Reflexivity in Credit Markets*

Robin Greenwood  Samuel G. Hanson
Harvard University and NBER  Harvard University and NBER

Lawrence J. Jin
California Institute of Technology

June 2021

Abstract

Reflexivity is the idea that investors’ biased beliefs affect market outcomes, and that market outcomes in turn affect investors’ biased beliefs. We develop a behavioral model of the credit cycle featuring such a two-way feedback loop. In our model, investors form beliefs about firms’ creditworthiness, in part, by extrapolating past default rates. Investor beliefs influence firms’ actual creditworthiness because firms that can refinance maturing debt on favorable terms are less likely to default in the short run—even if fundamentals do not justify investors’ generosity. Our model is able to match many features of credit booms and busts, including the imperfect synchronization of credit cycles with the real economy, the negative relationship between past credit growth and the future returns on risky bonds, and “calm before the storm” periods in which firm fundamentals have deteriorated but the credit market has not yet turned.

* A previous version of this paper circulated under the title “A Model of Credit Market Sentiment.” We are grateful to Nicholas Barberis, Jonathan Ingersoll, Gordon Liao, Yueran Ma, Raghuram Rajan, George Soros, Andrei Shleifer, Jeremy Stein, Lawrence Summers, Adi Sunderam, Yao Zeng, and seminar participants at Brandeis University, Columbia University, the Federal Reserve Bank of San Francisco, the London Business School, the London School of Economics, Oxford University, the University of Massachusetts Amherst, the University of Michigan, the University of North Carolina at Chapel Hill, the University of Washington, the American Economic Association Annual Meetings, the FIRN Annual Asset Pricing Workshop, the LA Finance Day Conference, and the NBER Risks of Financial Institutions Summer Institute for their helpful comments. Greenwood and Hanson gratefully acknowledge funding from the Division of Research at Harvard Business School. Outside activities and other relevant disclosures are provided on the authors’ websites at their host institutions.
1 Introduction

Over the past decade, researchers in finance and economics have documented a number of new facts about the credit cycle. High credit growth is associated with both a higher probability of a future financial crisis and lower GDP growth (Schularick and Taylor [2012], López-Salido, Stein, and Zakrajšek [2017], Mian, Sufi, and Verner [2017]). Other research has shown that credit market returns are predictable, suggesting a role for investor sentiment in driving the credit cycle. Greenwood and Hanson (2013) show that periods of elevated corporate credit growth and low average borrower credit quality forecast low returns to credit. In a large panel of countries, Baron and Xiong (2017) find that high bank credit growth forecasts low returns to bank stocks. Greenwood, Hanson, Shleifer and Sørensen (2021) show that the combination of large credit expansions and asset price booms predict financial crises.

An underappreciated feature of the credit cycle is how disconnected it can be from the stock market or the broader macroeconomy in the short run. In post-war U.S. history, credit expansions and contractions have often followed a similar pattern. Credit grows slowly as the economy emerges from a recession, picks up steam, but continues to expand even as the overall economy cools. For example, in the upswing preceding the 2008 financial crisis, GDP growth peaked in March 2005, but credit growth peaked two years later in March 2007, a period when credit spreads were near historical lows. Put simply, the credit cycle seems to have some life of its own at short horizons. However, these disconnects pose a challenge for many well-known models of the credit cycle—e.g., Bernanke and Gertler (1989), Holmström and Tirole (1997), Bernanke, Gertler, Gilchrist (1999)—and even for more recent behavioral models like Bordalo, Gennaioli, and Shleifer (2018). Although credit market frictions amplify business cycle fluctuations in these models, the business cycle and the credit cycle are essentially one and the same.

In this paper, we present a new behavioral model of the credit cycle that is consistent with much of the accumulating evidence on credit cycles, but also speaks to periods of disconnect between credit markets and the fundamentals of the economy. A key feature of our model is “reflexivity,” the idea that there is a feedback loop between investors’ biased beliefs and market outcomes. In finance, the idea of reflexivity is most prominently associated with the investor
George Soros, who argued that “distorted views can influence the situation to which they relate because false views lead to inappropriate actions.” *(Financial Times, October 26, 2009).*

In credit markets, reflexivity arises because investors who overestimate the creditworthiness of a borrower refinance maturing debt on more favorable terms, thereby making the borrower less likely to default, at least in the short run.

In our baseline model, a representative firm invests in a sequence of one-period projects. Each project requires an upfront investment of capital, which the firm finances using short-term debt that it must refinance each period. Projects generate a random cash flow that varies exogenously according to the state of the economy. Debt financing is provided by extrapolative investors whose beliefs condition only on the firm’s recent repayment history. Following periods of low defaults, investors believe that debt is safe, and hence refinance maturing debt on attractive terms.

Because investors hold extrapolative beliefs based on defaults—which are endogenously determined in the model—and not on the exogenously given cash flow fundamentals, this leads to a dynamic feedback loop between investor beliefs and defaults. Current investor beliefs affect future defaults via the terms on which investors are willing to refinance debt today. Figure 1 illustrates this feedback loop. During credit booms, default rates are low, so investors believe that future default rates will continue to be low. In the near term, these beliefs can be self-fulfilling: the perception of low future defaults leads to elevated bond prices, which in turn, makes it easier for firms to refinance their maturing debt. Holding constant firms’ cash flows, cheaper debt financing leads to slower debt accumulation and a near-term decline in future defaults, which further reinforces investor beliefs. If cash flow fundamentals deteriorate, the backward-looking nature of investors’ beliefs may allow firms to skate by for some amount of time, a phenomenon that we refer to as the “calm before the storm.” Eventually, the reality of poor cash flow fundamentals catches up with firms, and defaults escalate.

Conversely, suppose that the economy has just been through a wave of defaults. Since investors over-extrapolate these recent outcomes, they believe that the likelihood of future defaults is high. Investor beliefs turn out to be partially self-fulfilling in the short run: their bearish beliefs

---

1 See Soros (1987, 2013) for an extensive discussion of reflexivity.
make it more expensive for firms to refinance existing debt, leading to an increased probability of default in the short run. In some circumstances, this can lead to default spirals in which a default leads to further investor pessimism and an extended spell of further defaults, much like what has been observed in instances of sovereign debt restructuring (Das, Papaioannou, and Trebesch [2012]).

In our model, transitions between credit booms and credit busts are ultimately caused by changes in cash flow fundamentals. However, because investors extrapolate past defaults and not cash flows, these transitions are not fully synchronized with changes in fundamentals, and can be highly path-dependent. For example, our model generates “calm before the storm” periods in which the fundamentals of the economy have turned, but credit markets are still buoyant. Such episodes are consistent with Krishnamurthy and Muir (2020), who show that credit spreads are typically too low in the years preceding financial crises.

What happens in the model if investors exogenously become more bullish? We show that such a shock to beliefs can be self perpetuating: firms are able to roll over debt at more attractive rates, which in turn makes default less likely in the near term. Lending to firms now appears to be even safer, leading investors to become even more optimistic, further reducing credit spreads. There is a limit to this self-perpetuation, however, because the firm can eventually become over-leveraged, triggering a default. Although the feedback loop between biased expectations and outcomes is always present, there are times when it is stronger. We use the model to characterize the conditions under which changes in investors’ biased expectations have the most powerful impact on market outcomes. This mechanism may help make sense of the 2020-2021 period in US credit markets, where improvements in investor sentiment may have saved the economy from a wave of defaults coming from poor fundamentals, as argued by Hanson, Stein, Sunderam, and Zwick (2020).

The model matches a number of facts about credit cycles that researchers have documented in recent years. First, rapid credit growth appears to be quite useful for predicting future financial crises and business cycle downturns (Schularick and Taylor [2012], Mian, Sufi, and Verner [2017], López-Salido, Stein, and Zakrajšek [2017]), a result that is consistent with our model because
outstanding credit will grow rapidly when sentiment is high but cash flow fundamentals are poor. Relatedly, economies that have experienced high credit growth are more fragile, in the sense that they are vulnerable to shocks (Krishnamurthy and Muir [2020]). Second, high credit growth predicts low future returns on risky bonds in a univariate forecasting regression (Greenwood and Hanson [2013], Baron and Xiong [2017], Muir [2019]), a result that obtains in our model because investors do not fully understand that when credit is growing rapidly they are often quickly heading towards a credit bust. In a multivariate forecasting specification, lower quality borrowing negatively predicts future returns and credit spreads positively predict future returns (Greenwood and Hanson [2013] and López-Salido, Stein, and Zakrajšek [2017]). In our model, credit spreads are typically too low just before the economy experiences a wave of defaults, consistent with the evidence in Krishnamurthy and Muir (2020). Third, when credit markets become highly overheated, our model generates negative conditional expected excess returns on risky bonds. This result, which is consistent with prior empirical evidence, is difficult to square with rational, risk-based models of credit cycles and motivates models like ours which prominently feature biased investor beliefs.

After developing our baseline model, we consider several extensions. First, we add a rational component to beliefs, such that investors hold a weighted average of extrapolative beliefs and forward-looking rational beliefs. While this extension adds complexity and introduces the possibility of multiple equilibria, it allows us to match the finding in Krishnamurthy and Muir (2020) that credit spreads spike on the eve of a crisis—a finding that our baseline model cannot explain. Adding this rational component sometimes increases the magnitude of reflexivity. For example, just before a crisis, the rational component of investor beliefs causues the bond price to fall. With a lower bond price, even small changes in irrational beliefs cause significant changes in the actual likelihood of future defaults.

The second extension unpacks extrapolative beliefs into two distinct components. Namely, our investors not only condition on incomplete information (that is, observing firm defaults but not firm fundamentals), but also use that information incorrectly. In the extension, we explore the path of debt, default, and prices, if investors could only observe past defaults, but used
this information in a Bayesian manner to form beliefs about firm fundamentals and debt levels. While such a model generates low credit spreads before a default, it counterfactually predicts that credit spreads remain low even after default.

The third extension addresses the question of firm maximization. The firms in our model simply follow a mechanical default rule that is a function of leverage. We consider a more elaborate model in which default is triggered by equity holders in the spirit of Leland (1994). This alternate setup is less analytically tractable than our main model, but delivers an interesting new prediction that appears to be borne out in the data, namely that leverage is procyclical, consistent with what Korajczyk and Levy (2003) find for constrained firms. This procyclicality of leverage can further exacerbate the “calm before the storm” phenomenon and the default spirals discussed above.

We describe briefly two additional extensions that consider the role of opportunistic debt issuance and multiple firms who face idiosyncratic cash flows, but leave the full exposition for the Appendix. The goal of these two extensions is to further improve the model’s realism.

Our paper has much in common with Austrian theories of the credit cycle, including Mises (1924) and Hayek (1925), as well as the accounts of booms, panics, and crashes by Minsky (1986) and Kindleberger (1978). More recently, the idea that investors may neglect tail risk in credit markets was developed theoretically by Greenwood and Hanson (2013), Gennaioli, Shleifer, and Vishny (2012, 2015), and Bordalo, Gennaioli, and Shleifer (2018). We also draw on growing evidence that investors extrapolate cash flows, past returns, or past crash occurrences (Barberis, Shleifer, and Vishny [1998], Greenwood and Shleifer [2014], Barberis, Greenwood, Jin, and Shleifer [2015, 2018], Jin [2015], Greenwood and Hanson [2015]). Most related here is Jin (2015), who presents a model in which investors’ perceptions of crash risk depend on recent experience. Bordalo, Gennaioli, and Shleifer (2018) provide a model of credit cycles in which extrapolative investor expectations play an important role and in which bond returns are predictable. Their model is similar to ours in several respects, but extrapolative expectations in their model are perfectly tied to cash flow fundamentals, rather than to endogenous credit
market outcomes.\(^2\) Also related is Krishnamurthy and Li (2020), who analyze to what degree a behavioral model of credit cycles, such as the one presented here, can quantitatively match the historical evidence.

We have emphasized that our model of the credit cycle hinges on a feedback loop between investor beliefs and credit market outcomes. To be sure, rational expectations models also feature two-way feedback between agents’ beliefs and market outcomes (Muth [1961]). In a rational expectations equilibrium, agents’ beliefs help determine market outcomes and agents’ beliefs are rationally consistent with those outcomes, giving rise to a fixed-point problem. By contrast, our model emphasizes the feedback between market outcomes and irrational beliefs. Furthermore, this feedback is dynamic in nature: investors’ biased beliefs influence future market outcomes, and past market outcomes shape investors’ subsequent biased beliefs. In this way, our paper fuses extrapolative expectations with one of the key ideas in the sovereign default literature: the rationally, self-fulfilling debt crisis (Calvo [1988] and Cole and Kehoe [2000]). In these forward-looking rational expectations models, if investors believe that a default is likely, they require high interest rates, increasing the debt burden and hence the true probability of default, thus validating the initial belief. By contrast, our behavioral model is capable of generating self-fulfilling dynamics—including both the “calm before the storm” and “default spiral” phenomena—even when beliefs are completely extrapolative and backward-looking.

In Section 2, we briefly summarize a number of stylized facts about the credit cycle, drawing on the papers cited above but also presenting some novel observations about the synchronicity of the credit cycle and the business cycle. Section 3 develops the baseline model and explains the two-way feedback mechanism that is at the heart of our model. In Section 4, we formally define reflexivity and explain its properties. We then discuss how the model can match a number of features of credit cycles that researchers have documented in recent years, such as the predictability of returns and low credit spreads before crises. Section 5 explores several extensions of the baseline model, and Section 6 concludes.

\(^2\)See also Coval, Pan, and Stafford (2014) who suggest that in derivatives markets, model misspecification only reveals itself in extreme circumstances, by which time it is too late. Bebchuk and Goldstein (2011) present a model in which self-fulfilling credit market freezes can arise because of interdependence between firms.
2 Motivating facts about the credit cycle

We begin by summarizing a set of stylized facts about credit cycles. The first four facts are drawn from previous work; the fifth is based on some new empirical work of our own.

Observation 1. Rapid credit growth predicts financial crises and business cycle downturns.

In a panel of 14 countries dating back to 1870, Schularick and Taylor (2012) show that rapid credit growth predicts financial crises. Schularick and Taylor (2012) interpret their evidence as suggesting that financial crises are episodes of “credit booms gone bust.” Mian, Sufi, and Verner (2017) show that rapid credit growth—and especially growth in household credit—predicts future declines in GDP growth in an panel of 30 countries from 1960 to 2012. López-Salido, Stein, and Zakrajšek (2017) show that frothy credit market conditions—proxied using declines in the credit quality of corporate borrowers and low credit spreads—predict low GDP growth in U.S. data from 1929 to 2015. López-Salido, Stein, and Zakrajšek (2017) attribute their findings to predictable reversals in credit market sentiment. Consistent with this view, using an international panel of 38 countries, Kirti (2018) shows that rapid credit growth that is accompanied by a deterioration in lending standards—i.e., by declining borrower credit quality—is associated with low future GDP growth. By contrast, when rapid credit growth is accompanied by stable lending standards, there is no such decline in future GDP growth.

More recently, Greenwood, Hanson, Shleifer and Sørensen (2021) show that the degree of crisis predictability rises significantly when large credit expansions are accompanied by asset price booms. That is, when credit growth is high and valuations are high (or correspondingly, credit spreads are low), the probability of a subsequent crisis is substantially elevated.

A corollary of Observation 1—i.e., that credit growth predicts financial crises—is that economies that have experienced high credit growth are more fragile. Krishnamurthy and Muir (2020) argue that the natural way to interpret Schularick and Taylor’s (2012) findings about credit growth and financial crises is that credit growth creates financial fragility. When a more leveraged economy is exogenously hit by a negative fundamental shock, such as a large decline in house prices, this results in a financial crisis. Moreover, Krishnamurthy and Muir (2020) find that credit spreads
spike on the eve of a financial crisis. Alternately, crises may be triggered by predictable reversals in credit market sentiment as argued by López-Salido, Stein, and Zakrajšek (2017). Consistent with this view, Krishnamurthty and Muir (2020) show that credit spreads are typically “too low” in the years preceding financial crises. The model we develop reflects both of these ideas: as leverage grows, the probability of a default becomes more and more sensitive to both changes in underlying fundamentals and to changes in biased investor beliefs.

**Observation 2.** Credit market overheating—signaled either by (i) a rapid growth in debt outstanding or (ii) by a decline in the credit quality of debt issuers set against the backdrop of (relatively) low credit spreads—predicts low future returns on risky bonds.

A growing literature has demonstrated that credit market overheating predicts low future returns on risky bonds. Greenwood and Hanson (2013) find that rapid growth in outstanding corporate credit is associated with low future returns on risky bonds in U.S. data. Muir (2019) finds the same pattern in an panel of 17 developed economics from 1870 to 2016. Adopting a similar intuition, Baron and Xiong (2017) show that bank credit expansion also predicts low bank equity returns—which are naturally tied to the returns on risky debt—in a panel of 20 developed economies from 1920 to 2012.

Greenwood and Hanson (2013) develop a more statistically powerful measure of credit market overheating based on the credit quality of corporate debt issuers. Their “high yield share” measure—the share of all corporate bond issuance in a given year that is from high-yield-rated firms—captures the intuition that when credit markets are overheated, low quality firms increase their borrowing to take advantage. Greenwood and Hanson (2013) show that declines in issuer credit quality predict low future corporate bond returns in a univariate sense. Furthermore, as emphasized by Greenwood and Hanson (2013) and López-Salido, Stein, and Zakrajšek (2017), issuer quality contains information about future bond returns beyond that contained in credit spreads. Specifically, in a multivariate regression specification, low-quality issuance negatively predicts future bond returns and credit spreads positively predict future returns.

Table 1 updates the data from Greenwood and Hanson (2013) and also considers a set of additional proxies for credit market overheating. The table shows annual return forecasting
regressions of the form:

\[ r_{t \rightarrow t+k}^{HY} = a + b \cdot Overheating_t + e_{t \rightarrow t+k}, \]  

(1)

where \( r_{t \rightarrow t+k}^{HY} \) denotes the log return on high yield bonds in excess of the log returns on like-maturity Treasuries over a \( k = 2 \)- or 3-year horizon beginning in year \( t \). Here \( Overheating_t \) is a proxy for credit market overheating, measured using data through the end of year \( t \). All of our data begin in 1983 and run through 2014.\(^3\)

Columns (1) and (5) show that the log high yield share (\( \log(HYS_t) \)) significantly predicts low future excess bond returns. A one standard deviation in \( \log(HYS_t) \) is associated with an 8.3 percentage point reduction in log excess bond returns over the next two years, and a 9.7 percentage point reduction over the next three years.

Columns (2) and (6) of Table 1 show that the same forecasting results hold when credit market overheating is measured using the growth in aggregate nonfinancial corporate credit outstanding, \( \text{Credit Growth}_t \). Aggregate nonfinancial corporate credit is the sum of nonfinancial corporate debt securities and loans from Table L103 of the Federal Reserve’s Financial Accounts of the U.S. A one standard deviation increase in \( \text{Credit Growth}_t \) forecasts a 7.4 percentage point reduction in excess bond returns over the next two years, and a 9.3 percentage point reduction over the next three years.

Table 1 shows results for two additional measures of credit market overheating. The first, \( \text{Easy Credit}_t \), is based on the Federal Reserve’s Senior Loan Officer Opinion Survey (SLOOS), and the second, \(-1 \times EBP_t\), is negative one times the Excess Bond Premium (\( EBP_t \)) from Gilchrist and Zakrajšek (2012).\(^4,5\) Table 1 shows that both of these additional measures of credit

---

\(^3\)For results over different time horizons and with additional controls, see Greenwood and Hanson (2013) who compute alternate proxies for issuer quality that extend back as far as 1926.

\(^4\)Every quarter, the Federal Reserve surveys senior loan officers of major domestic banks concerning their lending standards. The officers report whether they have eased or tightened lending standards in the past quarter. We construct a measure of credit market overheating, \( \text{Easy Credit}_t \), by taking the three-year average percentage of banks that have reported easing credit standards to firms in any given quarter. The idea behind this averaging procedure is that we want to capture the level of bankers’ beliefs about future creditworthiness, whereas the quarterly survey tracks changes from the previous quarter. The SLOOS begins in the first quarter of 1990, so this measure of overheating begins in December 1992. \( \text{Easy Credit}_t \) is 55% correlated with the high yield share (\( HYS_t \)) and 68% correlated with \( \text{Credit Growth}_t \).

\(^5\)Gilchrist and Zakrajšek’s (2012) \( EBP_t \) variable is a measure of average corporate credit spreads after deduct-
market overheating forecast low future returns on corporate bonds. To summarize, Table 1
confirms that periods of credit market overheating—periods featuring low credit quality debt
issuance, rapid growth in outstanding credit, loose credit standards, and tight credit spreads—
are followed by low subsequent returns on risky corporate bonds.

Observation 3. Significant credit market overheating is associated with negative expected
future returns on risky bonds.

The fact that corporate bond returns are predictable does not imply that corporate bonds
are occasionally mispriced. For instance, if the rationally-required return on risky corporate
bonds fluctuates over time—e.g., due to movements in investor risk aversion (Campbell and
Cochrane [1999]) or in the quantity of aggregate risk (Bansal and Yaron [2004], Gabaix [2012],
Wachter [2013])—then the level of credit spreads might forecast future returns on corporate
bonds. And, combining such fluctuations in rationally-required returns with the neoclassical
q-theory of investment, one might expect recent credit growth and declines in debt issuer quality
to forecast low future returns on risky corporate bonds (Greenwood and Hanson [2013], Santos
and Veronesi [2018], Gomes, Grotteria, and Wachter [2019]).

However, Greenwood and Hanson (2013) and Baron and Xiong (2017) present evidence that
conditional expected excess returns on risky corporate bonds and bank stocks become reliably
negative when credit markets appear to be significantly overheated—i.e., when many low quality
borrowers are able to obtain credit and when credit growth is rapid. Furthermore, these same
authors find that future risk is high when credit markets appear to be most overheated (see Muir
[2019] for further evidence on this point). These negative expected returns and the negative
conditional relationship between expected future risk and return are quite difficult to square
with rational risk-based models—even rational models with intermediation frictions—and are
powerful motivations for the behavioral approach we adopt in this paper.6

Observation 4. Variables that forecast returns on risky bonds often do not forecast returns on

6In models with intermediation frictions, changes in the health of intermediary balance sheets and the resulting
shifts in risk appetite play an important role in determining asset prices. See, for example, He and Krishnamurthy
equities, and vice versa. Moreover, episodes of credit market overheating tend to follow periods of tranquility in credit markets, namely periods when defaults are low and when the returns on risky bonds are high.

What outcomes are credit-market investors over-extrapolating? One view is that investors over-extrapolate some underlying set of economic fundamentals—e.g., firm cash flows or the state of broader macroeconomy. This view leads to behavioral version of the $q$-theory of investment (Greenwood and Hanson [2015], Gennaioli, Ma, and Shleifer [2016], Bordalo, Gennaioli, and Shleifer [2018]) and generally suggests that equity-market sentiment and credit market sentiment should be tightly linked over time. However, in the data, many measures that predict credit returns are not strong predictors of equity returns and vice versa (Greenwood and Hanson [2013] and Ma [2019]). This disconnect between equity and credit market sentiment recommends a more nuanced behavioral view in which equity and credit markets are partially segmented and investors in each market extrapolate past market-specific outcomes.

Consistent with the idea that credit market investors tend to over-extrapolate recent credit market outcomes, Greenwood and Hanson (2013) show that past defaults and credit returns play a dominant role in shaping credit market sentiment. They find that debt issuer quality tends to deteriorate following periods with low realized corporate defaults and high realized returns on risky corporate bonds. However, after controlling for these recent credit market outcomes, recent equity returns and macro variables have relatively little impact on debt issuer quality. These findings motivate our model where credit investors extrapolate past bond defaults, which themselves are not perfectly tied to firm fundamentals.\(^7\)

Table 2 presents additional evidence that periods of credit market overheating follow times when corporate defaults are low. We estimate time-series regressions of the form:

$$\text{Overheating}_t = a + b \cdot \text{Def}_t + c \cdot \text{Def}_{t-1} + e_t,$$  \hspace{1cm} (2)

\(^7\)Similarly, Greenwood and Shleifer (2014) show that past equity returns play an outsized role in shaping equity market sentiment, motivating the model in Barberis, Greenwood, Jin, and Shleifer (2015) where equity investors extrapolate past equity returns (as opposed to firm fundamentals).
where $Def_t$ denotes the default rate on high yield bonds in year $t$. We estimate this regression using the same four measures of credit market overheating from Table 1. Table 2 shows that there is a strong negative relationship between recent default rates and current credit market overheating. Some measures of overheating ($\log(HYS_t)$ and $-1 \times EBP_t$) are more highly correlated with most recent default rates, while others are also strongly correlated with lagged default rates ($Credit\ Growth_t$ and $Easy\ Credit_t$).

**Observation 5.** The credit cycle and the business cycle can be quite disconnected in the short run.

Consistent with the market-specific extrapolation view discussed above, the credit cycle can become quite disconnected from both the broader business cycle as well as equity market conditions in the short run.

Figure 2 plots the annual growth rate of U.S. GDP alongside the annual growth rate of outstanding debt at nonfinancial corporations, both expressed in real terms. In the upswing proceeding the 2008 financial crisis, GDP growth peaked in March 2005, but credit growth peaked two years later. This pattern of credit expansion at the end of an economic expansion is also apparent in the late 1990s, with credit growth rising only at the end of the business cycle. During downturns, the economy often recovers well before credit growth returns to normal rates. In the post-2008 economic recovery, real credit growth first reached 3% in 2013, several years after the economy began its recovery. Overall, the correlation between credit growth and GDP growth is only 43%.

Figure 3 further illustrates the disconnect between the credit cycle and the business cycle in U.S. data. Here, we provide additional perspective on the lack of synchronicity between the credit cycle and the business cycle. In particular, we show that credit growth tends to increase towards the end of a business cycle boom. In Panel A of Figure 3, we plot real GDP growth from trough to peak of the business cycle, by business cycle expansion quarter. As can be seen, GDP growth is high in the beginning of business cycle expansions, but after quarter five, it stabilizes and if anything, declines slightly in later quarters. By contrast, Panel B shows credit growth over the same periods. As the figure makes clear, credit expansion is particularly high in the
latter part of the business cycle.

3 A model of credit market sentiment

In this section, we develop our baseline model. We first describe firm borrowing behavior and then explain investor beliefs, collecting several preliminary results along the way. We then present a series of formal results and numerical simulations that trace out the model’s key implications for credit market dynamics.

3.1 Firm borrowing

Each period $t$, the representative firm invests in a one-period project that requires a fixed up-front cost of $c > 0$. The next period, the project generates a random cash flow, $x_{t+1}$, that follows an exogenously given $AR(1)$ process

$$x_{t+1} - \bar{x} = \rho(x_t - \bar{x}) + \varepsilon_{t+1},$$

where $\rho \in (0, 1)$, $\bar{x} > c$, and the fundamental cash flow shock $\varepsilon_{t+1} \sim N(0, \sigma_{\varepsilon}^2)$ is i.i.d. over time.

The firm issues one-period bonds in order to finance these projects. Each bond is a promise to pay back one dollar to investors in one period. At time $t$, the price of each bond is denoted $p_t$. The total face amount of debt outstanding at time $t$ is $F_t$, meaning that at time $t$, the firm must repay the face amount of debt issued the prior period $F_{t-1}$. The firm also must pay the cost $c$ to begin a new project and receives the cash flow $x_t$ from the prior period’s project. Finally, the firm can issue new bonds at a price of $p_t$. Assuming the firm does not default and does not pay dividends to equity holders at time $t$, the total face amount of bonds outstanding at time $t$ is

$$F_t = (F_{t-1} + c - x_t)/p_t,$$

which is obtained by equating sources and uses.\textsuperscript{8} This law of motion is consistent with the fact

\textsuperscript{8}We assume that the firm always decides to invest, even when expected cash flows tomorrow do not cover the current cost—i.e., when $c > \bar{x} + \rho(x_t - \bar{x})$. There are two possible interpretations of this assumption. First,
that nonfinancial leverage is typically counter-cyclical, falling in good times when $x_t$ is high and rising in bad times when $x_t$ is low (Korajczyk and Levy [2003] and Halling, Yu, and Zechner [2016]).

We assume a simple mechanistic default rule. One should interpret a default by the representative firm in our model as a credit market bust in which there is an economy-wide spike in corporate defaults. Specifically, if at any time $t$, $F_{t-1} + c - x_t$ rises above a threshold of $F$, the representative firm defaults. Formally, letting $D_t$ denote a binary variable that indicates whether or not a default occurs at time $t$, we have

$$D_t = 1\{F_{t-1} + c - x_t \geq \bar{F}\}.$$  \hspace{1cm} (5)

The “default boundary” is the line in the $(F_{t-1}, x_t)$ space where the default indicator switches on or off—i.e., the line $F_{t-1} = \bar{F} - c + x_t$.

In the event of default, the firm continues to operate. However, the firms’ existing equity holders are wiped out and the firm writes off a fraction of its debt, much like under Chapter 11 of the U.S. Bankruptcy Code. Specifically, if the firm defaults, a fraction $1 - \eta$ of the firm’s debt is written off, generating losses for existing bondholders, and the remaining fraction $\eta \in (0, 1)$ is refinanced at current market prices. Thus, if the firm defaults at time $t$, the amount of debt outstanding becomes

$$F_t = \eta(F_{t-1} + c - x_t)/p_t.$$  \hspace{1cm} (6)

Finally, we assume that if $F_{t-1} + c - x_t \leq F$, the firm sets $F_t = F/p_t$ and pays all residual cash flows to equity holders as a dividend. The idea underlying the lower barrier $F > 0$ for debt outstanding can be motivated via the pecking order theory of capital structure (Myers and
Majluf [1984]). Firms only raise external finance in the form of debt. And when there is available free cash flow, the firm first uses this cash flow to retire existing debt. However, once the face value of debt reaches a sufficiently low level, the firm chooses to pay out all available free cash flows to its equity holders.

In summary, our assumptions imply that the firm defaults and its debt is written down when outstanding debt grows too large relative to the firm’s current free cash flows \((x_t - c)\)—i.e., when \(F_{t-1} \geq \bar{F} + (x_t - c)\). Conversely, the firm stops paying down its debt and instead pays all residual cash flows to equity holders when outstanding debt grows small relative to the firm’s current free cash flows—i.e., when \(F_{t-1} \leq \bar{F} + (x_t - c)\). In between these upper and lower boundaries, all else equal, the firm’s debt outstanding grows faster when refinancing conditions are less favorable—i.e., when the price of new bonds \(p_t\) is lower—and when the ratio of the firm’s free cash flow to outstanding debt \(x_t / F_{t-1}\) is lower.

Before turning to investor beliefs, we note that, taking \(F_t\) as given, it is straightforward to compute the fully-rational, forward-looking probability of a default at time \(t+1\), which we denote by \(\lambda_t^R\). Given the cash flow process in equation (3) and the default rule in equation (5), a default will occur at time \(t+1\) if and only if cash flows at \(t+1\) bring the firm over its default threshold

\[
\bar{F} \leq F_t + c - x_{t+1} = F_t + c - \rho x_t - (1 - \rho)\bar{x} - \varepsilon_{t+1}.
\]

Thus, at time \(t\), the true probability of default on the promised bond payments at time \(t+1\) is

\[
\lambda_t^R = \Phi \left( \frac{F_t - \bar{F} + c - \rho x_t - (1 - \rho)\bar{x}}{\sigma_{\varepsilon}} \right),
\]

where \(\Phi(\cdot)\) denotes the cumulative distribution function of the standard Normal distribution.

### 3.2 Investor beliefs

There is a continuum of risk-neutral bond investors with zero rate of time preference. Investors’ belief at time \(t\) about the probability of a default at time \(t+1\) is denoted \(\lambda_t\). For a given \(\lambda_t\), the
bond price at time \( t \) is simply

\[
p_t = p(\lambda_t) = (1 - \lambda_t) + \lambda_t \eta,
\]

(8)

and \( \lambda_t \) may deviate from \( \lambda_t^R \), the belief that would be held by a forward-looking rational investor.

Relative to the price of \( (1 - \lambda_t^R) + \lambda_t^R \eta \) in a fully rational economy, the bond price in (8) is deviated by \( (1 - \eta)(\lambda_t^R - \lambda_t) \).

The default rule in equations (5) and (6) and the bond pricing equation (8) give rise to the following law of motion for the amount of debt outstanding:

\[
F_t = F(F_{t-1}, \lambda_t, x_t) = \begin{cases} 
  F/p(\lambda_t) & \text{if } F_{t-1} + c - x_t \leq F \\
  (F_{t-1} + c - x_t)/p(\lambda_t) & \text{if } F < F_{t-1} + c - x_t < F \\
  \eta(F_{t-1} + c - x_t)/p(\lambda_t) & \text{if } F \leq F_{t-1} + c - x_t
\end{cases}
\]

(9)

Since \( p(\lambda_t) \leq 1 \), it follows that we always have \( F_t \geq F \). Thus, \( F \) is indeed a lower barrier for the amount of debt outstanding.

The model is fully characterized by equations (3), (8), and (9), together with the specifications for \( \lambda_t \) which will be introduced below in equation (10).

**Specifying investor beliefs \( \lambda_t \).** We now introduce our specification for \( \lambda_t \). We assume that \( \lambda_t \) depends solely on past default realizations and past “sentiment” shocks unrelated to cash flow fundamentals. Specifically, \( \lambda_t \) follows the law of motion

\[
\lambda_t = \max \{0, \min \{1, \beta \lambda_{t-1} + \alpha D_t + \omega_t\}\},
\]

(10)

where \( 0 < \beta < 1 \) is a memory decay parameter, \( \alpha > 0 \) measures the incremental impact of a default event on backward-looking beliefs, and \( \omega_t \sim N(0, \sigma^2_\omega) \) is a random “sentiment” shock that is independent of the fundamental cash flow shock \( \varepsilon_t \). The min and max functions in equation

\[\text{10}^\text{There are two distinct notions of “credit market sentiment.” First, one might say that credit market sentiment is elevated when } \lambda_t \text{ is low—i.e., when future defaults are perceived as being unlikely. Alternately, one might say that credit market sentiment is elevated when } (\lambda_t^R - \lambda_t) \text{ is high—i.e., when investors underestimate the true likelihood of a future default.}\]
ensure that $\lambda_t \in [0,1]$ for all $t$. If $\{\lambda_s; s \leq t\}$ is always strictly between 0 and 1, we have

$$
\lambda_t = \sum_{j=0}^{\infty} \beta^j (\alpha D_{t-j} + \omega_{t-j}).
$$

In this case, beliefs are just a geometric moving average of past defaults and past sentiment shocks.

The specification for extrapolative beliefs in equation (10) is similar to specifications in Barberis, Greenwood, Jin, and Shleifer (2015, 2018), and Nagel and Xu (2020). Empirically, equation (10) is motivated by the findings in Greenwood and Hanson (2013) who present evidence that credit market investors tend to extrapolate recent credit market outcomes. They show that credit market sentiment rises following periods when default rates have fallen and the returns on high-yield bonds have been high. These results hold controlling for contemporaneous conditions in the macroeconomy and in the stock market, suggesting that it is credit market outcomes, not fundamentals, that are being extrapolated.

The following lemma explains how beliefs evolve over time.

**Lemma 1** Assume there are no sentiment shocks (i.e., $\omega_t = 0$ for all $t$), so the law of motion for beliefs is simply $\lambda_t = \max \{0, \min \{1, \beta \lambda_{t-1} + \alpha D_t\}\}$.

- If there is no default at time $t$, then we always have $\lambda_t \leq \lambda_{t-1}$, and $\lambda_t < \lambda_{t-1}$ if $\lambda_{t-1} > 0$—i.e., beliefs always become more optimistic when there is no default.

- If there is a default at time $t$, then we have two cases:
  - If $\alpha \geq 1 - \beta$, then $\lambda_t \geq \lambda_{t-1}$, and $\lambda_t > \lambda_{t-1}$ if $\lambda_{t-1} > 1$—i.e., extrapolative beliefs always become more pessimistic following a default. As a result, $\lambda_t$ will converge to 1 following a long sequence of defaults.
  - If $\alpha < 1 - \beta$, then $\lambda_t \geq \lambda_{t-1}$ as $\lambda_{t-1} \leq \alpha/(1-\beta)$. As a result, $\lambda_t$ will converge to $\alpha/(1-\beta) < 1$ following a long sequence of defaults.

**Proof.** See the Appendix for all proofs.
The dynamics of $\lambda_t$ are governed by $\alpha$, the sensitivity of beliefs to default, and by the rate of memory decay $(1 - \beta)$. We will later show that the potential for backward-looking beliefs to drive persistent default cycles is greatest when $\alpha$ is high and when $\beta$ is close to 1 so memory fades slowly. In this case, a default episode makes investors much more pessimistic, which in turn makes it more difficult for firms to refinance maturing debt.

4 Understanding Reflexivity

Our model captures the idea that, in credit markets, investors’ biased beliefs can impact financial reality. Past defaults and sentiment shocks affect investors’ beliefs about future defaults via equation (10). These biased beliefs then impact bond prices via equation (8). And, since bond prices influence the ease with which the firm can refinance its existing debt, they in turn affect the evolution of debt outstanding via equation (9) and hence the true probability of future defaults in equation (7). As a result, biased investor beliefs have the potential to become partially self-fulfilling.

In this section, we provide a set of analytical results to illustrate the key implications of the model. After illustrating the key dynamics under a baseline set of model parameters, we lay out four key implications: defining reflexivity, the “calm before the storm” phenomenon, the “default spiral” phenomenon, and the predictability of bond returns.

Figure 4 shows a typical sample path of simulated data, illustrating some of the key dynamics in the model. To generate this data, we use the following set of baseline parameters:

- **Cash flow dynamics:** $\bar{x} = 2.4$, $\rho = 0.8$, $\sigma_\epsilon = 0.5$.
- **Investment cost:** $c = 2$.
- **Default and dividend barriers:** $\bar{F} = 5$, $\underline{F} = 1.5$.
- **Write-off parameter:** $\eta = 0.5$.
- **Belief dynamics:** $\beta = 0.8$, $\alpha = 0.2$, $\sigma_\omega = 0.05$. 
While these parameters are only illustrative, they have a number of desirable properties based on simulating the model for 100,000 periods (each period is one year):

1. **The unconditional default probability is realistic.** Here the unconditional probability of default is 8.5%. As noted above, one should interpret a default by our representative firm as a “credit market bust” in which there is an economy-wide spike in corporate defaults. Thus, these parameters imply that roughly one in ten years corresponds to such a bust.

2. **The unconditional means of $\lambda_t$ and $\lambda_t^R$ are similar.** Here the average of $\lambda_t^R$ is 8.7% and the average of $\lambda_t$ is 12.2%. Beliefs on average, then, are reasonable. As a result, the means of $(\lambda_t^R - \lambda_t)$ and $r_{t+1}$, the realized rate of return on bonds from $t$ to $t+1$, are small. Here the mean of $(\lambda_t^R - \lambda_t)$ is −3% and the average annual bond return is 0.3%.

3. **The time-series correlation between $\lambda_t$ and $\lambda_t^R$ is strong.** While clearly imperfect, investors’ beliefs are reasonable over time. Specifically, we have $Corr(\lambda_t, \lambda_t^R) = 0.56$. That is, the investors’ beliefs are close to the fully-rational ideal.

4. **Relation of $\alpha$ and $\beta$.** The strength of the default spiral mechanism is increasing in both $\alpha$ and $\beta$. Specifically, if $\alpha > (1 - \beta)$ then $\lambda_t$ always rises when $D_t = 1$ and $\omega_t = 0$. However, if $\alpha \leq 1 - \beta$ then $\lambda_t$ can actually fall when $D_t = 1$ and $\omega_t = 0$. Since $\alpha = 1 - \beta$ in this calibration, default spirals are possible.

The panels in Figure 4 show the path of cash flows, debt outstanding, default, the bond price, and investor beliefs. When showing investor beliefs, we also show the beliefs $\lambda_t^R$ that a rational investor would have. Notice that the time-series distribution of $\lambda_t^R$ is bimodal: $\lambda_t^R$ is typically either close to zero or 1. This bimodal distribution is largely a function of the short-term nature of debt in our model: short-term debt is extremely safe until it suddenly becomes risky.

Although it is just a single sample path, Figure 4 clearly illustrates some key features of the model. Consider the pattern of cash flows, debt, and investor beliefs shown between periods 18 and 30. In period 18, there is a negative cash flow realization ($x_t < c$), which must be covered by increased debt outstanding (top right panel). This increases the actual probability of default.
substantially, but investors’ perceived probability of default (bottom right panel) falls. In period 20, a default happens, and investors beliefs about future defaults also rise. Investor’s pessimism is persistent, with $\lambda_t$ remaining elevated even at period 30, when the true probability of default has fallen to approximately zero.

4.1 Understanding reflexivity

To more systematically explore the dynamic behavior implied by the model, we report impulse-response functions which trace out the dynamic impact of shocks to underlying cash flow fundamentals and investor sentiment. Let $z_t = (x_t, F_{t-1}, \lambda_t)$ denote the model’s state vector and consider some model-implied quantity $y_t$. The response of $y_{t+j}$ following an impulse $\varepsilon_t = s_\varepsilon$ to cash flow fundamentals $x_t$ at time $t$ is:

$$
\Phi_y(j, z_{t-1}, \varepsilon_t = s_\varepsilon) = \mathbb{E}^R[y_{t+j}|z_{t-1}, \varepsilon_t = s_\varepsilon] - \mathbb{E}^R[y_{t+j}|z_{t-1}, \varepsilon_t = 0].
$$

Similarly, the response of $y_{t+j}$ following an impulse $\omega_t = s_\omega$ to investor belief $\lambda_t$ at time $t$ is:

$$
\Phi_y(j, z_{t-1}, \omega_t = s_\omega) = \mathbb{E}^R[y_{t+j}|z_{t-1}, \omega_t = s_\omega] - \mathbb{E}^R[y_{t+j}|z_{t-1}, \omega_t = 0].
$$

Due to the nonlinear nature of the model, these impulse response functions (IRFs) can be asymmetric in the sense that the response to a positive shock can look quite different than the response to a negative shock, i.e., $\Phi_y(j, z_{t-1}, \omega_t = -s_\omega) \neq -\Phi_y(j, z_{t-1}, \omega_t = s_\omega)$. The IRFs are also state-contingent in the sense that both $\Phi_y(j, z_{t-1}, \varepsilon_t = s_\varepsilon)$ and $\Phi_y(j, z_{t-1}, \omega_t = s_\omega)$ depend on the initial condition $z_{t-1}$.

Figure 5 shows the IRFs for an impulse to cash flows $x_t$ and to beliefs $\lambda_t$ at time 1.\textsuperscript{11} The initial condition in Figure 5 is $x_0 = 2.25$, $F_{-1} = 2.25$, and $\lambda_0 = 0.3$. These initial conditions ensure that firm leverage is low and cash flows are strong, such that changes in investor beliefs

\textsuperscript{11}To compute these IRFs, we shock $x_t$ or $\lambda_t$ up or down at $t = 1$ and then generate 10,000 random paths following this shock. We also generate 10,000 random paths in the absence of a shock at $t = 1$. The IRF is just the difference in outcomes between the average path following this shock and the average path in the absence of a shock.
are unlikely to have a big impact on the true likelihood of default. Nonetheless, the impulse-
responses in Figure 5 are highly asymmetric. Starting from this initial condition, bad shocks
to either fundamentals (a downward shock to firm cash flows $x_t$) or investor beliefs (an upward
shock to the perceived default likelihood $\lambda_t$) have a much larger and more persistent effects on
credit market outcomes than good shocks.\(^{12}\)

Figure 6 repeats the exercise, but this time alters the initial conditions such that the firm is
near financial distress, by choosing $x_0 = 1.6$, $F_{-1} = 3.4$, and $\lambda_0 = 0.3$. In this region, changes
in investor beliefs can have a large impact on the true probability of default. First, consider an
impulse to cash flows $x_t$ at time 1. Compared to the responses in Figure 5, the same impulse to
fundamentals now has a far larger impact. (Note the difference in the $y$-axis scales in Figures
5 and 6.) In this way, our model naturally captures the fact that the build up of debt creates
fragility as emphasized by Krishnamurthty and Muir (2020).

The differences between Figure 5 and Figure 6 capture a more general point: while the
feedback loop between biased beliefs and credit market outcomes is always present, there are
times when it is particularly strong. More precisely, we can describe the economy as being in a
“reflexive” state when the true probability of default is highly dependent on investor beliefs $\lambda_t$—
i.e., when $\partial \lambda_t^R / \partial \lambda_t$ is large. While $\lambda_t^R$ always depends positively on $\lambda_t$, there are “non-reflexive”
regions in which $\partial \lambda_t^R / \partial \lambda_t$ is quite small. For example, states where debt is low and the economy
has not experienced defaults for a long time are typically non-reflexive. However, there are also
“highly reflexive” regions where $\partial \lambda_t^R / \partial \lambda_t$ is large: here a change in beliefs—whether due to a
current default or a sentiment shock $\omega_t$—will have a large impact on the true probability of
default $\lambda_t^R$.

When will $\partial \lambda_t^R / \partial \lambda_t$ be large? Using equations (7), (8), and (9), we have

\[
\frac{\partial \lambda_t^R}{\partial \lambda_t} = \frac{\phi(dist_t)}{\sigma_x} \cdot \frac{F(F_{t-1}, \lambda_t, x_t)}{p(\lambda_t)} (1 - \eta) \tag{12}
\]

\(^{12}\)Following shocks to beliefs, the saw-tooth patterns arise, even in expectation, because of the jaggedness of
debt outstanding in individual sample paths due to our mechanistic default rule.
where $\phi(\cdot)$ is the standard Normal density and

$$dist_t = \frac{F(F_{t-1}, \lambda_t, x_t) - F + c - \rho x_t - (1 - \rho)\bar{x}}{\sigma_x}$$

(13)

is the expected distance-to-default at time $t + 1$. Thus, $\partial \lambda_t^R / \partial \lambda_t$ is likely to be large at time $t$ when $|dist_t|$ is small so the economy is expected to be close to the default boundary at time $t + 1$, when $F_t = F(F_{t-1}, \lambda_t, x_t)$ is large, and when $p(\lambda_t)$ is small.

Figure 7 describes how reflexivity varies as a function of economic conditions. Using our baseline set of parameters, the heatmap in Panel A shows how $\lambda_t^R$ varies as a function of $(x_t, F_{t-1})$ when $\lambda_t = 0.2$. The dashed white line shows the default boundary at time $t$, namely $F_{t-1} = \bar{F} - c + x_t$, so the default region is to the northwest of this boundary. When the economy crosses the default boundary, there is a default today ($D_t = 1$) and a fraction $(1 - \eta)$ of debt is written down, leading the probability of another default tomorrow to jump downward. As shown, there are two regions where $\lambda_t^R$ is sensitive to movements in $(x_t, F_{t-1})$. The first of these regions—the one towards the southwest—is where there is no default today ($D_t = 0$), but where changes in current cashflows ($x_t$) and past debt ($F_{t-1}$) have a large impact on the default probability tomorrow. The second of these regions—the one towards the northwest—is where there is a default today ($D_t = 1$) and where changes in $(x_t, F_{t-1})$ have a large impact on $\lambda_t^R$.

The heatmap in Panel B shows how $\partial \lambda_t^R / \partial \lambda_t$, the magnitude of reflexivity, varies as a function of $(x_t, F_{t-1})$ when $\lambda_t = 0.2$. There are two highly reflexive regions where $\partial \lambda_t^R / \partial \lambda_t$ is large. The highly reflexive region in the southwest is where there is no default today ($D_t = 0$) and where a small increase in $\lambda_t$ has a large impact on the likelihood of a default tomorrow $\lambda_t^R$. Intuitively, in this region, a small increase in the backward-looking investor belief ($\lambda_t$) triggers an increase in the firm’s equilibrium debt burden, pushing the firm towards the brink of default. The second reflexive region in the northwest is where there is a default today ($D_t = 1$) and where a small increase in $\lambda_t$ has a large impact on the likelihood of another default tomorrow. Panel C shows these two results together, plotting both $\lambda_t^R$ and $\partial \lambda_t^R / \partial \lambda_t$ versus $F_{t-1}$ when $\lambda_t = 0.2$ and $x_t = 1$. Here the firm defaults today whenever $F_{t-1} \geq 4$. The economy is a highly reflexive state when $F_{t-1}$ is near 2.8 (corresponding to the first reflexive region mentioned above) or near 6.8.
(corresponding to the second region).

In summary, highly reflexive states arise near the end of a long credit boom when investors are still bullish, but default is not yet imminent. In this case, a credit crisis can suddenly become far more likely if investors become slightly more bearish on credit (i.e., if \( \lambda_t \) rises slightly). Highly reflexive states also arise in the wake of a credit bust where investors are still bearish. Here the likelihood that the credit crisis persists can drop dramatically if investors become slightly more bullish on credit (i.e., if \( \lambda_t \) falls slightly).

Proposition 1 below summarizes these results, explaining how reflexivity varies as a function of investor beliefs \( \lambda_t \), debt levels \( F_{t-1} \), and cash flows \( x_t \). It also describes the maximum magnitude of reflexivity.

**Proposition 1 Reflexivity.** Consider the case \( F_t < F_{t-1} + c - x_t < \tilde{F} \), so there is no default at time \( t \) (\( D_t = 0 \)). There exists a cutoff point \( \hat{\lambda}(F_{t-1}, x_t) \) such that reflexivity \( \partial \lambda^R_t / \partial \lambda_t \) increases in \( \lambda_t \) when \( \lambda_t < \hat{\lambda} \) and decreases in \( \lambda_t \) when \( \lambda_t > \hat{\lambda} \). Moreover, there exists a cutoff point \( \hat{F}(x_t, \lambda_t) \) such that reflexivity increases in \( F_{t-1} \) when \( F_{t-1} < \hat{F} \) and decreases in \( F_{t-1} \) when \( F_{t-1} > \hat{F} \). Lastly, there exists a cutoff point \( \hat{x}(F_{t-1}, \lambda_t) \) such that reflexivity increases in \( x_t \) when \( x_t < \hat{x} \) and decreases in \( x_t \) when \( x_t > \hat{x} \). When \( D_t = 0 \), the upper bound of reflexivity is \( \hat{F}(1 - \eta)/(\sqrt{2\pi} \sigma x \eta^2) \); reflexivity can get arbitrarily close to this upper bound.

Next, consider the case \( F_{t-1} + c - x_t \geq \tilde{F} \), so there is a default at time \( t \) (\( D_t = 1 \)). There exists a cutoff point \( \tilde{\lambda}(F_{t-1}, x_t) \) such that reflexivity \( \partial \lambda^R_t / \partial \lambda_t \) increases in \( \lambda_t \) when \( \lambda_t < \tilde{\lambda} \) and decreases in \( \lambda_t \) when \( \lambda_t > \tilde{\lambda} \). Moreover, there exists a cutoff point \( \tilde{F}(x_t, \lambda_t) \) such that reflexivity increases in \( F_{t-1} \) when \( F_{t-1} < \tilde{F} \) and decreases in \( F_{t-1} \) when \( F_{t-1} > \tilde{F} \). Lastly, there exists a cutoff point \( \tilde{x}(F_{t-1}, \lambda_t) \) such that reflexivity increases in \( x_t \) when \( x_t < \tilde{x} \) and decreases in \( x_t \) when \( x_t > \tilde{x} \).

The explicit expressions of \( \hat{\lambda}(F_{t-1}, x_t), \hat{F}(x_t, \lambda_t), \hat{x}(F_{t-1}, \lambda_t), \tilde{\lambda}(F_{t-1}, x_t), \tilde{F}(x_t, \lambda_t) \) and \( \tilde{x}(F_{t-1}, \lambda_t) \) are provided in the Appendix.

---

13 We have also examined a model with an alternative distributional assumption about the noise in cash flows—we assume that the cash flow shock \( \varepsilon_{t+1} \) in equation (3) has a uniform distribution. We find that the qualitative properties regarding the degree of reflexivity—in particular, the two highly reflexive regions described above—remain robust to this alternative assumption. We provide further discussion of this model in the Appendix.
4.2 The “calm before the storm” phenomenon

An elevated level of credit market sentiment—in the sense of a lower level of $\lambda_t$—slows down the accumulation of debt in the face of deteriorating cash flows fundamentals, thereby delaying or even preventing future defaults altogether. We term this phenomenon the “calm before the storm.” Below we provide a general result regarding this phenomenon.

**Proposition 2 Calm before the storm.** For any initial level of debt outstanding $F_{t-1}$ and cash flow $x_t$, lowering investor beliefs $\lambda_t$ weakly delays the next default path by path—i.e., for any given time series of future cash flow and sentiment shocks—and strictly delays the next default in expectation.

To illustrate the “calm before the storm” phenomenon, Figure 8 depicts a sample path of the model using our baseline set of parameters. The cash flow fundamental $x_t$ is initially set to $x_0 = 1.5 < 2 = c$ and debt is set to $F_0 = 3.5$. We assume that all of the subsequent shocks are zero ($\varepsilon_t = \omega_t = 0$). Figure 8 plots cash flows $x_t$, debt outstanding $F_t$, the default indicator $D_t$, bond prices $p_t$, rational beliefs $\lambda^R_t$, and investors’ biased beliefs $\lambda_t$. We compare the model dynamics starting from low and high initial values of beliefs, $\lambda_0 (Low) = 0.15$ and $\lambda_0 (High) = 0.3$. As can be seen, the firm defaults at time 3 when $\lambda_0 = \lambda_0 (High)$ and not at all when $\lambda_0 = \lambda_0 (Low)$. Consistent with Proposition 2, more optimistic initial beliefs have the potential to delay or even prevent default in the face of poor fundamental cash flows.

This “calm before the storm” behavior is consistent with one of the findings in Krishnamurthy and Muir (2020) who examine the behavior of credit spreads around a large sample of financial crises in developed countries and document that credit spreads are “too low” in the years before financial crises. Specifically, in our model, credit spreads are indeed too low in the run-up to a default, but jump on the realization of a default. Note that in our baseline model, we do not match Krishnamurthy and Muir’s other finding that credit spreads spike just before a crisis. We later show that this can be matched by adding a small measure of rational beliefs to the model.

The calm before the storm phenomenon also helps make sense of what Gennaioli and Shleifer (2018) have dubbed the “quiet period” of the 2008 global financial crisis—the period between the
initial disruptions in housing and credit markets in the summer of 2007 and onset of a full-blown financial crisis in the fall of 2008. Indeed, as Gennaioli and Shleifer (2018) argue, if investors were fully forward-looking, one should have expected a more rapid deterioration of financial conditions in late 2007 rather than the slow slide into crisis that was witnessed.

4.3 The “default spiral” phenomenon

Once the storm hits the credit market, default extrapolation can generate a default spiral: extrapolative, backward-looking beliefs lead to a form of default persistence that is absent when beliefs are fully rational and forward-looking. Specifically, investor beliefs typically become more pessimistic following a default according to equation (10). This pushes down bond prices, raising debt outstanding, and increasing the likelihood of future defaults.

Persistent default spirals can arise even when fundamental cash flows are strong \((x_t > c)\) if (i) investor beliefs are sensitive to defaults (i.e., if \(\alpha\) is high) and (ii) the rate of memory decay is low (i.e., if \(\beta\) is high). We formalize this observation in the following proposition.

**Proposition 3 Default spirals.** Suppose \(F_{t-1} + c - x_t \geq \bar{F}\), so there is a default at time \(t\) \((D_t = 1)\). A larger \(\alpha\) or a larger \(\beta\) increases the likelihood of a future default. Further suppose that (i) \(c \leq x_{t+1} \leq c + (F_{t-1} + c - x_t - \bar{F})\) and (ii) \(\omega_t = 0\). Then, with certainty, default continues to occur at \(t + 1\) \((D_{t+1} = 1)\) if \(\alpha + \beta \lambda_{t-1} \geq (1 - \eta(F_{t-1} + c - x_t)/(\bar{F} + x_{t+1} - c))/\eta\).

Proposition 3 says that, even when fundamentals are strong \((c \leq x_{t+1} \leq c + (F_{t-1} + c - x_t - \bar{F}))\), a string of multiple defaults can occur when default extrapolation induces a negative feedback loop. In particular, if a default has a significant impact on investor beliefs (i.e., if \(\alpha\) is high) or if memory fades slowly (i.e., if \(\beta\) is high), then, following a default, backward-looking extrapolative beliefs keep bond prices low and the debt level high for many periods. As a result, there is a lengthy sequence of defaults. By contrast, if investor beliefs are largely rational and forward-looking, then the debt writedown that occurs upon default leads to an immediate decrease in the rationally-expected default rate \(\lambda_t^R\). As such, bond prices quickly recover following the default and the firm rapidly repays its debts.

25
This default spiral dynamic highlights the potential disconnect between the endogenous credit cycle and the underlying business cycle that is at the heart of our model. In particular, the extrapolative nature of investor beliefs can make the financial recovery from a crisis slower and more protracted than in a world with fully forward-looking investors. This dynamic means that a moderate improvement in cash flows can be insufficient to “rescue” credit markets from a depressed state. Moreover, the likely timing of the recovery is influenced by the degree of investor pessimism ($\lambda_t$): one needs a large improvement in cash flows to ensure a recovery when $\lambda_t$ is high. The crucial role that investor beliefs play in driving default spirals suggests that a favorable sentiment shock (i.e., a large negative draw of $\omega_t$) coming from a policy intervention may also be an effective way to help credit markets recover.

### 4.4 Bond return predictability

In our model, bond returns are predictable. To see this, note that an investor who buy bonds for a price of $p_t = p(\lambda_t)$ at time $t$ will receive a payment of $1 - (1 - \eta)D_{t+1}$ at time $t + 1$. Thus, the realized return on risky bonds from time $t$ to $t + 1$ is

$$r_{t+1} = \frac{1 - (1 - \eta)D_{t+1}}{p(\lambda_t)} - 1. \quad (14)$$

At any time $t$, investors believe that $\mathbb{E}_t[D_{t+1}] = \lambda_t$ and the bond price is $p(\lambda_t) = 1 - (1 - \eta)\mathbb{E}_t[D_{t+1}] = 1 - (1 - \eta)\lambda_t$. Thus, by construction, investors always perceive a zero expected return on bonds from time $t$ to $t + 1$—i.e., $\mathbb{E}_t[r_{t+1}] = 0$. However, since $\mathbb{E}_t^{R}[D_{t+1}] = \lambda_t^R$ from the vantage point of a rational econometrician, the rationally-expected return on bonds is

$$\mathbb{E}_t^{R}[r_{t+1}] = \frac{1 - (1 - \eta)\lambda_t^R}{p(\lambda_t)} - 1 = \frac{-(1 - \eta)(\lambda_t^R - \lambda_t)}{1 - (1 - \eta)\lambda_t^R + (1 - \eta)(\lambda_t^R - \lambda_t)}. \quad (15)$$

Expected bond returns are negative (positive) when investors are overly bullish (bearish) about default probabilities—i.e., $\mathbb{E}_t^{R}[r_{t+1}] \gtrless 0$ as $\lambda_t^R \gtrless \lambda_t$. For instance, in a “calm before the storm” scenario where firm fundamentals have deteriorated but extrapolative investors remain bullish, we have $\lambda_t^R > \lambda_t$ and $\mathbb{E}_t^{R}[r_{t+1}] < 0$. Conversely, in a default spiral scenario where investors
are over-estimating the likelihood of future defaults because they have just witnessed a default, $\lambda_t^R < \lambda_t$ and $\mathbb{E}_t^R [r_{t+1}] > 0$.

Using equation (15) we can ask how a small change in $\lambda_t^R$ impacts expected bond returns. Holding fixed $\lambda_t$, an increase in $\lambda_t^R$ is associated with a decline in expected returns:

$$\frac{\partial \mathbb{E}_t^R [r_{t+1}]}{\partial \lambda_t^R} = -\frac{(1-\eta)}{p(\lambda_t)} < 0.$$ (16)

If we interpret $\lambda_t^R$ as an inverse measure of issuing firms’ creditworthiness, then equation (16) is consistent with Greenwood and Hanson (2013) who find that a deterioration in the credit quality of issuing firms negatively predicts future debt returns in a univariate sense. Going further, and assuming we are not at the default boundary, Lemma 1 implies that, all else equal, a small increase in $F_{t-1}$ leads to a decline in $\mathbb{E}_t^R [r_{t+1}]$ and a small increase in $x_t$ leads to an increase in $\mathbb{E}_t^R [r_{t+1}]$. Intuitively, when investors are extrapolative, holding fixed their beliefs, $\lambda_t$, worse cash flow fundamentals and higher levels of leverage predict lower future bond returns.

What is more interesting and subtle is that changes in investor beliefs—i.e., movements in $\lambda_t$—have an ambiguous impact on expected bond returns due to the reflexive nature of credit markets. Holding fixed expected future debt repayments, more bearish investor beliefs (higher values of $\lambda_t$) lower bond prices, raising expected bond returns. This is the intuition we have from standard settings where beliefs do not impact security payoffs. However, there is a competing effect that arises in our model because investor beliefs about future defaults are partially self-fulfilling. Specifically, more bearish investor sentiment makes it more difficult for firms to refinance maturing debt, raising the true probability of default and lowering expected future debt repayments. And, in highly reflexive states where investor beliefs have a large impact on the true likelihood of default—i.e., where $\partial \lambda_t^R / \partial \lambda_t > 0$ is large—the latter effect can outweigh the former.

As a result, the total impact of a shift in $\lambda_t$ on expected returns is ambiguous: depending on which effect dominates, a small increase in $\lambda_t$ can either lead $\mathbb{E}_t^R [r_{t+1}]$ to rise or fall. Formally,
we have

\[
\frac{\partial \mathbb{E}^R_t[r_{t+1}]}{\partial \lambda_t} = \frac{\partial \mathbb{E}^R_t[r_{t+1}]}{\partial \lambda_t} \bigg|_{\lambda^R_t = \text{Constant}} + \frac{\partial \mathbb{E}^R_t[r_{t+1}]}{\partial \lambda^R_t} \times \frac{\partial \lambda^R_t}{\partial \lambda_t} \geq 0
\]

which is ambiguous. And, we are more likely to have \( \frac{\partial \mathbb{E}^R_t[r_{t+1}]}{\partial \lambda_t} < 0 \) in highly reflexive states where \( \frac{\partial \lambda^R_t}{\partial \lambda_t} > 0 \) is large—e.g., near the end of a long credit boom when investors are still bullish, but default is not yet imminent.

Figure 9 shows the potentially ambiguous relationship between \( \mathbb{E}^R_t[r_{t+1}] \) and \( \lambda_t \). The figure plots \( \mathbb{E}^R_t[r_{t+1}] \) and \( \lambda^R_t \) versus \( \lambda_t \) using our baseline parameter values for \((x_t, F_{t-1}) = (1.6, 3.4)\), which is a highly reflexive state. For \( \lambda_t \) less than 0.2, \( \lambda^R_t \) rises gradually with \( \lambda_t \) and \( \mathbb{E}^R_t[r_{t+1}] \) is increasing in \( \lambda_t \). In this range, the negative effect of \( \lambda_t \) on bond price outweighs the positive effect on \( \lambda^R_t \). For \( \lambda_t \) between 0.2 and 0.63, \( \lambda^R_t \) rises with \( \lambda_t \) while \( \mathbb{E}^R_t[r_{t+1}] \) is decreasing in \( \lambda_t \): here the positive effect of \( \lambda_t \) on \( \lambda^R_t \) outweighs its negative effect on price.

Our model is also consistent with the multivariate return forecasting regressions emphasized in Greenwood and Hanson (2013) and López-Salido, Stein, and Zakrajšek (2017). These authors estimate forecasting regressions of the form:

\[
r_{t+1} = \alpha + \beta_1 \cdot \text{credit-spread}_t + \beta_2 \cdot \text{low-quality-issuance}_t + \xi_{t+1},
\]

where \( \text{credit-spread}_t \) is a measure of credit spreads and \( \text{low-quality-issuance}_t \) is an inverse measure of the creditworthiness of issuing firms. They find that \( \beta_1 > 0 \) and \( \beta_2 < 0 \): all else equal, future bond returns are high when current credit spreads are wide and are low when current debt issuers are less creditworthy. While credit spreads, borrower credit quality, and expected debt returns are endogenous equilibrium outcomes in our model, our model implies a strong multivariate forecasting relationship of the form found in the data. Specifically, the model analog of \( \text{credit-spread}_t \) is \( 1 - p_t = (1 - \eta) \lambda_t \) and the model analog of \( \text{low-quality-issuance}_t \) is
\( \lambda_t^R \). Since

\[
\mathbb{E}_t^R[r_{t+1}] \approx -k \cdot (1 - \eta)(\lambda_t^R - \lambda_t) = k \cdot [\text{credit-spread}_t - (1 - \eta) \cdot \text{low-quality-issuance}_t],
\]

by equation \((15)\) for some constant \( k > 1 \), we have \( \beta_1 \approx k > 0 \) and \( \beta_2 \approx -(1 - \eta) k < 0 \).

Intuitively, holding fixed \( \text{low-quality-issuance}_t = \lambda_t^R \), higher credit spreads signal higher values of \( \lambda_t \)—i.e., more bearish investor sentiment—and, thus, higher expected bond returns. Conversely, holding fixed credit spreads, more \( \text{low-quality-issuance}_t \) signals a larger gap between true credit risk and the risk perceived by investors and, therefore, lower expected returns.

We collect these observations in Proposition 4.

**Proposition 4 Return predictability.** Bond returns are predictable with \( \mathbb{E}_t^R[r_{t+1}] \), decreasing in \( \lambda_t^R - \lambda_t \), and equal to zero when \( \lambda_t^R - \lambda_t = 0 \).

- **Holding fixed \( \lambda_t \), a small increase in \( \lambda_t^R \) is associated with a decrease in \( \mathbb{E}_t^R[r_{t+1}] \). As a result, if the economy is not at the default boundary at time \( t \), then, all else equal, \( \mathbb{E}_t^R[r_{t+1}] \) is increasing in \( x_t \) and is decreasing in \( F_{t-1} \). However, these relationships flip signs when the economy is at the default boundary.**

- **Holding fixed \( x_t \) and \( F_{t-1} \), a small increase in \( \lambda_t \) has an ambiguous effect on \( \mathbb{E}_t^R[r_{t+1}] \):
  - In non-reflexive states—where a small increase in \( \lambda_t \) has a small effect on \( \lambda_t^R \)—a small increase in \( \lambda_t \) leads to an increase in \( \mathbb{E}_t^R[r_{t+1}] \).
  - In highly reflexive states—where a small increase in \( \lambda_t \) has a large effect on \( \lambda_t^R \)—a small increase in \( \lambda_t \) leads to a decline in \( \mathbb{E}_t^R[r_{t+1}] \).

- **We have \( \beta_1 > 0 \) and \( \beta_2 < 0 \) for the following multivariate regression:**

\[
r_{t+1} = \alpha + \beta_1 \cdot \text{credit-spread}_t + \beta_2 \cdot \text{low-quality-issuance}_t + \xi_{t+1},
\]

where \( \text{credit-spread}_t \equiv 1 - p_t \) and \( \text{low-quality-issuance}_t \equiv \lambda_t^R \).
Another question we can address is whether rapid credit growth negatively forecasts the returns on risky bonds and other credit-sensitive instruments documented in Greenwood and Hanson (2013) and Baron and Xiong (2017) and suggested in Schularick and Taylor (2012).

Assuming that $F < F_{t-1} + c - x_t < F$, the change is debt outstanding at time $t$ is

$$\Delta F_t = \frac{F_{t-1} + c - x_t}{p_t} - F_{t-1}.$$ 

It is straightforward to show that $\partial \Delta F_t / \partial x_t < 0$, $\partial \Delta F_t / \partial F_{t-1} > 0$, and $\partial \Delta F_t / \partial \lambda_t > 0$ when the economy is not near the default boundary. Combining these results with those in Proposition 4, one would expect large values of $\Delta F_t$ to predict low future values of $r_{t+1}$. This occurs because changes in $x_t$ and $F_{t-1}$ have opposing effects on $E_t^R[r_{t+1}]$ and $\Delta F_t$. And, changes in $\lambda_t$ will have opposing effects on $\Delta F_t$ and $E_t^R[r_{t+1}]$ in reflexive states where $\partial \lambda_t^R / \partial \lambda_t$ is large. Therefore, the model yields an overall negative relationship between increases in debt $\Delta F_t$ and expected returns $E_t^R[r_{t+1}]$.

5 Model extensions

In this section, we present several extensions of the model. The objectives are to provide a deeper understanding of the role of extrapolative beliefs in generating the model’s implications, and to improve the model’s realism.

5.1 Mixed beliefs

The baseline model assumes bond investors have extrapolative beliefs based totally on past firm defaults. This simple structure allows us to match many stylized facts about the credit markets, including predictability of bond returns and periods of calm before a crisis. As we have pointed out, however, when investor beliefs only react to past defaults, the simple model does not allow us to match the observation in Krishnamurty and Muir (2020) that credit spreads spike just before the crisis. In this section, we extend the model by introducing a rational component into

\textsuperscript{14}We have confirmed this result using numerical simulations.
investor beliefs. As we show below, this expanded structure leads credit spreads to jump up before a crisis. Moreover, we show that mixed beliefs can sometimes increase the magnitude of reflexivity, because investors are partly forward-looking.

Specifically, we assume that a fraction $\theta \in [0, 1]$ of investors’ beliefs are extrapolative and backward-looking and the remaining fraction $1 - \theta$ are rational and forward-looking. Thus, we have:

$$\lambda_t^C = \theta \lambda_t^B + (1 - \theta) \lambda_t^R = \lambda_t^R - \theta (\lambda_t^R - \lambda_t^B),$$

(19)

where $\lambda_t^C$ denotes the combined beliefs. For the rest of the paper, $\lambda_t^B$ denotes the backward-looking extrapolative beliefs. When $\theta = 1$, this model reduces to the model in Section 3 and $\lambda_t^B = \lambda_t$. On the contrary, when $\theta = 0$, investor beliefs become fully rational. This formulation of beliefs is in the spirit of Fuster, Laibson, and Mendel (2010) and Fuster, Hebert, and Laibson (2011) who argue that many agents have “natural expectations” which are a combination of fully-rational expectations and extrapolative expectations.

With mixed beliefs, bond investors set the bond price at time $t$ as

$$p_t = p(\lambda_t^B, \lambda_t^R) = (1 - \lambda_t^C) + \lambda_t^C \eta = 1 - (1 - \eta) \lambda_t^R + (1 - \eta) \theta (\lambda_t^R - \lambda_t^B).$$

(20)

Thus, the law of motion for the amount of debt outstanding becomes:

$$F_t = F(F_{t-1}, \lambda_t^B, \lambda_t^R, x_t) = \begin{cases} 
F / p(\lambda_t^B, \lambda_t^R) & \text{if } F_{t-1} + c - x_t \leq F \\
(F_{t-1} + c - x_t) / p(\lambda_t^B, \lambda_t^R) & \text{if } F < F_{t-1} + c - x_t < F \\
\eta (F_{t-1} + c - x_t) / p(\lambda_t^B, \lambda_t^R) & \text{if } F \leq F_{t-1} + c - x_t \end{cases}.$$  

(21)

Recall from equation (7) that $\lambda_t^R$ depends on $F_t$. Meanwhile, equations (20) and (21) imply that $F_t$ depends on $\lambda_t^R$ when $\theta < 1$. Thus, with mixed beliefs, $\lambda_t^R$ and $F_t$ must be simultaneously determined in equilibrium.

Formally, combining equations (7) and (21), we see that, when $\theta < 1$, the equilibrium value
of $\lambda_t^R$ must solve the following fixed-point problem:

$$
\lambda_t^R = g(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t) \equiv \Phi\left(\frac{F(F_{t-1}, \lambda_t^B, \lambda_t^R, x_t) + c - F - \rho x_t - (1 - \rho)\bar{\lambda}}{\sigma}\right).
$$

(22)

Note from (21) that the bond price $p(\lambda_t^B, \lambda_t^R)$ does not determine whether the firm defaults or pays dividends at time $t$; only $F_{t-1}$ and $x_t$ determine these outcomes. This means that $g(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t)$ is a continuous and increasing function of $\lambda_t^R$ for given values of $(F_{t-1}, \lambda_t^B, x_t)$. Also note that $g(0 | F_{t-1}, \lambda_t^B, x_t) > 0$ and $g(1 | F_{t-1}, \lambda_t^B, x_t) < 1$. Therefore, $g(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t)$ is a continuous function that maps the unit interval into itself, so a fixed point always exists by Brouwer’s fixed-point theorem.$^{15,16}$

Allowing $\theta < 1$ better allows us to match the evidence in Krishnamurthy and Muir (2020) who examine the behavior of credit spreads around a large sample of financial crises in developed countries. Specifically, with a backward-looking component of investor beliefs, credit spreads are typically low in the run-up to a default. At the same time, the forward-looking component of investor beliefs leads credit spreads to jump on the eve of a default. We can draw the link to these results more formally by tracing out the model-implied expected path of credit spreads conditional on a financial crisis at time $\tau = 0$. Specifically, we take simulated data from the model and estimate regression specifications of the form:

$$
(1 - \eta) \lambda_t^C = a + \sum_{T=-T}^{T} b_{t} 1_{(D_{t+1} = 1)} + e_t.
$$

(23)

$^{15}$The simultaneous determination of $F_t$ and $\lambda_t^R$ introduces the potential for multiple equilibria which reflects a straightforward self-fulfilling-prophecy intuition. If the rational component of investor beliefs about future default probabilities is low (high), then current bond prices are high (low). As a result, the face value of debt that the firm must promise to repay tomorrow is low (high), leading to a true probability of default tomorrow that is indeed low (high). Unlike in classic bank run models (e.g., Diamond and Dybvig [1983] and Goldstein and Pauzner [2005]), multiple equilibria in our model do not arise from strategic complementarities between the financing decisions of individual short-run creditors: the investors in our model are nonstrategic price-takers. Instead, multiple equilibria arise for reasons similar to those in classic models of sovereign default (e.g., Calvo [1988] and Cole and Kehoe [2000]).

$^{16}$When multiple equilibria arise, we focus on the smallest $\lambda^R$ that solves $\lambda^R = g(\lambda^R | \cdot)$—i.e., the model’s “best” stable equilibrium. An equilibrium is “stable” if it is robust to a small perturbation in investors’ beliefs regarding the likelihood of a default tomorrow. In our setting, if $\partial g(\lambda^R^* | \cdot) / \partial \lambda^R < 1$, then $\lambda^R^*$ is stable; if $\partial g(\lambda^R^* | \cdot) / \partial \lambda^R > 1$, then $\lambda^R^*$ is unstable. Since $g(0 | \cdot) > 0$ and $g(1 | \cdot) < 1$, our model always has at least one stable equilibrium. Following the correspondence principle of Samuelson (1947), stable equilibria have local comparative statics that accord with common sense. For example, at a stable equilibrium, $\lambda_t^R$ is locally increasing in $F_{t-1}$ and decreasing in $x_t$. 

32
Here \((1 - \eta) \lambda_t^C = 1 - p_t\) is the model analog of the credit spread. And we set \(\theta = 0.5\) so investor beliefs are an equal mix of backward-looking extrapolative beliefs and forward-looking rational beliefs. We then plot the \(b_t\) regression coefficients versus event time \(\tau\) for this regression in Figure 10, effectively tracing out the model-implied expected path of credit spreads in event time conditional on a financial crisis at time \(\tau = 0\). For purposes of comparison, we repeat the exercise separately for the rational, forward-looking component of spreads, \((1 - \eta) \lambda_t^R\), and the extrapolative, backward-looking component, \((1 - \eta) \lambda_t^B\).

Figure 10 shows these results. As can be seen, credit spreads \((1 - \eta) \lambda_t^C = (1 - \eta) \left[ \theta \lambda_t^B + (1 - \theta) \lambda_t^R \right]\) jump up on the eve of a crisis at \(\tau = -1\) due to their rational forward-looking component \((1 - \eta) \lambda_t^R\). However, comparing the coefficients for \(\lambda_t^R\) and \(\lambda_t^C\), we see that credit spreads are typically “too narrow” prior to financial crises as argued in Krishnamurthy and Muir (2020). Furthermore, Figure 10 shows that credit spreads are usually “too wide” in the aftermath of crises.

Besides more realistic credit spreads dynamics, the model with mixed beliefs can generate a higher degree of reflexivity, compared to the model with fully extrapolative beliefs. To see why, note that in this extended model,

\[
\frac{\partial \lambda_t^R}{\partial \lambda_t^B} = \frac{\frac{\partial g(\lambda_t^R|F_{t-1}, \lambda_t^B, x_t)}{\partial \lambda_t^R}}{1 - \frac{\partial g(\lambda_t^R|F_{t-1}, \lambda_t^B, x_t)}{\partial \lambda_t^R}} = \frac{\frac{\phi(dist_t)}{\sigma_x} F(F_{t-1}, \lambda_t^B, \lambda_t^R, x_t) p(\lambda_t^B, \lambda_t^R)}{1 - \frac{\phi(dist_t)}{\sigma_x} F(F_{t-1}, \lambda_t^B, \lambda_t^R, x_t) p(\lambda_t^B, \lambda_t^R)} (1 - \eta) (1 - \theta). \tag{24}
\]

Equation (24) implies a hump-shaped relationship between \(\partial \lambda_t^R / \lambda_t^B\) and \(\theta\). As \(\theta\) goes towards zero, \(\partial \lambda_t^R / \lambda_t^B\) decreases towards zero. As \(\theta\) goes towards one, the denominator in (24), \(1 - \frac{\partial g(\lambda_t^R|F_{t-1}, \lambda_t^B, x_t)}{\partial \lambda_t^R}\), increases towards one, reducing \(\partial \lambda_t^R / \lambda_t^B\). Therefore, \(\partial \lambda_t^R / \lambda_t^B\) is maximized when \(\theta\) takes an intermediate value. Intuitively, right before a crisis, the rational component of investor beliefs causes the bond price to fall. With a lower bond price, a small change in the behavioral component will then cause a significant change in the actual likelihood of future defaults. This interaction between the rational and behavioral components of investor beliefs amplifies the role of extrapolative beliefs in generating reflexivity.

Figure 11 confirms this intuition. The left panel describes how reflexivity varies as a function
of \((x_t, F_{t-1})\) when \(\theta = 1\) (fully extrapolative beliefs), and the right panel describes how reflexivity varies as a function of \((x_t, F_{t-1})\) when \(\theta = 0.5\) (an equal weight of rational beliefs and extrapolative beliefs). This comparison shows that, with mixed beliefs, the degree of reflexivity is indeed higher. Moreover, the interaction between the rational and behavioral components of investor beliefs leads to “narrower” reflexive regions: when the amount of debt outstanding moves away from the default boundary, the rational component of beliefs leads to a faster recovery of bond prices, reducing the role of extrapolative beliefs in generating reflexivity.

5.2 Bayesian inference

The baseline model in Section 3 can explain many stylized facts about the credit markets. However, notice that the model simultaneously marries two assumptions. First, the bond investors extrapolate from past defaults when forming beliefs about future defaults. Second, investors fail to observe firm fundamentals directly. To isolate the role of incomplete information from that of incorrect inference, below we consider an alternative model in which investors do not directly observe firm fundamentals but are otherwise rational, using Bayes’ law to infer the firm’s financial position at each point in time. We first present the structure of this model and then illustrate its main differences with our behavioral model using numerical analysis. Additional details are included in the Appendix.

To preview the results, this amended model still produces the “calm before the storm” phenomenon in which fundamentals have turned poor but credit spreads remain low. This happens because in the absence of defaults, bond investors use Bayes’ law and infer that the firm has strong cash flows and low debt, hence keeping credit spreads low. Nonetheless, this model further predicts that periods of low credit spreads persist even after firm defaults, counter to findings in Krishnamurthy and Muir (2020). Second, by construction, this fully rational model fails to generate the variation in expected bond returns observed empirically.

The model setup is identical to the one presented in Section 3, except that we now assume investors form Bayesian inference about cash flows and firm debt. Specifically, we assume the bond investors believe that per-period cash flows can only take a finite number of values and use
Bayes’ law to assess probabilities associated with these values.\textsuperscript{17} Denote the set of possible cash flow values as \( \mathbf{x} \equiv \{ \hat{x}_i | 1 \leq i \leq K; \hat{x}_j < \hat{x}_k \text{ if } j < k \} \), and denote the transition probability from \( \hat{x}_i \) to \( \hat{x}_j \) by:

\[
\delta_{i,j} \equiv \Pr (x_{t+1} = \hat{x}_j | x_t = \hat{x}_i), \quad \forall i, j.
\]  

(25)

To tie transition probabilities to the true cash flow process in equation (3), we assign values for \( \delta_{i,j} \) as follows

\[
\delta_{i,j} = \begin{cases} 
\Phi \left( \frac{1}{2} (\hat{x}_j + \hat{x}_{j+1}); \rho(\hat{x}_i - \bar{x}) + \bar{x}, \sigma_{\epsilon} \right) & j = 1 \\
\Phi \left( \frac{1}{2} (\hat{x}_j + \hat{x}_{j+1}); \rho(\hat{x}_i - \bar{x}) + \bar{x}, \sigma_{\epsilon} \right) - \Phi \left( \frac{1}{2} (\hat{x}_{j-1} + \hat{x}_j); \rho(\hat{x}_i - \bar{x}) + \bar{x}, \sigma_{\epsilon} \right) & 1 < j < K \\
1 - \Phi \left( \frac{1}{2} (\hat{x}_{j-1} + \hat{x}_j); \rho(\hat{x}_i - \bar{x}) + \bar{x}, \sigma_{\epsilon} \right) & j = K 
\end{cases}
\]  

(26)

where \( \Phi(\cdot; \mu, \sigma) \) is the cumulative distribution function of a Normal random variable with mean \( \mu \) and standard deviation \( \sigma \).

When investors do not observe \( x_t \), they also do not observe firm debt \( F_t \). As such, they need to form beliefs about both \( x_t \) and \( F_t \). To reduce the computational complexity for this joint formation of beliefs, we classify different levels of \( \hat{F}_t \)—investors’ perceived debt level at time \( t \)—into bins that range from \( F \) to \( F_U \), where \( F_U > \bar{F} \).\textsuperscript{18} Specifically, we consider \( N \) equally spaced bins with the \( i \text{th} \) bin defined by

\[
[F + (i - 1)(F_U - F)/N, F + i(F_U - F)/N] \equiv [F_i, \bar{F}_i], \; 1 \leq i \leq N.
\]

(27)

We further assume that investors do not differentiate debt levels within a bin, and that these

\textsuperscript{17}It is for tractability that we have investors believe cash flows can only take a finite number of values. If, instead, investors believe that \( x_t \) follows a continuous AR(1) process, then belief updates about cash flows using default realizations—default realizations are binary variables—is technically challenging: a truncated Normal distribution is not Normal; moreover, when investors believe that \( x_t \) is continuous, the model becomes non-Markovian: the number of state variables grows exponentially over time.

\textsuperscript{18}Debt level can go above \( \bar{F} \); upon a firm default, the debt level does not always go below \( \bar{F} \) immediately. For numerical simulations presented later, we choose a sufficiently large \( F_U \).
debt levels all equal the mid-point of the bin

\[ \hat{F}_i \equiv F + (i - 0.5)(F^U - F)/N. \]  

(28)

This simplifying assumption allows us to solve the model while keeping investors rational; in our numerical analysis, we set \( N \) to be sufficiently large, so that the “coarse” beliefs become rather accurate. Given the above assumptions, investor beliefs are fully characterized by \( N \cdot K \) probabilities:

\[ P_{i,j}^t \equiv \text{Prob} \left( F_t = \hat{F}_i, x_t = \hat{x}_j \right), \quad \forall i, j, \]

(29)

at each point in time. Updating of investor beliefs \( \{P_{i,j}^t\} \) depends on whether a default occurs or not.

We use numerical analysis to examine the implications of this model. Figure 12 plots firm debt \( F_t \), cash flow \( x_t \), bond price \( p_t \), default indicator \( D_t \), as well as bond investors’ perceived default probability \( \lambda_t \) for a short sample path. Comparing to the behavioral model with extrapolative beliefs, both models produce low credit spreads in the run-up to a default. In the behavioral model, credit spreads jump up on the eve of a default (either immediately on default for \( \theta = 1 \) or soon before for \( \theta < 1 \)). On the contrary, in the Bayesian model, credit spreads remain low both before the default and a few periods after the default, and then begin to increase; this implication is inconsistent with the evidence from Krishnamurthy and Muir (2020).

Why do credit spreads remain low after defaults in the model with Bayesian inference? In Figure 12, a default occurs at \( t = 17 \), and investors immediately revise downward their beliefs about cash flows; from \( t = 16 \) to \( t = 17 \), the perceived cash flow distribution shifts significantly to the left, concentrating around very low values. At the same time, investors know that the firm has just written off \( (1 - \eta) \) fraction of its debt and hence will take some time to accumulate debt back up above the default barrier \( \bar{F} \). Therefore, investors believe that debt levels tend to be below \( \bar{F} \) in the near future. Moreover, investors are aware of the persistence of poor fundamentals. As such, the perceived debt distribution continues to move towards the default barrier over the next few periods, resulting in a peak in perceived default probability at \( t = 20 \).
To sum up, a fully rational model with bond investors updating beliefs about firm fundamentals using Bayesian inference can indeed help match some of the facts observed in the data, including the “calm before the storm” phenomenon in which credit spreads are too low in the years before a crisis. However, this model fails to match the dynamics of credit spreads in the aftermath of a crisis, and generates no variation in expected bond returns. It seems that some departures from fully rational beliefs are called for.

5.3 Endogenous default barrier

The firms in our model are not value maximizing in a conventional sense. They simply follow a default rule that is a function of leverage. In this section, we study an extension of the baseline model ($\theta = 1$) by allowing profit-maximizing equity holders to endogenously set firms’ default barrier. This alternative setup is less analytically tractable than the baseline model, but it delivers an interesting new prediction that appears to be borne out in the data, namely the procyclicality of leverage ratios, as documented by many studies (Adrian and Shin, 2010a,b). Moreover, this procyclicality of leverage ratios can further exacerbate reflexivity, the “calm before the storm” phenomenon, and the ”default spiral” phenomenon discussed in Section 4.

We consider three sources of profits and costs for equity holders. First, when $F < F_{t-1} + c - x_t < \bar{F}_t$—here $\bar{F}_t$ denotes the default barrier at time $t$—the equity holders share a $k$ fraction of the rollover gain/loss. In dollar amount, the rollover gain/loss equals $kF_{t-1}p_t - k(F_{t-1} + c - x_t)$. Second, when $F_{t-1} + c - x_t \leq F$, the equity holders receive a payout $F - (F_{t-1} + c - x_t)$ in the form of dividend. Lastly, when $F_{t-1} + c - x_t \geq \bar{F}_t$, a default occurs and the equity holders pay a cost of $C$.\textsuperscript{19} Given these profits and costs, the equity holders maximize their average per-period net profit by choosing an optimal default policy in the following form\textsuperscript{20}

$$\bar{F}_t = A + B \cdot \lambda_{t-1}. \quad (30)$$

\textsuperscript{19}This setup is in the spirit of Leland (1994) and He and Xiong (2012).
\textsuperscript{20}The default barrier $\bar{F}_t$ can be a more general function: it can depend on $\lambda_{t-1}$ in a nonlinear way, and it can depend on other variables. However, setting $\bar{F}_t$ to a linear function of $\lambda_{t-1}$ is a reasonable starting point.
The law of motion for the amount of debt outstanding is:

\[ F_t = F(F_{t-1}, \lambda_t, \lambda_{t-1}, x_t) = \begin{cases} 
  \frac{F}{p}(&\lambda_t) & \text{if } F_{t-1} + c - x_t \leq F \\
  (1 - k)(F_{t-1} + c - x_t)/p(\lambda_t) + kF_{t-1} & \text{if } F < F_{t-1} + c - x_t < \bar{F}(\lambda_{t-1}) \\
  \eta(F_{t-1} + c - x_t)/p(\lambda_t) & \text{if } \bar{F}(\lambda_{t-1}) \leq F_{t-1} + c - x_t 
\end{cases} \]

(31)

When \( \bar{F}_t \) is kept constant and when \( k = 0 \), the law of motion for the amount of debt outstanding reduces to the baseline case characterized in equation (9).

We now consider a numerical example. We keep all the baseline parameters and further set \( k = 0.5 \) and \( C = 5 \). We then search for the coefficients \( A \) and \( B \) that lead to the highest average profit for the equity holders; the average profit is computed by taking an average of simulated profits from 100,000 periods. We find that \( A = 10.73 \) and \( B = -0.23 \). A negative coefficient \( B \) means that leverage is procyclical: when bond investors are optimistic about future defaults during a credit boom, equity holders choose to raise the default barrier, allowing the firm to take higher leverage. We also find that, this procyclicality of leverage sometimes increases the magnitude of reflexivity.

### 5.4 Opportunistic debt issuance

One limitation of the mixed beliefs model presented in Section 5.1 is: there is no sense in which a credit boom that is triggered by elevated levels of credit market sentiment naturally sows the seeds of its own destruction. Instead, all else equal, high levels of credit sentiment—i.e., lower levels of \( \lambda^B_t \)—always lead to a slower accumulation of debt, reducing the likelihood of a future crisis.\(^2\) However, the boom-bust narratives in Kindleberger (1978) and Minsky (1986) suggest that one might instead think a large over-valuation of risky debt (i.e., a large gap between \( \lambda^R_t \) and \( \lambda^B_t \)) could lead to an opportunistic increase in bond supply from firms—with extrapolative investors underreacting to the resulting increase in firm leverage, thereby raising the risk of a

\[^2\text{To see this recall that, assuming } F < F_{t-1} + c - x_t < \bar{F}, \text{ the change in debt at time } t \text{ is } \Delta F_t = (F_{t-1} + c - x_t)/p_t - F_{t-1}, \text{ which is decreasing in the bond price } p_t \text{ and, thus, increasing in } \lambda^B_t.\]
future credit crisis. This Minsky-type dynamic does not occur in our baseline model.

In this section, we consider a model extension which allows firms to exploit the mispricing of debt by having opportunistic debt issuance. Specifically, instead of (4), we now assume

\[
F_t = \frac{F_{t-1} + c - x_t}{p_t} + M \times \left[ p_t - (1 - (1 - \eta) \lambda_t^R) \right]
\]

(32)

\[
= \frac{F_{t-1} + c - x_t}{1 - \text{credit-spread}_t} + M \times \left[ \text{credit-spread}_t^R - \text{credit-spread}_t \right]
\]

\[
\approx \frac{F_{t-1} + c - x_t}{1 - \text{credit-spread}_t} - \left( M/k \right) \mathbb{E}_t^R[r_{t+1}].
\]

Here \(M \geq 0\) controls the aggressiveness of firms’ opportunistic supply response in response to debt mispricing, \(\text{credit-spread}_t = 1 - p_t = (1 - \eta) \lambda_t^C\), \(\text{credit-spread}_t^R = (1 - \eta) \lambda_t^R\), and the final line follows from the fact that \(\mathbb{E}_t^R[r_{t+1}] \approx -k (1 - \eta) \theta (\lambda_t^R - \lambda_t^B) = -k \times \left[ \text{credit-spread}_t^R - \text{credit-spread}_t \right]\). In this extension, opportunistic debt issuance raises firm leverage, but does not affect the quantity and quality of firm investment projects which are assumed to be fixed. Specifically, equation (32) implies that, all else equal, firms opportunistically take on additional leverage when credit spreads are too low—i.e., when rationally-expected returns are low. As discussed in the Appendix, we can solve for the equilibrium level of \(\lambda_t^R\) as before using equation (22), but now setting \(F(F_{t-1}, \lambda_t^R, \lambda_t^B, x_t)\) to \((F_{t-1} + c - x_t)/p(\lambda_t^R, \lambda_t^B) + M \times \left[ p(\lambda_t^R, \lambda_t^B) - (1 - (1 - \eta) \lambda_t^R) \right]\) in equation (9) when \(\underline{F} < F_{t-1} + c - x_t < \bar{F}\).

With this modification, the impact on \(F_t\) of a change in \(\lambda_t^B\) has an ambiguous sign. Specifically, when the firm is nearing default and \((F_{t-1} + c - x_t)/p(\lambda_t^R, \lambda_t^B)\) is large, a decline in \(\lambda_t^B\) will lower \(F_t\) just as in our baseline model. However, if the opportunistic supply response is sufficiently large, then when the firm is far from default and \((F_{t-1} + c - x_t)/p(\lambda_t^R, \lambda_t^B)\) is small, a decline in \(\lambda_t^B\) can raise \(F_t\). In other words, favorable credit market sentiment can lead to a boom—rising debt issuance and a decline in credit quality—that sows the seeds of its own destruction by increasing future default probabilities.

\[ \underline{\text{Greenwood and Hanson (2013) introduce an earlier model along these lines, although their model does not feature reflexivity.}} \]

\[ \underline{\text{Going further, we could also allow the quantity and quality (e.g., as parameterized by } \pi \text{ and } c \text{) of firm investment projects to depend on debt mispricing. Doing so would likely amplify the real effects of credit booms and busts even further.}} \]
Using this modified model with $M = 5$, Figure 13 shows the IRFs for an impulse to cash flows $x_t$ and to beliefs $\lambda_t^B$. The initial condition is the same as in Figure 5: $x_0 = 2.25$, $F_{-1} = 2.25$, and $\lambda_0^B = 0.3$, which is a non-reflexive region. The responses following an impulse to cash flows are similar in Figures 5 and 13. However, in Figure 13, firms’ opportunistic supply response means that—in contrast to the baseline IRFs shown in Figure 5—a downward shock to $\lambda_t^B$ now triggers an increase in outstanding debt $F_t$ and, hence, a rise in firm leverage and the likelihood of a future default crisis. In summary, firms’ opportunistic response to credit market sentiment means that credit booms have the potential to sow the seeds of their own destruction.

### 5.5 Multiple firms

The final model extension features multiple firms facing idiosyncratic cash flow shocks. As we saw earlier with a single representative firm, defaults are binary events. Allowing for multiple firms naturally yields a continuous default rate for the economy and therefore more realistic model-implied dynamics. We describe only the key new assumptions and results here and leave a more complete treatment for the Appendix. Our setup here builds on the mixed beliefs extension in Section 5.1.

We assume that there are $N$ firms, $i = 1, 2, \ldots, N$. The exogenous cash flow of firm $i$, $x_{i,t}$, consists of two components:

$$x_{i,t} = x_t + z_{i,t}, \quad (33)$$

where the systematic component $x_t$ evolves according to equation (3) and the mean-zero, firm-specific component $z_{i,t}$ follows

$$z_{i,t} = \psi z_{i,t-1} + \xi_{i,t}, \quad (34)$$

where $\xi_{i,t} \sim \mathcal{N}(0, \sigma^2_\xi)$ is i.i.d. over time, independent across firms, and independent of the systematic cash flow shock ($\varepsilon_t$) and the aggregate sentiment shock ($\omega_t$). The rule for firm default is similar to that in the baseline model: if at any time $t$, $F_{i,t-1} + c - x_{i,t}$ rises above $\bar{F}$, then firm $i$ defaults.

For simplicity, we assume that all firms’ bonds are priced identically even though firms have
heterogeneous cash flows and debt levels. This assumption can be seen as a short-hand for the idea that investors cannot perfectly observe each firm’s cash flow $x_{it}$ and leverage $F_{it-1}$ and treat firms as a homogeneous category. The extrapolative component of investor beliefs, $\lambda_t^B$, is then specified as follows. Let $D_{it} = 1\{F_{it-1} + c - x_{it} \geq F\}$ be an binary variable indicating whether firm $i$ defaults at time $t$ and let $\bar{D}_t = N^{-1} \sum_{i=1}^N D_{it}$ denote the economy-wide default rate at time $t$. We then assume that

$$\lambda_t^B = \max \{0, \min \{1, \beta \lambda_{t-1}^B + \alpha \bar{D}_t + \omega_t\}\}, \quad (35)$$

As in Section 5.1, investor belief $\lambda_t^C$ is a weighted average of behavioral beliefs and rational beliefs, $\lambda_t^R$, where $\lambda_t^R = \mathbb{E}_t^R [\bar{D}_{t+1}]$.

In the Appendix, we show that in this setup a solution exists. In fact, one consequence of having firms with different cash flows and debt levels is that, at each point in time, only a fraction of firms is close to defaults, so defaults occur more gradually. Second, by construction, this setup generates belief contagion: past defaults and likely future defaults of each firm affect the bond price, which in turn impacts the leverage dynamics of all firms.

The model with multiple firms yields similar qualitative implications to the mixed beliefs model featuring a single representative firm. For instance, the model with multiple firms still features the “calm before the storm” and “default spiral” phenomena. Similarly, since the realized return on an equal-weighted portfolio of risky bonds from time $t$ to $t+1$ is

$$\bar{r}_{t+1} = 1 - (1 - \eta) \bar{D}_{t+1} - 1, \quad (36)$$

and, since $\mathbb{E}_t^R [\bar{D}_{t+1}] = \lambda_t^R$, we have

$$\mathbb{E}_t^R [\bar{r}_{t+1}] = \frac{1 - (1 - \eta) \lambda_t^R}{p(\lambda_t^B, \lambda_t^R)} - 1 = \frac{-(1 - \eta) \theta (\lambda_t^R - \lambda_t^B)}{1 - (1 - \eta) \lambda_t^R + (1 - \eta) \theta (\lambda_t^R - \lambda_t^B)}, \quad (37)$$

just as in the mixed beliefs model from Section 5.1. Thus, the model with multiple heterogeneous firms has similar qualitative implications for bond return predictability compared to the mixed beliefs model. Quantitatively, however, the model with multiple heterogeneous firms leads to time-series dynamics of default rates that are more empirically realistic.
We present a model of credit market cycles in which investors extrapolate past defaults. Our key contribution is to model reflexivity in credit markets, an endogenous two-way feedback loop between biased investor beliefs and credit market outcomes. This feedback mechanism is particularly germane in credit markets, because firms must return to the market to refinance maturing debts, and the terms on which debt is refinanced will impact the likelihood of future defaults.

As we have shown, the combination of extrapolative beliefs and reflexive dynamics can lead to large short-run disconnects between cash flow fundamentals and credit market outcomes, including “calm before the storm” and “default spiral” episodes. Extrapolative beliefs also naturally lead to bond return predictability. But what is most striking here is that changes in investor sentiment can have an ambiguous impact on expected bond returns due to the reflexive nature of credit markets. When investors become more bullish, in the short run this can predict positive returns, even if at longer horizons expected returns become more negative.

Our analysis leaves open at least three areas for future research. First, we have not allowed conditions in credit markets to explicitly affect the underlying cash flow fundamentals of the economy. However, as demonstrated by a growing macro-finance literature, the inability to access credit on reasonable terms following a credit market bust may exacerbate an incipient economic downturn. Relatedly, according to Austrian accounts of the credit cycles, as the credit boom grows, increasing amounts of capital are devoted to poor quality projects to the detriment of future macroeconomic fundamentals. Indeed, López-Salido, Stein, and Zakrajšek (2017) and Mian, Sufi, and Verner (2017) show that periods of credit market overheating forecast low economic growth. Incorporating these features into our model would likely further strengthen the feedback loop between investor sentiment and credit market outcomes. Recent papers by Krishnamurthy and Li (2020), Maxted (2021), and Bordalo, Gennaioli, Shleifer, and Terry (2020) suggest progress in this direction.

Second, we have been silent on issues of welfare and optimal policy, even though our model suggests a potential role for policy. During credit booms, high sentiment can prevent defaults
from occurring in the near future, which can be welfare-improving if fundamentals recover soon enough. Nonetheless, self-fulfilling beliefs during default spirals can be welfare-reducing, both because these deteriorating beliefs accelerate future default realizations and because they lead to a slow recovery in the presence of improving fundamentals. Accepting these take-aways at face value, our model suggests a role for policy in moderating investor beliefs.

Third, our model has relevance for the literature on sovereign debt crises, suggesting how one might incorporate extrapolative expectations into standard models of sovereign crises (Calvo [1988] and Cole and Kehoe [2000]). Specifically, the introduction of extrapolative expectations may help explain the kinds of “slow-moving debt crises” studied in Lorenzoni and Werning (2019). And, our extension with multiple firms may help capture the idea of belief-driven market contagion across sovereign borrowers, which may prove useful in understanding events like the 1997 Asian financial crisis and the post-2010 European debt crisis.
References


Myers Stewart C., and Nicholas S. Majluf (1984). Corporate financing and investment decisions when firms have information that investors do not have, Journal of Financial Economics 13, 187–221.


A Proofs

Proof of Lemma 1: Since $\beta < 1$, $\lambda_t$ weakly declines if there is no default at time $t$ and the decline is strict if $\lambda_{t-1} > 0$.

How do extrapolative beliefs typically react to a default at time $t$—i.e., if $D_t = 1$ and $\omega_t = 0$? In this case, $\lambda_t = \min\{1, \beta \lambda_{t-1} + \alpha\} > 0$. If $\alpha \geq (1 - \beta)$, $\lambda_t$ weakly increases following a default and the increase is strict if $\lambda_{t-1} < 1$. Specifically, if $\lambda_t < 1$, then $\lambda_t - \lambda_{t-1} = \alpha - (1 - \beta) \lambda_{t-1} > 0$ for all $\lambda_{t-1} \in [0, 1)$ since $\alpha \geq (1 - \beta)$. By contrast, if $\lambda_t = 1$, then we trivially have $\lambda_t - \lambda_{t-1} > 0$ for all $\lambda_{t-1} \in [0, 1)$. Thus, if $\alpha \geq (1 - \beta)$, extrapolative beliefs will converge to $\lambda_t = 1$ following a long sequence of defaults.

By contrast, if $\alpha < (1 - \beta)$, extrapolative beliefs will not always become more pessimistic following a default. Specifically, if $D_t = 1$ and $\omega_t = 0$, then we have $\lambda_t \geq \lambda_{t-1}$ as $\lambda_{t-1} \leq \alpha / (1 - \beta)$ and extrapolative beliefs will converge to $\lambda_t = \alpha / (1 - \beta) < 1$ following a long sequence of defaults.

Proof of Proposition 1 (Reflexivity): First, when $\bar{F} < F_{t-1} + c - x_t < \bar{F}$, the definition of reflexivity $\mathcal{R} \equiv \partial \lambda_t^R / \partial \lambda_t$ implies

$$\text{sgn} \left( \frac{\partial \mathcal{R}(F_{t-1}, \lambda_t, x_t)}{\partial \lambda_t} \right) = \text{sgn} \left( 1 - \frac{1}{2\sigma_x^2} \left( \frac{F_{t-1} + c - x_t}{p(\lambda_t)} + c - \bar{F} - \rho x_t - (1 - \rho)\bar{x} \right) \frac{(F_{t-1} + c - x_t)}{p(\lambda_t)} \right). \tag{38}$$

The sign of the expression in the second line of (38) is positive when $\lambda_t < \hat{\lambda}$ and negative when $\lambda_t > \hat{\lambda}$; here $\hat{\lambda}$ is solved through setting the argument of the sign function to zero:

$$\hat{\lambda}(F_{t-1}, x_t) \equiv (1 - \eta)^{-1} \left[ 1 - \frac{F_{t-1} + c - x_t}{4\sigma_x^2} \left( \frac{\sqrt{[\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c]^2 + 8\sigma_x^2}}{-[\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c]} \right) \right]. \tag{39}$$

As such, $\mathcal{R}$ increases in $\lambda_t$ when $\lambda_t < \hat{\lambda}$ and decreases in $\lambda_t$ when $\lambda_t > \hat{\lambda}$. Moreover,

$$\text{sgn} \left( \frac{\partial \mathcal{R}(F_{t-1}, \lambda_t, x_t)}{\partial F_{t-1}} \right) = \text{sgn} \left( 1 - \frac{1}{\sigma_x^2} \left( \frac{F_{t-1} + c - x_t}{p(\lambda_t)} + c - \bar{F} - \rho x_t - (1 - \rho)\bar{x} \right) \frac{(F_{t-1} + c - x_t)}{p(\lambda_t)} \right). \tag{40}$$

The sign of the expression in the second line of (40) is positive when $F_{t-1} < \check{F}$ and negative when $F_{t-1} > \check{F}$; here $\check{F}$ is solved through setting the argument of the sign function to zero:

$$\check{F}(x_t, \lambda_t) \equiv x_t - c + \frac{[\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c]p(\lambda_t) + \sqrt{[\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c]^2 p^2(\lambda_t) + 4p^2(\lambda_t)\sigma_x^2}}{2}. \tag{41}$$
As such, $\mathcal{R}$ increases in $F_{t-1}$ when $F_{t-1} < \bar{F}$ and decreases in $F_{t-1}$ when $F_{t-1} > \bar{F}$. Lastly,

$$
= \text{sgn} \left( \frac{\partial \mathcal{R}(F_{t-1}, \lambda_t, x_t)}{\partial x_t} \right) \cdot \left( 1 + \frac{p(\lambda_t)(F_{t-1} + c - x_t)}{p(\lambda_t)} - 1 \right).
$$

(42)

The sign of the expression in the second line of (42) is positive when $x_t < \hat{x}$ and negative when $x_t > \hat{x}$; here $\hat{x}$ is solved through setting the argument of the sign function to zero:

$$
\hat{x}(F_{t-1}, \lambda_t) \equiv F_{t-1} + c - \frac{p(\lambda_t)}{2(1 + \rho p(\lambda_t))} \left( \sqrt{[(1 - \rho)c + \rho F_{t-1} - (1 - \rho)\bar{F}]^2 + 4\sigma^2} - [(1 - \rho)c - \rho F_{t-1} - (1 - \rho)\bar{F}] \right).
$$

(43)

As such, $\mathcal{R}$ increases in $x_t$ when $x_t < \hat{x}$ and decreases in $x_t$ when $x_t > \hat{x}$. When $F < F_{t-1} + c - x_t < \bar{F}$, the definition of $\mathcal{R}$ immediately implies an upper bound for $\mathcal{R}$ at $\bar{F}(1 - \eta)/\sqrt{2\pi}\sigma\eta^2$. One way of getting $\mathcal{R}$ arbitrarily close to this upper bound is by having

$$
x_t = \frac{1}{\rho} \left( \frac{1 - \eta}{\eta} \cdot \bar{F} + c - (1 - \rho)\bar{x} \right), \quad F_{t-1} \not\nearrow \bar{F} + x_t - c, \quad \lambda_t \not\nearrow 1.
$$

(44)

The results for the case $F_{t-1} + c - x_t \geq \bar{F}$ can be similarly derived. The expressions of $\hat{\lambda}(F_{t-1}, x_t)$, $\bar{F}(x_t, \lambda_t)$, and $\hat{x}(F_{t-1}, \lambda_t)$ are given by

$$
\hat{\lambda}(F_{t-1}, x_t) \equiv (1 - \eta)^{-1} \left[ 1 - \frac{\eta(F_{t-1} + c - x_t)}{4\sigma^2} \left( \sqrt{[\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c]^2 + 8\sigma^2} - [\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c] \right) \right],
$$

(45)

$$
\bar{F}(x_t, \lambda_t) \equiv x_t - c + \frac{[\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c]p(\lambda_t) + \sqrt{[\bar{F} + \rho x_t + (1 - \rho)\bar{x} - c]^2 p^2(\lambda_t) + 4p^2(\lambda_t)\sigma^2}}{2\eta},
$$

(46)

and

$$
\hat{x}(F_{t-1}, \lambda_t) \equiv F_{t-1} + c - \frac{p(\lambda_t)}{2(\eta + \rho p(\lambda_t))} \left( \sqrt{[(1 - \rho)c - \rho F_{t-1} - (1 - \rho)\bar{F}]^2 + 4\sigma^2} - [(1 - \rho)c - \rho F_{t-1} - (1 - \rho)\bar{F}] \right).
$$

(47)

respectively.

**Proof of Proposition 2 (Calm before the storm):** We compare two sample paths, denoted $L$ and $H$, that differ only in their initial levels of $\lambda_t$. Specifically, suppose that $\lambda_t(L) < \lambda_t(H)$. Because shocks to cash flows and sentiment are exogenous, we have $x_{t+j}^L(L) = x_{t+j}^L(H)$ and $\omega_{t+j}^L(L) = \omega_{t+j}^L(H)$ for all $j \geq 0$. Because $\lambda^R_t$ and $F_t$ are always increasing in $\lambda_t$, we have $\lambda^R_t(L) < \lambda^R_t(H)$ and $F_t(L) < F_t(H)$. Since $F_t(L) < F_t(H)$, if there is a default at time $t+1$ in the $L$ path, then there is also a default at time $t+1$ in the $H$ path. However, we can have default in the $H$ path, but not in the $L$ path at time $t+1$.

Assume that there is no default at time $t+1$ along either the $L$ or $H$ paths. Then we have $\lambda_{t+1}(L) \leq \lambda_{t+1}(H)$ by equation (10) and the equality is strict so long as $0 < \lambda_{t+1}(H)$. Since $\lambda^R_{t+1}$ and $F_{t+1}$ are increasing in $\lambda_{t+1}$ and $F_t$, it also follows that $\lambda^R_{t+1}(L) \leq \lambda^R_{t+1}(H)$ and $F_{t+1}(L) \leq F_{t+1}(H)$.
$F_{t+1}(H)$ and these inequalities are strict when $0 < \lambda_{t+1}(H)$. Since $F_{t+1}(L) \leq F_{t+1}(H)$, if the first default occurs at time $t+2$ in the $L$ path, then first default also occurs at time $t+2$ in the $H$ path. However, we can have default in the $H$ path, but not in the $L$ path at $t+2$.

Proceeding inductively in this fashion, we see that, so long as there is no default along either path by time $t+j$, we have $\lambda_{t+j}(L) \leq \lambda_{t+j}(H)$ and $F_{t+j}(L) \leq F_{t+j}(H)$ and these inequalities are strict when $\lambda_{t+j}(H) > 0$. Thus, lowering the default rate $\lambda_t$ weakly delays the next future default stochastic path by stochastic path. And, averaging across these paths, lowering the default rate $\lambda_t$ strictly delays the next default in expectation.

**Proof of Proposition 3 (Default spiral):** Suppose $F_{t-1} + c - x_t \geq \bar{F}$. Then the true probability of default over the next period is

$$\lambda_t^R = \Phi \left( \frac{\eta(F_{t-1}+c-x_t) + c - \bar{F} - \rho x_t - (1-\rho)\bar{x}}{\sigma_\varepsilon} \right).$$

(48)

As such,

$$\frac{\partial \lambda_t^R}{\partial \alpha} = \phi(\cdot) \frac{\eta(F_{t-1}+c-x_t) (1-\eta)}{(1-(1-\eta)\lambda_t)^2} \frac{\partial \lambda_t}{\sigma_\varepsilon} \frac{\partial \alpha}{\partial \beta}.$$ (49)

where $\phi(\cdot)$ is the standard Normal density evaluated at the argument given in the previous equation. When $\alpha < 1 - \beta \lambda_{t-1}$, $\partial \lambda_t / \partial \alpha > 0$, so $\partial \lambda_t^R / \partial \alpha > 0$. When $\alpha \geq 1 - \beta \lambda_{t-1}$, $\partial \lambda_t / \partial \alpha = 0$. As a result, $\partial \lambda_t^R / \partial \alpha = 0$. Similarly,

$$\frac{\partial \lambda_t^R}{\partial \beta} = \phi(\cdot) \frac{\eta(F_{t-1}+c-x_t) (1-\eta)}{(1-(1-\eta)\lambda_t)^2} \frac{\partial \lambda_t}{\partial \beta}.$$ (50)

So $\partial \lambda_t^R / \partial \beta > 0$ if $\alpha < 1 - \beta \lambda_{t-1}$, and $\partial \lambda_t^R / \partial \beta = 0$ if $\alpha \geq 1 - \beta \lambda_{t-1}$.

Moreover, suppose that (i) $c \leq x_{t+1} \leq c + (F_{t-1} + c - x_t - \bar{F})$; (ii) $\omega_t = 0$; and (iii) $\alpha + \beta \lambda_{t-1} \geq (1 - \eta(F_{t-1} + c - x_t) / (\bar{F} + x_{t+1} - c)) / (1 - \eta)$. Then

$$F_t + c - x_{t+1} = \frac{\eta(F_{t-1} + c - x_t)}{p_t} + c - x_{t+1} \geq \bar{F}$$ (51)

and default occurs at $t+1$.

**Belief updating in the model with Bayesian inference:** Suppose investors’ prior beliefs at time $t$ are $\{P_{i,j}^t\}$. Then, at $t+1$, belief updating depends on whether a default occurs or not. Below we discuss these two cases separately.

**Case A: Default at $t+1$.** This is the case where $F_t + c - x_{t+1}$—observed by outside econometricians—is greater than or equal to $\bar{F}$. In this case, we write

$$\lambda_t = \sum_{(i,j,j') \in \mathbf{A}_t} P_{i,j}^t \cdot \delta_{j,j'},$$

(52)

where $\mathbf{A}_t \equiv \{(i,j,j')$ with $F_t + c - x_{j'} \geq \bar{F}\}$. All the $(i,j,j')$ triplets with $F_t + c - x_{j'} < \bar{F}$ are not included in $\mathbf{A}_t$ because they do not lead to default at $t+1$. Thus, their associated probabilities
do not contribute to investors’ posterior beliefs at $t + 1$. We then define

$$\hat{P}_{t+1}^{i,j,j'} \equiv \lambda_t^{-1} \cdot P_t^{i,j} \cdot \delta_{j,j'}$$

(53)

for $(i, j, j') \in A_t$, which corresponds to the joint probability of $F_t = \hat{F}_t$, $x_t = \hat{x}_j$, and $x_{t+1} = \hat{x}_{j'}$. How does $\hat{P}_{t+1}^{i,j,j'}$ map to a new bin $j'$ for firm debt at time $t + 1$? It depends on the bond price $p_{t+1} = (1 - \lambda_{t+1}) + \lambda_{t+1} \cdot \eta$ at $t + 1$. To set the bond price, we search for a fixed point. Specifically, we examine values of $p$ in the range between one and $\eta$, the upper and lower bounds of the bond price. For each $p$, we derive

$$\hat{\lambda}_{t+1} = \sum_{(i,j,j',j'') \in A_{t+1}(p)} \hat{P}_{t+1}^{i,j,j'} \cdot \delta_{j', j''},$$

(55)

where

$$A_{t+1}(p) \equiv \{(i, j, j', j'') \text{ with } (i, j, j') \in A_t \text{ and } \eta \frac{(F_i + c - x_{j'})}{p} + c - x_{j''} \geq \bar{F}\}.$$  

(56)

We look for the largest $p$ such that $\hat{p} \equiv (1 - \hat{\lambda}_{t+1}) + \hat{\lambda}_{t+1} \cdot \eta \geq p$. We then set the bond price at time $t + 1$ to $p_{t+1} = \hat{p}$. And we set the investors’ $t + 1$ perceived probability of future defaults to $\lambda_{t+1} = \hat{\lambda}_{t+1}$. Finally, we cluster debt levels into bins, so investors’ beliefs at $t + 1$ are

$$\hat{P}_{t+1}^{i,j,j'} = \sum_{(i,j,j') \in A_t} \hat{P}_{t+1}^{i,j,j'} \cdot 1_{\{(i, j, j', j'') \text{ with } (i, j, j') \in A_t \text{ and } \eta \frac{(F_i + c - x_{j'})}{p} + c - x_{j''} \geq \bar{F}\}}.$$  

(57)

Case B: No default at $t + 1$. This is the case where $F_t + c - x_{t+1}$—observed by outside econometricians—is less than $\bar{F}$. In this case, the definition of $\hat{P}_{t+1}^{i,j,j'}$ is

$$\hat{P}_{t+1}^{i,j,j'} \equiv (1 - \lambda_t)^{-1} \cdot P_t^{i,j} \cdot \delta_{j,j'}$$

(58)

for $(i, j, j') \in A_c^t$, where $A_c^t \equiv \{(i, j, j') \text{ with } F_i + c - x_{j'} < \bar{F}\}$. Moreover, $A_{t+1}(p)$ in the definition of $\hat{\lambda}_{t+1}$ is replaced by

$$A_{t+1}(p) \equiv \{(i, j, j', j'') \text{ with } (i, j, j') \in A_c^t \text{ and } \frac{(F_i + c - x_{j'})}{p} + c - x_{j''} \geq \bar{F}\}.$$  

(59)

And investors’ beliefs at $t + 1$ are

$$\hat{P}_{t+1}^{j',j''} = \sum_{(i,j,j') \in A_c^t} \hat{P}_{t+1}^{i,j,j'} \cdot 1_{\{(i, j, j', j'') \text{ with } (i, j, j') \in A_c^t \text{ and } \eta \frac{(F_i + c - x_{j'})}{p_{t+1}} + c - x_{j''} \geq \bar{F}\}}.$$  

(60)

Model with opportunistic supply response: Assuming the firm does not default or pay dividends at time $t$, one can think of the mixed beliefs model presented in Section 5.1 as reflecting the interplay between the demand and supply for risk bonds:
Demand for bonds: \[ p_t^D = 1 - (1 - \eta) \theta \lambda_t^B - (1 - \eta) (1 - \theta) \Phi \left( \frac{F_t + c - \bar{F} - \rho x_t - (1 - \rho)\bar{x}}{\sigma} \right) \]

Supply of bonds:

\[ p_t^S = \frac{F_{t-1} + c - x_t}{F_t}, \]

where \( F_t \) is the quantity of risky bonds issued at time \( t \). We have \( \partial p_t^D / \partial F_t < 0 \), so the demand for bonds is downward sloping as is standard. (Here this works through a change in fundamentals: the default probability increases as debt outstanding rises.) However, we also have \( \partial p_t^S / \partial F_t < 0 \): supply is also downward sloping, which is non-standard. This is because the supply of bonds is determined by firms’ binding sources-and-used constraint. Of course, it is the fact that both demand and supply slope downwards that makes investor beliefs potentially self-fulfilling. However, the fact that supply cannot be upward-sloping precludes the kind of opportunistic supply response that might might allow a credit boom to sow the seeds of its own destruction.

Once we introduce a opportunistic debt supply response, the equilibrium value of \( \lambda_t^R \) must solve the following fixed-point problem:

\[ \lambda_t^R = g_{opp}(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t) \equiv \Phi \left( \frac{F_{opp}(F_{t-1}, \lambda_t^B, \lambda_t^R, x_t) + c - \bar{F} - \rho x_t - (1 - \rho)\bar{x}}{\sigma} \right). \]  

(61)

Here

\[ F_{opp}(F_{t-1}, \lambda_t^B, \lambda_t^R, x_t) = \begin{cases} \frac{F}{p(\lambda_t^B, \lambda_t^R)} & \text{if } F_{t-1} + c - x_t \leq F \\ (F_{t-1} + c - x_t)/p(\lambda_t^B, \lambda_t^R) & \text{if } F < F_{t-1} + c - x_t < \bar{F} \\ M \times [p(\lambda_t^B, \lambda_t^R) - (1 - (1 - \eta) \lambda_t^R)] & \text{if } \bar{F} \leq F_{t-1} + c - x_t \end{cases} \]

where \( M \geq 0 \) controls the aggressiveness of the corporate supply response to debt mispricing.

When \( M > 0 \), \( g_{opp}(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t) \) need not be monotonically increasing in \( \lambda_t^R \). However, for a given value of \( (F_{t-1}, \lambda_t^B, x_t) \), \( g(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t) \) is still a continuous function that maps the unit interval into itself—i.e., \( g(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t) \in [0, 1] \) for any \( \lambda_t^R \in [0, 1] \)—so a fixed point always exists by Brouwer’s fixed-point theorem. As in the model with \( M = 0 \), we select the smallest \( \lambda_t^R \) that solves \( \lambda_t^R = g_{opp}(\lambda_t^R | F_{t-1}, \lambda_t^B, x_t) \).

Let \( F_t \) denote the equilibrium level of debt at time \( t \). Holding \( \lambda_t^R \) fixed, we now have

\[ \frac{\partial F_t}{\partial \lambda_t^B} = (1 - \eta) \theta \left( \frac{F_{t-1} + c - x_t}{1 - (1 - \eta) \lambda_t^R + (1 - \eta) \theta (\lambda_t^R - \lambda_t^B)} - M \right). \]  

(62)

The size of \( \partial F_t / \partial \lambda_t^B \) is ambiguous when \( M > 0 \). And, holding fixed \( \lambda_t^B \), we have

\[ \frac{\partial F_t}{\partial \lambda_t^R} = (1 - \eta) \left( 1 - \theta \right) \left( \frac{F_{t-1} + c - x_t}{1 - (1 - \eta) \lambda_t^R + (1 - \eta) \theta (\lambda_t^R - \lambda_t^B)} + \theta M \right) > 0. \]  

(63)
Thus, we have

$$\frac{\delta F_t}{\delta \lambda_t^B} = \frac{F_{t-1} + c - x_t}{1 - (1 - \eta) \lambda_t^R + (1 - \eta) \theta (\lambda_t^R - \lambda_t^B)} \left[ (1 - \theta) \frac{\partial \lambda_t^R}{\partial \lambda_t^B} + \theta \right] + (1 - \eta) \theta M \left[ \frac{\partial \lambda_t^R}{\partial \lambda_t^B} - 1 \right].$$

Thus, the sign of $\delta F_t/\delta \lambda_t^B$ can vary across the parameter space when $M > 0$. Specifically, for sufficiently large $M$, we will have $\delta F_t/\delta \lambda_t^B < 0$ in good times when $F_{t-1} + c - x_t$ is small and $\partial \lambda_t^R/\partial \lambda_t^B$ is small. Here the opportunistic respond dominates and supply is upward sloping. By contrast, we will have $\delta F_t/\delta \lambda_t^B > 0$ in bad times when $F_{t-1} + c - x_t$ is large and $\partial \lambda_t^R/\partial \lambda_t^B$ is large. Here supply is downward sloping.

**Model with multiple firms:** The law of motion for each firm’s outstanding bonds $F_{it}$ is

$$F_{it} = F \left( F_{it-1}, \lambda_t^B, \lambda_t^R, x_{it} \right) = \begin{cases} F/p(\lambda_t^B, \lambda_t^R) & \text{if } F_{it-1} + c - x_{it} \leq \bar{F} \\ (F_{it-1} + c - x_{it})/p(\lambda_t^B, \lambda_t^R) & \text{if } \bar{F} < F_{it-1} + c - x_{it} < \bar{F} \\ \eta(F_{it-1} + c - x_{it})/p(\lambda_t^B, \lambda_t^R) & \text{if } \bar{F} \leq F_{it-1} + c - x_{it} \end{cases}$$

(64)

where $p(\lambda_t^B, \lambda_t^R) = [1 - (1 - \eta)\lambda_t^R] + (1 - \eta)\theta(\lambda_t^R - \lambda_t^B)$ is the price of corporate bonds, $\lambda_t^R$ is given in Section 5.5, and

$$\lambda_t^R = g(\lambda_t^R | \{F_{it-1}\}_{i=1}^N, \{z_{it}\}_{i=1}^N, x_t, \lambda_t^B)$$

$$= \frac{1}{N} \sum_{i=1}^N \Phi \left( \frac{F(F_{it-1}, \lambda_t^B, \lambda_t^R, x_{it}) - \bar{F} + c - \rho x_t - (1 - \rho)\bar{x} - \psi z_{it}}{\sqrt{\sigma_{\varepsilon}^2 + \sigma_{\xi}^2}} \right).$$

(65)

Just as in the mixed beliefs model, equations (64) and (65) imply that the right hand side of (65) can be viewed as a continuous function of $\lambda_t^B$ that maps the unit interval into itself. Therefore, by Brouwer’s fixed-point theorem, a solution exists. But, in addition to $x_t$ and $\lambda_t^B$, the distributions of $\{F_{it-1}\}_{i=1}^N$ and $\{z_{it}\}_{i=1}^N$ now impact $\lambda_t^R$.

**An alternative distributional assumption about the noise in cash flows:** We now consider an alternative distributional assumption of the cash flow shock. We assume that $\varepsilon_{t+1}$ in equation (3) has a uniform distribution between $-\xi$ and $\xi$. In all other respects, the model is kept the same as the baseline model.

At time $t$, the true probability of default on the promised bond payments at time $t + 1$ is

$$\lambda_t^R = \frac{F_t - \bar{F} + c - \rho x_t - (1 - \rho)\bar{x} + \xi}{2\xi}$$

(66)
if the right hand side of (66) is between zero and one; if it is below zero, then \( \lambda^R_t = 0 \); and if it is above one, \( \lambda^R_t = 1 \). Reflexivity, defined as \( \frac{\partial \lambda^R_t}{\partial \lambda_t} \), is

\[
\frac{\partial \lambda^R_t}{\partial \lambda_t} = \begin{cases} 
\frac{F(F_{t-1}, \lambda_t, x_t)}{2\xi} \cdot \frac{(1 - \eta)}{p(\lambda_t)} & \text{if } -\xi < F_t - \bar{F} + c - \rho x_t - (1 - \rho)\bar{x} < \xi \\
0 & \text{otherwise}
\end{cases}
\]

(67)

When the cash flow shock has a uniform distribution, we continue to observe two reflexive regions: one is when there is no default today but the values of \( F_{t-1} \) and \( x_t \) are chosen such that a small change in \( \lambda_t \) has a large impact on the likelihood of a default tomorrow \( \lambda^R_t \), and the other is when there is a default today and a small change in \( \lambda_t \) has a large impact on the likelihood of another default tomorrow \( \lambda^R_t \).
Figure 1. The credit cycle.
Figure 2. The credit market cycle. Panel A plots the year-over-year growth in real GDP and the year-over-year growth in real credit outstanding (defined as the sum of loans and bonds) to nonfinancial corporate businesses from the Federal Reserve’s Financial Accounts of the United States. Panel B plots real year-over-year credit growth versus the corporate credit spread, measured as the yields on Moody’s seasoned Baa corporate bond yield minus the 10-year constant maturity Treasury yield. The data begin in 1952 and end in 2016.
Figure 3. Real GDP growth and credit growth as a function of business cycle expansion quarter. This figure plots real GDP growth and real credit growth—the growth in real nonfinancial corporate loans and bonds from the Financial Accounts of the United States—as a function of NBER business cycle expansion quarter. The data begin in 1952 and end in 2016.
Figure 4. Simulated data using baseline parameter values. This figure shows a typical path of simulated data using our baseline set of parameter values. Specifically, the parameters are $\bar{x} = 2.4$, $\rho = 0.8$, $\sigma_c = 0.5$, $c = 2$, $F = 1.5$, $\bar{F} = 5$, $\eta = 0.5$, $\beta = 0.8$, $\alpha = 0.2$, and $\sigma_\omega = 0.05$. We plot the evolution of cash flows ($x_t$), debt outstanding ($F_t$), the default indicator ($D_t$), bond prices ($p_t$), investor beliefs ($\lambda_t$), and forward-looking rational beliefs about future defaults ($\lambda^R_t$). Each period represents one year.
Figure 5. Model-implied impulse response functions in a non-reflexive region. The top panel shows the responses following a 0.5 up or down impulse to cash flows ($x_t$) at $t = 1$. The bottom panel shows the responses following a 0.25 up or down impulse to investor beliefs at $t = 1$. The initial condition in both cases is $x_0 = 2.25$, $F_{-1} = 2.25$, and $\lambda_0 = 0.3$. The model parameters are the same as those in Figure 4.
Figure 6. Model-implied impulse response functions in a reflexive region. The top panel shows the responses following a 0.5 up or down impulse to cash flows ($x_t$) at $t = 1$. The bottom panel shows the responses following a 0.25 up or down impulse to investor beliefs at $t = 1$. The initial condition in both cases is $x_0 = 1.6$, $F_{-1} = 3.4$, and $\lambda_0 = 0.3$. The model parameters are the same as those in Figure 4.
Figure 7. Reflexive regions. This figure illustrates the existence of reflexive regions in our model. The heatmap in Panel A plots \( \lambda_i^R vs. (x_t, F_{t-1}) \) for \( \lambda_i = 0.2 \). The heatmap in Panel B plots \( \frac{\partial \lambda_i^R}{\partial \lambda_i} vs. (x_t, F_{t-1}) \) for \( \lambda_i = 0.2 \). (The dashed white line in Panels A and B is the default boundary at time \( t \).) Finally, Panel C plots \( \lambda_i^R \) and \( \frac{\partial \lambda_i^R}{\partial \lambda_i} vs. F_{t-1} \) for \( \lambda_i = 0.2 \) and \( x_t = 1 \). (The vertical black line is the default boundary at time \( t \).) The model parameters are the same as those in Figure 4.
Figure 8. Calm before the storm. This figure illustrates the “calm before the storm” phenomenon. The figure depicts sample paths of the model with cash flows initially set to $x_0 = 1.5 < 2 = c$ and debt initially set to $F_0 = 3.5$. We compare the model dynamics starting from a low initial value of $\lambda_0(L) = 0.15$ and a high initial value $\lambda_0(H) = 0.3$. We assume all subsequent shocks are zero ($\varepsilon_t = \omega_t = 0$). Otherwise, the model parameters are the same as those in Figure 4.
Figure 9. Impact of backward-looking beliefs on the true default probability and expected returns. This figure plots the true default probability \( \lambda_t^R \) (blue) and rationally-expected returns \( E_t^R[r_{t+1}] \) (red) against backward-looking beliefs \( \lambda_t \) in a highly reflexive region of the state space—i.e., a region where \( \partial \lambda_t^R / \partial \lambda_t \) is large so changes in beliefs \( \lambda_t \) have a large impact on future defaults. Specifically, we set \( x_t = 1.6 < 2 = c \) and \( F_{t-1} = 3.4 \). The model parameters are the same as those in Figure 4: \( \bar{x} = 2.4 \), \( \rho = 0.8 \), \( \sigma_x = 0.5 \), \( c = 2 \), \( \bar{F} = 1.5 \), \( F = 5 \), \( \eta = 0.5 \), \( \beta = 0.8 \), \( \alpha = 0.2 \), and \( \sigma_\omega = 0.05 \).
Figure 10. Model-implied path of credit spreads around a financial crisis. This figure shows the model-implied expected path of credit spreads in “event time” conditional on the onset of a crisis at time $\tau = 0$, for the extended model featuring both rational and behavioral beliefs in equal measure ($\theta = 0.5$). Specifically, using 100,000 periods of simulated data assuming our baseline set of parameters, we estimate regressions of the form:

$$(1 - \eta)\lambda_i^Y = a + \sum_{t=1}^{T} b_i [I_{1_{\text{default}}}^i] + \epsilon_i,$$

for $Y \in \{B, R, C\}$. We plot the $b_i$ coefficients versus event time $\tau$ below. Since $\theta = 0.5$, the coefficients for $(1 - \eta)\lambda_i^Y$ are a 50:50 mixture of the coefficients for $(1 - \eta)\lambda_i^R$ and $(1 - \eta)\lambda_i^B$. 

![Graph showing the model-implied path of credit spreads](image)
Figure 11. Reflexivity: Impact of adding rational beliefs. This figure illustrates the existence of reflexive regions in the extended model featuring both rational and extrapolative beliefs. The top-panel heatmaps plot $\frac{\partial \lambda}{\partial \lambda} \frac{\partial F_t}{\partial t}$ vs. $(x_t, F_{t-1})$ for $\lambda^R = 0.2$, with the left-panel showing results when $\theta = 1$ (fully extrapolative beliefs) and the right-panel showing results when $\theta = 0.5$ (both rational and extrapolative beliefs). The bottom panels plot $\lambda^R$ and $\frac{\partial \lambda}{\partial \lambda} \frac{\partial F_t}{\partial t}$ vs. $F_{t-1}$ for $\lambda^R = 0.2$ and $x_t = 1$, with the left-panel showing results when $\theta = 1$ and the right-panel showing results when $\theta = 0.5$. 
Figure 12. Sample path in a model with Bayesian inference. This figure plots the evolution of cash flows ($x_t$), debt outstanding ($F_t$), the default indicator ($D_t$), bond prices ($p_t$), and investor beliefs ($\lambda_t$) for a sample path in the model with Bayesian inference. We keep the parameter values from the baseline model: $\bar{x} = 2.4$, $\rho = 0.8$, $\sigma_x = 0.5$, $c = 2$, $\xi = 1.5$, $\bar{F} = 5$, and $\eta = 0.5$. In addition, for Bayesian inference, we set $F_U = 10$, $x = \{1.2, 1.5, 1.8, 2.1, 2.4, 2.7, 3.0, 3.3, 3.6\}$ (so $K = 9$) and $N = 85$ (so the bin size for firm debt is 0.1). We specify initial values $F_0 = 3.5$ and $x_0 = 2.4$, and we set initial beliefs $P_i^{i,j} = 0.1 / ((1 - \eta)F^j K)$ for $11 \leq i \leq 35$ and $1 \leq j \leq K$. Finally, we set the true cash flows as follows: $x_t = 2.4$ for $1 \leq t \leq 10$; $x_t = 1.9$ for $t = 11$; $x_t = 1.4$ for $12 \leq t \leq 18$; $x_t = 1.9$ for $t = 19$; and $x_t = 2.4$ for $20 \leq t \leq 30$. Each period represents one year.
Figure 13. Allowing for an opportunistic supply response: Model-implied impulse response functions in a non-reflexive region. The top panel shows the responses following a 0.5 up or down impulse to cash flows \( x_t \) at \( t = 1 \). The bottom panel shows the responses following a 0.25 up or down impulse to backward-looking beliefs \( \lambda^B_t \) at \( t = 1 \). The initial condition in both cases is \( x_0 = 2.25 \), \( F_{-1} = 2.25 \), and \( \lambda^B_0 = 0.3 \). The market timing parameter is set to \( M = 5 \). The other model parameters are \( \bar{x} = 2.4 \), \( \rho = 0.8 \), \( \sigma_x = 0.5 \), \( c = 2 \), \( F = 1.5 \), \( \bar{F} = 5 \), \( \eta = 0.5 \), \( \beta = 0.8 \), \( \alpha = 0.2 \), \( \sigma_\omega = 0.05 \), and \( \theta = 0.5 \).
Table 1. Credit market overheating and future corporate bond returns. This table presents time-series regressions of the form:

\[ r_{t+k}^{HY} = a + b \cdot \text{Overheating}_t + e_{t+k}, \]

where Overheating\(_t\) is a proxy for credit market overheating in year \(t\). The data begin in 1983 and end in 2014. The dependent variable is the cumulative \(k = 2\)- or 3-year excess return on high-yield bonds over like-maturity Treasuries. HY\(_t\) is the fraction of nonfinancial corporate bond issuance with a high-yield rating from Moody’s, as defined in Greenwood and Hanson (2013). Credit Growth\(_t\) is the percentage change in outstanding corporate credit and is computed using Table L103 from the Flow of Funds. Easy Credit\(_t\) is the three-year average of the percentage of bank loan officers reporting a loosening of commercial lending standards from the Federal Reserve’s Senior Loan Office Opinion Survey. \(-1 \times EBP_t\) is negative one times the excess bond premium from Gilchrist and Zakrajšek (2012). The \(t\)-statistics for \(k\)-period forecasting regressions (in brackets) are based on Newey-West (1987) standard errors, allowing for serial correlation up to \(k\) lags.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\log(HYS_t))</td>
<td>-15.95</td>
<td></td>
<td></td>
<td></td>
<td>-18.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-3.06]</td>
<td></td>
<td></td>
<td></td>
<td>[-3.78]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credit Growth(_t)</td>
<td>-126.50</td>
<td></td>
<td></td>
<td></td>
<td>-158.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-2.28]</td>
<td></td>
<td></td>
<td></td>
<td>[-2.86]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy Credit(_t)</td>
<td>-0.57</td>
<td></td>
<td></td>
<td></td>
<td>-0.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-2.12]</td>
<td></td>
<td></td>
<td></td>
<td>[-4.03]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-1 \times EBP_t)</td>
<td>-19.29</td>
<td></td>
<td></td>
<td></td>
<td>-24.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[-2.43]</td>
<td></td>
<td></td>
<td></td>
<td>[-4.79]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-15.67</td>
<td>11.46</td>
<td>-0.27</td>
<td>3.25</td>
<td>-17.54</td>
<td>14.93</td>
<td>-1.26</td>
<td>4.51</td>
</tr>
<tr>
<td></td>
<td>[-2.18]</td>
<td>[3.85]</td>
<td>[-0.06]</td>
<td>[0.94]</td>
<td>[-2.29]</td>
<td>[4.42]</td>
<td>[-0.21]</td>
<td>[1.15]</td>
</tr>
<tr>
<td>(N)</td>
<td>29</td>
<td>29</td>
<td>20</td>
<td>29</td>
<td>28</td>
<td>28</td>
<td>19</td>
<td>28</td>
</tr>
<tr>
<td>(R)-squared</td>
<td>0.20</td>
<td>0.11</td>
<td>0.20</td>
<td>0.20</td>
<td>0.25</td>
<td>0.16</td>
<td>0.38</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Table 2. Credit market overheating and current and past default rates. This table presents the results from estimating time-series regressions of the form:

\[ Overheating_t = a + b \cdot Def_t + c \cdot Def_{t-1} + e_t, \]

where \( Def \) denotes the default rate on speculative grade bonds and \( Overheating \) is a proxy for credit market overheating. The data begin in 1983 and end in 2014. \( HYS_t \) is the fraction of nonfinancial corporate bond issuance with a high-yield rating from Moody’s, as defined in Greenwood and Hanson (2013). \( Credit Growth_t \) is the percentage change in outstanding corporate credit and is computed using Table L103 from the Flow of Funds. \( Easy Credit_t \) is the three-year average of the percentage of bank loan officers reporting a loosening of commercial lending standards from the Federal Reserve’s Senior Loan Office Opinion Survey. \(-1 \times EBP_t\) is negative one times the excess bond premium from Gilchrist and Zakrajšek (2012). The \( t \)-statistics (in brackets) are based on Newey-West (1987) standard errors, allowing for serial correlation up to 3 lags.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(( HYS_t ))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Credit Growth_t )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Easy Credit_t )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-1 \times EBP_t)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Def_t )</td>
<td>-0.113</td>
<td>-0.005</td>
<td>-3.425</td>
<td>-0.104</td>
</tr>
<tr>
<td>( Def_{t-1} )</td>
<td>0.009</td>
<td>-0.008</td>
<td>-2.152</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>[0.44]</td>
<td>[-3.47]</td>
<td>[-4.24]</td>
<td>[0.96]</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.734</td>
<td>0.118</td>
<td>18.076</td>
<td>-0.36</td>
</tr>
<tr>
<td></td>
<td>[-3.47]</td>
<td>[7.92]</td>
<td>[4.58]</td>
<td>[-2.21]</td>
</tr>
<tr>
<td>( N )</td>
<td>31</td>
<td>31</td>
<td>22</td>
<td>30</td>
</tr>
<tr>
<td>( R )-squared</td>
<td>0.400</td>
<td>0.436</td>
<td>0.813</td>
<td>0.426</td>
</tr>
</tbody>
</table>