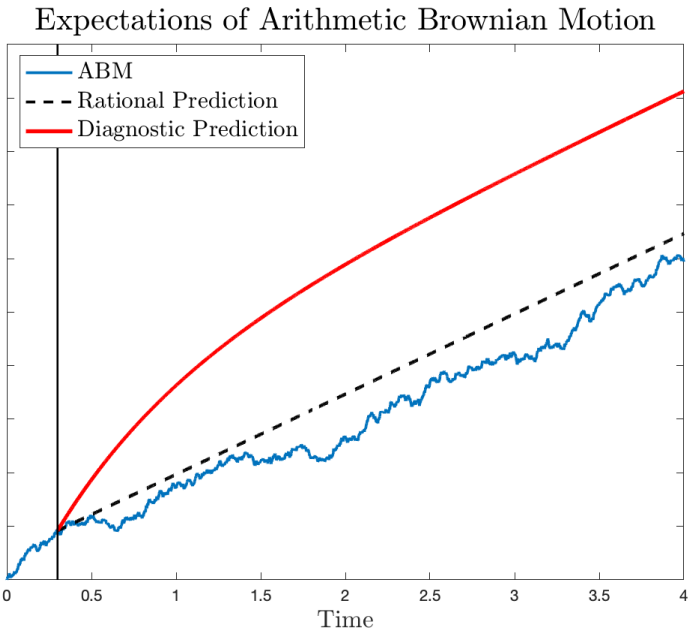


Figure 1: Diagnostic expectations of arithmetic Brownian motion (ABM).

Decomposing Diagnostic Expectations This paper's formulation of diagnostic expectations reconciles mechanical models of extrapolation with dynamic forward-looking expectations. This property is highlighted by the following decomposition of capital growth process $\frac{\widehat{dK_t}}{K_t}$:

$$\frac{\widehat{dK_t}}{K_t} = \underbrace{\frac{dK_t}{K_t}}_{\text{Rational}} + \underbrace{\theta \mathcal{I}_t dt}_{\text{Wedge}}.$$

The diagnostic agent's perception of the capital growth process can be separated into two components: a rational capital growth process and a diagnostic wedge capturing sentiment. The rational component is forward looking. The diagnostic wedge is a backward-looking function of past shocks, since $\mathcal{I}_t \equiv \int_0^t e^{-\kappa(t-s)} \sigma dZ_s$.

Diagnostic expectations differ from purely-mechanical models of extrapolative expectations (e.g., adaptive expectations) because diagnostic expectations depend on the true underlying process being predicted. This forward-looking element comes from the rational

component of diagnostic expectations. For example, a regime shift such as an unanticipated shock to depreciation rate δ would cause the diagnostic agent to update her prediction of $\frac{\widehat{dK_t}}{K_t}$ because the rational component $\frac{dK_t}{K_t}$ would change. However, the diagnostic wedge term means that diagnostic expectations are still subject to persistent errors.[11] Though the agent has the true model in the back of her mind (the rational component), future expectations are nonetheless biased toward states that are representative in light of new information (the diagnostic wedge). Thus, diagnostic expectations are characterized by the "kernel of truth" property: diagnostic expectations depend on the true economic process, yet they overreact to recent patterns in the data.

The above discussion specifies the extent to which diagnostic expectations are robust to the Lucas critique. Unlike mechanical models of extrapolative expectations, diagnostic expectations are forward looking and dependent on the true underlying economic model. This dependence on the true underlying process comes from the rational component of diagnostic expectations. Even with this forward-looking behavior, diagnostic expectations are still subject to persistent extrapolative errors due to the backward-looking diagnostic wedge.

2.3 The Financial Intermediary Sector

Individual Intermediaries There is a continuum of financial intermediaries. Each intermediary is run by a single banker. In order to purchase capital and housing, intermediaries must raise funds from the households. In particular, each intermediary is able to issue risk-free (instantaneous) debt and risky equity to households.[12] Equity issuance is subject to an endogenous equity capital constraint, such that each intermediary can issue up to ϵ_t of equity.

[11] Parameter κ governs the persistence of these errors.

[12] Because debt is instantaneous and asset prices are continuous, there will always exist an equity buffer that is large enough to absorb losses and ensure that debt is risk-free.

Let $d\tilde{R}_t$ denote the realized instantaneous return on the intermediary's equity. Equity capital constraint ϵ_t evolves according to:

$$\frac{d\epsilon_t}{\epsilon_t} = d\widetilde{R}_t. \tag{7}$$

Constraint ϵ_t should be thought of as the banker's "reputation." Poor investment returns hurt the banker's reputation, lowering the banker's ability to issue equity in the future. The banker does not consume. Instead, the banker has mean-variance preferences over reputation:

$$\widehat{\mathbb{E}}_t\left[\frac{d\epsilon_t}{\epsilon_t}\right] - \frac{\gamma}{2}\widehat{Var_t}\left[\frac{d\epsilon_t}{\epsilon_t}\right] = \widehat{\mathbb{E}}_t\left[d\widetilde{R}_t\right] - \frac{\gamma}{2}\widehat{Var_t}\left[d\widetilde{R}_t\right]. \tag{8}$$

Hat-notation is used to indicate that the banker has diagnostic expectations rather than rational expectations.

This reputation-based constraint behaves similarly to the more standard "net worth" constraint in which the banker's net worth fluctuates as a function of past performance, and the banker's ability to raise future capital depends on net worth. The benefit of the reputation-based constraint is that the banker has no net worth and does not consume, leaving the representative household to consume all of the economy's output. This is the typical assumption in quantitative macroeconomic models without financial frictions, and will allow for a more standard calibration of the macroeconomic model.[13]

The Aggregate Intermediary Sector Let $\mathcal{E}_t$ denote the maximum equity capital that can be raised by the aggregate intermediary sector. $\mathcal{E}_t$ evolves according to:

$$\frac{d\mathcal{E}_t}{\mathcal{E}_t} = d\widetilde{R}_t - \eta dt + d\psi_t. \tag{9}$$

[13]For details, see section 2.2 of He and Krishnamurthy (2019).

The first term captures that all bankers behave identically, and therefore aggregate equity capital evolves along with the reputation of each individual banker. The second term captures exogenous banker exit, which occurs at rate η. Exit is needed to ensure that bankers don't escape their equity constraint. The final term $d\psi_t \geq 0$ reflects entry into the intermediary sector. Entry occurs deep in crisis times when reputation is sufficiently low. Bank entry provides a boundary condition for the model, and details are provided in section 3.1.

2.4 The Household Sector

Consumption There is a unit measure of households. Households earn utility over consumption of the output good (c_t^y) and housing services (c_t^h). The output good is the numeraire. Since households do not hold housing directly, housing services must be rented at endogenous rental rate D_t.

The household maximizes the value function

$$\widehat{\mathbb{E}}\left[\int_0^\infty e^{-\rho t}\frac{C_t^{1-\gamma_h}}{1-\gamma_h}\right], \tag{10}$$

where C_t is a Cobb-Douglas consumption aggregator

$$C_t = (c_t^y)^{1-\phi}(c_t^h)^{\phi}. \tag{11}$$

Intratemporal maximization yields the following optimality condition:

$$\frac{c_t^y}{c_t^h} = \frac{1-\phi}{\phi}D_t. \tag{12}$$

Labor Income Households supply one unit of labor without disutility. There is an infinite marginal utility cost for providing labor beyond this point. Households receive a wage rate of $\mathcal{W}_t$ in exchange for a unit of labor. In equilibrium, households earn a share $1-\nu$ of output

as labor income:

$$\mathcal{W}_t = (1-\nu)AK_t.$$

Here I take this reduced-form wage equation as given. A simple microfoundation, based on Frankel (1962), is provided in Appendix A.1. In addition to diagnostic expectations, this labor income margin is where my model differs from HK.

Capital Goods Production Investment follows q-theory. There exists a capital goods producer who is responsible for undertaking capital investment. All profits from this producer are passed on to households. After capital is produced, it is sold directly to the intermediary sector at price q_t. The capital goods producer solves:

$$\max_{i_t} \quad q_t i_t K_t - \Phi(i_t, K_t),$$

which results in an equilibrium investment rate of

$$i_t = \delta + \frac{q_t - 1}{\xi}. \tag{13}$$

Equation (13) is important because it highlights the propagation of behavioral and financial frictions to the real economy. The economy's growth rate depends on investment rate i_t, which in turn depends on capital price q_t. To the extent that both financial and behavioral frictions affect the capital price q_t, these frictions will feed back into the endogenous growth rate of output.

2.5 Portfolio Choice and Asset Returns

Household Portfolio Choice Let W_t denote household wealth at time t. Households can invest in two assets: the debt and equity issued by financial intermediaries. Debt pays a risk-free return of r_t while equity pays a stochastic return of $d\widetilde{R}_t$. Reduced-from assumptions will now be made to ensure that intermediaries raise at least λW_t of debt funding from households. Households are not the focal-point of this model, and these simplifying assumptions will allow the equilibrium leverage of the financial sector to be regulated by exogenous parameter λ.

Each household is comprised of two members, an "equity member" and a "debt member." At the start of each period the household: (i) consumes, and (ii) splits its wealth W_t between the debt member and the equity member. A share λ of wealth is given to the debt member, who can only invest in the intermediaries' risk-free debt.[14] Share $1-\lambda$ of wealth is given to the equity member, who is free to invest in both the debt and the equity of the intermediary (but is not allowed to make levered investments). Investments pay off at time $t+dt$, and returns are pooled before this process is repeated at time $t+dt$.

The model will be calibrated such that the equity member always chooses to invest the maximum possible amount in F's equity in equilibrium.[15] Thus, equity members collectively invest their allocated wealth of $(1-\lambda)W_t$ in the intermediaries' equity, subject to the restriction that they do not purchase more than $\mathcal{E}_t$. If the capital constraint binds, equity members place their remaining wealth in bonds. I therefore define

$$E_t \equiv \min\{\mathcal{E}_t, W_t(1-\lambda)\} \tag{14}$$

as the total equity capital raised by the intermediary sector at time t.

The risk-free rate r_t is pinned down by the households' intertemporal optimization prob-

[14] One can think of λ as capturing household demand for liquid balances, though this is not formally modeled.

[15] This condition is verified as part of the model solution. For details, see Appendix E.3.

lem. Specifically,

$$r_t = \rho + \zeta \widehat{\mathbb{E}}_t \left[\frac{dc_t^y}{c_t^y} \right] - \frac{\zeta(\zeta+1)}{2} \widehat{Var}_t \left[\frac{dc_t^y}{c_t^y} \right], \tag{15}$$

where $\zeta = 1 - (1-\phi)(1-\gamma_h)$ can be interpreted as the inverse of the EIS.[16] Again, the hat-notation is used because expectations of the consumption process are diagnostic. Equation (15) is the standard consumption-based risk-free rate formula in continuous time.[17]

Intermediary Portfolio Choice Diagnostic expectations imply that households and intermediaries may not have correct beliefs about asset return processes. To start, I postulate that agents expect q_t and P_t to evolve according to the Itô processes:

$$\widehat{\frac{dq_t}{q_t}} = \widehat{\mu_t^q} dt + \widehat{\sigma_t^q} dZ_t \tag{16}$$

$$\widehat{\frac{dP_t}{P_t}} = \widehat{\mu_t^P} dt + \widehat{\sigma_t^P} dZ_t. \tag{17}$$

Perceived drift and volatility coefficients are endogenous and will be solved for in equilibrium.

Using price process (16), the return on an investment in productive capital is perceived to be:

$$\widehat{dR_t^k} = \frac{\nu A}{q_t} dt + \widehat{\frac{d(q_t K_t)}{q_t K_t}} - i_t dt = \left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta \mathcal{I} + \sigma \widehat{\sigma_t^q} \right) dt + \left(\sigma + \widehat{\sigma_t^q} \right) dZ_t. \tag{18}$$

The return on capital is comprised of a dividend component $\left(\frac{\nu A}{q_t} dt\right)$ and a capital gains

[16] Equating supply and demand in equilibrium, $c_t^h = 1$ since $H \equiv 1$. Setting $c_t^h = 1$ and plugging equation (11) into (10) gives $\widehat{\mathbb{E}} \left[\int_0^\infty e^{-\rho t} \frac{(c_t^y)^{(1-\phi)(1-\gamma_h)}}{1-\gamma_h} \right]$. Multiplying the value function by positive constant $\frac{1}{1-\phi}$ shows that we can think of the household as having CRRA preferences over just c_t^y, with a risk-aversion parameter of $\zeta = 1 - (1-\phi)(1-\gamma_h)$.

[17] Because the household sector is split into debt and equity members, additional assumptions are required in order to generate equation (15). In particular, if the household decides to save an additional dollar then that dollar must be invested in the risk-free bond. The benefit of making these assumptions is that it recovers the standard continuous-time formula for the risk-free rate. See footnote 5 of He and Krishnamurthy (2019) for details.

component $\left(\frac{\widehat{d(q_t K_t)}}{q_t K_t} - i_t dt\right)$. The dividend per unit of capital νA is the output that remains after labor wages are paid. Capital gains are accrued through growth in the price of capital and also through growth in the total of quantity of capital, excluding new investment made by capital goods producers.

Equation (18) illustrates how the perception of capital returns is biased. First, capital growth expectations are biased by $\theta \mathcal{I}_t$. Additionally, agents misperceive both the drift and volatility of price process dq_t.

Proceeding similarly, equation (17) can be used to derive the perceived return on an investment in housing:

$$\widehat{dR_t^h} = \frac{D_t}{P_t} dt + \frac{\widehat{dP_t}}{P_t} = \left(\frac{D_t}{P_t} + \widehat{\mu_t^P}\right) dt + \widehat{\sigma_t^P} dZ_t. \tag{19}$$

The dividend on housing is rental income D_t. Diagnostic expectations of capital growth create biased expectations of rental income growth. Since P_t is the present discounted value of future dividends, diagnostic expectations generate non-rational perceptions of price process dP_t.

Introducing some additional notation, let $\widehat{\pi_t^k} \equiv \left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta \mathcal{I} + \sigma \widehat{\sigma_t^q}\right) - r_t$ denote the perceived instantaneous risk premium on an investment in productive capital. Similarly, let $\widehat{\pi_t^P} \equiv \left(\frac{D_t}{P_t} + \widehat{\mu_t^P}\right) - r_t$ denote the perceived risk premium on an investment in housing. Using these definitions, equations (18) and (19) can be rewritten as follows:

$$\widehat{dR_t^k} = \left(\widehat{\pi_t^k} + r_t\right) dt + \widehat{\sigma_t^k} dZ_t$$
$$\widehat{dR_t^h} = \left(\widehat{\pi_t^h} + r_t\right) dt + \widehat{\sigma_t^h} dZ_t,$$

where $\widehat{\sigma_t^k} \equiv \sigma + \widehat{\sigma_t^q}$ and $\widehat{\sigma_t^h} \equiv \widehat{\sigma_t^P}$.

The banker makes portfolio choices in order to maximize her objective in (8). Let α_t^k and

α_t^h denote the intermediary sector's portfolio shares of capital and housing, respectively.[18] The intermediary's perceived return on equity is:

$$\widehat{d\widetilde{R}_t} = \alpha_t^k \widehat{dR_t^k} + \alpha_t^h \widehat{dR_t^h} + (1 - \alpha_t^k - \alpha_t^h) r_t dt. \tag{20}$$

From equation (8), the banker solves:

$$\max_{\alpha_t^k, \alpha_t^h} \left[r_t + \alpha_t^k \widehat{\pi_t^k} + \alpha_t^h \widehat{\pi_t^h} \right] - \frac{\gamma}{2} \left(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h} \right)^2, \tag{21}$$

which results in the optimality condition:

$$\frac{\widehat{\pi_t^k}}{\widehat{\sigma_t^k}} = \frac{\widehat{\pi_t^h}}{\widehat{\sigma_t^h}} = \gamma(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h}). \tag{22}$$

Equation (22) states that the intermediary chooses portfolio shares in order to equate the perceived Sharpe ratio on each asset to its risk aversion times its perceived portfolio risk. As the intermediary sector is required to bear additional risk on its portfolio, it will demand a higher Sharpe ratio as compensation.

Binding Constraints, Leverage, and Financial Crises Following HK, financial crises are defined as states in which the equity issuance constraint binds: $\mathcal{E}_t < W_t(1-\lambda)$. A binding constraint generates a sudden and dramatic increase in risk premia, which feeds through to lower asset prices and impair economic growth. When binding, intermediaries cannot raise sufficient equity capital and are forced to increase leverage in order to fund asset purchases. This increases balance sheet risk, which causes intermediaries to demand higher perceived Sharpe ratios as compensation.[19] Moreover, the binding constraint generates endogenous

[18] When $\alpha_t^k + \alpha_t^h > 1$ then the intermediary has a negative portfolio share in risk-free debt, meaning the intermediary is taking on leverage. The model will be calibrated such that the intermediary sector always takes on leverage in equilibrium.

[19] In particular, higher leverage implies higher α's on the RHS of equation (22).

amplification of negative capital shocks. This is because negative shocks cause the constraint to bind even more tightly, thereby increasing leverage and risk premia further. Finally, this endogenous amplification of shocks makes asset prices highly sensitive to e_t, which increases asset price volatilities $\widehat{\sigma_t^k}$ and $\widehat{\sigma_t^h}$. This too increases the Sharpe ratio demanded by the financial sector. All of these factors combine to generate large asset market crisis nonlinearities due to binding constraints on the intermediary sector's risk-bearing capacity.

Feedback from Behavioral Frictions to Financial Frictions The intermediary sector's capital capacity $\mathcal{E}_t$ evolves according to the realized return on equity $d\widetilde{R}_t$. However, equation (22) shows that assets are priced according to the perceived return process $\widehat{d\widetilde{R}_t}$. The difference between these two processes is:

$$\widehat{d\widetilde{R}_t} - d\widetilde{R}_t = \alpha_t^k(\widehat{dR_t^k} - dR_t^k) + \alpha_t^h(\widehat{dR_t^h} - dR_t^h).$$

To the extent that the perceived return process $\widehat{d\widetilde{R}_t}$ differs from the realized return process $d\widetilde{R}_t$, bankers will not be demanding the correct compensation for the risk that they bear. When bankers are excessively optimistic, future returns will tend to be lower than expected. These persistent forecast errors will cause the intermediary sector's capital capacity to deteriorate relative to expectations. Thus, *excessive optimism at time t feeds back to create future equity funding problems.* The reverse is true when bankers are excessively pessimistic about future returns.

2.6 Equilibrium

Definition 1. *(Diagnostic Expectations Equilibrium). A diagnostic expectations equilibrium is a set of prices* $\{q_t, P_t, D_t, r_t, \mathcal{W}_t\}$ *and decisions* $\{c_t^y, c_t^h, i_t, \alpha_t^k, \alpha_t^h\}$ *such that:*

1. *Given prices, decisions as specified by* (12), (13), (15), *and* (21) *are optimal under*

diagnostic expectations.

2. *The goods market and housing rental market clear:*

$$Y_t = AK_t = C_t^y + \Phi(i_t, K_t), \text{ and} \tag{23}$$
$$C_t^h = H \equiv 1.$$

3. *The equity issuance constraint is satisfied:*

$$E_t = \min\{\mathcal{E}_t, W_t(1-\lambda)\}$$

4. *Asset markets clear with intermediaries holding all capital and housing:*

$$q_t K_t = \alpha_t^k E_t, \text{ and} \tag{24}$$
$$P_t = \alpha_t^h E_t \tag{25}$$

5. *The total value of assets equals total household wealth:*

$$W_t = q_t K_t + P_t \tag{26}$$

It should be noted from the model of diagnostic expectations outlined in section 2.2 that diagnostic expectations generalize rational expectations — rationality is recovered by setting $\theta = 0$. This is formalized in the following definition, which will serve as a benchmark for later comparison.

Definition 2. *(Rational Expectations Equilibrium). A rational expectations equilibrium is a diagnostic expectations equilibrium in which* $\theta = 0$.

3 Solution and Calibration

3.1 Solution Strategy

In order to solve for an equilibrium I must characterize how any history of shocks $\{Z_s, s \leq t\}$ maps into equilibrium prices, asset allocations, and beliefs at time t such that (i) all agents maximize diagnostically-expected utility through their consumption and portfolio decisions, and (ii) markets clear. In particular, I consider Markov equilibria with state variables K_t, $\mathcal{E}_t$, and $\mathcal{I}_t$. K_t captures the overall size of the economy, $\mathcal{E}_t$ tracks the financial sector's capital capacity, and $\mathcal{I}_t$ gives the state of expectations relative to rationality. HK use K_t and $\mathcal{E}_t$ as state variables in their rational model. The innovation of this paper is to capture sentiment with state variable $\mathcal{I}_t$. When expectations are diagnostic it is not enough to know the current state of the economy (K_t and $\mathcal{E}_t$); extrapolation means that one must also know the path that the economy took to get to that point ($\mathcal{I}_t$).

Following HK, the solution can be simplified further by scaling the economy by K_t.[20] Specifically, let

$$e_t \equiv \frac{\mathcal{E}_t}{K_t}.$$

e_t captures the equity issuance capacity of the intermediary sector relative to the size of the economy as a whole. I will look for price functions of the form $p_t = \frac{P_t}{K_t} = p(e_t, \mathcal{I}_t)$ and $q_t = q(e_t, \mathcal{I}_t)$. The ability to scale the economy by K_t allows me to numerically solve for an equilibrium as a function of only two state variables: e_t and $\mathcal{I}_t$. e_t characterizes the severity of financial frictions and $\mathcal{I}_t$ characterizes the state of beliefs relative to rationality.

[20] This is possible because P_t is linear in K_t. The price of housing P_t is equal to the discounted value of future housing dividends. In equilibrium, the dividend on housing is given by $D_t = \frac{\phi}{1-\phi}\left[A - i_t + \frac{\xi}{2}(i_t - \delta)^2\right] K_t$, which is linear in K_t. Details of this dividend formula are provided in Appendix A.3.

Comparing Diagnostic and Rational Expectations Equilibria Within the class of Markov equilibria considered here, rational expectations equilibria (REE) can be neatly contextualized within the class of diagnostic expectations equilibria (DEE).[21] This is formalized in the following proposition:[22]

Proposition 1. *For any diagnostic expectations equilibrium that is Markov in* $\{K, e, \mathcal{I}\}$*, the price and policy functions for* $\{K, e, \mathcal{I} = 0\}$ *compose a rational expectations equilibrium that is Markov in* $\{K, e\}$*.*

Proof. When $\mathcal{I}_t = 0$, equation (6) implies that agents also perceives that $\widehat{d\mathcal{I}}_t = 0$ for all future t. So, when $\mathcal{I}_t = 0$ then all agents will make decisions that are optimal under the perception that $\mathcal{I} = 0$ in perpetuity. Decisions that are optimal when $\mathcal{I} = 0$ in perpetuity must also be optimal when $\theta = 0$ (REE), because in both cases $\mathcal{I}$ is perceived to have no further effect on the resulting equilibrium. □

Proposition 1 shows that every diagnostic expectations equilibrium nests its corresponding rational expectations equilibrium. This property makes comparisons between the DEE and the REE very simple, as one gets the REE "for free" when solving for the DEE.

One corollary of Proposition 1 is that the REE characterizes the long-run expectations of state variables e_t and $\mathcal{I}_t$ held by diagnostic agents.

Corollary 1. *The ergodic distribution over* $(e, \mathcal{I})$ *that is diagnostically expected by all agents in the model is equivalent to the stationary distribution over* $(e, \mathcal{I})$ *that would arise in the corresponding REE.*

Proof. Regardless of $\mathcal{I}_t$, all agents perceive that $\mathcal{I}_t$ decays deterministically to zero at rate κ (equation (6)). In the long run diagnostic agents therefore perceive that $\mathcal{I} = 0$, at which point Proposition 1 obtains. □

[21] While I have not found multiple equilibria in my numerical analysis, no claims of equilibrium uniqueness are made in this paper.

[22] A similar result should hold if one were to use BGS' discrete-time formulation of diagnostic expectations in a dynamic model.

Boundary Conditions Boundary conditions are needed to solve for price functions $q(e_t, \mathcal{I}_t)$ and $p(e_t, \mathcal{I}_t)$. As $e_t \to \infty$ financial frictions disappear and the equity issuance constraint ceases to affect the equilibrium price and policy functions. The scale invariance properties of the model imply that $\lim_{e_t \to \infty} q_e(e_t, \mathcal{I}_t) = \lim_{e_t \to \infty} p_e(e_t, \mathcal{I}_t) = 0$. In words, the price functions asymptote to a constant value in e_t as the financial sector moves further from its funding constraint.

A lower boundary is imposed by assuming that bankers enter the intermediary sector deep in crisis times. This is captured by the term $d\psi_t$ in equation (9). The model assumes that there exists an exogenous minimum reputation level $\underline{e}$ such that new bankers enter the market whenever e_t reaches $\underline{e}$. Thus, state variable e_t has a reflecting barrier at $\underline{e}$. Loosely, this captures either government intervention deep in crises or entry into the skill-intensive finance industry when the value of entry (the Sharpe ratio) is high enough. It is costly to create new intermediaries because new bankers must acquire the skills to operate capital. Specifically, the economy must destroy $\beta > 0$ units of capital K_t in order for entry to increase aggregate reputation parameter $\mathcal{E}_t$ by one unit.[23]

The assumption that banker entry imposes a reflecting barrier at $\underline{e}$ pins down boundary conditions for asset prices $q(e_t, \mathcal{I}_t)$ and $p(e_t, \mathcal{I}_t)$ at $e = \underline{e}$. Specifically, price functions q and P must have a zero derivative with respect to e_t at $\underline{e}$.[24] This implies $q_e(\underline{e}, \mathcal{I}_t) = 0$ and $p_e(\underline{e}, \mathcal{I}_t) = \frac{p(\underline{e}, \mathcal{I})\beta}{1+\underline{e}\beta}$. Details are provided in Appendix A.2.

3.2 Calibration

He and Krishnamurthy (2019) build a standard RBC model augmented with a financial intermediary sector. I introduce three new parameters to the HK model. These are the two

[23] The capital evolution equations (2) and (5) are altered at $\underline{e}$ to include this form of capital destruction.

[24] If this were not the case, an arbitrageur could make unbounded profits (almost surely) by betting on a unidirectional change in asset prices at the reflecting barrier. Note that the instantaneous price shock is of order $\sqrt{dt}$ while the risk-free rate an arbitrageur borrows at is only of order dt. Therefore the asset price shock dominates borrowing costs almost surely.

behavioral parameters θ and κ, and the labor income parameter ν. The model in this paper nests HK, and one can recover the HK model by setting $\theta = 0$ and $\nu = 1$.

The macroeconomic model behaves like a standard RBC model when e_t is far from the constraint, though frictions become quantitatively important near the crisis region. Here I follow HK in defining $e_{distress}$ as the 33^{rd} percentile value of e_t in the model's stationary distribution. $e_{distress}$ delineates "normal" periods from "distress" periods. In the calibration to follow, $e_{distress} = 0.65$. Parameters corresponding the model's RBC elements will be calibrated in the part of the state space where $e_t > e_{distress}$.

Table 1 presents the baseline calibration, and details are discussed below. When possible, I follow the parameter choices and/or calibration targets of He and Krishnamurthy (2019). In Table 1, parameter values that are marked with an asterisk in the "Choice" column are equivalent to the parameter values of HK. Asterisks in the "Target" column indicate parameters for which the value differs from HK, but the calibration target is the same. The only parameters for which neither the value nor the target aligns with HK are the three new parameters θ, κ, and ν.

RBC Parameters The household discount rate ρ, depreciation rate δ, and adjustment cost ξ are relatively standard RBC parameters. My calibration follows HK.

I set $A = 0.425$ and $\nu = 0.315$. Parameters A and ν are calibrated jointly to target the investment-to-capital ratio and the investment-to-output ratio. HK target an investment-to-capital ratio of approximately 9% in the non-distress states (note that investment rate i_t is the investment-to-capital ratio at time t). Given the calibration of $\delta = 0.1$, equation (13) shows that $q_t \approx 1$ is necessary to match this calibration target. The dividend on capital is $\frac{\nu A}{q_t}$, so ν and A must be jointly calibrated to generate a dividend commensurate with $q_t \approx 1$. To separately identify A and ν, I also target an investment-to-output ratio of 20%. Investment-to-output equals $\frac{i_t K_t}{A K_t}$, so the calibration that $i \approx 9\%$ pins down A accordingly.

	Parameter	Choice	Target
Panel A: Intermediation Parameters			
γ	Banker risk aversion	2^*	Mean non-crisis Sharpe ratio
λ	Debt ratio	0.75^*	Intermediary leverage
η	Banker exit rate	0.147	Prob(Crisis)*
$\underline{e}$	Entry barrier	0.082	Max $\mathcal{I} = 0$ Sharpe ratio*
β	Entry cost	2.8^*	Land price volatility
Panel B: Technology Parameters			
σ	Capital shock	$3\%^*$	Consumption volatility
δ	Depreciation rate	$10\%^*$	Literature
ξ	Adjustment cost	3^*	Literature
A	Productivity	0.425	Investment-capital ratio*
ν	Capital share	0.315	Investment-output ratio
Panel C: Household Preference Parameters			
ρ	Time discount rate	$2\%^*$	Literature
ζ	1/EIS	0.7	Interest rate volatility*
ϕ	Housing share	0.1675	Housing-to-wealth ratio*
Panel D: Diagnostic Expectations Parameters			
κ	Decay of new information	0.139	Slow-moving beliefs
θ	Diagnosticity	0.132	Magnitude of expectations bias
Unconditional Simulated Moments			
	Probability of Crisis		3.15%
	Volatility(r_t)		0.93%
	Volatility(Land Price Growth)		12.09%
Non-Distress Simulated Moments			
	Mean $\left(\frac{\text{Investment}}{\text{Capital}}\right)$		9.61%
	Mean $\left(\frac{\text{Investment}}{\text{Output}}\right)$		22.60%
	Mean (perceived Sharpe ratio)		52.01%
	Mean $\left(\frac{\text{Housing Wealth}}{\text{Total Wealth}}\right)$		46.07%
	Volatility(Consumption Growth)		2.32%
	Volatility(Investment Growth)		5.24%
	Volatility(Output Growth)		2.93%

Table 1: Baseline model calibration. The model is simulated at a quarterly frequency, and simulated growth rates are computed as log changes from $t-2$ to $t+2$.

Comparing this to the HK calibration, HK set $A = 0.133$. The HK model has no labor income, which is equivalent to shutting down the labor margin by setting $\nu = 1$.[25] $A = 0.133$ implies that HK have an investment-to-output ratio of approximately 67% and a consumption-to-output ratio of less than 30%. An important consequence of these counterfactual ratios is that the HK model struggles to match empirical moments on investment and consumption volatility, particularly in periods of financial distress. The interest rate depends on consumption volatility (equation (15)), and HK calibrate $\zeta = 0.13$ (EIS > 7) in order to prevent excessive interest rate volatility. I depart from the HK calibration because, as will be detailed shortly, the EIS becomes particularly important when growth expectations are biased. Though I sacrifice some parsimony of the original HK model by introducing a simple labor income margin, the benefit of doing so is that it allows my model to more accurately match the investment and consumption moments observed empirically.

Finally, parameter σ governs the volatility of exogenous shocks to capital. As in HK, I set $\sigma = 3\%$. HK report that from 1975 to 2015 the volatility of investment growth in non-distress periods was 5.79%, and the volatility of consumption growth was 1.24%. In the non-distress region of the model, setting $\sigma = 3\%$ generates an output growth volatility of 2.93%, an investment growth volatility of 5.24%, and a consumption growth volatility of 2.32%. The model generates too much consumption volatility and too little investment volatility, with the choice of $\sigma = 3\%$ attempting to strike a balance between these two inaccuracies.

Intermediation Parameters Parameter γ represents the bankers' risk aversion. I follow HK in setting $\gamma = 2$. This implies that bankers demand an average perceived Sharpe ratio of 0.52 in non-distress states.[26] The closest empirical counterpart is He et al. (2017), who

[25] My calibration of $\nu A = 0.315 \times 0.425 = 0.134$ yields a very similar dividend to the HK calibration of $\nu A = 1 \times 0.133 = 0.133$.

[26] Equation (22) shows that bankers demand a perceived Sharpe ratio equal to γ times their perceived portfolio risk. Thus, for any asset A the banker can potentially make a mistake both on the perceived Sharpe ratio of A (LHS of (22)) and also on the perceived portfolio risk (RHS of (22)). The bias in the perceived Sharpe ratio of A will depend on how diagnostic capital growth expectations affect the cash-flow

find that the capital ratio of financial intermediaries robustly prices cross-sectional expected returns for a variety of asset classes. They estimate a realized Sharpe ratio of 0.48 for assets intermediated by the financial sector.

Parameter λ controls the leverage of the financial sector when the capital constraint doesn't bind. Since the financial sector has assets of $P_t + q_t K_t = W_t$ and equity of E_t, equation (14) gives a market leverage value of $\frac{W_t}{E_t} = \frac{1}{1-\lambda}$ in non-crisis states. Again following HK, I set $\lambda = 0.75$ which generates a leverage ratio of 4 in non-distress states.

Crisis Parameters Financial crises are defined as states in which the equity issuance constraint binds. Banker exit rate parameter η is chosen to target a 3% crisis probability. I target the same crisis probability as HK, who observe that there have been three major U.S. financial crises in the past 100 years.

Parameters $\underline{e}$ and β control the lower boundary condition. In their model without sentiment, HK set $\underline{e}$ such that the Sharpe ratio at $\underline{e}$ is 6.5. To align this model's calibration with HK, I set $\underline{e}$ such that the perceived Sharpe ratio at $\underline{e}$ and $\mathcal{I} = 0$ is 6.5. $\underline{e}$ is set low enough that entry occurs very rarely, and therefore the precise choice of $\underline{e}$ will not quantitatively affect the dynamics that are the focus of this paper.

Parameter β governs the volatility of housing price P_t. This is because β determines the slope of P_t at the lower boundary in the crisis state, which in turn affects the slope of P_t throughout the distress region. HK estimate that the empirical volatility of land price growth has been 11.9% from 1975 to 2015. With $\beta = 2.8$, the volatility of land price growth in the model is 12.1%.

Household Parameters Parameter ϕ determines the relative value of housing services to the consumption good, which in turn pins down the rental rate D_t (see equation (12)). D_t is also the dividend on a housing investment, and P_t is the discounted value of these

expectations of asset A. This will be asset-specific.

future dividend flows. As in HK, I set ϕ to target a non-distress housing-to-wealth ratio of approximately 45%. This requires $\phi = 0.1675$.[27]

ζ is the inverse of the EIS, and determines the responsiveness of interest rates to expected consumption growth and volatility. When expectations are diagnostic, agents misperceive the growth rate of consumption. This implies that ζ governs the sensitivity of interest rates to variation in $\mathcal{I}_t$.

The reason that ζ plays an important role when expectations are diagnostic is that ζ modulates the extent to which sentiment gets incorporated into asset prices. When $\zeta = 1$, any bias in consumption growth expectations is passed one-for-one into the risk-free discount rate r_t (see equation (15)). An important implication is that when $\zeta = 1$, asset prices q_t and P_t are constant in $\mathcal{I}_t$. All bias in cash-flow expectations is exactly offset by the risk-free discount rate, leaving asset prices completely unresponsive to sentiment. When $\zeta < 1$, the interest rate responds less than one-for-one to changes in expected consumption growth. In this case, q_t and P_t are increasing in $\mathcal{I}_t$.

In the baseline calibration I set $\zeta = 0.7$. As in HK, I target a real interest rate volatility of approximately 1%. Appendix Figure 6 plots the ergodic distribution of the real interest rate in the baseline calibration. Since $\zeta < 1$, asset prices are increasing in $\mathcal{I}_t$. This will be important for generating the results in section 6.

Behavioral Parameters θ governs the extent to which expectations are biased by representativeness and κ governs the speed at which new information decays from $\mathcal{I}_t$. These two parameters need to be calibrated jointly. As an Ornstein-Uhlenbeck process, $\mathcal{I}_t$ has a Gaussian unconditional distribution with variance $\frac{\sigma^2}{2\kappa}$. This implies that the unconditional variance of the capital growth bias is $\frac{\theta^2\sigma^2}{2\kappa}$. Low values of κ (slow decay) must be accompanied

[27] HK set $\phi = 0.6$. My calibration has a much higher A than HK, meaning that my model generates a larger share of output goods relative to housing services. In order to prevent this from driving up the rental rate D_t paid on housing, my calibration requires a lower value of ϕ.

by low values of θ to prevent growth expectations from being unrealistically biased.

The existing literature offers little guidance for calibrating κ and θ. I pick values that seem reasonable, and parameter choices are intentionally made to highlight certain mechanisms in the model. I set $\kappa = 0.139$, equivalent to a half-life of 5 years.[28,29] This slow-moving sentiment is intended to capture prolonged periods of relatively positive and negative news, such as the Great Moderation, rather than high-frequency volatility.[30] Given κ, I set $\theta = 0.132$. θ is calibrated such that one standard deviation in $\mathcal{I}$ corresponds to an output growth bias of 0.75 p.p. Under this calibration, output growth expectations are biased by less than 1.5 p.p. 95% of the time. While this level of θ certainly makes expectations non-trivially biased, expectations are also not grossly irrational. Robustness to the results that follow is discussed in Appendix B.

Appendix F.1 outlines relevant properties of Ornstein-Uhlenbeck processes and provides numerical examples under the baseline calibration. Specifically, Appendix F.1 characterizes the conditional distribution of $\mathcal{I}_{t+\tau}$ given $\mathcal{I}_t$. It also characterizes the persistence of sentiment by providing a closed-form solution for the distribution of first hitting times, defined as the first time τ at which $\mathcal{I}_{t+\tau}$ crosses zero.

[28] This means 50% of the weight for $\mathcal{I}_t$ is placed on shocks that occurred in the past 5 years, 25% of the weight is placed on the shocks that occurred 5 to 10 years ago, 12.5% of the weight is placed on shocks that occurred 10 to 15 years ago, etc.

[29] Using a different specification than this paper, Bordalo et al. (2019a) estimate that diagnostic expectations incorporate the past 3 years of shocks. Wachter and Kahana (2019) is another behavioral model featuring expectations that are formed according to a slow-moving context. They calibrate their background context to have a half-life of just over 3 years, and argue this is in the range typically estimated in the memory literature. Using survey data from recent homebuyers, Case et al. (2012) argue that slow-moving long-term expectations were the critical driver of the 2000s housing bubble. Similarly, Malmendier and Nagel (2011) document that the financial market returns experienced over ones lifetime are predictive of future risk taking and beliefs about future returns.

[30] The method can easily be extended such that sentiment contains both a slow-moving component and a high-frequency component. However, this requires an additional state variable. Details are included in Appendix F.2.

4 Global Solution: Asset Prices, Policy Functions, and the Stationary Distribution

4.1 Prices, Policy Functions, and Forecast Errors

Select prices, policy functions, and forecast errors for the DEE under the baseline calibration are shown in Figure 2 (for more, see Appendix Figure 7). The horizontal axis lists (scaled) capital capacity $e_t = \frac{\mathcal{E}_t}{K_t}$. All panels plot three curves. The blue curve corresponds to depressed sentiment ($\mathcal{I}_t = -1.5SD$), the red curve corresponds to neutral sentiment ($\mathcal{I}_t = 0$), and the yellow curve corresponds to elevated sentiment ($\mathcal{I}_t = +1.5SD$).

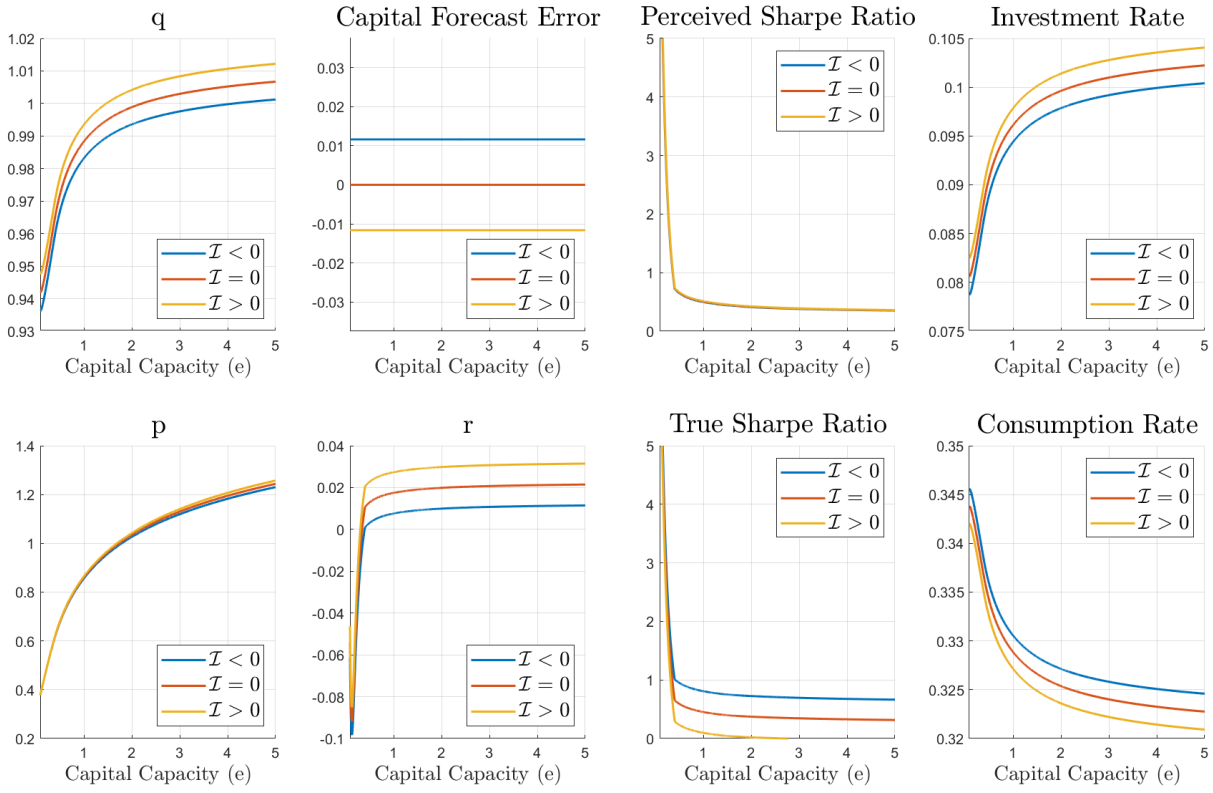


Figure 2: Selected price and policy functions under the baseline calibration. Each panel plots three curves, corresponding to $\mathcal{I} = -1.5SD$ (blue), $\mathcal{I} = 0$ (red), and $\mathcal{I} = +1.5SD$ (yellow).

The two leftmost panels of Figure 2 plot asset prices q_t and p_t. To understand the effect

of financial and behavioral frictions, it is instructive to review two benchmark economies for comparison: a "no financial frictions" economy (no equity issuance constraint) and a "no behavioral frictions" economy (rational expectations). In the "no financial frictions" benchmark, the scale-invariance properties of this model imply that the price functions are horizontal lines. In the rational expectations benchmark, a direct application of Proposition 1 gives that prices q_t and p_t are represented by only the red line ($\mathcal{I} = 0$).

Starting with the effect of financial frictions on asset prices, the q and p panels show that asset prices are sensitive to the intermediary sector's funding capacity e_t. In the crisis region (approximately $e_t < 0.4$), constraints on intermediary risk-bearing capacity cause asset prices to plummet. The Sharpe ratio panels similarly illustrate the nonlinear spike in risk premia that characterizes crisis times. Moving away from the constrained region, asset prices quickly rise as the intermediary sector restores its risk-bearing capacity. As e_t continues to increase, asset prices eventually asymptote to their value in the "no financial frictions" benchmark.

Asset prices exhibit what HK refer to as *anticipation effects*, meaning that asset prices start to fall well before the equity issuance constraint actually binds. These anticipation effects arise because bankers are not willing to pay a high price for an asset at time t if there is a chance that a binding constraint sometime in the future will cause that asset's price to crash.[31] One important implication is that financial frictions affect financial market and macroeconomic dynamics well before the constraint actually binds; the anticipation of a binding constraint at some point in the future is enough to drag down current asset prices. It is these anticipation effects that underlie the "financial distress" region of the model (defined

[31] In the model, the financial sector's diagnostic expectations of future risk premia will be low when asset prices are high, and vice-versa. This contradicts the evidence presented in Greenwood and Shleifer (2014), which shows that investor return expectations are negatively correlated with model-implied returns. However, Greenwood and Shleifer (2014) focus predominantly on the expectations of households whereas in the model, it is the beliefs of financial intermediaries that are relevant for pricing assets. Adam et al. (2018) provide evidence that professional investors have excess return expectations that covary negatively with the P/D ratio, while individual investors have excess return expectations that covary positively with the P/D ratio.

as e_t below the 33^{rd} percentile), where constraints don't necessarily bind but financial friction effects are still present.

The influence of diagnostic expectations on asset prices is illustrated by the separation of the yellow curve and the blue curve from the red curve. Holding e_t fixed, asset prices are increasing in $\mathcal{I}_t$. For capital, excessive optimism about capital growth increases the expected capital gains to be accrued from an investment in capital at time t. For housing, excessive optimism about the growth of output creates over-optimism about expected future rental payments (dividends). In both cases, elevated sentiment increases asset prices because bankers are willing to pay more in order to access what they expect to be higher future cash flows.

Examining expectations in further detail, the Capital Forecast Error panel shows diagnostic agents' instantaneous forecast error for the capital growth rate: $\mathbb{E}\frac{dK_t}{K_t} - \widehat{\mathbb{E}}\frac{dK_t}{K_t} = -\theta\mathcal{I}_t$. Capital is the fundamental upon which diagnostic expectations are based. In this dynamic equilibrium model, non-rational capital expectations also have broader equilibrium implications. For example, forecast errors for endogenous returns can be seen by comparing the Perceived Sharpe Ratio panel to the True Sharpe Ratio panel. Bankers make portfolio decisions according to the perceived Sharpe ratio. However, asset mispricing implies that forecast errors of capital growth are positively related to Sharpe ratio forecast errors. In equilibrium, diagnostic expectations transform the risk-return tradeoff faced by bankers, with excessive optimism eroding the realized risk premium earned by the financial sector.

The Investment Rate and Consumption Rate panels of Figure 2 illustrate the broader macroeconomic implications of financial and behavioral frictions. Recall that i_t is an increasing function of q_t, and that the growth rate of output depends on the endogenous rate of investment. The model with neither financial frictions nor behavioral frictions would feature a constant capital price q_t and therefore a constant rate of investment. Thus, *all variation in investment is due to financial and behavioral frictions.* Specifically, the Investment panel

illustrates that investment is low whenever either e_t is low or $\mathcal{I}_t$ is low.[32] Correspondingly, the Consumption Rate panel shows that the (scaled) rate of consumption $\frac{C_t^y}{K_t}$ is high whenever e_t is low or $\mathcal{I}_t$ is low. This follows from output market clearing in (23). For a given level of output, if investment is low then consumption must be high in order to clear the output market.

4.2 Ergodic Distribution

Figure 3 plots the economy's stationary distribution over e_t and $\mathcal{I}_t$. The ergodic distribution is solved for numerically using a *Kolmogorov forward equation*.[33] The economy is more likely to be in lighter-colored regions, while dark blue regions are rarely encountered. The dashed gray line marks where the constraint binds. The stochastic steady state occurs at $e_t = 0.87$ and $\mathcal{I}_t = 0$.

The unconditional correlation between e_t and $\mathcal{I}_t$ is approximately 0.7. This strong positive correlation arises in equilibrium because e_t and $\mathcal{I}_t$ both load positively on the same capital shocks. Consider a positive shock σdZ_t. The financial sector uses leverage to take a long position in capital, so a positive shock generates excess returns for the financial sector. This increases reputation and moves the financial sector away from its funding constraint. The same positive shock σdZ_t also represents positive new information. Positive new information makes future states with high capital more representative, thus increasing sentiment.

5 Financial Market Implications: Expectations and Crises

Now that the solution has been described, I proceed to illustrate the implications of the interaction that exists between behavioral frictions and financial frictions. To begin, I focus

[32] As was discussed in the calibration section, q_t is only increasing in $\mathcal{I}_t$ when $\zeta < 1$. However, even when $\zeta = 1$ variation in $\mathcal{I}_t$ still alters the evolution of e_t. In this way, behavioral frictions will still have an indirect effect on asset prices through the effect of sentiment on the evolution of e_t.

[33] Details are provided in Appendix E.2.

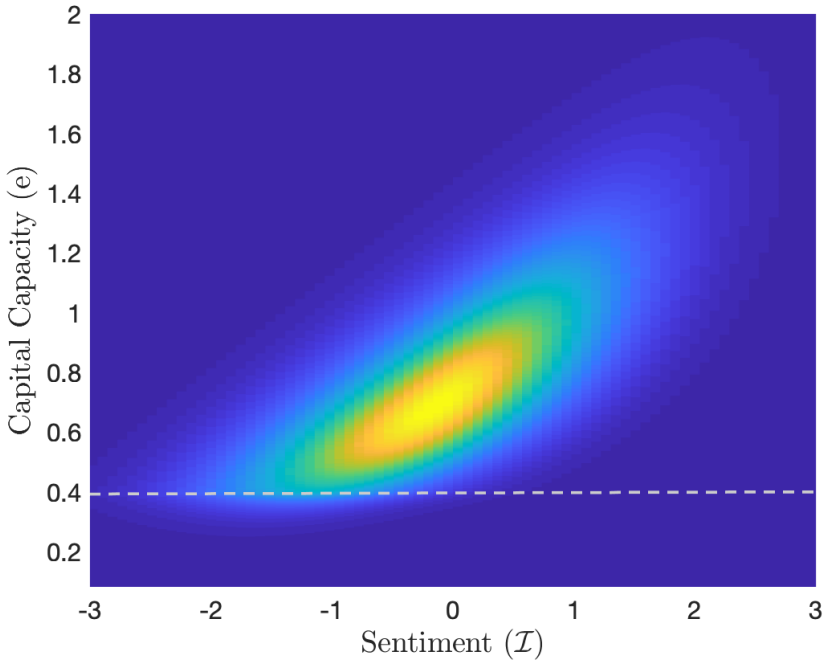


Figure 3: Ergodic distribution over state variables e and $\mathcal{I}$ under the baseline calibration. Sentiment is reported in standard deviation units. The gray dashed line marks the boundary of the crisis region.

on the financial market implications of this interaction.

5.1 Expectations and Financial Crises

5.1.1 Sentiment Conditional on Financial Distress

Elevated sentiment plays a direct role in triggering the amplification of financial distress into full-blown financial crises. This is in contrast to the REE, where state variable e_t alone is sufficient for characterizing the probability of future financial crises. To show this result, I conduct the following experiment. Assume that at time t the economy has a capital capacity of $e_t = e_{distress}$ (the 33$^{\text{rd}}$ percentile of e_t). For each level of sentiment $\mathcal{I}_t$, I calculate both the true and the diagnostically-expected probability that the economy finds itself in a crisis at some point in the next 1, 2, or 5 years.

These "hitting probabilities" can be calculated numerically using a *Kolmogorov backward*

equation.[34] Figure 4 graphs the results. The solid lines in Figure 4 plot the true probability of a crisis over different horizons. The dashed lines plot the probability of a crisis that is perceived by agents with diagnostic expectations. Following from Proposition 1, the dots intersecting the dashed line at $\mathcal{I}_t = 0$ indicate the crisis probabilities in the corresponding REE. Sentiment on the horizontal axis is reported in standard deviation units.

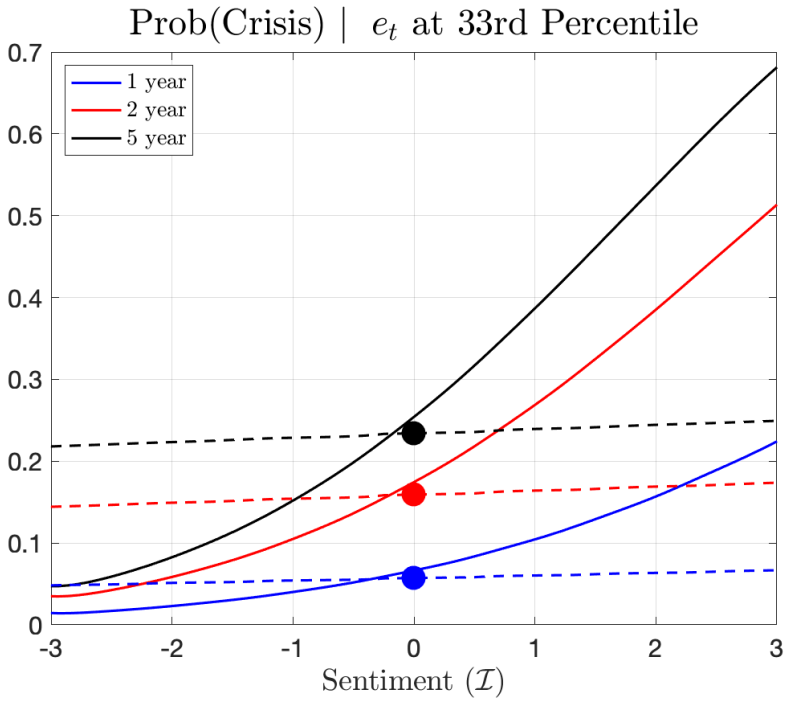


Figure 4: Crisis Hitting Probabilities. Solid lines report true probabilities in the DEE and dashed lines report perceived probabilities in the DEE. The dots indicate the crisis probabilities in the corresponding REE.

The perceived crisis probabilities (dashed lines) are all quite moderate.[35] The reason that crisis perceptions are low is that financial markets have a built-in defense mechanism – a rising risk premium – to combat financial crises. This is another way of stating the anticipation effects seen in Figure 2. As e_t moves closer to the crisis region the financial sector responds by demanding a larger risk premium on capital and housing. This increases

[34] Details are provided in Appendix E.2.

[35] As sentiment increases there is a mild upward slope in the perceived hitting probability. The reason for this is that a crisis occurs when $e_t < (1-\lambda)(p_t + q_t)$ (dividing equation (14) by K_t). Since p_t and q_t are increasing in $\mathcal{I}_t$, the RHS of this inequality is increasing in sentiment. This means that a given level of e_t (here, $e_t = e_{distress}$) is closer to the crisis region when $\mathcal{I}_t$ is high.

the return on the financial sector's equity, allowing bankers to rebuild their reputation and move away from the constraint. All agents in the model understand this defense mechanism, and therefore perceive that it takes a rare series of persistently negative shocks to move the financial sector from the distress region into the crisis region.

Perceptions are correct in the REE, so this same argument describes why the REE (dots) features a low probability of financial distress devolving into a crisis. The anticipation of a possible future crisis generates rising risk premia, which in turn quells the build-up of systemic risk in the first place. Thus, in this calibrated model the REE is unable to generate low risk prices concurrent with high crisis probabilities.

When expectations are diagnostic, elevated sentiment breaks the intermediary sector's crisis defense mechanism and sets the stage for future financial crises. When $\mathcal{I}_t > 0$ and expectations are excessively optimistic, bankers perceive that they are demanding a larger risk premium than they truly are. In particular, when $\mathcal{I}_t > 0$ then bankers borrow at an elevated interest rate and pay high prices to purchase capital and housing, which they are willing to do only because they have excessively-optimistic forecasts of the future cash flows offered by these investments. Bankers perceive that they are demanding an adequate risk premium for the risk on their balance sheet, but in reality they are not.[36] This implies that a more moderate series of negative shocks is sufficient to move the economy into the crisis region, explaining the neglected crash risk that occurs when $\mathcal{I}_t > 0$.

The reverse story explains why perceived crisis probabilities are inflated when $\mathcal{I}_t < 0$. In this case, excessive pessimism about future cash flows implies that the financial intermediary sector borrows at low interest rates and purchases assets cheaply. When cash flows end up being larger than expected, the financial sector makes returns in excess of expectations and finds itself quickly rebuilding its reputation.

[36]This is not to necessarily say that the risk premium is negative, only that it is not as large as perceived.

Stress Testing Sentiment-driven amplification of systemic risk is further illustrated through financial intermediary stress tests. In these stress tests I calculate the path of shocks required to generate wide-scale losses in the financial sector under different levels of initial sentiment. The financial sector is started with a capital capacity of $e_t = e_{distress}$ and sentiment $\mathcal{I}_t \in \{-2SD, -1SD, 0, +1SD, +2SD, \text{REE}\}$. Conditional on each starting location, a constant shock is fed into the system over the following 8 quarters in order to generate the 2-year bank equity loss listed in the left-hand column. Table 2 reports the 8-quarter compounded shock necessary for each level of bank equity losses.

2-yr. Bank Capital	$\mathcal{I}_t = -2SD$	$\mathcal{I}_t = -1SD$	$\mathcal{I}_t = 0$	$\mathcal{I}_t = +1SD$	$\mathcal{I}_t = +2SD$	REE
-2 %	-2.0 %	-1.4 %	-0.7 %	0.0 %	0.8 %	-0.8 %
-5 %	-3.4 %	-2.7 %	-1.9 %	-1.2 %	-0.4 %	-2.1 %
-10 %	-5.5 %	-4.8 %	-4.0 %	-3.2 %	-2.4 %	-4.1 %
-15 %	-7.6 %	-6.8 %	-6.0 %	-5.1 %	-3.8 %	-6.1 %
-25 %	-9.8 %	-8.6 %	-7.4 %	-6.2 %	-5.0 %	-7.3 %

Table 2: Bank Equity Losses and Sentiment.

As sentiment increases, smaller shocks are sufficient to generate a given level of capital losses. This is particularly true for large bank capital losses. Thus, elevated sentiment generates large, hidden, vulnerabilities in the financial sector which will not be reflected in the market price of risk.

To quantify the probability of these shock scenarios, note that 8-quarter compounded shocks are distributed approximately Lognormal$(0, 2\sigma^2)$. Consider the 25% equity loss scenario. When $\mathcal{I}_t = -2SD$ there is a 0.8% percent chance of realizing a series of shocks worse than -9.8%. When $\mathcal{I}_t = 0$, there is a 3.5% chance of realizing a series of shocks worse than -7.4%. When $\mathcal{I}_t = +2SD$, there is an 11.4% chance of realizing a series of shocks worse than -5.0%. These stress tests show that sentiment plays an essential role in predicting future financial crises given initial distress. These results also underscore the inability for rational expectations to generate predictable financial crises, as it is the breakdown of anticipation

effects driven by over-optimism that amplifies crisis risk.

Empirical Evidence Most of the empirical evidence on the triggers of financial crises studies the predictive power of growth in credit from the banking sector and/or credit spreads. Neither credit growth nor credit spreads have a direct counterpart in the model.[37] To map credit growth into the model, I follow Krishnamurthy and Muir (2017) who argue that credit growth is an empirical proxy for financial fragility. Since the purpose of state variable e_t is precisely to capture financial fragility, I will let e_t be the model's counterpart to credit growth.

To map credit spreads into the model, equation (22) establishes an asset pricing equation that will price any asset such that the financial sector has no incentive to take a net position in that asset. Details are included in Appendix D. The reduced-form result is that credit spreads are decreasing in both e_t and $\mathcal{I}_t$. Spreads are decreasing in e_t due to variation in the risk-bearing capacity of the financial sector. Spreads are decreasing in $\mathcal{I}_t$ because higher output growth expectations are assumed to lower the expected default rate of risky bonds.

Turning now to the empirical literature, the predictions of the DEE align closely with the empirical findings of Krishnamurthy and Muir (2017). Krishnamurthy and Muir (2017) study an international panel of credit spreads and macroeconomic outcomes and find that pre-crisis credit spreads are "too low" to be explained by rational models. Specifically, after controlling for credit growth they estimate that credit spreads are 23% lower in the five years before a financial crisis. In the model, controlling for credit growth means fixing e_t. Then, further conditioning on a future crisis is equivalent to conditioning on higher average values of $\mathcal{I}_t$. Since spreads are decreasing in $\mathcal{I}_t$, the model is able to replicate the finding that spreads are "too low" before financial crises.

[37] There is no margin for credit growth from the financial sector because businesses and financial intermediaries are collapsed into a single "intermediary sector" which holds capital directly. The model also has no credit spreads because there is only one type of bond, and it is risk-free.

Additionally, Krishnamurthy and Muir (2017) find that it is the interaction of low credit spreads and high credit growth that best predicts future financial crises. The model is again consistent with this empirical finding. A low value of e_t (high credit growth) is a necessary condition for a future financial crisis but, as Figure 4 and Table 2 highlight, the predictive power of e_t is significantly enhanced when combined with high values of $\mathcal{I}_t$ (low spreads). Further, the model suggests a mechanism through which financial market froth predicts future crises — the breakdown of anticipation effects due to asset mispricing.

The model is also consistent with direct evidence of neglected risk before the 2007-2008 Financial Crisis. Since October 1989, Robert Shiller and the Investor Behavior Project at Yale University have surveyed institutional investors to establish a U.S. Crash Confidence Index.[38] From October 1989 – December 2018, the Crash Confidence Index's mean value was 35.33%. The Index peaked at 57.95% in April 2006. In July 2008, just two months before the Lehman failure, the Crash Confidence Index was still at 38.37%. Foote et al. (2012) uncover a Lehman Brothers analyst report from 2005 documenting expectations for a variety of house price growth scenarios. Analysts assigned a 5% probability to the worst-case "Meltdown" scenario of -5% home price growth over the subsequent 3 years, followed by 5% home price growth thereafter. The realized path of home prices was significantly worse. House prices fell a total of 27% from their July 2006 peak to their February 2012 trough (S&P/Case-Shiller National Home Price Index). An analysis of structured finance products by Coval et al. (2009) reveals a similar neglect of tail risk by credit rating agencies and investors.

The potential for the financial sector to drag down the real economy was also underesti-

[38]Institutional investors are surveyed on "the probability of a catastrophic stock market crash in the U.S. ... in the next six months." The Crash Confidence Index is defined as the share of respondents who report a probability of less than 10%. In each wave, the institutional investors to be surveyed are sampled from the investment managers section of the Money Market Directory of Pension Funds and Their Investment Managers. For details, see https://som.yale.edu/faculty-research/our-centers-initiatives/international-center-finance/data/stock-market-confidence

mated by professional forecasters. In the summer of 2008 the Fed forecasting staff was asked to prepare a "severe financial stress" forecast. The prediction was a real GDP growth rate of -0.4% for the remainder of 2008, followed by a growth rate of 0.5% in 2009 and 2.6% in 2010 (Gennaioli and Shleifer, 2018). The true recession was far more severe.

5.1.2 Sentiment Globally

The results above document how elevated sentiment sets the stage for a future financial crisis and drives a wedge between the true and perceived probability of a crisis occurring. Here I leverage the model's global solution to examine this wedge over the entire state space.

Let $HP_5(e_t, \mathcal{I}_t)$ denote the true 5-year crisis hitting probability conditional on e_t and $\mathcal{I}_t$. Let $\widehat{HP_5}(e_t, \mathcal{I}_t)$ denote the perceived crisis hitting probability. Define

$$\text{CEW}_5(e_t, \mathcal{I}_t) \equiv HP_5(e_t, \mathcal{I}_t) - \widehat{HP_5}(e_t, \mathcal{I}_t),$$

where CEW stands for "Crisis Expectations Wedge." A positive CEW indicates neglected crash risk, and a negative CEW indicates that crises are overly representative.

Figure 5 plots $CEW_5(e_t, \mathcal{I}_t)$ over the entire state space of e_t, conditional on $\mathcal{I}_t \in \{-3SD, -2SD, ..., +2SD, +3SD\}$. The figure shows that CEW depends not only on sentiment, but also on the initial fragility of the financial sector. In particular, the sensitivity of CEW to sentiment is strongly non-monotonic in e_t.

When e_t is high and the intermediary sector is well-capitalized, $\text{CEW}_5 \approx 0$ and CEW is insensitive to $\mathcal{I}_t$. While sentiment still affects the future dynamics of e_t, the intermediary sector is far enough from the constraint that sentiment will unwind before the crisis region is reached. CEW becomes much more sensitive to sentiment as e_t moves closer to the crisis region. Once the intermediary sector starts to show distress, sentiment becomes a critical factor in determining whether or not a financial crisis materializes out of initial vulnerability.

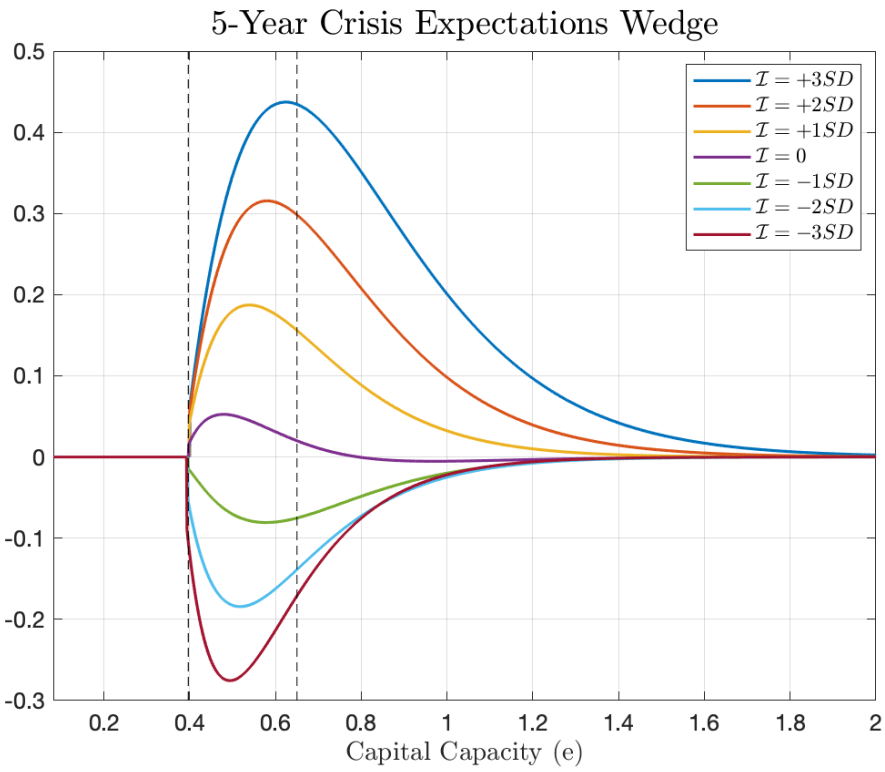


Figure 5: 5-year Crisis Expectations Wedge. The dashed vertical line at $e = 0.4$ marks the boundary for the crisis region, and the dashed vertical line at $e = 0.65$ marks $e_{distress}$.

However, as e_t moves even closer to the crisis region, CEW again becomes insensitive to $\mathcal{I}_t$. Near the crisis region, the economy is "one shock away" from a financial crisis regardless of the level of sentiment.[39] Finally, when the economy is actually in the crisis region then $HP_5 = \widehat{HP_5} = 1$, and $CEW_5 = 0$.

Empirical Evidence There is a small but growing literature characterizing when the bursting of asset price bubbles is and is not predictive of broader systemic risk. For example, the bursting of the dot-com bubble had a relatively muted impact on the real economy, whereas the bursting of the housing bubble resulted in the most severe recession since the Great Depression. Jordà et al. (2015) demonstrate that the bursting of an asset price bubble is predictive of a deep recession when the bubble coincides with a credit expansion. Brun-

[39] Over short intervals, shocks are of order $\sqrt{dt}$ whereas sentiment affects the drift of economy with order dt.

nermeier and Schnabel (2016) study 400 years of asset price bubbles and bursts, concluding that bursts are more severe when preceded by a lending boom. Similarly, Brunnermeier et al. (2017) estimate that the contribution of asset bubbles to future systemic risk depends on the vulnerability of the financial sector at the onset of the bubble. This evidence is consistent with the model's prediction that elevated sentiment is predictive of financial crises only when the intermediary sector is vulnerable.

Bubble-Pricking An important policy debate asks how financial regulators should optimally respond to emerging asset bubbles. My model includes neither monetary policy nor macroprudential regulation, and therefore one should be cautious in interpreting the following discussion. Nonetheless, Figure 5 contributes to this debate by characterizing where bubble-pricking can be most effective. In particular, the figure illustrates that there are three relevant regions to consider. First, when e_t is very high and the financial sector is far from its constraint, sentiment-driven asset price booms are not predictive of future systemic risk. In this region, the intermediary sector is well-capitalized and will be able to remain far from its constraint as sentiment unwinds. Second, there is an intermediate region where CEW is very sensitive to sentiment. It is this intermediate region where policymakers should likely focus, as this intermediate region is where policymakers can generate the largest reductions in crisis probabilities by unwinding sentiment. Third, for e_t near the crisis region it may be too late for policymakers to successfully intervene. When the intermediary sector is "one shock away" from a financial crisis, it may be best for policymakers to avoid being the ones who provide that final shock.

5.2 Stability from Beliefs

The analysis in section 5.1 looked at particular points in the $\{e_t, \mathcal{I}_t\}$ state space and calculated which parameter constellations are predictive of future financial crises. I now turn

to a slightly different question: does the DEE feature more financial crises than the corresponding REE? Under the baseline calibration, the answer is that the DEE features fewer financial crises than the REE. In short, diagnostic expectations are stabilizing.

Figure 6 shows this result visually. In the figure, the two orange lines (right axis) plot the marginal CDF of state variable e_t in the DEE and the REE. The blue curve (left axis) plots the CDF of the DEE divided by the CDF of the REE. When the blue curve is less than 1 for any particular value of e_t, this indicates that the DEE has a lower probability of having e_t below that point than the REE. The crisis region is marked by the dashed vertical line at $e_t \approx 0.4$. Since the blue curve crosses 1 above the crisis region, this indicates that the DEE features fewer financial crises than the REE.

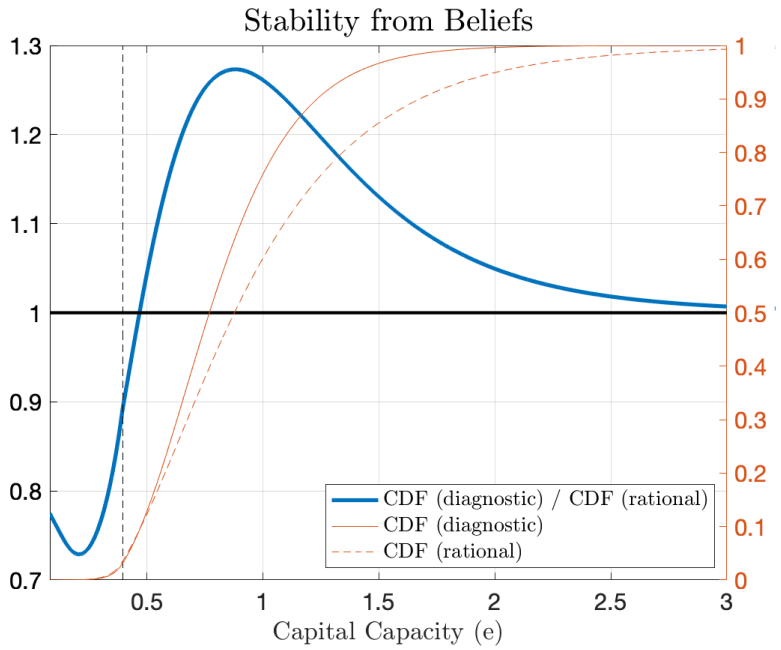


Figure 6: Comparison of the capital capacity (e) marginal distribution for the DEE and the REE. The dashed vertical line at $e \approx 0.4$ marks the boundary of the crisis region.

One can understand this *stability from beliefs* effect by superimposing the crisis likelihoods in Figure 4 onto the ergodic distribution shown in Figure 3. Figure 4 illustrates that the parts of the state space where the financial sector is distressed and sentiment is elevated are highly predictive of future financial crises. But, the ergodic distribution shows that this part

of the state space is rarely encountered. Instead, financial distress almost always coincides with excessive pessimism. Excessive pessimism decreases the probability of a future crisis because the financial sector earns a larger risk premium than perceived. This inhibits e_t from falling further, thereby providing a hedge against future financial crises.

Intuitively, this result can be understood as providing a rationale for Paul Samuelson's famous quip that "Wall Street indexes predicted nine out of the last five recessions" (Samuelson, 1966). It is precisely excessive pessimism in financial markets that leads to quick recovery and prevents the financial sector from slipping further towards a full-blown crisis.

This stability from beliefs result may appear somewhat at odds with the earlier result that elevated sentiment is highly predictive of financial crises. Indeed, much of the empirical literature has found that elevated sentiment is predictive of future financial market downturns and concluded – the model suggests incorrectly – from this finding that extrapolative expectations promote financial instability (Gennaioli and Shleifer, 2018). These two seemingly-contradictory results can be reconciled by recognizing that the former is a conditional prediction while the latter is an unconditional prediction. The model's conditional prediction is that periods in which the financial sector is vulnerable and sentiment is elevated are highly predictive of financial crises. However, the ergodic distribution illustrates that it is rare for the model to reach these states. The model's unconditional prediction is that beliefs are stabilizing, because periods of financial distress are highly correlated with depressed sentiment.

While this stability from beliefs result occurs under the baseline calibration, it can be overturned under alternative calibrations. In particular, diagnostic expectations can generate instability rather than stability when the magnitude of perceptual error is increased. Robustness is explored in Appendix B.2.

Empirical Evidence The stability from beliefs result highlights a benefit of economic models: they can be used to compare outcomes from two different data generating processes (DEE vs. REE). For the same reason, it is difficult to provide direct empirical evidence on stability from beliefs. Empirical studies can only examine the world as it is; they cannot study counterfactual data generating processes.

Nonetheless, it is still possible to test the mechanism underlying this result and the empirical evidence in Bordalo et al. (2018a) provides support for this mechanism. BGS analyze professional forecasts of the Baa-Treasury credit spread and conclude that the credit spread mean-reverts quicker than forecasters expect. Forecasters are over-pessimistic about future credit conditions when spreads are wide, and over-optimistic about future credit conditions when spreads are narrow. Since credit spreads are typically wide when the financial sector's risk-bearing capacity is impaired (Gilchrist and Zakrajšek, 2012), the BGS spread evidence is consistent with the model's prediction that recovery from financial distress is typically faster than diagnostically expected.

6 Macroeconomic Implications: Expectations and Growth

I now turn to the macroeconomic implications of diagnostic expectations. Whereas most of the results thus far relate to the behavioral triggers of financial crises, the focus here is instead on sentiment-driven macroeconomic fluctuations in the non-distress part of the state space. Expectations influence economic growth through investment, since output is a controlled diffusion with investment rate i_t alone controlling the growth rate of the economy: $\frac{dY_t}{Y_t} = \frac{dK_t}{K_t} = (i_t - \delta)dt + \sigma dZ_t$. The results to follow highlight the benefit of being able to apply diagnostic expectations to endogenous processes, as expectations about capital propagate through the dynamical system to alter the future evolution of capital and output.

The reader should note that the calibration of ζ will be important in this section, because

ζ governs how diagnostic expectations are passed into asset prices versus the risk-free rate. The baseline calibration sets $\zeta < 1$, implying that q_t is increasing in $\mathcal{I}_t$. Since investment rate i_t is a function of q_t, i_t is also increasing in sentiment when $\zeta < 1$.

To highlight the impact of sentiment in a conventional way, I study impulse-response functions for investment rate i_t following a one standard deviation shock to $\mathcal{I}$. Figure 7 plots three investment rate IRFs corresponding to economies with differing frictions. Appendix Figure 8 plots the corresponding IRFs for log-output, which is the integral of the investment rate IRFs in Figure 7.[40]

The black horizontal line at 0 is the investment rate IRF that obtains under rational expectations ($\theta = 0$). The dashed gray line plots the investment rate IRF in a model with behavioral frictions but no financial frictions. The red curve plots the investment rate IRF in the full model with both behavioral frictions and financial frictions. All three economies are started at their stochastic steady state before $\mathcal{I}_t$ is shocked to $+1SD$.[41] No further shocks are fed into the system.

The REE (black line) exhibits no dynamics following the sentiment shock. Next, the dashed gray line plots the investment rate IRF arising in a model with diagnostic expectations but no financial frictions. Because q_t is increasing in $\mathcal{I}_t$ under the baseline calibration, a shock to sentiment causes q_t to jump up. Investment is a function of q_t, so investment rate i_t also increases. As time passes and no further shocks hit the system, $\mathcal{I}_t$ decays back to zero at rate κ. As sentiment decays, q_t and i_t also return to their long-run values. Thus, the model with behavioral frictions but no financial frictions experiences a sentiment-driven investment boom, but no subsequent bust. Finally, the red curve plots the investment rate IRF in the full DEE with both behavioral frictions and financial frictions. As with the dashed gray line,

[40] Setting $\sigma dZ_t = 0$, this follows since $dlog(Y_t) = dk_t = (i_t - \delta - \frac{\sigma^2}{2})dt$.

[41] The rational economy (black curve) starts at its stochastic steady state of $e_0 = 0.94, \mathcal{I}_0 = 0$. The economy without financial frictions (gray curve) starts with $\mathcal{I}_0 = 0$. The initial level e_0 is irrelevant when there is no capital constraint. The full economy with both frictions (red curve) starts at $e_0 = 0.87, \mathcal{I}_0 = 0$.

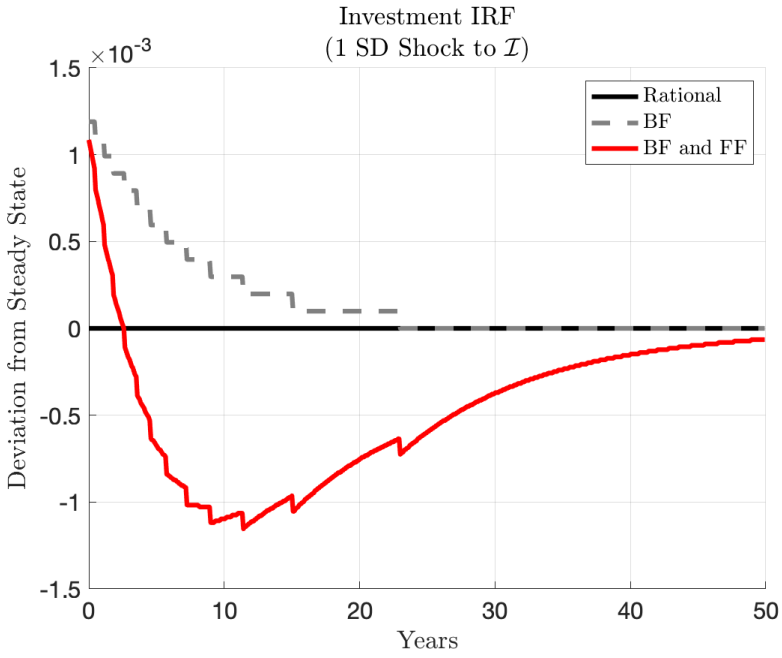


Figure 7: Investment Rate IRFs.

the red line jumps up at the time of the sentiment shock. However, the IRF declines rapidly and quickly turns negative, before slowly reverting back to the steady state from below. In short, the interaction of behavioral and financial frictions generates a boom-bust cycle in investment.

The initial investment boom arises because the positive shock to sentiment increases q_t. However, excessive optimism now feeds back into financial frictions by slowly eroding the balance sheet of intermediaries. As e_t declines so too does the risk appetite of the financial sector, and this decreases q_t and i_t. Because e_t reflects the cumulation of a series of past returns, the dynamics of e_t play out slowly over a long horizon. At first the sentiment shock dominates and increases q_t, causing a short-run boom. However, as sentiment unwinds the economy is left with a weakened financial sector. This generates a long-run bust, which slowly subsides as the financial sector rebuilds its capital capacity.

Empirical Evidence This boom-bust pattern in investment aligns with recent empirical studies of the credit cycle. Most directly, Gulen et al. (2019) find that elevated credit market

sentiment at time t correlates with a boom in corporate investment over the subsequent year, followed by a long-run contraction in corporate investment. Similarly, López-Salido et al. (2017) find that elevated credit-market sentiment at time t predicts lower GDP growth from year $t+2$ to $t+3$. At a more micro level, Greenwood and Hanson (2014) study the dry bulk shipping industry and argue that extrapolation of high current ship earnings produces irrationally high levels of investment followed by low future investment returns. Gennaioli et al. (2016) turn directly to survey evidence, documenting that large U.S. corporations tend to invest more when their CFOs are more optimistic, and that CFO expectations are irrationally extrapolative. Finally, the boom-bust cycle generated in Figure 7 is consistent with the observation of Greenwood et al. (2019) that the business cycle is disconnected from the credit cycle, with financial fragility arising at the tail-end of economic expansions.

7 Diagnostic Expectations in Continuous Time

I end by providing a microfoundation for the reduced-form expectations process outlined in section 2.2. The goal for the model of diagnostic expectations derived here is to be a Portable Extension of Existing Models ["PEEMish"](Rabin, 2013). Put differently, the expectations model is designed to be portable such that rational models can be augmented with diagnostic expectations using a single additional state variable.

Step 1: Defining the Background Context Following the terminology of Bordalo et al. (2018a), the first step is to define the *background context* for capital. The background context is the dynamic reference class against which diagnostic expectations are formed. It should be thought of as a counterfactual level of the log capital stock.

The background context reflects the absence of new information. In equation (3), variable $\mathcal{I}_t \equiv \sigma \int_0^t e^{-\kappa(t-s)} dZ_s$ was introduced as a measure of "new information." This leads to the

following definition of the background context.

Definition 3. *Let G_t^- denote the time-dependent "background context" of log capital k_t. G_t^- is defined as follows:*

$$G_t^- = k_t - \mathcal{I}_t.$$

As an Ornstein-Uhlenbeck process, $\mathcal{I}_t$ has an unconditional mean of zero. Additionally, $\mathbb{E}_t[\lim_{\tau\to\infty}(k_{t+\tau} - G_{t+\tau}^-)] = 0$, meaning that differences between k_t and G_t^- are not expected to persist permanently.[42]

Diagnostic expectations are applied to the log of capital for two reasons. Psychologically, it is consistent with Weber's Law that shocks are perceived as percentage changes rather than level changes. Mathematically, working with log capital ensures that variable $\mathcal{I}_t$ is stationary because the diffusion coefficient for log capital is constant.

Step 2: Modeling Expectations Given the Background Context Now that I've defined the background context G_t^-, the next step is to specify how agents form expectations. Because time is continuous, I need to specify diagnostic expectations over all future periods $t+\tau$, with $\tau > 0$. The current period is t. Let $h(k_{t+\tau}|k_t, e_t, \mathcal{I}_t)$ denote the true distribution of log capital at time $t+\tau$ conditional on current state variables (here log capital k_t is used as a state variable rather than capital level K_t). Let $h(k_{t+\tau}|G_t^-, e_t, \mathcal{I}_t)$ denote the true distribution of log capital at time $t+\tau$ conditional on current state variables e_t and $\mathcal{I}_t$, but now using counterfactual log capital level G_t^-.

Let $k'_{t+\tau}$ denote one possible realization of log capital at time $t+\tau$. Following BGS and Gennaioli and Shleifer (2010), the *representativeness* of future state $k'_{t+\tau}$ is given by the

[42] In this property, $\mathbb{E}_t$ is the true expectations operator at time t (not a distorted expectation).

following likelihood ratio:

$$\frac{h(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t)}{h(k'_{t+\tau}|G_t^-, e_t, \mathcal{I}_t)}. \tag{27}$$

The most representative states are the ones exhibiting the largest increase in likelihood based on recent objective news.

The representativeness heuristic biases expectations. However, one difficulty with equation (27) is that little is known about distributions $h(k_{t+\tau}|k_t, e_t, \mathcal{I}_t)$ and $h(k_{t+\tau}|G_t^-, e_t, \mathcal{I}_t)$ because k_t is an endogenous process.[43] This difficulty can be overcome by using an instantaneous prediction horizon of $\tau = dt$. Specifically, because k_t is an Itô Process it is instantaneously Gaussian. Taking $\tau \to dt$, $h(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t) \sim \mathcal{N}\left(k_t + \left[i(e_t, \mathcal{I}_t) - \delta - \frac{\sigma^2}{2}\right] dt, \sigma^2 dt\right)$ and $h(k'_{t+\tau}|G_t^-, e_t, \mathcal{I}_t) \sim \mathcal{N}\left(G_t^- + \left[i(e_t, \mathcal{I}_t) - \delta - \frac{\sigma^2}{2}\right] dt, \sigma^2 dt\right)$. For this reason I now define diagnostic expectations over prediction horizon $\tau = dt$. The prediction horizon will be extended iteratively in Step 3.

Diagnostic expectations overweight states of capital that are representative of new information. This is formalized by assuming that agents evaluate future levels of log capital according to the distorted density:

$$h_t^\theta(k'_{t+dt}|k_t, e_t, \mathcal{I}_t) = h(k'_{t+dt}|k_t, e_t, \mathcal{I}_t) \cdot \left[\frac{h(k'_{t+dt}|k_t, e_t, \mathcal{I}_t)}{h(k'_{t+dt}|G_t^-, e_t, \mathcal{I}_t)}\right]^{\theta dt} \frac{1}{Z}. \tag{28}$$

Equation (28) modifies a similar formula in BGS, with the key difference being that equation (28) defines expectations at $t + dt$ while the discrete-time formulation of BGS defines expectations at $t + 1$. In equation (28), the true conditional probability $h(k'_{t+dt}|k_t, e_t, \mathcal{I}_t)$ is distorted by the representativeness term in brackets.

The extent to which representativeness distorts expectations is governed by parameter θ. θ is scaled by the prediction horizon dt because representativeness should impose only an

[43] This is in contrast to BGS, where expectations are only formed over exogenous AR(N) processes.

infinitesimal distortion on the perceived distribution of k_{t+dt} over such a short horizon. The agent knows the current state k_t, and because k_t evolves continuously the agent also knows that k_{t+dt} must be "very close" to k_t.

Using equation (28), the following Proposition characterizes the perceived evolution capital:

Proposition 2. *A diagnostic agent perceives that capital evolves according to*

$$\frac{\widehat{dK_t}}{K_t} = (i(e_t, \mathcal{I}_t) - \delta)dt + \theta \mathcal{I}_t dt + \sigma dZ_t. \tag{29}$$

Proof. See Appendix C. □

Thus, Proposition 2 shows that at a prediction horizon of $\tau = dt$, judging by representativeness biases the perceived drift of K_t.

Step 3: The Evolution of Beliefs

In Step 1 I defined the background context G_t^- and in Step 2 I specified how an agent with diagnostic expectations predicts $\widehat{dK_t}$. The final step is to model both the perceived and realized dynamics of expectations over longer horizons. Because capital is endogenous, one only knows the instantaneous distribution of k_t and therefore future expectations are defined iteratively. In particular, by repeatedly applying the instantaneous Gaussian properties of k_t I can iteratively define expectations of the economy at $t + dt$, then $t + 2dt$, then $t + 3dt$, etc. This iterative procedure imposes that the law of iterated expectations holds with respect to distorted expectations, consistent with the BGS model.

The current time period is t. Diagnostic agents form expectations about future states by simulating the economy forward in their mind (possibly in a biased way). Let $\widehat{\mathcal{I}}_{t+\tau}$ denote the simulated perception of new information at time $t + \tau$. $\widehat{\mathcal{I}}_{t+\tau}$ evolves deterministically

and is defined as follows:

$$\begin{aligned}\widehat{\mathcal{I}}_{t+\tau} &= \int_0^t e^{-\kappa(t+\tau-s)}\sigma dZ_s, \text{ or equivalently} \\ &= e^{-\kappa\tau}\mathcal{I}_t.\end{aligned} \tag{30}$$

The psychology of diagnostic expectations is that the agent has in the back of her mind the true model, yet overweights states of the world that are representative of recent objective news. As the agent simulates the economy forward (in a biased manner) there is no new information hitting the economy and therefore no new shocks entering $\widehat{\mathcal{I}}_{t+\tau}$. Rather, the agent is just simulating forward the known model. New information that was representative at time t fades as the agent simulates further forward in time. Thus, equation (30) shows that as τ increases, news that is representative at time t fades and $\widehat{\mathcal{I}}_{t+\tau} \to 0$ deterministically.

Expectations are formed iteratively by simulating the economy forward state-by-state. Let $k'_{t+\tau}, e'_{t+\tau}, \widehat{\mathcal{I}}_{t+\tau}$ denote one possible realization of state variables at time $t+\tau$. Using equation (30), the simulated future background context can now be defined in an analogous fashion to Definition 3.

Definition 4. *Let $k'_{t+\tau}$ denote some simulated level of log capital at future time $t+\tau$. The simulated background context at time $t+\tau$ is defined as follows:*

$$\widehat{G}^-_{t+\tau} = k'_{t+\tau} - \widehat{\mathcal{I}}_{t+\tau}.$$

As above, the simulated future background context is defined to reflect the absence of new information.

Again proceeding in an analogous fashion to Step 2, at time $t+\tau$ the agent iteratively

forms expectations about $t+\tau+dt$ according to:

$$h_t^{\theta}(k'_{t+\tau+dt}|k'_{t+\tau}, e'_{t+\tau}, \widehat{\mathcal{I}}_{t+\tau}) = h(k'_{t+\tau+dt}|k'_{t+\tau}, e'_{t+\tau}, \widehat{\mathcal{I}}_{t+\tau}) \cdot \left[\frac{h(k'_{t+\tau+dt}|k'_{t+\tau}, e'_{t+\tau}, \widehat{\mathcal{I}}_{t+\tau})}{h(k'_{t+dt}|\widehat{G}^{-}_{t+\tau}, e'_{t+\tau}, \widehat{\mathcal{I}}_{t+\tau})}\right]^{\theta dt} \frac{1}{Z}. \tag{31}$$

As with Proposition 2, it follows that the diagnostic agent perceives that capital evolves according to:

$$\frac{\widehat{dK'_{t+\tau}}}{K'_{t+\tau}} = (i(e'_{t+\tau}, \widehat{\mathcal{I}}_{t+\tau}) - \delta)dt + \theta\widehat{\mathcal{I}}_{t+\tau}dt + \sigma dZ_t. \tag{32}$$

Future expectations in equation (32) should be contrasted with those of a rational agent who correctly believes that, given $e'_{t+\tau}$ and $\widehat{\mathcal{I}}_{t+\tau}$, the drift of $k'_{t+\tau}$ will be $i(e'_{t+\tau}, \widehat{\mathcal{I}}_{t+\tau}) - \delta$. Equation (32) specifies that the effect of diagnostic expectations on the perceived drift of capital at time $t+\tau$ fades as the agent simulates the evolution of the economy further into the future ($\tau \to \infty$). Diagnostic expectations capture the overweighting of states that are representative of current economic conditions. As the agent looks further temporally ahead, all states become less representative of economic conditions at time t so the bias caused by representativeness fades.

It is important to emphasize here that equation (32) only stipulates that the diagnostic agent's perception of drift converges to rationality as $\tau \to \infty$. Because the drift has a cumulative effect on the level of k_t, the diagnostic prediction of the level of $k_{t+\tau}$ can diverge increasingly from the rational prediction of $k_{t+\tau}$ as τ increases (e.g., Figure 1).

Summary This completes the microfoundation of the reduced-form beliefs process specified in section 2.2. Extensions are discussed in Appendix F.2. To summarize, expectations of the endogenous capital process are formed iteratively in order to make repeated use of the instantaneous Gaussian properties of $dk_{t+\tau}$. Step 2 defines how $\mathcal{I}_t$ affects the expected

evolution of the economy from t to $t + dt$. Step 3 then defines how $\mathcal{I}_t$ is simulated to evolve over time. In detail, Step 2 takes $k_t, e_t, \mathcal{I}_t$ as given and provides a perceived mapping into $\widehat{k}_{t+dt}$ and $\widehat{e}_{t+dt}$ given shock dZ_t. The hat-notation denotes that agents may not properly understand the evolution of these state variables. Step 3 takes $\mathcal{I}_t$ as given and provides $\widehat{\mathcal{I}}_{t+dt}$. Then, we can again apply Step 2 (now taking $\widehat{k}_{t+dt}$, $\widehat{e}_{t+dt}$, and $\widehat{\mathcal{I}}_{t+dt}$ as given) to calculate $\widehat{k}_{t+2dt}$ and $\widehat{e}_{t+2dt}$ given shocks dZ_t and dZ_{t+dt}. Applying step 3 again gives $\widehat{\mathcal{I}}_{t+2dt}$. This process is repeated to generate expectations at time $t + \tau$, for all $\tau > 0$.

I end by discussing why this model of diagnostic expectations can serve as a portable extension of existing rational models. First, equations (29) and (32) illustrate that state variable $\mathcal{I}_t$ alone is sufficient to characterize the state of expectations relative to rationality. Specifically, for any $\tau \geq 0$ diagnostic expectations distort the rationally-perceived drift by $\theta\widehat{\mathcal{I}}_{t+\tau}$. Second, the evolution of $\mathcal{I}_t$ is self-contained. Equation (4) shows that state variable $\mathcal{I}_t$ plus the shock dZ_t are sufficient to calculate $d\mathcal{I}_t$. It is these two attributes that make this formulation of diagnostic expectations portable: $\mathcal{I}_t$ alone characterizes expectations relative to rationality, and $\mathcal{I}_t$ is sufficient for its own evolution.

8 Conclusion

This paper develops a general equilibrium macroeconomic model that combines financial frictions with diagnostic expectations. This allows the model to jointly account for the macroeconomic effects of financial frictions as well as their behavioral triggers. When the financial sector is fragile, elevated sentiment amplifies systemic risk and sets the stage for financial crises. The model also predicts that these periods occur rarely. Unconditionally, diagnostic expectations inhibit financial crises because over-pessimism in times of financial distress allows the intermediary sector to earn large returns. Finally, the model generates boom-bust investment dynamics due to the interaction of diagnostic expectations and non-

linear financial frictions.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and wealth distribution in macroeconomics: A continuous-time approach," 2017.

Adam, Klaus, Dmitry Matveev, and Stefan Nagel, "Do survey expectations of stock returns reflect risk-adjustments?," 2018.

Adrian, Tobias and Markus K Brunnermeier, "CoVaR," *American Economic Review*, 2016, *106* (7), 1705–41.

_ **and Nina Boyarchenko**, "Intermediary Leverage Cycles and Financial Stability," *FRB of New York Staff Report*, 2012, (567).

Aghion, Philippe and Peter W Howitt, *The economics of growth*, MIT press, 2008.

Alili, Larbi, Pierre Patie, and Jesper Lund Pedersen, "Representations of the first hitting time density of an Ornstein-Uhlenbeck process," *Stochastic Models*, 2005, *21* (4), 967–980.

Barberis, Nicholas, "Psychology-Based Models of Asset Prices and Trading Volume," *Handbook of Behavioral Economics-Foundations and Applications 1*, 2018, p. 79.

_ **, Robin Greenwood, Lawrence Jin, and Andrei Shleifer**, "X-CAPM: An extrapolative capital asset pricing model," *Journal of financial economics*, 2015, *115* (1), 1–24.

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

Bernanke, Ben S, Mark Gertler, and Simon Gilchrist, "The financial accelerator in a quantitative business cycle framework," *Handbook of macroeconomics*, 1999, *1*, 1341–1393.

Bordalo, P, N Gennaioli, R La Porta, and A Shleifer, "Diagnostic expectations and stock returns," *Journal of Finance*, 2019.

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

_ , _ , _ , and Stephen J. Terry, "Real Credit Cycles," *Mimeo*, 2019.

_ , _ , Spencer Yongwook Kwon, and Andrei Shleifer, "Diagnostic bubbles," 2018.

_ , _ , Yueran Ma, and Andrei Shleifer, "Over-reaction in macroeconomic expectations," 2018.

Brunnermeier, Markus and Isabel Schnabel, "Bubbles and Central Banks: Historical Perspectives," in "Central Banks at a Crossroads–What Can We Learn from History?," Cambridge University, 2016, pp. 493–562.

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

_ and _ , "Macro, money, and finance: A continuous-time approach," in "Handbook of Macroeconomics," Vol. 2, Elsevier, 2016, pp. 1497–1545.

_ , Simon Rother, and Isabel Schnabel, "Asset Price Bubbles and Systemic Risk," 2017.

Case, Karl E, Robert J Shiller, and Anne Thompson, "What have they been thinking? Home buyer behavior in hot and cold markets," 2012.

Casscells, Ward, Arno Schoenberger, and Thomas B Graboys, "Interpretation by physicians of clinical laboratory results," *New England Journal of Medicine*, 1978, *299* (18), 999–1001.

Coval, Joshua, Jakub Jurek, and Erik Stafford, "The economics of structured finance," *Journal of Economic Perspectives*, 2009, *23* (1), 3–25.

Duffie, Darrell, *Dynamic asset pricing theory*, Princeton University Press, 2010.

Fahlenbrach, Rüdiger, Robert Prilmeier, and René M Stulz, "Why does fast loan growth predict poor performance for banks?," *The Review of Financial Studies*, 2017, *31* (3), 1014–1063.

Farhi, Emmanuel and Iván Werning, "Monetary policy, bounded rationality, and incomplete markets," 2017.

Foote, Christopher L, Kristopher S Gerardi, and Paul S Willen, "Why did so many people make so many ex post bad decisions? The causes of the foreclosure crisis," 2012.

Frankel, Marvin, "The production function in allocation and growth: a synthesis," *The American Economic Review*, 1962, *52* (5), 996–1022.

Fuster, Andreas, Benjamin Hebert, and David Laibson, "Natural expectations, macroeconomic dynamics, and asset pricing," *NBER Macroeconomics Annual*, 2012, *26* (1), 1–48.

_ , David Laibson, and Brock Mendel, "Natural expectations and macroeconomic fluctuations," *Journal of Economic Perspectives*, 2010, *24* (4), 67–84.

Gabaix, Xavier, "Behavioral macroeconomics via sparse dynamic programming," 2016.

_ , "A behavioral New Keynesian model," 2016.

García-Schmidt, Mariana and Michael Woodford, "Are low interest rates deflationary? A paradox of perfect-foresight analysis," *American Economic Review*, 2019, *109* (1), 86–120.

Gennaioli, Nicola and Andrei Shleifer, "What comes to mind," *The Quarterly journal of economics*, 2010, *125* (4), 1399–1433.

_ **and** _ , *A crisis of beliefs: Investor psychology and financial fragility*, Princeton University Press, 2018.

_ **,** _ **, and Robert Vishny**, "Neglected risks, financial innovation, and financial fragility," *Journal of Financial Economics*, 2012, *104* (3), 452–468.

_ **, Yueran Ma, and Andrei Shleifer**, "Expectations and investment," *NBER Macroeconomics Annual*, 2016, *30* (1), 379–431.

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

_ **,** _ **, and Andrea Prestipino**, "A macroeconomic model with financial panics," 2017.

Gilchrist, Simon and Egon Zakrajšek, "Credit spreads and business cycle fluctuations," *American Economic Review*, 2012, *102* (4), 1692–1720.

Greenwood, Robin and Andrei Shleifer, "Expectations of returns and expected returns," *The Review of Financial Studies*, 2014, *27* (3), 714–746.

_ **and Samuel G Hanson**, "Issuer quality and corporate bond returns," *The Review of Financial Studies*, 2013, *26* (6), 1483–1525.

_ **and** _ , "Waves in ship prices and investment," *The Quarterly Journal of Economics*, 2014, *130* (1), 55–109.

_ **, Samuel G. Hanson, and Lawrence J. Jin**, "Reflexivity in Credit Markets," *Working Paper*, 2019.

Gulen, Huseyin, Ion Mihai, and Stefano Rossi, "Credit Cycles and Corporate Investment," *Working Paper*, 2019.

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

_ **and** _ , "A Macroeconomic Framework for Quantifying Systemic Risk," *American Economic Journal: Macroeconomics*, 2019.

_ **, Bryan Kelly, and Asaf Manela**, "Intermediary asset pricing: New evidence from many asset classes," *Journal of Financial Economics*, 2017, *126* (1), 1–35.

Jin, Lawrence J, "A speculative asset pricing model of financial instability," *Available at SSRN 2524762*, 2015.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor, "Leveraged bubbles," *Journal of Monetary Economics*, 2015, *76*, S1–S20.

Kahneman, Daniel and Amos Tversky, "Subjective probability: A judgment of representativeness," *Cognitive psychology*, 1972, *3* (3), 430–454.

Karatzas, Ioannis and Steven E Shreve, "Stochastic Differential Equations," in "Brownian Motion and Stochastic Calculus," Springer, 1998, pp. 281–398.

Keynes, JM, "The General Theory of Employment, Interest and Money," 1936.

Kiyotaki, Nobuhiro and John Moore, "Credit cycles," *Journal of political economy*, 1997, *105* (2), 211–248.

Krishnamurthy, Arvind and Tyler Muir, "How Credit Cycles across a Financial Crisis," *Working Paper*, 2017.

López-Salido, David, Jeremy C Stein, and Egon Zakrajšek, "Credit-market sentiment and the business cycle," *The Quarterly Journal of Economics*, 2017, *132* (3), 1373–1426.

Maggiori, Matteo, "Financial intermediation, international risk sharing, and reserve currencies," *American Economic Review*, 2017, *107* (10), 3038–71.

Malmendier, Ulrike and Stefan Nagel, "Depression babies: do macroeconomic experiences affect risk taking?," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

Mian, Atif, Amir Sufi, and Emil Verner, "Household debt and business cycles worldwide," *The Quarterly Journal of Economics*, 2017, *132* (4), 1755–1817.

Minsky, Hyman P, "The financial instability hypothesis: An interpretation of Keynes and an alternative to "standard" theory," *Challenge*, 1977, *20* (1), 20–27.

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

Rabin, Matthew, "An approach to incorporating psychology into economics," *American Economic Review*, 2013, *103* (3), 617–22.

Samuelson, Paul, "Science and Stocks," *Newsweek*, 1966, *September*, 92.

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

Sims, Christopher A, "Implications of rational inattention," *Journal of monetary Economics*, 2003, *50* (3), 665–690.

Tversky, Amos and Daniel Kahneman, "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment.," *Psychological review*, 1983, *90* (4), 293.

Wachter, Jessica A and Michael J Kahana, "A retrieved-context theory of financial decisions," *Available at SSRN*, 2019.

Woodford, Michael, "Macroeconomic analysis without the rational expectations hypothesis," *Annu. Rev. Econ.*, 2013, *5* (1), 303–346.

A Model Details

A.1 Labor Income Microfoundation

The treatment here of Frankel (1962) follows from Aghion and Howitt (2008, Chapter 2). This is a model in which each individual producer faces decreasing returns to capital, but decreasing returns at the level of the individual producer are offset at the aggregate level through knowledge externalities.

At time t, there exists a measure $\mathbb{J}_t$ of bankers.[44] Each banker j operates $K_{j,t}$ of capital and hires $L_{j,t}$ units of labor at time t. The banker faces the production function

$$Y_{j,t} = \bar{A}_t K_{j,t}^{\nu} L_{j,t}^{1-\nu}, \tag{33}$$

where $\bar{A}_t$ is an aggregate productivity level. Aggregate productivity $\bar{A}_t$ is not constant. Due to knowledge spillovers it depends on the total amount of capital in the economy. In particular:

$$\bar{A}_t = A \left(\int_j K_{j,t} dj \right)^{\eta}. \tag{34}$$

Here, $\eta \geq 0$ controls the extent to which there are knowledge externalities in the economy.

Each banker will hire labor at wage $\mathcal{W}_t$ according to the following maximization problem:

$$\begin{aligned} L_{j,t} &= argmax_{\ell} \ \bar{A}_t K_{j,t}^{\nu} \ell^{1-\nu} - q_t K_{j,t} - \mathcal{W}_t \ell \\ 0 &= (1-\nu) \bar{A}_t \left(\frac{K_{j,t}}{\ell} \right)^{\nu} - \mathcal{W}_t \\ \ell &= K_{j,t} \left(\frac{\bar{A}_t (1-\nu)}{\mathcal{W}_t} \right)^{\frac{1}{\nu}}. \end{aligned} \tag{35}$$

[44] There is always a unit measure of households. The size of the financial intermediary sector varies over time due to banker entry and exit.

The next step is to impose market clearing. Specifically, $\int_j K_{j,t}dj = K_t$ and $\int_j L_{j,t}dj = 1$. Since all bankers are identical, this implies that $K_{j,t} = \frac{K_t}{\mathbb{J}_t}$ and $L_{j,t} = \frac{1}{\mathbb{J}_t}$. Plugging in, we have:

$$\begin{aligned}
\bar{A}_t &= AK_t^{\eta}, \text{ and} \\
Y_t &= \int_j Y_{j,t}dj = AK_t^{\eta} \int_j K_{j,t}^{\nu} L_{j,t}^{1-\nu} dj \\
&= AK_t^{\eta} \mathbb{J}_t \left(\frac{K_t}{\mathbb{J}_t}\right)^{\nu} \left(\frac{1}{\mathbb{J}_t}\right)^{1-\nu} \\
&= AK_t^{\eta+\nu}.
\end{aligned}$$

Here I impose a key parameter restriction to generate linearity:

$$\eta + \nu = 1. \tag{36}$$

This parameter restriction recovers an "AK economy" with the aggregate production function $Y_t = AK_t$, exactly as in the main text.

The final step is to use market clearing conditions to solve for the wage rate. Plugging these into equation (35):

$$\begin{aligned}
\frac{1}{\mathbb{J}_t} &= \frac{K_t}{\mathbb{J}_t} \left(\frac{AK_t^{\eta}(1-\nu)}{\mathcal{W}_t}\right)^{\frac{1}{\nu}} \\
\mathcal{W}_t^{\frac{1}{\nu}} &= K_t \left(AK_t^{\eta}(1-\nu)\right)^{\frac{1}{\nu}} \\
\mathcal{W}_t &= AK_t^{\eta+\nu}(1-\nu) \\
\mathcal{W}_t &= (1-\nu)AK_t,
\end{aligned} \tag{37}$$

where the last line again relies on the calibration that $\eta + \nu = 1$. Thus, the knowledge externalities model built here provides a simple microfoundation for the reduced-form labor

income process assumed in the main text.

A.2 Entry Details

As was outlined in the main text, there exists a reflecting barrier $\underline{e}$ deep in the crisis region at which capital is used to increase bank reputation $\mathcal{E}_t$. Adding details here, start at the boundary with an aggregate reputation of $\underline{\mathcal{E}} = \underline{e}K$. Consider a further shock to reputation of z which sends reputation to $\underline{\mathcal{E}} - z < \underline{e}K$. After this shock, the reflecting boundary implies that capital will immediately be converted into reputation to restore $\underline{e}$. Specifically, let x denote the amount of new reputation required to restore $\underline{e}$. x is given by:

$$\underline{e} = \frac{\underline{\mathcal{E}} - z + x}{K - \beta x}$$
$$x = \frac{z}{1 + \underline{e}\beta}.$$

Thus, βx capital will be destroyed in order to restore reputation to $\underline{e}$.

This equation for capital destruction can now be used to derive boundary conditions for price functions q and P. As was described in the main text, $q_e(\underline{e}, \mathcal{I}) = 0$. This is the lower boundary condition for q. For P, the condition that $P_e(\underline{e}, \mathcal{I}) = 0$ implies that $p = \frac{P}{K}$ will have a non-zero slope at $\underline{e}$.[45] Consider the above reputation shock of z. Immediately after this shock, the price of housing is $P = p\left(\underline{e} - \frac{z}{K}, \mathcal{I}'\right) K$. This must equal the price of housing immediately after capital is spent to re-build reputation, given by $P = p\left(\underline{e}, \mathcal{I}'\right)(K - \beta x)$. In the continuous-time limit with arbitrarily small shocks, $p\left(\underline{e} - \frac{z}{K}, \mathcal{I}'\right) K$ can be rewritten as

[45] This is due to capital destruction, which makes the denominator of $\frac{P}{K}$ change at the entry boundary.

$p(\underline{e}, \mathcal{I}')K + p_e(\underline{e}, \mathcal{I}')\, z$. Combining, we have:

$$p(\underline{e}, \mathcal{I}')K + p_e(\underline{e}, \mathcal{I}')\, z = p(\underline{e}, \mathcal{I}')\left(K - \beta \frac{z}{1 + \underline{e}\beta}\right)$$
$$p_e(\underline{e}, \mathcal{I}') = \frac{p(\underline{e}, \mathcal{I}')\beta}{1 + \underline{e}\beta}.$$

This gives the boundary condition for p at $\underline{e}$. Details on how these boundary conditions are imposed numerically are given in Appendix E.

A.3 Diagnostic Expectations Equilibrium Solution Details

Here I detail how to solve for a Diagnostic Expectations Equilibrium (DEE) that is Markov in state variables $e_t, \mathcal{I}_t$, and K_t. Brunnermeier and Sannikov (2016) provide a good treatment of the general techniques used in this section.

To start, I postulate that agents perceive that capital price q_t and scaled housing price p_t follow:

$$\widehat{\frac{dq_t}{q_t}} = \widehat{\mu_t^q} dt + \widehat{\sigma_t^q} dZ_t \tag{38}$$

$$\widehat{\frac{dp_t}{p_t}} = \widehat{\mu_t^p} dt + \widehat{\sigma_t^p} dZ_t. \tag{39}$$

Using these postulated price processes, the perceived laws of motion for the three state variables are:

$$\widehat{\frac{dK_t}{K_t}} = (i_t - \delta)dt + \theta \mathcal{I}_t dt + \sigma dZ_t \tag{40}$$

$$\begin{aligned}\widehat{\frac{de_t}{e_t}} = \widehat{d\widetilde{R}_t} - (i_t - \delta + \theta \mathcal{I}_t)dt + (\sigma^2 - \sigma(\alpha_t^h \widehat{\sigma_t^h} + \alpha_t^k \widehat{\sigma_t^k}))dt \\ - \eta dt + d\psi_t + (\alpha_t^h \widehat{\sigma_t^h} + \alpha_t^k \widehat{\sigma_t^k} - \sigma) dZ_t\end{aligned} \tag{41}$$

$$\widehat{\frac{d\mathcal{I}_t}{\mathcal{I}_t}} = -\kappa dt \tag{42}$$

Equations (40) and (42) were derived in the main text. Equation (41) can be derived using Itô's lemma to expand $e_t = \frac{\mathcal{E}_t}{K_t}$ under the perceived processes $\widehat{\frac{d\mathcal{E}_t}{\mathcal{E}_t}} = \widehat{d\widetilde{R}_t} - \eta dt + d\psi_t$ and $\widehat{\frac{dK_t}{K_t}} = (i_t - \delta)dt + \theta \mathcal{I}_t dt + \sigma dZ_t$. For simplicity, I will rewrite the perceived evolution of e_t as:

$$\widehat{\frac{de_t}{e_t}} = \widehat{\mu_t^e} dt + \widehat{\sigma_t^e} dZ_t.$$

Price Processes Given that I am searching for a Markov equilibrium, let $q(e, \mathcal{I})$ and $p(e, \mathcal{I})$ denote the two price functions. Applying Itô's lemma using the perceived laws of motion for the two state variables gives:

$$\widehat{\mu_t^q} = \frac{q_e(e_t, \mathcal{I}_t)}{q_t} e_t \widehat{\mu_t^e} - \frac{q_{\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \kappa \mathcal{I}_t + \frac{1}{2} \frac{q_{ee}(e_t, \mathcal{I}_t)}{q_t} (e_t \widehat{\sigma_t^e})^2 \tag{43}$$

$$\widehat{\sigma_t^q} = \frac{q_e(e_t, \mathcal{I}_t)}{q_t} e_t \widehat{\sigma_t^e} \tag{44}$$

$$\widehat{\mu_t^p} = \frac{p_e(e_t, \mathcal{I}_t)}{p_t} e_t \widehat{\mu_t^e} - \frac{p_{\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \kappa \mathcal{I}_t + \frac{1}{2} \frac{p_{ee}(e_t, \mathcal{I}_t)}{p_t} (e_t \widehat{\sigma_t^e})^2 \tag{45}$$

$$\widehat{\sigma_t^p} = \frac{p_e(e_t, \mathcal{I}_t)}{p_t} e_t \widehat{\sigma_t^e} \tag{46}$$

These formulas will prove useful throughout. Specifically, equations (43) and (45) are second-order PDEs that I will eventually solve numerically. As one can see from those two formulas, I need to pin down $\widehat{\mu_t^q}, \widehat{\mu_t^p}, \widehat{\mu_t^e}$, and $\widehat{\sigma_t^e}$, leaving only the price functions undetermined. This is where I now turn.

Market Clearing and Returns From goods market clearing equation (23):

$$\begin{aligned}
Y_t &= C_t^y + \Phi(i_t, K_t) \\
A &= \frac{C_t^y}{K_t} + i_t + \frac{\xi}{2}(i_t - \delta)^2 \\
&= \frac{1-\phi}{\phi} \frac{D_t}{K_t} + i_t + \frac{\xi}{2}(i_t - \delta)^2, \text{ using (12) and } C_t^h = 1 \\
\frac{D_t}{K_t} &= \frac{\phi}{1-\phi} \left[A - i_t - \frac{\xi}{2}(i_t - \delta)^2 \right].
\end{aligned}$$

Since equation (13) gives $i_t = \delta + \frac{q_t - 1}{\xi}$, I've now pinned down $\frac{D_t}{K_t}$ as a function of q_t.

D_t is the dividend paid on housing, and this expression can now be plugged into housing

returns as follows:

$$\begin{aligned}\widehat{dR_t^h} &= \frac{\widehat{dP_t} + D_t dt}{P_t} \\ &= \frac{d(\widehat{p(e_t, \mathcal{I}_t)K_t})}{p(e_t, \mathcal{I}_t)K_t} + \frac{D_t}{p_t K_t} \\ &= \frac{d(\widehat{p(e_t, \mathcal{I}_t)K_t})}{p(e_t, \mathcal{I}_t)K_t} + \frac{\frac{\phi}{1-\phi}\left[A - i_t - \frac{\xi}{2}(i_t - \delta)^2\right]}{p_t}.\end{aligned}$$

Applying Ito's Lemma to the first term:

$$\begin{aligned}\widehat{dR_t^h} &= \left[\widehat{\mu_t^p} + i_t - \delta + \theta\mathcal{I}_t + \sigma\widehat{\sigma_t^p} + \frac{\phi}{1-\phi}\frac{A - i_t - \frac{\xi}{2}(i_t - \delta)^2}{p(e_t, \mathcal{I}_t)}\right] dt + (\sigma + \widehat{\sigma_t^p})dZ_t \\ &= (\widehat{\pi_t^h} + r_t)dt + \widehat{\sigma_t^h} dZ_t.\end{aligned}$$

Equation (18) gives a similar process for capital returns. Repeating that equation here:

$$\begin{aligned}\widehat{dR_t^k} &= \left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta\mathcal{I}_t + \sigma\widehat{\sigma_t^q}\right) dt + (\sigma + \widehat{\sigma_t^q})dZ_t \\ &= (\widehat{\pi_t^k} + r_t)dt + \widehat{\sigma_t^k} dZ_t.\end{aligned}$$

The final return process to derive is the risk-free interest rate r_t. Starting again from market clearing:

$$\begin{aligned}C_t^y &= Y_t - \Phi(i_t, K_t) \\ &= Y_t - i_t K_t - \frac{\xi}{2}(i_t - \delta)^2 K_t \\ &= \left(A - \delta - \frac{q_t - 1}{\xi} - \frac{(q_t - 1)^2}{2\xi}\right) K_t, \text{ using (13)}.\end{aligned}$$

Now deriving the perceived evolution of C_t^y using Itô's Lemma:

$$\begin{aligned}
\widehat{dC_t^y} &= \left(A-\delta-\frac{q_t-1}{\xi}-\frac{(q_t-1)^2}{2\xi}\right)K_t((i_t-\delta+\theta\mathcal{I}_t)dt+\sigma dZ_t)-\left(\frac{\widehat{dq_t}}{\xi}+\frac{(q_t-1)}{\xi}\widehat{dq_t}\right)K_t \\
&= \left(A-\delta-\frac{q_t-1}{\xi}-\frac{(q_t-1)^2}{2\xi}\right)K_t((i_t-\delta+\theta\mathcal{I}_t)dt+\sigma dZ_t)-\frac{q_t\widehat{dq_t}}{\xi}K_t \\
\frac{\widehat{dC_t^y}}{C_t^y} &= (i_t-\delta+\theta\mathcal{I}_t)dt-\frac{\frac{1}{\xi}q_t^2\widehat{\mu_t^q}}{A-\delta-\frac{q_t-1}{\xi}-\frac{(q_t-1)^2}{2\xi}}dt+\left(\sigma-\frac{\frac{1}{\xi}q_t^2\widehat{\sigma_t^q}}{A-\delta-\frac{q_t-1}{\xi}-\frac{(q_t-1)^2}{2\xi}}\right)dZ_t
\end{aligned}$$

Plugging this into the interest rate formula (15):

$$r_t=\rho+\zeta\left[i_t-\delta+\theta\mathcal{I}_t-\frac{\frac{1}{\xi}q_t^2\widehat{\mu_t^q}}{A-\delta-\frac{q_t-1}{\xi}-\frac{(q_t-1)^2}{2\xi}}\right]-\frac{\zeta(\zeta+1)}{2}\left[\sigma-\frac{\frac{1}{\xi}q_t^2\widehat{\sigma_t^q}}{A-\delta-\frac{q_t-1}{\xi}-\frac{(q_t-1)^2}{2\xi}}\right]^2$$

Intermediary Optimality From equation (41), it is the case that $\alpha_t^k\widehat{\sigma_t^k}+\alpha_t^h\widehat{\sigma_t^h}=\widehat{\sigma_t^e}+\sigma$. Using this in equation (22) gives:

$$\frac{\widehat{\pi_t^h}}{\widehat{\sigma_t^h}}=\frac{\widehat{\pi_t^k}}{\widehat{\sigma_t^k}}=\gamma(\widehat{\sigma_t^e}+\sigma).$$

From the return processes $\widehat{dR_t^k}$ and $\widehat{dR_t^h}$ above, we have that $\widehat{\sigma_t^k}=\sigma+\widehat{\sigma_t^q}$ and $\widehat{\sigma_t^h}=\sigma+\widehat{\sigma_t^p}$. Starting with capital, here I combine the banker's optimality condition with the perceived returns on capital K_t :

$$\begin{aligned}
\gamma(\widehat{\sigma_t^e}+\sigma) &= \frac{\widehat{\pi_t^k}}{\widehat{\sigma_t^k}} \\
&= \frac{\left(\frac{\nu A}{q_t}+\widehat{\mu_t^q}-\delta+\theta\mathcal{I}_t+\sigma\widehat{\sigma_t^q}\right)-r_t}{\sigma+\widehat{\sigma_t^q}} \\
&= \frac{\left(\frac{\nu A}{q_t}+\widehat{\mu_t^q}-\delta+\theta\mathcal{I}_t+\sigma\frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}\right)-r_t}{\sigma+\frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}}, \text{ using equation (44)} \\
\gamma(\widehat{\sigma_t^e}+\sigma)\left(\sigma+\frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}\right) &= \left(\frac{\nu A}{q_t}+\widehat{\mu_t^q}-\delta+\theta\mathcal{I}_t+\sigma\frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}\right)-r_t \qquad (47)
\end{aligned}$$

Proceeding similarly for housing returns:

$$\begin{aligned}
\gamma(\widehat{\sigma_t^e}+\sigma) &= \frac{\widehat{\pi_t^h}}{\widehat{\sigma_t^h}} \\
&= \frac{\widehat{\mu_t^p}+i_t-\delta+\theta\mathcal{I}_t+\sigma\widehat{\sigma_t^p}+\frac{D_t}{p(e_t,\mathcal{I}_t)K_t}-r_t}{\sigma+\widehat{\sigma_t^p}} \\
&= \frac{\widehat{\mu_t^p}+i_t-\delta+\theta\mathcal{I}_t+\sigma\frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e}+\frac{D_t}{p(e_t,\mathcal{I}_t)K_t}-r_t}{\sigma+\frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e}}, \text{ using (46)} \\
\gamma(\widehat{\sigma_t^e}+\sigma)\left(\sigma+\frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e}\right) &= \widehat{\mu_t^p}+i_t-\delta+\theta\mathcal{I}_t+\sigma\frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e}+\frac{D_t}{p(e_t,\mathcal{I}_t)K_t}-r_t
\end{aligned} \tag{48}$$

Pinning down $\widehat{\sigma_t^e}$ Equations (47) and (48) express $\widehat{\mu_t^q}$ and $\widehat{\mu_t^p}$ in terms of prices, state variables, and $\widehat{\sigma_t^e}$. The final step is to pin down $\widehat{\sigma_t^e}$. Doing so:

$$\begin{aligned}
\widehat{\sigma_t^e} &= \alpha_t^h\widehat{\sigma_t^h}+\alpha_t^k\widehat{\sigma_t^k}-\sigma \\
&= \alpha_t^h(\sigma+\widehat{\sigma_t^p})+\alpha_t^k(\sigma+\widehat{\sigma_t^q})-\sigma \\
&= \frac{K_t}{E_t}\left[p_t(\sigma+\widehat{\sigma_t^p})+q_t(\sigma+\widehat{\sigma_t^q})\right]-\sigma, \text{ using (24) and (25)} \\
&= \frac{K_t}{E_t}\left[p_t(\sigma+\frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e})+q_t(\sigma+\frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e})\right]-\sigma \\
&= \frac{K_t}{E_t}\left[(p_t+q_t-\frac{E_t}{K_t})\sigma+(p_e(e_t,\mathcal{I}_t)+q_e(e_t,\mathcal{I}_t))e_t\widehat{\sigma_t^e}\right] \\
e_t\widehat{\sigma_t^e}\left(\frac{1}{e_t}-\frac{K_t}{E_t}p_e(e_t,\mathcal{I}_t)-\frac{K_t}{E_t}q_e(e_t,\mathcal{I}_t)\right) &= \frac{K_t}{E_t}(p_t+q_t-\frac{E_t}{K_t})\sigma \\
e_t\widehat{\sigma_t^e} &= \frac{K_t}{E_t}\frac{(p_t+q_t-\frac{E_t}{K_t})\sigma}{\frac{1}{e_t}-\frac{K_t}{E_t}p_e(e_t,\mathcal{I}_t)-\frac{K_t}{E_t}q_e(e_t,\mathcal{I}_t)} \\
e_t\widehat{\sigma_t^e} &= \frac{(p_t+q_t-\frac{E_t}{K_t})\sigma}{\frac{E_t}{K_t}\frac{1}{e_t}-p_e(e_t,\mathcal{I}_t)-q_e(e_t,\mathcal{I}_t)}.
\end{aligned} \tag{49}$$

Recall that $E_t=\min\{\mathcal{E}_t,(1-\lambda)(q_tK_t+p_tK_t)\}$, or equivalently $\frac{E_t}{K_t}=\min\{e_t,(1-\lambda)(q_t+p_t)\}$. Thus, equation (49) expresses $\widehat{\sigma_t^e}$ in terms of the two price functions and the two state variables e_t and $\mathcal{I}_t$.

Solving for Prices The final step is to solve for price functions $q(e_t, \mathcal{I}_t)$ and $p(e_t, \mathcal{I}_t)$. Specifically, they are given by the following system of second-order PDEs:

$$q_t \widehat{\mu_t^q} = q_e(e_t, \mathcal{I}_t) e_t \widehat{\mu_t^e} - q_{\mathcal{I}}(e_t, \mathcal{I}_t)\kappa \mathcal{I}_t + \frac{1}{2} q_{ee}(e_t, \mathcal{I}_t)(e_t \widehat{\sigma_t^e})^2$$
$$p_t \widehat{\mu_t^p} = p_e(e_t, \mathcal{I}_t) e_t \widehat{\mu_t^e} - p_{\mathcal{I}}(e_t, \mathcal{I}_t)\kappa \mathcal{I}_t + \frac{1}{2} p_{ee}(e_t, \mathcal{I}_t)(e_t \widehat{\sigma_t^e})^2$$

All terms in this system of second-order PDEs have now been expressed in terms of state variables e_t and $\mathcal{I}_t$, exogenous parameters, and price functions $p(e_t, \mathcal{I}_t)$ and $q(e_t, \mathcal{I}_t)$. Specifically, $\widehat{\mu_t^q}$ and $\widehat{\mu_t^p}$ are given by equations (47) and (48). $\widehat{\mu_t^e}$ is given by (41), noting that $\widehat{\widetilde{dR_t}}$ is itself a function of prices, $\widehat{\mu_t^q}$, $\widehat{\mu_t^p}$, $\widehat{\sigma_t^q}$ and $\widehat{\sigma_t^p}$. Equations (44) and (46) give $\widehat{\sigma_t^q}$ and $\widehat{\sigma_t^p}$ in terms of the price functions and $\widehat{\sigma_t^e}$. Equation (49) closes the loop by solving for $\widehat{\sigma_t^e}$ in terms of the two price functions.

This system of PDEs for price functions $p(e_t, \mathcal{I}_t)$ and $q(e_t, \mathcal{I}_t)$ is solved numerically. Details are provided in Appendix E.

A.4 True Laws of Motion

As with the perceived laws of motion, I begin by postulating that q_t and p_t truly evolve according to:

$$\frac{dq_t}{q_t} = \mu_t^q dt + \sigma_t^q dZ_t \tag{50}$$

$$\frac{dp_t}{p_t} = \mu_t^p dt + \sigma_t^p dZ_t. \tag{51}$$

Using these price processes, the true evolution of the three state variables is given by:

$$\frac{dK_t}{K_t} = (i_t - \delta)dt + \sigma dZ_t, \tag{52}$$

$$\begin{aligned}\frac{de_t}{e_t} = d\widetilde{R}_t - (i_t - \delta)dt + (\sigma^2 - \sigma(\alpha_t^h \sigma_t^h + \alpha_t^k \sigma_t^k))dt \\ - \eta dt + d\psi_t + (\alpha_t^h \sigma_t^h + \alpha_t^k \sigma_t^k - \sigma)dZ_t\end{aligned} \tag{53}$$

$$d\mathcal{I}_t = -\kappa \mathcal{I}_t dt + \sigma dZ_t \tag{54}$$

As above, equation (41) can be derived using Itô's lemma to expand $e_t = \frac{\mathcal{E}_t}{K_t}$ under the true processes $\frac{d\mathcal{E}_t}{\mathcal{E}_t} = d\widetilde{R}_t - \eta dt + d\psi_t$ and $\frac{dK_t}{K_t} = (i_t - \delta)dt + \sigma dZ_t$. For simplicity, I will rewrite the true evolution of e_t as:

$$\frac{de_t}{e_t} = \mu_t^e dt + \sigma_t^e dZ_t.$$

Price Processes The methods developed above in Appendix A.3 detail how to solve for price functions $q(e, \mathcal{I})$ and $p(e, \mathcal{I})$. Applying Itô's lemma to these functions using the true

laws of motion for the two state variables gives:

$$\mu_t^q = \frac{q_e(e_t, \mathcal{I}_t)}{q_t} e_t \mu_t^e - \frac{q_{\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \kappa \mathcal{I}_t + \frac{q_{e\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \sigma (e_t \sigma_t^e) + \frac{1}{2} \frac{q_{ee}(e_t, \mathcal{I}_t)}{q_t} (e_t \sigma_t^e)^2 + \frac{1}{2} \frac{q_{\mathcal{I}\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \sigma^2 \tag{55}$$

$$\sigma_t^q = \frac{q_e(e_t, \mathcal{I}_t)}{q_t} e_t \sigma_t^e + \frac{q_{\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \sigma \tag{56}$$

$$\mu_t^p = \frac{p_e(e_t, \mathcal{I}_t)}{p_t} e_t \mu_t^e - \frac{p_{\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \kappa \mathcal{I}_t + \frac{p_{e\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \sigma (e_t \sigma_t^e) + \frac{1}{2} \frac{p_{ee}(e_t, \mathcal{I}_t)}{p_t} (e_t \sigma_t^e)^2 + \frac{1}{2} \frac{p_{\mathcal{I}\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \sigma^2 \tag{57}$$

$$\sigma_t^p = \frac{p_e(e_t, \mathcal{I}_t)}{p_t} e_t \sigma_t^e + \frac{p_{\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \sigma \tag{58}$$

Market Clearing and Returns Following similar steps to above where I solved for perceived return process $\widehat{dR_t^h}$, the true housing return process dR_t^h is given by:

$$\begin{aligned}
dR_t^h &= \frac{dP_t + D_t dt}{P_t} \\
&= \frac{d(p(e_t, \mathcal{I}_t) K_t)}{p(e_t, \mathcal{I}_t) K_t} + \frac{D_t}{p_t K_t} \\
&= \left[\mu_t^p + i_t - \delta + \sigma \sigma_t^p + \frac{\phi}{1-\phi} \frac{A - i_t - \frac{\xi}{2}(i_t - \delta)^2}{p(e_t, \mathcal{I}_t)} \right] dt + (\sigma + \sigma_t^p) dZ_t \\
&= (\pi_t^h + r_t) dt + \sigma_t^h dZ_t.
\end{aligned}$$

The true process for capital returns is given by:

$$\begin{aligned}
dR_t^k &= \left(\frac{\nu A}{q_t} + \mu_t^q - \delta + \sigma \sigma_t^q \right) dt + (\sigma + \sigma_t^q) dZ_t \\
&= (\pi_t^k + r_t) dt + \sigma_t^k dZ_t
\end{aligned}$$

Pinning down σ_t^e Finally, I need to derive a formula for the true volatility of e_t. Doing so:

$$\begin{aligned}
\sigma_t^e &= \alpha_t^h \sigma_t^h + \alpha_t^k \sigma_t^k - \sigma \\
&= \alpha_t^h (\sigma + \sigma_t^p) + \alpha_t^k (\sigma + \sigma_t^q) - \sigma \\
&= \frac{K_t}{E_t}\left[p_t(\sigma + \widehat{\sigma_t^p}) + q_t(\sigma + \widehat{\sigma_t^q})\right] - \sigma, \text{ using (24) and (25)} \\
&= \frac{K_t}{E_t}\left[p_t(\sigma + \frac{p_e(e_t,\mathcal{I}_t)}{p_t} e_t\sigma_t^e + \frac{p_{\mathcal{I}}(e_t,\mathcal{I}_t)}{p_t}\sigma) + q_t(\sigma + \frac{q_e(e_t,\mathcal{I}_t)}{q_t} e_t\sigma_t^e + \frac{q_{\mathcal{I}}(e_t,\mathcal{I}_t)}{q_t}\sigma)\right] - \sigma \\
&= \frac{K_t}{E_t}\left[(p_t + q_t + p_{\mathcal{I}}(e_t,\mathcal{I}_t) + q_{\mathcal{I}}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma + (p_e(e_t,\mathcal{I}_t) + q_e(e_t,\mathcal{I}_t))e_t\sigma_t^e\right] \\
e_t\sigma_t^e\left[\frac{1}{e_t} - \frac{K_t}{E_t}p_e(e_t,\mathcal{I}_t) - \frac{K_t}{E_t}q_e(e_t,\mathcal{I}_t)\right] &= \frac{K_t}{E_t}(p_t + q_t + p_{\mathcal{I}}(e_t,\mathcal{I}_t) + q_{\mathcal{I}}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma \\
e_t\sigma_t^e &= \frac{K_t}{E_t}\frac{(p_t + q_t + p_{\mathcal{I}}(e_t,\mathcal{I}_t) + q_{\mathcal{I}}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma}{\frac{1}{e_t} - \frac{K_t}{E_t}p_e(e_t,\mathcal{I}_t) - \frac{K_t}{E_t}q_e(e_t,\mathcal{I}_t)} \\
e_t\sigma_t^e &= \frac{(p_t + q_t + \ p_{\mathcal{I}}(e_t,\mathcal{I}_t) + q_{\mathcal{I}}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma}{\frac{E_t}{K_t}\frac{1}{e_t} - p_e(e_t,\mathcal{I}_t) - q_e(e_t,\mathcal{I}_t)}.
\end{aligned} \tag{59}$$

Comparison with Diagnostic Expectations I end this section by comparing the processes that a diagnostic agent expects to the processes that truly arise in this economy. e_t is the state variable in this economy characterizing the severity of financial frictions. As is illustrated by equations (41) and (53), under diagnostic expectations agents misperceive both the drift and volatility of de_t. To show this in detail, I'll separately study the misperceptions of the drift component and the volatility component. Starting with the volatility component:

$$\widehat{\sigma_t^e} - \sigma_t^e = -\frac{(p_{\mathcal{I}}(e_t,\mathcal{I}_t) + q_{\mathcal{I}}(e_t,\mathcal{I}_t))\sigma}{\frac{E_t}{K_t} - e_t(p_e(e_t,\mathcal{I}_t) + q_e(e_t,\mathcal{I}_t))} \tag{60}$$

Equation (60) shows that the error agents make about the perceived volatility of e_t depends on $p_{\mathcal{I}}(e_t,\mathcal{I}_t)$ and $q_{\mathcal{I}}(e_t,\mathcal{I}_t)$. When the price functions are constant in $\mathcal{I}$ such that $p_{\mathcal{I}}(e_t,\mathcal{I}_t) =$

$q_{\mathcal{I}}(e_t, \mathcal{I}_t) = 0$, then it will be the case that $\widehat{\sigma_t^e} = \sigma_t^e$. In the baseline calibration with $\zeta < 1$, price functions are increasing in $\mathcal{I}$. When $\zeta < 1$, $\widehat{\sigma_t^e} > \sigma_t^e$; i.e., agents under-perceive the true volatility of e_t. This under-perception of volatility arises because shocks to capital have the unanticipated effect of altering $\mathcal{I}_t$ (which agents instead perceive to evolve deterministically). When prices q_t and P_t are increasing in $\mathcal{I}$ then capital shocks create unanticipated asset price appreciation when $dZ_t > 0$ and unanticipated asset price depreciation when $dZ_t < 0$.

Next I'll study the misperceptions that agents have about the drift of e_t. To keep the algebra clean, I'll drop all Itô correction terms. Fully:

$$\begin{aligned}
\widehat{\mu_t^e} - \mu_t^e &\approx \mathbb{E}\left[\widehat{d\widetilde{R}_t} - d\widetilde{R}_t\right] - \theta \mathcal{I}_t dt \\
&\approx \alpha_t^k \left(\widehat{\mu_t^q} + \theta \mathcal{I}_t - \mu_t^q\right) dt + \alpha_t^h \left(\widehat{\mu_t^p} + \theta \mathcal{I}_t - \mu_t^p\right) dt - \theta \mathcal{I}_t dt \\
&\approx \alpha_t^k \left(\frac{q_e(e_t, \mathcal{I}_t)}{q_t} e_t (\widehat{\mu_t^e} - \mu_t^e) + \theta \mathcal{I}_t\right) dt + \alpha_t^h \left(\frac{p_e(e_t, \mathcal{I}_t)}{p_t} e_t (\widehat{\mu_t^e} - \mu_t^e) + \theta \mathcal{I}_t - \mu_t^p\right) dt - \theta \mathcal{I}_t dt \\
\widehat{\mu_t^e} - \mu_t^e &\approx -\frac{(p_t + q_t - \frac{E_t}{K_t}) \theta \mathcal{I}_t}{\frac{E_t}{K_t} - e_t (p_e(e_t, \mathcal{I}_t) + q_e(e_t, \mathcal{I}_t))}, \text{ using similar steps as above} \qquad (61)
\end{aligned}$$

Note that $\frac{E_t}{K_t} = \min\{e_t, (1-\lambda)(q_t + p_t)\}$ and therefore $p_t + q_t - \frac{E_t}{K_t} > 0$. Equation (61) has the intuitive result that the error agents make about the perceived drift of e_t depends on $\mathcal{I}_t$. When $\mathcal{I}_t > 0$, agents are over-optimistic about the drift of e_t. When $\mathcal{I}_t < 0$, the reverse is true and e_t grows faster than agents expect. This shows why the level of sentiment $\mathcal{I}_t$ is so crucial for predicting crises.

B Robustness

This section examines robustness to behavioral parameters θ and κ, and the EIS parameter ζ. Specifically, I analyze the following six parameter perturbations: (i) $\theta \times SD(\mathcal{I}) = 1.5\%$, (ii) $\theta \times SD(\mathcal{I}) = \frac{0.75\%}{2}$, (iii) a $\mathcal{I}$ half-life of 20 years, (iv) a $\mathcal{I}$ half-life of 1 year, (v) $\zeta = 1$, and (vi) $\zeta = \frac{1}{2}$. The baseline calibration sets $\theta \times SD(\mathcal{I}) = 0.75\%$, a $\mathcal{I}$ half-life of 5 years, and $\zeta = 0.7$.

For perturbations to θ and ζ, all other parameters are kept at their baseline calibration in Table 1. For perturbations to the half-life of sentiment (κ), I change θ accordingly to ensure that $\theta \times SD(\mathcal{I})$ remains equal to 0.75%. This isolates the effect of sentiment persistence.[46]

[46] Otherwise, changes to κ would affect both the persistence of sentiment and also the magnitude of the bias in capital growth expectations. This is because changes to κ affect $Var(\mathcal{I}_t) = \frac{\sigma^2}{2\kappa}$.

B.1 Expectations of Financial Crises

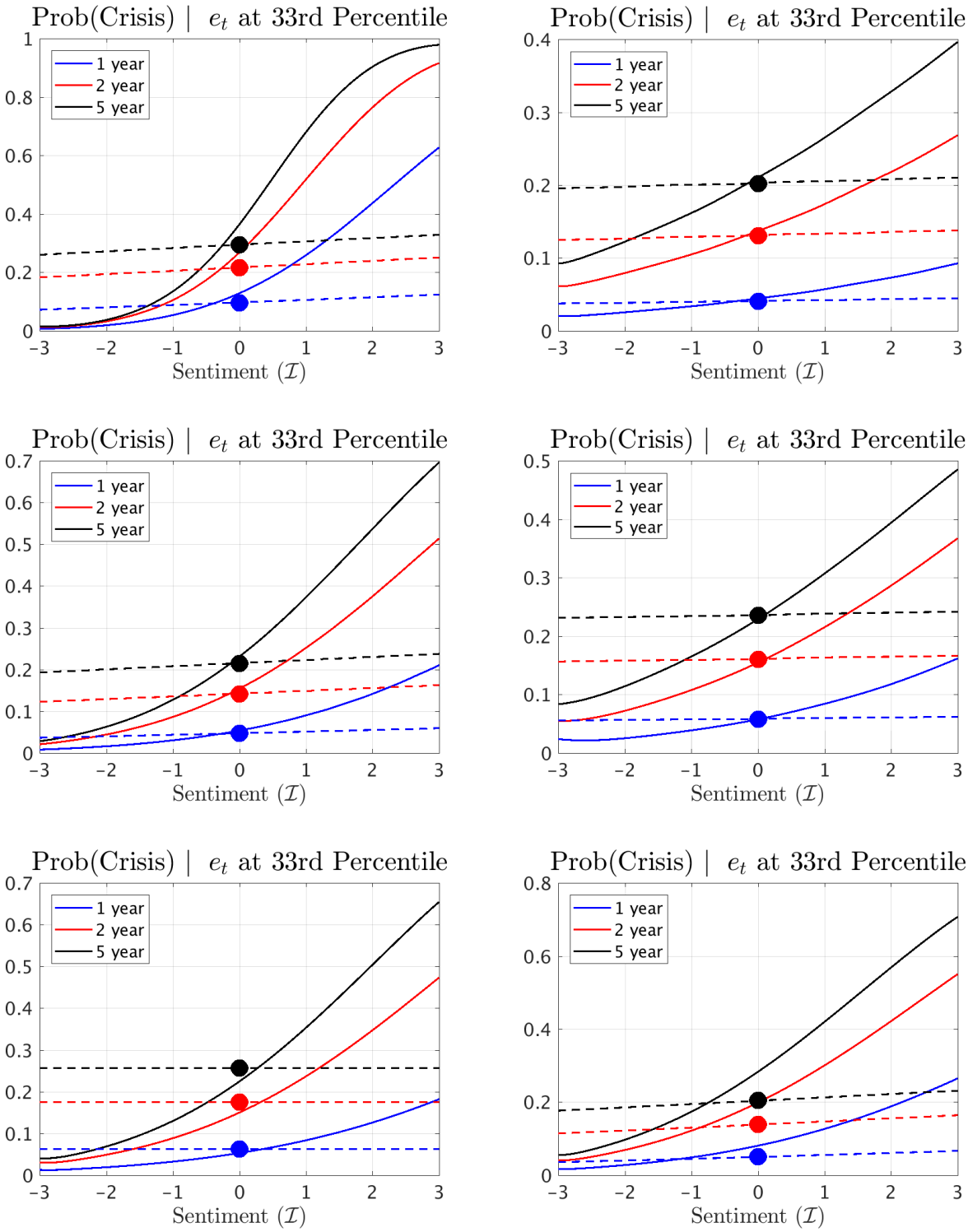


Figure 1: Crisis Hitting Probabilities. In order from left-to-right, top-to-bottom: $\theta \times SD(\mathcal{I}) = 1.5\%$, $\theta \times SD(\mathcal{I}) = \frac{0.75}{2}\%$, half-life $= 20$, half-life $= 1$, $\zeta = 1$, $\zeta = \frac{1}{2}$.

B.2 Stability from Beliefs

Appendix Table 3 lists crisis probabilities for the REE ($\theta = 0$) and the DEE across the six alternate calibrations listed above. Starting with θ, Appendix Table 3 evaluates the effect of doubling and halving θ. Note that the REE sets $\theta = 0$ by definition, so perturbations to parameter θ only alter the DEE. When θ is doubled, the crisis probability under the DEE increases above the REE crisis probability, thereby undoing the stability from beliefs effect. When θ is increased, periods of elevated and depressed sentiment correspond to larger asset pricing mistakes by the financial sector. This puts downward pressure on e_t when $\mathcal{I}_t$ is high, and upward pressure on e_t when $\mathcal{I}_t$ is low. Overall, this reduces the correlation between e_t and $\mathcal{I}_t$ to 0.1 (down from 0.7 in the baseline case). As was shown in Figure 4, the part of the state space where the financial sector is distressed and $\mathcal{I}_t > 0$ is the most predictive of future financial crises. By reducing the correlation between e_t and $\mathcal{I}_t$, the doubling of θ increases the probability that the economy will find itself in this part of the state space. This increases the probability of crises in the DEE.

The reverse is true when θ is halved from its baseline calibration. When θ is halved, diagnostic expectations become even more stabilizing relative to rationality. The intuition for this result is similar to the doubling of θ, except the effect here works in the opposite direction. When θ is halved, the correlation between e_t and $\mathcal{I}_t$ increases to 0.9. This makes it even less likely for the economy to find itself in a state where the financial sector is distressed and yet $\mathcal{I}_t > 0$.[47]

The persistence of sentiment is also important for determining the effect of diagnostic expectations on financial crises. Appendix Table 3 sets the half-life of new information to be 1 year and 20 years, compared to the baseline calibration of 5 years. Again, the REE

[47] Though not illustrated in Appendix Table 3, note that the relationship between θ and the probability of a crisis is non-monotonic. As $\theta \to 0$, the probability of a crisis in the DEE converges from below to the probability of a crisis in the REE.

is unaffected by perturbations to κ.[48] For the DEE, a large half-life of 20 years increases the probability of financial crises. In this case, $\mathcal{I}_t$ is highly persistent and this results in a correlation between e_t and $\mathcal{I}_t$ of only 0.1. Crisis probabilities increase accordingly. Alternatively, a shorter half-life of 1 year reduces the probability of crises relative to the baseline calibration. The shorter half-life means that $\mathcal{I}_t$ loads up more heavily on the most recent shocks to hit the system. This increases the correlation between e_t and $\mathcal{I}_t$ from 0.7 to 0.8.

Third, Appendix Table 3 evaluates the effect of parameter ζ on crisis probabilities. Recall that ζ is the inverse of the EIS, and has a baseline calibration value of $\zeta = 0.7$. Unlike perturbations to the behavioral parameters, perturbations to ζ affect both the REE and the DEE. Lowering ζ decreases the probability of crises in the REE, while raising ζ increases the probability of crises. This result for the REE is similar to the "volatility paradox" of Brunnermeier and Sannikov (2014). Lowering ζ increases the volatility of asset prices.[49] Higher volatility implies that each banker demands a higher risk premium. In equilibrium, this decreases the probability of financial crises even in the face of higher asset price volatility.

The effect of ζ on the probability of crises works in the opposite direction in the DEE. Under diagnostic expectations, a higher ζ lowers the probability of crises while a lower ζ increases the probability of crises. In the DEE, a lower ζ has a stronger effect on asset price volatility than in the REE. This is because asset prices become increasingly sensitive to $\mathcal{I}_t$ as ζ decreases.[50] However, diagnostic agents do not anticipate this sentiment-driven price volatility and therefore they do not demand a higher risk premium as compensation. This

[48] κ affects the evolution of $\mathcal{I}_t$, which is irrelevant when $\theta = 0$.

[49] The reason for this is that lowering ζ decreases the sensitivity of the risk-free rate to variation in e_t, and this means that asset prices must fluctuate instead. In detail, consider a decline in e_t which causes the financial sector to demand a higher risk premium. This increased risk premium can be realized either through a decline in the risk-free rate r_t or a decline in asset prices. Thus, variation in r_t can insulate asset prices from variation in e_t.

[50] As ζ decreases the interest rate r_t becomes less sensitive to changes in expected consumption growth. Consider an increase in sentiment. This increases the perceived future cash flows on capital and housing. If the risk-free discount rate r_t is insensitive to variation in $\mathcal{I}_t$, then this increase in perceived cash flows will be passed through to higher asset prices q_t and P_t (as these are just the discounted value of future cash flows). Thus, a lower ζ creates larger asset-price variation in $\mathcal{I}_t$.

is because diagnostic agents perceive $\mathcal{I}_t$ to evolve deterministically, so they do not understand the indirect effect of shocks on $\mathcal{I}_t$ and the corresponding pass-through to asset price volatility.[51] Thus, in the DEE a lower ζ increases asset price volatility without increasing the corresponding risk premium, and this additional (uncompensated) risk increases the relative probability of financial crises.

	Rational Prob(Crisis)	**Diagnostic Prob(Crisis)**	$\frac{\textbf{Diagnostic Prob(Crisis)}}{\textbf{Rational Prob(Crisis)}}$
Baseline	3.40 %	3.15 %	0.93
$\theta \times SD(\mathcal{I}) = 1.5\%$	3.40 %	5.44 %	1.60
$\theta \times SD(\mathcal{I}) = \frac{0.75}{2}\%$	3.40 %	3.00 %	0.88
Half-life = 20	3.40 %	4.08 %	1.20
Half-life = 1	3.40 %	2.81 %	0.83
$\zeta = 1$	4.63 %	2.92 %	0.63
$\zeta = \frac{1}{2}$	2.18 %	3.28 %	1.50

Table 3: Robustness of Stability from Beliefs.

[51] Equation (60) in Appendix A.3 illustrates the perceptual error that diagnostic agents make with respect to the volatility of e_t. As equation (60) shows, this error is driven by the partial derivatives $q_{\mathcal{I}}(e_t, \mathcal{I}_t)$ and $p_{\mathcal{I}}(e_t, \mathcal{I}_t)$; i.e., the sensitivity of asset prices to $\mathcal{I}_t$.

B.3 Expectations and Growth

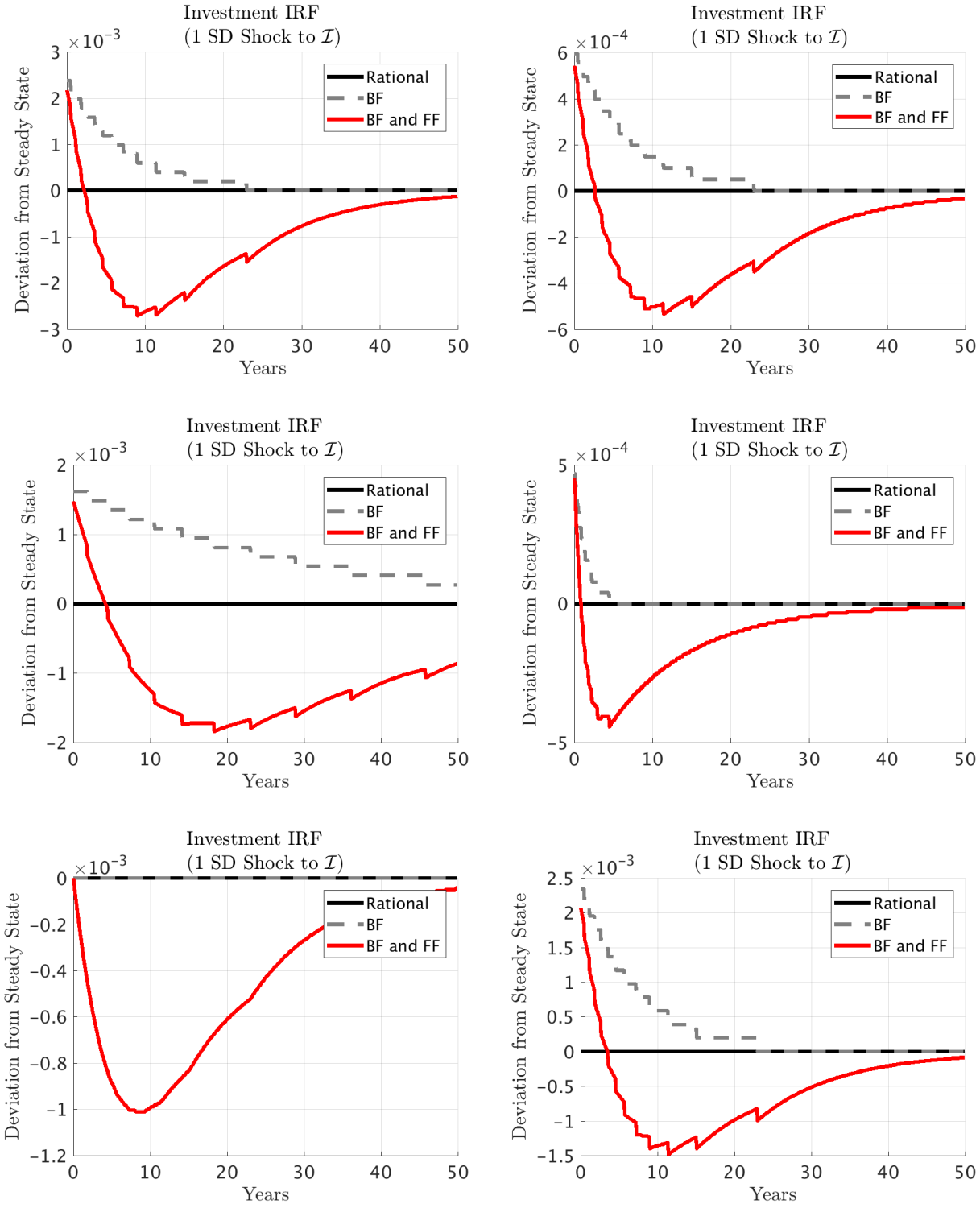


Figure 2: Investment Rate IRFs. In order from left-to-right, top-to-bottom: $\theta \times SD(\mathcal{I}) = 1.5\%$, $\theta \times SD(\mathcal{I}) = \frac{0.75}{2}\%$, half-life $= 20$, half-life $= 1$, $\zeta = 1$, $\zeta = \frac{1}{2}$.

C Proofs

C.1 Proof of Proposition 2

Consider the evolution of log capital k_t over a horizon of $\tau > 0$, holding the endogenous investment rate fixed at $\bar{i} = i_t$ from time t to $t + \tau$. Under fixed investment,

$$k_{t+\tau} = k_t + \int_t^{t+\tau} (\bar{i} - \delta - \frac{\sigma^2}{2})ds + \int_t^{t+\tau} \sigma dZ_s$$
$$k_{t+\tau} - k_t = +(\bar{i} - \delta - \frac{\sigma^2}{2})\tau + \sigma(Z_{t+\tau} - Z_t).$$

Since $\{Z_t\}$ is a standard Brownian motion:

$$k_{t+\tau} \sim \mathcal{N}(k_t + (\bar{i} - \delta - \frac{\sigma^2}{2})\tau, \sigma^2\tau).$$

For the arbitrary prediction horizon of τ, one can rewrite equation (28) as:

$$h_t^\theta(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t) = h(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t) \cdot \left[\frac{h(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t)}{h(k'_{t+\tau}|G_t^-, e_t, \mathcal{I}_t)}\right]^{\theta\tau} \frac{1}{Z}.$$

Given a fixed investment level of $\bar{i}$, this implies:

$$\begin{aligned} h^\theta(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t) &= \mathcal{N}(k_t + (\bar{i} - \delta - \frac{\sigma^2}{2})\tau, \sigma^2\tau) \cdot \left[\frac{\mathcal{N}(k_t + (\bar{i} - \delta - \frac{\sigma^2}{2})\tau, \sigma^2\tau)}{\mathcal{N}(G_t^- + (\bar{i} - \delta - \frac{\sigma^2}{2})\tau, \sigma^2\tau)}\right]^{\theta\tau} \frac{1}{Z} \\ &= \mathcal{N}(k_t + (\bar{i} - \delta - \frac{\sigma^2}{2})\tau + \theta(k_t - G_t^-)\tau, \sigma^2\tau). \end{aligned}$$

The BGS Internet Appendix provides algebraic details for this step. Equivalently, the agent perceives that

$$\widehat{k_{t+\tau}} - k_t = \int_t^{t+\tau} (\bar{i} - \delta - \frac{\sigma^2}{2})ds + \int_t^{t+\tau} \theta(k_t - G_t^-)ds + \int_t^{t+\tau} \sigma dZ_s.$$

Finally, taking the limit as $\tau \to dt$ gives

$$\widehat{dk_t} = (i_t - \delta - \frac{\sigma^2}{2})dt + \theta(k_t - G_t^-)dt + \sigma dZ_t.$$

Definition 3 can be used to plug $\mathcal{I}_t$ in for $k_t - G_t^-$. Finally, applying Itô's lemma to $K_t = exp(k_t)$ yields:

$$\frac{\widehat{dK_t}}{K_t} = (i_t - \delta)dt + \theta\mathcal{I}_t dt + \sigma dZ_t.$$

This completes the proof.

D Credit Spreads

Under diagnostic expectations, bankers set asset prices such that the perceived Sharpe ratio is equalized across all assets. In the baseline model bankers take a long position in two assets: capital and housing. However, asset pricing equation (22) can be used more generally to price any asset such that bankers have no incentive to take a net position in that asset. Details are provided below.

D.1 Short-Term Bond Spreads

The most straightforward way to study bond spreads is to price an instantaneous-maturity risky "corporate bond". A similar strategy is also employed by He and Krishnamurthy (2019) (see footnotes 14 and 15). Assume that there exists an instantaneous bond with a Poisson default rate that is a decreasing function of expected output growth, denoted $d(\widehat{\mathbb{E}}\frac{dK_t}{K_t})$. This assumption captures the empirical finding that default risk is countercyclical (Gilchrist and Zakrajšek, 2012).[52] For simplicity, the recovery rate is 0% in the case of default. I assume that this risky bond exists in zero net supply, and that it requires sophistication such that the financial sector must intermediate the holding of this corporate bond. Since the corporate bond can only be held by the banks, and all banks are identical, this bond must be priced such that it is held neither long nor short in equilibrium.

To simplify notation, let $\widehat{d}_t$ denote the perceived default rate at time t. Let the yield on this bond be denoted $y_t^r dt$. The perceived risk premium is given by $\pi_t^r = y_t^r - \widehat{d}_t - r_t$, and the Poisson default assumption implies that the Sharpe ratio is $\frac{y_t^r - \widehat{d}_t - r_t}{\widehat{d}_t}$.[53]

[52] I use expected output growth rather than realized output growth to abstract from the high-frequency Brownian noise that exists in realized output growth.

[53] For an instantaneous bond with price p^r, one can think of the perceived price as following $\frac{dp_t^r}{p_t^r} = y_t^r dt - \widehat{dJ}_t$, where $\widehat{J}_t$ is a Poisson jump process with a time-dependent jump intensity of $\widehat{d}_t$. Poisson jumps in the price process capture default, as $dp_t^r = -p_t^r$. The instantaneous expected log-return is given by $(1 - \widehat{d}_t dt) y_t^r dt + \widehat{d}_t dt(-1) = y_t^r dt - \widehat{d}_t dt$, dropping higher-order terms. Similarly, the instantaneous variance of log-returns is $(1 - \widehat{d}_t dt)(y_t^r dt - (y_t^r dt - \widehat{d}_t dt))^2 + \widehat{d}_t dt(-1 - (y_t^r dt - \widehat{d}_t dt))^2 = \widehat{d}_t dt$, again dropping terms

The next step is to price this corporate bond. Adding this asset to the banker's optimization problem:

$$\max_{\alpha_t^k,\alpha_t^h,\alpha_t^r} \left[r_t + \alpha_t^k \widehat{\pi_t^k} + \alpha_t^h \widehat{\pi_t^h} + \alpha_t^r \widehat{\pi_t^r}\right] - \frac{\gamma}{2}\left(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h} + \alpha_t^r \widehat{d}_t\right)^2,$$

where α_t^r denotes the banker's portfolio position in the corporate bond at time t. The FOC for α_t^r is:

$$\widehat{\pi_t^r} = \gamma\left(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h} + \alpha_t^r \widehat{d}_t\right)\widehat{d}_t$$

Setting $\alpha_t^r = 0$ in equilibrium and rearranging yields a Sharpe ratio of:

$$\frac{\widehat{\pi_t^r}}{\widehat{d}_t} = \frac{y^r - \widehat{d}_t - r_t}{\widehat{d}_t} = \gamma\left(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h}\right), \tag{62}$$

and a short-term spread of:

$$y^r - r_t = \widehat{d}_t\left[1 + \gamma\left(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h}\right)\right]. \tag{63}$$

The credit spread is a function of the perceived default risk $\widehat{d}_t$, plus a wedge capturing the risk premium.

Equation (63) is a numerically tractable measure of credit spreads because r_t α_t^k, α_t^h, $\widehat{\sigma_t^k}$, and $\widehat{\sigma_t^h}$ have all already been calculated as part of the numerical solution. Put differently, the short-term spread in equation (63) comes "for free". The downside of the credit spread in equation (63) is that it is constructed based on instantaneous debt. The majority of empirical work on the predictive power of spreads uses long-maturity bonds (e.g., the Baa-Treasury spread).[54]

of order smaller than dt.

[54]There are some short-term spreads which have been the focus of empirical work, such as the TED spread.

D.2 Long-Term Bond Spreads

Though more involved numerically, a similar strategy can be used to price long-term bonds. The pricing of long-term spreads aligns more closely with the existing literature studying the forecasting power of credit spreads (Gilchrist and Zakrajšek, 2012; Krishnamurthy and Muir, 2017; López-Salido et al., 2017).

Here, I assume that there exist long-term "corporate bonds" packaged into an index. For simplicity I assume that these are zero-coupon bonds, and pay a face value of \$100 15 years in the future, subject to default.[55] For each individual bond, default has a Poisson arrival rate of $d(\mathbb{E}\frac{dK_t}{K_t}, e_t, Rating)$, where the default rate is assumed to be a decreasing function of average economic growth, the financial sector's capital capacity, and the company's credit rating. As above, the assumption that default risk is countercyclical aligns with Gilchrist and Zakrajšek (2012). Default is assumed to covary negatively with the risk-bearing capacity of the financial sector to capture financing becoming tougher to obtain when the financial sector is fragile and intermediation is disrupted. Firm quality, denoted *Rating*, is introduced so that bonds of different risk levels can be priced.

I assume that there are an infinite number of these corporate bonds which are then packaged into indices based on *Rating*. This assumption implies that the default rate on each bond index is locally deterministic. Put differently, each index's instantaneous default rate is a deterministic function of the point in the state space, but the way that the economy evolves over the state space over the proceeding 15 years is stochastic.

I price risky corporate bond indices under the assumption that they exist in zero net supply and that only the financial sector can hold these indices. This ensures the financial sector will take no net position in these assets.

Ignoring dependence on the credit rating, at time $\tau \in [0, 15]$ let the price of the risky bond

[55] The Moody's Aaa and Baa bond indices are composed of bonds with maturities over 20 years. However, coupon payments will lower the duration of these bonds. This is why a 15-year maturity is chosen for the zero-coupon bond.

index be denoted p^r_τ such that the yield is $y^r_\tau = \frac{\log(100/p^r_\tau)}{\tau}$.[56] I postulate that the perceived price process for a risky bond index is:

$$\frac{\widehat{dp^r_\tau}}{p^r_\tau} = \widehat{\mu^r_\tau} dt + \widehat{\sigma^r_\tau} dZ_\tau.$$

Again I use the assumption that the financial sector takes no net position in the risky bond indices. The maximization problem of the financier yields:

$$\frac{\widehat{\mu^r_\tau} - \widehat{d_\tau} - r_t}{\widehat{\sigma^r_t}} = \gamma \left(\alpha^k_\tau \widehat{\sigma^k_\tau} + \alpha^h_\tau \widehat{\sigma^h_\tau} \right). \tag{64}$$

Unlike the instantaneous bond case, the yield spread is no longer an immediate consequence of the asset pricing equation. Instead, p^r_τ needs to be solved for numerically. Additionally, note that the only volatility in the above Sharpe ratio is price volatility. Default does not show up in the denominator because of the assumption that the perceived default rate is locally deterministic.

Price functions are solved for in a similar way to the price process solution in Appendix A.3. Assuming that $\widehat{\mu^r_\tau}$ and $\widehat{\sigma^r_\tau}$ are functions of e_τ and $\mathcal{I}_\tau$, an application of Itô's lemma yields:

$$\widehat{\mu^r_\tau} = \frac{\partial_\tau p^r_\tau}{p^r_\tau} + \frac{\partial_e p^r_\tau}{p^r_\tau} e_\tau \widehat{\mu^e_\tau} - \frac{\partial_\mathcal{I} p^r_\tau}{p^r_\tau} \kappa \mathcal{I}_\tau + \frac{1}{2} \frac{\partial_{ee} p^r_\tau}{p^r_\tau} (e_\tau \widehat{\sigma^e_\tau})^2 \tag{65}$$

$$\widehat{\sigma^r_\tau} = \frac{\partial_e p^r_\tau}{p^r_\tau} e_\tau \widehat{\sigma^e_\tau}. \tag{66}$$

To ease notation, I drop reliance on state variables e_t and $\mathcal{I}_t$ in equations (65) and (66). Additionally, I use the short-hand notation $\partial_a f$ for $\frac{\partial f}{\partial a}$.

Using equations (64) and (66) to plug in for $\widehat{\mu^r_\tau}$, equation (65) becomes a second-order

[56] The yield is $p^r_\tau \cdot exp(\tau \cdot y^r_\tau) = 100$. Taking logs and rearranging gives the desired result.

PDE where the only unknown is the price function $p^r_\tau(e_\tau, \mathcal{I}_\tau)$.[57] The general numerical methods for solving PDEs such as (65) is outlined in Appendix E, and Achdou et al. (2017) provides details.

Numerical Example Here I provide a simple numerical example of long-term bond spreads under the baseline calibration. I assume that there are two levels for *Rating*, corresponding loosely to "Aaa" and "Baa". The default function $d(\mathbb{E}\frac{dK_t}{K_t}, e_t, Rating)$ is set such that the average default rate on Aaa bonds is 0.24%, and the average default rate on Baa bonds is 0.53%. The default function is a decreasing function of average output growth, financing conditions, and bond quality. In order to generate downward-sloping spreads, one needs for the distance between $d(\widehat{\mathbb{E}}\frac{dK_t}{K_t}, e_t, Baa)$ and $d(\widehat{\mathbb{E}}\frac{dK_t}{K_t}, e_t, Aaa)$ to be decreasing in $\mathcal{I}_t$.[58]

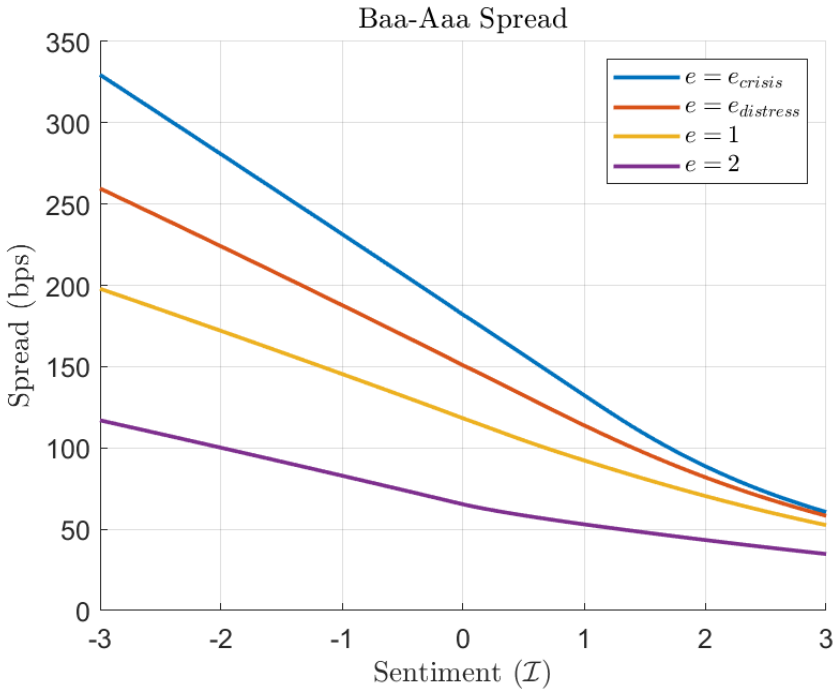


Figure 3: Long-Term Credit Spreads, Numerical Example.

[57] Assuming that $\widehat{\mu^e_\tau}$ and $\widehat{\sigma^e_\tau}$ have already been solved for in the main solution.
[58] For the specific calibration, see the online code.

E Numerical Methods

Before turning to the numerical methods, additional notation is required. The state space $\{e, \mathcal{I}\}$ can be thought of as a two-dimensional Itô process, denoted $\mathcal{S}$. Agents perceive that $\mathcal{S}$ evolves according to:

$$\widehat{d\mathcal{S}_t} = \begin{bmatrix} \widehat{de_t} \\ \widehat{d\mathcal{I}_t} \end{bmatrix} = \begin{bmatrix} e_t \widehat{\mu_t^e} \\ -\kappa \mathcal{I}_t \end{bmatrix} dt + \begin{bmatrix} e_t \widehat{\sigma_t^e} \\ 0 \end{bmatrix} dZ_t, \tag{67}$$

where Z_t is a one-dimensional Brownian motion. The true evolution of $\mathcal{S}$ is:

$$d\mathcal{S}_t = \begin{bmatrix} de_t \\ d\mathcal{I}_t \end{bmatrix} = \begin{bmatrix} e_t \mu_t^e \\ -\kappa \mathcal{I}_t \end{bmatrix} dt + \begin{bmatrix} e_t \sigma_t^e \\ \sigma \end{bmatrix} dZ_t. \tag{68}$$

The evolution of both $\widehat{d\mathcal{S}_t}$ and $d\mathcal{S}_t$ is subject to a reflecting barrier in the e_t dimension at $\underline{e}$.

Let $\mathcal{A}$ denote the infinitesimal generator of $\mathcal{S}_t$, and let $\widehat{\mathcal{A}}$ denote the infinitesimal generator of $\widehat{\mathcal{S}}_t$. This notation is not necessary, but will make for a more succinct description of the methods below.

E.1 Solving for Price Functions

As is shown in Appendix A.3, price functions $q(e_t, \mathcal{I}_t)$ and $p(e_t, \mathcal{I}_t)$ compose a system of second-order PDEs which must be solved for numerically:

$$\begin{aligned} q(e_t, \mathcal{I}_t)\widehat{\mu_t^q} &= q_e(e_t, \mathcal{I}_t) e_t \widehat{\mu_t^e} - q_{\mathcal{I}}(e_t, \mathcal{I}_t)\kappa \mathcal{I}_t + \frac{1}{2} q_{ee}(e_t, \mathcal{I}_t)(e_t \widehat{\sigma_t^e})^2 \\ &= \widehat{\mathcal{A}} q_t \end{aligned} \tag{69}$$

$$\begin{aligned} p(e_t, \mathcal{I}_t)\widehat{\mu_t^p} &= p_e(e_t, \mathcal{I}_t) e_t \widehat{\mu_t^e} - p_{\mathcal{I}}(e_t, \mathcal{I}_t)\kappa \mathcal{I}_t + \frac{1}{2} p_{ee}(e_t, \mathcal{I}_t)(e_t \widehat{\sigma_t^e})^2 \\ &= \widehat{\mathcal{A}} p_t \end{aligned} \tag{70}$$

In order to solve for these price functions I employ finite-difference methods. The numerical methods appendix of Achdou et al. (2017) provides an excellent reference, and I assume knowledge of their appendix here.

Algorithm I solve for price functions using two nested while-loops. In the outer loop, I iterate over lower boundary $\underline{e}$. In the inner loop, I take $\underline{e}$ as given and iterate over price functions p and q until convergence. Details for the inner loop are provided below. The outer loop continues to iterate over $\underline{e}$ until the resulting Sharpe ratio at $e = \underline{e}$, $\mathcal{I} = 0$ is close to the calibration target in Table 1.

I create a discretized grid over state variables e and $\mathcal{I}$.[59] Let subscript i denote the gridpoints in the e-dimension, and let subscript j denote the gridpoints in the $\mathcal{I}$-dimension. The algorithm is as follows. Let $n = 1, 2, ...$ track the current loop iteration.

1. Guess functions $q^0_{i,j}$ and $p^0_{i,j}$ at each point $\{i, j\}$ in the two-dimensional grid over e and $\mathcal{I}$.

2. Solve for $\widehat{\mu^q_{i,j}}$, $\widehat{\mu^p_{i,j}}$, $\widehat{\mu^e_{i,j}}$, and $\widehat{\sigma^e_{i,j}}$ using the previous iteration's price functions of $q^{n-1}_{i,j}$ and $p^{n-1}_{i,j}$ (or the initial guess) . To do so, use equation (49) to solve for $\widehat{\sigma^e_{i,j}}$, (47) to solve for $\widehat{\mu^q_{i,j}}$, (48) to solve for $\widehat{\mu^p_{i,j}}$, and finally (41) to solve for $\widehat{\mu^e_{i,j}}$. Next, construct the discretized infinitesimal generator, denoted $\widehat{\mathbf{A}}^{n-1}$, using $\widehat{\mu^e_{i,j}}$ and $\widehat{\sigma^e_{i,j}}$.[60] Note that $\widehat{\mathbf{A}}$ features a reflecting barrier at $\underline{e}$.[61]

3. Use an implicit scheme to solve for price functions $q^n_{i,j}$ and $p^n_{i,j}$:

$$\frac{q^n_{i,j} - q^{n-1}_{i,j}}{\Delta} + \widehat{\mu^q_{i,j}}\, q^n_{i,j} = \widehat{\mathbf{A}}^{n-1} q^n_{i,j}, \text{ which implies}$$
$$q^n_{i,j} = \left(\frac{1}{\Delta}\mathbf{I} + \widehat{\mu^q_{i,j}} - \widehat{\mathbf{A}}^{n-1}\right)^{-1} \left(\frac{1}{\Delta} q^{n-1}_{i,j}\right)$$

[59] The grid over e is non-uniform. See the numerical methods appendix of Achdou et al. (2017) for details.
[60] The "$n-1$" notation is used because $\widehat{\mathbf{A}}$ is constructed using the price functions from iteration $n-1$.
[61] Again, details in Achdou et al. (2017).

and similarly[62]

$$p_{i,j}^n = \left(\frac{1}{\Delta}\mathbf{I} + \widehat{\mu_{i,j}^p} - \widehat{\mathbf{A}}^{n-1}\right)^{-1} \left(\frac{1}{\Delta}p_{i,j}^{n-1}\right)$$

These equations implicitly define $q_{i,j}^n$ and $p_{i,j}^n$ as functions of information from iteration $n-1$. Parameter Δ is the step size, and governs how quickly the price functions are updated. Convergence is not guaranteed, so Δ cannot be set too large.

4. Evaluate $||q_{i,j}^n - q_{i,j}^{n-1}||$ and $||p_{i,j}^n - p_{i,j}^{n-1}||$ (using a sup-metric). If price functions have converged within a pre-specified tolerance, stop. If not, go to step 2 and repeat.

Once the algorithm has fully converged, I use the final values of $q_{i,j}$ and $p_{i,j}$ to solve for the realized evolution of state variables e_t and $\mathcal{I}_t$ and price functions p_t, q_t (very similar to step 2).

E.2 Kolmogorov Equations

E.2.1 Kolmogorov Forward Equation

Let $g_t(e, \mathcal{I})$ denote a probability density function over e and $\mathcal{I}$. The Kolmogorov forward equation (KF, also known as the Fokker-Planck equation) gives $\frac{\partial}{\partial t}g_t$. In words, the KF describes the time-evolution of distribution g_t. The stationary distribution is any distribution $\bar{g}$ that solves $\frac{\partial}{\partial t}\bar{g} = 0$.

The numerical methods appendix of Achdou et al. (2017) provides a great treatment of KF equations. Readers should refer to this appendix for details. The benefit of the finite-difference methods used to solve for price functions q_t and p_t is that the KF equation comes

[62] Footnote 13 of the Numerical Appendix of Achdou et al. (2017) shows that a reflecting barrier imposes that the derivative at the reflecting barrier is 0. This is correct for q, as $q_e(\underline{e}, \mathcal{I}) = 0$. However, this is not correct for p, since $p_e(\underline{e}, \mathcal{I}) = \frac{p(\underline{e},\mathcal{I})\beta}{1+\underline{e}\beta}$ (see Appendix A.2 for details). With a bit of algebra, one can show that $p(\underline{e}, \mathcal{I}) = \frac{\frac{B\underline{e}}{\gamma\sigma} - q(\underline{e},\mathcal{I})}{1+\frac{B\underline{e}\beta}{\gamma\sigma(1+\underline{e}\beta)} - \frac{\underline{e}\beta}{1+\underline{e}\beta}}$. I impose this lower boundary directly when solving for $p_{i,j}^n$.

"for free". Let $g^t_{i,j}$ denote a discretized distribution over e_t and $\mathcal{I}_t$ at time t. The perceived evolution of $g^t_{i,j}$ is given by:

$$\frac{g^{t+\Delta t}_{i,j} - g^t_{i,j}}{\Delta t} = (\widehat{\mathbf{A}})^{\mathrm{T}} g^{t+\Delta t}_{i,j} \quad \Longrightarrow \quad g^{t+\Delta t} = \left(\mathbf{I} - \Delta t (\widehat{\mathbf{A}})^{\mathrm{T}}\right)^{-1} g^t_{i,j},$$

and the true evolution of $g^t_{i,j}$ is given by:

$$\frac{g^{t+\Delta t}_{i,j} - g^t_{i,j}}{\Delta t} = (\mathbf{A})^{\mathrm{T}} g^{t+\Delta t}_{i,j} \quad \Longrightarrow \quad g^{t+\Delta t} = \left(\mathbf{I} - \Delta t (\mathbf{A})^{\mathrm{T}}\right)^{-1} g^t_{i,j}.$$

Numerical Difficulty Relative to the methods detailed in Achdou et al. (2017), there is one additional difficulty here which is that state variables e_t and $\mathcal{I}_t$ load on the same Brownian shock dZ_t. The reason this creates a difficulty is that the infinitesimal generator for the true process (discretized as $\mathbf{A}$) involves cross-partial derivatives. In particular, for any function $f(\mathcal{S}_t)$, one of the Itô correction terms is $\sigma e_t \sigma^e_t \times \frac{\partial^2}{\partial e \partial \mathcal{I}} f(\mathcal{S}_t)$. The discretization of $\frac{\partial^2}{\partial e \partial \mathcal{I}} f(\mathcal{S}_t)$ looks like:

$$\frac{\frac{f_{i+1,j+1} - f_{i-1,j+1}}{2\Delta e} - \frac{f_{i+1,j-1} - f_{i-1,j-1}}{2\Delta e}}{2\Delta \mathcal{I}}$$

The reason this is an issue is that the discretized KF equation is only guaranteed to preserve the positivity of density $g_{i,j}$ when the diagonal terms of $\mathbf{A}$ are weakly negative (capturing outflows) and the off-diagonal terms are weakly positive (capturing inflows). However, the above discretization of the cross-partial will put negative entries on off-diagonal elements $i-1, j+1$ and $i+1, j-1$. Positivity is no longer guaranteed. However, in practice I've found that this issue is abated in my numerical solution with a dense grid. Solving for the stationary distribution using simulation yields quantitatively similar results.

E.2.2 Kolmogorov Backward Equation

The Kolmogorov backward equation (KB) is used to derive the "hitting probabilities" plotted in Figure 4. Intuitively, the KB can be used to ask: "given the state at time T, how did the economy get here?" One starts with a terminal condition at time T, and then solves backward in time from that terminal condition. Hence, it is termed a "backward equation."

As above, $\mathcal{A}$ denotes the infinitesimal generator of $\mathcal{S}_t$, and $\widehat{\mathcal{A}}$ denotes the infinitesimal generator of $\widehat{\mathcal{S}}_t$. Let $\bar{e}_{crisis}(\mathcal{I})$ denote the boundary of the crisis region for sentiment level $\mathcal{I}$.

For $t \in [0, T]$, let $f(e_t, \mathcal{I}_t, t)$ denote the true probability that the economy currently at $\{e_t, \mathcal{I}_t\}$ enters the crisis region between time t and T. Similarly, let $\widehat{f}(e, \mathcal{I}, t)$ denote the perceived probability. The goal is to solve for $f(e_0, \mathcal{I}_0, 0)$ and $\widehat{f}(e_0, \mathcal{I}_0, 0)$. Conditional probability function $f(e_t, \mathcal{I}_t, t)$ is the solution to:

$$
\begin{aligned}
&0 = f_t(e_t, \mathcal{I}_t, t) + \mathcal{A}f(e_t, \mathcal{I}_t, t), \text{ subject to boundary conditions} \\
&(i)\ f(e_t, \mathcal{I}_t, t) = 1 \text{ if } e_t \leq \bar{e}_{crisis}(\mathcal{I}_t), \text{ and} \\
&(ii)\ f(e_T, \mathcal{I}_T, T) = 0 \text{ if } e_t > \bar{e}_{crisis}(\mathcal{I}_t)
\end{aligned}
$$

This result is stated without proof. Informally, the differential equation $0 = f_t(e_t, \mathcal{I}_t, t) + \mathcal{A}f(e_t, \mathcal{I}_t, t)$ arises from applying Itô's lemma to $f(e_t, \mathcal{I}_t, t)$ and setting the drift of the resulting expression equal to 0. The drift is set to 0 in the non-crisis region due to the law of iterated expectations: $f(e_t, \mathcal{I}_t, t) = \mathbb{E}_t[f(e_{t+dt}, \mathcal{I}_{t+dt}, t+dt)]$ and therefore $\mathbb{E}_t[df(e_t, \mathcal{I}_t, t)] = 0$. The first boundary condition sets f equal to 1 whenever the crisis region has been hit. The second boundary condition is a terminal condition which assigns $f(e_T, \mathcal{I}_T, T) = 0$ if the crisis region is not hit at terminal period T. $f(e_t, \mathcal{I}_t, t)$ is solved backwards from this terminal condition.[63]

[63] Most proofs of the KB equation rely on the continuity of f_T (e.g., Duffie (2010)). This condition is not met, since f_T jumps discretely from 0 to 1 at the crisis boundary. Numerically this is not truly an issue, as one can approximate the discontinuous function f_T arbitrarily closely with a continuous sigmoid.

Similarly, $\widehat{f}(e_t, \mathcal{I}_t, t)$ is the solution to:

$$
\begin{aligned}
0 &= \widehat{f}_t(e_t, \mathcal{I}_t, t) + \widehat{\mathcal{A}}\widehat{f}(e_t, \mathcal{I}_t, t), \text{ subject to boundary conditions} \\
&(i)\ \widehat{f}(e_t, \mathcal{I}_t, t) = 1 \text{ if } e_t \leq \bar{e}_{crisis}(\mathcal{I}_t), \text{ and} \\
&(ii)\ \widehat{f}(e_T, \mathcal{I}_T, T) = 0 \text{ if } e_T > \bar{e}_{crisis}(\mathcal{I}_T)
\end{aligned}
$$

The only difference is that the perceived infinitesimal generator $\widehat{\mathcal{A}}$ is used rather than the true infinitesimal generator.

Numerically, the discretized versions of both $\mathcal{A}$ and $\widehat{\mathcal{A}}$ were already generated when solving for price functions. Finite-difference methods can then be used to solve backward for f and $\widehat{f}$ starting from the terminal condition.

E.3 Verifying Equity Household Portfolio Choice

In the main text I assume that the equity households will choose to invest the maximal amount into the equity of the financial sector. This assumption must be verified ex-post given the resulting model solution.

The household maximizes the value function given by (10), where $C_t = (c_t^y)^{1-\phi}(c_t^h)^{\phi} = (c_t^y)^{1-\phi}$, since $c^h = 1$ in equilibrium. From (10), the household has a per-period utility flow of $\frac{(c_t^y)^{(1-\phi)(1-\gamma_h)}}{1-\gamma_h}$. Applying an affine transformation to the utility function by multiplying it by $\frac{1}{1-\phi}$ (this is always positive since $\phi \in (0,1)$), we see that the household can be represented as having standard power utility with CRRA coefficient $\zeta = 1-(1-\phi)(1-\gamma_h)$. Additionally, market clearing gives that $c_t^y = AK_t - \Phi(i_t, K_t) = \left(A - \delta - \frac{q_t - 1}{\xi} - \frac{(q_t-1)^2}{2\xi}\right) K_t$.

Let A be an arbitrary asset with perceived mean return $\widehat{\mu^A}$ and volatility $\widehat{\sigma^A}$. The

standard asset pricing equation for a CRRA agent in continuous time is:

$$\frac{\widehat{\mu^A} - r_t}{\widehat{\sigma^A}} = \zeta \widehat{\sigma_t^{cy}},$$

where $\widehat{\sigma_t^{cy}}$ is the perceived volatility of $\frac{dc_t^y}{c_t^y}$. Thus, the household demands a perceived Sharpe ratio of $\zeta\widehat{\sigma_t^{cy}}$ for arbitrary asset A. Note that this equation will not hold with equality when there are portfolio restrictions placed on the household, as there are with the equity household's investment in the financial sector's equity.

An investment in the financial sector's equity earns the perceived return $\widehat{d\widetilde{R}_t}$. This can be decomposed into a perceived risk premium of $\alpha^k\widehat{\pi^k} + \alpha^h\widehat{\pi^h}$ and perceived risk $\alpha^k\widehat{\sigma^k} + \alpha^h\widehat{\sigma^h}$. It is the banker that prices capital and housing because households are restricted from direct holding. From equation (22), we can see that the banker demands a risk premium on each asset of:

$$\widehat{\pi_t^k} = \gamma(\alpha_t^k\widehat{\sigma_t^k} + \alpha_t^h\widehat{\sigma_t^h})\widehat{\sigma_t^k}$$
$$\widehat{\pi_t^h} = \gamma(\alpha_t^k\widehat{\sigma_t^k} + \alpha_t^h\widehat{\sigma_t^h})\widehat{\sigma_t^h}$$

For the portfolio as a whole, this implies:

$$\begin{aligned}\alpha^k\widehat{\pi^k} + \alpha^h\widehat{\pi^h} &= \gamma(\alpha_t^k\widehat{\sigma_t^k} + \alpha_t^h\widehat{\sigma_t^h})(\alpha_t^k\widehat{\sigma_t^k} + \alpha_t^h\widehat{\sigma_t^h}) \\ &= \gamma(\alpha_t^k\widehat{\sigma_t^k} + \alpha_t^h\widehat{\sigma_t^h})^2.\end{aligned}$$

Going back to the equity household's portfolio decision, the equity household will invest the maximum in the financial sector's equity whenever the perceived Sharpe ratio on this

investment is weakly greater than $\zeta\widehat{\sigma_t^{cy}}$:

$$\frac{\gamma(\alpha_t^k\widehat{\sigma_t^k}+\alpha_t^h\widehat{\sigma_t^h})^2}{\alpha_t^k\widehat{\sigma_t^k}+\alpha_t^h\widehat{\sigma_t^h}} > \zeta\widehat{\sigma_t^{cy}}, \text{ or equivalently}$$
$$\gamma(\alpha_t^k\widehat{\sigma_t^k}+\alpha_t^h\widehat{\sigma_t^h}) > \zeta\widehat{\sigma_t^{cy}}$$

This condition can be verified numerically over the entire state space.

Intuitively, we can expect this to hold whenever $\gamma > \zeta$. The rate of consumption $\frac{C_t^y}{K_t}$ is countercyclical (see Figure 2), and therefore consumption growth is less volatile than output growth (σ). Since the financial sector is taking a levered position in risky assets, we can typically expect that equity returns are far riskier than consumption growth; i.e., $\alpha_t^k\widehat{\sigma_t^k}+\alpha_t^h\widehat{\sigma_t^h} > \widehat{\sigma_t^{cy}}$.[64] If so, we just need $\gamma > \zeta$ to ensure that the above condition holds.

[64] In the baseline calibration, Figure 2 shows that $\sigma_t^k > \sigma$ and $\sigma_t^h > \sigma$. Plus, $\alpha_t^k > 1$ and $\alpha_t^h > 1$. Thus, we can expect the condition $\alpha_t^k\widehat{\sigma_t^k}+\alpha_t^h\widehat{\sigma_t^h}) > \sigma > \widehat{\sigma_t^{cy}}$ to easily hold in the baseline calibration.

F Diagnostic Expectations

F.1 Ornstein-Uhlenbeck (OU) Properties

Conditional and Unconditional Distributions Repeating equation (3), new information parameter $\mathcal{I}_t$ is defined as:

$$\mathcal{I}_t = \int_0^t e^{-\kappa(t-s)}\sigma dZ_s.$$

I start by characterizing the conditional distribution of $\mathcal{I}_{t+\tau}$ given $\mathcal{I}_t$. $\mathcal{I}_{t+\tau}$ is given by:

$$\begin{aligned}\mathcal{I}_{t+\tau} &= \int_0^{t+\tau} e^{-\kappa(t+\tau-s)}\sigma dZ_s \\ &= \int_0^{t} e^{-\kappa(t+\tau-s)}\sigma dZ_s + \int_t^{t+\tau} e^{-\kappa(t+\tau-s)}\sigma dZ_s \\ &= e^{-\kappa\tau}\mathcal{I}_t + \int_t^{t+\tau} e^{-\kappa(t+\tau-s)}\sigma dZ_s.\end{aligned}$$

Since $\{Z_{t+\tau}\}$ is a standard Brownian motion with independent Gaussian increments, $\mathcal{I}_{t+\tau}|\mathcal{I}_t$ is itself Gaussian. The mean of $\mathcal{I}_{t+\tau}|\mathcal{I}_t$ is given by $e^{-\kappa\tau}\mathcal{I}_t$. By Itô isometry, the variance of $\mathcal{I}_{t+\tau}|\mathcal{I}_t$ is $\frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa\tau}\right)$.

The unconditional distribution of the OU process $\mathcal{I}$ is obtained by taking $\tau \to \infty$. In the limit, the unconditional distribution of $\mathcal{I}$ is Gaussian with mean 0 and variance $\frac{\sigma^2}{2\kappa}$. A textbook treatment is provided in Karatzas and Shreve (1998).

Distribution of First Hitting Times The *first hitting time* of a stochastic process is defined as the time at which a stochastic process first crosses some threshold. The distribution of first hitting times for an OU process can be analytically defined in the special case where the threshold is the OU process' mean. For the mean-zero OU process $\mathcal{I}_t$, this means that one can analytically characterize the persistence of sentiment. Put differently, for any

initial level of $\mathcal{I}_t$, the time $t+\tau$ at which the OU process first returns to 0 has a known distribution. The density of first hitting times is given by:

$$h(\tau|\mathcal{I}_t) = \frac{|\mathcal{I}_t|}{\sigma\sqrt{2\pi}} \left(\frac{\kappa}{\sinh(\kappa\tau)}\right)^{\frac{3}{2}} \exp\left(-\frac{\kappa\mathcal{I}_t^2 e^{-\kappa\tau}}{2\sigma^2\sinh(\kappa\tau)} + \frac{\kappa\tau}{2}\right). \tag{71}$$

This is stated without proof, and details are provided in Alili et al. (2005).

Importance of κ for Sentiment Persistence Let $\mathcal{I}_t^{SD} = \frac{\mathcal{I}_t}{\frac{\sigma}{\sqrt{2\kappa}}}$ denote sentiment normalized by its unconditional standard deviation. Upon normalizing $\mathcal{I}_t$ by its standard deviation, volatility parameter σ drops out of the equation for both the conditional distribution of $\mathcal{I}_{t+\tau}^{SD}$ and the distribution of first hitting times. These distributions become a function of only the persistence parameter κ. This is an alternative way of showing that κ is the critical parameter governing the persistence of sentiment.

In detail, the conditional distribution of $\mathcal{I}_{t+\tau}^{SD}$ given $\mathcal{I}_t^{SD}$ is Gaussian with mean $e^{-\kappa\tau}\mathcal{I}_t^{SD}$ and variance $(1-e^{-2\kappa\tau})$. Similarly, equation (71) can be re-written as:

$$h(\tau|\mathcal{I}_t^{SD}) = \frac{\kappa|\mathcal{I}_t^{SD}|}{2\sqrt{\pi}} \left(\frac{1}{\sinh(\kappa\tau)}\right)^{\frac{3}{2}} \exp\left(-\frac{(\mathcal{I}_t^{SD})^2 e^{-\kappa\tau}}{4\sinh(\kappa\tau)} + \frac{\kappa\tau}{2}\right). \tag{72}$$

Equation (72) shows that the distribution of first hitting times for $\mathcal{I}_t^{SD}$ depends only on the initial level of $\mathcal{I}_t^{SD}$ and on κ, but is independent of σ.

OU Figures: Baseline Calibration

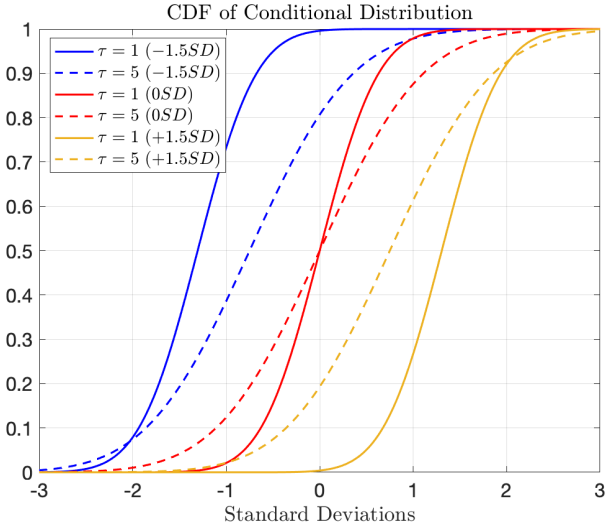


Figure 4: Conditional Distributions. CDF of $\mathcal{I}_{t+\tau}^{SD}$ given $\mathcal{I}_t^{SD} \in \{-1.5SD, 0SD, +1.5SD\}$. 1 and 5-year horizons are plotted.

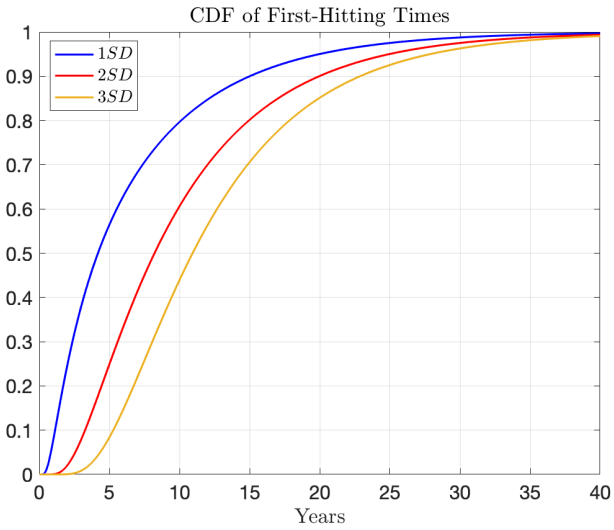


Figure 5: First Hitting Times. CDF of first hitting times for $\mathcal{I}_t \in \{\pm 1SD, \pm 2SD, \pm 3SD\}$.

F.2 Extensions

Objective versus Subjective Shocks The new information measure $\mathcal{I}_t = \int_0^t e^{-\kappa(t-s)}\sigma dZ_s$ is based on objective capital quality shocks σdZ_t. This choice is consistent with BGS, who argue that overreaction to objective news is more consistent with the psychology of diagnostic expectations.

However, it is just as tractable to define new information based on subjective capital quality shocks. An agent with a bias of $\theta\mathcal{I}_t$ will perceive a subjective shock of $\sigma\widetilde{dZ}_t = \sigma dZ_t - \theta\mathcal{I}_t$. In this case, one can define the new information measure as

$$\mathcal{I}_t = \int_0^t e^{-\kappa(t-s)}\sigma\widetilde{dZ}_s.$$

This measure of subjective new information evolves according to:

$$\begin{aligned} d\mathcal{I}_t &= -\kappa\mathcal{I}_t dt + \sigma\widetilde{dZ}_t \\ &= -(\kappa+\theta)\mathcal{I}_t + \sigma dZ_t. \end{aligned} \tag{73}$$

Comparing equation (4) to equation (73), the only difference when using subjective shocks is that $\mathcal{I}_t$ decays more quickly. This is intuitive — an overoptimistic agent will interpret incoming shocks with a negative bias (and vice-versa for an overpessimistic agent), thus leading to a faster unwinding of sentiment.

Multiple Frequencies of Extrapolation In the main text it was assumed that sentiment evolves at just a single frequency, with shocks fading from $\mathcal{I}_t$ at rate κ. However, it may be more psychologically valid to have sentiment operate over multiple frequencies. Indeed, the empirical literature on extrapolative expectations has found extrapolation over a variety of frequencies. At a low frequency, Malmendier and Nagel (2011) find that long-term risk

attitudes are shaped by macroeconomic experiences. Alternatively, Greenwood and Shleifer (2014) find that extrapolative expectations of future stock market returns depends strongly on returns experienced in the past quarter.

The model can easily be modified to capture multiple frequencies of extrapolation. For illustration, consider an agent whose sentiment is a function of a slow-moving component and a fast-moving component. The agent has a low-frequency measure of new information:

$$\mathcal{I}_t^L = \int_0^t e^{-\kappa^L(t-s)}\sigma dZ_s,$$

and a high-frequency measure of new information:

$$\mathcal{I}_t^H = \int_0^t e^{-\kappa^H(t-s)}\sigma dZ_s,$$

with $\kappa^L < \kappa^H$. The overall measure of new information can then be defined as:

$$\mathcal{I}_t = \theta^L \cdot \mathcal{I}_t^L + \theta^H \cdot \mathcal{I}_t^H,$$

where θ^L and θ^H determine the relative strength of low-frequency and high-frequency sentiment. Aggregate measure $\mathcal{I}_t$ has a true and perceived evolution of:

$$\begin{aligned} d\mathcal{I}_t &= \theta^L d\mathcal{I}_t^L + \theta^H d\mathcal{I}_t^H \\ &= \theta^L(-\kappa^L\mathcal{I}_t^L + \sigma dZ_t) + \theta^H(-\kappa^H\mathcal{I}_t^H + \sigma dZ_t), \text{ and} \\ \widehat{d\mathcal{I}_t} &= -\theta^L\kappa^L\mathcal{I}_t^L - \theta^H\kappa^H\mathcal{I}_t^H \end{aligned}$$

The downside of including multiple frequencies of sentiment is that each frequency requires its own state variable. For this reason, only a single frequency was used in the main text.

F.3 Discussion: Naivete of Non-Rationality

The model of diagnostic expectations fundamentally assumes that diagnostic agents are naive about their non-rationality. Starting with the discrete-time model of BGS, it is a generic property of the model that expectations held at $t+1$ are inconsistent with expectations held at t. Thus, it would take just one period for a sophisticated agent to learn the non-rationality of her beliefs.

To see this property, I'll start in discrete time and follow the notation of BGS. Let ω_t be an AR(1) process $\omega_t = b\omega_{t-1} + \epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$. At time t, the diagnostic agent perceives that given ω_{t+1}, $h_t^\theta(\omega_{t+2}|\omega_{t+1}) = \mathcal{N}(b\omega_{t+1}, \sigma^2)$. At time $t+1$, the diagnostic agent instead perceives that $h_{t+1}^\theta(\omega_{t+2}|\omega_{t+1}) = \mathcal{N}(b(\omega_{t+1} + \theta\epsilon_{t+1}), \sigma^2)$. Whenever $\epsilon_{t+1} \neq 0$, these beliefs are inconsistent with one another.

In this paper's continuous-time model, a consequence of arbitrarily short period lengths is that a sophisticated agent could learn in an instant that she is not rational.[65] This highlights the assumptions implicit in the psychology of diagnostic expectations. Diagnostic agents have the true model in the back of their minds and believe that they are making rational predictions. Therefore agents never understand the need – i.e., they lack the metacognition – to re-examine their non-rational beliefs.

Nonetheless, the model of diagnostic expectations developed here could be extended to allow for internally consistent (though misspecified) beliefs. To do so, two changes are required. First, new information variable $\mathcal{I}_t$ must be defined based on subjective shocks:

$$\mathcal{I}_t = \int_0^t e^{-\kappa(t-s)} \sigma \widetilde{dZ}_s.$$

[65] In addition to studying the inconsistency of her beliefs, the agent could learn of her non-rationality by comparing her volatility predictions to reality. Under the Brownian shocks assumed in this paper, the volatility of a process can be observed instantly because the agent observes an arbitrary number of Brownian shocks over any finite time interval. Since diagnostic agents misperceive the volatility of endogenous processes whenever $\zeta \neq 1$ (discussed in Appendix A.3), they could also learn instantly that their volatility predictions are incorrect.

This specification of $\mathcal{I}_t$ is discussed above. Second, the agent must form expectations about the future evolution of capital as if $\mathcal{I}_t$ evolves according to:

$$\widehat{d\mathcal{I}_t} = -\kappa \mathcal{I}_t + \sigma dZ_t.$$

This specification says that the diagnostic agent simulates the economy forward (in a biased way), and she overreacts to simulated future shocks in addition to actual realized shocks. Under this specification, given any perceived path of shocks from time t to time $t+\tau$, the agent's expectations at time $t+\tau$ will be consistent with what she believed at time t she would believe at time $t+\tau$ given the perceived path of shocks. In this case the only remaining bias is the time-varying capital growth bias, which cannot be observed instantaneously.

F.4 Comparison to Bordalo et al. (2018a)

Here I show that the discrete-time analogue of this paper's formulation of diagnostic expectations nests the model of Bordalo et al. (2018a) when applied to exogenous AR(1) processes. Following the notation of BGS, let ω_t be an AR(1) process $\omega_t = b\omega_{t-1} + \epsilon_t$, where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$.

Discrete-Time Setup In discrete time, the analogue of new information measure $\mathcal{I}_t$ is given by:

$$\mathcal{I}_t = \sum_{j=0}^{\infty} \mathcal{K}^j \epsilon_{t-j},$$

where $\mathcal{K} = e^{-\kappa}$ is the discount factor governing the speed of new information decay. The background context is:

$$G_t^- = \omega_t - \mathcal{I}_t.$$

As the agent simulates the economy forward from period t:

$$\widehat{\mathcal{I}}_{t+\tau} = \sum_{j=0}^{\infty} \mathcal{K}^{j+\tau} \epsilon_{t-j}.$$

Analogous to the continuous-time formulation, for any future value $\omega'_{t+\tau}$ the simulated future background context is:

$$\widehat{G}_{t+\tau}^- = \omega'_{t+\tau} - \widehat{\mathcal{I}}_{t+\tau}.$$

Expectations of $\omega_{t+\tau}$ Now using this discrete-time framework, I'll start by expressing ω_t in changes:

$$\omega_t = b\omega_{t-1} + \epsilon_t$$

$$\omega_t - \omega_{t-1} = (b-1)\omega_{t-1} + \epsilon_t$$

Thus, we can think of the "drift" of ω_t as $(b-1)\omega_t$. For any $\mathcal{I}_t$, the diagnostic agent therefore expects that

$$\begin{aligned} h^\theta(\omega_{t+1}) &\sim \mathcal{N}(\omega_t + (b-1)\omega_t + \theta(\omega_t - G_t^-), \sigma^2) \\ &= \mathcal{N}(b\omega_t + \theta\mathcal{I}_t, \sigma^2) \\ &= \mathcal{N}(b\omega_t + \theta\left(\sum_{j=0}^{\infty} \mathcal{K}^j \epsilon_{t-j}\right), \sigma^2) \end{aligned}$$

Simulating forward, for any $\omega'_{t+\tau}$ the perceived evolution from $t+\tau$ to $t+\tau+1$ is given by:

$$\begin{aligned} h^{\theta}(\omega_{t+\tau+1}|\omega'_{t+\tau}) &\sim \mathcal{N}(\check{\omega}_{t+\tau} + (b-1)\omega'_{t+\tau} + \theta(\omega'_{t+\tau} - \widehat{G}^{-}_{t+\tau}), \sigma^2) \\ &= \mathcal{N}(b\omega'_{t+\tau} + \theta\widehat{\mathcal{I}}_{t+\tau}, \sigma^2) \\ &= \mathcal{N}(b\omega'_{t+\tau} + \theta\left(\sum_{j=0}^{\infty}\mathcal{K}^{j+\tau}\epsilon_{t-j}\right), \sigma^2) \end{aligned}$$

BGS Equivalence To see that this nests the AR(1) framework of BGS, take $\mathcal{K} \to 0$. When $\mathcal{K} = 0$, we recover that $G^{-}_{t} = \omega_t - \epsilon_t = b\omega_{t-1}$. In other words, background context G^{-}_{t} doesn't incorporate the most recent shock, but fully incorporates all shocks of further lags. Additionally, $\widehat{G}^{-}_{t+\tau} = \omega'_{t+\tau}$ for all $\tau \geq 1$.

Consider this paper's iterative framework for defining expectations. From t to $t+1$, expectations are biased by $\theta(\omega_t - G^{-}_{t}) = \theta\epsilon_t$. While there is no further bias from $t+1$ onward in the perceived drift of $\omega_{t+\tau}$ for $\tau \geq 1$, the perceived level of $\omega_{t+\tau}$ will still be biased.

In more detail, at time t the agent believes that the distribution at time $t+1$ is $\mathcal{N}(b\omega_t + \theta\epsilon_t, \sigma^2)$. Because there is no additional bias in the perceived drift, the distribution at time $t+2$ is $\mathcal{N}(b\left(b\omega_t + \theta\epsilon_t\right), b^2\sigma^2 + \sigma^2)$, the distribution at time $t+3$ is $\mathcal{N}(b^2\left(b\omega_t + \theta\epsilon_t\right), b^4\sigma^2 + b^2\sigma^2 + \sigma^2)$, etc. Generically, the distribution at time $t+\tau$ is $\mathcal{N}(b^{\tau-1}\left(b\omega_t + \theta\epsilon_t\right), \sum_{i=0}^{\tau-1} b^{2i}\sigma^2)$. This is identical to BGS, with the one difference that my parameter θ is their parameter θ/b.

Moving from Discrete to Continuous Time In going from discrete to continuous time, it no longer makes sense to set $\mathcal{K} = 0$; beliefs must be more slow-moving. When $\mathcal{K} = 0$ only the most recent shock matters. But, the concept of "only the most recent shock" is not robust to the length of the period. This is especially true in continuous time, where the "most recent piece of news" is an instantaneous Brownian increment. Thus, in this paper I instead adopt a sentiment function that is based on multiple lags of past shocks (i.e., $\mathcal{K} > 0$).

G Appendix Figures

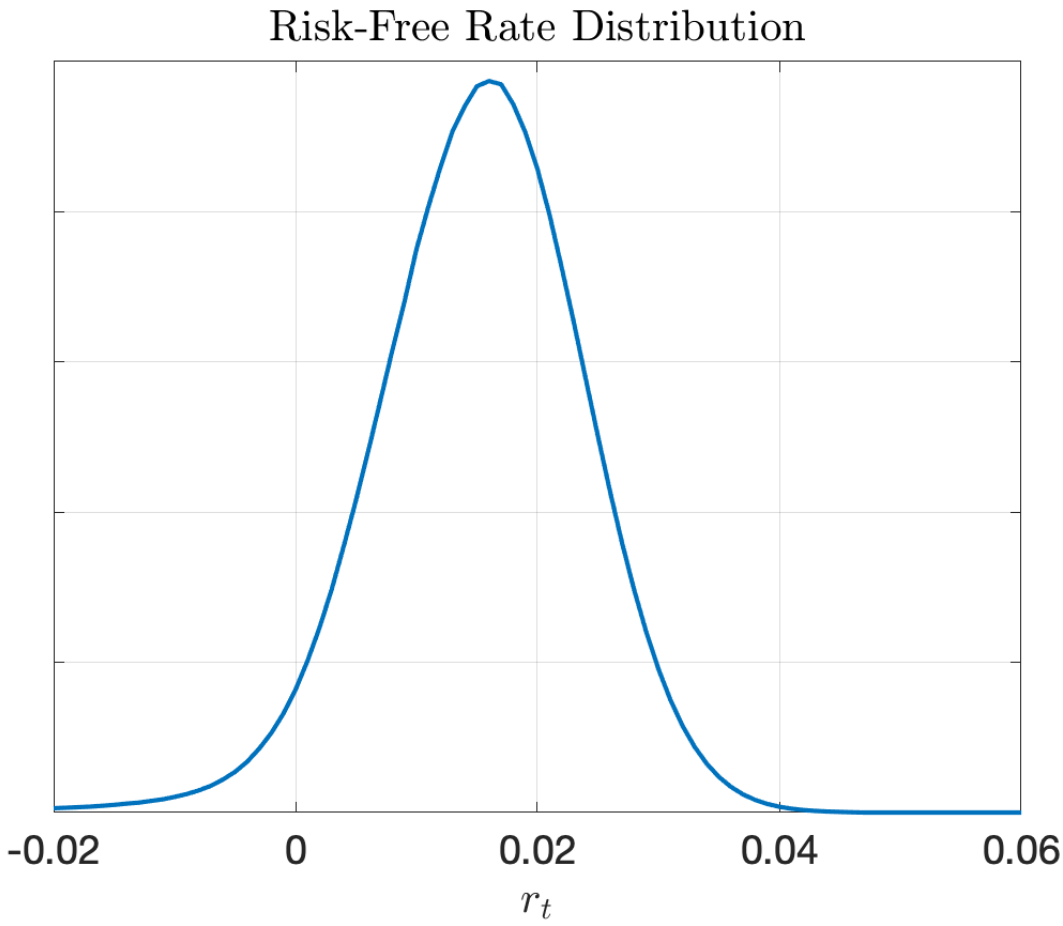


Figure 6: Ergodic distribution of interest rate r_t under the baseline calibration.

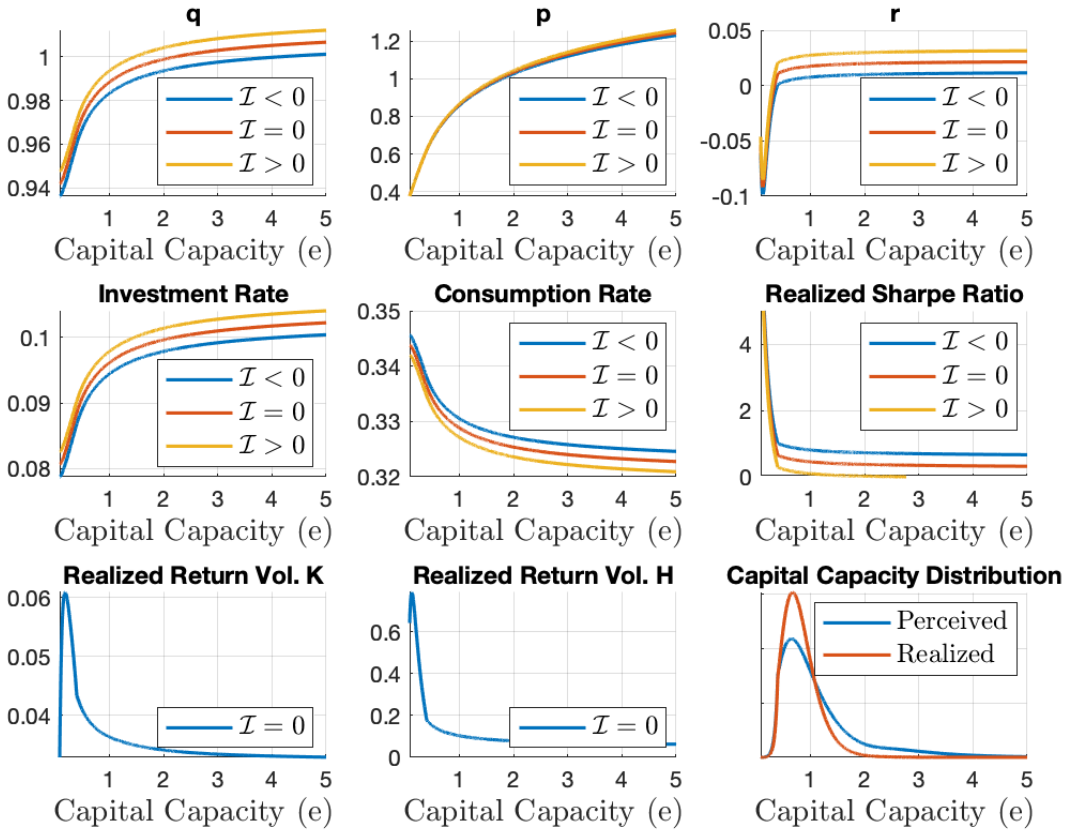


Figure 7: Selected price and policy functions under the baseline calibration. Most panels plot three curves, corresponding to $\mathcal{I} = -1.5SD$ (blue), $\mathcal{I} = 0$ (red), and $\mathcal{I} = +1.5SD$ (yellow).

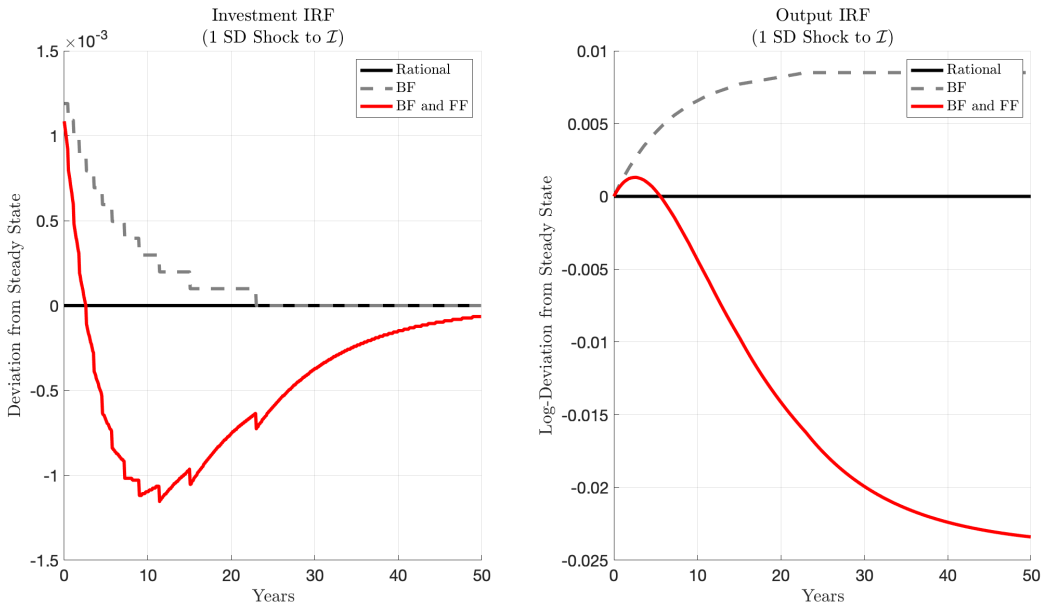


Figure 8: Investment and Output IRFs. The left panel replicates Figure 7. The right panel plots the associated output IRF in these three economies. Note that the output IRF is the integral of the investment IRF. Because this is an AK economy, output deviations are persistent.