A Macro-Finance Model with Sentiment*

Job Market Paper

Latest Version: Available Here

Peter Maxted

Harvard University

November 13, 2020

Abstract

This paper integrates diagnostic expectations into a general equilibrium macroeconomic model with a financial intermediary sector. Diagnostic expectations are a forward-looking model of extrapolative expectations that overreact to recent news. Frictions in financial intermediation produce nonlinear spikes in risk premia and slumps in investment during periods of financial distress. The interaction of sentiment with financial frictions generates a short-run amplification effect followed by a long-run reversal effect, termed the feedback from behavioral frictions to financial frictions. The model features sentiment-driven financial crises characterized by low pre-crisis risk premia and neglected risk. The conflicting short-run and long-run effect of sentiment produces boom-bust investment cycles. The model also identifies a stabilizing role for diagnostic expectations. Under the baseline calibration, financial crises are less likely to occur when expectations are diagnostic than when they are rational.

*Email: petermaxted@g.harvard.edu. I would like to thank Francesca Bastianello, Pedro Bordalo, John Campbell, Daniele d'Arienzo, Emmanuel Farhi, Paul Fontanier, Xavier Gabaix, Nicola Gennaioli, Thomas Graeber, Samuel Hanson, Zhiguo He (discussant), Arvind Krishnamurthy, Spencer Yongwook Kwon, David Laibson, Lu Liu, Matteo Maggiori, Andrei Shleifer, Jakob Ahm Sørensen, Jeremy Stein, Ludwig Straub, Yeji Sung (discussant), Stephen Terry, Emil Verner, and participants at the NBER Behavioral Finance Conference for helpful discussion and comments.

1 Introduction

Models of financial frictions successfully replicate crisis-driven spikes in risk premia and the transmission of financial sector vulnerabilities into aggregate downturns. However, the rational expectations version of these models struggles to match the empirical pattern of low risk premia in the lead-up to crises. Baron and Xiong (2017) find that bank equity is overvalued prior to financial crises due to overoptimistic beliefs. Krishnamurthy and Muir (2020) show that credit spreads are regularly "too low" before financial crises, and that frothy financial markets predict future crises. This conclusion is strengthened in Greenwood et al. (2020a), who document that the combination of rapid credit and asset price growth is highly predictive of future financial crises. All three papers conclude from this pre-crisis evidence that systemic crash risk is neglected during the buildup to crises.

These findings are not unique to crises. Patterns of excessive optimism preceding economic downturns appear consistently (Greenwood and Hanson, 2013, López-Salido et al., 2017). This is supported by direct expectations data from professional forecasters, which displays cyclical overreaction to recent macro-financial trends (Mian et al., 2017, Bordalo et al., 2018b, Bordalo et al., 2020).

In this paper I propose a macro-finance model that is consistent with this evidence of non-rational expectations. To do so, I incorporate diagnostic expectations into a macroeconomic model with a financial intermediary sector. Frictions in financial intermediation allow the model to capture the nonlinear behavior of risk premia during crises and the asymmetric effect of financial fragility on macroeconomic dynamics. However, under rational expectations the model relies on a long sequence of negative shocks to set crises in motion. By introducing behavioral frictions to this model of financial frictions, this paper is able to investigate the sentiment-driven triggers of financial fragility. The contribution of this paper is its characterization of the endogenous financial market and business cycle dynamics that are generated by the interaction of sentiment and financial frictions.

The economic model is based on He and Krishnamurthy (2019). This is a continuous-time RBC model augmented with a financial intermediary sector. Intermediaries are subject to an occasionally binding equity issuance constraint. In non-distress periods the model behaves similarly to a frictionless RBC model. A sequence of poor returns moves intermediaries closer to their constraint and leads to financial distress. In distress periods the nonlinearities arising from financial frictions become quantitatively important. Financial crises are triggered in the tail state where the constraint binds, causing risk premia to spike and asset prices to plummet.

I depart from rational expectations by introducing behavioral frictions to this model of financial intermediary frictions. This paper develops a method for extending rational models with a continuous-time variant of diagnostic expectations. Diagnostic expectations are a forward-looking model of extrapolative expectations in which agents overweight future states that are representative of recent news (Bordalo et al., 2018a). When recent shocks have tended to be positive, agents are overoptimistic about future economic growth. The reverse holds for negative shocks. Diagnostic expectations do not add an independent source of shocks to the model. Instead, diagnostic expectations alter the way that shocks drive the economy in equilibrium.

The interaction of diagnostic expectations with frictions in financial intermediation produces conflicting short-run and long-run dynamics. A sequence of positive shocks alleviates financial frictions. This increases asset prices and investment. In the short run, diagnostic expectations amplify this effect. Positive shocks induce overoptimism about fundamentals, which further elevates asset prices and investment. In the long run, this paper identifies a novel *feedback from behavioral frictions to financial frictions* that reverses the short-run effect. Elevated sentiment induces a progressive erosion of intermediary risk-bearing capacity as expectations are disappointed.

I present three results arising from the interaction of behavioral and financial frictions. First, the model produces sentiment-driven financial crises. Overoptimism dislocates asset prices from fundamentals. The inflation of asset prices initiates a

feedback from behavioral frictions to financial frictions that impairs the risk-bearing capacity of intermediaries and heightens financial fragility in the background of seemingly low-risk environments. Consistent with the empirical findings outlined above, sentiment-driven crises feature low pre-crisis risk premia and neglected crash risk.

Second, the interaction of sentiment with financial frictions produces boom-bust fluctuations in investment and output growth. These cyclical business cycle dynamics are driven by the conflicting short-run and long-run effect of diagnostic expectations. Sentiment-driven amplification in the short run begets its own financial-frictions-driven reversal in the long run.

Third, though diagnostic expectations amplify business cycles, diagnostic expectations can simultaneously stabilize financial cycles. Under the baseline calibration, financial crises are less likely to occur when expectations are diagnostic than when they are rational. This possibility sits in direct opposition to the typical narrative that extrapolative expectations create financial fragility (e.g., the Financial Instability Hypothesis of Minsky (1977)). Since sentiment tracks recent economic shocks, sentiment is typically depressed during periods of financial distress. The long-run feedback effect of depressed sentiment increases intermediary returns relative to expectations. This repairs intermediary balance sheets and reduces the potential for financial distress to erupt into a full-blown crisis.

By interacting sentiment and financial frictions, this paper identifies a long-run feedback from behavioral frictions to financial frictions that is a key driver of equilibrium dynamics. I explore three predictions to evaluate whether this feedback effect improves the model's empirical fit. First, the feedback effect produces long-run reversals. I examine the persistence of financial distress, the price-dividend ratio, and the investment-output ratio. In all three cases, the shorter persistence produced by diagnostic expectations improves the model's fit. Second, the feedback effect implies that elevated sentiment generates financial market overheating that can trigger the emergence of financial fragility. Once this fragility has been triggered, however, it

can persist even after sentiment has subsided. I use this prediction to identify a new fact about financial markets in the buildup to crises: the first eruption of a financial crisis is often preceded by frothy financial markets, but this is rarely the case for residual "double-dip" crises. Third, I evaluate the prediction of sentiment-driven financial crises by applying the model to the 2007-2008 Financial Crisis. This exercise suggests that overoptimism in the mid-2000s was critical for exacerbating financial market vulnerability prior to the failure of Lehman Brothers.

The analysis in this paper relies on two methodological contributions. First, this paper develops a method for applying diagnostic expectations to an endogenous economic growth process. I use this method to generalize the He and Krishnamurthy (2019) model with non-rational expectations. This allows for the study of beliefs without compromising the equilibrium dynamics on which sentiment can interact. Indeed, a key takeaway from this paper is that both behavioral and financial frictions are necessary for understanding the evolution of the economy around financial crises.

Second, this paper highlights the benefits of using global solution methods for characterizing the full equilibrium impact of beliefs on economic dynamics. The analysis of sentiment-driven expansions and slumps fundamentally requires evaluating the cyclical effects of expectations away from the steady state. Global solution methods characterize the complete nonlinear dynamical system.

Related Literature. This paper builds on a large financial frictions literature. Seminal work includes Kiyotaki and Moore (1997) and Bernanke et al. (1999). Recent research often uses continuous-time methods to analyze global dynamics in models with nonlinearities. Examples include Adrian and Boyarchenko (2012), He and Krishnamurthy (2013, 2019), Brunnermeier and Sannikov (2014), Di Tella (2017), Maggiori (2017), and Moreira and Savov (2017).

There is a growing theoretical literature on credit cycles and the behavioral triggers of crises. Bordalo et al. (2018a) develop the original model of diagnostic expectations. Bordalo et al. (2019a) quantify the business cycle implications of diagnostic expec-

tations in a partial equilibrium heterogeneous-firm model. Gennaioli and Shleifer (2018) summarize the research on diagnostic expectations and present a belief-driven narrative of the 2007-2008 Financial Crisis. The authors identify understanding the interaction between beliefs and financial frictions as a key "open problem," which this paper aims to address.

The contribution of this paper is the analysis of diagnostic expectations jointly with frictions in financial intermediation. This allows the model to characterize how sentiment shapes disruptions in financial intermediation, and how this propagates to the real economy. In more recent work, Krishnamurthy and Li (2020) extend the financial crisis model of Li (2020) with non-rational expectations of bank-run shocks in order to quantitatively dissect the mechanisms driving "crisis cycles." Greenwood et al. (2019), Gertler et al. (2020), and Wachter and Kahana (2020) also study the effect of beliefs on booms and busts in financial markets. While this paper contributes to a growing literature studying the behavioral drivers of financial crises, rational explanations have also been forwarded to account for certain pre-crisis patterns (Gomes et al., 2018, Farboodi and Kondor, 2020). Barberis (2018) surveys the research on extrapolative expectations and their application in macro-finance.

This paper also contributes to a budding literature on general equilibrium macroeconomic theory augmented with behavioral features. A small set of recent examples includes Fuster et al. (2012), Woodford (2013), Hirshleifer et al. (2015), Gabaix (2016), Adam and Merkel (2019), Caballero and Simsek (2020), and Farhi and Werning (2020).

2 Macro-Finance Model

This paper embeds diagnostic expectations into the macro-finance model of He and Krishnamurthy (2019, henceforth HK). The HK model is one of the first quantitative papers in the continuous-time macro-finance literature, and it successfully replicates the downside macroeconomic risks precipitated by disruptions in financial intermedi-

ation. However, because the HK model features rational expectations, risk premia in this model are always high when financial crises are likely, whereas in the data they are often low (e.g., Baron and Xiong, 2017, Krishnamurthy and Muir, 2020).

I extend HK in two ways. Most importantly, I generalize the model to allow for diagnostic expectations. I also introduce a simple labor income margin to improve the model's quantitative fit. I adopt HK's notation when possible.

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 assets in positive supply: productive capital K_t and housing H.[1] The housing supply is fixed and normalized to $H \equiv 1$. The price of capital is denoted q_t, and the price of housing is denoted P_t. Asset prices are endogenous and will be determined in equilibrium.

Only intermediaries can directly hold K_t and H.[2] Intermediaries fund these purchases by issuing debt and equity to households. Each intermediary faces an "equity capital constraint" that restricts its ability to raise equity funding. This is the key financial friction. When binding, intermediaries must replace their equity funding with additional debt funding.

Output flow Y_t is produced according to an "AK" production technology

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

Parameter A determines the productivity of capital. Capital evolves endogenously according to:

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

[1] He and Krishnamurthy (2019) use two assets to improve the model's quantitative predictions. Appendix B.3 provides further discussion.

[2] The model can be extended to allow households to directly hold a share of K_t and H, but He and Krishnamurthy (2019) conclude that the quantitative impact is negligible.

i_t is the endogenous rate of capital installation at time t, δ is the exogenous depreciation rate, and $\{Z_t\}$ is a standard Brownian motion. The term σdZ_t is a capital quality shock. Capital quality shocks 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 is given by $\Phi(i_t, K_t) = i_t K_t + \frac{\xi}{2}(i_t - \delta)^2 K_t$.

2.2 Diagnostic Expectations

Overview. All agents have diagnostic expectations about the log capital stock.[3] Capital is the fundamental in this economy — capital alone determines output ($Y_t = AK_t$), and capital quality shocks are the only source of uncertainty.

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) In the context of expectations, the reference class reflects the absence of new information. Representative future states are those that become more likely to occur in light of incoming data.

This paper features two innovations on the original diagnostic expectations model of Bordalo et al. (2018a). First, diagnostic expectations are cast in continuous time. Second, the methodology developed here allows diagnostic expectations to be applied to the endogenous capital process.[4] This will imply not only that recent economic performance affects expectations (a standard feature of extrapolative expectations), but also that expectations feed back into the dynamical system to alter the future evolution of the capital process on which expectations are formed.

This section presents the reduced-form specification of diagnostic expectations as applied to the macroeconomic model. Appendix A.1 provides a microfoundation. This paper's specification is designed portably so that diagnostic expectations can

[3] Log capital evolves according to $dk_t = (i_t - \delta - \frac{\sigma^2}{2})dt + \sigma dZ_t$.

[4] The Bordalo et al. (2018a) model only applies to exogenous AR(N) processes.

be added to rational models using a single additional state variable. Even though I introduce an additional state variable to the HK model, I do not introduce any additional shocks. Capital quality shocks remain the sole driving shock in the model.

Expectations of Capital. The psychology of diagnostic expectations is as follows. Agents have in the back of their mind all necessary information to form correct expectations. However, limited and selective memory means that representative states come to mind more easily. Representative states are those that are diagnostic of incoming data, which in this model corresponds to recent capital quality shocks. This is formalized in the following measure of "recent 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 at rate κ as shocks occur further in the past.[5] $\mathcal{I}_t > 0$ when recent shocks have tended to be positive, and $\mathcal{I}_t < 0$ when recent shocks have tended to be negative. $\mathcal{I}_t$ drifts back to 0 at rate κ in the absence of new shocks.

Throughout, I will use hat notation to denote the beliefs of diagnostic agents. The current period is t, and let $\tau \geq 0$ denote a prediction horizon. Diagnostic agents believe that capital in period $t + \tau$ evolves according to:

$$\frac{\widehat{dK}_{t+\tau}}{K_{t+\tau}} = (i_{t+\tau} - \delta + \theta \mathcal{I}_t e^{-\kappa\tau})dt + \sigma dZ_{t+\tau}. \tag{4}$$

Parameter θ governs the extent to which agents judge by representativeness. $\theta = 0$ recovers rationality. When $\theta > 0$, expectations of capital growth are biased toward states that are diagnostic of recent information. For this reason, information parameter $\mathcal{I}_t$ will be referred to as "sentiment" henceforth.

[5] $\mathcal{I}_t$ is an Ornstein-Uhlenbeck process. In discrete time, any individual shock can itself represent new information. In continuous time, I integrate over the past sequence of shocks because each individual shock σdZ_s has only an infinitesimal effect on the capital stock. My specification is similar to the definition of sentiment in Barberis et al. (2015).

Consider equation (4) when $\tau = 0$: $\frac{\widehat{dK}_t}{K_t} = (i_t - \delta + \theta\mathcal{I}_t)dt + \sigma dZ_t$. Diagnostic expectations bias the perceived growth rate of capital by $\theta\mathcal{I}_t$. At more distant horizons ($\tau > 0$), this bias dissipates at rate κ. Information that is diagnostic of economic conditions at time t slowly dims as the agent forms expectations about the economy's evolution in the more distant future.

Equation (4) should be interpreted as an "as if" process. Diagnostic agents do not consciously calculate the evolution of capital in a biased way. Agents have the true model in their memory database, but are exceedingly drawn to representative states. $\mathcal{I}_t$ is an unconscious internal parameter that characterizes the state of representativeness at time t.

Figure 1 provides an illustrative example of diagnostic expectations applied to an arithmetic Brownian motion (ABM).[6] The blue line plots the realized sample path of the ABM up to time t. The dashed black line plots the rational prediction of the ABM's future path. The solid red line plots the diagnostic prediction. Because recent shocks have tended to be positive, diagnostic expectations are biased upward.

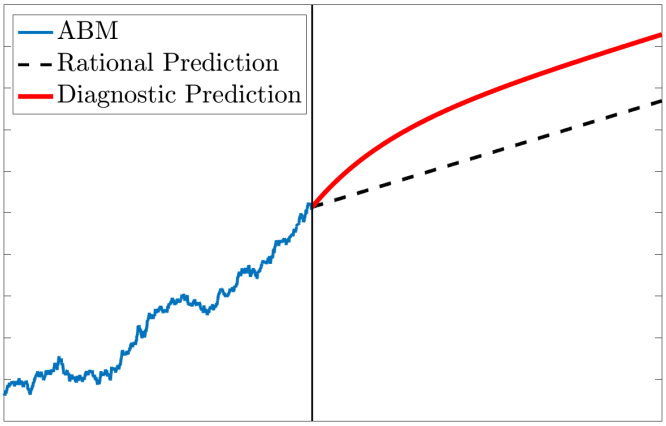


Figure 1: **Diagnostic expectations of arithmetic Brownian motion**. The blue line plots the sample path of an arithmetic Brownian motion (ABM). The solid red line plots diagnostic expectations of the ABM's future evolution, and the dashed black line plots rational expectations. The calibration is illustrative.

[6]Log capital k_t would follow arithmetic Brownian motion if i_t were constant.

Decomposing Diagnostic Expectations. This paper's specification of diagnostic expectations reconciles dynamic, forward-looking, expectations with mechanical models of extrapolation. A decomposition of the perceived capital process highlights this property:

$$\frac{\widehat{dK}_{t+\tau}}{K_{t+\tau}} = \underbrace{\frac{dK_{t+\tau}}{K_{t+\tau}}}_{\text{Rational}} + \underbrace{\theta \mathcal{I}_t e^{-\kappa\tau} dt}_{\text{Wedge}}. \tag{5}$$

The rational component of expectations 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 are characterized by the "kernel of truth" property: expectations depend on the true economic process, but overreact to recent patterns in the data.[7]

2.3 The Financial Intermediary Sector

Individual Intermediaries. There is a continuum of financial intermediaries, each run by a single banker. Intermediaries raise funds from households by issuing risk-free (instantaneous) debt and risky equity. Equity issuance is subject to a constraint. Each intermediary can issue up to ϵ_t of equity. Constraint ϵ_t evolves as follows:

$$\frac{d\epsilon_t}{\epsilon_t} = d\widetilde{R}_t. \tag{6}$$

$d\widetilde{R}_t$ denotes the instantaneous return on the intermediary's equity at time t. Following HK, constraint ϵ_t can be thought of as the intermediary's "reputation." Poor investment returns damage the intermediary's reputation and inhibit its ability to issue equity in the future.

The banker does not consume. Instead, the banker has mean-variance preferences

[7] The decomposition highlights the extent to which diagnostic expectations are robust to the Lucas critique. Diagnostic expectations are forward looking and dependent on the underlying economic model. However, diagnostic expectations are still subject to persistent extrapolative errors due to the diagnostic wedge.

over their intermediary's reputation:

$$\widehat{\mathbb{E}}_t\left[\frac{d\epsilon_t}{\epsilon_t}\right] - \frac{\gamma}{2}\widehat{Var}_t\left[\frac{d\epsilon_t}{\epsilon_t}\right] = \widehat{\mathbb{E}}_t\left[d\widetilde{R}_t\right] - \frac{\gamma}{2}\widehat{Var}_t\left[d\widetilde{R}_t\right]. \tag{7}$$

Hat-notation is used in equation (7) to indicate that the banker has diagnostic expectations of the return process $d\widetilde{R}_t$.

The reputational constraint behaves similarly to the standard net-worth constraint in which an intermediary's ability to raise capital depends on its net worth. The benefit of the reputational constraint is that it produces a more conventional calibration of the HK model.[8]

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

$$\frac{d\mathcal{E}_t}{\mathcal{E}_t} = d\widetilde{R}_t - \eta dt + d\psi_t. \tag{8}$$

All intermediaries behave identically. The term $d\widetilde{R}_t$ implies that the aggregate constraint evolves with the reputation of each individual intermediary. Parameter η governs the exogenous exit rate of intermediaries. Exit is needed to ensure that intermediaries do not escape their equity issuance constraint in equilibrium. The term $d\psi_t \geq 0$ reflects entry into the banking sector. Entry occurs deep in crisis times when reputation is sufficiently low, and establishes a boundary condition for the model. Details are provided in Appendix B.4.

2.4 The Household Sector

Consumption. There is a unit measure of households. Households consume 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 are rented at price D_t.

[8] With the reputation-based constraint, bankers do not consume. This allows the representative household to consume all of the economy's output, as is standard in macroeconomic models without financial frictions. For further details, see Section I.B of He and Krishnamurthy (2019).

Households maximize the value function

$$\widehat{\mathbb{E}}\left[\int_t^\infty e^{-\rho(s-t)}\frac{C_s^{1-\gamma_h}}{1-\gamma_h}ds\right], \tag{9}$$

where C_t is a Cobb-Douglas consumption aggregator $C_t = (c_t^y)^{1-\phi}(c_t^h)^\phi$. Intratemporal maximization yields:

$$\frac{c_t^y}{c_t^h} = \frac{1-\phi}{\phi}D_t. \tag{10}$$

Labor Income. Households can supply up to one unit of labor, without disutility, at wage $\mathcal{W}_t$. In equilibrium, households earn share $1-\nu$ of output as labor income:

$$\mathcal{W}_t = (1-\nu)AK_t. \tag{11}$$

Here I take this wage equation as given. A microfoundation is provided in Appendix B.1. In addition to diagnostic expectations, this stylized labor income margin is where my model differs from HK. The benefit of introducing labor income is that it produces more realistic consumption and investment output shares. See Appendix B.2 for a full discussion.

Capital Production. Investment follows q-theory. There exists a capital producer who is responsible for investment. The capital producer solves $\max_{i_t} q_t i_t K_t - \Phi(i_t, K_t)$. All profits are passed on to households. This results in an equilibrium investment rate of:

$$i_t = \delta + \frac{q_t - 1}{\xi}. \tag{12}$$

Equation (12) highlights the propagation of behavioral and financial frictions from financial markets to the real economy. The economy's growth rate depends on i_t, so these frictions influence economic growth through their effect on q_t.

2.5 Portfolio Choice and Asset Returns

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

Each household is split into a "debt member" and an "equity member." The debt member can only invest in the risk-free debt of intermediaries. The equity member is free to purchase intermediary equity (but cannot make levered investments). At the start of each period, the debt member is given share λ of wealth and the equity member is given share $1-\lambda$. Investments pay off at time $t+dt$, and returns are pooled before this process is repeated.

The model will be calibrated such that equity members collectively invest their allocated wealth of $(1-\lambda)W_t$ in intermediary equity, subject to the restriction that they do not purchase more than $\mathcal{E}_t$.[9] If the constraint binds, equity members place their remaining wealth in bonds. The total equity capital raised by the intermediary sector at time t is therefore

$$E_t \equiv \min\{\mathcal{E}_t,\ (1-\lambda)W_t\}. \tag{13}$$

Risk-free rate r_t is pinned down by households' intertemporal optimization:

$$r_t = \rho + \zeta\widehat{\mathbb{E}}_t\left[\frac{dc_t^y}{c_t^y}\right] - \frac{\zeta(\zeta+1)}{2}\widehat{Var}_t\left[\frac{dc_t^y}{c_t^y}\right]. \tag{14}$$

Parameter $\zeta = 1-(1-\phi)(1-\gamma_h)$ can be interpreted as the inverse of the elasticity of intertemporal substitution (EIS). Again, hat notation is used because household

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

expectations of the consumption process are diagnostic. Equation (14) is the standard consumption-based risk-free rate formula in continuous time.[10]

Intermediary Portfolio Choice. Diagnostic agents may not have correct beliefs about equilibrium asset returns. I postulate that agents expect q_t and P_t to evolve according to:

$$\frac{\widehat{dq_t}}{q_t} = \widehat{\mu_t^q} dt + \widehat{\sigma_t^q} dZ_t, \tag{15}$$

$$\frac{\widehat{dP_t}}{P_t} = \widehat{\mu_t^P} dt + \widehat{\sigma_t^P} dZ_t. \tag{16}$$

These endogenous processes will be determined in equilibrium.

Using (15), the return on an investment in capital is perceived to be:

$$\widehat{dR_t^k} = \left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta \mathcal{I}_t + \sigma \widehat{\sigma_t^q} \right) dt + \left(\sigma + \widehat{\sigma_t^q} \right) dZ_t. \tag{17}$$

The perceived return on capital consists of a dividend component $\left(\frac{\nu A}{q_t} dt\right)$ and a capital gains component $\left(\frac{\widehat{d(q_t K_t)}}{q_t K_t} - i_t dt\right)$. Equation (17) illustrates how diagnosticity affects expectations of capital returns. First, there is a direct effect: capital growth expectations are biased by $\theta \mathcal{I}_t$.[11] Second, diagnostic agents misjudge how the economy evolves in equilibrium. This introduces an indirect effect in which diagnostic agents misperceive the endogenous drift and volatility of q_t.

Proceeding similarly, the perceived return on housing is:

$$\widehat{dR_t^h} = \left(\frac{D_t}{P_t} + \widehat{\mu_t^P} \right) dt + \widehat{\sigma_t^P} dZ_t. \tag{18}$$

[10] To generate equation (14), it is assumed that any marginal savings are given to the debt member. The benefit of this assumption is that it recovers the standard continuous-time risk-free rate formula. See footnote 5 of He and Krishnamurthy (2019) for details.

[11] For empirical evidence on the extrapolation of fundamentals by financial market professionals, see Greenwood and Hanson (2013), Fahlenbrach et al. (2017), Bordalo et al. (2018b), Bordalo et al. (2019b, 2020), Gulen et al. (2019), Nagel and Xu (2019), and Pflueger et al. (2019).

The dividend on housing is given by rental income D_t. Price P_t is the present discounted value of these cash flows. Diagnosticity biases expectations of housing rent growth, which produces non-rational expectations of price process P_t.

Let $\widehat{\pi_t^k} \equiv \left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta \mathcal{I}_t + \sigma \widehat{\sigma_t^q}\right) - r_t$ denote the perceived risk premium on capital. Let $\widehat{\pi_t^h} \equiv \left(\frac{D_t}{P_t} + \widehat{\mu_t^P}\right) - r_t$ denote the perceived risk premium on housing. Equations (17) and (18) can be rewritten as follows:

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

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

Let α_t^k and α_t^h denote the intermediary's portfolio share of capital and housing, respectively. The intermediary's perceived return on equity is:

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

From the objective in equation (7), the intermediary solves:

$$\max_{\alpha_t^k, \alpha_t^h} \left[r_t + \alpha_t^k \widehat{\pi_t^k} + \alpha_t^h \widehat{\pi_t^h}\right] - \frac{\gamma}{2} \left(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h}\right)^2. \tag{19}$$

This results in the optimality condition:

$$\frac{\widehat{\pi_t^k}}{\widehat{\sigma_t^k}} = \frac{\widehat{\pi_t^h}}{\widehat{\sigma_t^h}} = \gamma(\alpha_t^k \widehat{\sigma_t^k} + \alpha_t^h \widehat{\sigma_t^h}). \tag{20}$$

Equation (20) says that intermediaries choose portfolio shares in order to equate the perceived Sharpe ratio on each asset to their risk aversion times their perceived portfolio risk. When intermediaries are required to bear additional risk, they demand higher Sharpe ratios as compensation.

2.6 Summary: Financial Frictions and Behavioral Frictions

Financial Frictions: Constraints and Crises. Following HK, crises are defined as states in which the equity issuance constraint binds: $\mathcal{E}_t < (1-\lambda)W_t$. When the constraint binds, the economy exhibits a dramatic increase in risk premia, a collapse in asset prices, and impaired economic growth. Crisis nonlinearities arise for two reasons, as can be seen with equation (20). First, a binding equity issuance constraint means that intermediaries are forced to increase leverage in order to fund asset purchases.[12] Second, a binding constraint endogenously amplifies the sensitivity of asset prices to negative shocks. Negative shocks cause the constraint to bind even more tightly, thereby increasing leverage and risk premia even further. Increased leverage is reflected in portfolio shares α_t^k and α_t^h. The amplification of shocks is reflected in volatility coefficients $\widehat{\sigma_t^k}$ and $\widehat{\sigma_t^h}$. As equation (20) shows, both of these effects increase the risk premia that intermediaries demand.

Short-Run Behavioral Frictions: Amplification. A series of positive shocks alleviates financial frictions. This raises asset prices and promotes investment. The reverse holds for negative shocks. Diagnostic expectations amplify the short-run impact of shocks, since shocks simultaneously shift sentiment about economic fundamentals. Bordalo et al. (2020) provide direct evidence of this effect, documenting that stock market analysts' expectations of aggregate earnings are overoptimistic, and asset prices are inflated, following positive fundamental news.

Long-Run Behavioral Frictions: Reversal. This paper identifies a novel long-run effect of behavioral frictions that works in opposition to the short-run effect. The long-run effect is called the *feedback from behavioral frictions to financial frictions*.

The intermediary sector's capital capacity evolves with realized returns $d\widetilde{R}_t$. However, intermediaries price assets according to the perceived return process $\widehat{d\widetilde{R}_t}$. When

[12] The model predicts that the market leverage of financial intermediaries is countercyclical. For empirical evidence, see He et al. (2017).

the perceived return process differs from the true return process, market prices will not reflect fundamentals. In the case of elevated sentiment, persistent forecast errors lower realized returns and cause intermediaries' capital capacity to deteriorate relative to expectations. Thus, overoptimism induces a gradual tightening of financial frictions. Alternatively, excessive pessimism gradually relaxes financial frictions. This long-run feedback from behavioral frictions to financial frictions will be a key mechanism driving many of the model's main predictions.

3 Equilibrium and Model Calibration

3.1 Equilibrium

Definition 1. ***Diagnostic Expectations Equilibrium (DEE).*** *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* (10), (12), (14), *and* (19) *are optimal under diagnostic expectations.*

2. *The goods market and housing rental market clear (using* C^y *and* C^h *to indicate aggregate household consumption):*

$$\begin{aligned} Y_t = AK_t = C_t^y + \Phi(i_t, K_t), \ \textit{and} \\ C_t^h = H \equiv 1. \end{aligned} \tag{21}$$

3. *The equity issuance constraint is satisfied:*

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

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

$$q_t K_t = \alpha_t^k E_t, \text{ and} \tag{22}$$

$$P_t = \alpha_t^h E_t. \tag{23}$$

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

$$W_t = q_t K_t + P_t.$$

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 (REE).*** *A rational expectations equilibrium is a diagnostic expectations equilibrium for $\theta = 0$.*

Solution Strategy. I consider Markov equilibria in state variables K_t, $\mathcal{E}_t$, and $\mathcal{I}_t$. K_t scales the size of the economy, $\mathcal{E}_t$ is the financial sector's capital capacity, and $\mathcal{I}_t$ characterizes sentiment. HK use K_t and $\mathcal{E}_t$ as state variables in their rational model. The innovation of this paper is to capture behavioral frictions with state variable $\mathcal{I}_t$. When expectations are extrapolative it is not enough to know the current state of the economy (K_t and $\mathcal{E}_t$); one must also know the path taken to get there ($\mathcal{I}_t$).

The solution can be simplified further by scaling the economy by K_t. Define

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

e_t captures the capital capacity of the intermediary sector relative to the size of the overall economy. I 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 model is solved numerically as a function of e_t and $\mathcal{I}_t$: e_t characterizes financial frictions, and $\mathcal{I}_t$ characterizes behavioral frictions.

In the class of Markov equilibria considered here, each diagnostic expectations equilibrium nests its corresponding rational expectations equilibrium. When solving for a DEE, this means that the REE comes "for free." Formally:

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

Proof. Equation (4) specifies that when $\mathcal{I}_t = 0$, agents act as if $\mathcal{I} = 0$ 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. □

3.2 Calibration

The HK model is a standard RBC model augmented with a financial intermediary sector. The economy behaves like an RBC model when e_t is far from the constraint, and intermediary frictions become quantitatively important near the crisis region. I follow HK in defining $e_{distress}$ as the 33$^{\text{rd}}$ percentile value of e_t in the model's stationary distribution. $e_{distress}$ separates "normal" periods from periods of financial distress.

Table 1 presents the baseline calibration. I use the parameters and/or calibration targets of He and Krishnamurthy (2019) when possible. 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 target is the same. The only parameters for which neither the value nor the target aligns with HK are the three new parameters. These are behavioral parameters θ and κ, and labor income parameter ν.

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

Parameter σ governs the volatility of capital quality shocks. 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

	Parameter	Choice	Target
Panel A: Intermediation Parameters			
γ	Banker Risk Aversion	2^*	Mean Sharpe Ratio
λ	Debt Ratio	0.75^*	Intermediary Leverage
η	Bank Exit Rate	0.13	Prob(Crisis)*
$\underline{e}$	Lower Entry Barrier	0.081	Max $\mathcal{I}=0$ Sharpe Ratio*
β	Entry Cost	2.8^*	Land Price Volatility
Panel B: Technology Parameters			
σ	Capital Shock Volatility	$3\%^*$	C and I Volatility
δ	Depreciation Rate	$10\%^*$	Literature
ξ	Adjustment Cost	3^*	Literature
A	Productivity	$\frac{1}{3}$	Consumption-Output Ratio*
ν	Capital Share	0.41	Investment-Capital Ratio
Panel C: Household Preference Parameters			
ρ	Time Discount Rate	$2\%^*$	Literature
$1/\zeta$	EIS	1.5	Interest Rate Volatility*
ϕ	Housing Expenditure Share	0.2	Housing-Wealth Ratio*
Panel D: Diagnostic Expectations Parameters			
κ	Decay of New Information	0.139	Corr$(e,\mathcal{I})$
θ	Diagnosticity	0.132	Bordalo et al. (2018b)
Unconditional Simulated Moments			
	Mean$\left(\frac{\text{Investment}}{\text{Capital}}\right)$		9.85%
	Mean$\left(\frac{\text{Consumption}}{\text{Output}}\right)$		70.43%
	Mean(Realized Sharpe Ratio)		0.48
	Mean(Realized Intermediary Risk Premium)		15.30%
	Probability of Crisis		3.22%
	Volatility(Land Price Growth)		11.16%
	Volatility(Interest Rate)		0.87%
	Corr$(e,\ \mathcal{I})$		0.78
Non-Distress Simulated Moments			
	Volatility(Investment Growth)		4.85%
	Volatility(Consumption Growth)		2.19%
	Volatility(Output Growth)		2.96%
	Mean$\left(\frac{\text{Housing Wealth}}{\text{Total Wealth}}\right)$		45.37%

Table 1: **Baseline calibration**. Model-generated moments are calculated by simulating the model at a monthly frequency. Growth rates are computed as log changes from month $t-6$ to month $t+6$.

1.24%. In the model, $\sigma = 3\%$ generates investment growth volatility of 4.8% and consumption growth volatility of 2.2% in non-distress periods. The model features too much consumption volatility and too little investment volatility, with $\sigma = 3\%$ balancing the two inaccuracies.

Parameters A and ν are calibrated jointly to target the average investment-capital ratio and consumption-output ratio. The targeted investment-capital ratio is 10% so

that average investment matches depreciation. Since the dividend yield on capital is $\frac{\nu A}{q_t}$, ν and A are calibrated to generate a capital price q commensurate with a 10% investment-capital ratio. To separately identify A and ν, I also target an average consumption-output ratio of 70%. The consumption-output ratio equals $\frac{AK_t - \Phi(i_t, K_t)}{AK_t} \approx 1 - \frac{i_t}{A}$. An investment-capital ratio of 10% pins down $A = \frac{1}{3}$.

Intermediation Parameters. Parameter γ represents the bankers' risk aversion. As in HK, I set $\gamma = 2$. This generates an average realized Sharpe ratio of 0.48, and an average realized intermediary risk premium of 15.30%. This aligns with He et al. (2017), who estimate an average Sharpe ratio of 0.48 and an average return of 13% for assets intermediated by the financial sector.

Intermediary leverage is governed by λ. Since intermediaries have assets of $P_t + q_t K_t = W_t$ and equity of E_t, equation (13) gives a market leverage value of $\frac{W_t}{E_t} = \frac{1}{1-\lambda}$ in non-crisis states. Again following HK, I set $\lambda = 0.75$. This generates a leverage ratio of 4 when the constraint does not bind.

Crisis Parameters. Crises are defined as states in which the equity issuance constraint binds. Bank exit rate η targets a 3% crisis probability.

Parameters $\underline{e}$ and β control the lower boundary condition, represented by $d\psi_t$ in equation (8) (details in Appendix B.4). Parameter $\underline{e}$ is the minimum level of capital capacity at which new entry occurs, and β is the cost of entry into the intermediary sector. HK set $\underline{e}$ such that the Sharpe ratio at $\underline{e}$ is 6.5 ($\underline{e}$ is set low enough that entry occurs rarely). Accordingly, I set $\underline{e}$ such that the perceived Sharpe ratio at $\underline{e}$ and $\mathcal{I} = 0$ is 6.5. Parameter β determines the slope of house price P_t at $\underline{e}$, which in turn affects the volatility of P_t throughout the distress region. As in HK, I set $\beta = 2.8$. HK estimate that the empirical volatility of land price growth from 1975 to 2015 is 11.9%. In the model, $\beta = 2.8$ generates land price growth volatility of 11.2%.

Household Parameters. Parameter ϕ governs the relative value of housing services to the output good. This determines the rental rate D_t (see equation (10)). P_t

is the discounted value of these rental payments. I set $\phi = 0.2$ to target a non-distress housing-wealth ratio of 45%. $\phi = 0.2$ also generates a housing services to total consumption expenditure ratio that is consistent with NIPA consumption data (Davis and Van Nieuwerburgh, 2015).

ζ is the inverse of the EIS, and determines the responsiveness of the risk-free rate to expected consumption growth and volatility. When expectations are diagnostic, agents misperceive the equilibrium consumption process. Thus, ζ also governs the sensitivity of the risk-free rate to variation in sentiment. The EIS plays an important role when expectations are diagnostic, because the EIS regulates the extent to which sentiment gets incorporated into asset prices. When $\zeta = 1$, any bias in growth expectations is passed one-for-one into discount rate r_t (see equation (14)). An important implication is that when $\zeta = 1$, asset prices q_t and P_t are independent of $\mathcal{I}_t$ — all bias in cash-flow expectations is exactly offset by the risk-free rate. When $\zeta < 1$, r_t responds less than one-for-one to biased growth expectations. In this case, q_t and P_t are increasing in $\mathcal{I}_t$.

I set the EIS equal to 1.5. This is a standard choice in the finance literature (e.g., Bansal and Yaron, 2004), and, as in HK, generates real interest rate volatility of roughly 1%. Since $\zeta < 1$, asset prices are increasing in $\mathcal{I}_t$. This will be important for generating the results in Section 5.2.

Behavioral Parameters. θ governs the extent to which expectations are biased by representativeness. κ governs the persistence of sentiment. θ and κ are calibrated jointly using two targets. The first target aligns the magnitude of the expectations bias with the estimates of Bordalo et al. (2018b): one standard deviation in sentiment generates an output growth bias of 0.75 percentage points.[13] The second target matches the model's unconditional correlation between state variables e_t and $\mathcal{I}_t$ to the correlation between intermediary capital and sentiment estimated empirically. To estimate this correlation in the data, I measure e_t with the "Intermediary Capital

[13]Note that $Var(\theta\mathcal{I}_t) = \frac{\theta^2\sigma^2}{2\kappa}$. This calibration target sets $\theta\frac{\sigma}{\sqrt{2\kappa}} = 0.0075$.

Ratio" of He et al. (2017). $\mathcal{I}_t$ can be calculated using expectations data, given κ and θ. Full calibration details are provided in Appendix A.2.

These two calibration targets produce $\theta = 0.132$ and $\kappa = 0.139$. $\kappa = 0.139$ implies that sentiment is slow moving, with a half-life of 5 years.[14] Slow-moving sentiment captures prolonged periods of relatively positive and negative news, such as the Great Moderation, rather than high-frequency volatility.[15] Appendix B.6 examines robustness to parameters κ, θ, and ζ.

4 Global Solution

4.1 Prices, Policy Functions, and Forecast Errors

Select price and policy functions for the DEE are shown in Figure 2.[16] The horizontal axis lists capital capacity $e = \frac{\mathcal{E}}{K}$. All panels plot three curves. The blue curve corresponds to depressed sentiment ($\mathcal{I} = -1.5SD$), the red curve corresponds to neutral sentiment ($\mathcal{I} = 0$), and the yellow curve corresponds to elevated sentiment ($\mathcal{I} = +1.5SD$).

The two leftmost panels of Figure 2 show asset prices $q(e, \mathcal{I})$ and $p(e, \mathcal{I})$. Financial frictions make asset prices sensitive to capital capacity e. In the crisis region (approximately $e < 0.4$), the binding constraint causes asset prices to plummet. The Sharpe ratio panels illustrate the nonlinear spike in risk premia that characterizes crisis times. Moving right, asset prices rise as intermediaries' risk-bearing capacity increases. Asset prices asymptote for high values of e as financial frictions dissipate.

Asset prices exhibit what HK refer to as "anticipation effects": asset prices start to fall well before the equity issuance constraint binds. Anticipation effects arise in equilibrium because forward-looking intermediaries are unwilling to support elevated asset prices in the face of mounting systemic risk.[17] Anticipation effects mean that

[14] Using a different specification than this paper, Bordalo et al. (2019b) estimate that the diagnostic expectations of stock market analysts incorporate the past three years of shocks.

[15] The beliefs model can be extended such that sentiment contains both a slow-moving component and a high-frequency component. However, this requires an additional state variable.

[16] See Appendix Figure 6 for more.

[17] Intermediaries perceive that risk premia will be low when asset prices are high, and vice-versa.

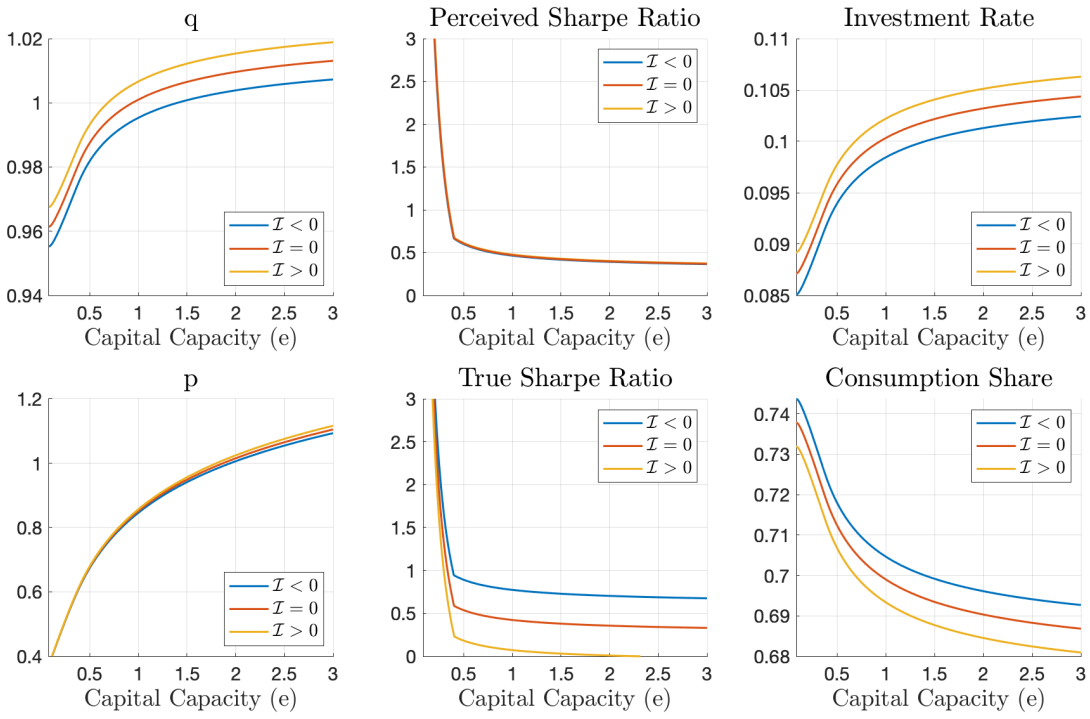


Figure 2: **Selected price and policy functions**. The horizontal axis lists capital capacity $e = \frac{\mathcal{E}}{K}$. Each panel plots three curves, corresponding to $\mathcal{I} = -1.5SD$ (blue), $\mathcal{I} = 0$ (red), and $\mathcal{I} = +1.5SD$ (yellow).

financial frictions affect macro-financial dynamics well before the constraint binds. This is captured by the "financial distress" region of the model, defined as e below its 33$^{\text{rd}}$ percentile, where constraints don't necessarily bind but financial friction effects are still present.[18]

Sentiment generates additional variation in asset prices. Asset prices are increasing in $\mathcal{I}$ because sentiment inflates expectations of cash-flow growth. Specifically, asset prices move so that sentiment gets "priced in" in equilibrium. This is illustrated by the Perceived Sharpe Ratio panel, which shows that variation in risk premia is driven almost entirely by e. The consequence of pricing assets based on non-rational

This goes against the evidence presented in Greenwood and Shleifer (2014). However, Greenwood and Shleifer (2014) study total returns rather than excess returns, and focus predominantly on household expectations. In the model, it is the beliefs of financial intermediaries that are relevant for pricing assets. Adam et al. (2020) find that professional investors have excess return expectations that covary negatively with the price-dividend ratio.

[18]Consistent with anticipation effects, Baron et al. (2020) document that bank equity declines predict output gaps, even when panics do not materialize.

expectations is shown in the True Sharpe Ratio panel. Elevated sentiment lowers realized risk premia, while depressed sentiment raises realized risk premia.

The investment and consumption panels show the propagation of financial and behavioral frictions to the real economy. The investment rate is high whenever either e or $\mathcal{I}$ is high. The consumption share $\frac{C^y}{Y}$ moves in the opposite manner. This follows from output market clearing in equation (21).

4.2 Stationary Distribution

Figure 3 plots the economy's stationary distribution over state variables e and $\mathcal{I}$. The ergodic distribution is solved numerically using a Kolmogorov forward equation. The economy is more likely to be in lighter-colored areas, while dark blue regions are rarely encountered. The dotted gray line marks the boundary of the crisis region.

The unconditional correlation between e_t and $\mathcal{I}_t$ is 0.78. This tight correlation arises in equilibrium because e_t and $\mathcal{I}_t$ both load positively on the same shocks. For example, a positive shock increases e_t by generating large returns for the financial sector, and also increases sentiment by making future states with high levels of capital more representative.

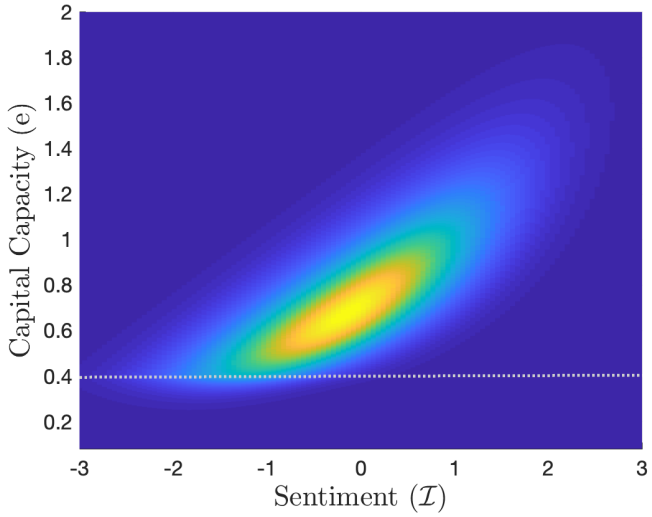


Figure 3: **Ergodic distribution**. Sentiment is reported in standard deviation units. The gray dotted line marks the boundary of the crisis region.

5 Results: Financial Market and Macroeconomic Dynamics

5.1 Sentiment-Driven Financial Crises

A critical shortcoming of rational models of financial crises is that they struggle to generate periods in which the probability of a crisis is high and yet risk premia are low. With diagnostic expectations, elevated sentiment can amplify financial fragility in the background of low risk premium environments.

To show this result, I use a Kolmogorov backward equation to calculate both the true and the diagnostically expected probability that the economy finds itself in a crisis at some point in the next three years. Each panel in Figure 4 below plots three contour lines, corresponding to a 10%, 30%, and 50% crisis probability. The left-hand panel plots the probability of a crisis that is perceived by agents with diagnostic expectations. The right-hand panel plots the true probability. The solid gray area marks the crisis region. For reference, the three dotted lines correspond to the 25^{th}, 50^{th}, and 75^{th} percentile level of e_t.

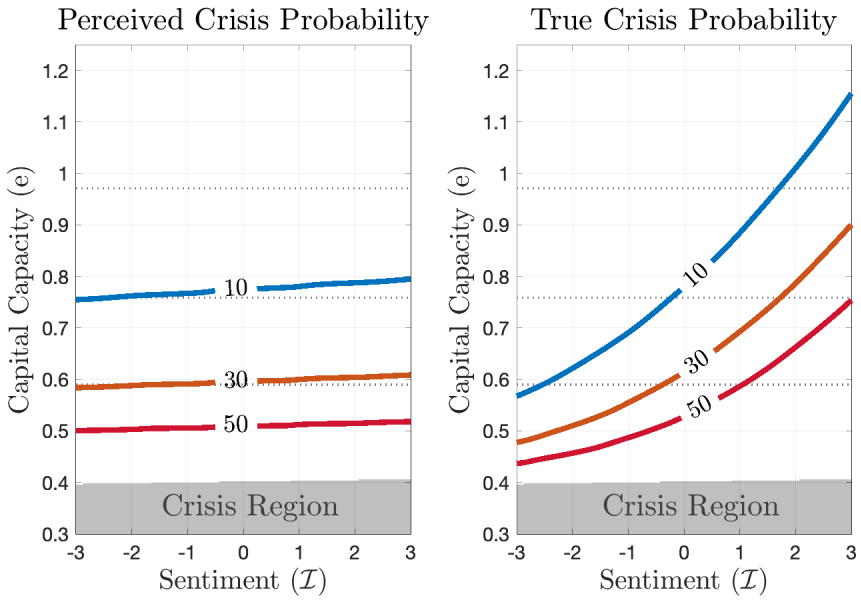


Figure 4: **Crisis hitting probabilities**. The three colored curves are contour lines corresponding to a 10%, 30%, and 50% probability that the economy enters a crisis in the next three years. The left-hand panel reports perceived crisis probabilities, and the right-hand panel reports true crisis probabilities. Sentiment is reported in standard deviation units.

Sentiment has almost no impact on perceived crisis probabilities.[19] Diagnostic expectations of fundamentals are incorporated into prices, so intermediaries perceive that crash risk is driven almost entirely by e_t. Perceived crash risk is high only near the crisis region, and fragility is quickly attenuated as the financial sector strengthens its capital capacity.

In the right-hand panel, the tilting of the contour lines highlights the buildup of undetected systemic risk that is triggered by overoptimism. When $\mathcal{I}_t > 0$, intermediaries borrow at elevated interest rates and pay inflated prices to purchase capital and housing. As expectations disappoint, the feedback from behavioral frictions to financial frictions causes intermediary balance sheets to deteriorate. Because this heightened fragility is an endogenous consequence of overoptimistic beliefs, it is neglected by intermediaries. Thus, diagnostic expectations allow the model to generate periods in which actual crash risk is high and yet risk premia are low.

The reverse story explains why perceived crisis probabilities are too high when $\mathcal{I}_t < 0$. Excessive pessimism allows the intermediary sector to borrow at low interest rates and purchase assets cheaply. When cash flows end up being larger than expected, the feedback effect works in the opposite direction and intermediaries quickly rebuild their capital capacity.

Figure 4 shows that when expectations are diagnostic, periods in which intermediary balance sheets appear to be strong can still be associated with heightened crisis risk. For example, when $\mathcal{I} = 0$ and the intermediary sector is at its median level of capital capacity, the three-year crisis probability is roughly 10%. This probability jumps to 30% when sentiment is elevated by 1.5 standard deviations. This result aligns with Greenwood et al. (2020b), who estimate that the three-year probability of a crisis reaches 40% when financial fragility is accompanied by elevated asset prices.

[19] Perceived crisis probabilities slope upward slightly. The reason is that a crisis is defined as $e_t < (1-\lambda)(p_t + q_t)$ (see equation (13)). Since the right-hand side of this inequality is increasing in sentiment, a given level of e_t is closer to the crisis region when $\mathcal{I}_t$ is high.

Behavioral Frictions Before Crises, Financial Frictions in Crises. Figure 5 explores the dynamic impact of behavioral and financial frictions around crises. Time 0 marks when the economy first enters a crisis. The blue curve plots the average realized (delevered) risk premium earned by intermediaries. The orange curve plots the average risk premium that intermediaries perceive they are demanding. Figure 5 reports the delevered risk premium to control for the variation in intermediary leverage that occurs when the equity issuance constraint binds. For reference, the horizontal line at 3.72% marks the unconditional mean realized risk premium.

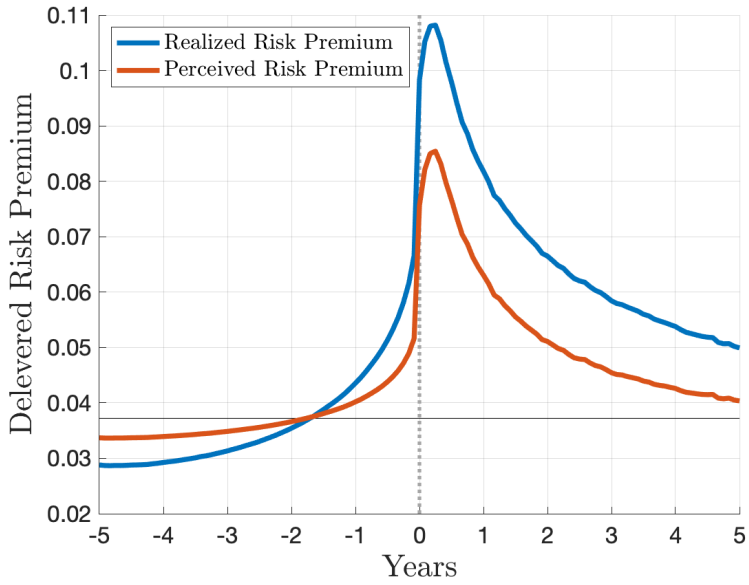


Figure 5: **Risk premia around crises**. Crises occur at time 0. The blue line plots the average realized risk premium earned by financial intermediaries around crises. The orange line plots the average risk premium perceived by intermediaries. The unconditional mean realized risk premium is 3.72%, as marked by the thin horizontal line. This analysis is inspired by Krishnamurthy and Muir (2020).

Elevated sentiment creates excessively low risk premia during the crisis buildup, consistent with the credit spread evidence of Krishnamurthy and Muir (2020). On average, realized risk premia are 15% below perceived risk premia at $t = -5$. Due to the feedback from behavioral frictions to financial frictions, this overoptimism amplifies financial fragility by triggering the depletion of intermediary capital capacity.

Figure 5 shows that sentiment-driven crises are generated by a persistently ele-

vated *level* of sentiment, not a sudden and dramatic *change* of sentiment. In fact, realized risk premia catch up to perceived risk premia approximately 1.5 years before crises hit, a pattern also shown in Krishnamurthy and Muir (2020). More broadly, this aligns with an emerging empirical finding that crises are the result of a slow-building erosion of financial sector resilience rather than a sudden sentiment shock (Baron et al., 2020).

The spike at $t = 0$ illustrates that both behavioral frictions and financial frictions are needed to replicate the full path of risk premia around crises. Diagnostic expectations produce low pre-crisis risk premia and neglected crash risk. However, slow-moving sentiment alone cannot generate the spike in risk premia caused by the binding constraint. Ex-ante behavioral frictions set the stage for crisis nonlinearities driven by financial frictions.

The model's post-crisis patterns reverse those of the pre-crisis period. The post-crisis period is characterized by excessive pessimism. Risk premia therefore sit above fundamentals in the aftermath of crises.[20]

5.2 Boom-Bust Investment Cycles

I now turn to studying sentiment-driven macroeconomic fluctuations. Diagnostic expectations affect investment rate i_t, which in turn controls the growth rate of output: $\frac{dY_t}{Y_t} = (i_t - \delta)dt + \sigma dZ_t$.

Before proceeding, I emphasize that the calibration of ζ is important here. ζ governs how diagnostic expectations are passed into asset prices versus the risk-free rate. The baseline calibration sets $\zeta < 1$. This means that q_t is increasing in $\mathcal{I}_t$. Accordingly, i_t is also increasing in sentiment.

I use impulse-response functions to study the response of investment to economic shocks. To capture periods of booms and malaise, I simulate the model at a monthly frequency and feed in a three-year sequence of either positive or negative shocks. These monthly shocks result in a one standard deviation cumulative shock over three

[20] See Muir (2017) for empirical evidence of high post-crisis risk premia.

years.[21] Shocks are turned off thereafter.

The investment IRFs are provided in Figure 6. The red curve plots the DEE and the dashed black curve plots the REE. Both economies start in their stochastic steady state at $t = -3$.[22]

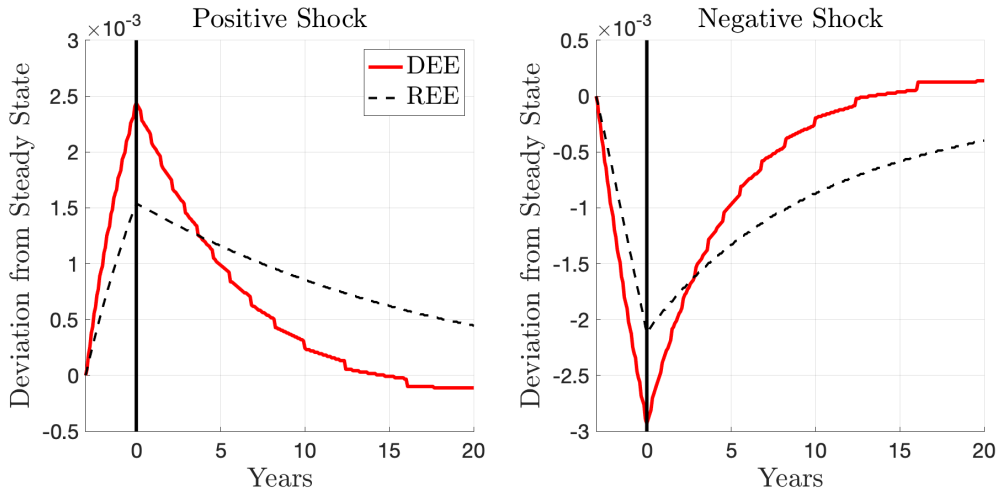


Figure 6: **Investment rate IRFs**. Starting from the stochastic steady state at $t = -3$, a sequence of positive/negative shocks is fed into the system until $t = 0$. The solid red line plots the response of i_t in the DEE. The dashed black line plots the response of i_t in the REE.

Comparing the DEE to the REE delineates the competing short- and long-run effects of sentiment. Diagnostic expectations promote boom-bust investment dynamics: short-run momentum is amplified, followed by steeper reversals. In the case of positive shocks, the sentiment-driven boom begets its own financial-frictions-driven bust. In the case of negative shocks, the sentiment-driven bust begets its own financial-frictions-driven boom.

First consider investment dynamics in the REE. For positive shocks, the boom from $t = -3$ to $t = 0$ increases e_t. This elevates capital price q_t and investment rate

[21] The monthly shock is set to $\pm\frac{\sigma\sqrt{3}}{12\times 3}$. Summing gives a 3-year cumulative shock of $\pm\sigma\sqrt{3}$.

[22] By turning shocks off from $t = 0$ onward, the IRFs in Figure 6 plot the investment response under one specific sequence of shocks. An alternative is to calculate the expected future path of investment: $\mathbb{E}_0[i_\tau]$. In this nonlinear model these two alternatives are not equivalent. Appendix Figure 9 plots the corresponding IRFs for $\mathbb{E}_0[i_\tau]$. The expected future investment rate is calculated using the Feynman-Kac formula.

i_t. Since e_t and i_t are both above their steady-state levels when the shocks stop at $t = 0$, they proceed to drift slowly back to the steady state. The opposite holds for negative shocks.

Turning to the DEE, consider the boom-bust pattern of the positive shock case. Positive shocks from $t = -3$ to $t = 0$ elevate $\mathcal{I}_t$ in addition to e_t. This causes a sharper investment boom in the short run. However, the feedback from behavioral frictions to financial frictions implies that excessive optimism decreases the subsequent returns earned by intermediaries. Over time, this erodes intermediary balance sheets and reduces long-run investment.

The negative shock case produces a bust-boom pattern in the DEE. Negative shocks depress sentiment, which amplifies the short-run drop in investment. But, the sentiment-driven bust also increases the future returns earned by intermediaries. As sentiment recovers, the economy is left with a stronger financial sector that is able to support higher levels of investment.[23]

Appendix Figure 7 studies the IRFs of an economy with diagnostic expectations but without financial frictions. Alone, slow-moving diagnostic expectations do not generate steep reversals in investment. The long-run reversals shown in Figure 6 rely on the feedback from behavioral frictions to financial frictions.

Recent empirical evidence supports the pattern of boom-bust investment cycles. Gulen et al. (2019) find that elevated credit-market sentiment in year t correlates with a boom in corporate investment over the subsequent year, followed by a long-run contraction. López-Salido et al. (2017) estimate that elevated credit-market sentiment in year t predicts lower GDP growth in year $t+2$. The long-run feedback from behavioral frictions to financial frictions is also consistent with the observation of Greenwood et al. (2019) that financial fragility arises at the end of economic expansions.

[23]Figure 6 shows a mildly asymmetric investment response to positive versus negative shocks. Larger initial shocks produce larger asymmetries. This is illustrated in Appendix Figure 8, which plots the IRFs that result after doubling the magnitude of the initial impulse.

5.3 Financial Market Stability from Beliefs

This paper identifies a stabilizing role for beliefs. Under the baseline calibration, financial crises are less likely to occur when expectations are diagnostic than when expectations are rational.

Figure 7 shows this result visually. The red and dashed black curves (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. The dotted vertical line at $e \approx 0.4$ marks the boundary of the crisis region. The blue curve crosses 1 to the right of the crisis region. This indicates that the probability of being in a crisis is lower in the DEE than the REE. The blue curve remains above 1 for large values of e, indicating that the DEE also features fewer periods of pronounced financial sector strength.

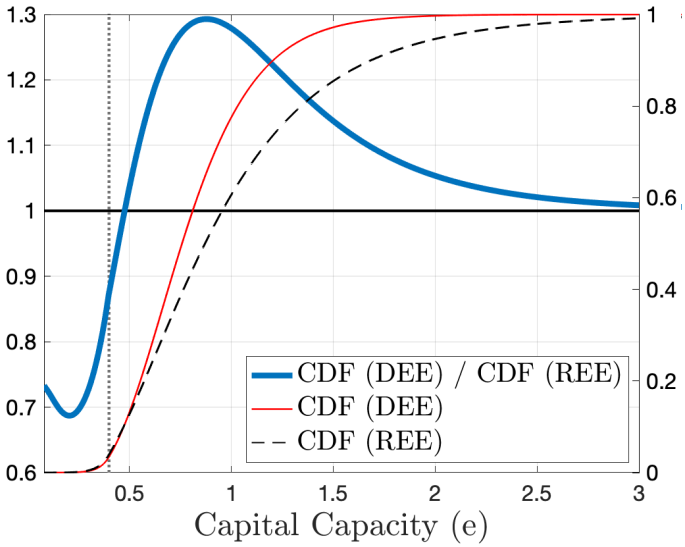


Figure 7: **Financial market stability from beliefs**. The solid red curve and the dashed black curve (right axis) plot the marginal CDF of capital capacity (e) in the DEE and the REE, respectively. The blue curve (left axis) divides the CDF of the DEE by the CDF of the REE. The dotted vertical line at $e \approx 0.4$ marks the boundary of the crisis region.

The stabilizing effect of beliefs can be understood by superimposing the crisis likelihoods in Figure 4 onto the ergodic distribution in Figure 3. Figure 4 illustrates that

the joint occurrence of financial distress and elevated sentiment is highly predictive of future financial crises. But, the ergodic distribution shows that this part of the state space is rarely encountered. Instead, financial distress typically coincides with excessive pessimism. When sentiment is pessimistic, the long-run reversal effect of diagnostic expectations protects intermediaries from financial crises.

Financial market stability may appear at odds with the earlier results of this paper. For example, Section 5.2 documents that diagnostic expectations amplify business cycles. The finding here is that diagnostic expectations stabilize financial markets. These results are intimately linked: the way that the economy avoids a financial crisis is by going through a sentiment-driven recession. In the case of negative shocks, depressed sentiment amplifies the investment bust in the short run. However, the long-run reversal effect means that depressed sentiment simultaneously reduces systemic risk by increasing the subsequent returns earned by intermediaries.

This finding of financial market stability may also appear to contradict the earlier result of sentiment-driven crises. Indeed, much of the empirical literature has found that elevated sentiment is predictive of financial market downturns, and concluded from this finding that extrapolative expectations promote financial instability. The model shows that such a conclusion does not necessarily follow from the evidence of sentiment-driven crises.

To reconcile sentiment-driven crises with financial market stability, note that the former is a conditional prediction while the latter is an unconditional prediction. The model's conditional prediction is that systemic risk is amplified when financial distress and elevated sentiment occur jointly. However, it is rare for the economy to reach these fragile states. The model's unconditional prediction is that beliefs can stabilize the financial sector, because financial distress is strongly correlated with depressed sentiment.

The goal of this section is not to assert that extrapolative expectations necessarily create financial stability. Rather, this section shows that the stabilizing effect of

extrapolation is a legitimate theoretical possibility, and one that the literature has neglected to date. Though diagnostic expectations prevent financial crises under the baseline calibration, this result can be overturned under alternate calibrations in which the magnitude of perceptual error is increased. Robustness is explored in Appendix B.6.

The identification of a stabilizing role for beliefs highlights a benefit of economic models. The model can be used to compare outcomes from different data generating processes (DEE versus REE). For the same reason, it is difficult to provide direct empirical evidence on financial market stability from beliefs. Nonetheless, expectations data is consistent with the channel of unanticipated reversals in financial market conditions. Bordalo et al. (2018a) analyze professional forecasts of the Baa-Treasury credit spread, finding that periods of financial distress reverse faster than forecasters expect. Similarly, Pflueger et al. (2019) document that market risk mean-reverts faster than analysts, options prices, and loan officers expect.

6 Evaluating Diagnosticity: The Feedback Effect

The feedback from behavioral frictions to financial frictions is a key mechanism underlying the model's equilibrium dynamics. This section tests three predictions that arise from this feedback. First, the feedback effect produces long-run reversals in economic conditions. This implies that the economy exhibits less persistence under diagnostic expectations than under rational expectations. Second, I use the model's prediction about how behavioral and financial frictions interact to identify a new fact about which crises are preceded by frothy financial markets. Third, I apply the model to the 2007-2008 Financial Crisis in order to assess the role of diagnostic expectations in shaping the evolution of the crisis.

6.1 Prediction 1: Long-Run Reversals and Economic Persistence

I start by comparing the persistence of macro-financial processes in the model and the data. Since the model's calibration does not target measures of persistence ex-

plicitly, an indicator of success is the extent to which the long-run reversals channel of diagnostic expectations helps to align the model with empirical moments.

In the experiments below, the calibration of the REE is identical to the calibration of the DEE, except for $\theta = 0$. I choose not to recalibrate the REE in order to pinpoint the effect of θ. Results are essentially unchanged if the REE is recalibrated (see Appendix B.7).

The Persistence of Financial Fragility. I begin by comparing the persistence of financial fragility in the DEE and the REE. The long-run reversal channel of diagnostic expectations means that intermediaries will recover from financial crises more quickly under diagnostic expectations than under rational expectations. Since crises are generated by a sequence of negative shocks, sentiment in the DEE is typically overpessimistic following crises. Due to the feedback from behavioral frictions to financial frictions, excessive pessimism stimulates intermediaries' recovery from crises.

This crisis-recovery effect is shown in Table 2. Starting from the time that the economy first enters a crisis, Table 2 lists the average number of years that it takes for capital capacity e_t to recover to its X^{th} percentile.

e_t Percentile	DEE	REE
5	0.39	0.42
10	0.85	1.08
25	2.22	3.72
50	4.85	10.42
75	9.74	26.41

Table 2: **Average crisis recovery time (in years)**. The DEE and the REE are simulated at a monthly frequency. For both equilibria, this table lists the average time (in years) that it takes for e_t to recover from a financial crisis to its X^{th} percentile. Median recovery times are reported in Appendix Table 6.

Recovery is faster under diagnostic expectations than under rational expectations. For example, in the DEE it takes an average of 4.9 years for intermediaries to recover from a crisis to their median level of capitalization, compared to 10.4 years in the REE. The relative difference in recovery times is larger for higher capital-capacity

thresholds because the feedback from behavioral frictions to financial frictions is a long-run effect that develops over time by increasing the drift of e_t.

Though there is no direct empirical counterpart to the analysis in Table 2, the data is suggestive of recovery times that are more consistent with the DEE. In financial markets, Krishnamurthy and Muir (2020) collect 150 years of credit spread data across 19 countries, and find that credit spreads recover to their mean value between 4 and 5 years after a financial crisis. Muir (2017) shows that the majority of stock-market losses in financial crises are recovered within 5 years. For the broader macroeconomy, Jordà et al. (2013) find in an international panel of 50 financial crises that real GDP per capita recovers between 4 and 5 years after a financial crisis. Reinhart and Rogoff (2014) study the recovery from 63 advanced economy financial crises, and calculate an average trough to recovery time of 4.4 years.

Macro-Financial Autocorrelations. The model focuses on financial crises, and abstracts from many macroeconomic considerations at the business cycle frequency. With this caveat in mind, the long-run reversal property can also be examined for broad macro-financial aggregates. I estimate the autocorrelation of the dividend-price ratio and the investment-output ratio using the Jordà-Schularick-Taylor Macrohistory Database over 17 developed countries from 1950 – 2016 (Jordà et al., 2019).[24] Each ratio is standardized at the country level and then pooled.

Figure 8 compares the estimated autocorrelation of the dividend-price ratio and the investment-output ratio (blue) to the corresponding autocorrelation in the DEE (red) and the REE (dashed black). The DEE is broadly consistent with the data, particularly for longer horizons. This suggests that the long-run reversals produced by the feedback from behavioral frictions to financial frictions improve the model's empirical fit. For robustness, Appendix Figure 10 plots these autocorrelations using

[24]These ratios are used because they are cointegrated. This helps to circumvent issues such as unit roots in investment and dividends. The concept of dividend-price cointegration is standard in financial economics. For the investment-output ratio, see Cochrane (1994). Post-WWII data is used to account for structural changes to the finance system (Schularick and Taylor, 2012).

just U.S. data. Bordalo et al. (2020) provide additional evidence of long-run reversals in the dividend-price ratio, documenting that variation in the dividend-price ratio is driven by long-term growth expectations that overreact to news.

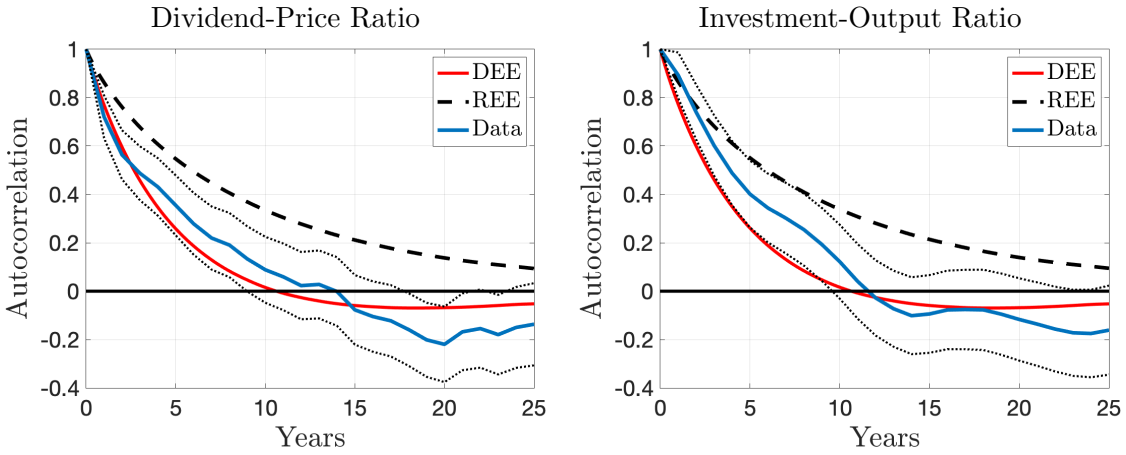


Figure 8: **Persistence: data and model**. The left-hand panel plots the empirical autocorrelation of the dividend-price ratio (blue) as well as the model-implied autocorrelation in the DEE (red) and the REE (dashed black). The right-hand panel conducts an equivalent analysis for the investment-output ratio. The 95% confidence interval is calculated using Bartlett's formula.

6.2 Prediction 2: Which Crises are Preceded by Overheating?

Some crises erupt following periods of relatively robust financial activity (e.g., the 2007-2008 Financial Crisis), while other crises emerge after prolonged periods of financial distress (e.g., the 2011 Eurozone Crisis). The model with diagnostic expectations predicts that these two types of crises should have different levels of pre-crisis froth in financial markets. In the model, the feedback from behavioral frictions to financial frictions implies that the initial trigger of financial distress is often elevated sentiment and financial market overheating. However, once financial distress has been triggered, the economy can persist in this vulnerable state — sometimes leading to residual crises — even after sentiment has subsided.

To formalize this "vulnerabilities-plus-triggers" framework (López-Salido et al., 2017), I simulate the model and compare the path of sentiment prior to First Crises versus Residual Crises. I define a Residual Crisis, or double-dip, as a crisis that

was preceded by an earlier crisis within the past five years. All other crises are First Crises.[25] In the model, 57% of First Crises are preceded by elevated sentiment ($\mathcal{I} > 0$) in at least one of the three years preceding the crisis. Alternatively, only 11% of Residual Crises are preceded by elevated sentiment.

I evaluate this prediction empirically using the Greenwood et al. (2020a, henceforth GHSS) dataset.[26] GHSS compile annual data on credit and asset price growth across 42 countries from 1950 – 2016, and define "R-zones" as years in which credit growth is above its 80^{th} percentile and asset price growth is above its 67^{th} percentile. Since credit growth is a well-known signal of financial fragility (e.g., Schularick and Taylor, 2012), and R-zones feature rapid credit growth accompanied by elevated asset prices, GHSS interpret R-zones as an indicator of overheating in financial markets. Consistent with the results in Figure 4, GHSS find that R-zones are highly predictive of financial crises over a three-year horizon.

The GHSS dataset contains 34 First Crises and 9 Residual Crises. 79% of First Crises (27/34) are preceded by an R-zone in the prior three years, compared to only 11% of Residual Crises (1/9). Thus, the data is consistent with the model's prediction that First Crises are often preceded by frothy financial markets, while Residual Crises are not. This finding also highlights the importance of developing models of financial crises that include both behavioral and financial frictions. Behavioral frictions help the model generate periods of pre-crisis froth, while slow-moving financial frictions allow for prolonged periods of financial fragility and occasional residual crises.

[25] I use the following crisis definition in order to align my model with the Greenwood et al. (2020a) dataset that I will analyze (details in the next paragraph). A crisis is defined as the equity issuance constraint binding for two consecutive months, and a crisis persists until e_t recovers to its 20^{th} percentile. When I annualize the simulated data, this definition produces the same unconditional crisis probability as in the Greenwood et al. (2020a) data (3.1%). 83% of crises are First Crises in the model, compared to 79% in the data.

[26] I thank the authors for sharing their data to make this analysis possible.

6.3 Prediction 3: Elevated Sentiment and the 2007-2008 Financial Crisis

This section applies the model to the 2007-2008 Financial Crisis in order to evaluate the channels through which sentiment influenced the evolution of the crisis.

Measuring Sentiment in the Data. An open question in the diagnostic expectations literature is how to measure sentiment in the data. Proposition 2 provides a simple method. Sentiment can be constructed using forecast errors of economic growth.

Proposition 2. *Let* $\sigma \widehat{dZ}_t = \frac{dY_t}{Y_t} - \widehat{\mathbb{E}}_t \frac{dY_t}{Y_t} = -\theta \mathcal{I}_t dt + \sigma dZ_t$ *denote the economic growth forecast error at time* t*. Sentiment* $\mathcal{I}_t$ *can be rewritten in terms of forecast errors as follows:*

$$\mathcal{I}_t = \int_0^t e^{(-\kappa+\theta)(t-s)} \sigma \widehat{dZ}_s. \tag{24}$$

Proof. See Appendix B.5. □

A feature of equation (24) is that it does not require parametric assumptions about the data generating process for the economy. Given a calibration of κ and θ, all that equation (24) requires is forecast errors of output growth.

Figure 9 uses the Survey of Professional Forecasters (SPF) to calculate sentiment from 1970 through 2018.[27] To initialize the calculation, I assume that sentiment equals zero in January 1970. Since SPF forecasts are collected at a quarterly frequency, Figure 9 is calculated using a discrete-time analogue of equation (24). Full details are provided in Appendix A.2.

Sentiment in Figure 9 captures what Kindleberger (1978) refers to as "displacement": financial crises are preceded by large positive shocks to economic fundamentals.[28] Sentiment builds rapidly during the 1990s economic boom. Though sentiment

[27] Expectations data from macroeconomic professionals is most consistent with the model, since sophisticated intermediaries are responsible for pricing risky assets.

[28] Cao and L'Huillier (2018) observe that the three deepest recessions in developed countries all occurred roughly 10 years after periods of rapid technological innovation.

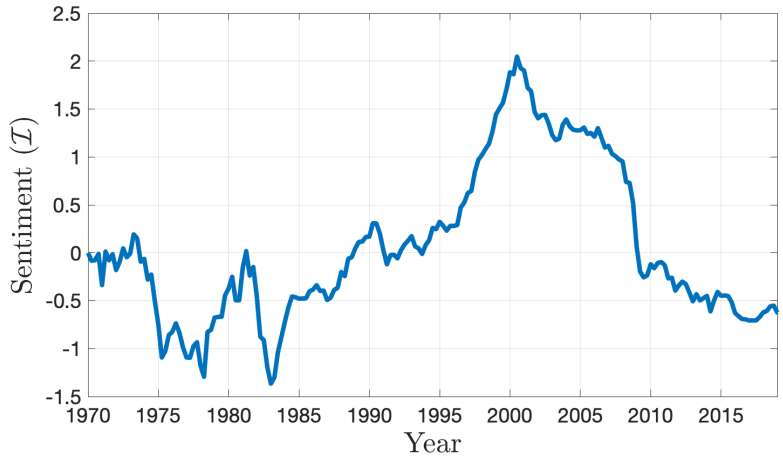


Figure 9: **SPF-measured sentiment**. Sentiment $\mathcal{I}_t$ is measured using the median SPF forecast error from 1970 through 2018 under the model's baseline calibration. Sentiment is reported in standard deviation units.

dips with the bursting of the dot-com bubble in the early 2000s, sentiment remains elevated until the financial crisis.

Simulating the Financial Crisis. I now use the model to evaluate the effect elevated sentiment on the 2007-2008 Financial Crisis. To simulate the model, I need to choose a model initialization and a sequence of shocks. Typically, the modeler chooses the path of shocks that best aligns their model with the data. This paper takes a more restrictive approach by calculating shocks externally from SPF forecast errors. That is, rather than asking whether any sequence of shocks can align the model with the data, I ask whether a specific sequence of shocks — one that is consistent with expectations — can align the model with the data.

Given the measure of sentiment in Figure 9, capital quality shocks can be backed out of forecast errors: $\sigma dZ_t = \sigma \widehat{dZ}_t + \theta \mathcal{I}_t dt$. I use this procedure to calculate the shocks implied by SPF data, and simulate the model conditional on these shocks. Under both rational and diagnostic expectations, forecast errors place empirical bounds on what can reasonably be considered a shock. This paper takes seriously the restrictions that expectations data provide (Manski, 2004).

I consider two initializations. First, I start the model in 1992Q1, right as the U.S.

is emerging from a minor financial crisis.[29] This initialization takes a long-run view of the 2007-2008 Financial Crisis in order to understand the full effect of the 1990s IT boom on the subsequent crisis. I initialize e_t at the boundary of the crisis region, and $\mathcal{I}_t$ at the 1992Q1 value shown in Figure 9.

The long-run simulation from 1992 through 2018 is shown in Appendix Figure 5. I compare the simulation to the data using the He et al. (2017) Intermediary Capital Ratio, which serves as an empirical counterpart to capital capacity e_t. In this single-shock model with shocks taken externally from SPF forecast errors, the correlation between e_t and the Intermediary Capital Ratio is 0.78 in the DEE. Using the same set of shocks, the correlation is only 0.41 in the REE. The difference in fit is driven by a divergence between the DEE and the REE over the mid-2000s. In the DEE, elevated sentiment from the 1990s boom leads to an erosion of capital capacity prior to 2007. In the REE, the financial sector remains well-capitalized throughout the mid-2000s (further details below).

Though diagnostic expectations allow the model to qualitatively replicate the empirical profile of intermediary capitalization over this boom-bust period, this long-run simulation points to a failure of the model. Using the shocks implied by SPF forecast errors, the model cannot jointly fit the 1990s boom and the 2007-2008 Financial Crisis. Though the DEE comes much closer than the REE, neither produces a financial crisis in 2008: the sequence of positive shocks realized throughout the 1990s places the financial sector too far above the crisis region for the negative shocks in 2007 and 2008 to generate a full-blown crisis.

This formal failure is not surprising. This test tasks the model with replicating almost 30 years of macro-financial dynamics using a single shock process that is calculated externally from SPF forecast errors. The dot-com crash triggered significant turmoil in financial markets, but had relatively muted effects on the real economy. Because the model features only one shock process, it cannot capture such differenti-

[29] See, for example, the crisis dating of Romer and Romer (2017) and Baron et al. (2020).

ation between financial markets and the real economy.[30]

In light of this failure, the second initialization starts at the trough of the tech bubble's collapse in 2002Q4. e_t is initialized to $e_{distress}$ (the 33$^{\text{rd}}$ percentile of e_t) in order to capture the tightness of financial conditions at that time. The initialization of e_t can be thought of as applying a single MIT shock to the financial sector in order to capture the differential effect of the dot-com crash on financial markets. $\mathcal{I}_t$ is initialized to the 2002Q4 value shown in Figure 9. Proceeding from this starting point, again using the shocks calculated from the SPF, the results of the simulation are plotted in Figure 10.

The top panel of Figure 10 shows the time path of capital capacity e_t in the DEE (red, left axis) and the REE (dashed black, left axis). The gray area marks the crisis region. For comparison, the blue curve (right axis) plots the Intermediary Capital Ratio of He et al. (2017).

The key result from the simulation is that only the DEE produces a financial crisis. This is not to say that it is impossible to generate a financial crisis in the REE, but, rather, that doing so requires large negative shocks that are inconsistent with empirical forecast errors.[31] This highlights the amplification of financial fragility that elevated sentiment can produce. In particular, the divergence of the e_t profiles from 2002 through 2008 shows how a sentiment-driven crisis evolves. Realized shocks are mild over this period, so capital capacity in the REE recovers quickly from initial distress. In the DEE, elevated sentiment implies that balance-sheet vulnerability persists throughout the mid-2000s. This leaves intermediaries exposed to the negative shocks that hit in 2008.

The middle panel evaluates the model's ability to replicate patterns in risk premia.

[30] Though sacrificing parsimony, this suggests that the model's fit can be improved with additional shocks that differentially affect the financial sector and the real economy.

[31] He and Krishnamurthy (2019) reach a similar conclusion when applying their rational model to the 2007-2008 Financial Crisis, and show that "hidden leverage" is important for understanding the crisis. The goal of my model with diagnostic expectations is to complement and build on this conclusion by microfounding non-rational beliefs and systematizing the channels through which sentiment can endogenously trigger neglected risk in financial markets.

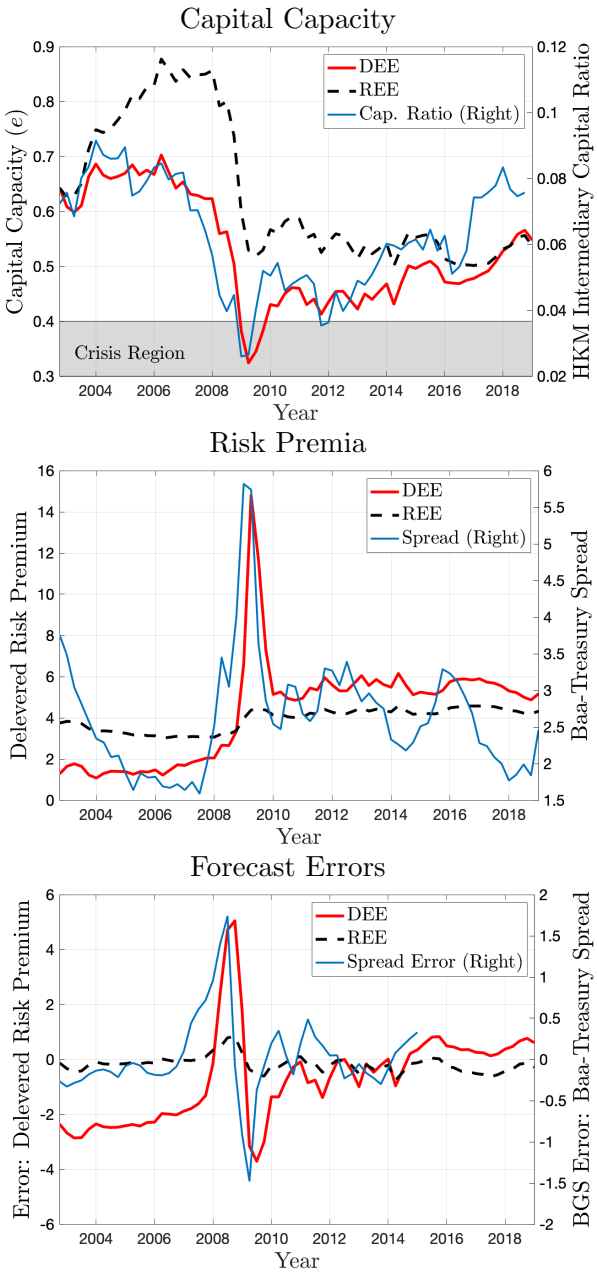


Figure 10: **Simulating the 2007-2008 Financial Crisis**. The sequence of shocks implied by SPF forecast errors is fed into the model from 2002Q4 through 2018Q4. The top panel plots the path of capital capacity e_t in the DEE and the REE (left axis), as well as the corresponding empirical measure from He et al. (2017) (right axis). The middle panel plots the intermediary risk premium in the DEE and the REE (left axis), and the Baa – 10-year Treasury spread (right axis). The bottom panel plots the forecast error of the risk premium averaged over quarters $t+1$ through $t+4$ in the DEE and the REE (left axis), and compares it to the credit spread forecast error reported in Bordalo et al. (2018a) (right axis).

This panel plots the realized risk premium earned by intermediaries in the DEE (red, left axis) and the REE (dashed black, left axis), as well as the Baa – 10-year Treasury spread for comparison (blue, right axis). The DEE broadly reproduces the path of risk premia around the crisis. Pre-crisis risk premia are low. This is consistent with the narrow spreads observed prior to the crisis, which many argue were due to neglected risk (e.g., Gennaioli and Shleifer, 2018). Once the crisis hits, risk premia spike. Because the crisis de-biases expectations, post-crisis risk premia remain persistently higher than pre-crisis risk premia. Alternatively, because the REE never enters a crisis, it exhibits almost no variation in risk premia over the simulated period.

The bottom panel assesses how these reversals in risk premia are reflected in beliefs. The red and dashed black curves plot forecast errors in the DEE and the REE, respectively (left axis). Specifically, this figure plots intermediary forecast errors of the delevered risk premium averaged over quarters $t+1$ through $t+4$. The blue curve (right axis) plots the corresponding forecast error of the Baa – Treasury spread, as reported in Bordalo et al. (2018a).[32] As in the data, the DEE generates large positive forecast errors heading into the crisis, and large negative forecast errors exiting the crisis. The REE produces only modest forecast errors, because risk premia exhibit little variation away from the crisis region in the REE.

7 Conclusion

This paper develops a general equilibrium macroeconomic model that combines frictions in financial intermediation with diagnostic expectations. The model examines how the interplay of behavioral and financial frictions drives macro-financial dynamics. When the financial sector is distressed, elevated sentiment amplifies systemic risk and sets the stage for financial crises. For the broader macroeconomy, the conflicting short-run and long-run effect of diagnostic expectations generates endogenous boom-bust patterns in investment and output growth. Even with this business cycle

[32] This data was kindly provided by the authors.

amplification, the long-run reversal effect of diagnostic expectations can inhibit financial crises. Empirical tests support the feedback from behavioral frictions to financial frictions as a channel that improves the model's fit of macro-financial dynamics.

This paper takes a first step toward integrating diagnostic expectations into models of financial frictions, and there are many avenues which may be fruitful for future research. First, the model lends itself to the study of policy interventions in financial markets with non-rational intermediaries. Second, the discussion in Section 6.3 suggests that extending this paper's framework, either with additional shocks or additional model features, may help to bring the model closer to the data. Third, it will also be interesting to examine how a richer belief process, for example one containing both a slow-moving and a high-frequency component, affects equilibrium dynamics.

References

Adam, Klaus and Sebastian Merkel, "Stock Price Cycles and Business Cycles," *SSRN 3455237*, 2019.

_ , Dmitry Matveev, and Stefan Nagel, "Do Survey Expectations of Stock Returns Reflect Risk Adjustments?," *Journal of Monetary Economics*, 2020.

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

Bansal, Ravi and Amir Yaron, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 2004, *59* (4), 1481–1509.

Barberis, Nicholas, "Psychology-Based Models of Asset Prices and Trading Volume," *Handbook of Behavioral Economics-Foundations and Applications 1*, 2018, pp. 79–175.

_ , 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.

_ , Emil Verner, and Wei Xiong, "Banking Crises Without Panics," *SSRN 3116148*, 2020.

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, 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.

_ , _ , Rafael La Porta, and Andrei Shleifer, "Diagnostic Expectations and Stock Returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

_ , _ , _ , and _ , "Expectations of Fundamentals and Stock Market Puzzles," *Mimeo*, 2020.

_ , _ , Yueran Ma, and Andrei Shleifer, "Over-reaction in Macroeconomic Expectations," *NBER w24932*, 2018.

Brunnermeier, Markus K. and Yuliy Sannikov, "A Macroeconomic Model with a Financial Sector," *American Economic Review*, February 2014, *104* (2), 379–421.

Caballero, Ricardo J. and Alp Simsek, "A Risk-Centric Model of Demand Recessions and Speculation," *The Quarterly Journal of Economics*, 2020, *Forthcoming.*

Cao, Dan and Jean-Paul L'Huillier, "Technological Revolutions and the Three Great Slumps: A Medium-Run Analysis," *Journal of Monetary Economics*, 2018, *96*, 93–108.

Cochrane, John H., "Permanent and Transitory Components of GNP and Stock Prices," *The Quarterly Journal of Economics*, 1994, *109* (1), 241–265.

Davis, Morris A. and Stijn Van Nieuwerburgh, "Housing, Finance, and the Macroeconomy," *Handbook of Regional and Urban Economics*, 2015, *5*, 753–811.

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.

Farboodi, Maryam and Péter Kondor, "Rational Sentiments and Economic Cycles," *NBER w27472*, 2020.

Farhi, Emmanuel and Iván Werning, "Taming a Minsky Cycle," *Mimeo*, 2020.

Fuster, Andreas, Benjamin Hebert, and David Laibson, "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing," *NBER Macroeconomics Annual*, 2012, *26* (1), 1–48.

Gabaix, Xavier, "A Behavioral New Keynesian Model," *NBER w22954*, 2016.

Gennaioli, Nicola and Andrei Shleifer, *A Crisis of Beliefs: Investor Psychology and Financial Fragility*, Princeton University Press, 2018.

Gertler, Mark, Nobuhiro Kiyotaki, and Andrea Prestipino, "A Macroeconomic Model with Financial Panics," *The Review of Economic Studies*, 2020, *87* (1), 240–288.

Gomes, João F., Marco Grotteria, and Jessica Wachter, "Foreseen Risks," *NBER w25277*, 2018.

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 Lawrence J. Jin, "Reflexivity in Credit Markets," *NBER w25747*, 2019.

_ , Samuel G Hanson, Andrei Shleifer, and Jakob Ahm Sørensen, "Predictable Financial Crises," June 2020, (27396).

_ , _ , _ , and _ , "Predictable Financial Crises," *Mimeo*, 2020.

Gulen, Huseyin, Ion Mihai, and Stefano Rossi, "Credit Cycles and Corporate Investment," *SSRN 3369295*, 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*, October 2019, *11* (4), 1–37.

_ , Bryan Kelly, and Asaf Manela, "Intermediary Asset Pricing: New Evidence from Many Asset Classes," *Journal of Financial Economics*, 2017, *126* (1), 1–35.

Hirshleifer, David, Jun Li, and Jianfeng Yu, "Asset Pricing in Production Economies with Extrapolative Expectations," *Journal of Monetary Economics*, 2015, *76*, 87–106.

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M. Taylor, "The Rate of Return on Everything, 1870–2015," *The Quarterly Journal of Economics*, 2019, *134* (3), 1225–1298.

_ , Moritz Schularick, and Alan M. Taylor, "When Credit Bites Back," *Journal of Money, Credit and Banking*, 2013, *45* (2), 3–28.

Kindleberger, Charles P., *Manias, Panics and Crashes: a History of Financial Crises*, Basic Books, 1978.

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," *NBER w23850*, 2020.

_ and Wenhao Li, "Dissecting Mechanisms of Financial Crises: Intermediation and Sentiment," *NBER w27088*, 2020.

Li, Wenhao, "Public Liquidity and Financial Crises," *SSRN 3175101*, 2020.

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.

Manski, Charles F., "Measuring Expectations," *Econometrica*, 2004, *72* (5), 1329–1376.

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.

Muir, Tyler, "Financial Crises and Risk Premia," *The Quarterly Journal of Economics*, 2017, *132* (2), 765–809.

Nagel, Stefan and Zhengyang Xu, "Asset Pricing with Fading Memory," *NBER w26255*, 2019.

Pflueger, Carolin, Emil Siriwardane, and Adi Sunderam, "Financial Market Risk Perceptions and the Macroeconomy," *NBER w24529*, 2019.

Reinhart, Carmen M. and Kenneth S. Rogoff, "Recovery from Financial Crises: Evidence from 100 Episodes," *American Economic Review*, 2014, *104* (5), 50–55.

Romer, Christina D. and David H. Romer, "New Evidence on the Aftermath of Financial Crises in Advanced Countries," *American Economic Review*, 2017, *107* (10), 3072–3118.

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.

Tella, Sebastian Di, "Uncertainty Shocks and Balance Sheet Recessions," *Journal of Political Economy*, 2017, *125* (6), 2038–2081.

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," *SSRN 3333248*, 2020.

Woodford, Michael, "Macroeconomic Analysis without the Rational Expectations Hypothesis," *Annual Review of Economics*, 2013, *5* (1), 303–346.

Appendix

Appendix Contents. Appendix A presents the continuous-time specification of diagnostic expectations used in this paper. Appendix B provides model details, proofs, and additional results. Appendix C describes the equilibrium derivation, and Appendix D outlines the numerical methods used to solve the model. Appendix E gives additional details and extensions to this paper's model of diagnostic expectations, including a discrete-time formulation.

A Diagnostic Expectations Appendix

A.1 Diagnostic Expectations in Continuous Time

This section provides a microfoundation for the reduced-form expectations process outlined in Section 2.2. A goal for this paper's model of diagnostic expectations is to be a portable extension of existing models ["PEEMish"](Rabin, 2013). The expectations model is designed such that rational models can be augmented with diagnostic expectations using a single additional state variable.

Diagnostic expectations are applied to the log of the capital stock. Log capital evolves according to $dk_t = (i_t - \delta - \frac{\sigma^2}{2})dt + \sigma dZ_t$. Diagnostic expectations are applied to log 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 $\mathcal{I}_t$ is stationary because the diffusion coefficient for log capital is constant.

Step 1: Defining the Background Context. Following the terminology of Bordalo et al. (2018a, henceforth BGS), the first step is to define the "background context" for capital. The background context is a counterfactual level of the log capital stock. It forms the dynamic "reference class" used to characterize representativeness.

The background context reflects the absence of recent information. Equation (3) introduces $\mathcal{I}_t \equiv \int_0^t e^{-\kappa(t-s)}\sigma dZ_s$ as a measure of recent information. This implies the following definition of the background context.

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

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

Step 2: Modeling Expectations Given the Background Context. The next step is to specify how agents form expectations. Because time is continuous, expectations must be specified for future periods $t+\tau$, for all $\tau > 0$. 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 state variables k_t, e_t, and $\mathcal{I}_t$. 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 following likelihood ratio:

$$\frac{h(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t)}{h(k'_{t+\tau}|G_t^-, e_t, \mathcal{I}_t)}. \tag{25}$$

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

One difficulty with equation (25) 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.[33] This difficulty is overcome by using an instantaneous prediction horizon of $\tau = dt$. Because k_t is an Itô Process it is instantaneously Gaussian. Taking $\tau \to dt$, $h(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t) = \mathcal{N}\left(k_t + \left[i(e_t, \mathcal{I}_t) - \delta - \frac{\sigma^2}{2}\right] dt, \sigma^2 dt\right)$ and $h(k'_{t+\tau}|G_t^-, e_t, \mathcal{I}_t) = \mathcal{N}\left(G_t^- + \left[i(e_t, \mathcal{I}_t) - \delta - \frac{\sigma^2}{2}\right] dt, \sigma^2 dt\right)$. I now define diagnostic expectations over prediction horizon $\tau = dt$. The prediction horizon will be extended iteratively in Step 3.

Diagnostic expectations overweight states that are representative of recent news.

[33] This is in contrast to BGS, where expectations are only specified for exogenous AR(N) processes.

This is formalized by assuming that agents evaluate future levels of log capital according to the distorted density:

$$h_t^{\theta}(k'_{t+dt}|k_t, e_t, \mathcal{I}_t) = h(k'_{t+dt}|k_t, e_t, \mathcal{I}_t) \cdot \left[\frac{h(k'_{t+dt}|k_t, e_t, \mathcal{I}_t)}{h(k'_{t+dt}|G_t^-, e_t, \mathcal{I}_t)}\right]^{\theta dt} \frac{1}{Z}. \tag{26}$$

Equation (26) modifies a similar formula in BGS. The key adjustment for continuous time is that equation (26) defines expectations at $t + dt$, while the discrete-time formulation of BGS defines expectations at $t+1$. In equation (26), 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 infinitesimal distortion on the perceived distribution of capital over an infinitesimally short horizon. Otherwise, the agent would expect that k_t jumps discontinuously from t to $t + dt$.

Using equation (26), the following proposition characterizes the perceived evolution of capital.

Proposition 3. *A diagnostic agent perceives that capital evolves according to*

$$\frac{\widehat{dK_t}}{K_t} = (i_t - \delta)dt + \theta \mathcal{I}_t dt + \sigma dZ_t. \tag{27}$$

Proof. See Appendix B.5. □

Judging by representativeness biases the perceived growth rate of capital by $\theta\mathcal{I}_t$.

Step 3: The Evolution of Beliefs. Step 1 defines the background context G_t^- and Step 2 specifies diagnostic expectations of $\widehat{dK_t}$. This step models the dynamics of expectations over longer horizons. Because capital is endogenous, only the instantaneous distribution of k_t is known. Future expectations are therefore defined iteratively. In particular, repeated applications of the instantaneous Gaussian properties of k_t can be used to 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.

Diagnostic agents form expectations by simulating the economy forward state-by-state. As the diagnostic agent simulates the economy forward from time t, the internal representativeness parameter at simulated future time $t + \tau$ is given by:

$$\begin{aligned} \mathcal{I}^S_{t+\tau} &\equiv \int_0^t e^{-\kappa(t+\tau-s)} \sigma dZ_s, \text{ or equivalently} \\ &= e^{-\kappa\tau} \mathcal{I}_t. \end{aligned} \tag{28}$$

The superscript S, for simulated, is used to signify that $\mathcal{I}^S_{t+\tau}$ is the agent's unconscious internal representativeness state as the agent forms expectations of the economy in period $t + \tau$. Information that was representative at time t decays at rate κ as the perceived economy is simulated forward in time.

Let $k'_{t+\tau}$ and $e'_{t+\tau}$ denote one possible realization of log capital and capital capacity at time $t + \tau$. Using equation (28), the simulated background context at $t + \tau$ 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 $t+\tau$. Given $k'_{t+\tau}$, the simulated background context at time $t + \tau$ is defined as follows:*

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

As above, the simulated future background context reflects the absence of recent 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:

$$h^\theta_t(k'_{t+\tau+dt}|k'_{t+\tau}, e'_{t+\tau}, \mathcal{I}^S_{t+\tau}) = h(k'_{t+\tau+dt}|k'_{t+\tau}, e'_{t+\tau}, \mathcal{I}^S_{t+\tau}) \cdot \left[\frac{h(k'_{t+\tau+dt}|k'_{t+\tau}, e'_{t+\tau}, \mathcal{I}^S_{t+\tau})}{h(k'_{t+\tau+dt}|G'^{-}_{t+\tau}, e'_{t+\tau}, \mathcal{I}^S_{t+\tau})}\right]^{\theta dt} \frac{1}{Z}. \tag{29}$$

It follows that the diagnostic agent perceives that capital evolves according to:

$$\widehat{\frac{dK'_{t+\tau}}{K'_{t+\tau}}} = (i_{t+\tau} - \delta)dt + \theta \mathcal{I}^S_{t+\tau} dt + \sigma dZ_{t+\tau}. \tag{30}$$

Future expectations in equation (30) should be contrasted with those of a rational agent who correctly believes that capital evolves according to $\frac{dK'_{t+\tau}}{K'_{t+\tau}} = (i_{t+\tau} - \delta)dt + \sigma dZ_{t+\tau}$. Since $\mathcal{I}^S_{t+\tau} = e^{-\kappa\tau}\mathcal{I}_t$, equation (30) specifies that the effect of diagnostic expectations on the perceived growth rate of capital fades at rate κ 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, information that is diagnostic of economic conditions at time t steadily dims.

It is worth noting that equation (30) only stipulates that the diagnostic agent's perception of $k_{t+\tau}$'s drift converges to rationality as $\tau \to \infty$. Since the drift has a cumulative effect on the level of $k_{t+\tau}$, diagnostic expectations of the level of $k_{t+\tau}$ can diverge increasingly from rational expectations 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 given in Appendix E.2. Appendix E.1 discusses how the discrete-time analogue of this paper's expectations model relates to the original BGS model.

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 $\mathcal{I}^S_t$ evolves as expectations are simulated forward. 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 $\mathcal{I}^S_{t+dt}$. Then, Step 2 is applied again (now taking $\widehat{k}_{t+dt}$, $\widehat{e}_{t+dt}$, and $\mathcal{I}^S_{t+dt}$ 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 $\mathcal{I}^S_{t+2dt}$. This process is repeated to generate expectations at time $t+\tau$, for all $\tau > 0$.

Diagnostic 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. A comparison of equations (3) and (28) shows that diagnostic agents do not perceive that future capital quality shocks will alter the bias of their future expectations.[34]

I end by discussing why this model of diagnostic expectations can serve as a portable extension of existing rational models. First, equations (27) and (30) illustrate that state variable $\mathcal{I}_t$ alone is sufficient to characterize the state of expectations relative to rationality. Second, the evolution of $\mathcal{I}_t$ is self-contained. $\mathcal{I}_t$ can be expressed in differential form as $d\mathcal{I}_t = -\kappa \mathcal{I}_t dt + \sigma dZ_t$. Thus, state variable $\mathcal{I}_t$ plus the shock σdZ_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.

A.2 Diagnostic Expectations Calibration and Application

θ Calibration. The baseline calibration sets θ such that one standard deviation in $\mathcal{I}$ corresponds to an output growth bias of 0.75 percentage points. The magnitude of this bias aligns with the estimates in Bordalo et al. (2018a, henceforth BGS), Bordalo et al. (2018b, BGMS), Bordalo et al. (2019b, BGLS), Bordalo et al. (2019a, BGST), and d'Arienzo (2020).

Using data from 1968Q4 through 2016Q4, BGMS assume that the realized annual growth rate of real GDP follows an AR(1) process: $x_t = \rho x_{t-1} + u_t$. BGMS estimate $\rho = 0.87$ and $\sigma_u = 1.10$. Let θ_D denote the representativeness parameter for the discrete-time specification developed in BGS. The basic model of diagnostic expectations applied to an AR(1) process predicts that in period t, the forecast of x_{t+1} is

[34] Put differently, the realized future information parameter $\mathcal{I}_{t+\tau}$ is a random variable at time t whereas $\mathcal{I}^S_{t+\tau}$ is deterministic at time t.

biased by $\theta_D \rho u_t$ (see BGS). Using the BGMS estimates of ρ and σ_u, a one standard deviation output growth bias of 0.75 percentage points corresponds to $\theta_D = 0.78$.[35] This is consistent with the estimates of θ_D provided in BGS ($\theta_D = 0.91$), BGMS (estimates vary, with a single collective estimate of $\theta_D = 0.50$), BGLS ($\theta_D = 0.90$), BGST ($\theta_D = 1.08$), and d'Arienzo (2020) (two estimates: $\theta_D = 0.47$ and $\theta_D = 0.70$).

Calculating Sentiment with Expectations Data. This section details how to construct an empirical measure of sentiment given a calibration of κ and θ. This construction is used to calibrate κ. It is also used to calculate the empirical measure of sentiment shown in Figure 9, and to calculate the SPF-implied shocks used in Figure 10.

Sentiment is constructed using the median forecast of real GDP growth from the Survey of Professional Forecasters (SPF).[36] Let FE_t^{SPF} denote the SPF forecast error in quarter t. I define the forecast error in quarter t as the realized GDP growth rate in quarter t minus the median quarter $t-2$ forecast of the GDP growth rate in quarter t. The forecast error is defined with a two-quarter lagged prediction to account for the slow incorporation of shocks into GDP statistics (e.g., Lehman declared bankruptcy in 2008Q3, but this was not reflected in U.S. GDP until 2008Q4).

Proposition 2 provides a method for calculating sentiment using subjective shocks to economic growth. SPF forecast data is collected quarterly while equation (24) is written in continuous time. Equation (24) can be discretized at a quarterly frequency as follows:

$$\mathcal{I}_t = \sum_{j=0}^{t-1} \left(\mathcal{K} + \frac{\theta}{4} \right)^j FE_{t-1-j}^{SPF}, \tag{31}$$

where $\mathcal{K} = e^{-\kappa/4}$. A derivation of equation (31) is provided below.

Equation (31) needs to be initialized with a "period 0." I assume that sentiment

[35] Setting $0.75 = \theta(0.87)(1.1)$ yields the desired result.

[36] Expert forecasts are used for consistency with the model, as discussed in footnote 27.

is equal to 0 in 1970Q1. Starting from the initial condition, equation (31) provides a method for constructing sentiment using quarterly SPF forecast errors. These forecast errors, along with the measure of sentiment in Figure 9, are also used to calculate the shocks used in Figure 10.

κ Calibration. Sentiment persistence parameter κ is calibrated to align the model-implied correlation between e_t and $\mathcal{I}_t$ with the data. I've just outlined a technique for calculating sentiment from SPF forecast data given a calibration of κ and θ. He et al. (2017) provide an empirical measure corresponding to e_t over the period 1970Q1 to 2018Q3. For any given κ and θ, this means that I can calculate the correlation between e_t and $\mathcal{I}_t$ in both the data and the model. This analysis is presented in Figure 1 below, which suggests that sentiment should have a half-life of approximately 5 years.[37]

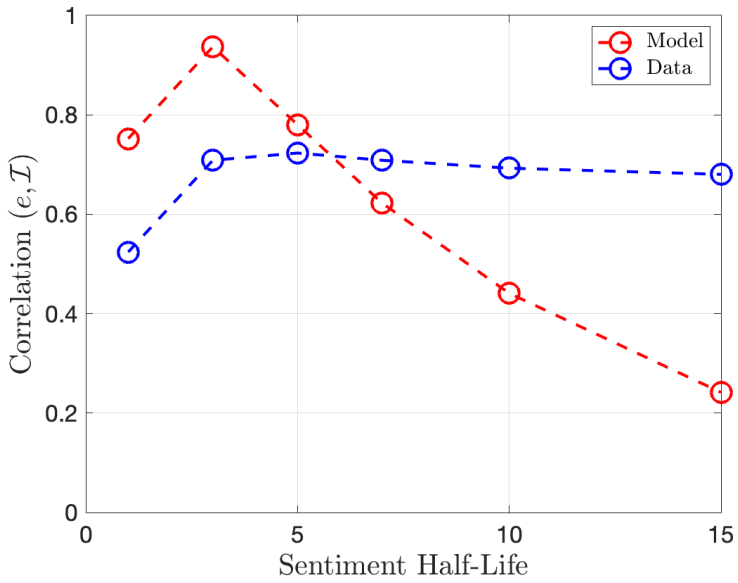


Figure 1: **Calibration of κ.** The red curve plots the correlation between e_t and $\mathcal{I}_t$ in the model as a function of κ (taking θ as given from the diagnostic expectations literature). The blue curve plots the empirical correlation between e_t (from He et al. (2017)) and $\mathcal{I}_t$ (calculated from SPF forecast errors given κ and θ).

[37] As κ is varied, θ is always set to $\frac{(\sqrt{2\kappa})0.0075}{\sigma}$. This maintains the other calibration target that one standard deviation in $\mathcal{I}$ corresponds to an output growth bias of 0.75 percentage points.

Discretization of the Model: Derivation of Equation (31). I partially rewrite the model in discrete time. The discrete-time model is written at a quarterly frequency for consistency with SPF data. Subscript t's denote model periods, such that period $t+1$ occurs one quarter after period t.

Define the capital law of motion as $K_{t+1} = K_t \exp(\upsilon_t + \sigma\epsilon_{t+1})$, where $\sigma\epsilon_t$ are quarterly capital quality shocks, υ_t captures investment and depreciation, and $\epsilon_t \sim \mathcal{N}(0, \frac{1}{4})$. The capital law of motion can be approximated as follows:

$$\frac{K_{t+1} - K_t}{K_t} = \frac{Y_{t+1} - Y_t}{Y_t} \approx \upsilon_t + \sigma\epsilon_{t+1}.$$

Next, I introduce the analogous discrete-time definition of sentiment:

$$\mathcal{I}_t = \sum_{j=0}^{t-1} \mathcal{K}^j \sigma\epsilon_{t-j}, \tag{32}$$

where $\mathcal{K} = e^{-\kappa/4}$. Under diagnostic expectations, agents expect that output (or capital) evolves approximately as follows:[38]

$$\widehat{\mathbb{E}}_t\left[\frac{Y_{t+1} - Y_t}{Y_t}\right] \approx \mathbb{E}_t\left[\frac{Y_{t+1} - Y_t}{Y_t}\right] + \frac{\theta\mathcal{I}_t}{4}.$$

Analogous to equation (5), diagnostic expectations of GDP growth consist of a rational component plus a diagnostic wedge. The dt factor in equation (5) becomes $\frac{1}{4}$ here since the model is discretized at a quarterly frequency.

[38] In logs, the law of motion for capital is $k_{t+1} = k_t + \upsilon_t + \sigma\epsilon_{t+1}$. Let $G_t^- = k_t - \mathcal{I}_t$. In discrete time, equation (26) becomes:

$$h_t^\theta(k'_{t+1}|k_t, e_t, \mathcal{I}_t) = h(k'_{t+1}|k_t, e_t, \mathcal{I}_t) \cdot \left[\frac{h(k'_{t+1}|k_t, e_t, \mathcal{I}_t)}{h(k'_{t+1}|G_t^-, e_t, \mathcal{I}_t)}\right]^{\frac{\theta}{4}} \frac{1}{Z}.$$

The main difference relative to equation (26) is that the power term of θdt becomes $\frac{\theta}{4}$ here, since the model is specified at a quarterly frequency. Since $k_{t+1}|k_t, e_t, \mathcal{I}_t \sim \mathcal{N}\left(k_t + \upsilon_t, \frac{\sigma^2}{4}\right)$ and $k_{t+1}|G_t^-, e_t, \mathcal{I}_t \sim \mathcal{N}\left(G_t^- + \upsilon_t, \frac{\sigma^2}{4}\right)$, a similar argument to that of Appendix B.5 gives $\widehat{\mathbb{E}}_t\left[k_{t+1} - k_t\right] = \upsilon_t + \frac{\theta\mathcal{I}_t}{4}$. This implies $\widehat{\mathbb{E}}_t\left[\frac{Y_{t+1}-Y_t}{Y_t}\right] \approx \mathbb{E}_t\left[\frac{Y_{t+1}-Y_t}{Y_t}\right] + \frac{\theta\mathcal{I}_t}{4}$.

At a quarterly frequency, subjective shocks are defined as

$$\sigma\widehat{\epsilon}_{t+1} \equiv \frac{Y_{t+1} - Y_t}{Y_t} - \widehat{\mathbb{E}}_t \left[\frac{Y_{t+1} - Y_t}{Y_t} \right].$$

Unlike objective shocks $\sigma\epsilon_t$, these subjective shocks are directly observable with expectations data. Using the above approximations:

$$\sigma\widehat{\epsilon}_{t+1} \approx \sigma\epsilon_{t+1} - \frac{\theta \mathcal{I}_t}{4}.$$

Equation (32) can now be redefined in terms of subjective shocks. It follows from equation (32) that $\mathcal{I}_t = \mathcal{K}\mathcal{I}_{t-1} + \sigma\epsilon_t$. Plugging in the approximation of subjective shocks gives

$$\mathcal{I}_t \approx \left(\mathcal{K} + \frac{\theta}{4} \right) \mathcal{I}_{t-1} + \sigma\widehat{\epsilon}_t.$$

Iterating backward and using the initial condition that $\mathcal{I}_0 = 0$ yields:

$$\mathcal{I}_t \approx \sum_{j=0}^{t-1} \left(\mathcal{K} + \frac{\theta}{4} \right)^j \sigma\widehat{\epsilon}_{t-j}.$$

Equation (31) is recovered by noting that $\sigma\widehat{\epsilon}_t = FE_{t-1}$. This holds since $\sigma\widehat{\epsilon}_t$ is the forecast error realized from the start of period $t-1$ to the start of period t.[39]

[39] In other words, $\sigma\widehat{\epsilon}_t$ is the forecast error that is realized over quarter $t-1$.

B Model Details

B.1 Labor Income Microfoundation

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

At time t, there exists a measure $\mathbb{J}_t$ of intermediaries.[40] Each intermediary j operates $K_{j,t}$ units of capital and hires $L_{j,t}$ units of labor at time t. The intermediary 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 endogenous aggregate productivity level. Due to knowledge spillovers, $\bar{A}_t$ depends on the total amount of capital in the economy:

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

Parameter $\varsigma \in [0, 1]$ controls the level of knowledge externalities.

Let $\mathcal{W}_t$ denote the wage rate. Intermediaries hire labor as follows:

$$L_{j,t} = \underset{\ell}{\text{argmax}} \;\; \bar{A}_t K_{j,t}^{\nu} \ell^{1-\nu} - \mathcal{W}_t \ell$$

The optimal labor choice is:

$$L_{j,t} = K_{j,t} \left(\frac{\bar{A}_t (1-\nu)}{\mathcal{W}_t} \right)^{\frac{1}{\nu}}. \tag{35}$$

The next step is to impose market clearing. Specifically, $\int_j K_{j,t} dj = K_t$ and

[40] In the model, there is always a unit measure of households. The size of the financial intermediary sector varies over time due to banker entry and exit.

$\int_j L_{j,t} dj = 1$. Since all intermediaries are identical, $K_{j,t} = \frac{K_t}{\mathbb{J}_t}$ and $L_{j,t} = \frac{1}{\mathbb{J}_t}$. Imposing market clearing gives:

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

The following parameter restriction generates aggregate linearity:

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

Assumption (36) recovers an "AK" economy with a linear aggregate production function $Y_t = AK_t$, exactly as in the main text.

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

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

where the last line follows from (36). The knowledge externalities model built here provides a simple microfoundation for equation (11) in the main text. The benefits of introducing a labor income margin are discussed below.

B.2 Quantitative Benefits of the Labor Income Margin

This model's benchmark calibration sets $A = \frac{1}{3}$ and $\nu = 0.41$. The original HK model, which does not feature labor income, sets $A = 0.133$.[41] Since the average investment rate is roughly 10%, $A = 0.133$ implies that investment typically accounts for more than $\frac{2}{3}$ of the economy's output in the original HK model. Consumption accounts for less than $\frac{1}{3}$ of output. Though some parsimony of the original HK model is lost, the benefit of introducing a simple labor income margin is that it allows for a more realistic consumption-output share.

Generating a realistic consumption-output ratio yields two benefits. First, it allows for a more standard calibration of EIS parameter ζ. In the HK model, consumption accounts for only a small share of output. This implies that changes to the investment rate will cause consumption growth to swing wildly. This is particularly true in periods of financial distress, when the investment rate is sensitive to e_t. HK calibrate $\zeta = 0.13$ (EIS > 7) in order to prevent these swings in consumption from generating excessive interest rate volatility.

Second, the labor income extension generates a more realistic ratio of housing expenditures to total consumption. Recall that aggregate consumption is a Cobb-Douglas aggregator over the output good and housing services: $C_t = (c_t^y)^{1-\phi}(c_t^h)^{\phi}$. Parameter ϕ governs the ratio of housing expenditures to total consumption. To match the housing-wealth ratio of 45%, HK set $\phi = 0.6$. This implies that housing services compose 60% of expenditures.

In order to match the same housing-wealth ratio of 45% in this paper, I calibrate $\phi = 0.2$.[42] In addition to matching the housing-wealth ratio, $\phi = 0.2$ is also consistent with the ratio of housing expenditures to total consumption observed empirically.[43]

[41] The labor income channel in this model can be shut down by setting $\nu = 1$.

[42] Matching the same housing-wealth ratio requires the equilibrium value of rental payments (D_t) to remain similar to the original HK model. Due to the labor income extension, my model features a larger share of output goods relative to housing services (i.e., $A \uparrow$). From equation (10), more output goods implies that a lower calibration of ϕ is needed in order to maintain D_t.

[43] For example, Table 2.4.5 of the 2019 NIPA reports total PCE of \$14563 (billion). Housing and utilities account for \$2670 (billion). The housing share of total consumption is 18%. This share is

B.3 Why Two Assets

At first glance it is puzzling that the model includes two assets, K_t and H, since these assets are perfectly conditionally correlated. The model attempts to jointly match key macroeconomic and financial market data. As a macroeconomic model it aims to generate empirically-plausible levels of investment volatility. As a finance model, enough asset price volatility is needed to produce quantitatively significant nonlinearities during periods of financial distress. These two goals present a well-known problem. Market values of capital are much more volatile than investment, both across firms and over time. In a standard q-theory model where investment is closely linked with asset prices, these two facts can only be reconciled with unreasonably high adjustment costs (e.g., Campbell, 2017, Ch. 7).

By introducing two assets, HK circumvent this issue. The two leftmost panels of Figure 2 show that p_t is more sensitive to e_t than q_t. This is because the supply of houses is fixed while K_t is procyclical.[44] Investment i_t is a function of q_t, so the lower variance of q_t allows for the model to match empirical investment volatility under reasonable adjustment costs. Since the intermediary holds both types of assets, the additional volatility provided by p_t generates an intermediary pricing kernel that is volatile enough to produce significant nonlinearities in financial intermediation. As Section 3.2 highlights, the two-asset model does a good job of matching empirical investment volatility as well as the overall risk-return profile of intermediary equity.

stable over time (Davis and Van Nieuwerburgh, 2015).

[44] In particular, recall that P_t is the present-discounted value of perceived future rental payments. When e_t is low, investment rate i_t is also expected to be low for a long period of time. This implies that the expected growth rate of output — and therefore rental payments — is expected to be low, too. Through this investment channel, the growth rate of rental payments is highly correlated with e_t. Hence, housing price p_t is sensitive to e_t.

B.4 Boundary Conditions

Boundary conditions are needed to solve for price functions $q(e, \mathcal{I})$ and $p(e, \mathcal{I})$. As $e \to \infty$ the equity issuance constraint ceases to affect the equilibrium price and policy functions. This implies $\lim_{e\to\infty} q_e(e, \mathcal{I}) = \lim_{e\to\infty} p_e(e, \mathcal{I}) = 0$.

A lower reflecting boundary is imposed by assuming entry into the intermediary sector deep in crisis times (term $d\psi_t$ in equation (8)). There exists an exogenous minimum reputation level $\underline{e}$ such that new intermediaries enter the economy whenever e_t hits $\underline{e}$.[45] Entry is costly because new intermediaries must acquire the skills to operate capital. Specifically, the economy must destroy $\beta > 0$ units of capital in order for entry to increase aggregate reputation $\mathcal{E}_t$ by one unit.[46] The assumption of a reflecting barrier at $\underline{e}$ implies that prices q and P must have a zero derivative with respect to e at $\underline{e}$. If this were not the case, an arbitrageur could bet on a unidirectional change in asset prices at the reflecting barrier. This implies $q_e(\underline{e}, \mathcal{I}) = 0$ and $p_e(\underline{e}, \mathcal{I}) = \frac{p(\underline{e},\mathcal{I})\beta}{1+\underline{e}\beta}$.

To derive the lower boundary condition for p, 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:[47]

$$x = \frac{z}{1 + \underline{e}\beta}.$$

Shock z requires the destruction of βx capital to restore reputation to $\underline{e}$.

This capital destruction equation can be used to derive the boundary condition for price function P. The condition that $P_e(\underline{e}, \mathcal{I}) = 0$ implies that $p = \frac{P}{K}$ will have

[45] Loosely, this captures (unmodeled) government intervention deep in crises.

[46] Equations (2) and (4) are altered at $\underline{e}$ to include this form of capital destruction.

[47] x is defined implicitly by $\underline{e} = \frac{\underline{\mathcal{E}}-z+x}{K-\beta x}$. The numerator of this equation is the level of initial reputation minus the shock and plus x of new reputation. The denominator is the initial level of capital minus the βx of capital destroyed to produce x reputation. Capital will be destroyed until $\underline{e}$ (the minimum level of $e = \mathcal{E}/K$) is restored.

a non-zero slope at $\underline{e}$.[48] 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 rebuild 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 $p(\underline{e}, \mathcal{I}')K - p_e\left(\underline{e}, \mathcal{I}'\right) z$. Combining:

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

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

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

B.5 Proofs

Proof of Proposition 2. By definition, $\mathcal{I}_t \equiv \int_0^t e^{-\kappa(t-s)}\sigma dZ_s$. Subjective shocks are defined as $\sigma\widehat{dZ_t} = -\theta\mathcal{I}_t dt + \sigma dZ_t$. The law of motion for $\mathcal{I}_t$ is given by $d\mathcal{I}_t = -\kappa\mathcal{I}_t dt + \sigma dZ_t$. Plugging in the definition of subjective shocks gives

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

Let $f(\mathcal{I}_t, t) = \mathcal{I}_t e^{(\kappa-\theta)t}$. Using Itô's lemma:

$$\begin{aligned} df(\mathcal{I}_t, t) &= e^{(\kappa-\theta)t}\left(d\mathcal{I}_t\right) + (\kappa-\theta)\mathcal{I}_t e^{(\kappa-\theta)t}dt \\ &= e^{(\kappa-\theta)t}\left((-\kappa+\theta)\mathcal{I}_t dt + \sigma\widehat{dZ_t}\right) + (\kappa-\theta)\mathcal{I}_t e^{(\kappa-\theta)t}dt \\ &= e^{(\kappa-\theta)t}\sigma\widehat{dZ_t}. \end{aligned}$$

Given an initial condition of $f(\mathcal{I}_0, 0) = 0$, we get:

$$\begin{aligned} f(\mathcal{I}_t, t) &= \int_0^t df(\mathcal{I}_s, s) \\ \mathcal{I}_t e^{(\kappa-\theta)t} &= \int_0^t e^{(\kappa-\theta)s}\sigma\widehat{dZ_s} \\ \mathcal{I}_t &= \int_0^t e^{(-\kappa+\theta)(t-s)}\sigma\widehat{dZ_s}. \end{aligned}$$

This completes the proof.

Proof of Proposition 3 (in Appendix A.1). Consider the evolution of log capital k_t over a horizon of $\tau > 0$, holding investment fixed at $\bar{i} = i_t$. With fixed investment:

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

Since $\{Z_t\}$ is a standard Brownian motion, $k_{t+\tau} \sim \mathcal{N}\left(k_t + \tau(\bar{i} - \delta - \frac{\sigma^2}{2}),\ \sigma^2\tau\right)$.

For the arbitrary prediction horizon of τ, one can rewrite equation (26) 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:

$$h_t^\theta(k'_{t+\tau}|k_t, e_t, \mathcal{I}_t) = \mathcal{N}\left(k_t + \tau(\bar{i} - \delta - \frac{\sigma^2}{2}) + \theta(k_t - G_t^-)\tau,\ \sigma^2\tau\right).$$

The Bordalo et al. (2018a) Appendix provides algebraic details for this step. Equivalently, the agent perceives that

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

In the limit as $\tau \to dt$:

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

Definition 3 gives $\mathcal{I}_t = k_t - G_t^-$. Applying Itô's lemma to $K_t = exp(k_t)$ yields:

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

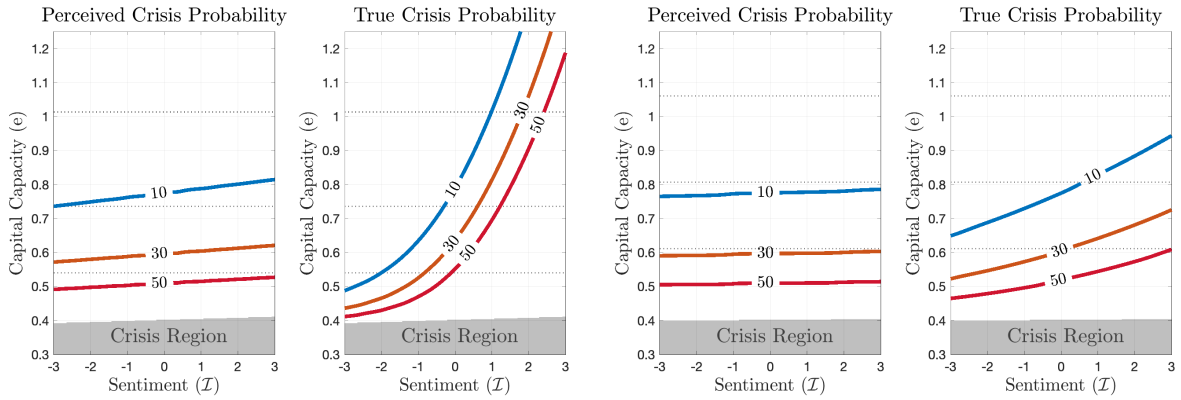
This completes the proof.

B.6 Robustness

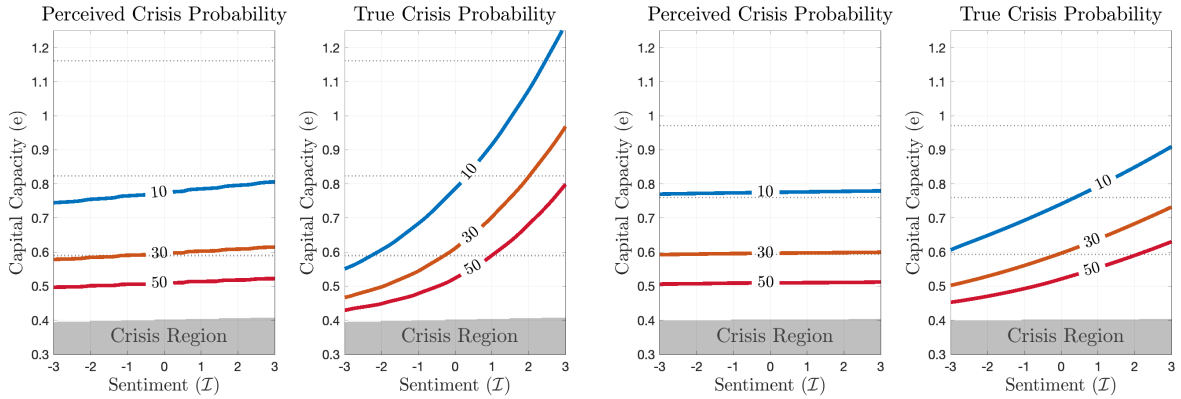
This section examines robustness to behavioral parameters θ and κ, and the EIS parameter ζ. The baseline calibration sets $\theta \times SD(\mathcal{I}) = 0.75\%$, the half-life of $\mathcal{I}$ to 5 years, and $\zeta = \frac{2}{3}$ (EIS = 1.5). I analyze the following six parameter perturbations: (i) $\theta \times SD(\mathcal{I}) = 1.5\%$, (ii) $\theta \times SD(\mathcal{I}) = \frac{0.75\%}{2}$, (iii) sentiment half-life of 20 years, (iv) sentiment half-life of 1 year, (v) $\zeta = 1$, and (vi) $\zeta = \frac{1}{2}$.

For perturbations to θ and ζ, all other parameters are kept at their baseline calibration in Table 1. For perturbations to the persistence 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 persistence.

(a) $\theta \times SD(\mathcal{I}) = 1.5\%$ (b) $\theta \times SD(\mathcal{I}) = \frac{0.75}{2}\%$



(c) Half-life = 20 (d) Half-life = 1



(e) $\zeta = 1$ (f) $\zeta = \frac{1}{2}$

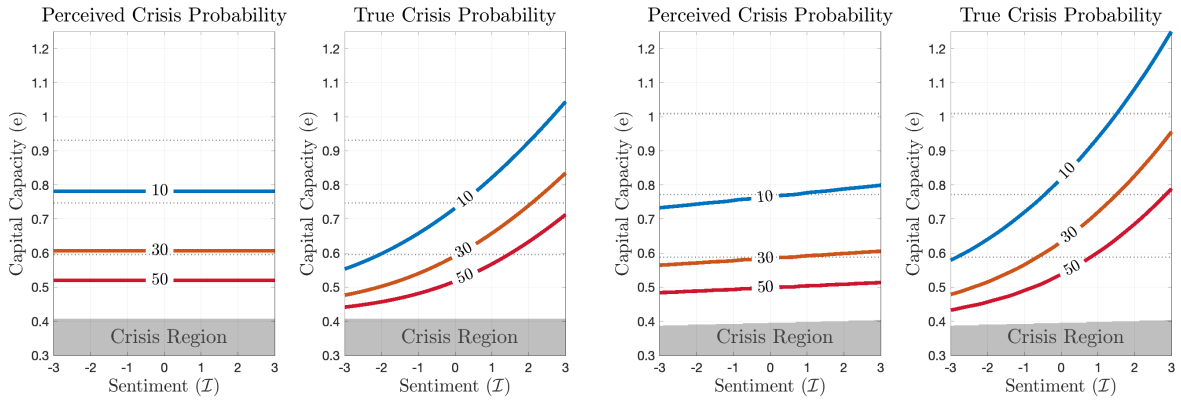


Figure 2: **Crisis hitting probabilities**. See Figure 4 for a description. In order from left-to-right, top-to-bottom: $\theta \times SD(\mathcal{I}) = 1.5\%$, $\theta \times SD(\mathcal{I}) = \frac{0.75}{2}\%$, half-life = 20, half-life = 1, $\zeta = 1$, $\zeta = \frac{1}{2}$.

Boom-Bust Investment Cycles.

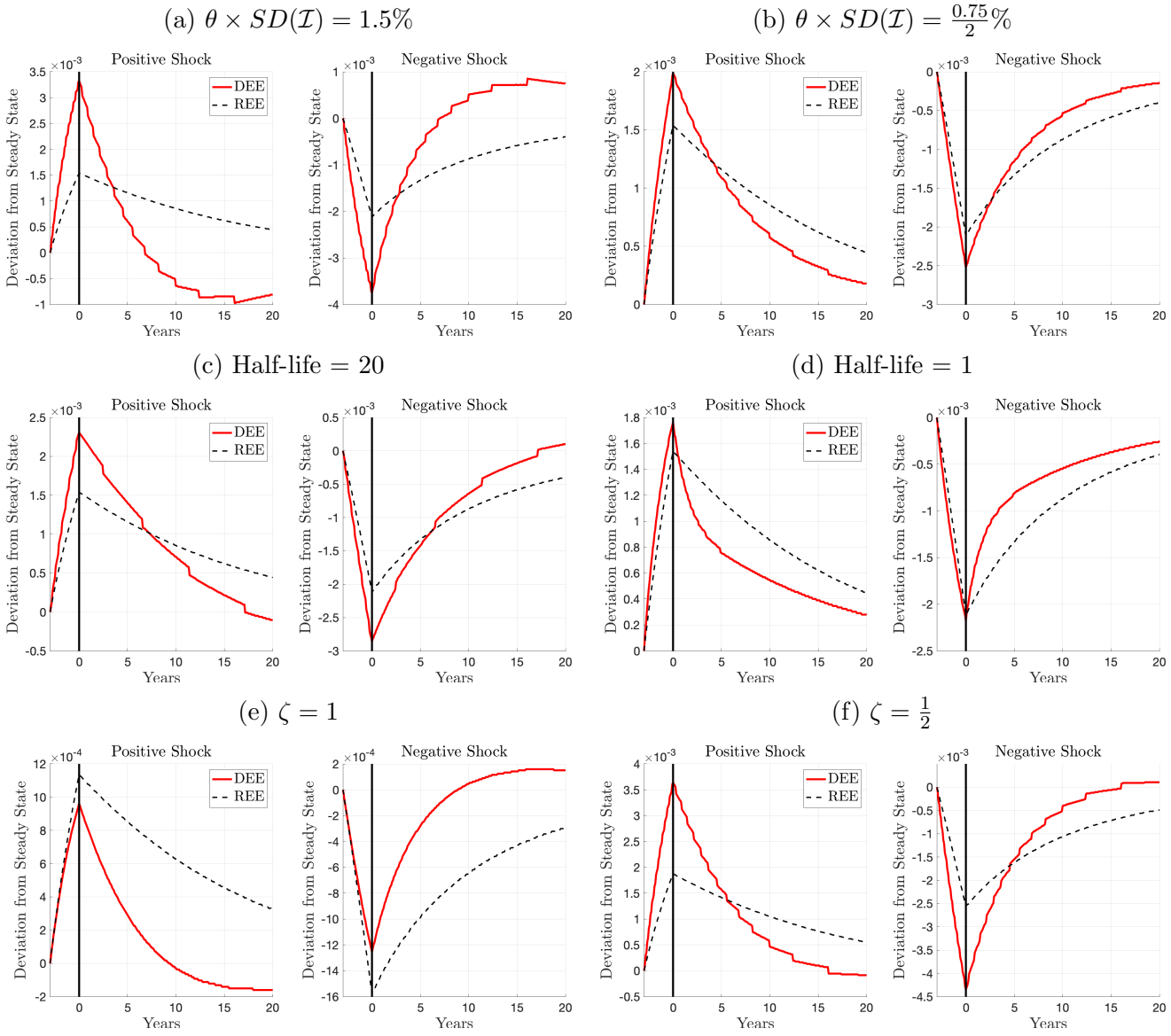


Figure 3: **Investment rate IRFs**. See Figure 6 for a description. In order from left-to-right, top-to-bottom: $\theta \times SD(\mathcal{I}) = 1.5\%$, $\theta \times SD(\mathcal{I}) = \frac{0.75}{2}\%$, half-life = 20, half-life = 1, $\zeta = 1$, $\zeta = \frac{1}{2}$.

Financial Market Stability from Beliefs. Appendix Table 3 lists crisis probabilities for the DEE and the REE across the six alternate calibrations.

	DEE Prob(Crisis)	**REE Prob(Crisis)**	$\frac{\textbf{DEE Prob(Crisis)}}{\textbf{REE Prob(Crisis)}}$
Baseline	3.22 %	3.83 %	0.84
$\theta \times SD(\mathcal{I}) = 1.5\%$	6.33 %	3.83 %	1.65
$\theta \times SD(\mathcal{I}) = \frac{0.75\%}{2}$	3.10 %	3.83 %	0.81
Half-life = 20	4.71 %	3.83 %	1.23
Half-life = 1	2.74 %	3.83 %	0.72
$\zeta = 1$	2.80 %	5.28 %	0.53
$\zeta = \frac{1}{2}$	3.43 %	2.80 %	1.23

Table 3: **Financial market stability from beliefs: robustness**.

First, Appendix Table 3 evaluates the effect of doubling and halving θ. The doubling of θ makes diagnostic expectations destabilizing relative to rationality. With higher θ, intermediaries make larger asset pricing mistakes. This puts downward pressure on e_t when $\mathcal{I}_t$ is high, and vice-versa. Overall, the doubling of θ reduces the correlation between e_t and $\mathcal{I}_t$ to 0.2. This makes it more likely that the economy enters the part of the state space where sentiment is elevated and the financial sector is distressed. This increases the probability of crises in the DEE. Alternatively, diagnostic expectations become more stabilizing when θ is halved. Lowering θ increases the correlation between e_t and $\mathcal{I}_t$ to 0.9. This makes it even less likely for the economy to enter states where the financial sector is distressed and yet $\mathcal{I}_t > 0$.[49]

The persistence of sentiment also matters for the effect of diagnostic expectations on financial crises. Crises become more likely to occur as the persistence of sentiment increases. Again, this effect operates through the correlation between e_t and $\mathcal{I}_t$. When sentiment has a half-life of 1 year, the correlation is 0.75. This drops to 0.1 for a half-life of 20 years. Decreasing κ reduces the correlation between e_t and $\mathcal{I}_t$ by making sentiment load less and less on the most recent shocks.

Finally, Appendix Table 3 evaluates the effect of ζ on financial crises. ζ is the

[49]Though not illustrated in Appendix Table 3, the relationship between θ and the probability of crises 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.

inverse of the EIS. 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. This result mirrors the "volatility paradox" of Brunnermeier and Sannikov (2014). Lowering ζ increases asset price volatility, so intermediaries demand higher risk premia.[50] In equilibrium, this decreases the probability of crises.

ζ works in the opposite direction in the DEE. Under diagnostic expectations, lower ζ increases the probability of crises. In the DEE, asset prices become increasingly sensitive to $\mathcal{I}_t$ as ζ decreases.[51] However, diagnostic agents do not anticipate this sentiment-driven price volatility and therefore they do not demand a higher risk premium as compensation. Since lowering ζ increases asset price volatility without increasing the corresponding risk premium, this additional (uncompensated) risk increases the probability of financial crises.

[50] Lowering ζ decreases the sensitivity of the risk-free rate to variation in e_t. This means that asset prices must fluctuate instead. For example, consider a decline in e_t which causes intermediaries to demand a higher risk premium. This 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.

[51] As ζ decreases the interest rate r_t becomes less sensitive to changes in expected consumption growth. Consider an increase in sentiment. This increases perceived future cash flows. If 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. Thus, lowering ζ creates larger asset price variation in $\mathcal{I}_t$.

B.7 Empirical Tests in the Recalibrated REE

In Section 6.1 the calibration of the REE is identical to the calibration of the DEE, except for $\theta = 0$. Here, I recalibrate the REE in accordance with Table 1. The analysis of Section 6.1 is repeated using the recalibrated REE.

Updated REE Calibration. Table 4 presents the updated calibration of the REE. Four parameters are changed from the main text: η, $\underline{e}$, β, and ϕ. Bank exit rate η is updated to re-establish a crisis probability of 3%. Lower entry barrier $\underline{e}$ is changed to maintain a maximum Sharpe ratio of 6.5. Entry cost β is updated to maintain land price growth volatility of roughly 11.9%. ϕ is updated to maintain a housing-wealth ratio of 45%.

	Parameter	**Choice**	**Target**
Updated Parameters			
η	Bank Exit Rate	0.139	Prob(Crisis)
$\underline{e}$	Lower Entry Barrier	0.112	Max $\mathcal{I} = 0$ Sharpe Ratio
β	Entry Cost	4.5	Land Price Volatility
ϕ	Housing Expenditure Share	0.204	Housing-Wealth Ratio
Unconditional Simulated Moments		REE (New)	REE (Old)
	Mean$\left(\frac{\text{Investment}}{\text{Capital}}\right)$	10.04%	9.93%
	Mean$\left(\frac{\text{Consumption}}{\text{Output}}\right)$	69.88%	70.43%
	Mean(Realized Sharpe Ratio)	0.54	0.52
	Mean(Realized Intermediary Risk Premium)	15.93%	14.63%
	Probability of Crisis	3.05%	3.83%
	Volatility(Land Price Growth)	11.64%	10.08
	Volatility(Interest Rate)	0.46%	0.41%
Non-Distress Simulated Moments			
	Volatility(Investment Growth)	3.49%	3.98%
	Volatility(Consumption Growth)	2.75%	2.57%
	Volatility(Output Growth)	2.97%	2.98%
	Mean$\left(\frac{\text{Housing Wealth}}{\text{Total Wealth}}\right)$	44.60%	46.79%

Table 4: **REE updated calibration**. See Table 1 for a description. Column REE (New) lists the simulated moments in the recalibrated REE. Column REE (Old) lists the simulated moments for the REE calibration used in the main text.

Persistence in the Recalibrated REE. Appendix Table 5 uses the recalibrated REE to reproduce Table 2 of the main text. The recalibrated REE features a slightly lower persistence of financial distress, but results are similar.

e_t Percentile	REE (New)	REE (Old)
5	0.46	0.42
10	1.13	1.08
25	3.74	3.72
50	10.11	10.42
75	24.84	26.41

Table 5: **Average crisis recovery time (in years)**. See Table 2 for a description. This table compares the persistence of financial fragility in the two REE calibrations.

Appendix Figure 4 uses the recalibrated REE to reproduce Figure 8 of the main text. Autocorrelations in the recalibrated REE are presented with solid black lines. The two calibrations produce very similar results.

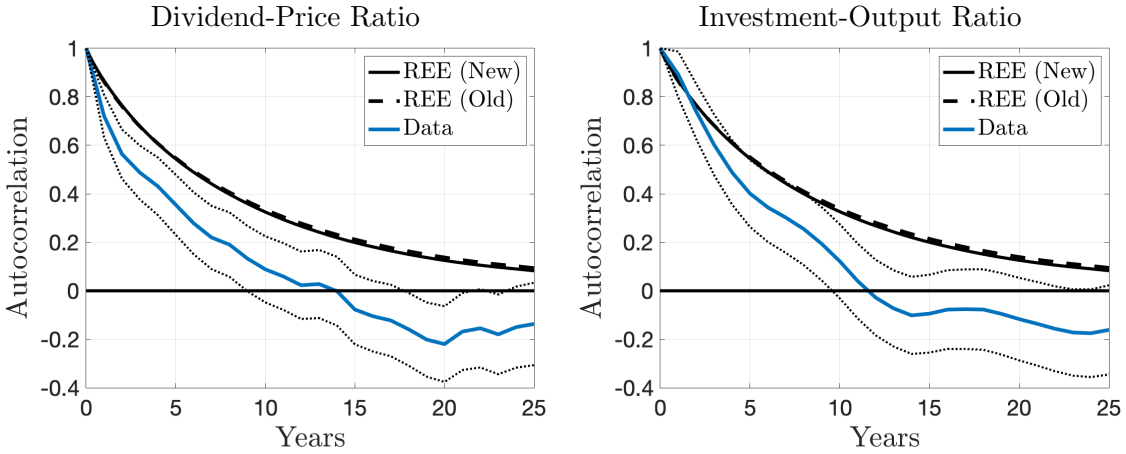


Figure 4: **Persistence: data and model**. See Figure 8 for a description. This figure compares the persistence of macro-financial aggregates in the two REE calibrations.

B.8 Simulating the 2007-2008 Financial Crisis: Additional Analysis

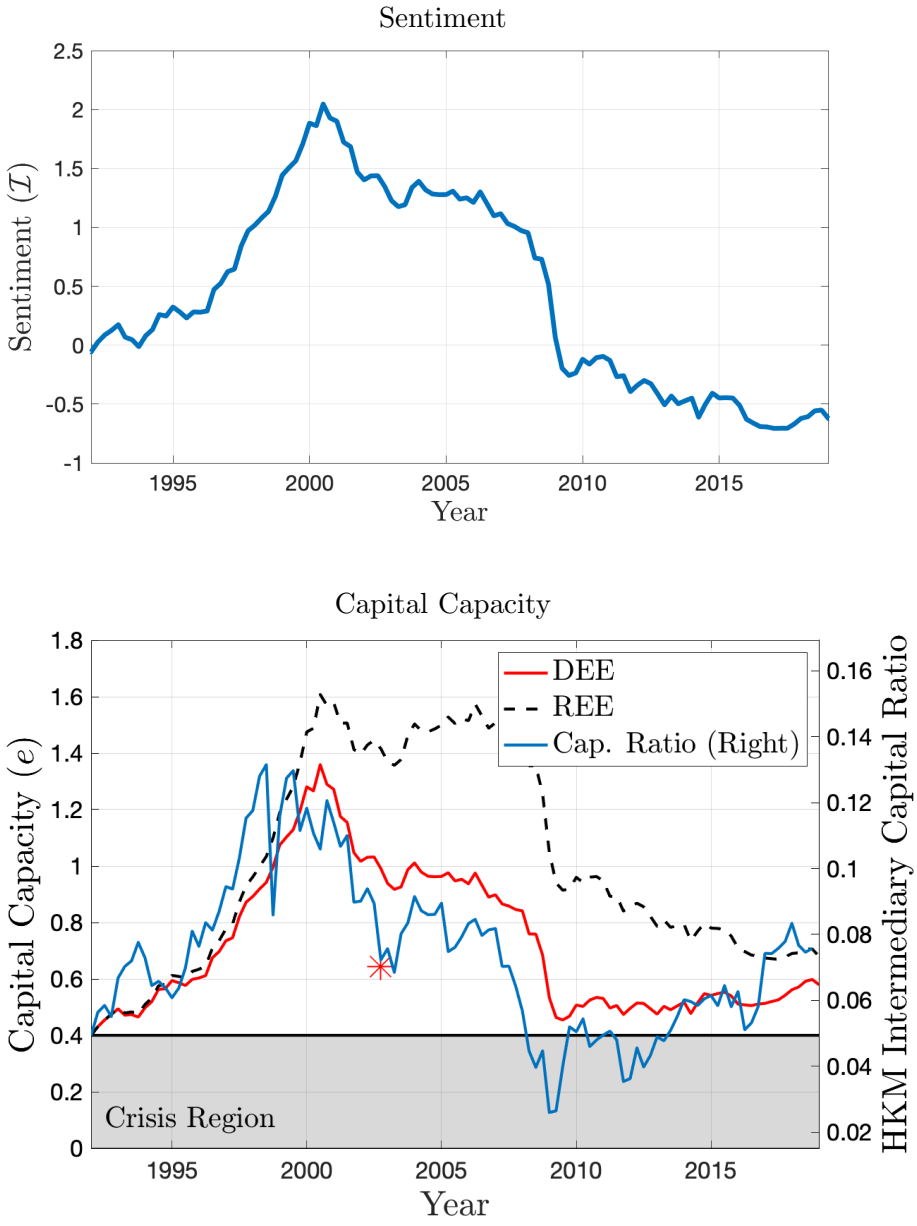


Figure 5: **A long-run analysis of the 2007-2008 Financial Crisis**. See Section 6.3 for a description. For reference, the top panel shows the path of sentiment from 1992 through 2018. The bottom panel plots the path of capital capacity e_t in the DEE and the REE (left axis), as well as the corresponding empirical measure from He et al. (2017) (right axis). The red star marks the e_t initialization of the second simulation, as shown in Figure 10 of the main text.

B.9 Additional Tables and Figures

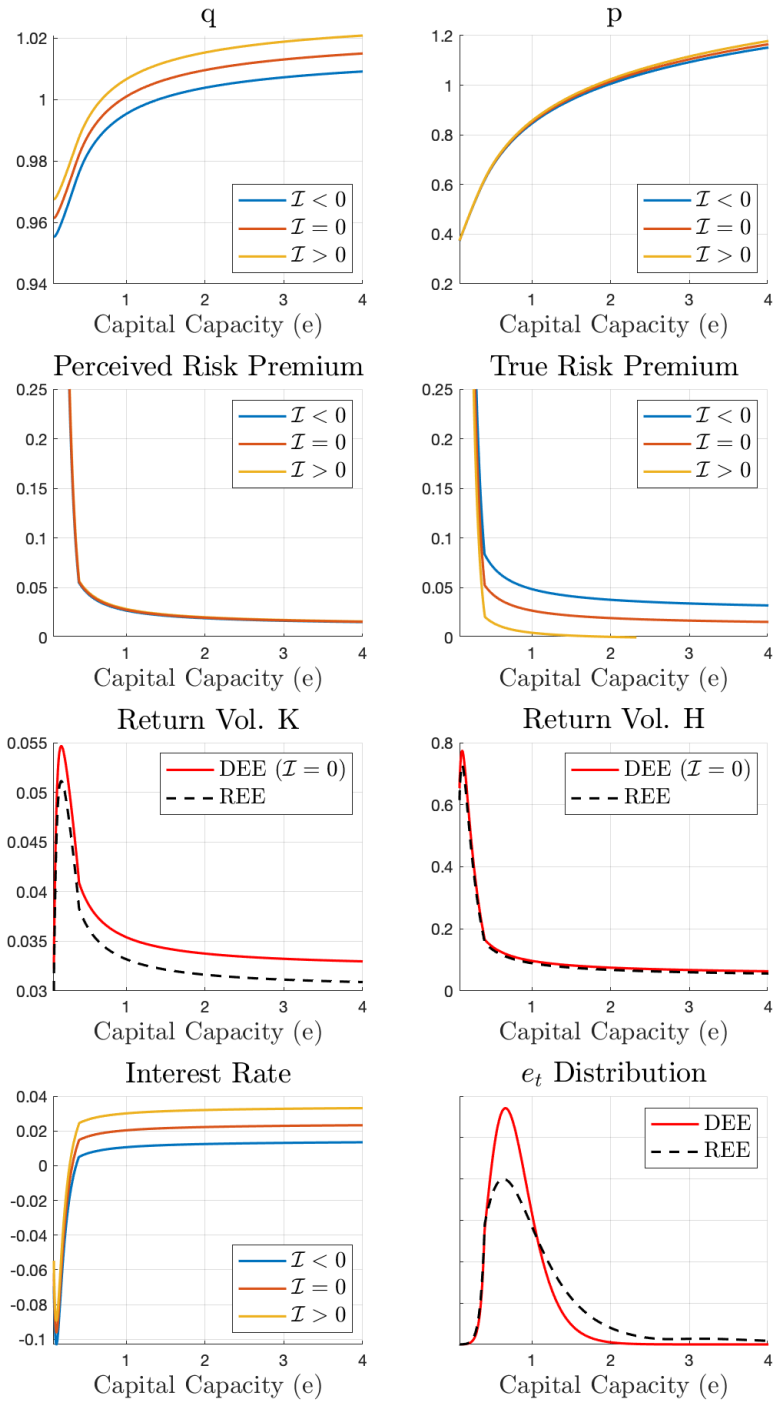


Figure 6: **Price and policy functions**. See Figure 2 for a description. This chart also contains information about delevered risk premia, asset volatilities (σ^k and σ^h), and interest rate r_t.

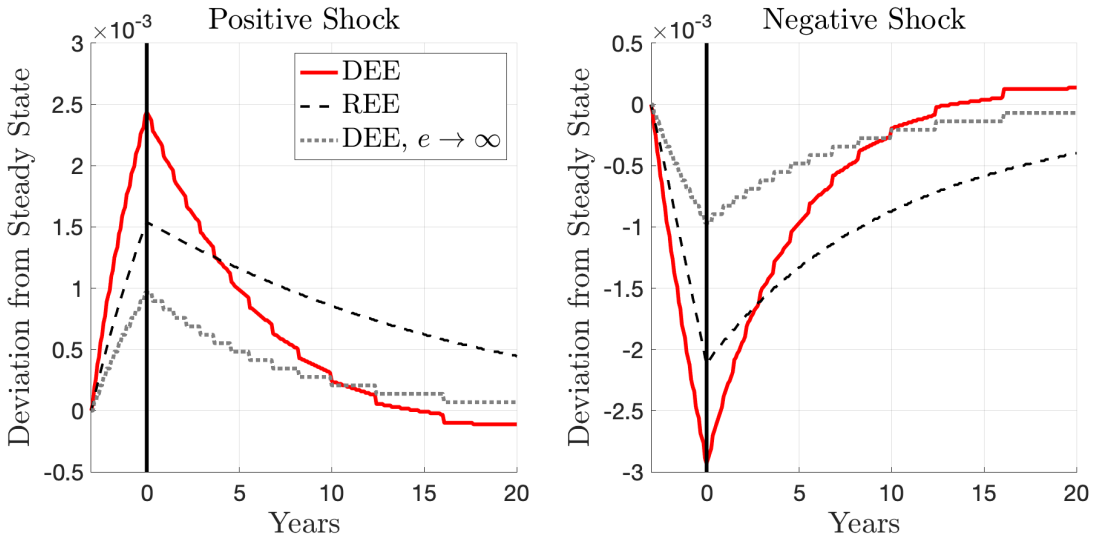


Figure 7: **Investment rate IRFs**. See Figure 6 for a description. This figure also includes the investment IRFs of an economy with diagnostic expectations but no financial frictions (dotted gray curve). Sentiment produces short-run momentum, but there is no feedback from behavioral frictions to financial frictions to generate sharp reversals thereafter.

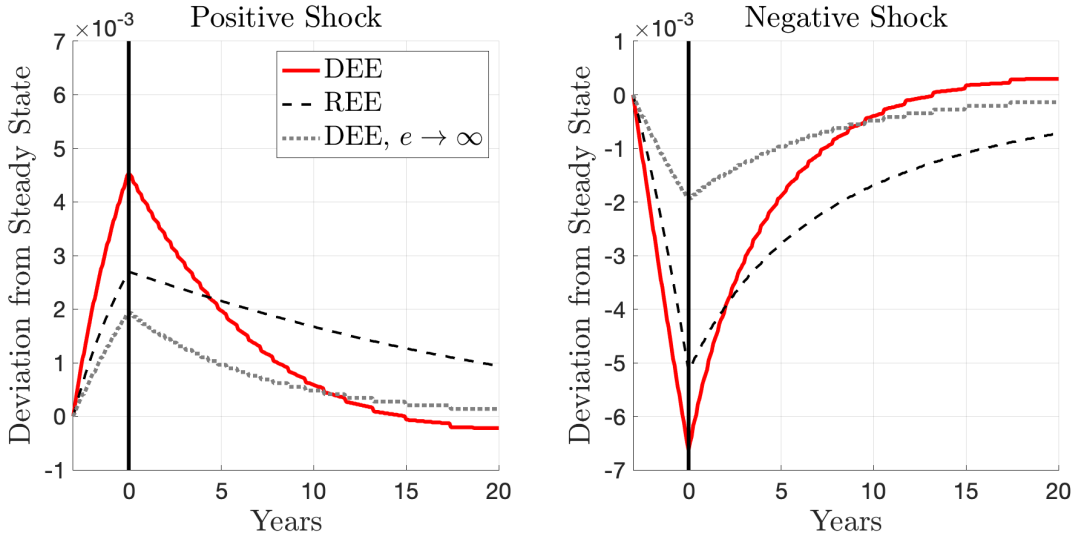


Figure 8: **Investment rate IRFs:** $2SD$ **shock**. See Figure 6 for a description. This figure replicates the investment rate IRF analysis, but doubles the size of the initial impulse. Due to the model's nonlinearity, the larger shock sequence produces a starker asymmetry between the positive and negative shock cases.

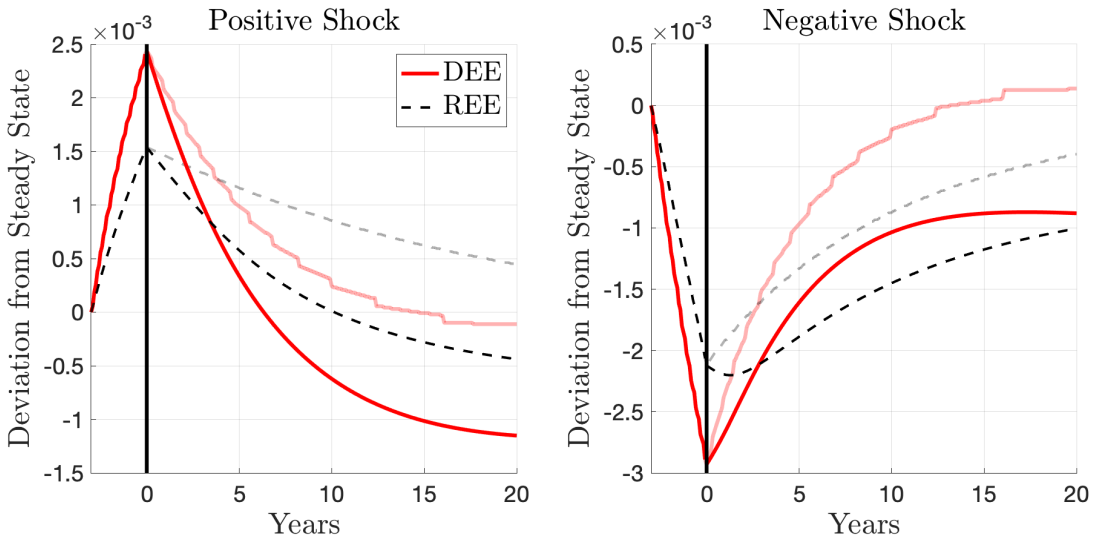


Figure 9: **Investment rate IRFs:** $\mathbb{E}_0[i_\tau]$. The baseline analysis in Figure 6 sets shocks to 0 for all $\tau \geq 0$ (shown here with transparent lines for reference). This figure instead plots the expected future investment rate $\mathbb{E}_0[i_\tau]$. Because this model is nonlinear, these two approaches are not equivalent. In particular, nonlinear financial frictions mean that the long-run expected investment rate, $\lim_{\tau\to\infty} \mathbb{E}_0[i_\tau]$, is less than the investment rate in the stochastic steady state. The expected investment rate is calculated using the Feynman-Kac formula (see Appendix D.3 for numerical details).

e_t Percentile	DEE	REE
5	0.25	0.25
10	0.67	0.83
25	1.83	2.83
50	4.00	8.08
75	8.17	20.08

Table 6: **Median crisis recovery time (in years)**. See Table 2 for a description. This table lists the median time (in years) that it takes for e_t to recover from a financial crisis to its X^{th} percentile.

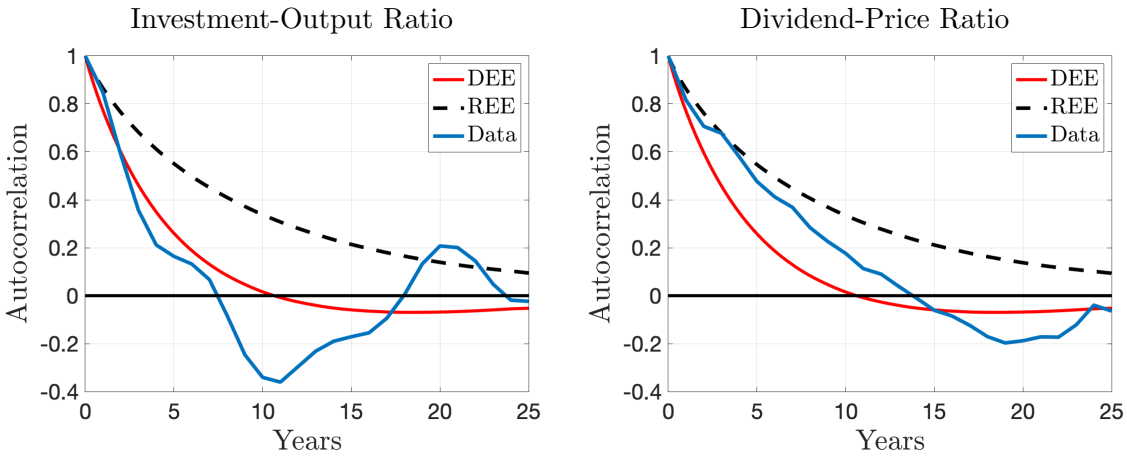


Figure 10: **Persistence: data and model**. This figure replicates Figure 8 using only U.S. data from 1950 – 2016.

C 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 an excellent treatment of the techniques used here.

C.1 The Diagnostic Expectations Equilibrium

I postulate that agents perceive that q_t and p_t evolve as follows:

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

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

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

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

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

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

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

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

Price Processes. 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. Equations (43) and (45) are second-order PDEs for the price functions, which I will solve numerically. To do so, 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 (21):

$$\begin{aligned}
Y_t &= C_t^y + \Phi(i_t, K_t) \\
A &= \frac{C_t^y}{K_t} + i_t + \frac{\xi}{2}(i_t - \delta)^2 \\
&= \frac{1-\phi}{\phi} \frac{D_t}{K_t} + i_t + \frac{\xi}{2}(i_t - \delta)^2, \text{ using (10) and } C_t^h = 1 \\
\frac{D_t}{K_t} &= \frac{\phi}{1-\phi} \left[A - i_t - \frac{\xi}{2}(i_t - \delta)^2 \right].
\end{aligned}$$

Since equation (12) gives $i_t = \delta + \frac{q_t - 1}{\xi}$, this pins 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:

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

Applying Itô's Lemma to the first term:

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

Equation (17) gives a similar process for capital returns. Repeating (17) 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:

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

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

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

Plugging this into the interest rate formula (14):

$$r_t = \rho + \zeta\left[i_t - \delta + \theta\mathcal{I}_t - \frac{\frac{1}{\xi}q_t^2(\widehat{\mu_t^q} + \frac{1}{2}\widehat{\sigma_t^q}^2 + \sigma\widehat{\sigma_t^q})}{A - \delta - \frac{q_t-1}{\xi} - \frac{(q_t-1)^2}{2\xi}}\right] - \frac{\zeta(\zeta+1)}{2}\left[\sigma - \frac{\frac{1}{\xi}q_t^2\widehat{\sigma_t^q}}{A - \delta - \frac{q_t-1}{\xi} - \frac{(q_t-1)^2}{2\xi}}\right]^2 \tag{47}$$

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 (20) gives:

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

Recall $\widehat{\sigma_t^k} = \sigma + \widehat{\sigma_t^q}$ and $\widehat{\sigma_t^h} = \sigma + \widehat{\sigma_t^p}$. Combining the banker's optimality condition with the perceived return on capital:

$$\begin{aligned}
\gamma(\widehat{\sigma_t^e} + \sigma) &= \frac{\widehat{\pi_t^k}}{\widehat{\sigma_t^k}} \\
&= \frac{\left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta\mathcal{I}_t + \sigma\widehat{\sigma_t^q}\right) - r_t}{\sigma + \widehat{\sigma_t^q}} \\
&= \frac{\left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta\mathcal{I}_t + \sigma\frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}\right) - r_t}{\sigma + \frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}}, \text{ using equation (44)} \\
\gamma(\widehat{\sigma_t^e} + \sigma)\left(\sigma + \frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}\right) &= \left(\frac{\nu A}{q_t} + \widehat{\mu_t^q} - \delta + \theta\mathcal{I}_t + \sigma\frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e}\right) - r_t \quad (48)
\end{aligned}$$

Proceeding similarly for housing returns:

$$\begin{aligned}
\gamma(\widehat{\sigma_t^e} + \sigma) &= \frac{\widehat{\pi_t^h}}{\widehat{\sigma_t^h}} \\
&= \frac{\widehat{\mu_t^p} + i_t - \delta + \theta\mathcal{I}_t + \sigma\widehat{\sigma_t^p} + \frac{D_t}{p_t K_t} - r_t}{\sigma + \widehat{\sigma_t^p}} \\
&= \frac{\widehat{\mu_t^p} + i_t - \delta + \theta\mathcal{I}_t + \sigma\frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e} + \frac{D_t}{p_t K_t} - r_t}{\sigma + \frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e}}, \text{ using (46)} \\
\gamma(\widehat{\sigma_t^e} + \sigma)\left(\sigma + \frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e}\right) &= \widehat{\mu_t^p} + i_t - \delta + \theta\mathcal{I}_t + \sigma\frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e} + \frac{D_t}{p_t K_t} - r_t
\end{aligned} \tag{49}$$

Pinning down $\widehat{\sigma_t^e}$. Equations (48) and (49) 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 $\widehat{\sigma_t^e}$:

$$\begin{aligned}
\widehat{\sigma_t^e} &= \alpha_t^h\widehat{\sigma_t^h} + \alpha_t^k\widehat{\sigma_t^k} - \sigma \\
&= \alpha_t^h(\sigma + \widehat{\sigma_t^p}) + \alpha_t^k(\sigma + \widehat{\sigma_t^q}) - \sigma \\
&= \frac{K_t}{E_t}\left[p_t(\sigma + \widehat{\sigma_t^p}) + q_t(\sigma + \widehat{\sigma_t^q})\right] - \sigma, \text{ using (22) and (23)} \\
&= \frac{K_t}{E_t}\left[p_t(\sigma + \frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\widehat{\sigma_t^e}) + q_t(\sigma + \frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\widehat{\sigma_t^e})\right] - \sigma \\
&= \frac{K_t}{E_t}\left[(p_t + q_t - \frac{E_t}{K_t})\sigma + (p_e(e_t,\mathcal{I}_t) + q_e(e_t,\mathcal{I}_t))e_t\widehat{\sigma_t^e}\right] \\
e_t\widehat{\sigma_t^e}\left(\frac{1}{e_t} - \frac{K_t}{E_t}(p_e(e_t,\mathcal{I}_t) + q_e(e_t,\mathcal{I}_t))\right) &= \frac{K_t}{E_t}(p_t + q_t - \frac{E_t}{K_t})\sigma \\
e_t\widehat{\sigma_t^e} &= \frac{K_t}{E_t}\frac{(p_t + q_t - \frac{E_t}{K_t})\sigma}{\frac{1}{e_t} - \frac{K_t}{E_t}p_e(e_t,\mathcal{I}_t) - \frac{K_t}{E_t}q_e(e_t,\mathcal{I}_t)} \\
e_t\widehat{\sigma_t^e} &= \frac{(p_t + q_t - \frac{E_t}{K_t})\sigma}{\frac{E_t}{K_t}\frac{1}{e_t} - p_e(e_t,\mathcal{I}_t) - q_e(e_t,\mathcal{I}_t)}.
\end{aligned} \tag{50}$$

Recall that $E_t = \min\{\mathcal{E}_t, (1-\lambda)(q_t K_t + p_t K_t)\}$, or equivalently $\frac{E_t}{K_t} = \min\{e_t, (1-\lambda)(q_t + p_t)\}$. Thus, equation (50) 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. I can now solve for price functions $q(e, \mathcal{I})$ and $p(e, \mathcal{I})$. 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. Specifically, r, $\widehat{\mu_t^q}$, and $\widehat{\mu_t^p}$ are given by equations (47), (48) and (49). $\widehat{\mu_t^e}$ is given by (41), noting that $\widehat{\widetilde{dR_t}}$ is itself a function of prices r, $\widehat{\mu_t^q}$, $\widehat{\mu_t^p}$, $\widehat{\sigma_t^q}$ 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 (50) closes the loop by solving for $\widehat{\sigma_t^e}$ in terms of the two price functions.

This system of PDEs is solved numerically. Details are in Appendix D.

C.2 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:

$$\frac{dq_t}{q_t} = \mu_t^q dt + \sigma_t^q dZ_t \tag{51}$$
$$\frac{dp_t}{p_t} = \mu_t^p dt + \sigma_t^p dZ_t. \tag{52}$$

The true evolution of the three state variables is:

$$\frac{dK_t}{K_t} = (i_t - \delta)dt + \sigma dZ_t, \tag{53}$$
$$\frac{de_t}{e_t} = (r_t + \alpha_t^h \pi_t^h + \alpha_t^k \pi_t^k)dt - (i_t - \delta)dt + (\sigma^2 - \sigma(\alpha_t^h \sigma_t^h + \alpha_t^k \sigma_t^k))dt$$
$$- \eta dt + d\psi_t + (\alpha_t^h \sigma_t^h + \alpha_t^k \sigma_t^k - \sigma)dZ_t \tag{54}$$
$$d\mathcal{I}_t = -\kappa \mathcal{I}_t dt + \sigma dZ_t \tag{55}$$

As above, equation (54) 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 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 show how to solve for price functions $q(e, \mathcal{I})$ and $p(e, \mathcal{I})$. Applying Itô's lemma to these price functions using the true laws of motion for the two state variables gives:

$$\mu_t^q = \frac{q_e(e_t, \mathcal{I}_t)}{q_t} e_t \mu_t^e - \frac{q_{\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \kappa \mathcal{I}_t + \frac{q_{e\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \sigma(e_t \sigma_t^e) + \frac{1}{2} \frac{q_{ee}(e_t, \mathcal{I}_t)}{q_t} (e_t \sigma_t^e)^2 + \frac{1}{2} \frac{q_{\mathcal{I}\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \sigma^2 \tag{56}$$

$$\sigma_t^q = \frac{q_e(e_t, \mathcal{I}_t)}{q_t} e_t \sigma_t^e + \frac{q_{\mathcal{I}}(e_t, \mathcal{I}_t)}{q_t} \sigma \tag{57}$$

$$\mu_t^p = \frac{p_e(e_t, \mathcal{I}_t)}{p_t} e_t \mu_t^e - \frac{p_{\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \kappa \mathcal{I}_t + \frac{p_{e\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \sigma(e_t \sigma_t^e) + \frac{1}{2} \frac{p_{ee}(e_t, \mathcal{I}_t)}{p_t} (e_t \sigma_t^e)^2 + \frac{1}{2} \frac{p_{\mathcal{I}\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \sigma^2 \tag{58}$$

$$\sigma_t^p = \frac{p_e(e_t, \mathcal{I}_t)}{p_t} e_t \sigma_t^e + \frac{p_{\mathcal{I}}(e_t, \mathcal{I}_t)}{p_t} \sigma \tag{59}$$

Market Clearing and Returns. Following similar steps as above, the true housing return process is given by:

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

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. Solving for the true volatility of e_t:

$$\begin{aligned} \sigma_t^e &= \alpha_t^h \sigma_t^h + \alpha_t^k \sigma_t^k - \sigma \\ &= \alpha_t^h(\sigma + \sigma_t^p) + \alpha_t^k(\sigma + \sigma_t^q) - \sigma \\ &= \frac{K_t}{E_t}\left[p_t(\sigma + \sigma_t^p) + q_t(\sigma + \sigma_t^q)\right] - \sigma, \text{ using (22) and (23)} \\ &= \frac{K_t}{E_t}\left[p_t(\sigma + \frac{p_e(e_t,\mathcal{I}_t)}{p_t}e_t\sigma_t^e + \frac{p_\mathcal{I}(e_t,\mathcal{I}_t)}{p_t}\sigma) + q_t(\sigma + \frac{q_e(e_t,\mathcal{I}_t)}{q_t}e_t\sigma_t^e + \frac{q_\mathcal{I}(e_t,\mathcal{I}_t)}{q_t}\sigma)\right] - \sigma \\ &= \frac{K_t}{E_t}\left[(p_t + q_t + p_\mathcal{I}(e_t,\mathcal{I}_t) + q_\mathcal{I}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma + (p_e(e_t,\mathcal{I}_t) + q_e(e_t,\mathcal{I}_t))e_t\sigma_t^e\right] \end{aligned}$$

$$e_t\sigma_t^e\left[\frac{1}{e_t} - \frac{K_t}{E_t}p_e(e_t,\mathcal{I}_t) - \frac{K_t}{E_t}q_e(e_t,\mathcal{I}_t)\right] = \frac{K_t}{E_t}(p_t + q_t + p_\mathcal{I}(e_t,\mathcal{I}_t) + q_\mathcal{I}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma$$

$$\begin{aligned} e_t\sigma_t^e &= \frac{K_t}{E_t}\frac{(p_t + q_t + p_\mathcal{I}(e_t,\mathcal{I}_t) + q_\mathcal{I}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma}{\frac{1}{e_t} - \frac{K_t}{E_t}p_e(e_t,\mathcal{I}_t) - \frac{K_t}{E_t}q_e(e_t,\mathcal{I}_t)} \\ e_t\sigma_t^e &= \frac{(p_t + q_t + \ p_\mathcal{I}(e_t,\mathcal{I}_t) + q_\mathcal{I}(e_t,\mathcal{I}_t) - \frac{E_t}{K_t})\sigma}{\frac{E_t}{K_t}\frac{1}{e_t} - p_e(e_t,\mathcal{I}_t) - q_e(e_t,\mathcal{I}_t)}. \end{aligned} \tag{60}$$

C.3 Verifying "Equity Member" Portfolio Choice

The main text assumes that the "equity member" will invest the maximal amount into the equity of the financial sector. This assumption must be verified ex-post given the resulting equilibrium.

The household maximizes the value function in equation (9), where $C_t = (c_t^y)^{1-\phi}(c_t^h)^\phi = (c_t^y)^{1-\phi}$, since $c^h = 1$ in equilibrium. From (9), the household accrues utility flow $\frac{(c_t^y)^{(1-\phi)(1-\gamma_h)}}{1-\gamma_h}$. Multiplying the utility function by $\frac{1}{1-\phi}$ (a positive affine transformation) shows that in equilibrium the household can be represented with power utility preferences, using relative risk aversion coefficient $\zeta = 1 - (1-\phi)(1-\gamma_h)$. Market

clearing gives $c_t^y = AK_t - \Phi(i_t, K_t) = \left(A - \delta - \frac{q_t - 1}{\xi} - \frac{(q_t-1)^2}{2\xi}\right) K_t$.

Let a be an arbitrary asset with perceived mean return $\widehat{\mu^a}$ and volatility $\widehat{\sigma^a}$. In equilibrium, CRRA utility implies:

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

where $\widehat{\sigma_t^{cy}}$ is the perceived volatility of $\frac{dc_t^y}{c_t^y}$ (a formula is provided in Appendix C). Thus, the household demands a perceived Sharpe ratio of $\zeta\widehat{\sigma_t^{cy}}$. Note that this equation need not hold with equality when there are portfolio restrictions placed on the household.

An investment in intermediary equity earns a perceived risk premium of $\alpha^k\widehat{\pi^k} + \alpha^h\widehat{\pi^h}$, with a perceived risk of $\alpha^k\widehat{\sigma^k} + \alpha^h\widehat{\sigma^h}$. From equation (20), intermediaries demand a risk premium of:

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

For the portfolio as a whole, this implies:

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

The equity member will invest all wealth in intermediary equity whenever the perceived Sharpe ratio on this investment is weakly greater than $\zeta\widehat{\sigma_t^{cy}}$:

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

This condition is verified numerically over the entire state space.

D Numerical Methods

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

$$\widehat{d\mathcal{S}_t} = \begin{bmatrix} \widehat{de_t} \\ d\mathcal{I}_t^S \end{bmatrix} = \begin{bmatrix} e_t \widehat{\mu_t^e} \\ -\kappa \mathcal{I}_t \end{bmatrix} dt + \begin{bmatrix} e_t \widehat{\sigma_t^e} \\ 0 \end{bmatrix} dZ_t, \tag{61}$$

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

$$d\mathcal{S}_t = \begin{bmatrix} de_t \\ d\mathcal{I}_t \end{bmatrix} = \begin{bmatrix} e_t \mu_t^e \\ -\kappa \mathcal{I}_t \end{bmatrix} dt + \begin{bmatrix} e_t \sigma_t^e \\ \sigma \end{bmatrix} dZ_t. \tag{62}$$

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

To simplify notation, let $\mathcal{A}$ denote the infinitesimal generator of $\mathcal{S}_t$. Let $\widehat{\mathcal{A}}$ denote the infinitesimal generator of $\widehat{\mathcal{S}}_t$.

D.1 Solving for Price Functions

As shown in Appendix C, price functions $q(e, \mathcal{I})$ and $p(e, \mathcal{I})$ compose a system of second-order PDEs:

$$\begin{aligned} q_t \widehat{\mu_t^q} &= q_e(e_t, \mathcal{I}_t) e_t \widehat{\mu_t^e} - q_{\mathcal{I}}(e_t, \mathcal{I}_t)\kappa \mathcal{I}_t + \frac{1}{2} q_{ee}(e_t, \mathcal{I}_t)(e_t \widehat{\sigma_t^e})^2 \qquad (63) \\ &= \widehat{\mathcal{A}} q_t \\ p_t \widehat{\mu_t^p} &= p_e(e_t, \mathcal{I}_t) e_t \widehat{\mu_t^e} - p_{\mathcal{I}}(e_t, \mathcal{I}_t)\kappa \mathcal{I}_t + \frac{1}{2} p_{ee}(e_t, \mathcal{I}_t)(e_t \widehat{\sigma_t^e})^2 \qquad (64) \\ &= \widehat{\mathcal{A}} p_t \end{aligned}$$

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. I assume knowledge of these methods 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 $\mathcal{I}$.[52] 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 price functions $q^0_{i,j}$ and $p^0_{i,j}$ at each grid point $\{i, j\}$.

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 (50) to solve for $\widehat{\sigma^e_{i,j}}$, (48) to solve for $\widehat{\mu^q_{i,j}}$, (49) to solve for $\widehat{\mu^p_{i,j}}$, and (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 $\widehat{\sigma^e_{i,j}}$.[53] Note that $\widehat{\mathbf{A}}$ features a reflecting barrier at $\underline{e}$.[54]

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

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

[52] The grid over e is non-uniform.

[53] "$n-1$" notation is used because $\widehat{\mathbf{A}}$ is constructed using the price functions from iteration $n-1$.

[54] Implementation details are in Achdou et al. (2017).

Similarly,[55]

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

These equations define $q^n_{i,j}$ and $p^n_{i,j}$ 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 should not be set too large.

4. If price functions have converged within a pre-specified tolerance, stop. If not, go to step 2 and repeat.

Once the algorithm has 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 $\mathcal{I}_t$. The realized evolution of price functions p_t and q_t can also be derived (similar to step 2).

D.2 Kolmogorov Equations

Kolmogorov Forward Equation. Readers should refer to the numerical appendix of Achdou et al. (2017) for details. Let $g_t(e, \mathcal{I})$ denote a probability density function over e and $\mathcal{I}$. The Kolmogorov forward equation (KF) gives $\frac{\partial}{\partial t}g_t$. A stationary distribution is a distribution $\bar{g}$ that solves $\frac{\partial}{\partial t}\bar{g} = 0$.

A benefit of finite-difference methods is that the KF equation comes "for free." Let $g^t_{i,j}$ denote a discretized distribution over e and $\mathcal{I}$ at time t. The perceived evolution of $g^t_{i,j}$ is given by:

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

[55] 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 B.4 for details). In the numerical solution for p, discretized matrix $\widehat{\mathbf{A}}$ is amended to ensure that $p^n_{1,j} = p^n_{2,j} - \frac{p^{n-1}_{1,j}\beta}{1+\underline{e}\beta} \times \Delta e$, where Δe denotes the grid increment in the e-dimension.

The true evolution of $g_{i,j}^t$ is given by:

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

Kolmogorov Backward Equation. The Kolmogorov backward equation (KB) is used to derive the hitting probabilities in Figure 4.

Let $\bar{e}_{crisis}(\mathcal{I})$ denote the upper boundary of the crisis region for sentiment level $\mathcal{I}$. For $s \in [0, T]$, let $f(e_s, \mathcal{I}_s, s)$ denote the true probability that the economy currently at $\{e_s, \mathcal{I}_s\}$ enters the crisis region between time s and T. Similarly, let $\widehat{f}(e_s, \mathcal{I}_s, s)$ denote the perceived probability. The goal is to solve for $f(e, \mathcal{I}, 0)$ and $\widehat{f}(e, \mathcal{I}, 0)$. Conditional probability function $f(e, \mathcal{I}, s)$ is the solution to:

$$\begin{aligned} 0 &= \frac{\partial f(e, \mathcal{I}, s)}{\partial s} + \mathcal{A} f(e, \mathcal{I}, s), \text{ subject to boundary conditions} \\ &(i)\ f(e, \mathcal{I}, s) = 1 \text{ if } e \leq \bar{e}_{crisis}(\mathcal{I}), \text{ and} \\ &(ii)\ f(e, \mathcal{I}, T) = 0 \text{ if } e > \bar{e}_{crisis}(\mathcal{I}) \end{aligned}$$

This result is stated without proof. Informally, the differential equation $0 = \frac{\partial f(e,\mathcal{I},s)}{\partial s} + \mathcal{A} f(e, \mathcal{I}, s)$ arises from applying Itô's lemma to $f(e, \mathcal{I}, s)$ 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: $f(e_s, \mathcal{I}_s, s) = \mathbb{E}_s[f(e_{s+dt}, \mathcal{I}_{s+dt}, s + dt)]$ and therefore $\mathbb{E}_s[df(e_s, \mathcal{I}_s, s)] = 0$. The first boundary condition sets f equal to 1 whenever the crisis region is hit. The second boundary condition is a terminal condition which assigns $f(e, \mathcal{I}, T) = 0$ if the crisis region is not hit at terminal period T. f is solved backwards from this terminal condition.

Using the perceived generator $\widehat{\mathcal{A}}$, function $\widehat{f}(e, \mathcal{I}, s)$ is the solution to:

$$0 = \frac{\partial \widehat{f}(e, \mathcal{I}, s)}{\partial s} + \widehat{\mathcal{A}}\widehat{f}(e, \mathcal{I}, s), \text{ subject to boundary conditions}$$
$$(i)\ \widehat{f}(e, \mathcal{I}, s) = 1 \text{ if } e \leq \bar{e}_{crisis}(\mathcal{I}), \text{ and}$$
$$(ii)\ \widehat{f}(e, \mathcal{I}, T) = 0 \text{ if } e > \bar{e}_{crisis}(\mathcal{I})$$

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.

D.3 Feynman-Kac Equation

The Feynman-Kac equation is used to calculate the expected investment rate profile in Appendix Figure 9. A straightforward application of the formula implies that the conditional expectation:

$$u(e, \mathcal{I}, 0) = \mathbb{E}\left[i(e_\tau, \mathcal{I}_\tau) \mid e_0 = e, \mathcal{I}_0 = \mathcal{I}\right]$$

satisfies a partial differential equation:

$$0 = \frac{\partial u(e, \mathcal{I}, s)}{\partial s} + \mathcal{A}u(e, \mathcal{I}, s),$$

subject to the terminal condition that $u(e, \mathcal{I}, \tau) = i(e, \mathcal{I})$.

Numerically, the discretized version of $\mathcal{A}$ was already generated when solving for price functions. Finite-difference methods can then be used to solve backward for $u(e, \mathcal{I}, 0)$ starting from the terminal condition.

E Diagnostic Expectations: Additional Details

E.1 Expectations of an AR(1): Comparison to Bordalo et al. (2018a)

Here I show how the discrete-time analogue of this paper's formulation of diagnostic expectations relates to 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, information measure $\mathcal{I}_t$ is given by:

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

where $\mathcal{K}$ is the discount factor governing the speed of information decay.[56] The background context is:

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

As the agent simulates the economy forward from period t:

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

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

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

[56] If the discrete-time model is written with a period frequency of Δ years, then $\mathcal{K} = e^{-\kappa\Delta}$.

Expectations of $\omega_{t+\tau}$. In this discrete-time framework, I start by expressing ω_t in changes:

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

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

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

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

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

Generally, in this AR(1) context the diagnostic agent will have the following perceptions about the distribution of $\omega_{t+\tau}$:

$$
\begin{aligned}
h^{\theta}(\omega_{t+1}|\omega_t) &\sim \mathcal{N}\left(b\omega_t + \theta\mathcal{I}_t, \sigma^2\right) \\
h^{\theta}(\omega_{t+2}|\omega_t) &\sim \mathcal{N}\left(b^2\omega_t + b\theta\mathcal{I}_t + \mathcal{K}\theta\mathcal{I}_t, \sigma^2 + b^2\sigma^2\right) \\
&\vdots \\
h^{\theta}(\omega_{t+\tau}|\omega_t) &\sim \mathcal{N}\left(b^{\tau}\omega_t + \theta\mathcal{I}_t\left(\sum_{i=0}^{\tau-1} b^i\mathcal{K}^{\tau-1-i}\right),\ \sigma^2\sum_{i=0}^{\tau-1} b^{2i}\right), \text{ or equivalently} \\
h^{\theta}(\omega_{t+\tau}|\omega_t) &\sim \mathcal{N}\left(b^{\tau}\omega_t + \theta\mathcal{I}_t\left(\frac{b^{\tau}-\mathcal{K}^{\tau}}{b-\mathcal{K}}\right),\ \sigma^2\sum_{i=0}^{\tau-1} b^{2i}\right)
\end{aligned}
$$

BGS Equivalence. To reproduce the AR(1) framework of BGS, take $\mathcal{K} \to 0$. In the limit, $G_t^- = \omega_t - \epsilon_t = b\omega_{t-1}$. In other words, background context G_t^- does not incorporate the most recent shock, but fully incorporates all shocks of further lags. Looking forward, $G'^-_{t+\tau} = \omega'_{t+\tau}$ 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 $\omega_{t+\tau}$, 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 $\mathcal{N}(b\left(b\omega_t + \theta\epsilon_t\right), \sigma^2 + b^2\sigma^2)$, the distribution at time $t+3$ is $\mathcal{N}(b^2\left(b\omega_t + \theta\epsilon_t\right), \sigma^2 + b^2\sigma^2 + b^4\sigma^2)$, etc. Using the formula above as $\mathcal{K} \to 0$, the distribution at time $t+\tau$ is $\mathcal{N}(b^{\tau}\omega_t + b^{\tau-1}\theta\epsilon_t,\ \sigma^2\sum_{i=0}^{\tau-1} b^{2i})$. This is almost identical to BGS, with the one difference being that parameter θ here is their parameter θ/b.

Moving from Discrete to Continuous Time. To pass from discrete to continuous time, one needs to set $\mathcal{K} > 0$. When $\mathcal{K} \to 0$ only the most recent shock matters. But, the concept of "only the most recent shock" varies with the length of the period.

This is especially important when passing to continuous time, where the most recent shock is an instantaneous Brownian increment. Thus, in this paper I instead adopt a sentiment function $\mathcal{I}_t$ that is based on multiple lags of past shocks (i.e., $\mathcal{K} > 0$).

E.2 Extensions of the Baseline Model

Objective versus Subjective Shocks. Information measure $\mathcal{I}_t \equiv \int_0^t e^{-\kappa(t-s)}\sigma dZ_s$ is based on objective 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 also tractable to define $\mathcal{I}_t$ based on subjective capital quality shocks. An agent with a bias of $\theta\mathcal{I}_t$ will perceive a subjective shock at time t of $\sigma\widehat{dZ}_t = -\theta\mathcal{I}_t dt + \sigma dZ_t$. In this case, one can define the new information measure as

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

This measure of subjective new information evolves according to:

$$\begin{aligned} d\mathcal{I}_t &= -\kappa\mathcal{I}_t dt + \sigma\widehat{dZ}_t \\ &= -(\kappa+\theta)\mathcal{I}_t dt + \sigma dZ_t. \end{aligned} \tag{65}$$

Note that the baseline measure of information in equation (3) evolves according to $d\mathcal{I}_t = -\kappa\mathcal{I}_t dt + \sigma dZ_t$. Comparing this to equation (65), $\mathcal{I}_t$ decays more quickly when defined in terms of subjective shocks. This is because an overoptimistic agent will interpret incoming shocks with a negative bias (and vice-versa for an overpessimistic agent), leading to a faster unwinding of sentiment.

Multiple Frequencies of Extrapolation. The main text assumes that sentiment evolves at a single frequency, with shocks fading from $\mathcal{I}_t$ at rate κ. However, it is likely more psychologically realistic to have sentiment operate over multiple frequencies. The empirical literature on extrapolative expectations has documented extrapolation

over a variety of horizons. At a low frequency, Malmendier and Nagel (2011) find that long-term risk attitudes are shaped by lifetime macroeconomic experiences. Alternatively, Greenwood and Shleifer (2014) find that stock market expectations depend strongly on returns experienced in the past quarter.

The expectations model can 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 be defined as:

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

Parameters θ^L and θ^H determine the relative strength of low-frequency and high-frequency sentiment.

The downside of including multiple sentiment frequencies is that each frequency requires its own state variable. The main text uses a single frequency for parsimony.

E.3 Properties of Sentiment $\mathcal{I}_t$

Sentiment parameter $\mathcal{I}_t$ is an Ornstein-Uhlenbeck (OU) process. This section sketches some useful mathematical properties of OU processes. A textbook treatment is provided by Karatzas and Shreve (1998).

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

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

The conditional distribution of $\mathcal{I}_{t+\tau}$ given $\mathcal{I}_t$ is characterized as follows:

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

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

The unconditional distribution of the OU process is obtained by taking $\tau \to \infty$. The unconditional distribution of $\mathcal{I}$ is Gaussian with mean 0 and variance $\frac{\sigma^2}{2\kappa}$.

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 characterized analytically in the special case where the threshold is the OU process' mean. So, from any initial level of sentiment $\mathcal{I}_t$, the time $t + \tau$ at which the OU process first returns to 0 has a known distribution, with a density given by:

$$h(\tau|\mathcal{I}_t) = \frac{|\mathcal{I}_t|}{\sigma\sqrt{2\pi}} \left(\frac{\kappa}{\sinh(\kappa\tau)}\right)^{\frac{3}{2}} \exp\left(-\frac{\kappa\mathcal{I}_t^2 e^{-\kappa\tau}}{2\sigma^2 \sinh(\kappa\tau)} + \frac{\kappa\tau}{2}\right). \tag{66}$$

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

Importance of κ for Sentiment Persistence. Let $\mathcal{I}_t^{SD} = \frac{\mathcal{I}_t}{\left(\frac{\sigma}{\sqrt{2\kappa}}\right)}$ 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 shows 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 (66) can be rewritten as:

$$h(\tau|\mathcal{I}_t^{SD}) = \frac{\kappa|\mathcal{I}_t^{SD}|}{2\sqrt{\pi}} \left(\frac{1}{\sinh(\kappa\tau)}\right)^{\frac{3}{2}} \exp\left(-\frac{(\mathcal{I}_t^{SD})^2 e^{-\kappa\tau}}{4\sinh(\kappa\tau)} + \frac{\kappa\tau}{2}\right). \tag{67}$$

Equation (67) 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 κ, but is independent of σ.

Ornstein-Uhlenbeck Figures: Baseline Calibration.

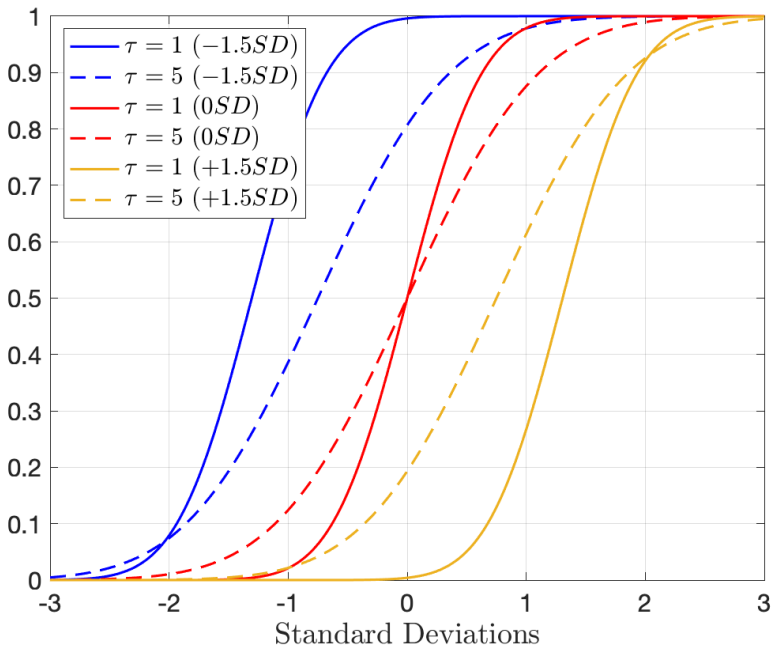


Figure 11: **Conditional CDF**. CDF of $\mathcal{I}_{t+\tau}^{SD}|\mathcal{I}_t^{SD} \in \{0SD, \pm 1.5SD\}$ for $\tau \in \{1, 5\}$.

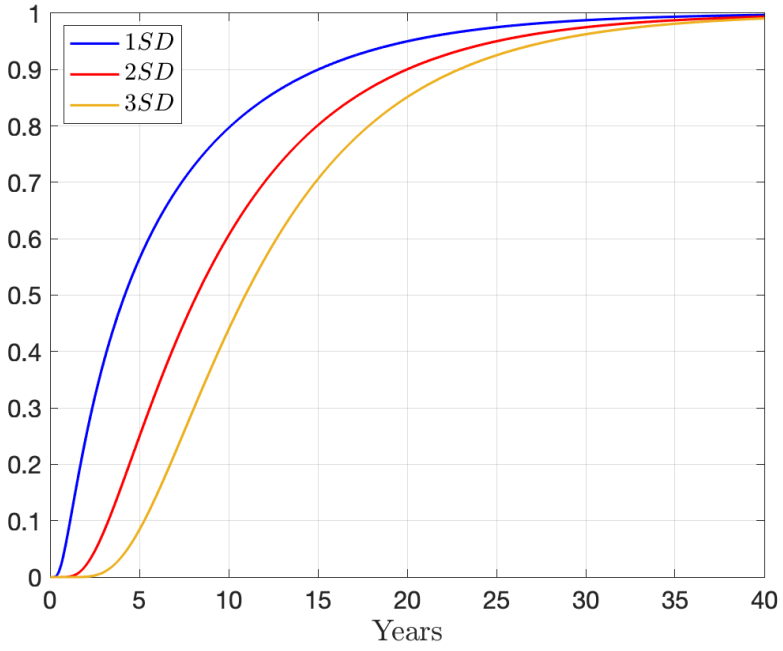


Figure 12: **1st hitting times**. Hitting time CDF for $\mathcal{I}_t \in \{\pm 1SD, \pm 2SD, \pm 3SD\}$.

Appendix References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *NBER w23732*, 2017.

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.

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.

_ , _ , Rafael La Porta, and Andrei Shleifer, "Diagnostic Expectations and Stock Returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

_ , _ , Yueran Ma, and Andrei Shleifer, "Over-reaction in Macroeconomic Expectations," *NBER w24932*, 2018.

Brunnermeier, Markus K. and Yuliy Sannikov, "A Macroeconomic Model with a Financial Sector," *American Economic Review*, February 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.

Campbell, John Y., *Financial Decisions and Markets: A Course in Asset Pricing*, Princeton University Press, 2017.

d'Arienzo, Daniele, "Maturity Increasing Over-reaction and Bond Market Puzzles," *Mimeo*, 2020.

Davis, Morris A. and Stijn Van Nieuwerburgh, "Housing, Finance, and the Macroeconomy," *Handbook of Regional and Urban Economics*, 2015, *5*, 753–811.

Frankel, Marvin, "The Production Function in Allocation and Growth: A Synthesis," *The American Economic Review*, 1962, *52* (5), 996–1022.

Gennaioli, Nicola and Andrei Shleifer, "What Comes to Mind," *The Quarterly Journal of Economics*, 2010, *125* (4), 1399–1433.

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

He, Zhiguo, Bryan Kelly, and Asaf Manela, "Intermediary Asset Pricing: New Evidence from Many Asset Classes," *Journal of Financial Economics*, 2017, *126* (1), 1–35.

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

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

Rabin, Matthew, "An Approach to Incorporating Psychology into Economics," *American Economic Review*, 2013, *103* (3), 617–22.